Algebraic Coding Theory

RWTH Aachen, WS 2024
RWTH Aachen, SS 2022
RWTH Aachen, SS 2019
Universität Duisburg-Essen, WS 2011
RWTH Aachen, SS 2006

Jürgen Müller

Contents

I Shannon

0 Introduction

(0.1) Communication. 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 like, 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, or parts might even be completely erased, when sent through the channel. How is it possible to recover all or at least 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 $w = [x_1, \dots, x_n]$ consisting of $n \in \mathbb{N}$ symbols $x_i \in \mathcal{X}$ 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 $u \in \mathcal{X}^*$. Hence $\mathcal{X}^*$ is a monoid, being called the **free monoid** over $\mathcal{X}$. Moreover, we have $l(vw) = l(v) + l(w)$, hence 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 $\mathbb{Z}_q := \{0, \dots, q-1\}$ for $q \in \mathbb{N}\setminus\{1\}$ are used; for example $q = 10$, whose elements are called **digits**. 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 $\{\mathtt{0}, \dots, \mathtt{9}, \mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{D}, \mathtt{E}, \mathtt{F}\}$ being in bijection with $\mathbb{Z}_{16}$ are used. For interchange of written texts the **Latin** alphabet $\{\mathtt{A}, \dots, \mathtt{Z}\}$ 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 **encoded** into words over an alphabet $\mathcal{X}$ suitable for the chosen channel, and words have to be **decoded** again after transmition. Thus, most generally, a **code** is a (finite) subset $\emptyset \neq \mathcal{C} \subseteq \mathcal{X}^*$; then $\mathcal{C}$ is interpreted as the set of all words over $\mathcal{X}$ representing sensible information.

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%

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 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 symbol. 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**, and $[x_7, \ldots, x_9]$ is the **title code**. Finally, x_{10} is a **check symbol** fulfilling $x_{10} = \sum_{i=1}^{9} i x_i \in \mathbb{Z}_{11}$. Hence a valid ISBN-10 is an element of the $\mathbb{Z}_{11}$-subspace $\{[x_1, \ldots, x_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} i x_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$.

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 symbols 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]. a) 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 $x_1, \dots, x_{12}$ are information symbols, where $[x_1, \dots, x_3]$ is the **country code**, $x_1 = 4$ referring to Germany, $[x_4, \dots, x_7]$ is the **company code**, and $[x_8, \dots, x_{12}]$ is the **article code**. Finally, 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 $\{[x_1, \dots, x_{13}] \in \mathbb{Z}_{10}^{13}; \sum_{i=1}^{6}(x_{2i-1} + 3x_{2i}) + x_{13} = 0 \in \mathbb{Z}_{10}\} \subseteq \mathbb{Z}_{10}^{13}$.

b) 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 symbol 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 symbol start with white and end with black, where for even type the width patterns are just those for odd type read backwardly. The negative type code for each symbol 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 symbol 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 symbols $x_2, \dots, x_7$ are depicted by odd and even type codes, then there is an infix 01010, then the symbols $x_8, \dots, x_{12}$ are depicted by negative type codes, and finally there is a postfix 101. The choice of the odd and even type codes for $x_2, \dots, x_7$ 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 $[x_1, x_2; x_3, x_4; x_5, \dots, x_{12}; x_{13}, \dots, x_{22}] \in \mathbb{Z}_{26}^2 \times \mathbb{Z}_{10}^{20}$ 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, $x_1, x_2; x_5, \dots, x_{12}; x_{13}, \dots, x_{22}$ 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** $[x_5, \dots, x_{12}]$, and the 10-digit **bank account number** $[x_{13}, \dots, x_{22}]$, where the latter is possibly filled up by leading zeroes; the word $[x_5, \dots, x_{22}]$ is also called the **Basic Bank Account Number (BBAN)**.

Table 2: EAN bar code.

symbols	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	− + + − +−

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 $v \equiv 1 \pmod{97}$.

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 $[x_3, x_4] \notin \{[0, 0], [0, 1], [9, 9]\}$ 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 '`DE`68 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 $vw^{\mathrm{tr}} := \sum_{i=1}^{n} x_i w_i = 0 \in \mathbb{Z}_q$.

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 $v'w^{\mathrm{tr}} = (v' - v)w^{\mathrm{tr}} = (x'_j - x_j)w_j \in \mathbb{Z}_q$, 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 $\mu_{w_j} \colon \mathbb{Z}_q \to \mathbb{Z}_q \colon x \mapsto xw_j$ is injective, or equivalently bijective.

Lemma. For $y \in \mathbb{Z}_q$ the map $\mu_y \colon \mathbb{Z}_q \to \mathbb{Z}_q \colon x \mapsto xy$ 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 $\mathbb{Z}_q$.

Proof. 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 $j \in \{1, \ldots, n-1\}$.

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 parity check codes, see (1.4); here 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 $j \in \{1, \ldots, n\}$.

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	$\alpha^4\beta$	$(1,2)(3,5)$

---

b) Let π_i be injective, for $i \in \{1,\dots,n\}$. We consider adjacent transposition errors for $n \geq 2$: Let $v := [x_1,\dots,x_n] \in G^n$ be valid and let $v' := [x_1,\dots,x_{j+1},x_j,\dots,x_n] \in G^n$ such that $x_{j+1} \neq x_j$ for some $j \in \{1,\dots,n-1\}$. Let $y_i := x_i^{\pi_i} \in G$ for $i \in \{1,\dots,n\}$ 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 $y_1^{-1},\dots,y_{j-1}^{-1}$, and from the right by $y_n^{-1},\dots,y_{j+2}^{-1}$, 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 $g \neq h \in G$ and $j \in \{1,\dots,n-1\}$.

(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 $g \neq h \in D_{10}$.

The serial numbers on the former German currency **Deutsche Mark (DM)** are formed as follows: The alphabet is $\mathcal{X} := \{\texttt{0},\dots,\texttt{9,A,D,G,K,L,N,S,U,Y,Z}\}$, and words $[x_1,\dots,x_{10};x_{11}] \in \mathcal{X}^{11}$ have length 11, where $x_1,\dots,x_{10}$ 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.

Theorem: Paige [1947]. The abelian group G has a complete map if either $|G|$ is odd or G has at least two involutions.

Proof. It is surprisingly difficult to prove this completely, so that we only give a partial proof, encompassing the accessible pieces:

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 $g \neq h \in G$, is equivalent to $g^{\tau-1} \neq h^{\tau-1}$, for $g \neq h \in G$, 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 characterizing the groups having complete maps:

Theorem. The following assertions are equivalent:
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.

Proof. 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 $S \leq N_G(S)$ centralizes $Z(S)$, we conclude that $C \cap Z(S)$

has odd length, so that $C\backslash Z(S)$ has even length, entailing $\prod_{z\in C\backslash Z(S)} z \in [G,G]$. 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 $i \in \{1,\dots,d\}$, 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 $\{g_1,\dots,g_{|G|}\}$ 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 $\prod_{g\in G} g \equiv \prod_{g\in G} gg^\sigma \equiv \prod_{g\in G} g \cdot \prod_{g\in G} g^\sigma \equiv (\prod_{g\in G} g)^2 \pmod{[G,G]}$, 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 $G \to G\colon g \mapsto gg^\sigma = g^2$ 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. ♯

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 infer that $|G|$ is odd if $d = 0$, that G has a non-trivial cyclic Sylow 2-subgroup if $d = 1$, and 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 Source coding

(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 $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$ such that $\mathcal{A} \cap \mathcal{B} = \emptyset$.

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 of x should be the smaller the more often x occurs. Moreover, for independent events their information contents should add up, while the associated probabilities multiply. Hence letting $\mathcal{S} := \{a \in \mathbb{R}; 0 < a \leq 1\}$, 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 $a, b \in \mathcal{S}$. Then the **information content** of a possible elementary event 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 $\eta(a + b) = \iota(\exp(a + b)) = \iota(\exp(a)\exp(b)) = \iota(\exp(a)) + \iota(\exp(b)) = \eta(a) + \eta(b)$. Letting $\alpha := -\eta(-1) \in \mathbb{R}$, we get $\eta(-n) = -\alpha n$, for $n \in \mathbb{N}_0$, thus $\eta(-\frac{n}{m}) = -\alpha \cdot \frac{n}{m}$, for $m \in \mathbb{N}$. Hence η being continuous we infer that $\eta(a) = \alpha a$, for $a \leq 0$. Thus from $\iota(\exp(a)) = \eta(a) = \alpha a = \alpha \ln(\exp(a))$ we get $\iota(a) = \alpha \ln(a)$, for $a \in \mathcal{S}$. Since ι is strongly decreasing we have $\alpha < 0$, so that $\iota(a) \geq 0$. Conversely, for any $\alpha < 0$ the map $\mathcal{S} \to \mathbb{R}_{\geq 0}\colon a \mapsto \alpha \ln(a)$ 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 for $a \in \mathcal{S}$ we let

$$\iota(a) := -\frac{\ln(a)}{\ln(2)} = -\log_2(a) = \log_2(\frac{1}{a}) \in \mathbb{R}_{\geq 0}.$$

(2.2) Entropy. Let $\mathcal{X} = \{x_1, \ldots, x_q\}$ be an alphabet with probability distribution μ. The **average information content** or **entropy** or **uncertainty** of $\mathcal{X}$, letting $p_i := \mu(x_i) \in \mathbb{R}_{\geq 0}$ for $i \in \{1, \ldots, q\}$, and $\mathcal{I} := \{i \in \{1, \ldots, q\}, p_i > 0\}$, is

the expected value of the information content associated with μ:

$$H(\mathcal{X}) = H(\mu) := -\sum_{i\in\mathcal{I}} p_i \log_2(p_i) \in \mathbb{R}_{\geq 0}.$$

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 $i \in \{1,\dots,q\}$, 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 $i \in \{1,\dots,q\}$, and $p_j = 0$ for $j \neq i$, that is μ is concentrated in x_i for some $i \in \{1,\dots,q\}$.

Proposition. We have $H(\mathcal{X}) \leq \log_2(|\mathcal{X}|)$, with equality if and only if $\mathcal{X}$ carries the uniform distribution.

Proof. The **Jensen inequality**, applied to the (strictly) concave logarithm function, says that for $\lambda_1,\dots,\lambda_q > 0$ such that $\sum_{i=1}^q \lambda_i = 1$, and $\alpha_1,\dots,\alpha_q > 0$ we have $\sum_{i=1}^q \lambda_i \log_2(\alpha_i) \leq \log_2(\sum_{i=1}^q \lambda_i\alpha_i)$, with equality (by strictness) if and only if $\alpha_1 = \dots = \alpha_q$.

We have $H(\mathcal{X}) = -\sum_{i\in\mathcal{I}} p_i\log_2(p_i) = \sum_{i\in\mathcal{I}} p_i\log_2(\frac{1}{p_i}) \leq \log_2(\sum_{i\in\mathcal{I}} p_i\cdot\frac{1}{p_i}) = \log_2(\sum_{i\in\mathcal{I}} 1) \leq \log_2(q) = \log_2(|\mathcal{X}|)$. Moreover, if equality holds then have $|\mathcal{I}| = q$, and thus $\sum_{i=1}^q p_i\log_2(\frac{1}{p_i}) = \log_2(\sum_{i=1}^q p_i\cdot\frac{1}{p_i})$ entails $p_1 = \dots = p_q = \frac{1}{q}$; conversely, in the latter case we get $H(\mathcal{X}) = -\frac{1}{q}\cdot\sum_{i=1}^q \log_2(\frac{1}{q}) = \log_2(q)$. ♯

Example: The binary alphabet. We consider the binary alphabet $\mathbb{Z}_2 = \{0,1\}$ with elementary probabilities $\mu(0) = p$ and $\mu(1) = 1-p$ for some $0 \leq p \leq 1$. Then the average information content equals $H(\mu) = H(p) := -p\log_2(p) - (1-p)\log_2(1-p)$. Differentiating yields $\partial_p H = -\log_2(p) + \log_2(1-p) = \log_2(\frac{1-p}{p})$, for $0 < p < 1$. Since $H(0) = H(1) = 0$ and $H(p) > 0$, for $0 < p < 1$, 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 $\mathbb{Z}_2$ 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**:

(2.3) Theorem: Shannon [1948]. Let $\mathcal{X} = \{x_1,\dots,x_q\}$ be an alphabet with probability distribution μ.

a) Let $\gamma\colon \mathcal{X} \to (\mathbb{Z}_2)^* \setminus \{\epsilon\}$ be any injective and **prefix-free** encoding, that is for $v \in \gamma(\mathcal{X})$ and $w \in (\mathbb{Z}_2)^* \setminus \{\epsilon\}$ we have $vw \notin \gamma(\mathcal{X})$. Then for the average

length of the code words in $\gamma(\mathcal{X})$ we have

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

b) If $H(\mathcal{X}) > 0$, then there is γ_0 as above such that $H(\gamma_0) < H(\mathcal{X}) + 1$.

Proof. a) i) Let $l_i := l(\gamma(x_i)) \in \mathbb{N}$, for $i \in \{1, \ldots, q\}$. We first show the **Kraft–McMillan inequality** [1949, 1956], saying that $\sum_{i=1}^{q} 2^{-l_i} \leq 1$:

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 $\gamma(x_i) \in \mathbb{Z}_2^{l_i}$ as their prefix. Since γ is prefix-free the latter sets of words are pairwise disjoint. Thus, since there are 2^{l_q} words in $\mathbb{Z}_2^{l_q}$, we get $\sum_{i=1}^{q} 2^{l_q - l_i} \leq 2^{l_q}$, hence $\sum_{i=1}^{q} 2^{-l_i} \leq 1$. ♯

ii) Let $p_i := \mu(x_i) \in \mathbb{R}_{\geq 0}$, for $i \in \{1, \ldots, q\}$. We show the **Gibbs inequality [1875]**, saying that for $\alpha_1, \ldots, \alpha_q > 0$ such that $\sum_{i=1}^{q} \alpha_i = 1$ we have $H(\mathcal{X}) \leq -\sum_{i=1}^{q} p_i \log_2(\alpha_i)$, with equality if and only if $\alpha_i = p_i$, for $i \in \{1, \ldots, q\}$:

Letting $\mathcal{I} := \{i \in \{1, \ldots, q\}, p_i > 0\}$, applying the Jensen inequality again 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 $\frac{\alpha_1}{p_1} = \cdots = \frac{\alpha_q}{p_q}$; in the latter case $\sum_{i=1}^{q} \alpha_i = 1 = \sum_{i=1}^{q} p_i$ yields $\alpha_i = p_i$, for $i \in \{1, \ldots, q\}$. ♯

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

b) We consider the **Shannon-Fano encoding**: We may assume that $p_1 \geq p_2 \geq \cdots \geq p_q \geq 0$, where since $H(\mathcal{X}) > 0$ we have $1 > p_1 \geq p_2 > 0$. In order to cover impossible symbols as well, we let $r := \max\{i \in \{2, \ldots, q\}; p_i > 0\}$. (Actually, the Shannon-Fano encoding only refers to the case $r = q$.)

For $i \in \{1, \ldots, r\}$ let $k_i \in \mathbb{N}$ such that $2^{-k_i} \leq p_i < 2^{-k_i+1}$, and for $i \in \{r+1, \ldots, q\}$ let $k_i := k_r + (i - r)$. Hence we have $1 \leq k_1 \leq k_2 \leq \cdots \leq k_r < k_{r+1} < \cdots < k_q$, and $k_i < 1 - \log_2(p_i) \leq k_i + 1$, for $i \in \{1, \ldots, r\}$.

Let $s_i := \sum_{j=1}^{i-1} p_j$, for $i \in \{1, \ldots, r+1\}$; hence we have $0 = s_1 < s_2 < \cdots < s_r < s_{r+1} = 1$. Now let $\gamma_0(x_i) \in \mathbb{Z}_2^{k_i}$ be given as the binary expansion of s_i truncated after position k_i, for $i \in \{1, \ldots, r\}$, while for $i \in \{r+1, \ldots, q\}$ we use the expansion of $2^i + (\sum_{j=1}^{r} 2^{-j})$; in particular, letting the expansion of $s_{r+1} = 1$ being defined as $(.111\ldots)_2$, truncating after position k_{r+1} yields $\gamma_0(x_{r+1})$.

Since for $1 \leq h < i \leq r+1$ we have $s_i - s_h = \sum_{j=h}^{i-1} p_j \geq p_h \geq 2^{-k_h}$ we infer that $\gamma_0(x_i)$ and $\gamma_0(x_h)$ differ on at least one of the first k_h positions. Moreover,

the same holds for $1 \leq h \leq r$ and $r+1 \leq i \leq q$, and by construction for $r+1 \leq h < i \leq q$ as well. Thus we conclude that γ_0 is both injective and prefix-free, that is an admissible encoding. Finally, for the average word length of γ_0 we get $H(\gamma_0) = \sum_{i=1}^{q} p_i k_i = \sum_{i=1}^{r} p_i k_i < \sum_{i=1}^{r} p_i(1 - \log_2(p_i)) = 1 + H(\mathcal{X})$. ♯

(2.4) Remark. a) The quantity $H(\gamma)$ 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 μ_n, which is obtained from the uniform distribution on $\mathbb{Z}_2$ by choosing the symbols in the words independently, we get $\mu_n(w) = \frac{1}{2^n}$, for $w \in \mathbb{Z}_2^n$, thus $\iota(w) = -\log_2(\frac{1}{2^n}) = \log_2(2^n) = n = l(w)$. Hence summing over $\mathbb{Z}_2^n$ yields $H(\mathbb{Z}_2^n) = -2^n \cdot \frac{1}{2^n} \cdot \log_2(\frac{1}{2^n}) = \log_2(2^n) = n$.

Thus $H(\gamma)$ is the average information content of the code words in $\gamma(\mathcal{X})$, with respect to the uniform distribution μ_n on $\mathbb{Z}_2^n$, for $n \in \mathbb{N}$, and Shannon's Theorem says that this cannot possibly be strictly smaller than the average information content of the original alphabet $\mathcal{X}$.

b) The second part, saying that there are prefix-free injective encodings having average word length bounded above by $H(\mathcal{X})+1$, shows that, whenever $H(\mathcal{X}) > 0$, the lower bound is attained up to a factor of $1 + \frac{1}{H(\mathcal{X})}$. Thus, replacing $\mathcal{X}$ by $\mathcal{X}^n$, where the entries are chosen independently, so that by (3.2) below we have $H(\mathcal{X}^n) = n \cdot H(\mathcal{X})$, we get a factor of $1 + \frac{1}{n \cdot H(\mathcal{X})} \to 1$, for $n \to \infty$, entailing that the lower bound actually is attained asymptotically.

A prefix-free injective encoding of $\mathcal{X}$ is called **optimal** if its average word length is best possible amongst all such encodings of $\mathcal{X}$. The Shannon-Fano encoding (a top to bottom approach) is not necessarily optimal, but the **Huffman encoding** [1952] (a bottom to top approach) always is, see Exercise (23.9).

3 Channel coding

(3.1) Noise. We describe the standard model for discrete noisy channels: The data consists of symbols in an alphabet $\mathcal{X}$ such that $q := |\mathcal{X}|$, sent with probability distribution $\mu_{\mathcal{X}}$, and being distorted by the channel, so that the received symbols in $\mathcal{Y} = \mathcal{X}$ carry the probability distribution $\mu_{\mathcal{Y}}$. The noise is described by the conditional probability distribution $\mu_{\mathcal{Y}|\mathcal{X}}$, thus we have $\mu_{\mathcal{Y}}(j) = \sum_{i \in \mathcal{X}} \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}|i}(j)$, for $j \in \mathcal{Y}$.

The **symmetric** channel with error probability $0 \leq p \leq \frac{q-1}{q}$ is given by $\mu_{\mathcal{Y}|i}(j) = \frac{p}{q-1}$ and $\mu_{\mathcal{Y}|i}(i) = 1 - p$, for $i \in \mathcal{X}$ and $i \neq j \in \mathcal{Y}$. In other words, writing $\mu_{\mathcal{X}} = [\mu_{\mathcal{X}}(i); i \in \mathcal{X}]$ and $\mu_{\mathcal{Y}} = [\mu_{\mathcal{Y}}(j); j \in \mathcal{Y}]$, we have $\mu_{\mathcal{Y}} = \mu_{\mathcal{X}} \cdot M_q(p)$, where $\mu_{\mathcal{Y}|\mathcal{X}}$ is given by the transition matrix

$$M_q(p) = [\mu_{\mathcal{Y}|i}(j)]_{i,j} = (1 - \frac{pq}{q-1}) \cdot E_q + \frac{p}{q-1} \cdot J_q,$$

where $J_q \in \mathbb{Q}^{q\times q}$ is the matrix all of whose entries are equal to 1. For example, for the **symmetric binary** channel we have $M_2(p) = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$.

We have $\det(M_q(p)) = (1-\frac{pq}{q-1})^{q-1}$; in particular $M_q(p) \in \mathrm{GL}_q(\mathbb{R})$ for $p < \frac{q-1}{q}$:

Since $\mathrm{rk}_{\mathbb{Q}}(J_q) = 1$ and J_q has q as an eigenvalues, the characteristic polynomial of J_q is given as $\det(XE_q - J_q) = \chi_{J_q}(X) = X^{q-1}(X-q) \in \mathbb{Q}[X]$. Hence from $M_q(p) = \frac{-p}{q-1} \cdot \big((q-\frac{q-1}{p}) \cdot E_q - J_q\big)$ we get $\det(M_q(p)) = (\frac{-p}{q-1})^q \cdot \chi_{J_q}(q-\frac{q-1}{p}) = (\frac{-p}{q-1})^q \cdot (-\frac{q-1}{p}) \cdot (q-\frac{q-1}{p})^{q-1} = \big(\frac{-p}{q-1} \cdot (q-\frac{q-1}{p})\big)^{q-1} = (1-\frac{pq}{q-1})^{q-1}$. ♯

The **quiet** channel is given by $p = 0$, hence from $M_q(0) = E_q$ we get $\mu_{\mathcal{Y}} = \mu_{\mathcal{X}}$. Note that if the channel is not quiet then we have $\mu_{\mathcal{Y}}(j) \neq 0$ for $j \in \mathcal{Y}$.

The **completely noisy** channel is given by $p = \frac{q-1}{q}$, hence from $M_q(p) = \frac{1}{q} \cdot J_q$ we get $\mu_{\mathcal{Y}} = \frac{1}{q} \cdot [1, \ldots, 1]$, that is the uniform distribution, independently of $\mu_{\mathcal{X}}$.

If the channel is not completely noisy, that is $p < \frac{q-1}{q}$, then if $\mathcal{X}$ carries the uniform distribution we get $\mu_{\mathcal{Y}} = \frac{1}{q} \cdot [1, \ldots, 1] \cdot M_q(p) = \frac{1}{q} \cdot [1, \ldots, 1]$, that is $\mathcal{Y}$ carries the uniform distribution as well, independently of p. Moreover, since in this case $M_q(p) \in \mathrm{GL}_q(\mathbb{R})$ is bijective, if $\mathcal{Y}$ carries the uniform distribution, then $\mathcal{X}$ necessarily carries the uniform distribution as well. In conclusion, $\mathcal{X}$ carries the uniform distribution if and only if $\mathcal{Y}$ does so.

Finally, if the channel is not quiet, that is $p > 0$, we provide the transition matrix $\widetilde{M}_q(p)$ for the conditional probability $\mu_{\mathcal{X}|\mathcal{Y}}$:

Bayes's Theorem says that $\mu_{\mathcal{X}|j}(i)\mu_{\mathcal{Y}}(j) = \mu_{\mathcal{X}\times\mathcal{Y}}(i,j) = \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}|i}(j)$, hence $\widetilde{M}_q(p)^{\mathrm{tr}} \cdot \mathrm{diag}[\mu_{\mathcal{Y}}(j)]_j = \mathrm{diag}[\mu_{\mathcal{X}}(i)]_i \cdot M_q(p)$. From $\mu_{\mathcal{Y}}(j) \neq 0$, for $j \in \mathcal{Y}$, we get

$$\widetilde{M}_q(p) = \mathrm{diag}[\mu_{\mathcal{Y}}(j)]_j^{-1} \cdot M_q(p) \cdot \mathrm{diag}[\mu_{\mathcal{X}}(i)]_i.$$

In particular, if $\mathcal{X}$ carries the uniform distribution, thus $\mathcal{Y}$ carrying the uniform distribution as well, then $\widetilde{M}_q(p) = M_q(p)$; note that this also holds for $p = 0$. ♯

(3.2) Capacity. 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 afforded by 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 noise. This is the unavoidable cost of transport through the noisy channel, so that 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.

Proposition. We have $H(\mathcal{X}\times\mathcal{Y}) = H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y})$, and $H(\mathcal{X}\times\mathcal{Y}) \leq H(\mathcal{X}) + H(\mathcal{Y})$, with equality if and only if $\mathcal{X}$ and $\mathcal{Y}$ are independent.

Proof. i) We first show that $H(\mathcal{X}\times\mathcal{Y}) \leq H(\mathcal{X}) + H(\mathcal{Y})$: 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 $i \in \mathcal{X}$ 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

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

By Jensen's inequality this entails

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

Now 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 $i \in \mathcal{X}$ and $j \in \mathcal{Y}$. In this case we get $1 = \sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j) = m\cdot\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{X}\times\mathcal{Y}}(i,j) = m$, saying that $\mu_{\mathcal{X}\times\mathcal{Y}}(i,j) = \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j)$.

ii) We now show that $H(\mathcal{X}\times\mathcal{Y}) = H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y})$: 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

$$\begin{aligned} H(\mathcal{X}\times\mathcal{Y}) &= -\textstyle\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{X}\times\mathcal{Y}}(i,j)\log_2(\mu_{\mathcal{X}\times\mathcal{Y}}(i,j)) \\ &= -\textstyle\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{Y}}(j)\mu_{\mathcal{X}|j}(i)(\log_2(\mu_{\mathcal{Y}}(j)) + \log_2(\mu_{\mathcal{X}|j}(i))) \\ &= -\textstyle\sum_{j\in\mathcal{Y}} \big(\mu_{\mathcal{Y}}(j)\cdot\sum_{i\in\mathcal{X}} \mu_{\mathcal{X}|j}(i)\log_2(\mu_{\mathcal{X}|j}(i))\big) \\ &\quad -\textstyle\sum_{j\in\mathcal{Y}} \big(\mu_{\mathcal{Y}}(j)\log_2(\mu_{\mathcal{Y}}(j))\cdot\sum_{i\in\mathcal{X}} \mu_{\mathcal{X}|j}(i)\big) \\ &= H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y}), \end{aligned}$$

where we have used the fact that $\sum_{i\in\mathcal{X}} \mu_{\mathcal{X}|j}(i) = 1$. ♯

Thus the capacity $C(\mathcal{X}|\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X}\times\mathcal{Y}) \geq 0$ is non-negative 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.

(3.3) Capacity of symmetric channels. We consider the symmetric channel working over an alphabet $\mathcal{X} = \mathcal{Y}$ such that $q := |\mathcal{X}|$, with error probability

$0 \leq p \leq \frac{q-1}{q}$. Using the transition matrix $M_q(p)$ describing $\mu_{\mathcal{Y}|\mathcal{X}}$, we obtain

$$\begin{aligned} H(\mathcal{Y}|\mathcal{X}) &= -\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}}\mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}|i}(j)\log_2(\mu_{\mathcal{Y}|i}(j)) \\ &= -\sum_{i\in\mathcal{X}}\mu_{\mathcal{X}}(i)\cdot\big((1-p)\log_2(1-p)+(q-1)\cdot\tfrac{p}{q-1}\log_2(\tfrac{p}{q-1})\big) \\ &= -\big(p\log_2(\tfrac{p}{q-1})+(1-p)\log_2(1-p)\big)\cdot\sum_{i\in\mathcal{X}}\mu_{\mathcal{X}}(i) \\ &= p\log_2(q-1)-p\log_2(p)-(1-p)\log_2(1-p). \end{aligned}$$

Thus the capacity of the symmetric channel is $C_q(p) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X}) = H(\mathcal{Y}) - p\log_2(q-1) + p\log_2(p) + (1-p)\log_2(1-p)$.

In particular, for the quiet channel, that is $p = 0$, we have $\mu_{\mathcal{Y}} = \mu_{\mathcal{X}}$, and recalling that $\lim_{p\to 0^+}(p\log_2(p)) = 0$ we get $H(\mathcal{Y}|\mathcal{X}) = 0$, entailing $C_q(0) = H(\mathcal{Y}) = H(\mathcal{X})$, saying that the average information content of $\mathcal{X}$ is transported without loss through the channel.

For the completely noisy channel, that is $p = \frac{q-1}{q}$, the alphabet $\mathcal{Y}$ carries the uniform distribution in any case, hence we have $H(\mathcal{Y}) = \log_2(|\mathcal{Y}|) = \log_2(q)$, and $H(\mathcal{Y}|\mathcal{X}) = \log_2(q)$ as well; indeed, in this case the conditional entropy on $\mathcal{Y}$ should be independent of $\mathcal{X}$. This entails $C_q(\frac{q-1}{q}) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X}) = 0$, saying that no information is transported through the channel.

In general, we have $0 \leq H(\mathcal{Y}) \leq \log_2(|\mathcal{Y}|)$, where the maximum is attained precisely for the uniform distribution on $\mathcal{Y}$. Hence the maximum capacity of the symmetric channel is given as

$$C_q^{\max}(p) = \log_2(q) - p\log_2(q-1) + p\log_2(p) + (1-p)\log_2(1-p).$$

Moreover, if the channel is not completely noisy, that is $p < \frac{q-1}{q}$, then $\mathcal{Y}$ carrying the uniform distribution is equivalent to $\mathcal{X}$ carrying the uniform distribution. Thus, the maximum capacity of the channel is attained if and only if $\mathcal{X}$ has maximum average information content.

(3.4) 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 $\mathcal{X}^n$, and thus the uniform distribution on $\mathcal{C}$ as well, the information content of a code word $v \in \mathcal{C}$, viewed as an element of $\mathcal{C}$ is given as $\iota_{\mathcal{C}}(v) = -\log_2(\frac{1}{m}) = \log_2(m)$, while viewed as an element of $\mathcal{X}^n$ we have $\iota_{\mathcal{X}^n}(v) = -\log_2(\frac{1}{q^n}) = n\log_2(q)$. Hence the relative information content of the code words in $\mathcal{C}$ is given by the **information rate**

$$\rho(\mathcal{C}) = \rho_{\mathcal{X}^n}(\mathcal{C}) := \frac{\iota_{\mathcal{C}}(v)}{\iota_{\mathcal{X}^n}(v)} = \frac{\log_2(m)}{n\log_2(q)} = \frac{\log_q(m)}{n\log_q(q)} = \frac{\log_q(m)}{n}.$$

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 at all. Thus the larger $\rho(\mathcal{C})$ is, the better $\mathcal{C}$ is, in terms of information content.

Example. i) 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 such that $k := \dim_{\mathbb{F}_q}(\mathcal{C})$, then we have $\rho(\mathcal{C}) = \frac{\log_q(q^k)}{n} = \frac{k}{n}$.

ii) If $\mathcal{C} := \{v \in \mathbb{Z}_q^n; vw^{\mathrm{tr}} = 0 \in \mathbb{Z}_q\}$ is a parity check code over $\mathbb{Z}_q$, with respect to weights $w = [w_1, \ldots, w_n] \in \mathbb{Z}_q^n$, where $w_n \in \mathbb{Z}_q^*$, for any $[x_1, \ldots, x_{n-1}] \in \mathbb{Z}_q^{n-1}$ 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 $\rho(\mathcal{C}) = \frac{\log_q(q^{n-1})}{n} = \frac{n-1}{n}$.

(3.5) Maximum likelihood decoding. 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:

We again consider the symmetric channel with error probability $0 \leq p < \frac{q-1}{q}$, working over an alphabet $\mathcal{X} = \mathcal{Y}$ such that $q := |\mathcal{X}|$. We assume that $\mathcal{X}$ carries the uniform distribution, or equivalently that $\mathcal{Y}$ carries the uniform distribution, so that both the transition matrix describing $\mu_{\mathcal{X}|\mathcal{Y}}$ and the transition matrix describing $\mu_{\mathcal{Y}|\mathcal{X}}$ are equal to $M_q(p)$; see (3.1).

Let $\mathcal{C} \subseteq \mathcal{X}^n$ be a block code of length $n \in \mathbb{N}$. Since $\mathcal{X}$ carries the uniform distribution, and distinct positions are independent, $\mathcal{X}^n$ carries the uniform distribution as well, and thus so does $\mathcal{C}$.

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 we let $\mu_{\mathcal{X}^n|v}(c) = \max\{\mu_{\mathcal{X}^n|v}(w) \in \mathbb{R}; w \in \mathcal{C}\}$, being called **maximum likelihood (ML) decoding**. This is turned into combinatorics as follows:

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**. Thus for $w \in \mathcal{X}^n$ we have

$$\mu_{\mathcal{X}^n|v}(w) = \big(\frac{p}{q-1}\big)^{d(v,w)} \cdot (1-p)^{n-d(v,w)} = (1-p)^n \cdot \Big(\frac{p}{(1-p)(q-1)}\Big)^{d(v,w)}.$$

If $0 < p < \frac{q-1}{q} < 1$, then since $\frac{p}{1-p} = \frac{1}{1-p} - 1$ is increasing with p increasing, we conclude that $0 < \frac{p}{(1-p)(q-1)} < \frac{(q-1)q}{q(q-1)} = 1$, implying that $\big(\frac{p}{(1-p)(q-1)}\big)^a$ is strictly decreasing with $a \geq 0$ increasing; if $p = 0$ then we have $\mu_{\mathcal{X}^n|v}(w) = \delta_{v,w}$ anyway. Thus, in any case, 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 condition does not determine c uniquely, we revert to **partial decoding** by only allowing for **unique nearest neighbor decoding**, and otherwise mark v as an **erasure**.

If $c \in \mathcal{C}$ is sent, let $0 \leq \epsilon_c \leq 1$ be the probability that c is not recovered by the above decoding process. The expected value $\epsilon(\mathcal{C}) := \frac{1}{|\mathcal{C}|} \cdot \sum_{c \in \mathcal{C}} \epsilon_c$ is called the

average error probability of $\mathcal{C}$. Thus the smaller $0 \leq \epsilon(\mathcal{C}) \leq 1$ is the better $\mathcal{C}$ is, as far as decoding is concerned.

The question whether there are 'good' (binary) block codes $\mathcal{C}$, in the sense of having a large information rate $\rho(\mathcal{C})$ and a small average error probability $\epsilon(\mathcal{C})$ 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 following is a pure existence proof giving no clue how to actually find 'good' codes. Moreover, we are assuming that the codes under consideration carry the uniform distribution, while source coding might produce codes for which this condition does not at all hold. Anyway:

(3.6) Proposition: Chernoff inequality. Let $\mathcal{X}$ be an alphabet with probability distribution μ, let $X_i \colon \mathcal{X} \to \mathbb{Z}_2$, for $i \in \{1, \ldots, n\}$ and $n \in \mathbb{N}$, be independent random variables such that $\mu_{X_i}(1) = p$, and let $X := \sum_{i=1}^{n} X_i \colon \mathcal{X} \to \mathbb{N}_0$.

Then X is binomially distributed, such that $\mu(X = d) = \binom{n}{d} \cdot p^d (1-p)^{n-d}$, for $d \in \{0, \ldots, n\}$, and for $0 \leq \epsilon \leq 1$ we have

$$\mu\big(X \geq (1+\epsilon)pn\big) \leq \exp(-\frac{1}{4}\epsilon^2 pn).$$

Proof. The first assertion is a matter of counting. To show the second one, we may assume that $0 < p < 1$ and $0 < \epsilon < \frac{1-p}{p}$. Note that the left hand side is piecewise constant and decreasing, while the right hand side is continuous and strictly decreasing, so that we may safely exclude the case $(1+\epsilon)p = 1$.

i) Now, since X has non-negative values, for $t \in \mathbb{R}$ we have the following special case of the **Markov inequality** (where this is non-trivial only for $t > 0$):

$$t \cdot \mu(X \geq t) = \sum_{x \in \mathcal{X}} \mu(x) t \delta_{X(x) \geq t} \leq \sum_{x \in \mathcal{X}} \mu(x) X(x) \delta_{X(x) \geq t} \leq \sum_{x \in \mathcal{X}} \mu(x) X(x),$$

where the right hand side is the expected value $E(X)$ of X. Moreover, we have

$$E(\exp(tX)) = E(\exp(t \cdot \sum_{i=1}^{n} X_i)) = E(\prod_{i=1}^{n} \exp(tX_i)) = \prod_{i=1}^{n} E(\exp(tX_i)),$$

where $E(\exp(tX_i)) = (1-p) + p\exp(t)$, hence $E(\exp(tX)) = \big(1 + (\exp(t)-1)p\big)^n$. Thus we get

$$\begin{aligned} \mu\big(X \geq (1+\epsilon)pn\big) &= \mu\big(\exp(tX) \geq \exp\big(t(1+\epsilon)pn\big)\big) \\ &\leq \exp\big(-t(1+\epsilon)pn\big) \cdot E(\exp(tX)) \\ &= \exp\big(-t(1+\epsilon)pn\big) \cdot \big(1 + (\exp(t)-1)p\big)^n. \end{aligned}$$

ii) In order to obtain an optimal bound, we have to minimize the right hand side. Letting $f(t) := \exp\big(-t(1+\epsilon)p\big) \cdot \big(1 + (\exp(t)-1)p\big)$, derivation with

respect to t yields

$$\partial_t f = \exp\big(-t(1+\epsilon)p\big)\cdot\big(-(1+\epsilon)p(1+(\exp(t)-1)p)+p\exp(t)\big).$$

Hence we have $\partial_t f = 0$ if and only if $\exp(t) = \frac{(1+\epsilon)(1-p)}{1-(1+\epsilon)p} = 1+\frac{\epsilon}{1-(1+\epsilon)p}$. Since both $\lim_{t\to\infty} f(t) = \lim_{t\to\infty} \frac{1+(\exp(t)-1)p}{\exp\big(t(1+\epsilon)p\big)} = \lim_{t\to\infty} \frac{p\exp(t)}{(1+\epsilon)p\exp(t(1+\epsilon)p)} = \frac{1}{1+\epsilon}\cdot \lim_{t\to\infty}\exp(t)^{1-(1+\epsilon)p} = \infty$ and $\lim_{t\to-\infty} f(t) = \lim_{t\to-\infty} \frac{1+(\exp(t)-1)p}{\exp(t(1+\epsilon)p)} = \infty$, we infer that f has its unique minimum at $t = \ln\big(1+\frac{\epsilon}{1-(1+\epsilon)p}\big)$.

iii) Letting t be as above, the right hand side becomes $\exp\big(g(\epsilon)\cdot n\big)$, where

$$g(\epsilon) := -p(1+\epsilon)\ln(1+\epsilon) + \big(1-(1+\epsilon)p\big)\ln\big(\frac{1-p}{1-(1+\epsilon)p}\big).$$

We have $g(0) = 0$, and derivation with respect to ϵ yields

$$\partial_\epsilon g = -p\ln(1+\epsilon) - p\ln(1-p) + p\ln\big(1-(1+\epsilon)p\big).$$

We have $\partial_\epsilon g(0) = 0$, and derivation now yields $\partial_\epsilon^2 g = \frac{-p}{(1+\epsilon)(1-(1+\epsilon)p)}$. Hence Taylor's Theorem yields $g(\epsilon) = \frac{\epsilon^2}{2}\cdot\partial_\epsilon^2 g(\delta)$, for some $0\le\delta\le\epsilon$. Recalling that $\epsilon\le 1$, from $\partial_\epsilon^2 g(\delta) \le \frac{-p}{1+\delta} \le \frac{-p}{1+\epsilon} \le \frac{-p}{2}$ we finally get $g(\epsilon) \le -\frac{1}{4}\epsilon^2 p$. ♯

(3.7) Theorem: Shannon [1948]. For 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 $0\le p<\frac{1}{2}$ we have:

For any $0<\rho<1+p\log_2(p)+(1-p)\log_2(1-p) = C_2^{\max}(p)$ and $\epsilon>0$ there is a block code $\mathcal{C}\subseteq\mathcal{X}^n$, for some $n\in\mathbb{N}$, such that $\rho(\mathcal{C})\ge\rho$ and $\epsilon(\mathcal{C})<\epsilon$.

Proof. If $p=0$, then we may take $\mathcal{C}=\mathcal{X}$, thus $n=1$, fulfilling $\rho(\mathcal{X})=1$ and $\epsilon(\mathcal{X})=0$, 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 $|\Gamma_n| = \binom{2^n}{m} > 0$. Hence for $\mathcal{C}\in\Gamma_n$ we have $\rho(\mathcal{C}) = \frac{\log_2(m)}{n} = \frac{\lceil\rho n\rceil}{n}\ge\rho$. Now let $\mathcal{C}_0\in\Gamma_n$ such that $\epsilon(\mathcal{C}_0)$ is minimal. Hence $\epsilon(\mathcal{C}_0)$ is bounded above by the expected value $E_{\Gamma_n}(\epsilon)$, with respect 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 $d\in\{0,\dots,n\}$ errors is given by the binomial distribution $\beta_n(d) = \binom{n}{d}\cdot p^d(1-p)^{n-d}$. Note that the equation $n^n = (d+(n-d))^n \ge \binom{n}{d}d^d(n-d)^{n-d}$ yields $\binom{n}{d}\le\frac{n^n}{d^d(n-d)^{n-d}}$.

For $n\in\mathbb{N}$ let $0<a_n<1$ such that $\lim_{n\to\infty}a_n = 0$ and $\lim_{n\to\infty}(na_n^2) = \infty$; for example we may let $a_n := \frac{1}{\log_2(n+1)}$. Moreover, let $\delta = \delta_n := \lfloor(1+a_n)np\rfloor\in\mathbb{N}_0$; then we have $\lim_{n\to\infty}\frac{\delta_n}{n} = p < \frac{1}{2}$, hence we have $\delta_n\le\frac{n}{2}-1$, for $n\gg 0$. Now the Chernoff inequality yields, for $n\gg 0$:

$$\mu(\beta_n>\delta_n) = \mu\big(\beta_n\ge(1+a_n)np\big)\le\exp(-\frac{1}{4}a_n^2np) < \epsilon.$$

Let $\mathcal{B}_\delta(v) := \{w \in \mathcal{X}^n; d(v,w) \leq \delta_n\}$ be the **sphere** with **radius** $\delta = \delta_n$ around $v \in \mathcal{X}^n$. Since binomial coefficients $\binom{n}{d}$ are strictly increasing for $d \in \{0, \dots, \frac{n}{2}\}$, for $n \gg 0$ we get

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

Given $\mathcal{C} \in \Gamma_n$, we decode by unique nearest neighbor decoding with respect to δ_n, 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 the characteristic function of $\mathcal{B}_\delta(c)$, that is $\chi_c(v) = 1$ if and only if $d(v,c) \leq \delta_n$. Moreover, let $\varphi_c \colon \mathcal{Y}^n \to \mathbb{N}_0$ be given by

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

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 the error probability $\epsilon_c \leq \sum_{v \in \mathcal{Y}^n} \mu_{\mathcal{Y}^n|c}(v) \varphi_c(v)$. From

$$\sum_{v \in \mathcal{Y}^n} \mu_{\mathcal{Y}^n|c}(v)(1 - \chi_c(v)) = \sum_{v \in \mathcal{Y}^n \setminus \mathcal{B}_\delta(c)} \mu_{\mathcal{Y}^n|c}(v) = \mu(\beta_n > \delta_n) < \epsilon,$$

for $n \gg 0$, for the average error probability of $\mathcal{C}$ we get

$$\epsilon(\mathcal{C}) = \frac{1}{m} \cdot \sum_{c \in \mathcal{C}} \epsilon_c < \epsilon + \frac{1}{m} \cdot \sum_{v \in \mathcal{Y}^n} \sum_{c \in \mathcal{C}} \sum_{c' \in \mathcal{C} \setminus \{c\}} \mu_{\mathcal{Y}^n|c}(v) \chi_{c'}(v).$$

Hence averaging over all $\binom{2^n}{m}$ subsets of Γ_n of cardinality m, 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

$$E_{\Gamma_n}(\epsilon) < \epsilon + \frac{1}{m} \cdot \frac{m(m-1)}{2^n(2^n-1)} \cdot \sum_{v \in \mathcal{Y}^n} \sum_{c \in \mathcal{C}} \big(\mu_{\mathcal{Y}^n|c}(v) \cdot \sum_{c' \in \mathcal{C} \setminus \{c\}} \chi_{c'}(v)\big).$$

For $v \in \mathcal{Y}^n$ we have $\sum_{c \in \mathcal{X}^n} \chi_c(v) = b_\delta$, and

$$\frac{1}{2^n} = \mu_{\mathcal{Y}^n}(v) = \sum_{c \in \mathcal{X}^n} \mu_{\mathcal{Y}^n|c}(v) \mu_{\mathcal{X}^n}(c) = \frac{1}{2^n} \cdot \sum_{c \in \mathcal{X}^n} \mu_{\mathcal{Y}^n|c}(v)$$

shows that $\sum_{c \in \mathcal{X}^n} \mu_{\mathcal{Y}^n|c}(v) = 1$. Thus we get

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

for $n \gg 0$. This entails

$$\frac{1}{n}\log_2(\epsilon(\mathcal{C}_0)-\epsilon) < \frac{1}{n}\big(\lceil \rho n\rceil - n - 1 + \log_2(n)\big) - \frac{\delta_n}{n}\log_2(\frac{\delta_n}{n}) - (1-\frac{\delta_n}{n})\log_2(1-\frac{\delta_n}{n}),$$

where the right hand side tends to $\rho - 1 - p\log_2(p) - (1-p)\log_2(1-p) = \rho - C_2^{\max}(p) < 0$. Hence there is $\alpha > 0$ such that $\epsilon(\mathcal{C}_0) < \epsilon + 2^{-n\alpha}$, for $n \gg 0$. ♯

II Hamming

4 Block codes

(4.1) Hamming distance. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, and let $n \in \mathbb{N}$. Letting $v = [x_1, \dots, x_n] \in \mathcal{X}^n$ and $w = [y_1, \dots, y_n] \in \mathcal{X}^n$, then $d(v,w) := |\{i \in \{1,\dots,n\}; x_i \neq y_i\}| \in \{0,\dots,n\}$ is called their **Hamming distance**; recall that we have already used this for $q = 2$ in (3.5).

Lemma. The Hamming distance defines a discrete metric on $\mathcal{X}^n$.

Proof. We have positive definiteness $d(v,w) \in \mathbb{R}_{\geq 0}$, where $d(v,w) = 0$ if and only if $v = w$; symmetry $d(v,w) = d(w,v)$; and the triangle inequality:

Let $u := [z_1, \dots, z_n] \in \mathcal{X}^n$. Then from $\{i \in \{1,\dots,n\}; x_i \neq z_i\} = \{i \in \{1,\dots,n\}; y_i = x_i \neq z_i\} \dot\cup \{i \in \{1,\dots,n\}; y_i \neq x_i \neq z_i\} \subseteq \{i \in \{1,\dots,n\}; y_i \neq z_i\} \cup \{i \in \{1,\dots,n\}; x_i \neq y_i\}$ we get $d(v,u) \leq d(v,w) + d(w,u)$. ♯

An **isometry** of $\mathcal{X}^n$ is a **distance-preserving** map $g\colon \mathcal{X}^n \to \mathcal{X}^n$, that is $d(v,w) = d(v^g, w^g)$, for $v, w \in \mathcal{X}^n$; it follows from positive definiteness that any isometry is injective, hence is bijective. Thus the set $\mathrm{Isom}(\mathcal{X}^n)$ of all isometries of $\mathcal{X}^n$ forms a group, called the **isometry group** of $\mathcal{X}^n$.

Proposition. We have $\mathrm{Isom}(\mathcal{X}^n) \cong (\mathcal{S}_\mathcal{X})^n \rtimes \mathcal{S}_n =: \mathcal{S}_\mathcal{X} \wr \mathcal{S}_n$.

Proof. Given permutations $\pi_i \in \mathcal{S}_\mathcal{X}$, for $i \in \{1,\dots,n\}$, this yields an isometry $[\pi_1, \dots, \pi_n]$ acting component-wise, and any permutation in $\mathcal{S}_n$ induces an isometry by permuting the components. Hence $G := \mathrm{Isom}(\mathcal{X}^n)$ contains the semidirect product $S := (\mathcal{S}_\mathcal{X})^n \rtimes \mathcal{S}_n$ as a subgroup; in particular, S and hence G act transitively on $\mathcal{X}^n$.

Let $0 \in \mathcal{X}$ be a fixed element, and let $H := \mathrm{Stab}_G([0,\dots,0])$; hence we have $[G\colon H] = q^n$. Let $T := (\mathcal{S}_{\mathcal{X}\setminus\{0\}})^n \rtimes \mathcal{S}_n \leq S \cap H$; hence we have $[S\colon T] = [\mathcal{S}_\mathcal{X}\colon \mathcal{S}_{\mathcal{X}\setminus\{0\}}]^n = q^n$. For $v = [x_1, \dots, x_n] \in \mathcal{X}^n$ let $\mathrm{supp}_0(v) := \{i \in \{1,\dots,n\}; x_i \neq 0\}$. (This is reminiscent of the notion of support to be coined later.) Now let $h \in H$.

Let first $v \neq w \in \mathcal{X}^n$ such that $|\mathrm{supp}_0(v)| = |\mathrm{supp}_0(w)| = 1$. If $\mathrm{supp}_0(v) = \mathrm{supp}_0(w)$ then from $d(v^h, w^h) = d(v, w) = 1$ we conclude that $\mathrm{supp}_0(v^h) = \mathrm{supp}_0(w^h)$, while if $\mathrm{supp}_0(v) \neq \mathrm{supp}_0(w)$ then from $d(v^h, w^h) = d(v, w) = 2$ we conclude that $\mathrm{supp}_0(v^h) \neq \mathrm{supp}_0(w^h)$. Hence h induces a permutation of the components of $\mathcal{X}^n$, and permutations within the components. In other words the action of h on $\{v \in \mathcal{X}^n; |\mathrm{supp}_0(v)| = 1\}$ is induced by an element of T.

Now let $v = [x_1, \ldots, x_n] \in \mathcal{X}^n$ such that $\mathrm{supp}_0(v) = \{i_1, \ldots, i_s\}$, 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 T describing its action on $\{v \in \mathcal{X}^n; |\mathrm{supp}_0(v)| = 1\}$. Hence H is isomorphic to a subgroup of T. Hence from $H \leq T \leq S \leq G$ and $[G\colon H] = q^n = [S\colon T]$ we get $G = S$ and $H = T$. ♯

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 $\mathrm{Isom}_n(\mathbb{F}_q) \leq \mathrm{GL}_n(\mathbb{F}_q)$ of all linear isometries is called the **linear isometry group** of $\mathbb{F}_q^n$.

Lemma. An $\mathbb{F}_q$-linear map $g\colon \mathbb{F}_q^n \to \mathbb{F}_q^n$ is an isometry if and only if g is **weight-preserving**, that is $\mathrm{wt}(v^g) = \mathrm{wt}(v)$, for $v \in \mathbb{F}_q^n$.

Proof. If g is an isometry, then we have $\mathrm{wt}(v^g) = d(v^g, 0_n) = d(v, 0_n) = \mathrm{wt}(v)$; conversely, if g is weight-preserving, then we have $d(v^g, w^g) = \mathrm{wt}(v^g - w^g) = \mathrm{wt}((v - w)^g) = \mathrm{wt}(v - w) = d(v, w)$, for $v, w \in \mathbb{F}_q^n$. ♯

Proposition. We have $\mathrm{Isom}_n(\mathbb{F}_q) \cong (\mathbb{F}_q^*)^n \rtimes \mathcal{S}_n =: \mathbb{F}_q^* \wr \mathcal{S}_n$.

Proof. We have $d(v + u, w + u) = d(v, w)$, for $u, v, w \in \mathbb{F}_q^n$, thus we have $d(v, w) = d(v - w, 0_n) = \mathrm{wt}(v - w)$. Since $\mathrm{Isom}_n(\mathbb{F}_q)$ fixes $0_n \in \mathbb{F}_q^n$, we infer that $\mathrm{wt}(v^g) = \mathrm{wt}(v)$, for $g \in \mathrm{Isom}_n(\mathbb{F}_q)$. Hence for the i-th unit vector $e_i = [0, \ldots, 0, 1, 0, \ldots, 0] \in \mathbb{F}_q^n$, for some $i \in \{1, \ldots, n\}$, 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 $\mathrm{diag}[x_1, \ldots, x_n] \cdot P_\pi \in (\mathbb{F}_q^*)^n \rtimes \mathcal{S}_n \leq \mathrm{GL}_n(\mathbb{F}_q)$, 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 $\mathrm{Isom}_n(\mathbb{F}_q) = (\mathbb{F}_q^*)^n \rtimes \mathcal{S}_n \leq \mathrm{GL}_n(\mathbb{F}_q)$ is the subgroup of all monomial matrices. ♯

In other words, $\mathrm{Isom}_n(\mathbb{F}_q)$ is the normalizer in $\mathrm{GL}_n(\mathbb{F}_q)$ of the split maximal torus $(\mathbb{F}_q^*)^n$, where $\mathrm{Isom}_n(\mathbb{F}_q)/(\mathbb{F}_q^*)^n \cong \mathcal{S}_n$ is the Weyl group of $\mathrm{GL}_n(\mathbb{F}_q)$. In particular, $\mathrm{Isom}_n(\mathbb{F}_q)$ acts transitively on the sets $\{v \in \mathbb{F}_q^n; \mathrm{wt}(v) = i\}$.

(4.2) 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** $m := |\mathcal{C}| \in \mathbb{N}$.

If $m = 1$ then $\mathcal{C}$ is called **trivial**. Block codes $\mathcal{C}, \mathcal{C}' \subseteq \mathcal{X}^n$ are called **equivalent**, if there is an isometry $g \in \mathrm{Isom}(\mathcal{X}^n)$ such that $\mathcal{C}^g = \mathcal{C}'$; the **automorphism group** $\mathrm{Aut}(\mathcal{C})$ of $\mathcal{C}$ is the group of all isometries $g \in \mathrm{Isom}(\mathcal{X}^n)$ 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.

Proposition: Singleton bound [1964]. For any non-trivial (n, m, d)-code $\mathcal{C}$ we have $\log_q(m) \leq n - d + 1$.

Proof. We consider the **puncturing** map $\alpha\colon \mathcal{X}^n \to \mathcal{X}^{n-d+1}\colon [x_1, \ldots, x_n] \mapsto [x_1, \ldots, x_{n-d+1}]$ with respect to the last $d-1$ components. Since for $v \neq w \in \mathcal{C}$ 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 $m = |\mathcal{C}| \leq |\mathcal{X}^{n-d+1}| = q^{n-d+1}$. (Note that likewise we may choose any $d-1$ components instead of the tail ones.) ♯

If we have equality $d - 1 = n - \log_q(m)$, then $\mathcal{C}$ is called **maximum distance separable (MDS)**; 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; if $q \in \{2, 3, 4\}$ then $\mathcal{C}$ is called **binary**, **ternary** or **quaternary**, respectively.

Let $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \in \mathbb{N}_0$ be the **dimension** of $\mathcal{C}$, and let $d = d(\mathcal{C})$ 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 $d - 1 \leq n - k$.

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 $\mathrm{wt}(\mathcal{C}) := \infty$.

Lemma. We have $\mathrm{wt}(\mathcal{C}) = d(\mathcal{C})$.

Proof. We may assume that $\mathcal{C}$ is non-trivial. Since $\mathrm{wt}(v) = d(v, 0_n) \geq d(\mathcal{C})$, for $0_n \neq v \in \mathcal{C}$, we have $\mathrm{wt}(\mathcal{C}) \geq d(\mathcal{C})$. And for $v \neq w \in \mathcal{C}$ 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 $d(\mathcal{C}) \geq \mathrm{wt}(\mathcal{C})$. ♯

Linear codes $\mathcal{C}, \mathcal{C}' \leq \mathbb{F}_q^n$ are called **linearly equivalent**, if there is a linear isometry $g \in \mathrm{Isom}_n(\mathbb{F}_q)$ 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 linear isometries $g \in \mathrm{Isom}_n(\mathbb{F}_q)$ such that $\mathcal{C}^g = \mathcal{C}$. Note that linearly equivalent codes have the same dimension and the same minimum weight.

(4.3) Packing radius. 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 (3.7).

Let $\mathcal{C}$ be an (n, m, d)-code over $\mathcal{X}$. Then $\mathcal{C}$ is called **e-error correcting**, for some $e \in \{0, \ldots, n\}$, if $\mathcal{B}_e(v) \cap \mathcal{B}_e(w) = \emptyset$, for $v \neq w \in \mathcal{C}$. The maximum $e \in \{0, \ldots, n\}$ such that $\mathcal{C}$ is e-error correcting is called the **packing radius** $e(\mathcal{C})$ of $\mathcal{C}$. In particular, we have $e(\mathcal{C}) = n$ if and only if $\mathcal{C}$ is trivial.

Proposition: Sphere packing bound [Hamming, 1960]. We have

$$m \cdot \sum_{i=0}^{e(\mathcal{C})} \binom{n}{i} \cdot (q-1)^i \leq q^n = |\mathcal{X}^n|.$$

Proof. The spheres $\mathcal{B}_{e(\mathcal{C})}(v)$, for $v \in \mathcal{C}$, are pairwise disjoint. ♯

b) We consider a slightly weaker notion: Letting $e = e(\mathcal{C})$, let $f \in \{e, \ldots, n\}$ such that $\mathcal{B}_e(v) \cap \mathcal{B}_f(w) = \emptyset$, for $v \neq w \in \mathcal{C}$; in this case $\mathcal{C}$ is called **f-error detecting**. Assume that $f \geq e+2$; then let $v \neq w \in \mathcal{C}$ such that $u \in \mathcal{B}_{e+1}(v) \cap \mathcal{B}_{e+1}(w)$; hence there is $t \in \mathcal{B}_e(v)$ such that $d(t,u) = 1$, thus the triangle inequality implies $t \in \mathcal{B}_{e+2}(w)$, a contradiction. Hence actually we do not have too much choice, namely $f \in \{e, e+1\}$.

The above notions have the following relevance for decoding purposes: Let still $e = e(\mathcal{C})$ and let $\mathcal{C}$ be 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, we have $d(u,w) \geq e+1$ for $v \neq w \in \mathcal{C}$. Thus, if at most f errors have occurred, it is detected how many of them have occurred (if any at all). Moreover, if at most e errors have occurred, then u is correctly decoded by unique nearest neighbor decoding; while if $f = e+1$ and precisely $e+1$ errors have occurred, then u is not necessarily uniquely nearest neighbor decodable.

c) Let now $\mathcal{C}$ be non-trivial. Then $d = d(\mathcal{C}) \in \mathbb{N}$ is related to the error correction and detection properties of $\mathcal{C}$ as follows:

Proposition. i) Let $e \in \{0, \ldots, n\}$. Then $\mathcal{C}$ is e-error correcting if and only if $2e+1 \leq d$; thus the packing radius of $\mathcal{C}$ equals $e(\mathcal{C}) = \lfloor \frac{d-1}{2} \rfloor$.
ii) Let $e = e(\mathcal{C})$ and $f \in \{e, e+1\}$. Then $\mathcal{C}$ is f-error detecting if and only if $2f \leq d$; thus $\mathcal{C}$ is $(e+1)$-error detecting if and only if d is even, where $e+1 = \frac{d}{2}$.

Proof. i) Assume that $u \in \mathcal{B}_e(v) \cap \mathcal{B}_e(w)$, for some $v \neq w \in \mathcal{C}$; then the triangle inequality implies $d \leq d(v,w) \leq d(v,u) + d(u,w) \leq 2e$.

Conversely, assume that $d \leq 2e$, and let $v \neq w \in \mathcal{C}$ such that $d(v,w) = d$; then let $u \in \mathcal{X}^n$ such that $d(v,u) = \lfloor \frac{d}{2} \rfloor$ and $d(u,w) = \lfloor \frac{d+1}{2} \rfloor$. Distinguishing the cases d even and d odd, we observe that $d(v,u) + d(u,w) = d$, and both $\lfloor \frac{d}{2} \rfloor \leq e$ and $\lfloor \frac{d+1}{2} \rfloor \leq e$, entailing $u \in \mathcal{B}_e(v) \cap \mathcal{B}_e(w)$.

ii) We may assume that $f = e + 1$. Assume that $u \in \mathcal{B}_e(v) \cap \mathcal{B}_{e+1}(w)$, for some $v \neq w \in \mathcal{C}$; then the triangle inequality implies $d \leq d(v,w) \leq d(v,u) + d(u,w) \leq 2e + 1 = 2f - 1$.

Conversely, assume that $d \leq 2f - 1 = 2e + 1$, and let $v \neq w \in \mathcal{C}$ such that $d(v,w) = d$; then let $u \in \mathcal{X}^n$ such that $d(v,u) \leq e$ and $d(u,w) \leq e + 1$, as well as $d(v,u) + d(u,w) = d$, hence $u \in \mathcal{B}_e(v) \cap \mathcal{B}_{e+1}(w)$. ♯

Example: Parity check codes. Let $q \geq 2$, let $n \geq 2$, and let $\mathcal{C} := \{v \in \mathbb{Z}_q^n; vw^{\mathrm{tr}} = 0 \in \mathbb{Z}_q\}$ be a parity check code, where $w = [w_1, \ldots, w_n] \in (\mathbb{Z}_q^*)^n$. Hence for any $[x_1, \ldots, x_{n-1}] \in \mathbb{Z}_q^{n-1}$ and $x_{n-1} \neq x'_{n-1} \in \mathbb{Z}_q$, there are $x_n, x'_n \in \mathbb{Z}_q$ such that both $v := [x_1, \ldots, x_{n-2}, x_{n-1}, x_n] \in \mathcal{C}$ and $[x_1, \ldots, x_{n-2}, x'_{n-1}, x'_n] \in \mathcal{C}$, thus we have $d(\mathcal{C}) \leq 2$.

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

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

b) If $c(\mathcal{C}) = e(\mathcal{C}) = \lfloor \frac{d-1}{2} \rfloor =: e$, 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 $m \cdot \sum_{i=0}^{e} \binom{n}{i} \cdot (q-1)^i = q^n$.

As for the existence of perfect codes, if d is even then $c(\mathcal{C}) \geq f = e + 1$ implies that there are none in this case. The picture changes if d is odd, that is $e = \frac{d-1}{2}$, but still perfect codes are rare; fulfilling the Hamming bound is not sufficient. In particular, $\mathcal{C}$ is perfect with $d = 1$, that is $e = 0$, if and only if $\mathcal{C} = \mathcal{X}^n$.

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: Repetition codes. Let $\mathcal{C} := \{[x,\dots,x] \in \mathcal{X}^n; x \in \mathcal{X}\}$. Then $\mathcal{C}$ has information rate $\rho(\mathcal{C}) = \frac{\log_q(q)}{n} = \frac{1}{n}$ and minimum distance $d(\mathcal{C}) = n$, that is an (n,q,n)-code; it is $\lfloor \frac{n-1}{2} \rfloor$-error correcting, if n is even it is $\frac{n}{2}$-error detecting.

In particular, if $\mathcal{X} = \mathbb{F}_2$ then we have $\mathcal{C} = \{0_n, 1_n\} \leq \mathbb{F}_2^n$, an $[n,1,n]$-code. Since for $v \in \mathbb{F}_2^n$ 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 (5.3) and (6.2), by shortening with respect to the 4-th component, see (5.8).)

Thus $\mathcal{C}$ has dimension $k = \dim_{\mathbb{F}_2}(\mathcal{C}) = 3$ and minimum distance $d = d(\mathcal{C}) = \mathrm{wt}(\mathcal{C}) = 3$, that is $\mathcal{C}$ is a $[6,3,3]$-code. Using $q=2$ and $e = e(\mathcal{C}) = \frac{d-1}{2} = 1$, 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 $u = [0,0,0,*,*,*]$. Moreover, we may assume that $\mathrm{wt}(u) \geq 3$, which leaves $u = [0,0,0,1,1,1]$. For the latter we indeed have $d(u,[0,0,1,1,1,0,]) = 2$. ♯

(4.5) Theorem: van Lint [1971], Tietäväinen, Zinovev–Leontev [1973]. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a perfect $(n,m,2e+1)$-code, where $e \in \mathbb{N}$; in particular q is a prime power, and 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 (16.1);
iii) the **ternary Golay** $[11,6,5]$**-code** $\mathcal{G}_{11}$, see (16.2).

b) If $e = 1$, then $n = \frac{q^k-1}{q-1}$ and $m = q^{n-k}$, for some $k \geq 2$. 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 (perfect) non-linear codes having the parameters of Hamming codes, and their classification still is an open problem. ♯

5 Linear codes

(5.1) 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, while 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 map $\mathcal{C} \to \mathbb{F}_q^k \colon [z_1, \ldots, z_n] \mapsto [z_1, \ldots, z_k]$ onto the first k positions is a bijection; because of this the encoding is called **systematic** on the information symbols.

(5.2) 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 $\mathbb{F}_q^n$. The latter is symmetric and non-degenerate, and we have $\langle vM, w\rangle = \langle v, wM^{\mathrm{tr}}\rangle$, for $v, w \in \mathbb{F}_q^n$ and $M \in \mathbb{F}_q^{n\times n}$.

Given a code $\mathcal{C} \leq \mathbb{F}_q^n$, its orthogonal space $\mathcal{C}^\perp := \{v \in \mathbb{F}_q^n; \langle v, w\rangle = 0 \in \mathbb{F}_q \text{ for } w \in \mathcal{C}\} \leq \mathbb{F}_q^n$ 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**, if equality $\mathcal{C} = \mathcal{C}^\perp$ holds 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 necessarily even.

b) If $G \in \mathbb{F}_q^{k\times n}$ is a generator matrix of $\mathcal{C}$, then we have $\mathcal{C}^\perp = \{v \in \mathbb{F}_q^n; Gv^{\mathrm{tr}} = 0_k^{\mathrm{tr}}\} = \{v \in \mathbb{F}_q^n; vG^{\mathrm{tr}} = 0_k\}$. Hence if $H \in \mathbb{F}_q^{(n-k)\times n}$ is a generator matrix of $\mathcal{C}^\perp$, then $\mathcal{C} = (\mathcal{C}^\perp)^\perp = \{v \in \mathbb{F}_q^n; vH^{\mathrm{tr}} = 0_{n-k}\} = \ker(H^{\mathrm{tr}}) \leq \mathbb{F}_q^n$. Thus H is called a **check matrix** of $\mathcal{C}$, and instead of using a generator matrix, $\mathcal{C}$ can likewise 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 $\mathcal{C}^\perp$, being called a **standard** check matrix of $\mathcal{C}$: We have $\mathrm{rk}_{\mathbb{F}_q}(H) = n - k$, and $HG^{\mathrm{tr}} = [-A^{\mathrm{tr}} \mid E_{n-k}] \cdot \left[\frac{E_k}{A^{\mathrm{tr}}}\right] = -A^{\mathrm{tr}} \cdot E_k + E_{n-k} \cdot A^{\mathrm{tr}} = 0 \in \mathbb{F}_q^{(n-k)\times k}$.

Proposition. 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 $\mathcal{C}^\perp$.

Proof. It suffices to show one direction. If $\mathcal{C}'$ is linearly equivalent to $\mathcal{C}$, then there is a monomial matrix $M \in \mathrm{Isom}_n(\mathbb{F}_q)$ such that $\mathcal{C}' = \mathcal{C} \cdot M$. Then we have $\mathcal{C}' \cdot (HM^{-\mathrm{tr}})^{\mathrm{tr}} = \mathcal{C} \cdot MM^{-1}H^{\mathrm{tr}} = \{0\}$. Since $\mathrm{rk}_{\mathbb{F}_q}(HM^{-\mathrm{tr}}) = \mathrm{rk}_{\mathbb{F}_q}(H) = n - k$, we conclude that $HM^{-\mathrm{tr}} \in \mathbb{F}_q^{(n-k)\times n}$ is a check matrix of $\mathcal{C}'$. Thus $HM^{-\mathrm{tr}}$ is a generator matrix of $(\mathcal{C}')^\perp$. Since H is a generator matrix of $\mathcal{C}^\perp$, this implies that $(\mathcal{C}')^\perp = \mathcal{C}^\perp \cdot M^{-\mathrm{tr}}$, where $M^{-\mathrm{tr}} \in \mathrm{Isom}_n(\mathbb{F}_q)$ indeed. ♯

(5.3) Example: Hamming code. Let the binary **Hamming code** $\mathcal{H} \leq \mathbb{F}_2^7$ be given by the generator matrix $G \in \mathbb{F}_2^{4\times 7}$, 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 (5.4) below, we get $d = d(\mathcal{H}) = 3$, thus $\mathcal{H}$ is a $[7, 4, 3]$-code, hence $e = e(\mathcal{H}) = \frac{d-1}{2} = 1$; note that d cannot be read off directly from the generator matrix G.

Moreover, we have $m \cdot \sum_{i=0}^{e} \binom{7}{i} = 16 \cdot (1 + 7) = 128 = 2^7 = |\mathbb{F}_2^7|$, thus by the

Hamming bound we conclude that $\mathcal{H}$ is perfect; see also (6.2) and (4.5).

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

(5.4) Generalized check matrices a) Let $\mathcal{C} \leq \mathbb{F}_q^n$ be an $[n, k, d]$-code, let $\mathbb{F}_q \subseteq \mathbb{F}$ be a finite field extension of degree $f := [\mathbb{F} \colon \mathbb{F}_q] \in \mathbb{N}$, and let $H \in \mathbb{F}^{r \times n}$, for some $r \in \mathbb{N}_0$, be a **generalized check matrix** of $\mathcal{C}$, that is $\mathcal{C} = \ker(H^{\mathrm{tr}}) \cap \mathbb{F}_q^n$ is the associated **subfield subcode** of $\ker(H^{\mathrm{tr}}) \leq \mathbb{F}^n$.

Theorem. Let $\mathcal{C}$ be non-trivial, hence we have $d \in \mathbb{N}$. 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.

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 its columns; note that such a d' indeed exists.

In order to show that $d \geq d'$, we may assume that $d' \geq 2$. Let $0 \neq v \in \mathbb{F}_q^n$ such that $\mathrm{wt}(v) \leq d' - 1$. Hence $vH^{\mathrm{tr}} \in \mathbb{F}^r$ 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, 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 $vH^{\mathrm{tr}} = 0 \in \mathbb{F}^r$, thus we have $v \in \mathcal{C}$, implying that $d \leq d'$. ♯

b) So far, we do not require that H has full $\mathbb{F}$-rank. But by taking an $\mathbb{F}$-basis of the row space of H we may assume that $\mathrm{rk}_{\mathbb{F}}(H) = r$.

Then, in the particular case $\mathbb{F} = \mathbb{F}_q$ we get $r = \mathrm{rk}_{\mathbb{F}_q}(H) = n - \dim_{\mathbb{F}_q}(\ker(H^{\mathrm{tr}})) = n - k$, so that H indeed is a check matrix of $\mathcal{C}$. Thus this yields a new proof of the Singleton bound $d - 1 \leq r = n - k$ for linear codes.

In general, since $\mathcal{C}$ is $\mathbb{F}_q$-free, we have $\dim_{\mathbb{F}}(\mathcal{C} \otimes_{\mathbb{F}_q} \mathbb{F}) = \dim_{\mathbb{F}_q}(\mathcal{C})$. Hence for any $\mathbb{F}$-subspace $V \leq \mathbb{F}^n$ from $(V \cap \mathbb{F}_q^n) \otimes_{\mathbb{F}_q} \mathbb{F} \leq V$ we obtain $\dim_{\mathbb{F}_q}(V \cap \mathbb{F}_q^n) \leq \dim_{\mathbb{F}}(V)$; but note that equality not necessarily holds.

Proposition. We have $n - fr \leq k \leq n - r$.

Proof. i) The second inequality follows from $k = \dim_{\mathbb{F}_q}(\mathcal{C}) = \dim_{\mathbb{F}}(\mathcal{C} \otimes_{\mathbb{F}_q} \mathbb{F}) \leq \dim_{\mathbb{F}}(\ker(H^{\mathrm{tr}})) = n - \mathrm{rk}_{\mathbb{F}}(H) = n - r$.

ii) We turn to the first inequality: Let $\mathcal{B} := \{\beta_1, \ldots, \beta_f\} \subseteq \mathbb{F}$ be an $\mathbb{F}_q$-basis, let $[\gamma_1, \ldots, \gamma_n] \in \mathbb{F}^n$ be a row of G, and let $\gamma_j = \sum_{k=1}^{f} c_{kj}\beta_k$ be its decomposition into $\mathcal{B}$, where $c_{kj} \in \mathbb{F}_q$. Then, given $[x_1, \ldots, x_n] \in \mathbb{F}_q^n$, the check condition

$$0 = \sum_{j=1}^{n} x_j\gamma_j = \sum_{j=1}^{n}\sum_{k=1}^{f} x_j c_{kj}\beta_k = \sum_{k=1}^{f} \big(\sum_{j=1}^{n} x_j c_{kj}\big)\beta_k$$

is equivalent to the f-set of conditions $\sum_{j=1}^{n} x_j c_{kj} = 0$, for $k \in \{1, \ldots, f\}$.

Hence the **blow-up matrix** $H^\vee \in \mathbb{F}_q^{fr\times n}$, obtained by replacing any entry by the column of its coefficients with respect to $\mathcal{B}$, is a generalized check matrix of $\mathcal{C}$ as well, thus $k = \dim_{\mathbb{F}_q}(\mathcal{C}) = \dim_{\mathbb{F}_q}(\ker(H^{\vee\mathrm{tr}})) = n - \mathrm{rk}_{\mathbb{F}_q}(H^\vee) \geq n - fr$. ♯

(5.5) 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 word $v \in \mathbb{F}_q^n$, we proceed as follows: Let $w \in \mathcal{C}$ be the code word sent, and let $u := v - w \in \mathbb{F}_q^n$ be the associated **error vector**, thus we have $d(v, w) = \mathrm{wt}(u)$. For the syndromes we have $vH^{\mathrm{tr}} = (w+u)H^{\mathrm{tr}} = uH^{\mathrm{tr}} \in \mathbb{F}_q^{n-k}$, that is $\overline{v} = \overline{u} \in \mathbb{F}_q^n/\mathcal{C}$. Hence v is uniquely nearest neighbor decodable if and only if the coset $\overline{v} \in \mathbb{F}_q^n/\mathcal{C}$ 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 $\overline{v} \in \mathbb{F}_q^n/\mathcal{C}$ having minimum weight are called its **coset leaders**. In general coset leaders are not unique, but 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 $\overline{v} \in \mathbb{F}_q^n/\mathcal{C}$ possesses an element $u \in \mathbb{F}_q^n$ of weight $\mathrm{wt}(u) \leq e$, then u is its unique coset leader.

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**, in general is an NP-hard problem. Hence the aim is to find codes having fast decoding algorithms.

(5.6) Example: Syndrome decoding of 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 weight vector $0 \neq w :=$

$[w_1, \ldots, w_n] \in \mathbb{F}_q^n$, where up to linear equivalence we may assume that $w_n = 1$, that is w is in standard form; hence $k := \dim_{\mathbb{F}_q}(\mathcal{C}) = n-1$. 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 $i \in \{1, \ldots, n\}$.

The standard generator matrix is $G := [E_{n-1} \mid -[w_1, \ldots, w_{n-1}]^{\mathrm{tr}}] \in \mathbb{F}_q^{(n-1)\times n}$, saying that $[x_1, \ldots, x_{n-1}] \in \mathbb{F}_q^{n-1}$ is encoded into $[x_1, \ldots, x_{n-1}; -\sum_{i=1}^{n-1} x_i w_i] \in \mathbb{F}_q^n$. For $v = [x_1, \ldots, x_n] \in \mathbb{F}_q^n$ we get the syndrome $vw^{\mathrm{tr}} = \sum_{i=1}^n x_i w_i \in \mathbb{F}_q$.

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. Moreover, $v = [x_1, \ldots, x_n] \in \mathbb{F}_2^n$ has syndrome $vw^{\mathrm{tr}} = \sum_{i=1}^n x_i \in \mathbb{F}_2$, 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 $n \geq 2$, the code $\mathcal{C}$ is an $[n, n-1, 2]$-code, and is the unique such code: Any such code has a weight vector $w \in \mathbb{F}_2^n$ without zero entries, thus $w = 1_n$.

We have $\mathbb{F}_2^n = \overline{0} \,\dot\cup\, \overline{v}$, 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 $\mathbb{F}_2^n \setminus \mathcal{C}$, hence none of the words in $\mathbb{F}_2^n \setminus \mathcal{C}$ is uniquely nearest neighbor decodable.

(5.7) Example: Syndrome decoding of repetition codes. a) Let $\mathcal{C}$ be given by the standard generator matrix $G := [1_n] \in \mathbb{F}_q^n$, hence we have $k := \dim_{\mathbb{F}_q}(\mathcal{C}) = 1$ and $d(\mathcal{C}) = n$. Then $x \in \mathbb{F}_q$ is encoded into $[x, \ldots, x] \in \mathbb{F}_q^n$.

The standard check matrix is $H := [-1_{n-1}^{\mathrm{tr}} \mid E_{n-1}] \in \mathbb{F}_q^{(n-1)\times n}$. 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. Then $\mathcal{C}$ is an $[n, 1, n]$-code, in fact the unique one. The code $\mathcal{C}$ 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 of $\mathcal{C}$ is the weight vector of the even-weight code, hence $\mathcal{C}^\perp$ is the even-weight code of length n.

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

i) For n odd, coset leaders are uniquely given as $[0; w]$ if $\mathrm{wt}(w) \leq \frac{n-1}{2} = e(\mathcal{C}) =: e$, and $[0; w] + 1_n$ if $\mathrm{wt}(w) \geq \frac{n+1}{2} = e + 1$; in both cases the coset leaders have weight at most e.

Thus, if $\mathrm{wt}(v) \leq e$ and $x_1 = 0$, then v has syndrome $[x_2, \ldots, x_n]$, and is decoded to $v + [0; x_2, \ldots, x_n] = 0_n$; if $x_1 = 1$, then v has syndrome $[x_2 + 1, \ldots, x_n + 1]$, and is decoded to $v + ([0; x_2 + 1, \ldots, x_n + 1] + 1_n) = 0_n$.

If $\mathrm{wt}(v) \geq e + 1$ and $x_1 = 0$, then v is decoded to $v + ([0; x_2, \ldots, x_n] + 1_n) = 1_n$; if $x_1 = 1$, then v is decoded to $v + [0; x_2 + 1, \ldots, x_n + 1] = 1_n$.

ii) For n even, coset leaders are uniquely given as $[0;w]$ if $\mathrm{wt}(w) \leq \frac{n}{2}-1 = e(\mathcal{C}) =: e$, and $[0;w]+1_n$ if $\mathrm{wt}(w) \geq \frac{n}{2}+1 = e+2$, where in both cases the coset leaders have weight at most e. But for $\mathrm{wt}(w) = \frac{n}{2} = e+1$ we have $\mathrm{wt}([0;w]) = \mathrm{wt}([0;w]+1_n) = \mathrm{wt}(w)$, in which case coset leaders are not unique.

Thus, if $\mathrm{wt}(v) \leq e$ and $x_1 = 0$, then v is decoded to $v + [0;x_2,\ldots,x_n] = 0_n$; if $x_1 = 1$, then v is decoded to $v + ([0;x_2+1,\ldots,x_n+1]+1_n) = 0_n$.

If $\mathrm{wt}(v) \geq e+2$ and $x_1 = 0$, then v is decoded to $v + ([0;x_2,\ldots,x_n]+1_n) = 1_n$; if $x_1 = 1$, then v is decoded to $v + [0;x_2+1,\ldots,x_n+1] = 1_n$.

But if $\mathrm{wt}(v) = e+1$ then v is not uniquely nearest neighbor decodable. ♯

(5.8) 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 $\mathcal{C}^\bullet := \{[x_1,\ldots,x_{n-1}] \in \mathbb{F}_q^{n-1}; [x_1,\ldots,x_n] \in \mathcal{C}\} \leq \mathbb{F}_q^{n-1}$.

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 $[x_1,\ldots,x_n] \in \mathcal{C}$, then $k^\bullet = k$, amounting to deleting a check symbol.

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 $[x_1,\ldots,x_n] \in \mathcal{C}$, then $d^\bullet = d$.

ii) Extending by adding a check symbol yields the code $\widehat{\mathcal{C}} := \{[x_1,\ldots,x_{n+1}] \in \mathbb{F}_q^{n+1}; [x_1,\ldots,x_n] \in \mathcal{C}, \sum_{i=1}^{n+1} x_i = 0\} \leq \mathbb{F}_q^{n+1}$. Thus we have $\widehat{k} := \dim_{\mathbb{F}_q}(\widehat{\mathcal{C}}) = k$, and $\widehat{\mathcal{C}}$ has check matrix

$$\widehat{H} := \left[\begin{array}{c|c} H & 0_{n-k}^{\mathrm{tr}} \\ \hline 1_n & 1 \end{array}\right] \in \mathbb{F}_q^{(n+1-k)\times(n+1)}.$$

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 last row of $\widehat{H}$ amounts to $\mathrm{wt}(v) \in \mathbb{N}_0$ even, for $v \in \widehat{\mathcal{C}}$. 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.

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

b) i) Expurgating by throwing away certain code words 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 $d' \geq d$.

We have $k-1 \leq k' := \dim_{\mathbb{F}_q}(\mathcal{C}') \leq k$. If $k' = k$ then we have $\mathcal{C}' = \mathcal{C}$, while if $k' = k-1$ then $\mathcal{C}' < \mathcal{C}$ has check matrix

$$H' := \left[\frac{H}{1_n}\right] \in \mathbb{F}_q^{(n-k+1)\times n}.$$

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 code words 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 $\widetilde{d} \leq d$.

We have $k \leq \widetilde{k} := \dim_{\mathbb{F}_q}(\widetilde{\mathcal{C}}) \leq k+1$, where $\widetilde{k} = k+1$ if and only if $1_n \notin \mathcal{C}$. In this case $\widetilde{\mathcal{C}}$ has generator matrix

$$\widetilde{G} := \left[\frac{G}{1_n}\right] \in \mathbb{F}_q^{(k+1)\times n}.$$

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 $v \in \mathbb{F}_2^n$, 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.

iii) If $1_n \in \mathcal{C}$, then we have $1_n \in \mathcal{C}'$ if and only if $\gcd(q,n) > 1$, thus augmenting the expurgated code again yields $\mathcal{C}' \leq \widetilde{(\mathcal{C}')} = \langle \mathcal{C}', 1_n\rangle_{\mathbb{F}_q} \leq \mathcal{C}$; hence we get $\widetilde{(\mathcal{C}')} = \mathcal{C}$ if $\gcd(q,n) = 1$, and $\widetilde{(\mathcal{C}')} = \mathcal{C}'$ if $\gcd(q,n) > 1$.

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 again 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 composition of n**-expurgation** and subsequent puncturing, where the former is given as $\mathcal{C}^{(n)} := \{[x_1, \ldots, x_n] \in \mathcal{C}; x_n = 0\} \leq \mathbb{F}_q^n$, from which puncturing yields the code $\mathcal{C}^\circ := (\mathcal{C}^{(n)})^\bullet = \{[x_1, \ldots, x_{n-1}] \in \mathbb{F}_q^{n-1}; [x_1, \ldots, x_n] \in \mathcal{C}, x_n = 0\} \leq \mathbb{F}_q^{n-1}$. Hence for the minimum distance of $\mathcal{C}^\circ$ we have $d^\circ \geq d$.

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

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

6 Geometric codes

(6.1) Hamming codes [1950]. a) 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 (4.5).

Proposition. Any $[n, n-k, 3]$-code over $\mathbb{F}_q$ is linearly equivalent to $\mathcal{H}_k$.

Proof. Let $H \in \mathbb{F}_q^{k\times n}$ be a check matrix of the code in question. 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 $\mathbf{P}^{k-1}(\mathbb{F}_q)$. ♯

b) 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 $\{1,\dots,q^k-1\}$ 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 with respect to $\mathcal{H}_k$. Since there are $q^{n-(n-k)} - 1 = q^k - 1$ such cosets, this covers all of them. Thus the non-trivial coset leaders are precisely the vectors $xe_1,\dots,xe_n \in \mathbb{F}_q^n$, where $x \in \mathbb{F}_q^*$ and $e_i \in \mathbb{F}_q^n$ is the i-th unit vector, for $i \in \{1,\dots,n\}$.

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 $i \in \{1,\dots,n\}$, letting $H_1 := [1] \in \mathbb{F}_2^{1\times 1}$, we

recursively get

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

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 (5.3); moreover, the $[6,3,3]$-code obtained by shortening $\mathcal{H}_3$ with respect to the 4-th component has been considered in (4.4).

In particular, we observe that for $k \geq 2$ any row of H_k has weight 2^{k-1}, which is even, hence we have $1_n \in \mathcal{H}_k$; thus 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 $\mathcal{H}_k^\perp \leq \mathcal{H}_k$, saying that $\mathcal{H}_k^\perp$ is weakly self-dual: Letting $H_k = [w_1, \ldots, w_k] \in \mathbb{F}_2^{k\times n}$, we get $\langle w_i, w_i\rangle = \mathrm{wt}(w_i) = 2^{k-1} = 0 \in \mathbb{F}_2$, for $i \in \{1,\ldots,k\}$, 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 some of the modifications described in (5.8), see Table 4:

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 $\langle 1_{n+1}, [w_i; 0]\rangle = \mathrm{wt}(w_i) = 0 \in \mathbb{F}_2$, for $i \in \{1,\ldots,k\}$, 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 $(\widehat{\mathcal{H}}_k)^\perp \leq \widehat{\mathcal{H}}_k$, saying that $(\widehat{\mathcal{H}}_k)^\perp$ is weakly self-dual. In particular, since $\dim((\widehat{\mathcal{H}}_3)^\perp) = 4 = \dim(\widehat{\mathcal{H}}_3)$, we conclude that $\widehat{\mathcal{H}}_3$ is a self-dual $[8,4,4]$-code.

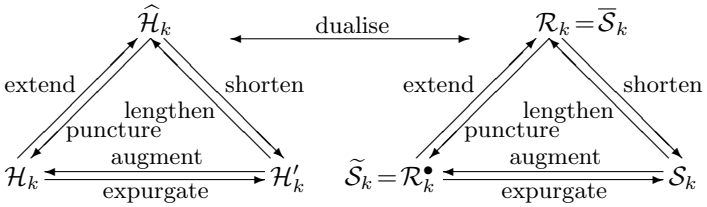
ii) Expurgating $\mathcal{H}_k \leq \mathbb{F}_2^n$ yields the **even-weight Hamming code** $\mathcal{H}'_k \leq \mathbb{F}_2^n$. Since $1_n \in \mathcal{H}_k$ and n is odd, we conclude that $\mathcal{H}'_k$ is an $[n, n-k-1, d']$-code, where $d' \geq 3$, and augmenting $\mathcal{H}'_k$ yields $\widetilde{(\mathcal{H}'_k)}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

For $k \geq 3$ we have $d' = 4$: Since $d' \geq 3$ is even, it suffices to show that $d' \leq 4$. To this end, we consider the 4-subset $\{e_1^{\mathrm{tr}}, e_2^{\mathrm{tr}}, e_3^{\mathrm{tr}}, (e_1+e_2+e_3)^{\mathrm{tr}}\} \subseteq \mathbb{F}_2^{k\times 1}$ of columns of H_k. Summing up the associated columns of $H'_k := \left[\begin{array}{c} H_k \\ \hline 1_n \end{array}\right]$ yields $0 \in \mathbb{F}_2^{(k+1)\times 1}$, thus $\mathcal{H}'_k$ has an $\mathbb{F}_2$-linearly dependent 4-subset of columns. ♯

iii) Finally, 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 $\overline{\mathcal{H}'_k} = \widehat{\mathcal{H}'_k} = \widehat{\mathcal{H}}_k \leq \mathbb{F}_2^{n+1}$ again.

(6.3) Simplex codes. 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

Table 4: Modified binary Hamming codes and simplex codes.



associated **simplex code**. We show that $\mathcal{S}_k$ is an **equidistant** $[n, k, q^{k-1}]$-code:

Lemma. Any $0 \neq v \in \mathcal{S}_k$ has weight $\mathrm{wt}(v) = q^{k-1}$.

Proof. Let $H_k = [w_1, \ldots, w_k]$ again. Then there is $0 \neq [x_1, \ldots, x_k] \in \mathbb{F}_q^k$ such that $v = \sum_{i=1}^k x_i w_i \in \mathbb{F}_q^n$. Now the j-th entry of v, for $j \in \{1, \ldots, n\}$, equals $\langle v, e_j \rangle = \sum_{i=1}^k x_i \langle w_i, e_j \rangle$. Hence the j-th entry of v is zero if and only if the vector $[\langle w_i, e_j \rangle]_i \in \mathbb{F}_q^k$ is an element of $U_v := \langle [x_1, \ldots, x_k] \rangle_{\mathbb{F}_q}^\perp \leq \mathbb{F}_q^k$. The vector $[\langle w_i, e_j \rangle]_i^{\mathrm{tr}} \in \mathbb{F}_q^{k \times 1}$ coincides with the j-th column of H_k. Since $\dim_{\mathbb{F}_q}(U_v) = k - 1$, by construction of H_k there are precisely $\frac{q^{k-1}-1}{q-1}$ columns belonging to U_v. Thus we get $\mathrm{wt}(v) = n - \frac{q^{k-1}-1}{q-1} = \frac{q^k-1}{q-1} - \frac{q^{k-1}-1}{q-1} = q^{k-1}$. ♯

Note that $\sum_{i=0}^{k-1} \lceil \frac{q^{k-1}}{q^i} \rceil = \sum_{i=0}^{k-1} q^{k-i-1} = \sum_{i=0}^{k-1} q^i = \frac{q^k-1}{q-1} = n$ shows that $\mathcal{S}_k$ fulfills the Griesmer bound, see (7.2) below. Likewise $q^k(q^{k-1} - \frac{q^k-1}{q-1} \cdot \frac{q-1}{q}) = q^{k-1} = d$ shows that $\mathcal{S}_k$ fulfills the Plotkin bound, see (7.1) below; actually the latter also implies that $\mathcal{S}_k$ is an equidistant code. (In the sequel of this section we will have to make use of the Griesmer bound already, although its discussion is postponed to the next section.)

Proposition. Any $[n, k, q^{k-1}]$-code over $\mathbb{F}_q$ is linearly equivalent to $\mathcal{S}_k$.

Proof. Let $\mathcal{C}$ be the code in question. We consider $\mathcal{C}^\perp \leq \mathbb{F}_q^n$, having a check matrix $H \in \mathbb{F}_q^{(n-k) \times n}$ of $\mathcal{C}$ as a generator matrix. Letting $d^\perp := d(\mathcal{C}^\perp) \in \mathbb{N}$, we show that $d^\perp \geq 3$: To this end, we assume to the contrary that $d^\perp \leq 2$.

If $d^\perp = 1$, then we may assume that $[0, \ldots, 0, 1] \in \mathcal{C}^\perp$. Thus any word in $\mathcal{C}$ has zero n-th entry, hence the shortened code $\mathcal{C}^\circ$ is an $[n-1, k, q^{k-1}]$-code. Then the Griesmer bound yields $n - 1 \geq \sum_{i=0}^{k-1} \lceil \frac{q^{k-1}}{q^i} \rceil = \frac{q^k-1}{q-1} = n$, a contradiction.

If $d^\perp = 2$, then we may assume that $[0, \ldots, 0, 1, 1] \in \mathcal{C}^\perp$. Thus any word in $\mathcal{C}$ is of the form $[*, \ldots, *, -x, x] \in \mathbb{F}_q^n$, for some $x \in \mathbb{F}_q$. This implies 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 $k-1 \le k^{\circ\circ} \le k$ and $d^{\circ\circ} \ge q^{k-1}$. Hence the Griesmer bound implies $n-2 \ge \sum_{i=0}^{k^{\circ\circ}-1} \lceil \frac{d^{\circ\circ}}{q^i} \rceil \ge \sum_{i=0}^{k^{\circ\circ}-1} \lceil \frac{q^{k-1}}{q^i} \rceil = \frac{q^k - q^{k-k^{\circ\circ}}}{q-1} = n - \frac{q^{k-k^{\circ\circ}}-1}{q-1} \ge n-1$, a contradiction again.

Hence we have $d^\perp \ge 3$. Since $\dim_{\mathbb{F}_q}(\mathcal{C}^\perp) = n-k > 0$, and any $[n, n-k, 3]$-code fulfills the Hamming bound, we conclude that $d^\perp = 3$. Hence $\mathcal{C}^\perp$, being a $[n, n-k, 3]$-code, is linearly equivalent to the Hamming code $\mathcal{H}_k = \mathcal{S}_k^\perp$. ♯

(6.4) Binary simplex codes and Reed-Muller codes. Keeping the notation of (6.3), let $q := 2$, hence for $k \ge 2$ 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 $\mathcal{S}_2 = \{0_3, [0,1,1], [1,0,1], [1,1,0]\} \le \mathbb{F}_2^3$ can be seen (abusively) 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 some of the modifications described in (5.8), see Table 4: Since the words 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 having $H'_k := \left[\begin{array}{c} H_k \\ \hline 1_n \end{array}\right] \in \mathbb{F}_2^{(k+1)\times n}$ as a generator matrix, thus $(\mathcal{H}'_k)^\perp = \widetilde{\mathcal{S}}_k$ is the augmented simplex code. Expurgating $\widetilde{\mathcal{S}}_k$ yields $(\widetilde{\mathcal{S}}_k)' = (\mathcal{S}_k \,\dot\cup\, (1_n + \mathcal{S}_k))' = \mathcal{S}_k$ again.

Since for $0 \ne v \in \mathcal{S}_k$ 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 $\widetilde{\mathcal{S}}_k$ is a $[2^k-1, k+1, 2^{k-1}-1]$-code. 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** $\mathcal{R}_k := (\widehat{\mathcal{H}}_k)^\perp \le \mathbb{F}_2^{n+1}$ with generator matrix $\widehat{H}_k = \left[\begin{array}{c|c} H_k & 0_k^{\mathrm{tr}} \\ \hline 1_n & 1 \end{array}\right] \in \mathbb{F}_2^{(k+1)\times(n+1)}$.

Thus $\mathcal{R}_k$ is obtained by extending $\widetilde{\mathcal{S}}_k$, so that $\mathcal{R}_k = \widehat{\widetilde{\mathcal{S}}}_k = \overline{\mathcal{S}}_k$; note that $1_{2^k} \in \mathcal{R}_k$. Shortening $\mathcal{R}_k$ yields $\mathcal{R}_k^\circ = (\overline{\mathcal{S}}_k)^\circ = (\widehat{\widetilde{\mathcal{S}}}_k)^\circ = (\widetilde{\mathcal{S}}_k)' = \mathcal{S}_k$ again.

Since $\mathcal{R}_k$ extends $\widetilde{\mathcal{S}}_k$, which has minimum distance $2^k - 1$, we conclude that $\mathcal{R}_k$ has minimum distance 2^{k-1}, thus $\mathcal{R}_k$ is a $[2^k, k+1, 2^{k-1}]$-code. In particular, from $\widetilde{\mathcal{S}}_2 = \mathbb{F}_2^3$ we infer that $\mathcal{R}_2$ is the even-weight $[4,3,2]$-code, while we conclude that $\mathcal{R}_3 = \widehat{\mathcal{H}}_3^\perp = \widehat{\mathcal{H}}_3$ is the extended Hamming $[8,4,4]$-code.

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 $\mathcal{R}_k^\bullet = (\overline{\mathcal{S}}_k)^\bullet = (\widehat{\widetilde{\mathcal{S}}}_k)^\bullet = \widetilde{\mathcal{S}}_k$ again, and extending $\mathcal{R}_k^\bullet$ yields $\widehat{(\mathcal{R}_k^\bullet)} = \widehat{\widetilde{\mathcal{S}}}_k = \overline{\mathcal{S}}_k = \mathcal{R}_k$ again. ♯

Proposition. Any binary $[2^k, k+1, 2^{k-1}]$-code is linearly equivalent to $\mathcal{R}_k$.

Proof. Let $\mathcal{C}$ be the code in question. We use induction on $k \in \mathbb{N}$, where $\mathcal{R}_1 := \mathbb{F}_2^2$ is the unique $[2,2,1]$-code. Hence 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 $d^\circ \geq 2^{k-1}$. Assume that $d^\circ > 2^{k-1}$, then we have $\sum_{i=0}^{k-1} \lceil \frac{d^\circ}{2^i} \rceil > \sum_{i=0}^{k-1} \frac{2^{k-1}}{2^i} = \sum_{i=0}^{k-1} 2^i = 2^k - 1$, contradicting the Griesmer bound. Hence we have $d^\circ = 2^{k-1}$.

Thus we may assume that $\mathcal{C}^\circ = \mathcal{S}_k$. This also shows that shortening with respect to any component yields a code linearly equivalent to the simplex code, thus any word in $\mathcal{C} \setminus \{0_{2^k}, 1_{2^k}\}$ has even weight 2^{k-1}. Hence we have $\widehat{(\mathcal{C}^\circ)} \leq \mathcal{C}$.

We proceed to show that $1_{2^k} \in \mathcal{C}$; then we conclude that $\mathcal{C} = \langle \widehat{(\mathcal{C}^\circ)}, 1_{2^k} \rangle_{\mathbb{F}_2} = \widehat{\mathcal{S}}_k + \langle \widehat{1_{2^k-1}} \rangle_{\mathbb{F}_2} = \widehat{\widetilde{\mathcal{S}}}_k = \overline{\mathcal{S}}_k = \mathcal{R}_k$. To this end, let

$$G = \left[\begin{array}{c|c} 1_{2^{k-1}} & 0_{2^{k-1}} \\ \hline * \mid 0_k^{\mathrm{tr}} & G^* \end{array}\right] \in \mathbb{F}_q^{(k+1)\times 2^k}$$

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 the rows of $G^* \in \mathbb{F}_q^{k\times 2^{k-1}}$.

Assume that $d^* < 2^{k-2}$, then letting $0 \neq v \in \mathcal{C}^*$ have minimum weight, we infer that there is $[w \mid v] \in \mathcal{C}$ such that, possibly by adding $[1_{2^{k-1}} \mid 0_{2^{k-1}}]$, we may assume that $\mathrm{wt}(w) \leq 2^{k-2}$; thus we have $\mathrm{wt}([w \mid v]) < 2^{k-1}$, a contradiction. (This is reminiscent of the **Helgert-Stinaff construction** used in (7.2).)

Assume that $2^{k-1} \geq d^* > 2^{k-2}$, then $\sum_{i=0}^{k-1} \lceil \frac{d^*}{2^i} \rceil = \lceil \frac{d^*}{2^{k-1}} \rceil + \sum_{i=0}^{k-2} \lceil \frac{d^*}{2^i} \rceil > 1 + \sum_{i=0}^{k-2} \frac{2^{k-2}}{2^i} = 1 + \sum_{i=0}^{k-2} 2^i = 2^{k-1}$, contradicting the Griesmer bound. Hence we have $d^* = 2^{k-2}$, so that $\mathcal{C}^*$ is a $[2^{k-1}, k, 2^{k-2}]$-code.

Thus we may assume by induction that $\mathcal{C}^* = \mathcal{R}_{k-1}$. Hence we have $1_{2^{k-1}} \in \mathcal{C}^*$, thus $\mathcal{C}$ contains a word $[*; 0 \mid 1_{2^{k-1}}] + [1_{2^{k-1}} \mid 0_{2^{k-1}}] = [*; 1 \mid 1_{2^{k-1}}]$, which has weight at least $2^{k-1}+1$, hence equals 1_{2^k}. ♯

(6.5) Remark. The Reed-Muller codes $\mathcal{R}_k$ are **Hadamard** codes, being defined by **Hadamard matrices of Sylvester type**, see Exercise (24.18), 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 $\delta(\mathcal{R}_k) := \frac{d(\mathcal{R}_k)}{2^k} = \frac{2^{k-1}}{2^k} = \frac{1}{2}$ this outweighs their small information rate $\rho(\mathcal{R}_k) = \frac{\dim_{\mathbb{F}_2}(\mathcal{R}_k)}{2^k} = \frac{k+1}{2^k}$, 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 small 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.

(6.6) 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.

Proposition. Let $\mathcal{C}'$ be an $[n, k', d']$-code, and let $\mathcal{C}''$ be an $[n, k'', d'']$-code, both over the field $\mathbb{F}_q$ with q elements. Then their **Plotkin sum**

$$\mathcal{C} := \mathcal{C}' \ltimes \mathcal{C}'' := \{[v \mid v + w] \in \mathbb{F}_q^{2n}; v \in \mathcal{C}', w \in \mathcal{C}''\} \leq \mathbb{F}_q^{2n}$$

is a $[2n, k' + k'', \min\{2d', d''\}]$-code.

Proof. The $\mathbb{F}_q$-linear map $\mathcal{C}' \oplus \mathcal{C}'' \to \mathcal{C} \colon [v, w] \mapsto [v \mid v + w]$ 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 $\mathrm{wt}(u) = \mathrm{wt}(v) + \mathrm{wt}(v + w) \geq \mathrm{wt}(v) + (\mathrm{wt}(v) + \mathrm{wt}(w) - 2 \cdot |\mathrm{supp}(v) \cap \mathrm{supp}(w)|) \geq \mathrm{wt}(w) \geq d''$. ♯

We are now prepared to present a straightforward construction of the (binary) **Reed-Muller code** $\mathcal{R}_k^{(r)} \leq \mathbb{F}_2^{2^k}$ **of order** $r \in \mathbb{N}_0$, for $k \in \mathbb{N}_0$ such that $k \geq r$:

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 $k \geq 2$, and recalling that $\mathcal{R}_k^{(0)}$ and $\mathcal{R}_k^{(k)}$ are already defined, 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 $\dim_{\mathbb{F}_2}(\mathcal{R}_k^{(k)}) = 2^k = \sum_{i=0}^{k}\binom{k}{i}$. For $k \geq 2$ and $r \in \{1, \dots, k-1\}$ we get $\dim_{\mathbb{F}_2}(\mathcal{R}_k^{(r)}) = \sum_{i=0}^{r}\binom{k-1}{i} + \sum_{i=0}^{r-1}\binom{k-1}{i} = 1 + \sum_{i=0}^{r-1}\left(\binom{k-1}{i+1} + \binom{k-1}{i}\right) = 1 + \sum_{i=0}^{r-1}\binom{k}{i+1} = \sum_{i=0}^{r}\binom{k}{i}$,

We have $d(\mathcal{R}_k^{(0)}) = 2^k$ and $d(\mathcal{R}_k^{(k)}) = 1$. For $k \geq 2$ and $r \in \{1, \dots, k-1\}$ we get $d(\mathcal{R}_k^{(r)}) = \min\{2 \cdot d(\mathcal{R}_{k-1}^{(r)}), d(\mathcal{R}_{k-1}^{(r-1)})\} = \min\{2 \cdot 2^{k-r-1}, 2^{k-r}\} = 2^{k-r}$. ♯

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 $k \geq 2$.

(6.7) Boolean functions. We present an alternative construction of higher order Reed-Muller codes. To do so, we need some preparations first:

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 $[x_1, \dots, x_k] \in \mathbb{F}_2^k$ with the integer $\sum_{i=1}^{k} x_i \cdot 2^{i-1} \in \{0, \dots, 2^k - 1\}$, 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 its values.

Identifying the values 0 and 1 with the Boolean values 'false' and 'true', respectively, 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 $i \in \{1, \dots, k\}$ let $p_i\colon \mathbb{F}_2^k \to \mathbb{F}_2$ be the projection onto the i-th component, thus we have $p_1 = [0, 1, \dots, 0, 1] \in \mathbb{F}_2^{2^k}$, $p_2 = [0, 0, 1, 1, \dots, 0, 0, 1, 1] \in \mathbb{F}_2^{2^k}$, and so forth, up to $p_k = [0, \dots, 0, 1, \dots, 1] \in \mathbb{F}_2^{2^k}$. For $\mathcal{I} \subseteq \{1, \dots, k\}$ let $p_{\mathcal{I}} := \prod_{i \in \mathcal{I}} p_i \in \mathbb{F}_2^{2^k}$, where $p_\emptyset := 1_{2^k}$, and $|\mathcal{I}| \in \{0, \dots, k\}$ 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 $\mathcal{I} \neq \{1, \dots, k\}$ the vector $p_{\mathcal{I}} \in \mathbb{F}_2^{2^k}$ has even weight, while $p_{\{1,\dots,k\}} = [0, \dots, 0, 1] \in \mathbb{F}_2^{2^k}$.

Then, using the disjunctive normal form of Boolean logic, any Boolean function can be written as $\sum_{\mathcal{I} \subseteq \{1,\dots,k\}} a_{\mathcal{I}} p_{\mathcal{I}}$, where $a_{\mathcal{I}} \in \mathbb{F}_2$. Since there are $2^{2^k} = |\mathbb{F}_2^{2^k}|$ such sums, we deduce that these are pairwise distinct, thus $\{p_{\mathcal{I}} \in \mathbb{F}_2^{2^k}; \mathcal{I} \subseteq \{1, \dots, k\}\}$ is an $\mathbb{F}_2$-basis of the space of Boolean functions in k variables.

Example. 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 $\{3, 4\} \subseteq \{0, \dots, 7\}$, 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

$$\begin{array}{rcrcl} & & p_3 & = & [0,0,0,0,1,1,1,1], \\ & & p_2 & = & [0,0,1,1,0,0,1,1], \\ & & p_1 & = & [0,1,0,1,0,1,0,1], \\ p_{\{1,2\}} & = & p_1p_2 & = & [0,0,0,1,0,0,0,1], \\ p_{\{1,3\}} & = & p_1p_3 & = & [0,0,0,0,0,1,0,1], \\ p_{\{2,3\}} & = & p_2p_3 & = & [0,0,0,0,0,0,1,1] \end{array}$$

Thus we get $p_3 + p_{\{1,2\}} + p_{\{1,3\}} + p_{\{2,3\}} = [0,0,0,1,1,0,0,0] = p$.

(6.8) Higher order Reed-Muller codes again. Now let $r \in \{0,\dots,k\}$, and let $\mathcal{F}_k^{(r)} := \langle p_{\mathcal{I}} \in \mathbb{F}_2^{2^k}; |\mathcal{I}| \leq r \rangle_{\mathbb{F}_2} \leq \mathbb{F}_2^{2^k}$ be the linear code spanned by the Boolean functions in k variables of degree at most r. In particular, we have $\dim_{\mathbb{F}_2}(\mathcal{F}_k^{(r)}) = \sum_{i=0}^{r} \binom{k}{i}$, but we have no clue about its minimum distance.

Proposition. We have $\mathcal{F}_k^{(r)} = \mathcal{R}_k^{(r)}$, the (binary) Reed-Muller code of order r.

Proof. We have $\mathcal{F}_k^{(0)} = \langle 1_{2^k} \rangle_{\mathbb{F}_2} = \mathcal{R}_k^{(0)}$, and $\dim_{\mathbb{F}_2}(\mathcal{F}_k^{(r)}) = \sum_{i=0}^{k} \binom{k}{i} = 2^k$ implies that $\mathcal{F}_k^{(k)} = \mathbb{F}_2^{2^k} = \mathcal{R}_k^{(k)}$. Now we proceed by induction on $k \in \mathbb{N}$:

For $k \geq 2$ and $r \in \{1,\dots,k-1\}$, any $p \in \mathcal{F}_k^{(r)}$ can be written as $p = \sum_{\mathcal{I} \subseteq \{1,\dots,k\}, |\mathcal{I}| \leq r} a_{\mathcal{I}} p_{\mathcal{I}} = \sum_{k \notin \mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I}} + p_k \cdot \sum_{k \in \mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I} \setminus \{k\}}$. The first summand $\sum_{k \notin \mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I}}$ 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 $[v \mid v]$. In the second summand, $\sum_{k \in \mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I} \setminus \{k\}}$ 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 $\mathcal{F}_k^{(r)} = \mathcal{R}_{k-1}^{(r)} \ltimes \mathcal{R}_{k-1}^{(r-1)} = \mathcal{R}_k^{(r)}$. ♯

Accordingly, the (binary) **punctured Reed-Muller code** $(\mathcal{R}_k^{(r)})^\bullet$ **of order** $r \in \mathbb{N}_0$, for $k \in \mathbb{N}$ such that $k \geq r$, is given by puncturing $\mathcal{F}_k^{(r)}$ in the first component, that is the component associated with the point $0_k \in \mathbb{F}_2^k$.

Moreover, this allows us to read off further properties of $\mathcal{R}_k^{(r)} = \mathcal{F}_k^{(r)}$:

i) 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 $k \geq r \geq 0$.

ii) For $k \geq 1$, since $\mathcal{R}_k^{(k-1)}$ has an $\mathbb{F}_2$-basis consisting of even-weight vectors, and $\sum_{i=0}^{k-1} \binom{k}{i} = 2^k - 1$, it follows that $\mathcal{R}_k^{(k-1)}$ is the even-weight $[2^k, 2^k - 1, 2]$-code.

iii) We consider the associated dual codes, noting that $(\mathcal{R}_k^{(k)})^\perp = (\mathbb{F}_2^{2^k})^\perp = \{0\}$:

Proposition. For $k > r \geq 0$ we have $(\mathcal{R}_k^{(r)})^\perp = \mathcal{R}_k^{(k-r-1)}$.

Proof. 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 $\langle p, q \rangle = \mathrm{wt}(pq) = 0 \in \mathbb{F}_2$. 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, for $k \geq 1$, from $(\mathcal{R}_k^{(k-1)})^\perp = \mathcal{R}_k^{(0)} = \langle 1_{2^k} \rangle_{\mathbb{F}_2}$ we indeed recover the even-weight code $\mathcal{R}_k^{(k-1)} = (\langle 1_{2^k} \rangle_{\mathbb{F}_2})^\perp$.

For $k \geq 2$, from $(\mathcal{R}_k^{(k-2)})^\perp = \mathcal{R}_k^{(1)}$, which is linearly equivalent to $\mathcal{R}_k$, we infer that $\mathcal{R}_k^{(k-2)}$ is linearly equivalent to $(\mathcal{R}_k)^\perp = \widehat{\mathcal{H}}_k$, the extended Hamming code.

III MacWilliams

7 Bounds for codes

(7.1) 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 $m \cdot (d - n \cdot \frac{q-1}{q}) \leq d$.

If equality holds then $d(v, w) = d$, for $v \neq w \in \mathcal{C}$, that is $\mathcal{C}$ is **equidistant**.

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 $v \neq w \in \mathcal{C}$, 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 $m_j := [m_{ij}]_i \in \mathbb{R}^q$ and $1_q \in \mathbb{R}^q$, yields

$$m^2 = \big(\sum_{i=1}^{q} m_{ij}\big)^2 = (1_q \cdot m_j)^2 \leq (1_q \cdot 1_q)(m_j \cdot m_j) = q \cdot \sum_{i=1}^{q} m_{ij}^2,$$

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

$$\Delta = \sum_{j=1}^{n} \sum_{i=1}^{q} m_{ij}(m - m_{ij}) = \sum_{j=1}^{n} \big(m^2 - \sum_{i=1}^{q} m_{ij}^2\big) \leq \sum_{j=1}^{n} m^2(1 - \frac{1}{q}) = nm^2 \cdot \frac{q-1}{q}.$$

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 $\Delta = m(m - 1)d$, thus $\mathcal{C}$ is equidistant. ♯

Note that if $\frac{d}{n} \leq \frac{q-1}{q}$, then the above inequality is fulfilled for all $m \in \mathbb{N}$, hence giving no obstruction at all in this case.

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

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 $d-1 \leq n-k \leq n-2$, that is $d < n$. We show that the **residual code** $\mathcal{C}^*$ generated by the rows of G^* is an $[n-d, k-1, d^*]$-code such that $d^* \geq \lceil \frac{d}{q} \rceil$:

i) We first show that $\mathrm{rk}_{\mathbb{F}_q}(G^*) = k-1$, so that G^* is a generator matrix of $\mathcal{C}^*$:

Assume to the contrary that $\mathrm{rk}_{\mathbb{F}_q}(G^*) \leq k-2$, then there is $w \in \mathbb{F}_q^d$ such that up to row operations may assume that

$$[G^{**} \mid G^*] = \left[\begin{array}{c|c} w & 0_{n-d} \\ \hline * & * \end{array}\right] \in \mathbb{F}_q^{(k-1)\times n}.$$

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 $0 \neq [w - x \cdot 1_d \mid 0_{n-d}] \in \mathcal{C}$ such that $\mathrm{wt}(w - x \cdot 1_d) < d$, for some $x \in \mathbb{F}_q$, a contradiction as well. ♯

ii) We show that the minimum weight of $\mathcal{C}^*$ 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 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 we get

$$n \geq d + \sum_{i=0}^{k-2} \lceil \frac{\lceil \frac{d}{q} \rceil}{q^i} \rceil \geq \lceil \frac{d}{q^0} \rceil + \sum_{i=0}^{k-2} \lceil \frac{d}{q^{i+1}} \rceil = \sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil. \qquad ♯$$

The Griesmer bound for linear codes improves the (non-linear) Plotkin bound:

The former entails $n \geq d \cdot \sum_{i=0}^{k-1} \frac{1}{q^i} = \frac{d}{q^{k-1}} \cdot \frac{q^k-1}{q-1}$, or equivalently $d(q^k-1) \leq nq^{k-1}(q-1)$, that is $q^k(d - n \cdot \frac{q-1}{q}) \leq d$, which is the Plotkin bound. ♯

Proposition. If $\mathcal{C}$ is an MDS code such that $k \geq 2$, then we have $d \leq q$; saying that the minimum distance of MDS codes is severely restricted.

Proof. Assume to the contrary that $d > q$, then we have $n \geq \sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil \geq d + 2 + \sum_{i=2}^{k-1} 1 = d + k = (n - k + 1) + k = n + 1$, a contradiction. ♯

(7.3) Theorem: Gilbert bound [1952]. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, and let $n, m, d \in \mathbb{N}$. (Recall that we neither assume that q is a prime power, nor that m is a q-power, but we may assume that $0 \in \mathcal{X}$.)

If $m \cdot |\mathcal{B}_{d-1}(0_n)| \leq q^n$, then there is an (n, m, d')-code over $\mathcal{X}$ such that $d' \geq d$.

b) Let $n, k, d \in \mathbb{N}$. If $q^k \cdot |\mathcal{B}_{d-1}(0_n)| < q^{n+1}$, 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 in this case $\mathcal{C}$ has infinite minimum distance.

As long as $|\mathcal{C}| < m$ we have $|\mathcal{C}| \cdot |\mathcal{B}_{d-1}(0_n)| < q^n = |\mathcal{X}^n|$, hence $\bigcup_{v \in \mathcal{C}} \mathcal{B}_{d-1}(v) \subset \mathcal{X}^n$ is a proper subset. Thus there is $w \in \mathcal{X}^n$ having distance at least d from any element of $\mathcal{C}$. Hence, since $d(\mathcal{C}) \geq d$, this also holds for $\mathcal{C} \dot\cup \{w\}$.

b) Since $|\mathcal{B}_n(0_n)| = q^n$, the assumption implies that $d \leq n$. We proceed by induction on k: If $k = 1$ then the repetition $[n, 1, n]$-code is as desired. Hence we may assume that $k \geq 2$.

There is an $[n, k-1, d']$-code $\mathcal{C} \leq \mathbb{F}_q^n$ such that $d' \geq d$. Since $q^{k-1} \cdot |\mathcal{B}_{d-1}(0_n)| < q^n$ entails that $\bigcup_{v \in \mathcal{C}} \mathcal{B}_{d-1}(v) \subset \mathbb{F}_q^n$ is a proper subset, there is $w \in \mathbb{F}_q^n$ having distance at least d from any element of $\mathcal{C}$. In particular, we have $\mathrm{wt}(w) \geq d$, and $aw \in \mathbb{F}_q^n$ has the same distance properties from $\mathcal{C}$, for $a \in \mathbb{F}_q^*$. Thus we have $\mathcal{C} \cap \langle w \rangle_{\mathbb{F}_q} = \{0\}$, hence for $\mathcal{C}^+ := \mathcal{C} + \langle w \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^n$ we have $\dim_{\mathbb{F}_q}(\mathcal{C}^+) = k$.

We get $d(aw + v, bw + v') = d((a - b)w, v' - v) \geq d$, for $a, b \in \mathbb{F}_q$ and $v, v' \in \mathcal{C}$ such that $a \neq b$ or $v \neq v'$, showing that $d(\mathcal{C}^+) \geq d$. ♯

(7.4) Theorem: Gilbert-Varshamov bound [1952, 1957]. Let $n, k, d \in \mathbb{N}$ such that $n \geq 2$ and $d \geq 2$. If $q^k \cdot |\mathcal{B}_{d-2}(0_{n-1})| < q^n$, then there is an $[n, k, d']$-code over $\mathbb{F}_q$ such that $d' \geq d$.

Proof. The assumption implies that $k < n$, and moreover, since we have $|\mathcal{B}_{n-k}(0_{n-1})| \geq |\mathcal{B}_{n-k}(0_{n-k})| = q^{n-k}$, it implies that $d - 1 \leq n - k$. (Thus the Singleton bound is fulfilled.) We construct an $\mathbb{F}_q$-generating set $B_n \subseteq \mathbb{F}_q^{n-k}$, such that any $(d-1)$-subset thereof is $\mathbb{F}_q$-linearly independent; then $(B_n)^{\mathrm{tr}} \in \mathbb{F}_q^{(n-k)\times n}$ is the check matrix of a code as desired:

To do so, we proceed successively to find subsets $B_{n-k} \subseteq B_{n-k+1} \subseteq \cdots \subseteq B_j \subseteq \cdots \subseteq B_n$ of cardinality $|B_j| = j$, such that any $(d-1)$-subset of B_j is $\mathbb{F}_q$-linearly independent. To start with, let $B_{n-k} \subseteq \mathbb{F}_q^{n-k}$ be an $\mathbb{F}_q$-basis.

For $j \in \{n-k, \ldots, n-1\}$, the number of vectors in $\mathbb{F}_q^{n-k}$ being an $\mathbb{F}_q$-linear combination of at most $d-2$ elements of 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 $w \in \mathbb{F}_q^{n-k}$ such that any $(d-1)$-subset of $B_{j+1} := B_j \dot\cup \{w\}$ is $\mathbb{F}_q$-linearly independent. ♯

The (linear) Gilbert-Varshamov bound 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 $d'' \geq d+1$. This indeed is an improvement, since $d'' \geq d+1 \geq 2$ implies that the punctured code $\mathcal{C}^\bullet \leq \mathbb{F}_q^n$ has $\mathbb{F}_q$-dimension k and fulfills $d(\mathcal{C}^\bullet) \geq d''-1 \geq d$. ♯

(7.5) 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 the repetition $[n,1,n]$-code entails that $K_q(n,d) \in \{1,\ldots,n\}$ is well-defined. Moreover, we have $K_q(n,d+1) \leq K_q(n,d)$. For $d=1$ the $[n,n,1]$-code $\mathbb{F}_q^n$ shows that $K_q(n,1) = n$; and for $d=n$ the Singleton bound implies $k \leq 1$, showing that $K_q(n,n) = 1$.

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

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 a zero component we get the $[n,k,d]$-code $\{[x_1,\ldots,x_{n-1},0] \in \mathbb{F}_q^n; [x_1,\ldots,x_{n-1}] \in \mathcal{C}^\bullet\} \leq \mathbb{F}_q^n$, ♯

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 as follows (where (i)–(iii) hold for non-linear codes):

i) Singleton bound $k \leq n-d+1$,
ii) Hamming bound $q^k \cdot \sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} \cdot (q-1)^i \leq q^n$,
iii) Plotkin bound $q^k \cdot (d - n \cdot \frac{q-1}{q}) \leq d$,
iv) Griesmer bound $\sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil \leq n$.

Lower bounds on $k \geq 0$ for an $[n,k,d']$-code over $\mathbb{F}_q$ to exist, where $d' \geq d$, thus lower bounds on $K_q(n,d)$, are as follows (where (v) holds for non-linear codes):

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

A non-trivial code $\mathcal{C} \leq \mathbb{F}_q^n$ such that $\dim_{\mathbb{F}_q}(\mathcal{C}) = K_q(n, d(\mathcal{C}))$ 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 the minimum distance is severely restricted.

Actually, in general $K_q(n,d)$ is not at all known precisely. Improving the known lower and upper bounds, aiming at determining $K_q(n,d)$ affirmatively, and finding, and possibly classifying, optimal codes continue to be major problems of combinatorial coding theory.

(7.6) Example: Binary codes of length 13 and minimum distance 5. Let $q := 2$, let $n := 13$, and let $d := 5$. Then we get the following necessary conditions for the existence of an $[n,k,d]$-code:

i) The Singleton bound yields $k \leq 13 - 5 + 1 = 9$.

iv) The Griesmer bound yields $\sum_{i=0}^{k-1} \lceil \frac{5}{2^i} \rceil = 5+3+2+(k-3) \leq 13$, thus $k \leq 6$.

iii) Since $5 - 13 \cdot \frac{1}{2} < 0$ the Plotkin bound does not yield immediately.

But by the lower bounds given below we may assume that $k \geq 4$, so that we may first extend a $[13,k,5]$-code to a $[14,k,6]$-code, and then shorten it three times to obtain a $[11,k-3,6]$-code (possibly by additionally puncturing and adding zero components). Now the Plotkin bound yields $2^{k-3} \cdot (6 - 11 \cdot \frac{1}{2}) = 2^{k-4} \leq 6$, thus $k \leq 6$ again.

ii) The Hamming bound yields $2^k \cdot \sum_{i=0}^{2} \binom{13}{i} = 2^k \cdot 92 \leq 2^{13} = 8192$, thus $k \leq 6$.

But now assume that there is a $[13,6,5]$-code. Then the Helgert-Stinaff construction yields a residual $[8,5,d']$-code, where $d' \geq \lceil \frac{5}{2} \rceil = 3$. Hence there is a $[8,5,3]$-code (possibly by puncturing and adding zero components). Now the Hamming bound yields $2^5 \cdot \sum_{i=0}^{1} \binom{8}{i} = 2^5 \cdot 9 = 288 \leq 2^8 = 256$, a contradiction.

Thus we have $k \leq 5$, yielding the upper bound $K_2(13,5) \leq 5$. Conversely, we get the following sufficient conditions for the existence of an $[n,k,d]$-code:

v) The Gilbert bound reads $2^k \cdot \sum_{i=0}^{4} \binom{13}{i} = 2^k \cdot 1093 \leq 2^{13}$, hence $K_2(13,5) \geq 2$.

vi) The linear Gilbert bound reads $2^k \cdot 1093 < 2^{14} = 16384$, hence $K_2(13,5) \geq 3$.

vii) The Gilbert-Varshamov bound reads $2^k \cdot \sum_{i=0}^{3} \binom{12}{i} = 2^k \cdot 299 < 2^{13}$, hence we get the lower bound $K_2(13,5) \geq 4$.

In conclusion we have $4 \leq K_2(13,5) \leq 5$. We show that $K_2(13,5) = 5$ indeed:

In view of the Helgert-Stinaff construction, we consider the extended Hamming $[8,4,4]$-code $\widehat{\mathcal{H}}_3 \leq \mathbb{F}_2^8$, see (6.2), having generator and check matrices as follows:

$$\widehat{G}_3 := \left[\begin{array}{cccccccc} . & . & . & 1 & 1 & 1 & 1 & . \\ . & 1 & 1 & 1 & 1 & . & . & . \\ 1 & 1 & . & . & 1 & 1 & . & . \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right], \quad \widehat{H}_3 := \left[\begin{array}{cccccccc} . & . & . & 1 & 1 & 1 & 1 & . \\ . & 1 & 1 & . & . & 1 & 1 & . \\ 1 & . & 1 & . & 1 & . & 1 & . \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right].$$

Now let $\mathcal{C} \leq \mathbb{F}_2^{13}$ be the code generated by

$$G := \left[\begin{array}{ccccc|cccccccc} 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . & . & . \\ \hline . & . & 1 & . & . & . & . & . & 1 & 1 & 1 & 1 & . \\ . & . & . & 1 & . & . & 1 & 1 & 1 & 1 & . & . & . \\ . & . & . & . & 1 & 1 & 1 & . & . & 1 & 1 & . & . \\ . & . & . & . & . & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right] \in \mathbb{F}_2^{5\times 13},$$

the lower right hand corner consisting of the matrix $\widehat{G}_3$. 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 $\mathrm{wt}(v) = \mathrm{wt}(v') + \mathrm{wt}(v'') \geq 1 + 4 = 5$. ♯

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) **Nadler** $(12, 2^5, 5)$**-code**.

8 Asymptotic bounds

(8.1) 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 strictly increasing length whose relative minimum distance 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 $\kappa_q(1) = 0$. Actually, $\kappa_q(\delta)$ is continuous [Manin, 1981]. 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 5.

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

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

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

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 $\delta \notin \{0, \frac{q-1}{q}\}$; recall that $\kappa_q(\delta)$ is decreasing.

i) Let first $\frac{q-1}{q} < \delta \leq \delta(\mathcal{C})$. The Plotkin bound $q^k \cdot (d - n \cdot \frac{q-1}{q}) \leq d$ yields $q^k \cdot (\frac{d}{n} - \frac{q-1}{q}) \leq \frac{d}{n}$, thus $q^k \leq \frac{\delta(\mathcal{C})}{\delta(\mathcal{C}) - \frac{q-1}{q}} = 1 + \frac{q-1}{q} \cdot \frac{1}{\delta(\mathcal{C}) - \frac{q-1}{q}} \leq 1 + \frac{q-1}{q} \cdot \frac{1}{\delta - \frac{q-1}{q}}$ Hence k is bounded above, implying $\rho(\mathcal{C}) = \frac{k}{n} \to 0$, for $n \to \infty$, thus $\kappa_q(\delta) = 0$.

ii) 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 $1 \leq n' \leq \frac{q(d-1)}{q-1} < \frac{n(d-1)}{d} < n$. 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 (possibly by puncturing and adding zero components).

We have $d - n' \cdot \frac{q-1}{q} = d - \frac{q-1}{q} \cdot \lfloor \frac{q(d-1)}{q-1} \rfloor \geq d - \frac{q-1}{q} \cdot \frac{q(d-1)}{q-1} = 1$. Hence the Plotkin bound yields $q^{k'} \leq \frac{d}{d - n' \cdot \frac{q-1}{q}} \leq d$. Thus we have $q^k \leq q^{k'} q^{n-n'} \leq d q^{n-n'}$, hence $k \leq n - n' + \log_q(d)$, thus $\rho(\mathcal{C}) = \frac{k}{n} \leq 1 - \frac{n'}{n} + \frac{1}{n} \cdot \log_q(d)$. Now 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, for $n \to \infty$, we infer that $\kappa_q(\delta) \leq 1 - \frac{q}{q-1} \cdot \delta$. ♯

(8.3) Bounds based on sphere packing. In order to proceed, we need a few preparations: Let $q \in \mathbb{N}$ such that $q \geq 2$, and for $0 < \alpha \leq \frac{q-1}{q} < 1$ 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 by letting $H_q(0) := 0$.

Up to a constant, $H_q(p)$ is the average conditional information content of a symmetric channel over an alphabet of cardinality q with error probability $0 \leq p \leq \frac{q-1}{q}$; see (3.3). For this reason H_q is also called the q-ary **entropy function**.

Lemma. Let $0 < \delta < \frac{q-1}{q}$, and let $[d_1, d_2, \ldots] \subseteq \mathbb{N}_0$ be a sequence such that, writing $d = d_n$, we have $\frac{d}{n} \to \delta$, for $n \to \infty$. Then, we have

$$\frac{1}{n} \cdot \log_q(|\mathcal{B}_d(0_n)|) = \frac{1}{n} \cdot \log_q\big(\sum_{i=0}^{d} \binom{n}{i} \cdot (q-1)^i\big) \to H_q(\delta), \quad \text{for} \quad n \to \infty.$$

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 terms $\frac{s}{n+1-s}$ 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 $\frac{j}{n+1-j} \leq q-1$ indeed. ♯

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 $j := d$ for $n \gg 0$.

Since $\frac{1}{n} \cdot \log_q(d+1) \leq \frac{1}{n} \cdot \log_q(n \cdot \frac{q-1}{q} + 1) \leq \frac{1}{n} \cdot \log_q(n+1) \to 0$, for $n \gg 0$, it suffices to show that $L_n := \frac{1}{n} \cdot \log_q\big(\binom{n}{d} \cdot (q-1)^d\big) \to H_q(\delta)$, for $n \to \infty$:

Stirling's formula $\lim_{n\to\infty} \frac{n! \cdot e^n}{n^n \cdot \sqrt{2\pi n}} = 1$ implies that $n! = (\frac{n}{e})^n \cdot \sqrt{2\pi n} \cdot (1 + \epsilon_n)$, for some sequence $\epsilon_n \to 0$, for $n \to \infty$. Thus we get, again for some $\epsilon_n \to 0$,

$$\log_q(n!) = (n + \frac{1}{2}) \log_q(n) - n \log_q(e) + \log_q(\sqrt{2\pi}) + \epsilon_n.$$

Thus using $\binom{n}{j} = \frac{n!}{j! \cdot (n-j)!}$ we obtain $\log_q(\binom{n}{j}) = (n+\frac{1}{2}) \log_q(n) - (j+\frac{1}{2}) \log_q(j) - (n-j+\frac{1}{2}) \log_q(n-j) - \log_q(\sqrt{2\pi}) + \epsilon_n$. This entails

$$\begin{aligned} L_n &= (1 + \tfrac{1}{2n}) \cdot \log_q(n) - (\tfrac{d}{n} + \tfrac{1}{2n}) \cdot \log_q(d) - (\tfrac{n-d}{n} + \tfrac{1}{2n}) \cdot \log_q(n-d) \\ &\quad + \tfrac{d}{n} \cdot \log_q(q-1) - \tfrac{1}{n} \cdot \log_q(\sqrt{2\pi}) + \epsilon_n. \end{aligned}$$

Since $d \leq n \cdot \frac{q-1}{q} \leq n$, for $n \gg 0$, we get

$$\begin{aligned} L_n &= \log_q(n) - \tfrac{d}{n} \cdot \log_q(n \cdot \tfrac{d}{n}) - \tfrac{n-d}{n} \cdot \log_q(n \cdot \tfrac{n-d}{n}) \\ &+ \tfrac{d}{n} \cdot \log_q(q-1) + \epsilon_n \\ &= -\tfrac{d}{n} \cdot \log_q(\tfrac{d}{n}) - (1 - \tfrac{d}{n}) \cdot \log_q(1 - \tfrac{d}{n}) + \tfrac{d}{n} \cdot \log_q(q-1) + \epsilon_n. \end{aligned}$$

Since $\frac{d}{n} \to \delta$, for $n \to \infty$, this finally yields

$$L_n \to \delta \log_q(q-1) - \delta \log_q(\delta) - (1-\delta) \log_q(1-\delta) = H_q(\delta). \qquad \sharp$$

We are now prepared to proceed to further asymptotic bounds:

Theorem: Asymptotic Hamming bound. For $0 \leq \delta \leq 1$ we have the upper bound $\kappa_q(\delta) \leq 1 - H_q(\frac{\delta}{2})$.

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 assuming that $d \geq 2 \cdot \lceil \frac{\delta}{2} \cdot n \rceil - 1$.

The Hamming bound $q^k \cdot |\mathcal{B}_{\lfloor \frac{d-1}{2} \rfloor}(0_n)| \leq q^n$ yields $k \leq n - \log_q(|\mathcal{B}_{\lfloor \frac{d-1}{2} \rfloor}(0_n)|) \leq n - \log_q(|\mathcal{B}_{\lceil \frac{\delta}{2} \cdot n \rceil - 1}(0_n)|)$. Thus, given a non-trivial $[n,k,d]$-code $\mathcal{C} \leq \mathbb{F}_q^n$, we infer that $\rho(\mathcal{C}) = \frac{k}{n} \leq 1 - \frac{1}{n} \cdot \log_q(|\mathcal{B}_{\lceil \frac{\delta}{2} \cdot n \rceil - 1}(0_n)|)$. Hence, since $\frac{1}{n} \cdot (\lceil \frac{\delta}{2} \cdot n \rceil - 1) \to \frac{\delta}{2} < \frac{1}{2} \leq \frac{q-1}{q}$, 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.

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

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 $|\mathcal{B}_{d-2}(0_{n-1})| < q^{n-k}$, then by the Gilbert-Varshamov bound there exists a non-trivial $[n,k,d]$-code $\mathcal{C} \leq \mathbb{F}_q^n$. By maximality we have $q^{k+1} \geq \frac{q^n}{|\mathcal{B}_{d-2}(0_{n-1})|}$, thus $k + 1 \geq n - \log_q(|\mathcal{B}_{d-2}(0_{n-1})|)$, entailing

$$\rho(\mathcal{C}) = \frac{k+1}{n} - \frac{1}{n} \geq -\frac{1}{n} + 1 - \frac{1}{n} \cdot \log_q(|\mathcal{B}_{d-2}(0_{n-1})|) \geq \frac{n-1}{n} - \frac{1}{n} \cdot \log_q(|\mathcal{B}_d(0_n)|).$$

Since $\frac{d}{n} \to \delta$, for $n \to \infty$, we get $\kappa_q(\delta) \geq 1 - H_q(\delta)$. ♯

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

The **asymptotic Elias-Bassalygo bound** [1967], being based on an improvement of the strategy used to prove the Plotkin bound, says that we have

$$\kappa_q(\delta) \leq 1 - H_q(\frac{q-1}{q} - \sqrt{\frac{q-1}{q} \cdot (\frac{q-1}{q} - \delta)}), \quad \text{for} \quad 0 \leq \delta \leq \frac{q-1}{q}.$$

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 we have

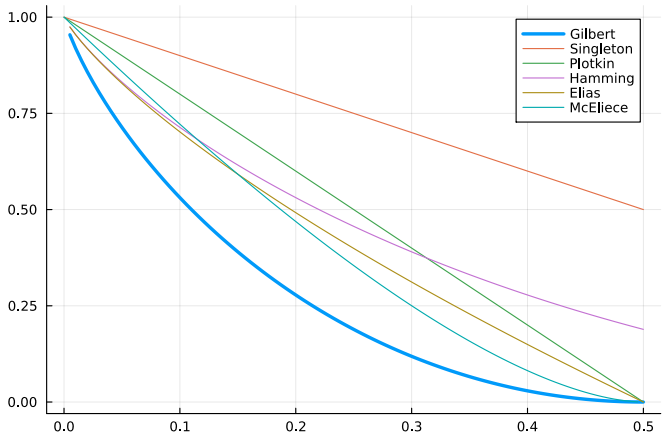
$$\kappa_q(\delta) \leq H_2(\frac{1}{2} - \sqrt{\delta(1-\delta)}), \quad \text{for} \quad 0 \leq \delta \leq \frac{1}{2},$$

and being the best asymptotic upper bound known.

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

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

Table 5: Asymptotic bounds for $q = 2$.



---

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.

(8.5) Asymptotically good codes. Surprisingly, it turns out to be extremely difficult to provide explicit sequences of codes reaching the asymptotic Gilbert-Varshamov bound, let alone improving it this way. Even the following much weaker notion is challenging:

A sequence $[\mathcal{C}_1, \mathcal{C}_2, \ldots]$ of codes over $\mathbb{F}_q$ of length n_i, such that $n_i \to \infty$ for $i \to \infty$, is called **(asymptotically) good**, if both $\limsup_{i\to\infty} \rho(\mathcal{C}_i) > 0$ and $\limsup_{i\to\infty} \delta(\mathcal{C}_i) > 0$; otherwise the sequence is called **(asymptotically) bad**.

In the latter case, the sequence does not provide any further insights at all into the values of κ_q. Actually, the optimal codes at hand are asymptotically bad:

Example. If $\mathcal{C}$ fulfills the Singleton bound, that is $\mathcal{C}$ is an MDS $[n, k, d]$-code over $\mathbb{F}_q$, then we have $d \leq q$, implying $\delta(\mathcal{C}) = \frac{d}{n} \to 0$ and $\rho(\mathcal{C}) = \frac{k}{n} = 1 + \frac{1}{n} - \frac{d}{n} \to 1$, for $n \to \infty$. Thus we recover $\kappa_q(0) = 1$.

Example. If $\mathcal{C}$ fulfills the Hamming bound, that is $\mathcal{C}$ is perfect, then we have (only) two infinite series of codes:

i) Firstly, $\mathcal{C} \leq \mathbb{F}_2^n$ is the binary repetition $[n, 1, n]$-code, where n is odd. Then we have $\delta(\mathcal{C}) = \frac{n}{n} = 1$, and $\rho(\mathcal{C}) = \frac{1}{n} \to 0$, for $n \to \infty$. Thus we recover $\kappa_q(1) \geq 0$.

ii) Secondly, we have the Hamming $[n, n-k, 3]$-code $\mathcal{H}_k$, for $k \geq 2$, where $n = \frac{q^k-1}{q-1}$. Then for $k \to \infty$ we have $\delta(\mathcal{H}_k) = \frac{3}{n} \to 0$ and $\rho(\mathcal{H}_k) = \frac{n-k}{n} = 1 - \frac{k(q-1)}{q^k-1} \to 1$. Thus we recover $\kappa_q(0) = 1$.

Example. The simplex $[n, k, q^{k-1}]$-code $\mathcal{S}_k$, for $k \geq 2$, where $n := \frac{q^k-1}{q-1}$, fulfills the Plotkin and Griesmer bounds. Then for $k \to \infty$ we have $\delta(\mathcal{S}_k) = \frac{q^{k-1}}{n} = \frac{q^{k-1}(q-1)}{q^k-1} \to \frac{q-1}{q}$ and $\rho(\mathcal{S}_k) = \frac{k}{n} = \frac{k(q-1)}{q^k-1} \to 0$. Thus we recover $\kappa_q(\frac{q-1}{q}) \geq 0$.

Remark. We discuss a few classes of codes with respect to their asymptotic behavior (which mostly we are not able to prove here unfortunately):

i) It is known that **BCH codes** are asymptotically bad. It still is an open problem whether or not **cyclic codes** are asymptotically good. But **self-dual codes** over any field, as well as **doubly-even self-dual binary codes**, are asymptotically good and reach the asymptotic Gilbert-Varshamov bound.

Actually, the proofs showing the above results on self-dual codes are non-constructive, as well as the ones for Goppa codes mentioned in the sequel. Although it was long considered doubtful to possibly construct a good infinite sequence of codes explicitly, such a construction was given by Justesen [1972]; but **Justesen codes** (which we are not able to treat here) do not reach the asymptotic Gilbert-Varshamov bound.

ii) More substantially, it is shown in (17.5) below that **Goppa codes** [1970] are asymptotically good and reach the asymptotic Gilbert-Varshamov bound.

The asymptotic Gilbert-Varshamov bound was long considered to be the best possible asymptotic lower bound. But using **geometric Goppa codes** [1981], and applying deep methods from the theory of **modular curves** (which we are not at all able to explain here), Tsfasman–Vladut–Zink [1982] provided a sequence of codes over $\mathbb{F}_{p^2}$, where $p \geq 7$, which exceeds the asymptotic Gilbert-Varshamov bound; these are still the asymptotically best codes known.

9 Weight enumerators

(9.1) Weight enumerators. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a code. For $i \in \mathbb{N}_0$ let $A_i = A_i(\mathcal{C}) := |\{v \in \mathcal{C}; \mathrm{wt}(v) = i\}| \in \mathbb{N}_0$; hence $A_0 \leq 1$, and $A_i = 0$ for $i \in \{1, \ldots, \mathrm{wt}(\mathcal{C}) - 1\}$, and $A_{\mathrm{wt}(\mathcal{C})} \geq 1$, and $A_i = 0$ for $i \geq n+1$, and $\sum_{i=0}^n A_i = |\mathcal{C}|$. The sequence $[A_0, \ldots, A_n]$ is called the **weight distribution** of $\mathcal{C}$.

Let $\{X, Y\}$ be indeterminates. The **(homogeneous) generating function**

$$A_{\mathcal{C}} := \sum_{i=0}^{n} A_i X^i Y^{n-i} = \sum_{v \in \mathcal{C}} X^{\mathrm{wt}(v)} Y^{n-\mathrm{wt}(v)} \in \mathbb{C}[X, Y]$$

is called the **(Hamming) weight enumerator** of $\mathcal{C}$. Hence $A_{\mathcal{C}}$ is homogeneous of total degree n and has non-negative integers as its coefficients. Note that

$$A_{\mathcal{C}}(-X, -Y) = \sum_{i=0}^{n} A_i (-X)^i (-Y)^{n-i} = (-1)^n \cdot \sum_{i=0}^{n} A_i X^i Y^{n-i} = (-1)^n \cdot A_{\mathcal{C}}.$$

By **dehomogenizing**, that is specializing $X \mapsto X$ and $Y \mapsto 1$, we obtain the **(ordinary) generating function**

$$A_{\mathcal{C}}(X, 1) = \sum_{i=0}^{n} A_i X^i = \sum_{v \in \mathcal{C}} X^{\mathrm{wt}(v)} \in \mathbb{C}[X].$$

Example. i) For the trivial code $\mathcal{C} := \{0_n\} \leq \mathbb{F}_q^n$ we get $A_{\mathcal{C}} = Y^n$.

For its dual $\mathcal{C}^\perp = \mathbb{F}_q^n$ we get $A_{\mathcal{C}^\perp} = \sum_{i=0}^{n} \binom{n}{i} (q-1)^i X^i Y^{n-i} = (Y + (q-1)X)^n$.

ii) 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

$$A_{\mathcal{C}^\perp} = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2i} X^{2i} Y^{n-2i} = \frac{1}{2} \cdot \sum_{i=0}^{n} \binom{n}{i} X^i Y^{n-i} + \frac{1}{2} \cdot \sum_{i=0}^{n} (-1)^i \binom{n}{i} X^i Y^{n-i},$$

thus $A_{\mathcal{C}^\perp} = \frac{1}{2} \cdot \big((Y-X)^n + (Y+X)^n \big) = \frac{1}{2} \cdot A_{\mathcal{C}}(Y - X, Y + X)$.

(9.2) 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 $\mathcal{C}$ is self-dual, then we have, recalling that $n = 2k$,

$$A_{\mathcal{C}}(X, Y) = A_{\mathcal{C}}\big(\frac{Y - X}{\sqrt{q}}, \frac{Y + (q-1)X}{\sqrt{q}} \big) \in \mathbb{C}[X, Y].$$

Proof. i) Let $\chi \colon \mathbb{F}_q \to \mathbb{C}^*$ be an **(additive) character** of $\mathbb{F}_q$, that is a group homomorphism from the additive group $\mathbb{F}_q^+$ to the multiplicative group $\mathbb{C}^*$.

There always is the **trivial character** $\mathbb{F}_q \to \mathbb{C}^* \colon a \mapsto 1$; but since $\mathbb{F}_q^+ \cong (\mathbb{Z}_p^+)^f$ as additive groups, where $p = \mathrm{char}(\mathbb{F}_q)$ and $q = p^f$, there always is a non-trivial character $\mathbb{F}_q \to \langle \zeta_p \rangle \leq \mathbb{C}^*$, where $\zeta_p \in \mathbb{C}^*$ is a primitive p-th root of unity.

Let V be a $\mathbb{C}$-vector space, and let $\omega\colon \mathbb{F}_q^n \to V$ be any map. Then the map

$$F_\chi(\omega)\colon \mathbb{F}_q^n \to V\colon v \mapsto \sum_{w\in\mathbb{F}_q^n} \chi(\langle v,w\rangle)\cdot\omega(w)$$

is called the **Hadamard transform** of ω. We show that for any non-trivial character χ of $\mathbb{F}_q$ we have

$$\sum_{v\in\mathcal{C}} F_\chi(\omega)(v) = q^k \cdot \sum_{w\in\mathcal{C}^\perp} \omega(w) \in V:$$

The left hand side equals

$$\begin{array}{rcl} \sum_{v\in\mathcal{C}} F_\chi(\omega)(v) & = & \sum_{v\in\mathcal{C}} \sum_{w\in\mathbb{F}_q^n} \chi(\langle v,w\rangle)\cdot\omega(w) \\ & = & \sum_{w\in\mathcal{C}^\perp} \big(\sum_{v\in\mathcal{C}} \chi(\langle v,w\rangle)\big)\cdot\omega(w) \\ & + & \sum_{w\in\mathbb{F}_q^n\setminus\mathcal{C}^\perp} \big(\sum_{v\in\mathcal{C}} \chi(\langle v,w\rangle)\big)\cdot\omega(w). \end{array}$$

Since $\chi(\langle v,w\rangle) = \chi(0) = 1 \in \mathbb{C}^*$, for $w \in \mathcal{C}^\perp$, the first summand becomes $q^k\cdot\sum_{w\in\mathcal{C}^\perp}\omega(w)$, which coincides with the right hand side of the above equation. Hence we have to show that the second summand vanishes:

We may assume that $\mathcal{C} \neq \{0\}$, or equivalently $\mathcal{C}^\perp \neq \mathbb{F}_q^n$. Then, 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 is surjective. Thus for $a \in \mathbb{F}_q$ we have $|\{v\in\mathcal{C}; \langle v,w\rangle = a\}| = \frac{q^k}{q} = q^{k-1}$. Hence the second summand above becomes

$$\sum_{w\in\mathbb{F}_q^n\setminus\mathcal{C}^\perp} \big(\sum_{v\in\mathcal{C}} \chi(\langle v,w\rangle)\big)\cdot\omega(w) = q^{k-1}\cdot\big(\sum_{a\in\mathbb{F}_q}\chi(a)\big)\cdot\big(\sum_{w\in\mathbb{F}_q^n\setminus\mathcal{C}^\perp}\omega(w)\big).$$

Hence it suffices to show that $\sum_{a\in\mathbb{F}_q}\chi(a) = 0 \in \mathbb{C}$:

Since χ is non-trivial, there is $b \in \mathbb{F}_q$ such that $\chi(b) \neq 1$. 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 $(\chi(b)-1))\cdot\sum_{a\in\mathbb{F}_q}\chi(a) = 0$, thus $\sum_{a\in\mathbb{F}_q}\chi(a) = 0$. ♯

ii) Let $\mathbb{C}[X,Y]_n$ be the $\mathbb{C}$-vector space of polynomials of total degree n, including the zero polynomial, let $\omega\colon \mathbb{F}_q^n \to \mathbb{C}[X,Y]_n\colon v \mapsto X^{\mathrm{wt}(v)}Y^{n-\mathrm{wt}(v)}$, and let $\delta\colon \mathbb{F}_q \to \{0,1\} \subseteq \mathbb{N}_0$ be defined by $\delta(0) = 0$, and $\delta(a) = 1$ for $a \neq 0$.

Thus, for any character χ, the Hadamard transform $F_\chi(\omega)\colon \mathbb{F}_q^n \to \mathbb{C}[X,Y]_n$ is given as follows, where $v = [x_1,\ldots,x_n] \in \mathbb{F}_q^n$:

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

If $x_i = 0$, or χ is trivial, then $\chi(ax_i) = \chi(0) = 1 \in \mathbb{C}^*$ shows that the associated factor equals

$$\sum_{a \in \mathbb{F}_q} X^{\delta(a)} Y^{1-\delta(a)} = Y + (q-1)X.$$

If $x_i \neq 0$ and χ is non-trivial, then using $\sum_{a \in \mathbb{F}_q} \chi(a) = 0 \in \mathbb{C}$ again, the associated factor becomes

$$\begin{aligned} \textstyle\sum_{a \in \mathbb{F}_q} \chi(ax_i) X^{\delta(a)} Y^{1-\delta(a)} &= Y + \big(\textstyle\sum_{a \in \mathbb{F}_q^*} \chi(ax_i)\big) \cdot X \\ &= Y + \big(\textstyle\sum_{a \in \mathbb{F}_q^*} \chi(a)\big) \cdot X \\ &= Y - \chi(0) \cdot X \\ &= Y - X. \end{aligned}$$

Thus $F_\chi(\omega)(v) = (Y-X)^{\mathrm{wt}(v)}(Y+(q-1)X)^{n-\mathrm{wt}(v)}$, whenever χ is non-trivial. (For the trivial character $\mathbf{1}$ we get $F_{\mathbf{1}}(\omega)(v) = (Y+(q-1)X)^n$, independently of $v \in \mathbb{F}_q^n$, which is not too interesting.) ♯

iii) In conclusion, by i) and ii), for any non-trivial character χ we get

$$\begin{aligned} q^k \cdot A_{\mathcal{C}^\perp}(X,Y) &= q^k \cdot \textstyle\sum_{w \in \mathcal{C}^\perp} X^{\mathrm{wt}(w)} Y^{n-\mathrm{wt}(w)} \\ &= q^k \cdot \textstyle\sum_{w \in \mathcal{C}^\perp} \omega(w) \\ &= \textstyle\sum_{v \in \mathcal{C}} F_\chi(\omega)(v) \\ &= \textstyle\sum_{v \in \mathcal{C}} (Y-X)^{\mathrm{wt}(v)} (Y+(q-1)X)^{n-\mathrm{wt}(v)} \\ &= A_{\mathcal{C}}(Y-X, Y+(q-1)X). \end{aligned}$$ ♯

Example: Simplex and Hamming codes. i) For $k \geq 2$, 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 we have

$$A_{\mathcal{S}_k} = Y^{\frac{q^k-1}{q-1}} + (q^k-1) X^{q^{k-1}} Y^{\frac{q^{k-1}-1}{q-1}} \in \mathbb{C}[X,Y].$$

Thus for the Hamming code $\mathcal{H}_k = \mathcal{S}_k^\perp \leq \mathbb{F}_q^n$ we get

$$A_{\mathcal{H}_k}(X,Y) = \frac{1}{q^k} \cdot A_{\mathcal{S}_k}(Y-X, Y+(q-1)X) \in \mathbb{C}[X,Y].$$

ii) We restrict ourselves to the binary case $q = 2$. Then we have $n = 2^k - 1$, hence $A_{\mathcal{S}_k} = Y^n + nX^{\frac{n+1}{2}} Y^{\frac{n-1}{2}}$, 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 $Y \mapsto 1$, yields

$$A_{\mathcal{H}_k}(X,1) = \frac{1}{n+1} \cdot ((1+X)^n + n(1-X)^{\frac{n+1}{2}} (1+X)^{\frac{n-1}{2}}) \in \mathbb{C}[X].$$

In particular, we get $A_{\mathcal{H}_2}(X,1) = \frac{1}{4}\cdot\big((1+X)^3+3(1-X)^2(1+X)\big) = 1+X^3$, showing again that $\mathcal{H}_2$ is the binary repetition code, and

$$A_{\mathcal{H}_3}(X,1) = \frac{1}{4}\cdot\big((1+X)^3+7(1-X)^4(1+X)^3\big) = 1+7X^3+7X^4+X^7.$$

iii) Thus for the self-dual extended Hamming $[8,4,4]$-code $\widehat{\mathcal{H}}_3 = (\widehat{\mathcal{H}}_3)^\perp$ by construction we get $A_{\widehat{\mathcal{H}}_3}(X,1) = 1+14X^4+X^8$, which by homogenizing yields $A_{\widehat{\mathcal{H}}_3} = Y^8+14X^4Y^4+X^8$. Now the MacWilliams identity indeed becomes

$$\begin{aligned}
A_{(\widehat{\mathcal{H}}_3)^\perp} &= A_{\widehat{\mathcal{H}}_3}(\tfrac{Y-X}{\sqrt{2}},\tfrac{Y+X}{\sqrt{2}})\\
&= \tfrac{1}{16}\cdot\big((Y+X)^8+14(Y-X)^4(Y+X)^4+(Y-X)^8\big)\\
&= \tfrac{1}{16}\cdot\big(\textstyle\sum_{i=0}^{8}\binom{8}{i}(1+(-1)^i)X^iY^{8-i}+14\cdot\sum_{i=0}^{4}\binom{4}{i}(-1)^iX^{2i}Y^{8-2i}\big)\\
&= \tfrac{1}{16}\cdot\big(\textstyle\sum_{i=0}^{4}(2\cdot\binom{8}{2i}+14\cdot(-1)^i\cdot\binom{4}{i})X^{2i}Y^{8-2i}\big)\\
&= Y^8+14Y^4X^4+X^8\\
&= A_{\widehat{\mathcal{H}}_3}.
\end{aligned}$$

10 Self-dual codes

(10.1) Polynomial algebras. To proceed we need help from the theory of polynomial invariants of finite groups. We collect the necessary facts, where we have to emphasize that we are working over a field of characteristic 0:

For $n\in\mathbb{N}$ let $R := \mathbb{C}[\mathcal{X}]$ be the polynomial $\mathbb{C}$-algebra in the indeterminates $\mathcal{X} := \{X_1,\dots,X_n\}$. Then $R = \bigoplus_{d\geq 0} R_d$ is an $\mathbb{N}_0$-graded algebra with respect to the total degree, where R_d is the $\mathbb{C}$-vector space of homogeneous polynomials of degree d, including the zero polynomial.

We shall need a criterion to decide whether a set $\{f_1,\dots,f_n\}\subseteq R$ is algebraically independent; note that R has Krull dimension $\dim(R)=n$, so that any algebraically independent subset of R has cardinality at most n:

The associated **Jacobian matrix** is defined as

$$J(f_1,\dots,f_n) = J_{\mathcal{X}}(f_1,\dots,f_n) := [\partial_{X_j}(f_i)]_{ij}\in R^{n\times n},$$

and $\det(J(f_1,\dots,f_n))\in R$ is called the associated **Jacobian determinant**.

Proposition: Jacobian criterion. The set $\{f_1,\dots,f_n\}\subseteq R$ is algebraically independent if and only if we have $\det(J(f_1,\dots,f_n))\neq 0$. ♯

(10.2) Invariant algebras. For $n\in\mathbb{N}$ let $R := \mathbb{C}[\mathcal{X}]$, where still $\mathcal{X} := \{X_1,\dots,X_n\}$, and let $\mathbf{G} := \mathrm{GL}_n(\mathbb{C})$. Then, letting $\mathbf{G}$ act trivially on $R_0\cong\mathbb{C}$, by the universal property of polynomial rings, the natural $\mathbb{C}$-linear action of $\mathbf{G}$ (from the right) on $R_1 = \bigoplus_{i=1}^n\langle X_i\rangle_{\mathbb{C}}$ extends uniquely to an action of $\mathbf{G}$ on R by graded $\mathbb{C}$-algebra automorphisms.

Given a finite subgroup $G \leq \mathbf{G}$, let

$$R^G := \mathrm{Fix}_R(G) = \{f \in R; f \cdot g = f \text{ for } g \in G\} \subseteq R.$$

Then, since G acts by graded algebra automorphisms, we conclude that $R^G = \bigoplus_{d \geq 0} R^G_d \subseteq R$ is a graded $\mathbb{C}$-subalgebra, being called the **invariant algebra** with respect to G. Moreover, R^G is a finitely generated $\mathbb{C}$-algebra again, and $R^G \subseteq R$ is a finite ring extension, so that $\dim(R^G) = \dim(R) = n$. Hence the question arises whether R^G might be a polynomial $\mathbb{C}$-algebra again, in this case necessarily in n indeterminates. This is answered as follows:

(10.3) Pseudo-reflection groups. For $n \in \mathbb{N}$ let $R := \mathbb{C}[\mathcal{X}]$, where still $\mathcal{X} := \{X_1, \ldots, X_n\}$, and let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite subgroup.

An element $s \in G$, hence having finite order, is called a **pseudo-reflection**, if its fixed point space fulfills $\dim_{\mathbb{C}}(\mathrm{Fix}(s)) = n-1$. A pseudo-reflection such that $s^2 = 1$ is called a **reflection**. Let $\mathcal{S}(G) \subseteq G$ be the set of its pseudo-reflections.

Theorem: Shephard–Todd [1954], Chevalley [1955], Serre [1967]. Let $G \leq \mathbf{G}$ be finite. Then the invariant algebra R^G is a polynomial algebra if and only if G is a **pseudo-reflection group**, that is $G = \langle \mathcal{S}(G) \rangle$. ♯

If G is a pseudo-reflection group, let $R^G = \mathbb{C}[f_1, \ldots, f_n]$, where $\{f_1, \ldots, f_n\}$ are (algebraically independent) homogeneous **basic invariants**, and let $d_i := \deg(f_i) \in \mathbb{N}$ be the associated **(fundamental) degrees**, for $i \in \{1, \ldots, n\}$; note that we may indeed assume to have homogeneous generators.

Proposition. a) The degrees $[d_1, \ldots, d_n]$ are uniquely defined up to order, independently of the particular choice of basic invariants. Moreover, we have

$$\prod_{i=1}^{n} d_i = |G| \quad \text{and} \quad \sum_{i=1}^{n} (d_i - 1) = |\mathcal{S}(G)|.$$

b) If $\{g_1, \ldots, g_n\} \subseteq R^G$ are algebraically independent and homogeneous such that $\prod_{i=1}^{n} \deg(g_i) = |G|$, then $\{g_1, \ldots, g_n\}$ are a set of basic invariants. ♯

(10.4) Weakly self-dual codes. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a linear code. We collect a few immediate properties related to weak self-duality over $\mathbb{F}_2$ and $\mathbb{F}_3$. To this end, $\mathcal{C}$ is called **r-divisible**, for some $r \geq 2$, if $r \mid \mathrm{wt}(v)$, for all $v \in \mathcal{C}$.

Proposition. a) Let $q = 2$ and $\mathcal{C} \leq \mathcal{C}^\perp$. Then $\mathcal{C}$ is 2-divisible; that is $\mathcal{C} = \mathcal{C}'$ is an even-weight code. Moreover, we have $1_n \in \mathcal{C}^\perp$; that is $\mathcal{C}^\perp = \widetilde{(\mathcal{C}^\perp)}$.

b) Let still $q = 2$. Then $\mathcal{C}$ is 4-divisible if and only if we have $\mathcal{C} \leq \mathcal{C}^\perp$ and $\mathcal{C}$ has a 4-divisible $\mathbb{F}_2$-basis; in this case $\mathcal{C}$ is called **doubly-even**.

c) Let $q = 3$ and $\mathcal{C} \leq \mathcal{C}^\perp$. Then $\mathcal{C}$ is 3-divisible.

Proof. For $v = [x_1, \ldots, x_n] \in \mathcal{C} \leq \mathcal{C}^\perp$ we have $0 = \langle v, v\rangle = \sum_{i=1}^n x_i^2 = \mathrm{wt}(v) \in \mathbb{F}_q$, both for $q = 2$ and $q = 3$. This proves c) and the first assertion in a).

Let now $q = 2$. Since any $v \in \mathcal{C} \leq \mathcal{C}^\perp$ has even weight, we have $\langle v, 1_n\rangle = \mathrm{wt}(v) = 0 \in \mathbb{F}_2$, so that $1_n \in \mathcal{C}^\perp$. This proves the second assertion in a). We now proceed to prove b):

Assume first that $\mathcal{C} \leq \mathcal{C}^\perp$ has a 4-divisible $\mathbb{F}_2$-basis. Then for $v, w \in \mathcal{C}$, letting $S := \mathrm{supp}(v)$ and $T := \mathrm{supp}(w)$, we get $0 = \langle v, w\rangle = |S \cap T| \in \mathbb{F}_2$, saying that $|S \cap T|$ is even. Moreover, we have $|\mathrm{supp}(v + w)| = |(S \setminus T) \dot\cup (T \setminus S)| = |S| + |T| - 2 \cdot |S \cap T|$, thus $v + w$ is 4-divisible if both v and w are so.

Assume finally that $\mathcal{C}$ is 4-divisible. Then for $v, w \in \mathcal{C}$, letting again $S := \mathrm{supp}(v)$ and $T := \mathrm{supp}(w)$, we have $2 \cdot |S \cap T| = |S| + |T| - |\mathrm{supp}(v + w)|$, so that $|S \cap T|$ is even. This implies $\langle v, w\rangle = |S \cap T| = 0 \in \mathbb{F}_2$. ♯

(10.5) The Gleason-Pierce group. Let $\mathcal{C} = \mathcal{C}^\perp \leq \mathbb{F}_q^n$ be a self-dual code. Hence we have $k := \dim_{\mathbb{F}_q}(\mathcal{C}) = \frac{n}{2}$, so that $n = 2k$ is even. Moreover, let $A_\mathcal{C} \in \mathbb{C}[X, Y]_n$ be its weight enumerator.

Then by self-duality the MacWilliams identity $A_\mathcal{C} = A_\mathcal{C}(\frac{Y-X}{\sqrt{q}}, \frac{Y+(q-1)X}{\sqrt{q}})$ holds; since n is even, we have $A_\mathcal{C} = A_\mathcal{C}(-X, -Y)$; and if $\mathcal{C}$ is r-divisible then we have $A_\mathcal{C} = A_\mathcal{C}(\zeta_r X, Y)$, where $\zeta_r \in \mathbb{C}$ is a primitive r-th root of unity.

In other words, letting $\Gamma_{q,r} := \langle -E_2, R, S\rangle = \langle R, \pm S\rangle \leq \mathbf{G} := \mathrm{GL}_2(\mathbb{C})$, where $R := \mathrm{diag}[\zeta_r, 1] \in \mathbf{G}$ and $S := \frac{1}{\sqrt{q}} \cdot \begin{bmatrix} -1 & 1 \\ q-1 & 1 \end{bmatrix} \in \mathbf{G}$, then $A_\mathcal{C} \in \mathbb{C}[X, Y]_n^{\Gamma_{q,r}}$.

Theorem: Gleason–Pierce [1967]. If $\mathcal{C} = \mathcal{C}^\perp$ is r-divisible for some $r \geq 2$, then $\Gamma_{q,r}$ is finite, or we have $A_\mathcal{C} = (Y^2 + (q-1)X^2)^{\frac{n}{2}}$.

Proof. We consider the projective line $\mathbf{P}^1(\mathbb{C}) \cong \widehat{\mathbb{C}} := \mathbb{C} \dot\cup \{\infty\}$, that is the Riemann sphere; writing $[x\colon y] \in \mathbf{P}^1(\mathbb{C})$ in homogeneous coordinates, for $x, y \in \mathbb{C}$ such that $[x, y] \neq [0, 0]$, the identification with $\widehat{\mathbb{C}}$ is given by $[x\colon y] \mapsto z := \frac{x}{y}$.

Then $\mathbf{G}$ acts on $\mathbf{P}^1(\mathbb{C})$ by $A\colon [x\colon y] \mapsto [(ax + cy)\colon (bx + dy)]$, for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{G}$; this action factors through $\overline{\mathbf{G}} := \mathrm{PGL}_2(\mathbb{C}) = \mathbf{G}/(\mathbb{C}^* \cdot E_2)$, and identifying with $\widehat{\mathbb{C}}$ yields the **(fractional) linear transformation** $A\colon z \mapsto \frac{az+c}{bz+d}$. Moreover, $\overline{\mathbf{G}}$ acts sharply 3-transitively on $\mathbf{P}^1(\mathbb{C})$; in particular, only the identity element of $\overline{\mathbf{G}}$ fixes any triple of pairwise distinct points.

We consider the finite set $\emptyset \neq \mathcal{Z} \subseteq \mathbf{P}^1(\mathbb{C})$ of zeroes of $A_\mathcal{C} = \sum_{i=0}^n A_i X^i Y^{n-i} \in \mathbb{C}[X, Y]_n^{\Gamma_{q,r}}$, which by $\Gamma_{q,r}$-invariance is $\Gamma_{q,r}$-stable. If $|\mathcal{Z}| \geq 3$, the finite non-empty set of 3-subsets of $\mathcal{Z}$ is $\Gamma_{q,r}$-stable as well, by sharp 3-transitivity implying that $\Gamma_{q,r}$ is finite. Hence we may assume that $|\mathcal{Z}| = 2$.

From $A_{\mathcal{C}}(0,1) = A_0 = 1$ we infer that $[0\colon 1] \notin \mathcal{Z}$. Letting $l \in \{1,\dots,n\}$ be maximal such that $A_l \neq 0$, we have $A_{\mathcal{C}} = Y^{n-l} \cdot (\sum_{i=0}^{l} A_i X^i Y^{l-i})$, entailing that there is $[1\colon 0] \neq z \in \mathcal{Z}$. Then we get $|z \cdot \langle R\rangle| = |\{\zeta_r^i z \in \mathbb{C}; i \in \mathbb{Z}_r\}| = r$, so that from $|\mathcal{Z}| = 2$ we infer that $r = 2$ and $[1\colon 0] \notin \mathcal{Z}$.

Thus we have $l = n$, and for $[1\colon 0] \neq [x\colon 1] \in \mathcal{Z}$ we have $[x\colon 1]\cdot\langle R\rangle = \{[\pm x\colon 1]\} = \mathcal{Z}$. Hence, since $A_0 = 1$, we get $A_{\mathcal{C}} = (Y^2 - \frac{1}{x^2}X^2)^{\frac{n}{2}}$. From $q^{\frac{n}{2}} = \sum_{i=0}^{n} A_i = A_{\mathcal{C}}(1,1) = (1 - \frac{1}{x^2})^{\frac{n}{2}}$ we get $1 - \frac{1}{x^2} = q\zeta_{\frac{n}{2}}^i$, for some $i \in \mathbb{Z}_{\frac{n}{2}}$. Since A_2 is a non-negative integer, this entails $i = 0$ and $A_{\mathcal{C}} = (Y^2 + (q-1)X^2)^{\frac{n}{2}}$. ♯

We first consider the exceptional case $A_{\mathcal{C}} = (Y^2 + (q-1)X^2)^{\frac{n}{2}}$: Then $A_{\mathcal{C}}$ is invariant even under $\Gamma_{q,2}$, not just $\Gamma_{q,1}$. Moreover, since $A_1(\mathcal{C}) = 0$ and $A_2(\mathcal{C}) = \frac{n}{2}\cdot(q-1)$ we have $\mathcal{C} = (\mathcal{C}_0)^{\oplus\frac{n}{2}}$, where $\mathcal{C}_0 := \langle [1,x]\rangle_{\mathbb{F}_q}$, for some $x \in \mathbb{F}_q^*$ such that $x^2 = -1$; the latter exists if and only if $4 \nmid q+1$.

(10.6) The Gleason-Pierce group, continued. Keeping the notation of (10.5), we are left with the case of $\Gamma_{q,r}$ being finite. We proceed to show when this happens; this essentially is a number theoretic property:

Recall that $\mathbb{Z}[\zeta_r]$ is the ring of integers in the algebraic number field $\mathbb{Q}(\zeta_r)$, which has degree $[\mathbb{Q}(\zeta_r)\colon \mathbb{Q}] = \varphi(r)$, where in turn φ denotes Euler's totient function, and let $N = N_{\mathbb{Q}(\zeta_r)/\mathbb{Q}}\colon \mathbb{Q}(\zeta_r) \to \mathbb{Q}$ be the associated norm map.

Lemma. i) If $r \geq 2$ is not a prime power, then $1 - \zeta_r \in \mathbb{Z}[\zeta_r]$ is a unit.
ii) If $r = p^f$, for some prime p and $f \geq 1$, then $N(1-\zeta_r) = p$.

Proof. i) Let $\Phi_r \in \mathbb{Z}[T]$ be the r-th cyclotomic polynomial. Then we obtain $N(1-\zeta_r) = \Phi_r(1) \in \mathbb{Z}$. Now we have $\sum_{i=0}^{r-1} T^i = \frac{T^r-1}{T-1} = \prod_{1\neq s\,|\,r} \Phi_s$, so that $\prod_{1\neq s\,|\,r} \Phi_s(1) = r$. In particular, we have $\prod_{i=1}^{f} \Phi_{p^i}(1) = p^f$. Taking the prime power parts of r into account, this yields $|\Phi_r(1)| = 1$ if r is not a prime power.

ii) If $r = p^f$ then $\Phi_r(T) = \Phi_p(T^{\frac{r}{p}})$, where $\Phi_p(T) = \sum_{i=0}^{p-1} T^i$, thus $\Phi_r(1) = p$. ♯

Theorem: Gleason–Pierce [1967]. The group $\Gamma_{q,r}$ is finite if and only if $r = 1$; or **i)** $q = 2$, $r = 2$; **ii)** $q = 2$, $r = 4$; **iii)** $q = 3$, $r = 3$; **iv)** $q = 4$, $r = 2$.

Proof. For the cases mentioned, $\Gamma_{q,1} = \langle \pm S\rangle$ is finitely generated and abelian, hence is finite; for the cases i)–iv) finiteness is checked explicitly using GAP. Hence we may assume that $r \geq 2$, and that $\Gamma_{q,r}$ is finite.

We have $RS = \frac{1}{\sqrt{q}} \cdot \begin{bmatrix} \zeta_r & -\zeta_r \\ q-1 & 1 \end{bmatrix} \in \Gamma_{q,r}$, which has characteristic polynomial $\chi_{RS} = T^2 + \frac{1-\zeta_r}{\sqrt{q}} \cdot T - \zeta_r \in \mathbb{C}[T]$. Since RS has finite order, its eigenvalues are roots of unity. Thus $\frac{1-\zeta_r}{\sqrt{q}}$ is an algebraic integer, hence $q \mid (1-\zeta_r)^2 \in \mathbb{Z}[\zeta_r]$.

Thus we conclude that $1-\zeta_r \in \mathbb{Z}[\zeta_r]$ is not a unit, so that $r = p^f$ is a prime power, for some $f \geq 1$. Moreover, from $q^{\varphi(r)} = N(q) \mid N(1-\zeta_r)^2 = p^2$, where $\varphi(r) = p^{f-1}(p-1)$, we conclude that we have the following four cases:

If $\varphi(r) = 1$, then $p = 2$ and $f = 1$, thus $r = 2$; moreover, we have $q = p = 2$ or $q = p^2 = 4$. If $\varphi(r) = 2$, then $q = p$; hence $q = p = 2$ and $f = 2$, thus $r = 4$; or $q = p = 3$ and $f = 1$, thus $r = 3$. ♯

(10.7) Weight enumerators of self-dual codes. We are now prepared to present the main result on weight enumerators of self-dual codes. This is governed by the group $\Gamma_{q,r}$ associated with the various types of codes:

The first case below refers to codes without divisibility condition; this also covers the exceptional codes in (10.5), which are 2-divisible, but $\Gamma_{q,2}$ is infinite for $q \notin \{2, 4\}$. The other cases below, named 'Type I–IV', refer to codes with divisibility condition, which correspond to the groups i)–iv) above.

Anyway: Since $(\pm S)^2 = E_2$, we infer that $\pm S$ are reflections; and R is a pseudo-reflection, for $r \geq 2$. Thus $\Gamma_{q,r}$ is a pseudo-reflection group. Hence, if $\Gamma_{q,r}$ is finite, then $\mathbb{C}[X,Y]^{\Gamma_{q,r}}$ is a (bivariate) polynomial $\mathbb{C}$-algebra, and for the associated degrees we have $(T-d_1)(T-d_2) = T^2 - (|\mathcal{S}(\Gamma_{q,r})|+2)\cdot T + |\Gamma_{q,r}| \in \mathbb{C}[T]$, so that the degrees can be determined from $|\Gamma_{q,r}|$ and $|\mathcal{S}(\Gamma_{q,r})|$.

Theorem: Gleason [1970]. For the weight enumerator $A_{\mathcal{C}} \in \mathbb{C}[X,Y]$ of a self-dual code $\mathcal{C} = \mathcal{C}^\perp \leq \mathbb{F}_q^n$ we have:

a) The weight enumerator $A_{\mathcal{C}}$ is a polynomial in

$$f := Y^2 + (q-1)X^2 \quad \text{and} \quad g := X(Y-X).$$

Recall that $f = A_{\mathcal{C}_0}$, where $\mathcal{C}_0 := \langle [1,x] \rangle_{\mathbb{F}_q}$, for some $x \in \mathbb{F}_q^*$ such that $x^2 = -1$; the latter exists if and only if $4 \nmid q+1$. But $f + (q-1)g = Y^2 + (q-1)XY$ is not the weight enumerator of a self-dual code.

b) 'Type I'. If $q = 2$, then $A_{\mathcal{C}}$ is a polynomial in

$$f := Y^2 + X^2 \quad \text{and} \quad g := X^2Y^2(Y^2-X^2)^2.$$

Note that $f = A_{\mathcal{C}_0}$, where $\mathcal{C}_0 := \langle [1,1] \rangle_{\mathbb{F}_2}$ is the repetition code of length 2, and $f^4 - 4g = Y^8 + 14X^4Y^4 + X^8 = A_{\widehat{\mathcal{H}}_3}$, where $\widehat{\mathcal{H}}_3$ is the extended Hamming code.

c) 'Type II'. If $q = 2$ and $\mathcal{C}$ is doubly-even, then $A_{\mathcal{C}}$ is a polynomial in

$$f := Y^8 + 14X^4Y^4 + X^8 \quad \text{and} \quad g := X^4Y^4(Y^4-X^4)^4.$$

Note that $f = A_{\widehat{\mathcal{H}}_3}$, where $\widehat{\mathcal{H}}_3$ is the extended Hamming code, and $f^3 - 42g = A_{\mathcal{G}_{24}}$, where $\mathcal{G}_{24} = \widehat{\mathcal{G}}_{23}$ is the extended binary Golay code, see (16.1).

d) 'Type III'. If $q = 3$, then $A_{\mathcal{C}}$ is a polynomial in

$$f := Y(Y^3 + 8X^3) \quad \text{and} \quad g := X^3(Y^3 - X^3)^3.$$

Note that $f = A_{\mathcal{H}_2}$, where $\mathcal{H}_2 = (\mathcal{H}_2)^\perp = \mathcal{S}_2$ is the Hamming code, coinciding with the simplex code, and $f^3 - 24g = A_{\mathcal{G}_{12}}$, where $\mathcal{G}_{12} = \widehat{\mathcal{G}}_{11}$ is the extended ternary Golay code, see (16.2).

e) 'Type IV'. If $q = 4$ and $\mathcal{C}$ is even, then $A_\mathcal{C}$ is a polynomial in

$$f := Y^2 + 3X^2 \quad \text{and} \quad g := X^2(Y^2 - X^2)^2.$$

Note that $f = A_{\mathcal{C}_0}$, where $\mathcal{C}_0 := \langle [1,1] \rangle_{\mathbb{F}_4}$ is the repetition code of length 2. But $f^3 - 9g = Y^6 + 45X^4Y^2 + 18X^6$ is not the weight enumerator of a self-dual quaternary code.

Proof. We again use GAP to obtain the group structures used in the sequel. In all cases, to show that $\{f, g\}$ is algebraically independent we verify that $\det(J(f, g)) \neq 0$ using GAP.

Moreover, since $\Gamma_{q,r}$ has an invariant hermitian scalar product, for any pseudo-reflection $Q \in \Gamma_{q,r}$, its eigenspace with respect to the exceptional eigenvalue is perpendicular to its reflection hyperplane. Hence $\Gamma_{q,r}$ acts transitively on the set of exceptional eigenspaces of the $\Gamma_{q,r}$-conjugates of Q.

a) The group $\Gamma_{q,1} = \langle \pm S \rangle \cong V_4$ is a real reflection group. We have $|\Gamma_{q,1}| = 4$ and $|\mathcal{S}(\Gamma_{q,1})| = |\{\pm S\}| = 2$, implying that $d_1 = d_2 = 2$.

We have the orbits $X \cdot \langle S \rangle = \{X, \frac{1}{\sqrt{q}}(Y - X)\}$ and $Y \cdot \langle S \rangle = \{Y, \frac{1}{\sqrt{q}}(Y + (q-1)X)\}$. Hence the products $g := X(Y - X)$ and $h := Y(Y + (q-1)X)$ are $\Gamma_{q,1}$-invariant, and so is $f := h - (q-1)g = Y^2 + (q-1)X^2$.

b) The group $\Gamma_{2,2} = \langle S, R \rangle \cong D_{16}$ is a real reflection group. We have $|\Gamma_{2,2}| = 16$ and $|\mathcal{S}(\Gamma_{2,2})| = 8$, implying that $d_1 = 2$ and $d_2 = 8$.

Let $F := Y^2 + X^2$ and $G := X^2Y^2(Y^2 - X^2)^2$ be the chosen basic invariants for the subgroup $\Gamma_{2,2} \leq \Gamma_{2,4}$. Then we have $F^4 - 4G = Y^8 + 14X^4Y^4 + X^8$.

Letting $f := Y^2 + X^2$, then f is both $\Gamma_{2,1}$-invariant and R-invariant.

The conjugacy class of the reflection R has 4 elements. Their eigenvectors with respect to the eigenvalue -1 are given as $\{[1,0], [1,1], [-1,1], [0,1]\}$, up to scalar multiples. Since $C_{\Gamma_{2,2}}(R) = \langle \pm R \rangle \cong V_4$, we get

$$[1,0] \cdot \Gamma_{2,2} = \{\pm[1,0], \frac{1}{\sqrt{2}}[\pm 1, \pm 1], \pm[0,1]\}.$$

Identifying $\mathbb{C}^2$ with $\mathbb{C}[X,Y]_1$, we get the orbit product $g := X^2Y^2(Y^2 - X^2)^2$, up to a scalar multiple, which is $\Gamma_{2,2}$-invariant.

c) The group $\Gamma_{2,4} = \langle R, S \rangle \cong 2.(4 \times \mathcal{S}_4)$ is a genuinely complex reflection group. We have $|\Gamma_{2,4}| = 192$ and $|\mathcal{S}(\Gamma_{2,4})| = 30$, implying that $d_1 = 8$ and $d_2 = 24$.

Let $F := Y^2 + X^2$ and $G := X^2Y^2(Y^2 - X^2)^2$ be the chosen basic invariants for the subgroup $\Gamma_{2,2} \leq \Gamma_{2,4}$. Then we have $F^4 - 4G = Y^8 + 14X^4Y^4 + X^8$.

Letting $f := Y^8 + 14X^4Y^4 + X^8$, then f is both $\Gamma_{2,2}$-invariant and R-invariant.

The conjugacy class of the pseudo-reflection R has 6 elements. Their eigenvectors with respect to the eigenvalue ζ_4 are given as $\{[1,0],[\pm 1,1],[\pm\zeta_4,1],[0,1]\}$, up to scalar multiples. Identifying $\mathbb{C}^2$ with $\mathbb{C}[X,Y]_1$, we get the product

$$g' := XY \cdot \prod_{i\in\mathbb{Z}_4} (Y - \zeta_4^i X) = XY(Y^4 - X^4).$$

Then we have $(g')^2 = (Y^2+X^2)^2 \cdot X^2Y^2(Y^2-X^2)^2 = F^2G$, showing that $(g')^2$ is $\Gamma_{2,2}$-invariant. Then $g := (g')^4 = X^4Y^4(Y^4-X^4)^4$ is R-invariant as well, thus is $\Gamma_{2,4}$-invariant.

d) The group $\Gamma_{3,3} = \langle R, S\rangle \cong 2.(2\times\mathcal{A}_4)$ is a genuinely complex reflection group. We have $|\Gamma_{3,3}| = 48$ and $|\mathcal{S}(\Gamma_{3,3})| = 14$, implying that $d_1 = 4$ and $d_2 = 12$.

Let $F := Y^2 + 2X^2$ and $G := X(Y - X)$ be the chosen basic invariants for the subgroup $\Gamma_{3,1} \leq \Gamma_{3,3}$. Then we have $F + 2G = Y(Y + 2X)$ and $F - 2G = Y^2 - 2XY + 4X^2$, as well as $F + G = Y^2 + XY + X^2$.

To find $\Gamma_{3,3}$-invariants of degree 4, by R-invariance we let $f := Y(Y^3+(aX)^3) = Y(Y + aX)(Y^2 - aXY + (aX)^2)$, for some $a \in \mathbb{C}$. Thus for $a = 2$ we have $f = (F + 2G)(F - 2G) = F^2 - 4G^2 = Y(Y^3 + 8X^3)$, which hence is $\Gamma_{3,1}$-invariant as well, thus is $\Gamma_{3,3}$-invariant.

To find $\Gamma_{3,3}$-invariants of degree 12 we observe that the conjugacy class of the pseudo-reflection R has 4 elements. Their eigenvectors with respect to the eigenvalue ζ_3 are given as $\{[1,0],[-1,1],[-E(3),1],[-E(3)^2,1]\}$, up to scalar multiples. Identifying $\mathbb{C}^2$ with $\mathbb{C}[X,Y]_1$, we get the product

$$g' := X \cdot \prod_{i\in\mathbb{Z}_3} (Y - \zeta_3^i X) = X(Y^3 - X^3) = X(Y - X)(Y^2 + XY + X^2).$$

Hence we have $g' = G(F + G)$, showing that g' is $\Gamma_{3,1}$-invariant. Then $g := (g')^3 = X^3(Y^3 - X^3)^3$ is R-invariant as well, thus is $\Gamma_{3,3}$-invariant.

e) The group $\Gamma_{4,2} = \langle R, S\rangle \cong D_{12}$ is a real reflection group. We have $|\Gamma_{4,2}| = 12$ and $|\mathcal{S}(\Gamma_{4,2})| = 6$, implying that $d_1 = 2$ and $d_2 = 6$.

Letting $f := Y^2 + 3X^2$, then f is both $\Gamma_{4,1}$-invariant and R-invariant.

The conjugacy class of the reflection R has 3 elements. Their eigenvectors with respect to the eigenvalue -1 are given as $\{[1,0],[1,1],[-1,1]\}$, up to scalar multiples. Since $C_{\Gamma_{4,2}}(R) = \langle \pm R\rangle \cong V_4$, we get $[1,0]\cdot\Gamma_{4,2} = \{\pm[1,0], \frac{1}{\sqrt{2}}[\pm 1,\pm 1]\}$. Identifying $\mathbb{C}^2$ with $\mathbb{C}[X,Y]_1$, we get the orbit product $g := X^2(Y^2 - X^2)^2$, up to a scalar multiple, which is $\Gamma_{4,2}$-invariant. ♯

(10.8) Extremal codes. The following applies to all the above types of self-dual codes, but here we restrict ourselves to Type I and Type II:

a) We consider self-dual binary codes $\mathcal{C} = \mathcal{C}^\perp \leq \mathbb{F}_2^n$. We have $n = 2k$, for some $k \in \mathbb{N}$. Letting $Z := X^2$, and specializing $Y \mapsto 1$, we let $f := 1 + Z \in \mathbb{C}[Z]$ and

$g := Z(1-Z)^2 \in \mathbb{C}[Z]$; thus $A_{\mathcal{C}}(\sqrt{Z},1) \in \langle f^{k-4j}g^j \in \mathbb{C}[Z]; j \in \{0,\ldots,N\}\rangle_{\mathbb{C}}$, where $N := \lfloor \frac{n}{8} \rfloor = \lfloor \frac{k}{4} \rfloor \in \mathbb{N}_0$.

Let ν_Z denote the valuation of $\mathbb{C}[Z]$ at the place $z = 0$. Since $\nu_Z(f) = 0$ and $f(0) = 1$, and $\nu_Z(g) = 1$ and $(g/Z)(0) = 1$, there are unique $a_1,\ldots,a_N \in \mathbb{Z}$ such that $h := f^k + \sum_{j=1}^{N} a_j f^{k-4j} g^j \in \mathbb{C}[Z]$ fulfills $\nu_Z(h-1) \geq N+1$. Rewriting in terms of X, we get $d_n^I := \nu_X(h-1) \geq 2 \cdot (\lfloor \frac{n}{8} \rfloor + 1)$.

Thus d_n^I is the maximum minimum distance a self-dual binary code $\mathcal{C}$ of length n possibly has. If $\mathcal{C}$ achieves this bound then $\mathcal{C}$ is called **extremal**; note that an extremal self-dual binary code might actually be doubly-even. By the above, the weight distribution of an extremal code of length n is uniquely determined; thus, although two such codes are not necessarily linearly equivalent, they are **formally equivalent**, that is they have the same weight distribution.

Actually, Mallows–Sloane [1973] have shown that $d_n^I = 2 \cdot (\lfloor \frac{n}{8} \rfloor + 1)$ holds. Moreover, Mallows–Sloane, Pless–Sloane [1973, 1975] have shown that extremal self-dual binary codes exist if and only if $n \in \{2,4,6,8,12,14,22,24\}$.

b) We consider doubly-even self-dual binary codes $\mathcal{C} = \mathcal{C}^{\perp} \leq \mathbb{F}_2^n$. We have $n = 8l$, for some $l \in \mathbb{N}$. Letting $Z := X^4$, and specializing $Y \mapsto 1$, we let $f := 1 + 14Z + Z^2 \in \mathbb{C}[Z]$ and $g := Z(1-Z)^4 \in \mathbb{C}[Z]$; thus $A_{\mathcal{C}}(\sqrt[4]{Z},1) \in \langle f^{l-3j}g^j \in \mathbb{C}[Z]; j \in \{0,\ldots,N\}\rangle_{\mathbb{C}}$, where $N := \lfloor \frac{n}{24} \rfloor = \lfloor \frac{l}{3} \rfloor \in \mathbb{N}_0$.

Since $\nu_Z(f) = 0$ and $f(0) = 1$, and $\nu_Z(g) = 1$ and $(g/Z)(0) = 1$, there are unique $a_1,\ldots,a_N \in \mathbb{Z}$ such that $h := f^l + \sum_{j=1}^{N} a_j f^{l-3j} g^j \in \mathbb{C}[Z]$ fulfills $\nu_Z(h-1) \geq N+1$. Rewriting in terms of X, we get $d_n^{II} := \nu_X(h-1) \geq 4 \cdot (\lfloor \frac{n}{24} \rfloor + 1)$.

Thus d_n^{II} is the maximum minimum distance a doubly-even self-dual binary code $\mathcal{C}$ of length n possibly has. Again, if $\mathcal{C}$ achieves this bound then $\mathcal{C}$ is called **extremal**, and the weight distribution of such a code is uniquely determined. Again, Mallows–Sloane [1973] have shown that $d_n^{II} = 4 \cdot (\lfloor \frac{n}{24} \rfloor + 1)$ holds.

Extremal doubly-even self-dual binary codes exist for finitely many n only: Actually, Mallows–Odlyzko–Sloane [1975] have shown that the weight enumerator of a putative extremal code has coefficient $A_{d_n^{II}+4} < 0$ for $n \gg 0$. A direct computation shows that this first happens for $l = 462$, that is $n = 3696$, for which we get $d_n^{II} = 620$ and $A_{d_n^{II}} \sim 9.6 \cdot 10^{168}$, while $A_{d_n^{II}+4} \sim -1.2 \cdot 10^{169}$. By Zhang [1999] this happens for all $l \geq 462$ such that $3 \mid l$, that is $n \geq 3696$ such that $24 \mid n$; and it happens for all $l \geq 492$, that is $n \geq 3936$.

On the other hand, extremal doubly-even self-dual binary codes seem to be very rare and are only known for the cases $l \in \{1,\ldots,8\} \dot{\cup} \{10,11,13,17\}$; the first gaps being $n = 72$ and $n = 96$ (which both are divisible by 24). It still is an open question for which n extremal doubly-even self-dual binary codes exist.

IV BCH

11 Cyclic codes

(11.1) 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 the permutation matrix $P_{(1,\ldots,n)} \in \mathrm{Aut}_{\mathbb{F}_q}(\mathcal{C})$, then $\mathcal{C}$ is called **cyclic**.

Example: Repetition codes. 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: Hamming code. The binary Hamming code $\mathcal{H} = \mathcal{H}_3 \leq \mathbb{F}_2^7$, whose elements are explicitly given in (5.3), is not cyclic, but applying the permutation matrix $P_{(3,4)(5,7,6)} \in \mathrm{Isom}_n(\mathbb{F}_q)$ 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

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

(11.2) Univariate polynomial rings. Let $\mathbb{F}_q[X]$ be the polynomial ring over $\mathbb{F}_q$ in the indeterminate X. Recall that $\mathbb{F}_q[X]$ is an Euclidean ring with respect to polynomial division, hence in particular is a principal ideal domain.

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

Then polynomial division yields $\mathbb{F}_q[X] = \mathbb{F}_q[X]_{<n} \oplus \langle X^n - 1\rangle$ as $\mathbb{F}_q$-vector spaces, where $\mathbb{F}_q[X]_{<n} := \bigoplus_{i=0}^{n-1} \mathbb{F}_q[X]_i = \bigoplus_{i=0}^{n-1} \langle X^i\rangle_{\mathbb{F}_q}$. Hence $\mathbb{F}_q[X]_{<n}$ is a set of representatives of $\overline{\mathbb{F}_q[X]}$, and $\overline{\ }\colon \mathbb{F}_q[X]_{<n} \to \overline{\mathbb{F}_q[X]}$ is an $\mathbb{F}_q$-isomorphism.

Let $\psi\colon \mathbb{F}_q^n \to \mathbb{F}_q[X]_{<n}\colon [c_0, \ldots, c_{n-1}] \mapsto \sum_{i=0}^{n-1} c_i X^i$, that is we consider the words in $\mathbb{F}_q^n$ as coefficients of polynomials in $\mathbb{F}_q[X]_{<n}$, and let $\overline{\psi} := \overline{\ } \circ \psi\colon \mathbb{F}_q^n \to \overline{\mathbb{F}_q[X]}\colon [c_0, \ldots, c_{n-1}] \mapsto \sum_{i=0}^{n-1} c_i \overline{X}^i$; thus both ψ and $\overline{\psi}$ 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 $\overline{\psi}(v) \cdot \overline{X} = (\sum_{i=0}^{n-1} c_i \overline{X}^i) \cdot \overline{X} = \overline{\sum_{i=0}^{n-1} c_i X^{i+1}} = c_{n-1}\overline{X}^0 + \sum_{i=1}^{n-1} c_{i-1}\overline{X}^i = \overline{\psi}(w) \in \overline{\mathbb{F}_q[X]}$, where $w := [c_{n-1}, c_0, \ldots, c_{n-2}] \in \mathbb{F}_q^n$. Thus the action of $P_{(1,\ldots,n)} \in \mathrm{Isom}_n(\mathbb{F}_q^n)$ on $\mathbb{F}_q^n$ is transported to multiplication with $\overline{X}$ on $\overline{\mathbb{F}_q[X]}$.

(11.3) Cyclic codes as ideals. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a linear code. Then $\mathcal{C}$ can be identified via ψ with the $\mathbb{F}_q$-subspace $\psi(\mathcal{C}) \leq \mathbb{F}_q[X]_{<n}$, and via $\overline{\psi}$ with the $\mathbb{F}_q$-subspace $\overline{\psi}(\mathcal{C}) \leq \overline{\mathbb{F}_q[X]}$. Moreover, $\mathcal{C}$ is cyclic if and only if $\overline{\psi}(\mathcal{C}) \leq \overline{\mathbb{F}_q[X]}$ is invariant under multiplication with $\overline{X}$, or equivalently under multiplication with $\overline{\mathbb{F}_q[X]}$, that is $\overline{\psi}(\mathcal{C}) \trianglelefteq \overline{\mathbb{F}_q[X]}$ is an ideal.

In this case, the preimage $\psi(\mathcal{C}) + \langle X^n - 1\rangle \subseteq \mathbb{F}_q[X]$ of $\overline{\psi}(\mathcal{C}) \subseteq \overline{\mathbb{F}_q[X]}$ with respect to $\overline{\ }$ 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 $\langle g\rangle = \psi(\mathcal{C}) + \langle X^n - 1\rangle \trianglelefteq \mathbb{F}_q[X]$, in particular implying $\langle \overline{g}\rangle = \overline{\psi}(\mathcal{C})$. 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.

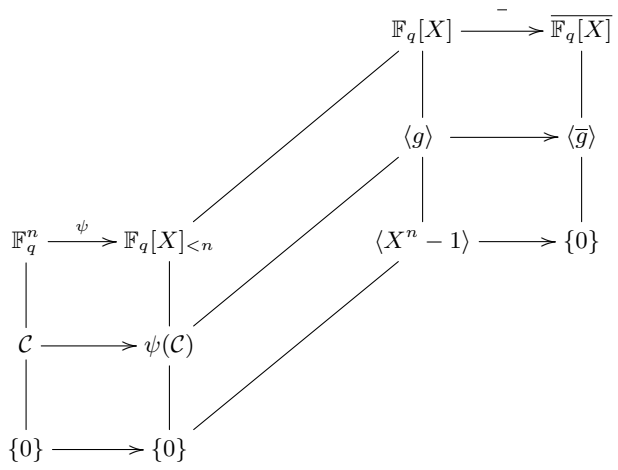
Conversely, any $g \in \mathbb{F}_q[X]$ such that $g \mid X^n - 1$ yields an ideal $\langle X^n - 1\rangle \subseteq \langle g\rangle \trianglelefteq \mathbb{F}_q[X]$, hence via $\overline{\psi}$ we get an ideal $\langle \overline{g}\rangle \trianglelefteq \overline{\mathbb{F}_q[X]}$, which in turn can 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 $X^n - 1 \in \mathbb{F}_q[X]$.

If $\mathcal{C} \leq \mathbb{F}_q^n$ is cyclic with generator polynomial $g \in \mathbb{F}_q[X]$, then for $v \in \mathbb{F}_q^n$ we have $v \in \mathcal{C}$ if and only if $g \mid \psi(v) \in \mathbb{F}_q[X]$. Moreover, if $\mathcal{C}$ is non-trivial then we have $g \in \psi(\mathcal{C}) \leq \mathbb{F}_q[X]_{<n}$, thus g is given as $\gcd(\psi(\mathcal{C}))$, or likewise as a non-zero polynomial of smallest degree in $\psi(\mathcal{C})$.

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 $\mathrm{lcm}(g, g') \in \mathbb{F}_q[X]$, respectively.

Example. i) For $g = 1 \in \mathbb{F}_q[X]$ we get $\langle \overline{1}\rangle = \overline{\mathbb{F}_q[X]}$, thus the associated code is $\mathcal{C} = \mathbb{F}_q^n$. For $h = X^n - 1 \in \mathbb{F}_q[X]$ we get $\langle \overline{X^n - 1}\rangle = \{0\} \trianglelefteq \overline{\mathbb{F}_q[X]}$, thus the associated code is the trivial code $\mathcal{C}^\perp = \{0\} \leq \mathbb{F}_q^n$.

Table 6: Cyclic codes.



---

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

Let $\mathcal{C}^\perp = \{[c_0,\dots,c_{n-1}] \in \mathbb{F}_q^n; \sum_{i=0}^{n-1} c_i = 0\}$ be the parity check code: 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 $f(1) = 0$, that is $X-1 \mid f$. Hence $\mathcal{C}^\perp$ corresponds to $\langle \overline{h}\rangle \trianglelefteq \overline{\mathbb{F}_q[X]}$, where $h := \psi([-1,1,0,\dots,0]) = X-1 \in \mathbb{F}_q[X]$ is the associated monic generator polynomial.

(11.4) 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 $k := \deg(g) \in \{0,\dots,n\}$. 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, generated by the **reversed polynomial**

$$h^* := X^{\deg(h)} \cdot h(X^{-1}) = \sum_{i=0}^{n-k} h_{n-k-i} X^i \in \mathbb{F}_q[X];$$

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

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

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 $\overline{\psi}(v) = \overline{gf} \in \overline{\mathbb{F}_q[X]}$ 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 $\psi(v) - gf = \psi(v) - g(qh+r) = \psi(v) - gr - (X^n-1)q \in \langle X^n - 1\rangle \trianglelefteq \mathbb{F}_q[X]$, which implies $\overline{\psi}(v) = \overline{gr} \in \overline{\mathbb{F}_q[X]}$.

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 $\{\overline{g}, \overline{g}\overline{X}, \dots, \overline{g}\overline{X}^{n-k-1}\} \subseteq \langle\overline{g}\rangle = \overline{\psi}(\mathcal{C}) \trianglelefteq \overline{\mathbb{F}_q[X]}$ is an $\mathbb{F}_q$-basis, which consists of the image under $\overline{\psi}$ of the rows of G.

b) Note that, since evaluation of polynomials into any commutative $\mathbb{F}_q$-algebra is an algebra homomorphism, and taking polynomial degrees is a homomorphism from the multiplicative monoid $\mathbb{F}_q[X]\setminus\{0\}$ to $\mathbb{N}_0^+$, we conclude that * is a monoid endomorphism of $\mathbb{F}_q[X]\setminus\{0\}$, that is we have $a^*b^* = (ab)^*$, for $0 \neq a, b \in \mathbb{F}_q[X]$.

Now, from $gh = X^n - 1$ we conclude that $h_0 \neq 0$, hence we have $\deg(h^*) = \deg(h) = n-k$, and from $g^*h^* = (gh)^* = (X^n-1)^* = -(X^n-1) \in \mathbb{F}_q[X]$ 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 $i \in \{1, \dots, n-k\}$ the i-th row of G is $v_i := [0, \dots, 0, g_0, \dots, g_k, 0, \dots, 0] \in \mathbb{F}_q^n$, where g_0 is the i-th entry, and for $j \in \{1, \dots, k\}$ the j-th row if H is $w_j := [0, \dots, 0, h_{n-k}, \dots, h_0, 0, \dots, 0] \in \mathbb{F}_q^n$, 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 $\langle v_i, w_j\rangle = 0$. ♯

Example: Hamming code. The non-zero elements of the Hamming code $\mathcal{H} \leq \mathbb{F}_2^7$, up to the linear equivalence applied in (11.1), consist of 1_7 and the cyclic shifts of $[1,1,0,1,0,0,0]$ and $[1,0,1,1,1,0,0]$.

Hence $\psi([1,1,0,1,0,0,0]) = X^3 + X + 1 \in \mathbb{F}_2[X]$ is the non-zero polynomial of smallest degree in $\psi(\mathcal{H})$. Thus $\mathcal{H}$ corresponds to $\overline{\psi}(\mathcal{H}) = \langle \overline{g} \rangle \trianglelefteq \mathbb{F}_2[X]$ with generator polynomial $g := X^3 + X + 1 \in \mathbb{F}_2[X]$; note that indeed we have $X^7 + 1 = (X+1)(X^3 + X + 1)(X^3 + X^2 + 1) \in \mathbb{F}_2[X]$.

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 $h^* := (X+1)(X^3+X+1) = X^4 + X^3 + X^2 + 1 \in \mathbb{F}_2[X]$. The generator and check matrices of $\mathcal{H}$ given in (11.1) reflect these facts; see also (12.5).

From $h^* = g \cdot (X+1) \in \mathbb{F}_2[X]$, and $X+1$ being a generator polynomial of the even-weight code $\mathcal{C}' \leq \mathbb{F}_2^n$, we conclude that $\mathcal{H}^\perp = \mathcal{H} \cap \mathcal{C}' = \{[c_0, \ldots, c_{n-1}] \in \mathcal{H}; \sum_{i=0}^{n-1} c_i = 0\}$; note that we recover the fact that $\mathcal{H}^\perp$ is weakly self-dual.

(11.5) Cyclic redundancy check (CRC) codes [Peterson, 1961]. a) 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 $\deg(g) = k$ 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 this generator matrix, a word $v = [a_0, \ldots, a_{n-k-1}] \in \mathbb{F}_q^{n-k}$ is encoded into $w = [b_0, \ldots, b_{k-1}; a_0, \ldots, a_{n-k-1}] \in \mathbb{F}_q^n$.

We have to find $[b_0, \ldots, b_{k-1}] \in \mathbb{F}_q^k$: We have $\psi(w) = \psi([b_0, \ldots, b_{k-1}]) + X^k \cdot \psi(v) \in \mathbb{F}_q[X]$. Polynomial division yields $X^k \cdot \psi(v) = qg + r$, for some $q, r \in \mathbb{F}_q[X]$ such that $r = 0$ or $\deg(r) < \deg(g) = k$. Since $w \in \mathcal{C}$ we have $g \mid \psi(w) = qg + r + \psi([b_0, \ldots, b_{k-1}])$, entailing $r + \psi([b_0, \ldots, b_{k-1}]) = 0$. This says that $\psi([b_0, \ldots, b_{k-1}])$ is the remainder of the shifted polynomial $-X^k \cdot \psi(v)$ upon polynomial division by g.

b) 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 $g \mid \psi(w) \in \mathbb{F}_q[X]$. Again polynomial division yields $\psi(w) = qg + r$, for some $q, r \in \mathbb{F}_q[X]$ 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 just 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 $l \leq k$; 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 is the j-th entry, for $j \in \{1, \ldots, n\}$, then $\psi(u) = X^{j-1} \cdot \psi([c_0, \ldots, c_{l-1}]) \in \mathbb{F}_q[X]$, hence from $\gcd(g, X) = 1$ and $\deg(g) = k$ we infer $g \nmid \psi(u)$, thus $u \notin \mathcal{C}$. ♯

ii) Since the parity check code is cyclic with generator polynomial $X - 1$, we conclude that $\mathcal{C} = \mathcal{C}' \leq \mathbb{F}_q^n$, the latter denoting the expurgated code, that is $\mathcal{C}$ is contained in the parity check code, if and only if $X - 1 \mid g \in \mathbb{F}_q[X]$.

For $q = 2$, if $\mathcal{C}$ is even-weight, then it detects all errors $u \in \mathbb{F}_2^n$ of odd weight: We have $\psi(u)(1) \equiv \mathrm{wt}(u) \not\equiv 0 \pmod 2$, hence $X + 1 \nmid \psi(u)$, thus $g \nmid \psi(u)$.

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 $\psi(u) = X^{i-1}(X^{j-i} + 1) \in \mathbb{F}_2[X]$. 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 $m \in \{1, \ldots, n - 1\}$. In other words, this holds if and only if g has an **n-primitive divisor**, that is an irreducible polynomial $f \in \mathbb{F}_q[X]$ such that $f \mid X^n - 1$, but $f \nmid X^m - 1$ for $m \in \{1, \ldots, n - 1\}$.

(11.6) Example: Universal Serial Bus (USB) [≥1996]. 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. A prominent example is the **Universal Serial Bus (USB)** data transmission standard:

i) The 'CRC-5-USB' polynomial $f := X^5 + X^2 + 1 \in \mathbb{F}_2[X]$ is used to add 5 check bits to 'token' packets consisting of 11 information bits, making up a code of length 16; thus storing a word needs 2 Bytes.

The polynomial $f \in \mathbb{F}_2[X]$ is irreducible, hence splits in $\mathbb{F}_{32}$, and thus divides $X^{31} + 1 \in \mathbb{F}_2[X]$. Since 31 is a prime, it follows that f is 31-primitive, entailing that we cannot do shorter than letting $n = 31$. Thus the code actually used is a 15-fold shortened cyclic code with generator polynomial f; note that the encoding and decoding algorithms are not affected by shortening.

ii) Similarly, for 'data' packets, having length 1023 Bytes, the 'CRC-16-USB' polynomial $g := X^{16} + X^{15} + X^2 + 1 = (X + 1)(X^{15} + X + 1) \in \mathbb{F}_2[X]$ is used to add 2 check Bytes to packets consisting of 1021 information Bytes.

Since $X + 1 \mid g$, the associated cyclic code is an even weight code.

The polynomial $g' := X^{15} + X + 1 \in \mathbb{F}_2[X]$ is the lexicographically smallest irreducible polynomial of degree 15, hence splits in $\mathbb{F}_{2^{15}}$, and thus divides $X^{32767} + 1 \in \mathbb{F}_2[X]$. Actually, g' is 32767-primitive, where $2^{15} - 1 = 32767 = 7 \cdot 31 \cdot 151$, entailing that we cannot do shorter than letting $n = 32767$. Thus the code of length $8184 = 8 \cdot 1023$ actually used is a 24583-fold shortened cyclic even-weight code with generator polynomial g.

(11.7) 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 $\psi(u) = X^j(X^5+1) \cdot \psi([c_0, \ldots, c_4]) \in \mathbb{F}_2[X]$, where the leftmost c_0 is the j-th entry, for $j \in \{0, \ldots, n-1\}$. Hence all adjacent transposition errors are detected if and only if $g \nmid \psi(u)$ for all error vectors u as above. Noting that we have the factorization $X^5 + 1 = (X+1)(X^4 + X^3 + X^2 + X + 1) \in \mathbb{F}_2[X]$, we conclude that the latter property holds if and only if $\gcd(g, X^5 + 1) = 1$.

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 $\mathbb{F}_{2^6} = \mathbb{F}_{64}$.

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

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

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.

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

For the RWTH-ID the 'CRC-5-USB' polynomial $g := X^5 + X^2 + 1 \in \mathbb{F}_2[X]$ has been chosen; 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 the word `L8BRX`, from Table 7 we get

$$\begin{aligned} h \;=\; & & & (1 + X^3 + X^4) \\ & + & X^5 \;\cdot\; & (X) \\ & + & X^{10} \;\cdot\; & (X + X^3 + X^4) \\ & + & X^{15} \;\cdot\; & (1 + X) \\ & + & X^{20} \;\cdot\; & (1 + X + X^2 + X^4). \end{aligned}$$

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.

12 BCH codes

(12.1) Roots of unity. a) Let $\mathbb{F}_q$ be the field with q elements, let $\mathbb{F}_q \subseteq \overline{\mathbb{F}}$ be a (fixed) algebraic closure, 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 $\gcd(\partial_X(X^n - 1), X^n - 1) = \gcd(nX^{n-1}, X^n - 1) = 1 \in \mathbb{F}_q[X]$, implying that $X^n - 1 \in \mathbb{F}_q[X]$ is square-free, that is a product of pairwise non-associate irreducible polynomials.

Thus $X^n - 1$ splits into pairwise distinct linear factors over $\overline{\mathbb{F}}$. Thus letting $\mathcal{V}_n := \mathcal{V}(X^n - 1) \subseteq \overline{\mathbb{F}}^*$ be the associated **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 $\mathcal{V}_n = \langle \zeta_n \rangle$.

Hence we have a group isomorphism $\mathbb{Z}_n \to \mathcal{V}_n \colon i \mapsto \zeta_n^i$, the left hand side written additively. 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 $i \in \mathbb{Z}_n^*$.

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

We have $\Gamma = \langle \varphi_q \rangle$, where $\varphi_q\colon \mathbb{F} \to \mathbb{F}\colon a \mapsto a^q$ is the associated Frobenius automorphism, having order $|\varphi_q| = \min\{i \in \mathbb{N}; \varphi_q^i = \mathrm{id}_{\mathbb{F}}\} = \min\{i \in \mathbb{N}; \zeta_n^{q^i} = \zeta_n\} = \min\{i \in \mathbb{N}; q^i = 1 \in \mathbb{Z}_n\} = |q|_{\mathbb{Z}_n^*}$, the order of $q \in \mathbb{Z}_n^*$. Identifying $\mathcal{V}_n$ with $\mathbb{Z}_n$, the group $\Gamma = \langle \varphi_q \rangle$ acts by $\varphi_q\colon \mathbb{Z}_n \to \mathbb{Z}_n\colon i \mapsto iq$.

The monic divisors $g \mid X^n - 1 \in \mathbb{F}[X]$ are described by their sets of zeroes $\mathcal{V}(g) \subseteq \mathcal{V}_n$, as $g = \prod_{\zeta \in \mathcal{V}(g)}(X - \zeta) \in \mathbb{F}[X]$. Since $\mathrm{Fix}_{\mathbb{F}}(\Gamma) = \mathbb{F}_q$, we have $g \in \mathbb{F}_q[X]$ if and only if $\mathcal{V}(g)$ is a union of Frobenius 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 as $\mu_i := \prod_{j \in i \cdot \Gamma}(X - \zeta_n^j) \in \mathbb{F}_q[X]$, for $i \in \mathbb{Z}_n$.

The polynomial $\mu_i \in \mathbb{F}_q[X]$ is the minimum polynomial of $\zeta_n^i \in \mathbb{F}$ over $\mathbb{F}_q$, hence we have $[\mathbb{F}_q(\zeta_n^i)\colon \mathbb{F}_q] = \deg(\mu_i) = |\zeta_n^i \cdot \Gamma| \mid |\Gamma| = [\mathbb{F}\colon \mathbb{F}_q]$. Thus we have equality $\deg(\mu_i) = |\Gamma|$ if and only if $\mathbb{F}_q(\zeta_n^i) = \mathbb{F}$. 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 does not necessarily act transitively on the set of primitive n-th roots of unity.

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 Frobenius orbits, next to $\{0\}$, are $\mathcal{O}' := \{1, 2, 4\}$ and $\mathcal{O}'' := \{3, 5, 6\}$.

This yields $X^7 + 1 = (X + 1) \cdot \prod_{i \in \mathcal{O}'}(X + \zeta_7^i) \cdot \prod_{i \in \mathcal{O}''}(X + \zeta_7^i) = \mu_0\mu_1\mu_3 = (X + 1) \cdot (X^3 + X + 1)(X^3 + X^2 + 1) \in \mathbb{F}_8[X]$, where $\mu_0, \mu_1, \mu_3 \in \mathbb{F}_2[X]$ are irreducible. Note that we do not specify which of the factors μ_1 and μ_3 has the chosen primitive 7-th root of unity ζ_7 as a zero.

(12.2) Zeroes of cyclic codes. a) Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q, n) = 1$, be a cyclic code with monic generator polynomial $g \in \mathbb{F}_q[X]$ of degree $k \in \{0, \ldots, n\}$. Let $\mathcal{V}(\mathcal{C}) := \mathcal{V}(g) \subseteq \mathcal{V}_n \subseteq \mathbb{F}_q(\zeta_n) =: \mathbb{F}$ be the **set of zeroes** of $\mathcal{C}$.

Hence $\mathcal{V}(\mathcal{C})$ is Frobenius stable. Moreover, any subset $\mathcal{V} \subseteq \mathcal{V}_n$ whose smallest Frobenius stable superset equals $\mathcal{V}(\mathcal{C})$, that is we have $\mathcal{V}(\mathcal{C}) = \mathcal{V} \cdot \Gamma$, is called a **defining set** of $\mathcal{C}$; 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 $g \mid \psi(v) \in \mathbb{F}_q[X]$. Since $g \in \mathbb{F}[X]$ is square-free, this is equivalent to $\mathcal{V}(\mathcal{C}) \subseteq \mathcal{V}(\psi(v)) \subseteq \mathcal{V}_n$, where $\mathcal{V}(\psi(v))$ is the set of zeroes of $\psi(v)$ being contained in $\mathcal{V}_n = \{\zeta_n^i; i \in \mathbb{Z}_n^*\}$; note that $\mathcal{V}(\psi(v))$ is Frobenius stable, and that $\psi(v)$ might have further zeroes or non-linear irreducible divisors. By taking Frobenius orbits, this in turn is equivalent to $\mathcal{V} \subseteq \mathcal{V}(\psi(v))$, for any defining set $\mathcal{V}$ of $\mathcal{C}$. Thus we have $\mathcal{C} = \{v \in \mathbb{F}_q^n; \psi(v)(\mathcal{V}) = \{0\}\}$, in other words

$$\mathcal{C} = \{[c_0, \ldots, c_{n-1}] \in \mathbb{F}_q^n; \sum_{i=0}^{n-1} c_i\zeta^i = 0 \in \mathbb{F} \text{ for } \zeta \in \mathcal{V}\}.$$

Moreover, we recover $\mathcal{V}(\mathcal{C}) = \bigcap_{v\in\mathcal{C}} \mathcal{V}(\psi(v))$.

b) We determine $\mathcal{V}(\mathcal{C}^\perp) \subseteq \mathcal{V}_n$: Letting $h \in \mathbb{F}_q[X]$ be the check polynomial associated with g, we have $h = \prod_{\zeta\in\mathcal{V}_n\setminus\mathcal{V}(\mathcal{C})}(X-\zeta) \in \mathbb{F}[X]$, thus we get $h^* = \prod_{\zeta\in\mathcal{V}_n\setminus\mathcal{V}(\mathcal{C})}(X-\zeta)^* = \prod_{\zeta\in\mathcal{V}_n\setminus\mathcal{V}(\mathcal{C})} -\zeta\cdot(X-\zeta^{-1}) \in \mathbb{F}[X]$, hence

$$\mathcal{V}(\mathcal{C}^\perp) = \mathcal{V}(h^*) = \mathcal{V}_n \setminus \mathcal{V}(\mathcal{C})^{-1}.$$

(12.3) Theorem. a) Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q,n) = 1$, be a cyclic code with set of zeroes $\mathcal{V}(\mathcal{C}) = \{\zeta_n^{a_1}, \ldots, \zeta_n^{a_k}\} \subseteq \mathcal{V}_n$, for $k \in \{0,\ldots,n\}$. Then a check matrix of $\mathcal{C}\otimes_{\mathbb{F}_q}\mathbb{F} \leq \mathbb{F}^n$, where $\mathbb{F} := \mathbb{F}_q(\zeta_n)$, is given as [Delsarte, 1975]

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

b) Let $\mathcal{V} \subseteq \mathcal{V}(\mathcal{C})$ be a defining set, and let $H(\mathcal{V})$ be the submatrix of $H(\mathcal{V}(\mathcal{C}))$ consisting of the rows corresponding to $\mathcal{V}$. Then we have $\mathcal{C} = \ker(H(\mathcal{V})^{\mathrm{tr}}) \cap \mathbb{F}_q^n$; in particular, $H(\mathcal{V}(\mathcal{C}))$ is a generalized check matrix of $\mathcal{C}$, in the sense of (5.4).

Proof. To abbreviate, we write $H := H(\mathcal{V}(\mathcal{C}))$. The submatrix of H consisting of its first k columns is a Vandermonde matrix associated with $\mathcal{V}(\mathcal{C})$, hence is invertible. Thus we have $\mathrm{rk}_{\mathbb{F}}(H) = k$, hence $\dim_{\mathbb{F}}(\ker(H^{\mathrm{tr}})) = n-k$.

Thus we infer $\dim_{\mathbb{F}_q}(\ker(H^{\mathrm{tr}}) \cap \mathbb{F}_q^n) \leq n-k$. Hence, since $\dim_{\mathbb{F}_q}(\mathcal{C}) = n-k$, to show a) and the statement in b) concerning H, it suffices to show $\mathcal{C} \subseteq \ker(H^{\mathrm{tr}})$:

Let $g = \sum_{i=0}^k g_i X^i := \prod_{\zeta\in\mathcal{V}(\mathcal{C})}(X-\zeta) \in \mathbb{F}[X]$ 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 associated generator matrix of $\mathcal{C}$, where $i \in \{1,\ldots,n-k\}$, and the j-th row $w_j \in \mathbb{F}^n$ of H, where $j \in \{1,\ldots,k\}$, we get

$$\langle v_i, w_j\rangle = \sum_{l=0}^k g_l \zeta_n^{(i+l-1)a_j} = \zeta_n^{(i-1)a_j}\cdot\sum_{l=0}^k g_l(\zeta_n^{a_j})^l = \zeta_n^{(i-1)a_j}\cdot g(\zeta_n^{a_j}) = 0 \in \mathbb{F}.$$

Finally, to conclude it remains to consider the submatrix $H(\mathcal{V})$: For any row $w \in \mathbb{F}^n$ of $H(\mathcal{V})$ there is a row $u \in \mathbb{F}^n$ of H such that $w = u^{q^i}$, for some $i \geq 0$, where the Frobenius automorphism is applied component-wise. For $v \in \mathbb{F}_q^n$ we have $v^q = v$, hence from $\langle v, w\rangle = \langle v, u^{q^i}\rangle = \langle v^{q^i}, u^{q^i}\rangle = \langle v, w\rangle^{q^i} \in \mathbb{F}$ we infer that $v \in \ker(H^{\mathrm{tr}})$ if and only if $v \in \ker(H(\mathcal{V})^{\mathrm{tr}})$. ♯

(12.4) Theorem. Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q,n) = 1$, be a cyclic code such that $\mathcal{V}(\mathcal{C})$ 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 $b \in \mathbb{Z}_n^*$, for some $\delta \in \{1,\ldots,n+1\}$. Then $\mathcal{C}$ has minimum distance $d(\mathcal{C}) \geq \delta$.

Proof. For $\delta = 1$ the consecutive set is empty, and $d(\mathcal{C}) \geq 1$ anyway; and for $\mathcal{C} = \{0\}$ we have $d(\mathcal{C}) = \infty$. Hence we may assume that $\delta \geq 2$ and $\mathcal{C} \neq \{0\}$.

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 submatrix $H(\zeta_n^a, \ldots, \zeta_n^{a+\delta-2}) \in \mathbb{F}^{(\delta-1)\times n}$ of H, where $\mathbb{F} := \mathbb{F}_q(\zeta_n)$, and show that any $(\delta-1)$-subset of its columns is $\mathbb{F}$-linearly independent:

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

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

Hence we have

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

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. ♯

(12.5) 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 the condition on k holds for $q = 2$.

Let $\mathcal{C} \leq \mathbb{F}_q^n$ be the cyclic code with defining set $\{\zeta_n\}$, that is $\mathcal{V}(\mathcal{C}) = \zeta_n \cdot \Gamma \subseteq \mathcal{V}_n$, thus 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 have $n - \dim_{\mathbb{F}_q}(\mathcal{C}) = \deg(g) = |\mathcal{V}(\mathcal{C})| = |\zeta_n \cdot \Gamma| = |\Gamma| = |q|_{\mathbb{Z}_n^*}$. Hence we have to show that $|q|_{\mathbb{Z}_n^*} = k$: We have $n \mid q^k - 1$, implying that $|q|_{\mathbb{Z}_n^*} \mid k$; moreover, we have $n = \sum_{i=0}^{k-1} q^i \geq q^{k-1}$, thus we have $n \nmid q^l - 1$, for $l \in \{1, \ldots, k-1\}$.

Finally, 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 $\delta - 1 = 2$ and step size $q-1$. Now $n = \sum_{i=0}^{k-1} q^i \equiv k \pmod{(q-1)}$ implies $\gcd(n, q-1) = \gcd(k, q-1) = 1$, hence we have $q - 1 \in \mathbb{Z}_n^*$. This entails $d(\mathcal{C}) \geq \delta = 3$. ♯

Example. The condition $\gcd(k, q-1) = 1$ cannot not be dispensed of, as the following example shows: Let $q := 3$ and $k := 2$; hence we have $n = \frac{q^2-1}{q-1} = 4$.

i) Let $\mathcal{C}$ be the cyclic code with defining set $\mathcal{V} = \{\zeta_4\}$; then we have $\mathcal{V}(\mathcal{C}) = \{\pm\zeta_4\} \subseteq \mathbb{F}_9$. The only (generalized) check condition is given by $[1, \zeta_4, -1, -\zeta_4] \in \mathbb{F}_9^4$, entailing $\mathcal{C} = \langle [1,0,1,0], [0,1,0,1] \rangle_{\mathbb{F}_3}$, which is a $[4,2,2]$-code.

ii) Let $\mathcal{H}_2$ be the associated ternary Hamming $[4,2,3]$-code. Assume that $\mathcal{H}_2$ is cyclic, then we have $|\mathcal{V}(\mathcal{H}_2)| = 2$. Since $X^4-1 = (X-1)(X+1)(X^2+1) \in \mathbb{F}_3[X]$, we have $\mathcal{V}(\mathcal{H}_2) = \{\pm\zeta_4\}$ or $\mathcal{V}(\mathcal{H}_2) = \{\pm 1\}$. In the first case we have already seen that this defines a $[4,2,2]$-code; in the second case we get the check matrix

$$\mathcal{H}(\pm 1) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \in \mathbb{F}_3^{2\times 4},$$

thus $\mathcal{C} = \langle [1,0,-1,0], [0,1,0,-1] \rangle_{\mathbb{F}_3}$, again a $[4,2,2]$-code; a contradiction. ♯

(12.6) BCH-Codes [Bose–Ray-Chaudhuri, 1960; Hocquenghem, 1959].
a) A cyclic code $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q, n) = 1$, having a (genuinely) consecutive defining 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. Hence for the minimum distance of $\mathcal{C}$ we have the **BCH bound** $d(\mathcal{C}) \geq \delta$.

In particular, for $\delta = 1$ we get $\mathcal{V}(\mathcal{C}) = \emptyset$, thus $\mathcal{C} = \mathbb{F}_q^n$; and for $\delta = n+1$ we get $\mathcal{V}(\mathcal{C}) = \mathcal{V}_n$, thus $\mathcal{C} = \{0\}$. But in general $\mathcal{C}$ 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 thus occurring is called the **Bose distance**.

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, 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 consider the minimum distance of BCH codes: In general, it might be strictly larger than the Bose distance; for example, the binary Golay code $\mathcal{G}_{23} \leq \mathbb{F}_2^{23}$ is a narrow sense BCH code having Bose distance 5 and minimum distance 7, see (16.1). But at least we have the following:

Theorem: Peterson [1967]. A narrow sense BCH code $\mathcal{C} \leq \mathbb{F}_q^n$ of designed distance $\delta \mid n$ has minimum distance $d(\mathcal{C}) = \delta$.

Proof. By the BCH bound we have to show that $d(\mathcal{C}) \leq \delta$. To this end, 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 $\{\zeta_n, \ldots, \zeta_n^{\delta-1}\} \subseteq \mathcal{V}(\sum_{i=0}^{\delta-1} X^{il})$.

This implies that $\mathcal{V}(\mathcal{C}) = \bigcup_{i=1}^{\delta-1}(\zeta_n^i \cdot \Gamma) \subseteq \mathcal{V}(\sum_{i=0}^{\delta-1} X^{il}) = \mathcal{V}(\psi(v))$, 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. We consider narrow sense primitive binary BCH codes, where for $k \in \{2,3,4\}$ and $n = 2^k - 1$ we have the following:

We have $X^3+1 = \mu_0 \cdot \mu_1 = (X+1)(X^2+X+1) \in \mathbb{F}_2[X]$, where the Frobenius orbits are given as $\mathbb{Z}_3 = \{0\} \dot\cup \{1,2\}$, and $X^7+1 = \mu_0 \cdot \mu_1\mu_3 = (X+1) \cdot (X^3+X+1)(X^3+X^2+1) \in \mathbb{F}_2[X]$, where the Frobenius orbits are given as $\mathbb{Z}_7 = \{0\} \dot\cup \{1,2,4\} \dot\cup \{3,5,6\}$; see (12.1).

Moreover, we have $X^{15}+1 = \mu_0 \cdot \mu_5 \cdot \mu_3 \cdot \mu_1\mu_7$, where the Frobenius orbits are given as $\mathbb{Z}_{15} = \{0\} \dot\cup \{1,2,4,8\} \dot\cup \{3,6,9,12\} \dot\cup \{5,10\} \dot\cup \{7,11,13,14\}$, and

$$\begin{aligned}
\mu_0 &= X+1;\\
\mu_5 &= \textstyle\prod_{i\in\{5,10\}}(X-\zeta_{15}^i) = \prod_{i\in\{1,2\}}(X-\zeta_3^i) = X^2+X+1;\\
\mu_3 &= \textstyle\prod_{i\in\{3,6,9,12\}}(X-\zeta_{15}^i) = \prod_{i\in\{1,2,3,4\}}(X-\zeta_5^i)\\
&= (X^4+X^3+X^2+X+1);\\
\mu_1 &= \textstyle\prod_{i\in\{1,2,4,8\}}(X-\zeta_{15}^i),\\
\mu_7 &= \textstyle\prod_{i\in\{7,11,13,14\}}(X-\zeta_{15}^i) = \prod_{i\in\{1,2,4,8\}}(X-\zeta_{15}^{-i}),\\
\mu_1\mu_7 &= (X^4+X+1)(X^4+X^3+1).
\end{aligned}$$

The associated narrow sense primitive binary BCH codes $\mathcal{C}$ are given in Table 8, where we indicate the Bose distance δ, the generator polynomial $g \in \mathbb{F}_2[X]$, the union $\mathcal{O}$ of Frobenius orbits associated with $\mathcal{V}(\mathcal{C})$, the dimension $\dim_{\mathbb{F}_2}(\mathcal{C}) = n - \deg(g) = n - |\mathcal{O}|$, and the actual minimum distance d.

In all cases given, except the trivial codes, we observe that $\delta = d$: The case $\delta = 1$ is trivial anyway; for $[k,\delta] \in \{[2,3],[3,7],[4,3],[4,5],[4,15]\}$ this follows from Peterson's Theorem; for $[k,\delta] \in \{[3,3],[4,7]\}$ this follows from (13.6) below; alternatively, for $\delta = 3$ the code in question, being defined by $\{\zeta_n\}$, is linearly equivalent to a Hamming code, having minimum distance $d = 3$; see (12.5).

(12.7) Reed-Solomon codes [1954]. a) We consider primitive BCH codes for $n := q-1$: In this case, 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 $\mathcal{C} \le \mathbb{F}_q^{q-1}$ is called a **Reed-Solomon code**.

Thus $\mathcal{V}(\mathcal{C})$ coincides with the 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 $\mathcal{C}$. Hence we have $k := \dim_{\mathbb{F}_q}(\mathcal{C}) = (q-1) - (\delta-1) = q-\delta$.

If $k \ge 1$, that is $\delta < q$, then from the Singleton and BCH bounds we get $\delta - 1 = (q-1) - k \ge d - 1 \ge \delta - 1$, where $d := d(\mathcal{C})$ is the minimum distance of $\mathcal{C}$, showing that $d = \delta$, implying that $\mathcal{C}$ is an MDS $[q-1, q-\delta, \delta]$-code.

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	∞

---

b) We describe a fast encoding procedure for $\mathcal{C}$, which already indicates the connection to the viewpoint of algebraic geometry:

We first observe that $\mathcal{C}^\perp$ is a Reed-Solomon code again, inasmuch

$$\begin{aligned}\mathcal{V}^\perp := \mathcal{V}(\mathcal{C}^\perp) = \mathbb{F}_q^* \setminus \mathcal{V}^{-1} &= \{\zeta_{q-1}^i \in \mathbb{F}_q^*; i \notin \{-a,\ldots,-a-\delta+2\}\} \\ &= \{\zeta_{q-1}^i \in \mathbb{F}_q^*; i \in \{1-a,\ldots,q-\delta-a\}\}.\end{aligned}$$

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 $\mathcal{C}$:

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

Thus for $v := [a_0,\ldots,a_{q-\delta-1}] \in \mathbb{F}_q^{q-\delta}$ the associated codeword $[w_1,\ldots,w_{q-1}] = w := v \cdot H(\mathcal{V}^\perp) \in \mathbb{F}_q^{q-1}$ is given as follows: For $j \in \{1,\ldots,q-1\}$ we get

$$w_j = \sum_{i=0}^{q-\delta-1} a_i(\zeta_{q-1}^{j-1})^{i-(a-1)} = \sum_{i=0}^{q-\delta-1} a_i(\zeta_{q-1}^{j-1})^{i+(q-a)} = (X^{q-a}\psi(v))(\zeta_{q-1}^{j-1}),$$

saying that w is obtained by evaluating $X^{q-a}\psi(v) \in \mathbb{F}_q[X]$ at all places of $\mathbb{F}_q^*$.

Thus running through all polynomials in $\mathbb{F}_q[X]_{<q-\delta}$ we get

$$\mathcal{C} = \{[\zeta_{q-1}^{(1-a)(j-1)} \cdot \psi(v)(\zeta_{q-1}^{j-1})]_j \in \mathbb{F}_q^{q-1}; v \in \mathbb{F}_q^{q-\delta}\} \leq \mathbb{F}_q^{q-1}.$$

(12.8) Remark: Shortening MDS codes. 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 such that $k \geq 2$; that is we have $d - 1 = n - k$. Then for the 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 $d - 1 \leq d^\circ - 1 \leq (n-1) - k^\circ \leq (n-1) - (k-1) = n - k$. Thus we have equality throughout, implying $k^\circ = k - 1$ and $d^\circ = d$, so that $\mathcal{C}^\circ$ is an MDS $[n-1, k-1, d]$-code as well. ♯

Now, starting with a Reed-Solomon $[q-1, q-\delta, \delta]$-code, successive shortening yields an MDS $[q-1-i, q-\delta-i, \delta]$-code, for $i \in \{0, \ldots, q-\delta-1\}$.

For example, for $q := 2^8 = 256$, hence $n = q - 1 = 255$, and designed distance $\delta = 5$, starting with the narrow sense Reed-Solomon $[255, 251, 5]$-code, thus having defining set $\{\zeta_{255}, \ldots, \zeta_{255}^4\}$, we get the 2-error correcting $[32, 28, 5]$- and $[28, 24, 5]$-codes over $\mathbb{F}_{256}$ being used in the following application:

(12.9) Example: The Audio Compact Disc [1982]. The **Red Book Standard**, 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 ~ 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 $9408 \cdot 300\text{nm} = 2822400\text{nm} \sim 2.8\text{mm}$.

13 Minimum distance of BCH codes

We have already remarked that the BCH bound for the minimum distance of a BCH code is not necessarily sharp. In view of this, we 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.

(13.1) 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 $i \in \{1,\dots,r\}$ and $i' \in \{1,\dots,s\}$. Note that $A * B$ in general does not have full rank, even if A and B have.

Theorem. Let $v \in \mathcal{C} := \ker((A * B)^{\mathrm{tr}}) \leq \mathbb{F}_q^n$, 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 $\mathcal{J} := \mathrm{supp}(v) \subseteq \{1,\dots,n\}$. Then we have

$$\mathrm{rk}_{\mathbb{F}_q}(A_{\mathcal{J}}) + \mathrm{rk}_{\mathbb{F}_q}(B_{\mathcal{J}}) \leq |\mathcal{J}| = \mathrm{wt}(v).$$

Proof. Let $v = [x_1,\dots,x_n]$, where we may assume that $\mathcal{J} = \{1,\dots,n\}$, that is $x_j \neq 0$, for $j \in \{1,\dots,n\}$. Letting $B' := B\cdot\mathrm{diag}[x_1,\dots,x_n] = [b_{i'j}x_j]_{i'j} \in \mathbb{F}_q^{s\times n}$, we have $\mathrm{rk}_{\mathbb{F}_q}(B) = \mathrm{rk}_{\mathbb{F}_q}(B') \in \mathbb{N}_0$. The condition

$$v\cdot(A * B)^{\mathrm{tr}} = [\sum_{j=1}^{n} a_{ij}b_{i'j}x_j]_{(i-1)s+i'} = 0 \in \mathbb{F}_q^{rs}$$

can be rewritten as $A \cdot B'^{\mathrm{tr}} = 0 \in \mathbb{F}_q^{r\times s}$. Thus the row space of A is orthogonal to the row space of B', hence we have $\mathrm{rk}_{\mathbb{F}_q}(A) \leq n - \mathrm{rk}_{\mathbb{F}_q}(B') = n - \mathrm{rk}_{\mathbb{F}_q}(B)$. ♯

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

(13.2) Theorem: Roos bound [1983]. For $n \in \mathbb{N}$ let $\mathcal{V}' \subseteq \mathcal{V}_n \subseteq \mathbb{F}_q(\zeta_n) =: \mathbb{F}$, where $\gcd(q,n) = 1$, be a consecutive set of length $\delta-1$, where $\delta \in \{2,\dots,n+1\}$. 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 $H(\mathcal{V}') * H(\mathcal{V}'') \in \mathbb{F}^{|\mathcal{V}'|\cdot|\mathcal{V}''|\times n}$ and $H(\mathcal{V}) \in \mathbb{F}^{|\mathcal{V}|\times n}$ have the same set of rows, we have $\mathcal{C} = \ker(H(\mathcal{V})^{\mathrm{tr}}) \cap \mathbb{F}_q^n = \ker((H(\mathcal{V}') * H(\mathcal{V}''))^{\mathrm{tr}}) \cap \mathbb{F}_q^n$. We apply the van-Lint–Wilson bound:

For $\emptyset \neq \mathcal{I} \subseteq \{1,\dots,n\}$, by the BCH bound we have $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}')_{\mathcal{I}}) = |\mathcal{I}|$, for $|\mathcal{I}| \leq \delta - 1$. Since we always have $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}'')_{\mathcal{I}}) \geq 1$, we get $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}')_{\mathcal{I}}) + \mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}'')_{\mathcal{I}}) > |\mathcal{I}|$, for $|\mathcal{I}| \leq \delta - 1$.

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 we get $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{W})_{\mathcal{I}}) = |\mathcal{I}|$, for $|\mathcal{I}| \leq |\mathcal{W}|$. Since deleting the rows of $H(\mathcal{W})$ corresponding to $\mathcal{W} \setminus \mathcal{V}''$ yields $H(\mathcal{V}'')$, we infer $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}'')_{\mathcal{I}}) \geq |\mathcal{I}| - |\mathcal{W}| + |\mathcal{V}''|$.

Since $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}')_{\mathcal{I}}) \geq \delta - 1$ for $|\mathcal{I}| \geq \delta$, for $\delta \leq |\mathcal{I}| \leq |\mathcal{W}|$ we get

$$\mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}')_{\mathcal{I}}) + \mathrm{rk}_{\mathbb{F}}(H(\mathcal{V}'')_{\mathcal{I}}) \geq \delta - 1 + |\mathcal{I}| - |\mathcal{W}| + |\mathcal{V}''|.$$

The right hand side exceeds $|\mathcal{I}|$ if and only if $|\mathcal{W}| \leq |\mathcal{V}''| + \delta - 2$, in which case $d(\mathcal{C}) \geq |\mathcal{I}| + 1$. 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 step size in $\mathbb{Z}_n^*$ such that $|\mathcal{V}''| + \delta - 2 \leq n$. 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 step size and having length $|\mathcal{V}''| + \delta - 2$. Recall that the rank estimates for Delsarte matrices also hold for generalized consecutive sets. ♯

(13.3) Example. Let $q := 2$ and $n := 35$. Then $2 \in \mathbb{Z}_{35}^*$ has order 12, and the Frobenius orbits 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 let explicitly

$$\begin{aligned}\mu_1 &:= X^{12} + X^{11} + X^{10} + X^8 + X^5 + X^4 + X^3 + X^2 + 1,\\ \mu_3 &= X^{12} + X^{10} + X^9 + X^8 + X^7 + X^4 + X^2 + X + 1;\\ \mu_5 &= X^3 + X + 1,\\ \mu_{15} &= X^3 + X^2 + 1;\\ \mu_7 &= X^4 + X^3 + X^2 + X + 1.\end{aligned}$$

Then it turns out that

$$g = \mu_1\mu_5\mu_7 \mid f := X^{28} + X^{16} + X^{14} + X^{11} + X^7 + X + 1 \in \mathbb{F}_2[X],$$

that is $f(\zeta_{35}^i) = 0$, for $i \in \{1, 5, 7\}$. Hence $\mathcal{C}$ has an element of weight 7. ♯

(13.4) Example. Let $q := 2$ and $n := 127 = 2^7 - 1$. Then $2 \in \mathbb{Z}_{127}^*$ has order 7, and the Frobenius orbits 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

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

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

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

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 $\mathcal{C}$, where the BCH bound yields $d \geq 16$:

i) We first show that $\mathcal{C}$ is 4-divisible:

We consider the localization $\mathbb{F}_2[X^{\pm 1}] := \mathbb{F}_2[X]_X \subseteq \mathbb{F}_2(X)$. Let $\overline{\ }\colon \mathbb{F}_2[X^{\pm 1}] \to \overline{\mathbb{F}_2[X]} := \mathbb{F}_2[X]/\langle X^{127}-1\rangle \cong \bigoplus_{i\in\mathbb{Z}_{127}} \mathbb{F}_2[X]/\langle X - \zeta_{127}^i\rangle$ be the extension of the natural epimorphism, and using 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 $g(X)g(X^{-1}) = 0 \in \overline{\mathbb{F}_2[X]}$.

Let $v = [a_0,\ldots,a_{126}] \in \mathcal{C}$, let $\mathcal{J} := \operatorname{supp}(v) \subseteq \mathbb{Z}_{127}$, let $s := |\mathcal{J}|$, and let $f := \psi(v) = \sum_{i=0}^{126} a_i X^i \in \mathbb{F}_2[X]$. Then we have $f(X^{-1}) = \sum_{j=0}^{126} a_j X^{-j} \in \mathbb{F}_2[X^{\pm 1}]$. 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 $k \in \mathbb{Z}_{127}$ we have $\sum_{j\in\mathbb{Z}_{127}} a_{k+j}a_j = 0 \in \mathbb{F}_2$.

Thus $|\{[i,j] \in \mathcal{J}^2; i = j + k\}|$ is even. For $k = 0$ this says that s is even. Since for $k \neq 0$ we have $-k \neq k \in \mathbb{Z}_{127}$, we get $4 \mid |\{[i,j] \in \mathcal{J}^2; i \in \{j \pm k\}\}|$, thus $4 \mid |\{[i,j] \in \mathcal{J}^2; i \neq j\}| = s^2 - s = s(s-1)$, implying that $4 \mid s$.

Alternatively, we may proceed as follows: Since $\mathcal{C}$ is weakly self-dual, it suffices to show 4-divisibility for an $\mathbb{F}_2$-basis of $\mathcal{C}$. In turn, $\mathcal{C}$ being cyclic, it suffices to ensure this for the codeword associated with the generator polynomial g. The latter is given explicitly below, yielding a codeword of weight 48.

ii) We apply 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 $\mathcal{O}' + \mathcal{O}''$. 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 apply the van-Lint–Wilson bound: Let $\mathbb{F} := \mathbb{F}_2(\zeta_n) = \mathbb{F}_{128}$.

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 sets $\{16, \ldots, 29\}$ and $\{16, \ldots, 44\}$ are consecutive of length 14 and 29, respectively, for $\emptyset \neq \mathcal{I} \subseteq \{1, \ldots, 127\}$ we get $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}')_{\mathcal{I}}) \geq |\mathcal{I}|$ for $|\mathcal{I}| \leq 14$, and $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}')_{\mathcal{I}}) \geq 14$ for $|\mathcal{I}| = 15$, and $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}')_{\mathcal{I}}) \geq |\mathcal{I}| - 2$ for $16 \leq |\mathcal{I}| \leq 29$.

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 $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}'')_{\mathcal{I}}) \geq |\mathcal{I}|$ for $|\mathcal{I}| \leq 2$, and $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}'')_{\mathcal{I}}) \geq 2$ for $3 \leq |\mathcal{I}| \leq 7$, and $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}'')_{\mathcal{I}}) \geq |\mathcal{I}| - 6$ for $8 \leq |\mathcal{I}| \leq 9$, and $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}'')_{\mathcal{I}}) = 3$ for $|\mathcal{I}| \geq 10$; note that $|\{-1, \ldots, 7\} \setminus \{-1, 0, 7\}| = 6$.

Thus we have $\mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}')_{\mathcal{I}}) + \mathrm{rk}_{\mathbb{F}}(H(\mathcal{P}'')_{\mathcal{I}}) > |\mathcal{I}|$ whenever $|\mathcal{I}| \leq 29$, thus 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

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

Moreover, it turns out that $g \mid f \in \mathbb{F}_2[X]$, where

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

that is $f(\zeta_{127}^i) = 0$, for $i \in \mathcal{Q}$. Hence $\mathcal{C}$ has an element of weight 32. ♯

Now we turn to criteria ensuring that the BCH bound is actually sharp. To this end, we need an auxiliary construction:

(13.5) Newton identities. 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, 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** of degree k.

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 $\partial_X h = \sum_{j=1}^m (-1)^j j e_{m,j} X^{j-1}$, and the product rule yields

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

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 following **Newton identities**, for $i \in \mathbb{N}$:

$$\begin{aligned} &\textstyle\sum_{j=1}^i (-1)^{j-1} e_{m,i-j} p_{m,j} = i e_{m,i}, && \text{for } i \in \{1, \ldots, m\}, \\ &\textstyle\sum_{j=i-m}^i (-1)^{j-1} e_{m,i-j} p_{m,j} = 0, && \text{for } i \geq m+1. \end{aligned}$$

(13.6) 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) We have $\mathbb{F} := \mathbb{F}_q(\zeta_n) = \mathbb{F}_{q^s}$, hence the Frobenius orbit $\{\zeta_n, \zeta_n^q, \ldots, \zeta_n^{q^{s-1}}\} \subseteq \mathcal{V}_n$. Thus the elements $\{\zeta_n, \zeta_n^q, \ldots, \zeta_n^{q^{t-1}}\}$ are pairwise distinct, implying the invertibility of the Vandermonde matrix

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

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 $[a_0, \ldots, a_{t-1}] \in \mathbb{F}^t$. Let

$$f := X^{q^t} + \sum_{i=0}^{t-1} a_i X^{q^i} \in \mathbb{F}[X]$$

be the associated **q-linearized polynomial** of degree q^t; note that we have $f(ax + y) = af(x) + f(y) \in \mathbb{F}$, for $x, y \in \mathbb{F}$ 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 $j \in \{0, \ldots, t-1\}$, hence $\mathcal{V} := \{1, \zeta_n, \ldots, \zeta_n^{t-1}\}$ consists of zeroes of f.

Moreover, $\mathcal{V} \subseteq \mathbb{F}$ is $\mathbb{F}_q$-linearly independent: Let $[b_0, \ldots, b_{t-1}] \in \mathbb{F}_q^t$ such that $\sum_{j=0}^{t-1} b_j \zeta_n^j = 0 \in \mathbb{F}$, then we get $\sum_{j=0}^{t-1} b_j (\zeta_n^{q^i})^j = 0$, for $i \in \{0, \ldots, t-1\}$, that is $A \cdot [b_0, \ldots, b_{t-1}]^{\mathrm{tr}} = 0 \in \mathbb{F}^{t \times 1}$, implying that $[b_0, \ldots, b_{t-1}] = 0$.

Letting $V := \langle \mathcal{V} \rangle_{\mathbb{F}_q} \leq \mathbb{F}$, we have $|V| = q^t$, and f being $\mathbb{F}_q$-linear implies that V consists of zeroes of f, implying that f splits over $\mathbb{F}$ as $f = \prod_{c \in V} (X - c) \in \mathbb{F}$.

ii) Let $m := q^t$. Then we get

$$f = X^{q^t} + \sum_{i=0}^{t-1} a_i X^{q^i} = \prod_{c \in V} (X - c) = \sum_{j=0}^{m} (-1)^{m-j} e_{m,m-j}(V) X^j \in \mathbb{F}[X].$$

Thus we have $e_{m,i}(V) \neq 0$ possibly only for $i = q^t - q^j$, where $j \in \{0, \ldots, t\}$, hence $ie_{m,i}(V) = 0$ for $i \in \{0, \ldots, m-2\}$. From the Newton identities we by induction on $i \in \mathbb{N}$ get

$$\sum_{c \in V \setminus \{0\}} c^i = \sum_{c \in V} c^i = p_{m,i}(V) = 0, \quad \text{for } i \in \{1, \ldots, m-2\}.$$

We have $\mathbb{F}^* = \langle \zeta_n \rangle$. Hence let $v = [x_0, \ldots, x_{n-1}] \in \{0,1\}^n \subseteq \mathbb{F}_q^n$ such that $V \setminus \{0\} = \{\zeta_n^i \in \mathbb{F}; i \in \mathbb{Z}_n, x_i = 1\}$. Then we have $\mathrm{wt}(v) = |V| - 1 = m - 1 = \delta$,

and for $j \in \{1, \ldots, \delta - 1\}$ we get

$$\psi(v)(\zeta_n^j) = \sum_{i \in \mathbb{Z}_n, x_i = 1} \zeta_n^{ji} = \sum_{c \in V \setminus \{0\}} c^j = 0.$$

Since $\mathcal{C}$ is associated with $\{\zeta_n, \ldots, \zeta_n^{\delta-1}\} \subseteq \mathcal{V}_n$ this implies $v \in \mathcal{C}$. ♯

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 $q^t - 1 = q \cdot q^{t-1} - 1 \leq q\delta - 1$. ♯

(13.7) Remark. Finally, we just mention the following theorem (whose proof requires tools we do not have at our disposal here):

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

V GOLAY

14 Quadratic residue codes

(14.1) 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 $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p^* \to \{\pm 1\}$ 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 following shows that given q there are infinitely many p such that $\left(\frac{q}{p}\right)=1$:

Theorem: Quadratic Reciprocity Law [GAUSS, 1796]**.**
i) We have $\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}$, that is $\left(\frac{2}{p}\right)=1$ if and only if $p\equiv\pm1 \pmod 8$.
ii) Let q be odd. Then we have $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right)=(-1)^{\frac{(p-1)(q-1)}{4}}$; that is we have $\left(\frac{q}{p}\right)=\left(\frac{p}{q}\right)$, except both $p,q\equiv -1 \pmod 4$. ♯

c) Let $\gamma_p:=\sum_{i\in\mathbb{Z}_p^*}\left(\frac{i}{p}\right)\zeta_p^i\in\mathbb{F}_q(\zeta_p)=:\mathbb{F}$ be the p-th **Gaussian sum**.

Lemma. i) We have $\gamma_p^2=\left(\frac{-1}{p}\right)\cdot p\in\mathbb{F}_q$; in particular we have $\gamma_p\neq 0$.
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 a bijection on $\mathbb{Z}_p^*$, 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 get $\sum_{j\in\mathbb{Z}_p}(\zeta_p^i)^j=0$, for $i\in\mathbb{Z}_p^*$. 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 $\mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F})=\langle\varphi_q\rangle$, we show that $\gamma_p^q=\gamma_p\in\mathbb{F}$: 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 $\gamma_p\in\mathbb{F}_q$.

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

$$\rho_p\eta_p = \prod_{i\in\mathbb{Z}_p^*}(X - \zeta_p^i) = \frac{X^p - 1}{X - 1} = \sum_{i=0}^{p-1} X^i \in \mathbb{F}_q[X].$$

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

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 $\psi(v)(1) = (\sum_{i=0}^{p-1} a_i X^i)(1) = 0$, that is $X - 1 \mid \psi(v) \in \mathbb{F}_q[X]$. 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

$$\widehat{\mathcal{C}^p} := \{[a_0, \ldots, a_{p-1}, a_\infty] \in \mathbb{F}_q^{p+1}; [a_0, \ldots, a_{p-1}] \in \mathcal{C}^p, a_\infty = \frac{\epsilon\gamma_p}{p} \cdot \sum_{i=0}^{p-1} a_i\},$$

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 (14.4) below. In particular, in the binary case $q = 2$ and in the ternary case $q = 3$ there is $\epsilon \in \{\pm 1\}$ such that the check condition becomes $a_\infty + \sum_{i=0}^{p-1} a_i = 0$, so that in both cases we recover the conventional extended code.

(14.3) 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 $\varphi_2 \in \mathrm{Aut}_{\mathbb{F}_2}(\mathbb{F}_8)$ has order 3. We conclude that $2 \in \mathcal{Q}_7$, that is $\left(\frac{2}{7}\right) = 1$. Hence the Frobenius orbits are $\mathcal{V}_7 = \{1\} \dot\cup \{\zeta_7^i; i \in \mathcal{Q}_7\} \dot\cup \{\zeta_7^i; i \in \mathcal{N}_7\}$, 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 (12.1).

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 $\widehat{\mathcal{C}} \leq \mathbb{F}_2^8$, where by (14.4) below we get $(\widehat{\mathcal{C}})^\perp = \widehat{\mathcal{C}}$, that is $\widehat{\mathcal{C}}$ is self-dual.

The code $\mathcal{C}$ has defining set $\{\zeta_7\}$, so that by (12.5) 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 $\widehat{\mathcal{H}}_3$.

(14.4) 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 $(\mathcal{Q}^p)^\perp = (\mathcal{Q}^p)'$, and $(\widehat{\mathcal{Q}}^p)^\perp = \widehat{\mathcal{Q}}^p$ with either choice of ϵ.
b) If $p \equiv 1 \pmod 4$, then we have $(\mathcal{Q}^p)^\perp = (\mathcal{N}^p)'$, and $(\widehat{\mathcal{Q}}^p)^\perp = \widehat{\mathcal{N}}^p$ with opposite choices of ϵ.

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

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 $\rho_p^* \sim \eta_p \in \mathbb{F}_q[X]$. Since $\mathcal{Q}^p$ has generator polynomial ρ_p, it follows that $(\mathcal{Q}^p)^\perp$ 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 $\eta_p^* \sim \eta_p$. It follow that $(\mathcal{Q}^p)^\perp$ 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 $1_p \in \mathbb{F}_q^p$ we have $\psi(1_p) = \sum_{i=0}^{p-1} X^i = \rho_p\eta_p$, thus $1_p \in \mathcal{C}^p$. But since $X-1 \nmid \psi(1_p)$ we have $1_p \notin (\mathcal{C}^p)'$. We recover $\mathcal{C}^p$ by augmenting $(\mathcal{C}^p)'$, 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 $\mathcal{C}^p$.

For $[a_0,\ldots,a_{p-1}] \in (\mathcal{C}^p)'$ we have $\sum_{i=0}^{p-1} a_i = 0$, hence $[a_0,\ldots,a_{p-1},0] \in \widehat{\mathcal{C}^p}$, and for $1_p \in \mathcal{C}^p$ we get $v := [1,\ldots,1,\epsilon\gamma_p] \in \widehat{\mathcal{C}^p}$. Hence we have the following generator matrix of $\widehat{\mathcal{C}^p}$:

$$\widehat{G} := \left[\begin{array}{c|c} G' & 0_{\frac{p-1}{2}}^{\mathrm{tr}} \\ \hline 1_p & \epsilon\gamma_p \end{array}\right] \in \mathbb{F}_q^{\frac{p+1}{2}\times(p+1)}.$$

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

If $p \equiv 1 \pmod 4$, then still $\langle \mathcal{Q}^p, (\mathcal{N}^p)' \rangle = \{0\}$, but now we get

$$\langle [1, \ldots, 1, \epsilon\gamma_p], [1, \ldots, 1, -\epsilon\gamma_p] \rangle = p - \gamma_p^2 = \big(1 - \left(\frac{-1}{p}\right)\big) \cdot p = 0 \in \mathbb{F}_q.$$

Hence we have $\widehat{\mathcal{Q}}^p \leq (\widehat{\mathcal{N}}^p)^\perp$ with opposite choices of ϵ, thus $\dim_{\mathbb{F}_q}(\widehat{\mathcal{Q}}^p) = \frac{p+1}{2} = \dim_{\mathbb{F}_q}((\widehat{\mathcal{N}}^p)^\perp)$ entails equality. ♯

(14.5) 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, so are $(\mathcal{Q}^p)'$ and $(\mathcal{N}^p)'$, and so are $\widehat{\mathcal{Q}}^p$ and $\widehat{\mathcal{N}}^p$ with either choice of ϵ.

b) For $v \in \mathcal{Q}^p \setminus (\mathcal{Q}^p)'$ and $d := \mathrm{wt}(v)$ 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 additionally $q = 2$ then we have $d \equiv 3 \pmod 4$. (Note that in the latter case we necessarily have $p \equiv -1 \pmod 8$.)

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 $i \in \mathbb{Z}_p^*$, we conclude that $\pi\colon \mathbb{Z}_p \to \mathbb{Z}_p\colon i \mapsto ij$ interchanges $\mathcal{Q}_p$ and $\mathcal{N}_p$, keeping 0 fixed. We consider the linear isometry given 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 $\psi(v^\pi) = \sum_{i\in\mathbb{Z}_p} a_{i\pi^{-1}} X^i = \sum_{i\in\mathbb{Z}_p} a_i X^{i\pi} = \sum_{i\in\mathbb{Z}_p} a_i X^{ij} \in \overline{\mathbb{F}_q[X]} := \mathbb{F}_q[X]/\langle X^p - 1\rangle$. Since evaluation at a p-th root of unity factors through $\overline{\mathbb{F}_q[X]}$, for $k \in \mathbb{Z}_p$ we get $\psi(v^\pi)(\zeta_p^k) = \sum_{i\in\mathbb{Z}_p} a_i \zeta_p^{ijk} = \sum_{i\in\mathbb{Z}_p} a_i (\zeta_p^{kj})^i = \psi(v)(\zeta_p^{kj}) \in \mathbb{F}_q(\zeta_p)$, thus $\zeta_p^k \in \mathcal{V}(\psi(v^\pi))$ if and only if $\zeta_p^{kj} \in \mathcal{V}(\psi(v))$, hence $v \in \mathcal{Q}^p$ 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 $\rho_p \mid \psi(v) \in \mathbb{F}_q[X]$, but $(X-1) \nmid \psi(v)$. Recalling $\psi(v^\pi)(\zeta_p^k) = \psi(v)(\zeta_p^{kj})$, for $k \in \mathbb{Z}_p$, we get $\eta_p \mid \psi(v^\pi)$, but $(X-1) \nmid \psi(v^\pi)$. Hence we have $\sum_{i=0}^{p-1} X^i = \rho_p\eta_p \mid \psi(v)\psi(v^\pi)$, but $X^p - 1 = (X-1)\cdot\rho_p\eta_p \nmid \psi(v)\psi(v^\pi)$.

Let $w \in \mathbb{F}_q^p$ be the vector associated with $\psi(v)\psi(v^\pi) \in \mathbb{F}_q[X]$. Then we have $w \neq 0$, and since $\sum_{i=0}^{p-1} X^i$ generates the repetition code in $\mathbb{F}_q^p$ we conclude that $w = [a, \ldots, a]$ for some $0 \neq a \in \mathbb{F}_q$, hence $\mathrm{wt}(w) = p$.

Now for $\psi(v)\psi(v^\pi)$ we get d^2 products of non-zero coefficients of $\psi(v)$ and $\psi(v^\pi)$, respectively. Hence $\psi(v)\psi(v^\pi)$ 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 $\psi(v)\psi(v^\pi)$, hence the latter has at most $d^2 - d + 1$ non-zero coefficients.

Finally, if additionally $q = 2$ 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 $\psi(v)\psi(v^\pi)$ cancel, for $i \in \mathbb{Z}_p^*$, 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 $d^2 - d + 1 \equiv p \equiv -1 \pmod 4$, hence $d(d-1) \equiv 2 \pmod 4$, which implies $d \equiv 3 \pmod 4$. ♯

For the next result we already need the Gleason-Prange Theorem (15.3) 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.

Corollary. We have $d((\mathcal{Q}^p)') = d(\mathcal{Q}^p) + 1$; in particular, the assertions of the square root bound 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 (15.3) 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 (15.3) 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 $v \in (\mathcal{Q}^p)'$. ♯

15 Automorphisms of quadratic residue codes

(15.1) Automorphisms of codes. a) We need an additional general piece of notation: Given a linear code $\mathcal{C} \leq \mathbb{F}_q^n$, let $A(\mathcal{C}) := \mathrm{Aut}_{\mathbb{F}_q}(\mathcal{C}) \leq \mathrm{Isom}_n(\mathbb{F}_q) \cong (\mathbb{F}_q^*)^n \rtimes \mathcal{S}_n$ be its linear automorphism group. Then let

$$P(\mathcal{C}) := A(\mathcal{C})/(A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n) \cong (A(\mathcal{C}) \cdot (\mathbb{F}_q^*)^n)/(\mathbb{F}_q^*)^n \leq \mathrm{Isom}_n(\mathbb{F}_q)/(\mathbb{F}_q^*)^n \cong \mathcal{S}_n$$

be the group of component permutations induced by $A(\mathcal{C})$, and let $\bar{}: A(\mathcal{C}) \to P(\mathcal{C})$ 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 $\mathrm{supp}(g) := \mathrm{supp}(\psi^{-1}(g)) \subseteq \mathbb{Z}_n$ be the **support** of g; note that $0 \in \mathrm{supp}(g)$. Moreover, let $\langle \mathrm{supp}(g) \rangle = \gamma\mathbb{Z}_n \trianglelefteq \mathbb{Z}_n$, 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 $\mathrm{diag}[a_{i+1}, \ldots, a_{n-1}, a_0, \ldots, a_i] \in A(\mathcal{C})$, 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 $i \in \mathrm{supp}(g)$. Transporting the action of D to $\mathbb{F}_q[X]_{<n}$, we have $g \cdot D = \sum_{i=0}^{k} a_i g_i X^i \in \psi(\mathcal{C})$, that is $g \mid g \cdot D$, entailing $g \cdot D \sim g \in \mathbb{F}_q[X]$, implying that $a_0 = a_i$. Thus the diagonal entries of D are constant along the cosets of $\gamma\mathbb{Z}_n$ in $\mathbb{Z}_n$.

Conversely, let $D := \mathrm{diag}[a_0, \ldots, a_{n-1}] \in (\mathbb{F}_q^*)^n$ have diagonal entries constant along the cosets of $\gamma\mathbb{Z}_n$ in $\mathbb{Z}_n$. We consider the $\mathbb{F}_q$-basis $\{X^j g \in \mathbb{F}_q[X]_{<n}; j \in \{0, \ldots, n-k-1\}\} \subseteq \psi(\mathcal{C})$. Then $(X^j g) \cdot D \in \psi(\mathcal{C})$, that is $g \mid (X^j g) \cdot D$. Since $X^j \mid (X^j g) \cdot D$ we infer that $(X^j g) \cdot D \sim X^j g \in \mathbb{F}_q[X]$, entailing $D \in A(\mathcal{C})$.

Hence in conclusion we have $A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n \cong (\mathbb{F}_q^*)^{[\mathbb{Z}_n : \gamma\mathbb{Z}_n]} \cong (C_{q-1})^\gamma$; in particular, we have $A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n = \mathbb{F}_q^* \cdot E_n$ if and only if $\gamma = 1$. ♯

(15.2) 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 $A(\mathcal{Q}^p) \cap (\mathbb{F}_q^*)^p = A((\mathcal{Q}^p)') \cap (\mathbb{F}_q^*)^p = \mathbb{F}_q^* \cdot E_p$.

For the extended QR code $\widehat{\mathcal{Q}}^p \leq \mathbb{F}_q^{p+1}$ we get: If $D := \mathrm{diag}[a_0, \ldots, a_{p-1}, a_\infty] \in A(\widehat{\mathcal{Q}}^p) \cap (\mathbb{F}_q^*)^{p+1}$, then since $(\widehat{\mathcal{Q}}^p)^\bullet = \mathcal{Q}^p$ we have $\mathrm{diag}[a_0, \ldots, a_{p-1}] \in A(\mathcal{Q}^p)$, 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 linear automorphism $\alpha \in A(\mathcal{Q}^p) \leq \mathrm{Isom}_p(\mathbb{F}_q)$ extends to a linear automorphism $\widehat{\alpha} \in A(\widehat{\mathcal{Q}}^p) \leq \mathrm{Isom}_{p+1}(\mathbb{F}_q)$; in other words we have $A(\mathcal{Q}^p) = \mathrm{Stab}_{A(\widehat{\mathcal{Q}}^p)}(\langle [0_p \mid 1] \rangle_{\mathbb{F}_q})$ and $P(\mathcal{Q}^p) = \mathrm{Stab}_{P(\widehat{\mathcal{Q}}^p)}(\infty)$:

Let first $p \equiv -1 \pmod 4$. Then $(\mathcal{Q}^p)^\perp = (\mathcal{Q}^p)'$, so that α restricts to a linear 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 $w \in (\mathcal{Q}^p)'$.

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 $\widehat{\alpha} \in \mathrm{Isom}_{p+1}(\mathbb{F}_q)$ 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 a linear automorphism extending α.

Let now $p \equiv 1 \pmod 4$. Then $(\mathcal{Q}^p)^\perp = (\mathcal{N}^p)'$, so that α yields a linear automor-

phism 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 $w \in (\mathcal{N}^p)'$.

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 $\widehat{\alpha} \in \mathrm{Isom}_{p+1}(\mathbb{F}_q)$ 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 a linear automorphism extending α, which hence yields an a linear automorphism of $\widehat{\mathcal{Q}}^p = (\widehat{\mathcal{N}}^p)^\perp$. ♯

(15.3) The Gleason-Prange Theorem. Let p be an odd prime and $q \neq p$ be a prime such that $\left(\frac{q}{p}\right) = 1$.

Lemma. 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 $i \in \mathbb{Z}_p^*$, and $b_0 := -\epsilon \left(\frac{-1}{p}\right) a_\infty$ 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 $\sigma^2 = \left(\frac{-1}{p}\right) \cdot \mathrm{id}_{\mathbb{F}_q^{p+1}}$, where $\mathrm{tr}(\sigma) = 0$ if $p \equiv -1 \pmod 4$, and $\mathrm{tr}(\sigma) = 2\left(\frac{\zeta_4}{p}\right)$ if $p \equiv 1 \pmod 4$; in the latter case $\zeta_4 \in \mathbb{F}_p^*$ denotes a primitive 4-th root of unity.

Proof. a) For $v := [a_0, \ldots, a_{p-1}, a_\infty]$ we write $w := \sigma(v) = [b_0, \ldots, b_{p-1}, b_\infty]$. 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

$$\sum_{j \in \mathbb{Z}_p} \alpha_j \zeta_p^{-ij} = \sum_{k \in \mathbb{Z}_p} (a_k \cdot \sum_{j \in \mathbb{Z}_p} \zeta_p^{(k-i)j}) = p a_i.$$

Now $v \in \widehat{\mathcal{Q}}^p$ is equivalent to $\alpha_j = 0$, for $j \in \mathcal{Q}_p$, and $a_\infty = \frac{\epsilon\gamma_p}{p} \cdot \sum_{i \in \mathbb{Z}_p} a_i = \frac{\epsilon\gamma_p \alpha_0}{p}$. Since σ is a monomial $\mathbb{F}_q$-linear map, we have $\sigma \in \mathrm{Isom}_{p+1}(\mathbb{F}_q)$. Hence we have to show that if $v \in \widehat{\mathcal{Q}}^p$ then $w \in \widehat{\mathcal{Q}}^p$ as well:

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

$$\beta := \sum_{i \in \mathbb{Z}_p^*} b_i = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-i^{-1}}{p}\right) a_{-i^{-1}} = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) a_i = \frac{1}{p} \cdot \sum_{j \in \mathbb{Z}_p} (\alpha_j \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^{-ij}).$$

For $j = 0$ we get $\sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) = 0$ in the inner sum, hence

$$\beta = \frac{1}{p} \left(\frac{-1}{p}\right) \cdot \sum_{j \in \mathbb{Z}_p^*} (\alpha_j \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-i}{p}\right) \zeta_p^{-ij}).$$

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

$$\sum_{i\in\mathbb{Z}_p} b_i = b_0 + \beta = \left(\frac{-1}{p}\right)(-\epsilon a_\infty - \gamma_p a_0 + \epsilon a_\infty) = -\gamma_p \left(\frac{-1}{p}\right) a_0.$$

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 turn to showing $y = 0$: We have

$$y = \left(\frac{-1}{p}\right) \cdot \sum_{k\in\mathbb{Z}_p^*} \Big(\sum_{i\in\mathbb{Z}_p^*} \left(\frac{ki^{-1}}{p}\right) \zeta_p^{ki^{-1}+ij} \Big) \left(\frac{k}{p}\right) \alpha_k,$$

where $\alpha_k = 0$ for $k \in \mathcal{Q}_p$. We show that the inner sum vanishes for $k \in \mathcal{N}_p$:

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

$$\left(\frac{\pi(ij)}{p}\right) = \left(\frac{ki^{-1}}{p}\right) = \left(\frac{k}{p}\right)\left(\frac{i}{p}\right) = -\left(\frac{i}{p}\right) = -\left(\frac{i}{p}\right)\left(\frac{j}{p}\right) = -\left(\frac{ij}{p}\right).$$

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 $i\in\{\pm\zeta_4\}\subseteq\mathbb{Z}_p^*$.

If $p\equiv -1 \pmod 4$, that is $\left(\frac{-1}{p}\right)=-1$, then $\det(\sigma)=\det(\pm\begin{bmatrix}0&1\\-1&0\end{bmatrix})^{\frac{p+1}{2}}=1$, where $\operatorname{tr}(\sigma)=0$ 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 $\det(\sigma)=\left(\frac{\zeta_4}{p}\right)^2\cdot\det(\pm\begin{bmatrix}0&1\\1&0\end{bmatrix})^{\frac{p-1}{2}}=1$, where $\operatorname{tr}(\sigma)=2\left(\frac{\zeta_4}{p}\right)$ and $\sigma^2=\mathrm{id}_{\mathbb{F}_q^{p+1}}$. ♯

Theorem: Gleason–Prange [1964]. 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 $\bar{}\colon S\to S/Z(S)=\overline{S}:=\mathrm{PSL}_2(\mathbb{F}_p)$, the latter being the associated projective special linear group, having order $\frac{1}{2}p(p-1)(p+1)$. Moreover, S acts naturally on the projective line $\mathbf{P}^1(\mathbb{F}_p)=\{x_0,\dots,x_{p-1},x_\infty\}$, 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 an embedding of $\overline{S}$ into $\mathcal{S}_{p+1}$. More precisely:

We have $x_it=\langle[1,i]\rangle_{\mathbb{F}_q}\cdot t=\langle[1,i]\cdot t\rangle_{\mathbb{F}_q}=\langle[1,i+1]\rangle_{\mathbb{F}_q}=x_{i+1}$, 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 $\bar{t}\in\overline{S}$ induces the p-cycle $(0,\dots,p-1)\in\mathcal{S}_{p+1}$.

We have $x_ir=\langle[1,i]\rangle_{\mathbb{F}_q}\cdot r=\langle[1,i]\cdot r\rangle_{\mathbb{F}_q}=\langle[c,ic^{-1}]\rangle_{\mathbb{F}_q}=\langle[1,ic^{-2}]\rangle_{\mathbb{F}_q}=x_{ic^{-2}}$, 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 $\bar{r}\in\overline{S}$ induces the permutation $(1,c^{-2},\dots,c^2)(c^{-1},c^{-3},\dots,c)\in\mathcal{S}_{p+1}$ of order $\frac{p-1}{2}$, permuting the squares and non-squares in $\mathbb{F}_p^*$, respectively.

We have $x_is=\langle[1,i]\rangle_{\mathbb{F}_q}\cdot s=\langle[1,i]\cdot s\rangle_{\mathbb{F}_q}=\langle[-i,1]\rangle_{\mathbb{F}_q}=\langle[1,-i^{-1}]\rangle_{\mathbb{F}_q}=x_{-i^{-1}}$, for $i\in\mathbb{Z}_p^*$, while $x_0s=\langle[1,0]\rangle_{\mathbb{F}_q}\cdot s=\langle[1,0]\cdot s\rangle_{\mathbb{F}_q}=\langle[0,1]\rangle_{\mathbb{F}_q})=x_\infty$ 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 $\bar{s}\in\overline{S}$ 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 of $\bar{s}$ if and only if $i^2=-1\in\mathbb{Z}_p$, thus there are fixed points if and only if $p\equiv 1 \pmod 4$, in which case the fixed points are x_i for $i\in\{\pm\zeta_4\}$.

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 triangular 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 $(0, \ldots, p-1) \in \mathcal{S}_{p+1}$. 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 $\tau \in A(\widehat{\mathcal{Q}}^p)$, 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 $(1, c^{-2}, \ldots, c^2)(c^{-1}, c^{-3}, \ldots, c) \in \mathcal{S}_{p+1}$. For $[a_0, \ldots, a_{p-1}] \in \mathbb{F}_q^p$ we have

$$\psi(\rho([a_0, \ldots, a_{p-1}])) = a_0 + \sum_{i \in \mathbb{Z}_p^*} a_{ic^{-2}} X^i = a_0 + \sum_{i \in \mathbb{Z}_p^*} a_i X^{ic^2} \in \mathbb{F}_q[X].$$

Thus for $[a_0, \ldots, a_{p-1}] \in \mathcal{Q}^p$ and $k \in \mathcal{Q}_p$, noting $kc^2 \in \mathcal{Q}_p$, we get

$$\psi(\rho([a_0, \ldots, a_{p-1}]))(\zeta_p^k) = a_0 + \sum_{i=1}^{p-1} a_i (\zeta_p^{kc^2})^i = 0,$$

thus $\rho([a_0, \ldots, a_{p-1}]) \in \mathcal{Q}_p$. Since the extension condition is trivially fulfilled, we conclude that $\rho \in A(\widehat{\mathcal{Q}}^p)$, 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, 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 natural 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 $\overline{s} \mapsto \overline{\sigma}$, $\overline{t} \mapsto \overline{\tau}$, $\overline{r} \mapsto \overline{\rho}$. ♯

Remark. Actually, it turns out that, up to three exceptions, we have $P(\widehat{\mathcal{Q}}^p) = \overline{S}$, in which 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 $\widehat{\mathcal{Q}}^p \cong \widehat{\mathcal{H}}_3$ and $\mathcal{Q}^p \cong \mathcal{H}_3$, where we get

$$A(\widehat{\mathcal{Q}}^p) \cong P(\widehat{\mathcal{Q}}^p) \cong \mathrm{AGL}_3(\mathbb{F}_2) \cong C_2^3 \rtimes \mathrm{SL}_3(\mathbb{F}_2) \cong C_2^3 \rtimes \mathrm{PSL}_2(\mathbb{F}_7),$$

hence $A(\mathcal{Q}^p) \cong P(\mathcal{Q}^p) \cong C_2^3 \rtimes (C_7 \rtimes C_3)$, see (14.3);

ii) $[q, p] = [2, 23]$, in which case we have $\widehat{\mathcal{Q}}^p \cong \mathcal{G}_{24}$ and $\mathcal{Q}^p \cong \mathcal{G}_{23}$, see (16.1);

iii) $[q, p] = [3, 11]$, in which case we have $\widehat{\mathcal{Q}}^p \cong \mathcal{G}_{12}$ and $\mathcal{Q}^p \cong \mathcal{G}_{11}$, see (16.2).

(15.4) Remark. We keep the notation of (15.3). Having dealt with the group $\overline{A}$ of component permutations induced by A, we now consider the group $A := \langle \sigma, \tau, \rho \rangle \leq A(\widehat{\mathcal{Q}}^p) \cap \mathrm{SL}_{p+1}(\mathbb{F}_q)$ itself. 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 $A \cap (\mathbb{F}_q^*)^{p+1} \leq \{\pm E_{p+1}\} =: Z \cong C_2$.

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 or a non-split extension.

We show that indeed $A \cong Z.\overline{S}$, where to decide whether the extension is split or non-split, we have to distinguish the cases $p \equiv \pm 1 \pmod 4$, and to determine the character χ of Δ, using the ordinary character table of S:

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 this we determine the character χ:

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 $\chi|_{\overline{S}}$: The group $\overline{S}$ has the following irreducible characters, subscripts denoting degrees: The trivial character χ_1, the Steinberg character χ_p, two (algebraically conjugate) characters $\chi_{\frac{p+1}{2}}^{\pm 1}$, as well as $\frac{p-1}{4}$ characters χ_{p-1}^i for certain $i \in \mathbb{Z}_{p+1}$, and $\frac{p-5}{4}$ characters χ_{p+1}^j 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 $[a, \ldots, a, b] = [a, \ldots, a, b] \cdot (\pm\sigma) = \pm[-\epsilon b, \ldots, \left(\frac{i}{p}\right) a, \ldots, -\epsilon a]$, 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 that $\chi|_{\overline{S}} = \chi_{p+1}^j$, 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 $\chi_{p+1}^j(r) = \zeta_{p-1}^j + \zeta_{p-1}^{-j}$, where $\zeta_{p-1} \in \mathbb{C}$ is a primitive $(p-1)$-st root of unity. Now $\chi(\rho) = 2$ entails $\zeta_{p-1}^j = \zeta_{p-1}^{-j} \in \{\pm 1\}$, implying $\frac{p-1}{2} \mid j$, an invalid parameter, a contradiction.

Thus, taking character degrees into account, and recalling that χ is rational, we conclude that $\chi|_{\overline{S}} = \chi_{\frac{p+1}{2}}^1 + \chi_{\frac{p+1}{2}}^{-1}$. Since $\left(\frac{c}{p}\right) = -1$ we have $\chi_{\frac{p+1}{2}}^{\pm 1}(r) = -1$, 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 $\chi_{\frac{p+1}{2}}^{\pm 1}(s) = \left(\frac{\zeta_4}{p}\right)$, 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 $A \cong Z.\overline{S}$. For $p \neq 3$, since in this case $\overline{S}$ is perfect, so that S is the unique Schur representation group of $\overline{S}$, we conclude that $A \cong S$. For $p = 3$, an explicit check shows that $A \cong S$ as well. For completeness we determine χ:

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 that $\chi = \chi^j_{p+1}$, for some $j \in \mathbb{Z}_{p-1}$. The element ρ of order $\frac{p-1}{2}$ belongs to the conjugacy class of S of $r^2 = \mathrm{diag}[c^2, c^{-2}]$. We have $\chi^j_{p+1}(r^2) = \zeta^j_{\frac{p-1}{2}} + \zeta^{-j}_{\frac{p-1}{2}}$, 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 $\zeta^j_{\frac{p-1}{2}} = 1$, implying $\frac{p-1}{2} \mid j$, an invalid parameter, a contradiction.

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

16 Golay codes

(16.1) Binary Golay codes [1949]. a) 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. 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 $\mathcal{V}_{23} = \{1\} \dot\cup \{\zeta^i_{23}; i \in \mathcal{Q}_{23}\} \dot\cup \{\zeta^i_{23}; i \in \mathcal{N}_{23}\}$, where

$$\mathcal{Q}_{23} = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}.$$

Thus we have $X^{23}+1 = (X+1)\cdot\prod_{i\in\mathcal{Q}_{23}}(X+\zeta^i_{23})\cdot\prod_{i\in\mathcal{N}_{23}}(X+\zeta^i_{23}) = \mu_0\mu_1\mu_5 \in \mathbb{F}_2(\zeta_{23})[X]$. Actually we have $X^{23}+1 = (X+1)\cdot g'g'' \in \mathbb{F}_2[X]$, where $g' := X^{11}+X^9+X^7+X^6+X^5+X+1$ and $g'' := X^{11}+X^{10}+X^6+X^5+X^4+X^2+1$, 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 in $\mathbb{F}_2^{12\times 23}$ is 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 $\mathcal{G}^\perp_{23}$ has generator polynomial $h^* = (X+1)^*\cdot(g'')^* = (X+1)\cdot g'$, showing that $\mathcal{G}^\perp_{23} = \mathcal{G}'_{23}$; the associated check matrix 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 $\mathcal{G}^\perp_{23} = \mathcal{G}'_{23} \leq \mathcal{G}_{23}$.

Extending yields the **extended binary Golay code** $\mathcal{G}_{24} := \widehat{\mathcal{G}}_{23} \leq \mathbb{F}_2^{24}$. Puncturing yields $\mathcal{G}_{24}^\bullet = (\widehat{\mathcal{G}}_{23})^\bullet = \mathcal{G}_{23}$, and we have $\mathcal{G}_{24}^\perp = \mathcal{G}_{24}$, that is $\mathcal{G}_{24}$ is self-dual.

b) 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 Frobenius orbit $\mathcal{Q}_{23}$, hence has Bose distance $\delta = 5$, so that the BCH bound yields $d(\mathcal{G}_{23}) \geq 5$, thus $d(\mathcal{G}_{24}) \geq 6$.

Moreover, the generator matrix of $\mathcal{G}_{23}$ consists of rows of weight 7, hence the extended code $\mathcal{G}_{24} = \widehat{\mathcal{G}}_{23}$ an $\mathbb{F}_2$-basis consisting of vectors of weight 8. Thus $\mathcal{G}_{24}$, being self-dual, is doubly-even, implying $d(\mathcal{G}_{24}) \geq 8$ and $d(\mathcal{G}_{23}) \geq 7$. Note that, alternatively, the square root bound yields $d(\mathcal{G}_{23}) \geq \lceil\sqrt{23}\rceil = 5$ as well, where additionally we have $d(\mathcal{G}_{23}) \equiv 3 \pmod 4$, hence $d(\mathcal{G}_{23}) \geq 7$.

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

This allows us to determine the weight distribution of $\mathcal{G}_{24}$: The latter is a doubly-even self-dual binary code, that is of Type II. Hence, by Gleason's Theorem, see (10.7), its weight enumerator belongs to $\langle f^3, g\rangle_{\mathbb{C}} \in \mathbb{C}[X, Y]$, where $f := Y^8 + 14X^4Y^4 + X^8$ and $g := X^4Y^4(Y^4 - X^4)^4$. Since $d(\mathcal{G}_{24}) = 8$ we infer that $A_{\mathcal{G}_{24}} = f^3 - 42g$, so that $\mathcal{G}_{24}$ is an extremal code of Type II, where

$$A_{\mathcal{G}_{24}}(X, 1) = 1 + 759X^8 + 2576X^{12} + 759X^{16} + X^{24} \in \mathbb{C}[X].$$

c) It can be shown that $\mathcal{G}_{24}$ is the unique binary $[24, 12, 8]$-code, up to linear equivalence. Its linear automorphism group $A(\mathcal{G}_{24}) = P(\mathcal{G}_{24}) \leq \mathrm{Isom}_{24}(\mathbb{F}_2) \cong \mathcal{S}_{24}$ is the **sporadic simple Mathieu group** M_{24} of 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 $A(\mathcal{G}_{23}) = P(\mathcal{G}_{23}) \leq \mathrm{Isom}_{23}(\mathbb{F}_2) \cong \mathcal{S}_{23}$, coinciding with $\mathrm{Stab}_{P(\mathcal{G}_{24})}(\infty) \leq P(\mathcal{G}_{24})$ of index 24, is the **sporadic simple Mathieu group** M_{23} of 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 (27.5) and (27.6), where we give a brief indication.

(16.2) Ternary Golay codes [1949]. a) 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 $\varphi_3 \in \Gamma := \mathrm{Aut}_{\mathbb{F}_3}(\mathbb{F}_3(\zeta_{11}))$ has order 5. 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

$$\mathcal{Q}_{11} = \{1, 3, 4, 5, 9\}.$$

Thus we have $X^{11} - 1 = (X-1) \cdot \prod_{i \in \mathcal{Q}'_{11}} (X - \zeta_{11}^i) \cdot \prod_{i \in \mathcal{N}_{11}} (X - \zeta_{11}^i) = \mu_0\mu_1\mu_2 \in$

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

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

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

$\mathbb{F}_3(\zeta_{11})[X]$. We have $X^{11}-1=(X-1)\cdot g'g'' \in \mathbb{F}_3[X]$, where $g' := X^5 - X^3 + X^2 - X - 1$ and $g'' := X^5 + X^4 - X^3 + X^2 - 1$, hence the latter are 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 $G \in \mathbb{F}_3^{6\times 11}$ is 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 $\mathcal{G}_{11}^\perp = \mathcal{G}'_{11}$; 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 $\mathcal{G}_{11}^\perp = \mathcal{G}'_{11} \leq \mathcal{G}_{11}$.

Extending yields the **extended ternary Golay code** $\mathcal{G}_{12} := \widehat{\mathcal{G}}_{11} \leq \mathbb{F}_3^{12}$. Puncturing yields $\mathcal{G}_{12}^\bullet = (\widehat{\mathcal{G}}_{11})^\bullet = \mathcal{G}_{11}$i, and we have $\mathcal{G}_{12}^\perp = \mathcal{G}_{12}$, that is $\mathcal{G}_{12}$ is self-dual.

b) 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 Frobenius orbit $\mathcal{Q}_{11}$, hence has Bose distance $\delta = 4$, so that the BCH bound yields $d(\mathcal{G}_{11}) \geq 4$, thus $d(\mathcal{G}_{12}) \geq 4$.

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

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

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

Moreover, $\mathcal{G}_{12}$ is self-dual, hence is 3–divisible. Thus we have $d(\mathcal{G}_{12}) \geq 6$, hence $d(\mathcal{G}_{11}) \geq 5$. Note that the square root bound only yields $d(\mathcal{G}_{11}) \geq \lceil\sqrt{11}\rceil = 4$.

Since the generator matrix of $\mathcal{G}_{11}$ consists of rows of weight 5, we conclude that $d(\mathcal{G}_{11}) = 5$. Similarly, since $\mathcal{G}_{12}$ has a generator matrix consisting of rows of weight 6, we get $d(\mathcal{G}_{11}) = 6$. Hence $\mathcal{G}_{11}$ is a $[11, 6, 5]$-code, $\mathcal{G}_{12}$ is a $[12, 6, 6]$-code, and $\mathcal{G}'_{11}$ is an $[11, 5, 6]$-code, Finally, 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 $e := 2 = \frac{5-1}{2}$; hence $\mathcal{G}_{11}$ is perfect, and $\mathcal{G}_{12}$ is quasi-perfect. ♯

This allows us to determine the weight distribution of $\mathcal{G}_{12}$: The latter is a self-dual ternary code, that is of Type III. Hence, by Gleason's Theorem, see (10.7), its homogeneous weight enumerator belongs to $\langle f^3, g\rangle_{\mathbb{C}} \in \mathbb{C}[X, Y]$, where $f := Y(Y^3 + 8X^3)$ and $g := X^3(Y^3 - X^3)^3$. Since $d(\mathcal{G}_{12}) = 6$ we infer that $A_{\mathcal{G}_{12}} = f^3 - 24g$, so that $\mathcal{G}_{12}$ is an extremal code of Type III, where

$$A_{\mathcal{G}_{12}}(X, 1) = 24X^{12} + 440X^9 + 264X^6 + 1.$$

c) It can be shown that $\mathcal{G}_{12}$ is the unique ternary $[12, 6, 6]$-code, up to linear equivalence. Its linear automorphism group $A(\mathcal{G}_{12}) \leq \mathrm{Isom}_{12}(\mathbb{F}_3) \cong \{\pm 1\}^{12} \rtimes \mathcal{S}_{12}$ is the non-split two-fold central extension $2.M_{12}$ of the **sporadic simple Mathieu group** M_{12} of 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 $A(\mathcal{G}_{11}) \leq \mathrm{Isom}_{11}(\mathbb{F}_3) \cong \{\pm 1\}^{11} \rtimes \mathcal{S}_{11}$, coinciding with $\mathrm{Stab}_{A(\mathcal{G}_{12})}(\{[0, \ldots, 0, \pm 1]\}) \leq A(\mathcal{G}_{12})$ of index 12, is the direct product $2 \times M_{11}$, where M_{11} is the **sporadic simple Mathieu group** of order 7920; hence $P(\mathcal{G}_{11}) \cong M_{11}$.

(16.3) 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

i) 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	$q_i \cdot 3^{-13}$	$\mu_i \cdot 3^{13}$		
0	$\frac{21}{100}$	$\frac{21}{100}$	$\binom{13}{0}$	=	1
1	$\frac{3}{25}$	$\frac{3}{650}$	$2 \cdot \binom{13}{1}$	=	26
2	$\frac{3}{25}$	$\frac{1}{2600}$	$2^2 \cdot \binom{13}{2}$	=	312
3	$\frac{3}{20}$	$\frac{3}{45760}$	$2^3 \cdot \binom{13}{3}$	=	2288

ii) 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 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 plane $\mathbf{P}^2(\mathbb{F}_3)$ has $\frac{3^3-1}{3-1} = 13$ elements. Thus with $3^{10} = 59049$ bets, out of $3^{13} = 1594323$, it is possible to guess at least $13-1 = 12$ of the outcomes correctly. Assuming 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, and placing 3^{10} bets, we get an expected reward rate

$$\begin{aligned} & \tfrac{1}{3^{10}} \cdot \tfrac{1}{3^3} \cdot \textstyle\sum_{v \in \mathbb{F}_3^{13}/\mathcal{H}_3} \big(\sum_{i=0}^3 |\{w \in v+\mathcal{H}_3; \mathrm{wt}(w) = i\}| \cdot q_i \big) \\ = \; & \tfrac{1}{3^{13}} \cdot \textstyle\sum_{i=0}^3 q_i \cdot \big(\sum_{v \in \mathbb{F}_3^{13}/\mathcal{H}_3} |\{w \in v+\mathcal{H}_3; \mathrm{wt}(w) = i\}| \big) \\ = \; & \tfrac{1}{3^{13}} \cdot \textstyle\sum_{i=0}^3 q_i \cdot |\{w \in \mathbb{F}_3^{13}; \mathrm{wt}(w) = i\}| \\ = \; & \tfrac{1}{3^{13}} \cdot \textstyle\sum_{i=0}^3 q_i \cdot (\mu_i \cdot 3^{13}) \\ = \; & \textstyle\sum_{i=0}^3 p_i, \end{aligned}$$

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.

iii) Another strategy, using expert knowledge, is as follows: If the outcome of a single match is for sure, then we may puncture with respect to this component, and end up in $\mathbb{F}_3^{12}$. 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	$q_i \cdot 3^{-13}$	$\mu_i' \cdot 3^{12}$		
0	$\frac{21}{100}$	$\frac{21}{100}$	$\binom{12}{0}$	=	1
1	$\frac{3}{25}$	$\frac{3}{650}$	$2 \cdot \binom{12}{1}$	=	24
2	$\frac{3}{25}$	$\frac{1}{2600}$	$2^2 \cdot \binom{12}{2}$	=	264
3	$\frac{3}{20}$	$\frac{3}{45760}$	$2^3 \cdot \binom{12}{3}$	=	1760

We get an expected reward rate of

$$\begin{aligned} \textstyle\sum_{i=0}^3 \mu_i' q_i &= \textstyle\sum_{i=0}^3 \frac{\mu_i'}{\mu_i} \cdot p_i \\ &= 3 \cdot \textstyle\sum_{i=0}^3 \frac{\binom{12}{i}}{\binom{13}{i}} \cdot p_i \\ &= 3 \cdot (p_0 + \tfrac{12}{13} \cdot p_1 + \tfrac{11}{13} \cdot p_2 + \tfrac{10}{13} \cdot p_3) \\ &= \tfrac{2097}{1300}. \end{aligned}$$

Hence providing an expected reward rate of ~ 1.61 this is a sensible winning strategy, only depending on expert knowledge, but not on using any code.

iv) If even the outcome of two matches is for sure, then we similarly end up in $\mathbb{F}_3^{11}$. Looking for perfect codes 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 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.

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	$q_i \cdot 3^{-13}$	$\mu_i'' \cdot 3^{11}$		
0	$\frac{21}{100}$	$\frac{21}{100}$	$\binom{11}{0}$	$=$	1
1	$\frac{3}{25}$	$\frac{3}{650}$	$2 \cdot \binom{11}{1}$	$=$	22
2	$\frac{3}{25}$	$\frac{1}{2600}$	$2^2 \cdot \binom{11}{2}$	$=$	220
3	$\frac{3}{20}$	$\frac{3}{45760}$	$2^3 \cdot \binom{11}{3}$	$=$	1320

We get an expected reward rate of

$$\begin{aligned}\sum_{i=0}^3 \mu_i'' q_i &= \sum_{i=0}^3 \tfrac{\mu_i''}{\mu_i} \cdot p_i \\ &= 3^2 \cdot \sum_{i=0}^3 \tfrac{\binom{11}{i}}{\binom{13}{i}} \cdot p_i \\ &= 3^2 \cdot (p_0 + \tfrac{11}{13} \cdot p_1 + \tfrac{55}{78} \cdot p_2 + \tfrac{15}{26} \cdot p_3) \\ &= \tfrac{2259}{520}.\end{aligned}$$

Hence providing an expected reward rate of ~ 4.34 this is an even better winning strategy, only depending on expert knowledge, but not on using any code.

Similar to the attack above, now using the code $\mathcal{G}_{11}$, averaging over the $3^{11-6} = 3^5$ cosets, and placing 3^6 bets, we get an expected reward rate

$$\begin{aligned}&\tfrac{1}{3^6} \cdot \tfrac{1}{3^5} \cdot \textstyle\sum_{v \in \mathbb{F}_3^{11}/\mathcal{G}_{11}} \big(\sum_{i=0}^3 |\{w \in v + \mathcal{G}_{11}; \mathrm{wt}(w) = i\}| \cdot q_i \big) \\ = \;& \tfrac{1}{3^{11}} \cdot \textstyle\sum_{i=0}^3 q_i \cdot |\{w \in \mathbb{F}_3^{11}; \mathrm{wt}(w) = i\}| \\ = \;& \tfrac{1}{3^{11}} \cdot \textstyle\sum_{i=0}^3 q_i \cdot (\mu_i'' \cdot 3^{11}) \\ = \;& \textstyle\sum_{i=0}^3 \mu_i'' q_i,\end{aligned}$$

which is precisely as good as random guessing the unknown outcomes. Note that this again is independent of the p_i, and of any properties of the code used.

VI Goppa

17 Goppa codes

(17.1) Generalized Reed-Solomon codes. a) Let $n \in \mathbb{N}$ be arbitrary, and let $k \in \{0, \ldots, n\}$. We choose pairwise distinct places $\alpha := [\alpha_1, \ldots, \alpha_n]$ of

$\mathbb{F}_q$, and a vector $v := [v_1, \ldots, v_n] \in (\mathbb{F}_q^*)^n$. Then the associated **generalized Reed-Solomon code** is given as, essentially by evaluating polynomials,

$$\mathcal{G} = \mathrm{GRS}_k(\alpha, v) := \{[v_1 f(\alpha_1), \ldots, v_n f(\alpha_n)] \in \mathbb{F}_q^n; f \in \mathbb{F}_q[X]_{<k}\} \leq \mathbb{F}_q^n.$$

Example: Reed-Solomon codes. Let $n := q-1$ and $k := q-\delta$, where $\delta \in \{1, \ldots, q\}$. Moreover, let $\alpha = [\zeta_{q-1}^{j-1}]_j \in \mathbb{F}_q^{q-1}$ and $v := [\zeta_{q-1}^{(1-a)(j-1)}]_j \in \mathbb{F}_q^{q-1}$, for some $a \in \mathbb{Z}_{q-1}$. Then indeed the associated generalized Reed-Solomon code

$$\{[\zeta_{q-1}^{(1-a)(j-1)} \cdot f(\zeta_{q-1}^{j-1})]_j \in \mathbb{F}_q^{q-1}; f \in \mathbb{F}_q[X]_{<q-\delta}\} \leq \mathbb{F}_q^{q-1}$$

coincides with the Reed-Solomon code of designed distance δ and parameter a.

b) Hence $\mathcal{G}$ is generated by the rows of the **alternant matrix** with respect to the places v and the functions $\mathbb{F}_q \to \mathbb{F}_q \colon x \mapsto \alpha^{i-1}x$, that is $G = G_k(\alpha, v) := [v_j \alpha_j^{i-1}]_{ij} = [\alpha_j^{i-1}]_{ij} \cdot \mathrm{diag}[v_j]_j \in \mathbb{F}_q^{k \times n}$, or more explicitly

$$G = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \\ v_1\alpha_1 & v_2\alpha_2 & \ldots & v_n\alpha_n \\ \vdots & \vdots & & \vdots \\ v_1\alpha_1^{k-1} & v_n\alpha_2^{k-1} & \ldots & v_n\alpha_n^{k-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \alpha_n^{k-1} & \ldots & \alpha_n^{k-1} \end{bmatrix} \cdot \mathrm{diag}[v_j]_j.$$

In particular, we have $v \in \mathrm{GRS}_k(\alpha, v)$ whenever $k \geq 1$. Moreover, we observe that any k-subset of columns of the left hand factor forms a Vandermonde matrix. In particular, we conclude that G has full $\mathbb{F}_q$-rank, so that $\dim_{\mathbb{F}_q}(\mathcal{G}) = k$, implying that G is a generator matrix of $\mathcal{G}$ indeed. Since $G_k(\alpha, v)$ is a submatrix of $G_{k+1}(\alpha, v)$, for $k \leq n-1$, we conclude that

$$\{0\} = \mathrm{GRS}_0(\alpha, v) < \mathrm{GRS}_1(\alpha, v) < \cdots < \mathrm{GRS}_{n-1}(\alpha, v) < \mathrm{GRS}_n(\alpha, v) = \mathbb{F}_q^n.$$

For $k \geq 1$, since any $0 \neq f \in \mathbb{F}_q[X]_{<k}$ has at most $k-1$ zeroes in $\mathbb{F}_q$, we infer that $d(\mathcal{G}) \geq n-k+1$. Thus the Singleton bound yields $n-k \geq n - \dim_{\mathbb{F}_q}(\mathcal{G}) \geq d(\mathcal{G}) - 1 \geq n-k$, implying equality throughout, so that $\mathcal{G}$ is an MDS $[n, k, n-k+1]$-code. (This also follows from the observation that all k-subsets of columns of G form an invertible matrix, see Exercise (24.10).)

Theorem. We have $\mathrm{GRS}_k(\alpha, v)^\perp = \mathrm{GRS}_{n-k}(\alpha, w) \leq \mathbb{F}_q^n$, for some $w \in \mathbb{F}_q^n$, where w can be chosen simultaneously for all $k \in \{0, \ldots, n\}$.

Proof. We may assume that $k \in \{1, \ldots, n-1\}$.

i) We first consider the case $k = n-1$, where $\mathrm{GRS}_{n-1}(\alpha, v)^\perp = \langle w \rangle_{\mathbb{F}_q}$, for some $0 \neq w := [w_1, \ldots, w_n] \in \mathbb{F}_q^n$. Then we have

$$w \cdot G_{n-1}(\alpha, v)^{\mathrm{tr}} = [w_i v_i]_i \cdot [\alpha_i^{j-1}]_{ij} = 0_{n-1}.$$

Assume that $w_i = 0$, for some $i \in \{1, \dots, n\}$. Then leaving out the i-th row of the right hand matrix yields a Vandermonde $((n-1) \times (n-1))$-matrix, giving rise to a full rank set of linear equations for the remaining entries $w_j v_j$. This entails $w_j v_j = 0$, and thus $w_j = 0$, for $j \neq i$. Thus $w = 0$, a contradiction.

Hence $w_i \neq 0$, for all $i \in \{1, \dots, n\}$, showing $\mathrm{GRS}_{n-1}(\alpha, v)^\perp = \mathrm{GRS}_1(\alpha, w)$.

ii) Let now $k \in \{1, \dots, n-1\}$ be arbitrary. We show that $\mathrm{GRS}_k(\alpha, v)^\perp = \mathrm{GRS}_{n-k}(\alpha, w)$, by showing that $G_{n-k}(\alpha, w) \cdot G_k(\alpha, v)^{\mathrm{tr}} = 0 \in \mathbb{F}_q^{(n-k)\times k}$:

We determine the entry in row i and column j of the latter matrix, where $i \in \{1, \dots, n-k\}$ and $j \in \{1, \dots, k\}$: The latter equals $\sum_{k=1}^n w_k \alpha_k^{i-1} \cdot v_k \alpha_k^{j-1} = \sum_{k=1}^n w_k v_k \alpha_k^{i+j-2}$, which coincides with the entry of $w \cdot G_{n-1}(\alpha, v)^{\mathrm{tr}}$ in position $i + j - 1 \in \{1, \dots, n-1\}$, and hence by i) equals zero. ♯

Corollary. For $k \in \{1, \dots, n-1\}$ we have $\mathrm{GRS}_k(\alpha, v) = \mathrm{GRS}_k(\alpha, v')$, where $v, v' \in (\mathbb{F}_q^*)^n$, if and only if $\langle v \rangle_{\mathbb{F}_q} = \langle v' \rangle_{\mathbb{F}_q}$.

Proof. We only have to show necessity, for which we prove that $\langle v \rangle_{\mathbb{F}_q}$ is uniquely defined by $\mathcal{G} := \mathrm{GRS}_k(\alpha, v)$: Let $\mathcal{G}^\perp = \mathrm{GRS}_{n-k}(\alpha, w)$ be the associated dual code. Then we have $v \cdot G_{n-1}(\alpha, w)^{\mathrm{tr}} = 0_{n-1}$, that is $[v_i w_i]_i \in \ker([\alpha_i^{j-1}]_{ij})$, where $[\alpha_i^{j-1}]_{ij} \in \mathbb{F}_q^{n\times(n-1)}$ has $\mathbb{F}_q$-rank $n-1$. Hence $\langle [v_i w_i]_i \rangle_{\mathbb{F}_q}$ is uniquely defined by $\mathcal{G}$, since $w \in (\mathbb{F}_q^*)^n$ this also holds for $\langle v \rangle_{\mathbb{F}_q}$. ♯

(17.2) Alternant codes [Helgert, 1974]**.** Generalized Reed-Solomon codes give rise to a large class of codes as follows: Let $\mathbb{F}_q \subseteq \mathbb{F}$ be a finite field extension of degree $f := [\mathbb{F} \colon \mathbb{F}_q] \in \mathbb{N}$, let $n \in \mathbb{N}$, let $k \in \{0, \dots, n\}$, let $\alpha := [\alpha_1, \dots, \alpha_n]$ be pairwise distinct places of $\mathbb{F}$, and let $v := [v_1, \dots, v_n] \in (\mathbb{F}^*)^n$.

Then the associated **alternant code** over $\mathbb{F}_q$ is defined as the subfield subcode of $\mathrm{GRS}_k(\alpha, v)^\perp \leq \mathbb{F}^n$, that is $\mathcal{A} := \mathcal{A}_k(\alpha, v) := \mathrm{GRS}_k(\alpha, v)^\perp \cap \mathbb{F}_q^n \leq \mathbb{F}_q^n$.

Hence $G := G_k(\alpha, v) \in \mathbb{F}^{k\times n}$ is a generalized check matrix of $\mathcal{A}$, in the sense of (5.4). Since any k-subset of columns of G is $\mathbb{F}$-linearly independent, we get the **Helgert bound** $d(\mathcal{A}) \geq k+1$. Moreover, we have $n - fk \leq \dim_{\mathbb{F}_q}(\mathcal{A}) \leq n - k$.

Example. We consider the field extension $\mathbb{F}_2 \subseteq \mathbb{F}_8 = \langle \zeta \rangle$, where $\zeta = \zeta_7$ has minimum polynomial $\mu_\zeta := X^3 + X + 1 \in \mathbb{F}_2[X]$. Let $n := 7$ and $k := 2$, where $\alpha := [\zeta^{i-1}]_i$, and let $v := 1_7$ and $v' := \alpha$. Then we have

$$G = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 & \zeta^3 & \zeta^4 & \zeta^5 & \zeta^6 \end{bmatrix} \text{ and } G' = \begin{bmatrix} 1 & \zeta & \zeta^2 & \zeta^3 & \zeta^4 & \zeta^5 & \zeta^6 \\ 1 & \zeta^2 & \zeta^4 & \zeta^6 & \zeta & \zeta^3 & \zeta^5 \end{bmatrix}.$$

Hence the associated alternant codes have minimum distance ≥ 3, are non-trivial, and have $\mathbb{F}_2$-dimension ≤ 5. More precisely:

i) In the second case, we observe that the second row of G' is the Frobenius image of the first row, hence is redundant. Thus blowing up α yields

$$H := \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{3\times 7},$$

the associated code being linearly equivalent to the Hamming $[7, 4, 3]$-code $\mathcal{H}_3$.

ii) In the first case, blowing up G, and ignoring two zero rows, yields $\left[\frac{1_7}{H}\right] \in \mathbb{F}_2^{4\times 7}$, now giving rise to the even-weight Hamming $[7, 3, 4]$-code $\mathcal{H}_3'$.

Example: BCH codes. Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q, n) = 1$, be a BCH code of designed distance $\delta \in \{1, \ldots, n+1\}$, and consecutive defining set $\mathcal{V} := \{\zeta_n^a, \ldots, \zeta_n^{a+\delta-2}\} \subseteq \mathbb{F}_q(\zeta_n) =: \mathbb{F}$ of length $\delta - 1$, where $a \in \mathbb{Z}_n$. Then $\mathcal{C}$ has generalized check matrix

$$H(\mathcal{V}) = [\zeta_n^{(a+i-1)(j-1)}]_{ij} = [\zeta_n^{(i-1)(j-1)}]_{ij} \cdot \operatorname{diag}[\zeta_n^{a(j-1)}]_j \in \mathbb{F}^{(\delta-1)\times n}.$$

Thus $\mathcal{C}$ coincides with the alternant code $\mathcal{A}_{\delta-1}(\alpha, v)$, where $\alpha = [\zeta_n^{j-1}]_j$ and $v = [\zeta_n^{a(j-1)}]_j$. Note that from $d(\mathcal{A}_{\delta-1}(\alpha, v)) \geq \delta$ we recover the BCH bound.

(17.3) Goppa codes [1970]. Generally speaking, alternant codes allow for kind of too much flexibility, inasmuch their parameters, the places α and the vector v, can be chosen independently. We consider an important subset of alternant codes, which arise from specifying relations between α and the vector v. To do so, we need some preparation first:

a) Let $\mathbb{F}$ be a field, let $\mathscr{O} := \mathbb{F}[X]$, and let $\mathscr{K} := \mathrm{Q}(\mathscr{O}) = \mathbb{F}(X)$ be the field of **rational functions** in the indeterminate X; recall that $\mathscr{O}$ is Euclidean, hence in particular is a principal ideal domain, and thus is a unique factorization domain.

Let $0 \neq g = \sum_{i=0}^k g_i X^i \in \mathscr{O}$, having degree $k := \deg(g) \in \mathbb{N}_0$. Then let $\mathscr{O}_{\langle g\rangle} := \{\frac{a}{b} \in \mathscr{K}; \gcd(ag, b) = 1\} \subseteq \mathscr{K}$ be the **localization** of $\mathscr{O}$ with respect to the multiplicatively closed set $\{b \in \mathscr{O} \setminus \{0\}; \gcd(b, g) = 1\} \subseteq \mathscr{O}$, having the ideal

$$g\mathscr{O}_{\langle g\rangle} = \left\{\frac{a}{b} \in \mathscr{O}_{\langle g\rangle}; \gcd(a, b) = 1, g \mid a\right\} \trianglelefteq \mathscr{O}_{\langle g\rangle} \subseteq \mathscr{K}.$$

Note that we do neither assume g to be non-constant nor to be irreducible. We have $g\mathscr{O}_{\langle g\rangle} = \mathscr{O}_{\langle g\rangle}$ if and only if g is constant. If g is irreducible, then $\mathscr{O}_{\langle g\rangle}$ is a discrete valuation ring in $\mathscr{K}$, thus is a local ring whose maximal ideal is $g\mathscr{O}_{\langle g\rangle}$.

b) Now let $\mathbb{F}_q \subseteq \mathbb{F}$ be a finite field extension, and let $\alpha := [\alpha_1, \ldots, \alpha_n]$ be pairwise distinct places of $\mathbb{F}$, where $n \in \mathbb{N}$, such that $g(\alpha_i) \neq 0$, for $i \in \{1, \ldots, n\}$.

Hence we have $\gcd(g, X - \alpha_i) = 1$, so that $\frac{1}{X-\alpha_i} \in \mathscr{O}_{\langle g\rangle}$. Let

$$\psi_\alpha \colon \mathbb{F}^n \to \mathscr{O}_{\langle g\rangle} \colon c = [c_1, \ldots, c_n] \mapsto \sum_{i=1}^n \frac{c_i}{X - \alpha_i}.$$

Thus ψ_α is $\mathbb{F}$-linear, its image consists of rational functions having poles of order ≤ 1 in the places α only, and ψ_α specifies the associated residues.

Then the **(classical) Goppa code** associated with the places α is defined as

$$\mathcal{G}(\alpha, g) := \psi_\alpha^{-1}(g\mathscr{O}_{\langle g\rangle}) \cap \mathbb{F}_q^n = \left\{c \in \mathbb{F}_q^n; g \mid \psi_\alpha(c) \in \mathscr{O}_{\langle g\rangle}\right\} \leq \mathbb{F}_q^n;$$

it is called **separable** if the **Goppa polynomial** g does not have multiple zeroes (in a splitting field); it is called **irreducible** if $g \in \mathscr{O}$ is irreducible. In particular, if g is constant, that is $k = 0$, then $\mathcal{G}(\alpha, g) = \mathbb{F}_q^n$. (Note that the roles of ψ_α and g are reminiscent of the roles played by the map ψ and the generator polynomial for cyclic codes.)

(17.4) Theorem. We keep the notation of (17.3).

i) For $k \geq n$ we have $\psi_\alpha^{-1}(g\mathscr{O}_{\langle g\rangle}) = \{0\}$, and hence $\mathcal{G}(\alpha, g) = \{0\}$.

ii) Let $v := [v_1, \ldots, v_n]$, where $v_i := \frac{1}{g(\alpha_i)} \in \mathbb{F}^*$. Then for $k \leq n$ we have $\psi_\alpha^{-1}(g\mathscr{O}_{\langle g\rangle}) = \mathrm{GRS}_k(\alpha, v)^\perp \leq \mathbb{F}^n$, and hence $\mathcal{G}(\alpha, g) = \mathcal{A}_k(\alpha, v) \leq \mathbb{F}_q^n$.

In particular have the **Goppa bound** $d(\mathcal{G}(\alpha, g)) \geq k + 1$.

Proof. We may assume that g is non-constant, that is $k \geq 1$. Then we have $g - g(\alpha_i) = (X - \alpha_i)f_i \in \mathscr{O}$, where $f_i \in \mathscr{O}$ is given as

$$f_i = \frac{g - g(\alpha_i)}{X - \alpha_i} = \sum_{r=1}^{k} g_r \cdot \frac{X^r - \alpha_i^r}{X - \alpha_i} = \sum_{r=0}^{k-1} \left(g_{r+1} \cdot \sum_{s=0}^{r} \alpha_i^{r-s} X^s\right) \in \mathscr{K},$$

which can be rewritten as $f_i = \sum_{s=0}^{k-1}(\sum_{r=s}^{k-1} g_{r+1}\alpha_i^{r-s}) \cdot X^s \in \mathscr{O}$. Moreover, we have $\deg(f_i) = k - 1$, and $\frac{1}{X-\alpha_i} = -v_i f_i \in \mathscr{O}_{\langle g\rangle}/g\mathscr{O}_{\langle g\rangle}$.

Let $c = [c_1, \ldots, c_n] \in \mathbb{F}^n$. Then $\psi_\alpha(c) \in g\mathscr{O}_{\langle g\rangle}$ if and only if $\sum_{i=1}^n c_i v_i f_i \in g\mathscr{O}_{\langle g\rangle}$. Since $f_i \in \mathscr{O}$ and $g\mathscr{O}_{\langle g\rangle} \cap \mathscr{O} = g\mathscr{O}$, this holds if and only if $\sum_{i=1}^n c_i v_i f_i \in g\mathscr{O}$, which since $\deg(f_i) = k - 1 < k = \deg(g)$ is equivalent to $\sum_{i=1}^n c_i v_i f_i = 0 \in \mathscr{O}$.

In turn, considering homogeneous components, the latter holds if and only if $\sum_{i=1}^n \sum_{r=s}^{k-1} g_{r+1}\alpha_i^{r-s} v_i c_i = 0$, for $s \in \{0, \ldots, k-1\}$, or equivalently

$$\sum_{i=1}^{n} \sum_{t=0}^{k-s-1} g_{s+t+1}\alpha_i^t v_i c_i = \sum_{t=0}^{s} \left(g_{k-s+t} \cdot \sum_{i=1}^{n} \alpha_i^t v_i c_i\right) = 0, \text{ for } s \in \{0, \ldots, k-1\}.$$

Running through increasing values of s, and recalling that $g_k \neq 0$, shows that this amounts to a uni-triangular system of equations for the entities $\sum_{i=1}^n \alpha_i^t v_i c_i$, for $t \in \{0, \ldots, k-1\}$. Thus we infer that the above is equivalent to having $\sum_{i=1}^n \alpha_i^t v_i c_i = 0$, for $t \in \{0, \ldots, k-1\}$, or equivalently $[\alpha_i^{t-1}]_{ti} \cdot \mathrm{diag}[v_i]_i \cdot c^{\mathrm{tr}} = 0_k^{\mathrm{tr}}$.

Since $G_n(\alpha, v) \in \mathrm{GL}_n(\mathbb{F})$, for $k \geq n$ we get $c = 0$; for $k \leq n$ we get the conditions $c \cdot G_k(\alpha, v)^{\mathrm{tr}} = 0_k$ for the generalized Reed-Solomon code $\mathrm{GRS}_k(\alpha, v)^\perp$. ♯

Example: Narrow sense BCH codes. Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q,n)=1$, be a narrow sense BCH code of designed distance $\delta \in \{1,\dots,n+1\}$, and consecutive defining set $\mathcal{V} := \{\zeta_n,\dots,\zeta_n^{\delta-1}\} \subseteq \mathbb{F}_q(\zeta_n) =: \mathbb{F}$ of length $\delta-1$. Reversing the order of $\mathcal{V}$, we get the generalized check matrix $H = [\zeta_n^{(\delta-i)(j-1)}]_{ij} \in \mathbb{F}^{(\delta-1)\times n}$.

Now, letting $\alpha := [\zeta_n^{1-j}]_j \in \mathbb{F}^n$, and choosing the Goppa polynomial $g := X^{\delta-1}$, we get $v := [g(\zeta_n^{1-j})^{-1}]_j = [\zeta_n^{(\delta-1)(j-1)}]_j \in \mathbb{F}^n$, and hence indeed

$$G_{\delta-1}(\alpha,v) = [\zeta_n^{(i-1)(1-j)+(\delta-1)(j-1)}]_{ij} = [\zeta_n^{(\delta-i)(j-1)}]_{ij} = H.$$

Note that BCH codes are not always Goppa codes, see Exercise (28.5).

(17.5) Goppa codes are good. We proceed to show that the class of irreducible Goppa codes over $\mathbb{F}_q$ reaches the asymptotic Gilbert-Varshamov bound. We need a little preparation first:

For $k \in \mathbb{N}$ let $I_q(k) \subseteq \mathbb{F}_q[X]$ be the set of monic irreducible polynomials over $\mathbb{F}_q$ of degree k, and let $\iota_q(k) := |I_q(k)| \in \mathbb{N}_0$. Then for $j \in \mathbb{N}$ we have

$$q^j = \deg(X^{q^j} - X) = \sum_{k \mid j} \sum_{f \in I_q(k)} \deg(f) = \sum_{k \mid j} k \cdot \iota_q(k).$$

Thus **Möbius inversion** with respect to divisibility, using the number theoretic **Möbius function** $\mu\colon \mathbb{N} \to \{0,\pm 1\}$, yields $\iota_q(k) = \frac{1}{k} \cdot \sum_{j \mid k} \mu(\frac{k}{j}) \cdot q^j$. In particular, using $\mu(1)=1$, for $k \geq 2$ we get

$$\iota_q(k) \geq \frac{1}{k} \cdot (q^k - \sum_{j=1}^{\frac{k}{2}} q^j) = \frac{1}{k} \cdot (q^k - \frac{q^{\frac{k+2}{2}} - q}{q-1}) > \frac{1}{k} \cdot (q^k - q^{\frac{k+2}{2}}) \geq 0.$$

Theorem. Let $0 < \delta < \frac{q-1}{q}$. Then there is a sequence of irreducible Goppa codes $\mathcal{G}_r$ over $\mathbb{F}_q$ of length $n = q^r$, for $r \in \mathbb{N}$, such that $\limsup_{r\to\infty} \delta(\mathcal{G}_r) \geq \delta$ and $\limsup_{r\to\infty} \rho(\mathcal{G}_r) \geq 1 - H_q(\delta)$.

Proof. i) Let $d = d_r := \lceil n\delta \rceil \in \mathbb{N}$, that is d is minimal such that $\frac{d}{n} \geq \delta$. Hence from $n\delta \leq d \leq n\delta + 1$ we get $\lim_{r\mapsto\infty} \frac{d}{n} = \delta$.

Moreover, let $k = k_r \in \mathbb{N}$ be minimal such that

$$b = b_n := d \cdot |\mathcal{B}_{d-1}(0_n)| = d \cdot \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i \leq n^k - n^{\frac{k+2}{2}};$$

hence we have $k \geq 3$. We may assume that n and in consequence d are large enough such that $\sum_{i=0}^{d-1} \binom{n}{i} > n^3$, so that we have $k \geq 4$; in particular $n \geq 3$.

Then we have $n^{k-2} \leq n^{k-1} - n^{\frac{k+1}{2}} < b < n^k$. Taking logarithms yields

$$\frac{r(k-2)}{n} < \frac{\log_q(b)}{n} = \frac{\log_q(d)}{n} + \frac{\log_q(|\mathcal{B}_{d-1}(0_n)|)}{n} < \frac{rk}{n}.$$

From $\lim_{r\mapsto\infty}\frac{\log_q(d)}{n}=\lim_{r\mapsto\infty}(\frac{\log_q(d)}{d}\cdot\frac{d}{n})=0$, and $\lim_{r\mapsto\infty}\frac{r}{n}=\lim_{r\mapsto\infty}\frac{r}{q^r}=0$, by (8.3) we get

$$\lim_{r\mapsto\infty}\frac{rk}{n}=\lim_{r\mapsto\infty}\frac{\log_q(|\mathcal{B}_{d-1}(0_n)|)}{n}=H_q(\delta)<1.$$

Hence for $r\gg 0$ we have $4\leq k<n=q^r$.

ii) We show that there is an irreducible Goppa $[n,k_r,d_r]$-code $\mathcal{G}_r$ over $\mathbb{F}_q$, such that $k_r\geq n-rk$ and $d_r\geq d$. Then we have $\limsup_{r\to\infty}\rho(\mathcal{G}_r)\geq\lim_{r\to\infty}\frac{n-rk}{n}=1-H_q(\delta)$ and $\limsup_{r\to\infty}\delta(\mathcal{G}_r)\geq\lim_{r\to\infty}\frac{d}{n}=\delta$, as desired.

To this end, let $\mathbb{F}:=\mathbb{F}_{q^r}=\{\alpha_a,\ldots,\alpha_n\}$, and let $\alpha=[\alpha_1,\ldots,\alpha_n]$. Hence we have to show that there is $g\in I_n(k)$ such that $d(\mathcal{G}(\alpha,g))\geq d$:

Let $h\in I_n(k)$ such that there is $0\neq c=[c_1,\ldots,c_n]\in\mathcal{G}(\alpha,h)$, where $J:=\operatorname{supp}(c)$ such that $j:=|J|<d$. Then we have $\psi_\alpha(c)=\sum_{i\in J}\frac{c_i}{X-\alpha_i}=\frac{a}{b}\in\mathscr{K}$, where $a,b\in\mathbb{F}[X]$ such that $\gcd(a,b)=1$ and $h\mid a$. Since we have $\psi_\alpha(c)\cdot\prod_{i\in J}(X-\alpha_i)\in\mathbb{F}[X]$, we conclude that $\deg(a)\leq j-1$. Thus there are at most $\lfloor\frac{j-1}{k}\rfloor$ such polynomials h such that $c\in\mathcal{G}(\alpha,h)$.

Hence, running through all possible choices of c, we have to exclude at most

$$\sum_{j=1}^{d-1}\lfloor\frac{j-1}{k}\rfloor\cdot\binom{n}{j}(q-1)^j\leq\frac{d}{k}\cdot\sum_{j=1}^{d-1}\binom{n}{j}(q-1)^j\leq\frac{b}{k}\leq\frac{1}{k}(n^k-n^{\frac{k+2}{2}})$$

polynomials h as above from $I_n(k)$. Since we have $\iota_n(k)>\frac{1}{k}(n^k-n^{\frac{k+2}{2}})$, we conclude that there is a polynomial $g\in I_n(k)$ as desired. ♯

18 Background: Varieties

(18.1) Affine and projective varieties. a) Let $\overline{\mathbb{F}}$ be an algebraically closed field. For $n\in\mathbb{N}_0$ let $\mathbf{A}^n=\mathbf{A}^n(\overline{\mathbb{F}}):=\overline{\mathbb{F}}^n$ be the n-dimensional **affine space** over $\overline{\mathbb{F}}$, its elements $x=[x_1,\ldots,x_n]\in\mathbf{A}^n$ are called **points**; in particular, $\mathbf{A}^1$ is called the **affine line**, and $\mathbf{A}^2$ is called the **affine plane**.

Let $\mathcal{X}:=\{X_1,\ldots,X_n\}$ be indeterminates over $\overline{\mathbb{F}}$, let $\mathscr{O}:=\overline{\mathbb{F}}[\mathcal{X}]$, and let $\mathscr{K}=\mathscr{O}_{\{0\}}=\mathrm{Q}(\mathscr{O})=\overline{\mathbb{F}}(\mathcal{X})$. Letting $\mathfrak{p}\lhd\mathscr{O}$ be a prime ideal, the set of zeroes

$$\mathbf{V}=\mathbf{V}(\mathfrak{p}):=\{x\in\mathbf{A}^n;f(x)=0\text{ for }f\in\mathfrak{p}\}\subseteq\mathbf{A}^n$$

is called an **(irreducible) affine variety**, and $\mathscr{R}_\mathbf{V}:=\mathscr{O}/\mathfrak{p}$ is called the associated **(affine) coordinate ring**.

Then $\mathscr{O}_\mathfrak{p}:=\{\frac{f}{g}\in\mathscr{K};g\notin\mathfrak{p}\}$ is a local ring having maximal ideal $\mathfrak{p}\mathscr{O}_\mathfrak{p}=\{\frac{f}{g}\in\mathscr{K};g\notin\mathfrak{p},f\in\mathfrak{p}\}\lhd\mathscr{O}_\mathfrak{p}$, and the residue field $\mathscr{K}_\mathbf{V}:=\mathscr{O}_\mathfrak{p}/\mathfrak{p}\mathscr{O}_\mathfrak{p}$ is called the **field of rational functions** on $\mathbf{V}$. Moreover, we have an embedding $\mathscr{R}_\mathbf{V}\subseteq\mathscr{K}_\mathbf{V}$, such that $(\mathscr{R}_\mathbf{V})_{\{0\}}=\mathrm{Q}(\mathscr{R}_\mathbf{V})=\mathscr{K}_\mathbf{V}$.

b) Let $\mathbf{P}^n = \mathbf{P}^n(\overline{\mathbb{F}})$ be the n-dimensional **projective space** over $\overline{\mathbb{F}}$, whose **points** are given as equivalence classes $x = [x_0 \colon x_1 \colon \ldots \colon x_n] \in (\mathbf{A}^{n+1} \setminus \{0\})/_\sim$, where $\sim$ is given by letting $x \sim y$ if $y = ax$ for some $a \in \overline{\mathbb{F}}^*$; in particular, $\mathbf{P}^1$ is called the **projective line**, and $\mathbf{P}^2$ is called the **projective plane**.

Let $\widetilde{\mathcal{X}} := \{X_0, X_1, \ldots, X_n\}$ be indeterminates over $\overline{\mathbb{F}}$, let $\mathscr{O} := \overline{\mathbb{F}}[\widetilde{\mathcal{X}}]$, and let $\mathfrak{p}_0 := \langle X_0, X_1, \ldots, X_n\rangle \lhd \mathscr{O}$ be the **irrelevant ideal**, which is maximal and homogeneous with respect to the natural grading. Letting $\mathfrak{p}_0 \neq \mathfrak{p} \lhd \mathscr{O}$ be a homogeneous prime ideal, the set of zeroes

$$\mathbf{V} = \mathbf{V}(\mathfrak{p}) := \{x \in \mathbf{P}^n; f(x) = 0 \text{ for } f \in \mathfrak{p} \text{ homogeneous}\} \subseteq \mathbf{P}^n$$

is called an **(irreducible) projective variety**, and $\mathscr{R}_{\mathbf{V}} := \mathscr{O}/\mathfrak{p}$ is called the associated **(homogeneous) coordinate ring**.

Let $\mathrm{Q}'(\mathscr{O}) \subseteq \mathrm{Q}(\mathscr{O})$ be the localization of $\mathrm{Q}(\mathscr{O})$ with respect to the multiplicative set consisting of the non-zero homogeneous polynomials. Then $\mathrm{Q}'(\mathscr{O})$ has a grading given by $\deg(\frac{f}{g}) := \deg(f) - \deg(g)$, where $f, g \in \mathscr{O}$ are homogeneous. Let $\mathscr{K} := \mathscr{O}_{(\{0\})} \subseteq \mathrm{Q}'(\mathscr{O})$ be the **homogeneous field of fractions** of $\mathscr{O}$, thus

$$\mathscr{K} = \big\{\frac{f}{g} \in \mathrm{Q}(\mathscr{O}); f, g \text{ homogeneous}, \deg(\frac{f}{g}) = 0\big\} = \overline{\mathbb{F}}(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}).$$

Then $\mathscr{O}_{(\mathfrak{p})} := \{\frac{f}{g} \in \mathscr{K}; g \notin \mathfrak{p}\}$ is a local ring having maximal ideal

$$\mathfrak{p}\mathscr{O}_{\mathfrak{p}} \cap \mathscr{O}_{(\mathfrak{p})} = \{\frac{f}{g} \in \mathscr{K}; g \notin \mathfrak{p}, f \in \mathfrak{p}\} \lhd \mathscr{O}_{(\mathfrak{p})},$$

and $\mathscr{K}_{\mathbf{V}} := \mathscr{O}_{(\mathfrak{p})}/(\mathfrak{p}\mathscr{O}_{\mathfrak{p}} \cap \mathscr{O}_{(\mathfrak{p})})$ is called the **field of rational functions** on $\mathbf{V}$. We still have $(\mathscr{R}_{\mathbf{V}})_{(\{0\})} = \mathscr{K}_{\mathbf{V}}$, but we do not have a map from $\mathscr{R}_{\mathbf{V}}$ to $\mathscr{K}_{\mathbf{V}}$.

Example. i) The full affine space $\mathbf{A}^n$ is an affine variety, having regular functions $\mathscr{O}_{\mathbf{A}^n} = \mathscr{O}$ and function field $\mathscr{K}_{\mathbf{A}^n} = \mathrm{Q}(\mathscr{O}) = \mathscr{K}$, hence $\dim(\mathbf{A}^n) = n$.

Any singleton set $\{x\} \subseteq \mathbf{A}^n$ is an affine variety, having regular functions $\mathscr{O}_{\{x\}} = \mathscr{O}/\mathfrak{p}_x \cong \overline{\mathbb{F}}$, where $\mathfrak{p}_x = \langle X_1 - x_1, \ldots, X_n - x_n\rangle \lhd \mathscr{O}$ is maximal, and function field $\mathscr{K}_{\{x\}} = \mathrm{Q}(\overline{\mathbb{F}}) = \overline{\mathbb{F}}$, hence $\dim(\{x\}) = 0$.

ii) The full projective space $\mathbf{P}^n$ is an projective variety, having function field $\mathscr{K}_{\mathbf{P}^n} = \mathscr{K}$, the graded field of fractions, hence we have $\dim(\mathbf{P}^n) = n$.

Any singleton set $\{x\} \subseteq \mathbf{P}^n$ is a projective variety, with associated maximal ideal $\mathfrak{p}_x = \langle x_0X_1 - x_1X_0, \ldots, x_0X_n - x_nX_0\rangle \lhd \mathscr{O}$, having function field $\mathscr{K}_{\{x\}} = \mathscr{O}_{(\mathfrak{p}_x)}/(\mathfrak{p}_x\mathscr{O}_{\mathfrak{p}_x} \cap \mathscr{O}_{(\mathfrak{p}_x)}) \cong \mathscr{O}/\mathfrak{p}_x \cong \overline{\mathbb{F}}$, hence we have $\dim(\{x\}) = 0$.

(18.2) Regular functions. Let $\mathbf{V}$ be a variety. Then, since $\overline{\mathbb{F}}$ is algebraically closed, the field $\mathscr{K}_{\mathbf{V}}$ of rational functions on $\mathbf{V}$ is **separably generated**, that is it contains a finitely generated rational function field K such that $K \subseteq \mathscr{K}_{\mathbf{V}}$ is a finite separable extension.

The transcendence degree of $\mathscr{K}_{\mathbf{V}}$ is called the **(Krull) dimension** $\dim(\mathbf{V}) := \operatorname{trdeg}(\mathscr{K}_{\mathbf{V}}) \in \mathbb{N}_0$ of $\mathbf{V}$; if $\dim(\mathbf{V}) = 1$ then $\mathbf{V}$ is called a **curve**.

Given $x \in \mathbf{V}$, and $\varphi := \frac{f}{g} \in \mathscr{K}_{\mathbf{V}}$ such that $g(x) \neq 0$, then φ is called **regular** at the point x. The set $\mathscr{O}_x \subseteq \mathscr{K}_{\mathbf{V}}$ of all regular functions at x is called the **local ring** at x, having maximal ideal $\mathfrak{m}_x := \{\varphi \in \mathscr{O}_x; \varphi(x) = 0\} \lhd \mathscr{O}_x$; hence we have the residue field $\overline{\mathbb{F}}_x := \mathscr{O}_x/\mathfrak{m}_x \cong \overline{\mathbb{F}}$. Moreover, we have **i)** $\mathscr{O}_x = (\mathscr{R}_{\mathbf{V}})_{\mathfrak{p}_x}$ in the affine case, and **ii)** $\mathscr{O}_x = (\mathscr{R}_{\mathbf{V}})_{(\mathfrak{p}_x)}$ in the projective case.

If $U \subseteq \mathbf{V}$ is an open subset (with respect to the Zariski topology) such that φ is regular in all points of U, then φ is called a **regular function** on U; recall that U is dense whenever $U \neq \emptyset$. Let $\mathscr{O}_U \subseteq \mathscr{K}_{\mathbf{V}}$ be the ring of all regular functions on U; in particular, we have $\mathscr{O}_\emptyset = \mathscr{K}_{\mathbf{V}}$. To describe the regular functions $\mathscr{O}_{\mathbf{V}}$ on all of $\mathbf{V}$, we have to distinguish the affine and projective cases:

i) Let $\mathbf{V}$ be affine. Since $\mathscr{O}$ induces polynomial functions on $\mathbf{V}$, we have $\mathscr{O}/\mathfrak{p} \subseteq \mathscr{O}_{\mathbf{V}}$, hence it follows from **Hilbert's Nullstellensatz** that $\mathscr{O}_{\mathbf{V}} = \mathscr{O}/\mathfrak{p} = \mathscr{R}_{\mathbf{V}}$.

ii) Let $\mathbf{V}$ be projective. Then any morphism from $\mathbf{V}$ to any (quasi-projective) variety is closed. (This is reminiscent of the topological notion of compactness.) In particular, any regular function $\mathbf{V} \to \mathbf{A}^1$ is constant, entailing $\mathscr{O}_{\mathbf{V}} = \overline{\mathbb{F}}$.

(18.3) Affine open subsets. We discuss the relation between affine and projective varieties: Let $\mathbf{V} = \mathbf{V}(\mathfrak{p}) \subseteq \mathbf{A}^n$ be an affine variety.

For $0 \neq f \in \mathscr{O} := \overline{\mathbb{F}}[\mathcal{X}]$ let $\widetilde{f} := X_0^{\deg(f)} \cdot f(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}) \in \widetilde{\mathscr{O}} := \overline{\mathbb{F}}[\widetilde{\mathcal{X}}]$ be its **homogenization**; then $\deg(\widetilde{f}) = \deg(f)$, and for completeness we let $\widetilde{0} := 0$.

Let $D_{X_0} := \mathbf{P}^n \setminus \mathbf{V}(X_0)$, which is open, hence is dense in $\mathbf{P}^n$, and let

$$\varphi_0 \colon \mathbf{A}^n \to D_{X_0} \colon [x_1, \ldots, x_n] \mapsto [1 \colon x_1 \colon \ldots \colon x_n].$$

Then φ_0 is a homeomorphism with inverse

$$\varphi_0^{-1} \colon D_{X_0} \to \mathbf{A}^n \colon [x_0 \colon x_1 \colon \ldots \colon x_n] \mapsto [\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}].$$

Let $\widetilde{\mathfrak{p}} := \langle \widetilde{f} \in \widetilde{\mathscr{O}}; f \in \mathfrak{p} \rangle \lhd \widetilde{\mathscr{O}}$, which is a prime ideal, and let $\widetilde{\mathbf{V}} := \mathbf{V}(\widetilde{\mathfrak{p}}) \subseteq \mathbf{P}^n$. Then we get $\varphi_0(\mathbf{V}) = \widetilde{\mathbf{V}} \cap D_{X_0}$, which is open, hence is dense in $\widetilde{\mathbf{V}}$; thus $\overline{\varphi_0(\mathbf{V})} = \widetilde{\mathbf{V}}$. The set $\widetilde{\mathbf{V}} \cap \mathbf{V}(X_0) = \widetilde{\mathbf{V}} \setminus \varphi_0(\mathbf{V})$ are called **points at infinity**.

For any $U \subseteq \mathbf{V}$ open the associated **comorphism**

$$\varphi_0^* \colon \mathscr{O}_{\widetilde{\mathbf{V}},U} \to \mathscr{O}_{\mathbf{V},U} \colon \frac{f(\widetilde{\mathcal{X}})}{g(\widetilde{\mathcal{X}})} \mapsto \frac{f(1,\mathcal{X})}{g(1,\mathcal{X})}$$

is a ring isomorphism, with inverse

$$(\varphi_0^{-1})^* \colon \mathscr{O}_{\mathbf{V},U} \to \mathscr{O}_{\widetilde{\mathbf{V}},U} \colon \frac{f(\mathcal{X})}{g(\mathcal{X})} \mapsto X_0^{\deg g - \deg(f)} \cdot \frac{\widetilde{f}(\widetilde{\mathcal{X}})}{\widetilde{g}(\widetilde{\mathcal{X}})}.$$

Thus φ_0 is an **open immersion**, so that $\mathbf{V}$ can be identified with the **affine open subset** $\varphi_0(\mathbf{V}) \subseteq \widetilde{\mathbf{V}}$. In particular, we have $\mathbf{A}^n \cong \varphi_0(\mathbf{A}^n) \subseteq \mathbf{P}^n$; for $n = 1$ we may write $\mathbf{P}^1 = \mathbf{A}^1 \dot\cup \{\infty\}$, where the latter point is $[0\colon 1] \in \mathbf{P}^1$.

(18.4) Regular points. a) Let $\mathbf{V}$ be a variety, and let $x \in \mathbf{V}$ be a point. Then the **cotangent space** $T_x^* = T_{\mathbf{V},x}^* := \mathfrak{m}_x/\mathfrak{m}_x^2$ of $\mathbf{V}$ at x is a finitely generated vector space over $\overline{\mathbb{F}}_x = \mathscr{O}_x/\mathfrak{m}_x \cong \overline{\mathbb{F}}$, such that $\dim_{\overline{\mathbb{F}}}(T_x^*) \geq \dim(\mathbf{V})$. The **(Zariski) tangent space** is defined as $T_x = T_{\mathbf{V},x} := \mathrm{Hom}_{\overline{\mathbb{F}}}(\mathfrak{m}_x/\mathfrak{m}_x^2, \overline{\mathbb{F}})$.

If $\dim_{\overline{\mathbb{F}}}(T_x^*) = \dim(\mathbf{V})$ holds, that is $\mathscr{O}_x$ is a **regular local ring**, then x is called a **regular** point of $\mathbf{V}$, otherwise x is called **singular**.

The set of regular points is a non-empty open, hence is a dense subset of $\mathbf{V}$. The variety $\mathbf{V}$ is called **smooth** or **non-singular** if all its points are regular. For example, the affine space $\mathbf{A}^n$ and the projective space $\mathbf{P}^n$ are smooth.

b) The intrinsic property of regularity can be checked as follows, as soon as an explicit embedding into some affine or projective space is given:

Let $\mathbf{V} = \mathbf{V}(\mathfrak{p})$ be a variety, and let $\mathfrak{p} = \langle F_1, \ldots, F_k\rangle \lhd \mathscr{O}$, for some $k \in \mathbb{N}$, where in the projective case the F_i are homogeneous. Abbreviating $\partial_j := \partial_{X_j}$ for $j \in \{0, \ldots, n\}$, the associated **Jacobian matrix** is **i)** in the affine case

$$J = J(F_1, \ldots, F_k) = J_{\mathcal{X}}(F_1, \ldots, F_k) := [\partial_j F_i]_{ij} \in \mathscr{O}^{k\times n},$$

and similarly **ii)** in the projective case $J := J_{\widetilde{\mathcal{X}}}(F_1, \ldots, F_k) \in \mathscr{O}^{k\times(n+1)}$.

Proposition: Zariski [1947]. We have $\dim_{\overline{\mathbb{F}}}(T_x^*) + \mathrm{rk}_{\overline{\mathbb{F}}}(J(x)) = n$. ♯

Thus $\mathrm{rk}_{\overline{\mathbb{F}}}(J(x)) \leq n - \dim(\mathbf{V})$, and $x \in \mathbf{V}$ is regular if and only if equality holds.

c) If $\mathbf{V} \subseteq \mathbf{A}^n$ is affine, this can be geometrically interpreted as follows: Let $x = [x_1, \ldots, x_n] \in \mathbf{V}$, and for $f \in \mathscr{O}_x$ let

$$D_x f := \sum_{j=1}^{n} (\partial_j f)(x_1, \ldots, x_n) \cdot (X_j - x_j) \in \mathfrak{m}_x \lhd \mathscr{O}_x.$$

Then the geometric tangent space of $\mathbf{V} \subseteq \mathbf{A}^n$ at x is the affine subspace defined by $\langle D_x(F_1), \ldots, D_x(F_k)\rangle \lhd \mathscr{O}$, thus we have $T_{\mathbf{V}\subseteq\mathbf{A}^n,x} = x + \ker(J^{\mathrm{tr}}(x)) \subseteq \mathbf{A}^n$, indeed being an affine space of $\overline{\mathbb{F}}$-dimension $n - \mathrm{rk}(J(x))$.

Now these locally defined spaces are 'bundled' globally as follows:

(18.5) Differential forms. a) Let $\mathbf{V}$ be an affine variety, and let $\mu_{\mathbf{V}}\colon \mathscr{O}_{\mathbf{V}} \otimes_{\overline{\mathbb{F}}} \mathscr{O}_{\mathbf{V}} \to \mathscr{O}_{\mathbf{V}}\colon f \otimes g \mapsto fg$, which is an $\mathscr{O}_{\mathbf{V}}$-module homomorphism. Then $\Omega_{\mathbf{V}} = \Omega_{\mathscr{O}_{\mathbf{V}}} := \ker(\mu_{\mathbf{V}})/\ker(\mu_{\mathbf{V}})^2$ is called the $\mathscr{O}_{\mathbf{V}}$-module of **Kähler differentials**.

The $\overline{\mathbb{F}}$-linear **total differential** $d\colon \mathscr{O}_{\mathbf{V}} \to \Omega_{\mathbf{V}}\colon f \mapsto f\otimes 1 - 1\otimes f$ is a **derivation**, that is fulfills the **product rule** $d(fg) = df \cdot g + dg \cdot f$; thus the elements of $\Omega_{\mathbf{V}}$ are also called **regular differential forms**.

For the full affine space $\mathbf{A}^n$ it turns out that $\Omega_{\mathbf{A}^n} = \Omega_{\mathscr{O}}$ is the free $\mathscr{O}$-module generated by $\{dX_1, \ldots, dX_n\}$, and we get $df = \sum_{j=1}^n \partial_j f \cdot dX_j$.

In general, if $\mathbf{V} = \mathbf{V}(\mathfrak{p})$, where $\mathfrak{p} = \langle F_1, \ldots, F_k\rangle \lhd \mathscr{O}$, for some $k \in \mathbb{N}$, then $\Omega_{\mathbf{V}}$ can be described as follows: Note that $\Omega_{\mathscr{O}} \otimes_{\mathscr{O}} \mathscr{O}_{\mathbf{V}} \cong \bigoplus_{j=1}^n \mathscr{O}_V \cdot dX_n$ is the free $\mathscr{O}_{\mathbf{V}}$-module generated by $\{dX_1, \ldots, dX_n\}$. Then we have

$$\Omega_{\mathbf{V}} \cong \big(\bigoplus_{j=1}^n \mathscr{O}_V \cdot dX_n\big)/\langle dF; F \in \mathfrak{p}\rangle_{\mathscr{O}_{\mathbf{V}}} = \big(\bigoplus_{j=1}^n \mathscr{O}_V \cdot dX_n\big)/\langle dF_1, \ldots, dF_k\rangle_{\mathscr{O}_{\mathbf{V}}}.$$

The derivation d extends to $\mathscr{K}_{\mathbf{V}} = \mathrm{Q}(\mathscr{O}_{\mathbf{V}})$, yielding the $\mathscr{K}_{\mathbf{V}}$-module $\Omega_{\mathscr{K}_{\mathbf{V}}} := \Omega_{\mathbf{V}} \otimes_{\mathscr{O}_{\mathbf{V}}} \mathscr{K}_{\mathbf{V}}$; the elements of $\Omega_{\mathscr{K}_{\mathbf{V}}}$ are called **rational differential forms**. Then we have $\dim_{\mathscr{K}_{\mathbf{V}}}(\Omega_{\mathscr{K}_{\mathbf{V}}}) = d := \dim(\mathbf{V})$.

If $x \in \mathbf{V}$ is regular, then there is an affine open neighborhood $U \subseteq \mathbf{V}$ such that Ω_U is a free $\mathscr{O}_U$-module of rank d; hence $\Omega_{\mathbf{V}}$ is said to be **locally free**. Moreover, if $\mathbf{V}$ is smooth, then $\Omega_{\mathbf{V}}$ is a (globally) free $\mathscr{O}_{\mathbf{V}}$-module of rank d.

b) We consider the local behavior at a regular point $x = [x_1, \ldots, x_n] \in \mathbf{V}$: In this case $\mathscr{O}_x$ is a regular local ring, where $\mathfrak{m}_x = \mathfrak{p}_x\mathscr{O}_x \lhd \mathscr{O}_x$ and $\overline{\mathbb{F}}_x := \mathscr{O}_x/\mathfrak{m}_x$, such that $T_x^* := \mathfrak{m}_x/\mathfrak{m}_x^2$ is an $\overline{\mathbb{F}}$-vector space of dimension d.

It turns out that $\Omega_x := (\Omega_{\mathbf{V}})_{\mathfrak{p}_x} = \Omega_{\mathbf{V}} \otimes_{\mathscr{O}} \mathscr{O}_x$ is a free $\mathscr{O}_x$-module of rank d, and thus $\Omega_{\overline{\mathbb{F}}_x} := \Omega_x \otimes_{\mathscr{O}_x} \overline{\mathbb{F}}_x$ is an $\overline{\mathbb{F}}$-vector space of dimension d. Let $d_x \colon \mathscr{O}_x \to \Omega_x$ and $d_{\overline{\mathbb{F}}_x} \colon \mathscr{O}_x \to \overline{\mathbb{F}}_x \to \Omega_{\overline{\mathbb{F}}_x}$ be the associated derivations. Using the $\overline{\mathbb{F}}$-isomorphism $\Omega_{\overline{\mathbb{F}}_x} \to T_x^* \colon d_{\overline{\mathbb{F}}_x}(X_j) \mapsto X_j - x_j$, we may identify $d_{\overline{\mathbb{F}}_x} f$ with $D_x f$, for $f \in \mathcal{O}_x$.

c) Let $\mathbf{V}$ be a projective variety, and let $U \subseteq \mathbf{V}$ be open. We aim at defining differential forms on U by a process called **gluing**; to this end recall that the affine open subsets are a basis of the Zariski topology:

Note first, that if $W \subseteq V \subseteq \mathbf{V}$ are affine open subsets, then by restriction of functions we get a **restriction map** $\Omega_V \to \Omega_W \colon \omega \mapsto \omega|_W$. Now let $\{U_i\}_{i\in\mathcal{I}}$ be an affine open covering of U, for some (finite) index set $\mathcal{I}$, and let $\omega_i \in \Omega_{U_i}$ be **compatible** in the sense that, for $i, j \in \mathcal{I}$ and any affine open subset $V \subseteq U_i \cap U_j$, we have $\omega_i|_V = \omega_j|_V$. Then the set $\{\omega_i\}_{i\in\mathcal{I}}$ defines a unique **regular differential form** on U. Let Ω_U be the $\mathscr{O}_U$-module of all forms obtained by this process, together with a derivation $d_U \colon \mathscr{O}_U \to \Omega_U$. ♯

Similarly, a **rational differential form** on $\mathbf{V}$ is obtained by gluing regular differential forms on open (affine) subsets of $\mathbf{V}$ (not necessarily covering $\mathbf{V}$). This yields the $\mathscr{K}_{\mathbf{V}}$-module $\Omega_{\mathscr{K}_{\mathbf{V}}}$, together with the derivation $d_{\mathbf{V}} \colon \mathscr{K}_{\mathbf{V}} \to \Omega_{\mathscr{K}_{\mathbf{V}}}$. The elements of the image $d_{\mathbf{V}}(\mathscr{K}_{\mathbf{V}}) \subseteq \Omega_{\mathscr{K}_{\mathbf{V}}}$ are called **exact** differential forms.

Finally, for a regular point $x \in U$ we get the $\mathscr{O}_x$-module $\Omega_{U,x} = \Omega_{V,x}$, where $V \subseteq U \subseteq \mathbf{V}$ is an affine open neighborhood of x.

19 Background: Curves

(19.1) Curves. a) Let $\mathbf{V}$ be a smooth projective curve. Then any $x \in \mathbf{V}$ is regular, that is $\dim_{\overline{\mathbb{F}}}(T_x^*) = 1$. Hence $\mathscr{O}_x \subseteq \mathscr{K}_{\mathbf{V}}$ is a **discrete valuation ring**, and any element of $\mathfrak{m}_x \setminus \mathfrak{m}_x^2$ is called a **local (uniformising) coordinate**.

Let $\nu_x \colon \mathscr{K}_{\mathbf{V}}^* \to \mathbb{Z}$ be the associated **(discrete) valuation**, also called a **place** of $\mathscr{K}_{\mathbf{V}}$, which for $0 \neq f \in \mathscr{O}_x$ is given by $\nu_x(f) := k \geq 0$ if $f \in \mathfrak{m}_x^k \setminus \mathfrak{m}_x^{k+1}$, and for $f \in \mathscr{K}_{\mathbf{V}} \setminus \mathscr{O}_x$ is given by $\nu_x(f) := -\nu_x(\frac{1}{f}) < 0$.

If $0 \neq f \in \mathscr{K}_{\mathbf{V}}$ such that $\nu_x(f) > 0$, then f is said to have a **zero of order** $\nu_x(f)$ at x; if $\nu_x(f) < 0$, then f is said to have a **pole of order** $-\nu_x(f)$ at x.

Since the set of points at which f is regular and $f(x) \neq 0$ is non-empty and open, hence is dense in $\mathbf{V}$, and thus has a finite complement in $\mathbf{V}$, we infer that $\nu_x(f) = 0$ for almost all points of $\mathbf{V}$.

b) We now bring morphisms into play: Let $\mathbf{W}$ be any curve (not necessarily projective nor smooth), and let $\varphi \colon \mathbf{V} \to \mathbf{W}$ be a non-constant morphism.

Theorem. Then **i)** $\mathbf{W}$ is projective, **ii)** φ is **finite**, thus is surjective with finite fibers, and **iii)** φ^* induces a finite field extension $\mathscr{K}_{\mathbf{W}} \cong \varphi^*(\mathscr{K}_{\mathbf{W}}) \subseteq \mathscr{K}_{\mathbf{V}}$. ♯

The **degree** of φ is defined as $\deg(\varphi) := [\mathscr{K}_{\mathbf{V}} \colon \mathscr{K}_{\mathbf{W}}] \in \mathbb{N}$.

(19.2) Divisors. a) Let $\mathbf{V}$ be a smooth projective curve. The free abelian group $\mathrm{Div}_{\mathbf{V}} := \bigoplus_{x \in \mathbf{V}} \mathbb{Z} \cdot (x)$ is called the group of **(Weil) divisors** of $\mathbf{V}$. The divisor (x), where $x \in \mathbf{V}$, is called a **prime divisor**.

Let $D = \sum_{x \in \mathbf{V}} m_x \cdot (x) \in \mathrm{Div}_{\mathbf{V}}$. The finite set $\mathrm{supp}(D) := \{x \in \mathbf{V}; m_x \neq 0\} \subseteq \mathbf{V}$ is called the **support** of D. Moreover, D is called **effective**, if $m_x \geq 0$ for all $x \in \mathbf{V}$. This yields a partial order, by letting $D \geq D'$ if $D - D'$ is effective.

We have the **degree** homomorphism $\deg \colon \mathrm{Div}_{\mathbf{V}} \to \mathbb{Z}$ of abelian groups, given by $\deg(D) := \sum_{x \in \mathbf{V}} m_x \cdot [\overline{\mathbb{F}}_x \colon \overline{\mathbb{F}}]$. Recalling that, since $\overline{\mathbb{F}}$ is algebraically closed, we have $[\overline{\mathbb{F}}_x \colon \overline{\mathbb{F}}] = 1$, we just get $\deg(D) = \sum_{x \in \mathbf{V}} m_x$. In particular, for prime divisors we get $\deg((x)) = [\overline{\mathbb{F}}_x \colon \overline{\mathbb{F}}] = 1$, so that the degree map is surjective.

Let $\mathrm{Div}_{\mathbf{V}}^0 \leq \mathrm{Div}_{\mathbf{V}}$ be the kernel of the degree homomorphism, the subgroup of divisors of degree 0, and let the **Néron-Severi group** be $\mathrm{Div}_{\mathbf{V}}/\mathrm{Div}_{\mathbf{V}}^0 \cong \mathbb{Z}$

b) Let $\varphi \colon \mathbf{V} \to \mathbf{W}$ be a non-constant morphism, where $\mathbf{W}$ is a smooth projective curve. Then for $z \in \mathbf{W}$ let $t_z \in \mathscr{O}_x$ be a local coordinate, and let $\varphi^{-1}(z) = \{y_1, \ldots, y_r\} \subseteq \mathbf{V}$, for some $r \in \mathbb{N}$; then we have $\varphi^*(t_z) \in \mathfrak{m}_{y_i} \lhd \mathscr{O}_{y_i}$.

The number $e_{y_i}(\varphi) := \nu_{y_i}(\varphi^*(t_z)) \in \mathbb{N}$ is called the **ramification index** of φ at the point y_i. If $e_{y_i}(\varphi) \geq 2$, then $y_i \in \mathbf{V}$ is called a **ramification point**, and $z = \varphi(y_i) \in \mathbf{W}$ is called a **branch point**; in this case, if $\mathrm{char}(\overline{\mathbb{F}}) \nmid e_{y_i}(\varphi)$ then the ramification is called **tame**, otherwise it is called **wild**.

For the prime divisor $(z) \in \mathrm{Div}_{\mathbf{W}}$ we get the effective divisor

$$\varphi^*((z)) := \sum_{i=1}^{r} e_{y_i}(\varphi) \cdot (y_i) = \sum_{i=1}^{r} \nu_{y_i}(\varphi^*(t_z)) \cdot (y_i) \in \mathrm{Div}_{\mathbf{V}}.$$

Theorem: Degree formula. Recalling $\deg((z)) = 1$ and $[\overline{\mathbb{F}}_{y_i} \colon \overline{\mathbb{F}}] = 1$, we get

$$\deg(\varphi^*((z))) = \sum_{i=1}^{r} e_{y_i}(\varphi) = \deg(\varphi). \qquad \sharp$$

By $\mathbb{Z}$-linearity this gives rise to a group homomorphism $\varphi^* \colon \mathrm{Div}_{\mathbf{W}} \to \mathrm{Div}_{\mathbf{V}}$, mapping effective divisors to effective divisors, where by the degree formula we have $\deg(\varphi^*(D)) = \deg(\varphi) \cdot \deg(D)$, for $D \in \mathrm{Div}_{\mathbf{W}}$.

(19.3) Picard groups. a) Let $\mathbf{V}$ be a smooth projective curve. For $f \in \mathscr{K}_{\mathbf{V}}^*$ let the **principal divisor** be defined as follows; note that $(f) \in \mathrm{Div}_{\mathbf{V}}$ indeed:

$$(f) := \sum_{x \in \mathbf{V}} \nu_x(f) \cdot (x) \in \mathrm{Div}_{\mathbf{V}}.$$

Moreover, $(f)_0 := \sum_{x \in \mathbf{V}(f)} \nu_x(f) \cdot (x) \in \mathrm{Div}_{\mathbf{V}}$ and $(f)_\infty := (\frac{1}{f})_0 \in \mathrm{Div}_{\mathbf{V}}$ are called the **zero** and **pole divisor** of f, respectively, so that $(f) = (f)_0 - (f)_\infty$. For example, if $a \in \overline{\mathbb{F}}^*$ is constant, then $(a) = 0 \in \mathrm{Div}_{\mathbf{V}}$, thus $\deg((a)) = 0$.

In order to state the following theorem, note that any element of $\mathscr{K}_{\mathbf{V}} \setminus \overline{\mathbb{F}}$ can be considered as a non-constant morphism $\mathbf{V} \to \mathbf{P}^1$, and as such has a degree.

Theorem. Let $\varphi = \frac{f}{g} \in \mathscr{K}_{\mathbf{V}} \setminus \overline{\mathbb{F}}$, where $f, g \in \mathscr{R}_{\mathbf{V}}$ are homogeneous. Then we have $\deg(\varphi) = \deg(f) = \deg(g)$ and $\deg((\varphi)) = 0$.

Proof. We consider the morphism $\varphi \colon \mathbf{V} \to \mathbf{P}^1 = \mathbf{A}^1 \,\dot\cup\, \{\infty\} \colon x \mapsto [g(x) \colon f(x)]$. We have $t_0 = t_{[1\colon 0]} = X$ and $t_\infty = t_{[0\colon 1]} = Y$, where $\varphi^*(t_0) = \varphi^*(X) = f$ and $\varphi^*(t_\infty) = \varphi^*(Y) = g$. Hence we get

$$\deg(\varphi^*((0))) = \deg\big(\sum_{x \in \varphi^{-1}(0)} \nu_x(f) \cdot (x)\big) = \sum_{x \in \mathbf{V}, f(x) = 0} \nu_x(f) = \deg(f),$$

and similarly

$$\deg(\varphi^*((\infty))) = \deg\big(\sum_{x \in \varphi^{-1}(\infty)} \nu_x(g) \cdot (x)\big) = \sum_{x \in \mathbf{V}, g(x) = 0} \nu_x(g) = \deg(g).$$

Thus the degree formula yields $\deg(\varphi) = \deg(f) = \deg(g)$. Moreover, since $(\varphi) = \varphi^*((0)) - \varphi^*((\infty)) \in \mathrm{Div}_{\mathbf{V}}$, this implies $\deg((\varphi)) = 0$. $\sharp$

This yields an injective group homomorphism $\mathscr{K}_{\mathbf{V}}^* \to \mathrm{Div}_{\mathbf{V}}^0 \leq \mathrm{Div}_{\mathbf{V}}$, whose image $(\mathscr{K}_{\mathbf{V}}^*) := \{(f) \in \mathrm{Div}_{\mathbf{V}}; f \in \mathscr{K}_{\mathbf{V}}^*\}$ is the subgroup of principal divisors.

b) The quotient group $\mathrm{Pic}_{\mathbf{V}} := \mathrm{Div}_{\mathbf{V}}/(\mathscr{K}_{\mathbf{V}}^*)$ is called the **divisor class group** or **Picard group** of $\mathbf{V}$; its elements are written as $[D]$. Divisors $D, D' \in \mathrm{Div}_{\mathbf{V}}$ are called **(linearly) equivalent** if $[D] = [D'] \in \mathrm{Pic}_{\mathbf{V}}$; we also write $D \sim D'$.

Since $(\mathscr{K}_{\mathbf{V}}^*) \leq \mathrm{Div}_{\mathbf{V}}^0$, the Picard group inherits the (surjective) degree homomorphism, and we let $\mathrm{Pic}_{\mathbf{V}}^0 := \mathrm{Div}_{\mathbf{V}}^0/(\mathscr{K}_{\mathbf{V}}^*)$ be its kernel, also called the **pure Picard group**. Hence the Néron-Severi group is $\mathrm{Pic}_{\mathbf{V}}/\mathrm{Pic}_{\mathbf{V}}^0 \cong \mathrm{Div}_{\mathbf{V}}/\mathrm{Div}_{\mathbf{V}}^0 \cong \mathbb{Z}$.

It turns out that $\mathrm{Pic}_{\mathbf{V}}^0$ carries the structure of an **abelian variety**, that is a projective variety carrying the structure of a (necessarily abelian) algebraic group, being called the **Picard variety** or **Jacobian variety** of $\mathbf{V}$.

c) If $\varphi\colon \mathbf{V} \to \mathbf{W}$ is a non-constant morphism of degree $d := \deg(\varphi) \in \mathbb{N}_0$, where $\mathbf{W}$ is a smooth projective curve, then we have $\varphi^*\colon (\mathscr{K}_{\mathbf{W}}^*) \to (\mathscr{K}_{\mathbf{V}}^*)$ and $\varphi^*\colon \mathrm{Div}_{\mathbf{W}}^0 \to \mathrm{Div}_{\mathbf{V}}^0$, yielding group homomorphisms $\varphi^*\colon \mathrm{Pic}_{\mathbf{W}} \to \mathrm{Pic}_{\mathbf{V}}$ and $\varphi^*\colon \mathrm{Pic}_{\mathbf{W}}^0 \to \mathrm{Pic}_{\mathbf{V}}^0$, by the degree formula entailing the group homomorphism

$$\varphi^*\colon \mathbb{Z} \cong \mathrm{Pic}_{\mathbf{W}}/\mathrm{Pic}_{\mathbf{W}}^0 \to \mathrm{Pic}_{\mathbf{V}}/\mathrm{Pic}_{\mathbf{V}}^0 \cong \mathbb{Z}\colon 1 \mapsto d.$$

(19.4) Theorem: Interpolation. Let $\mathbf{V}$ be a smooth projective curve, let $\{x_1, \ldots, x_n\} \subseteq \mathbf{V}$ be pairwise distinct, where $n \in \mathbb{N}$, and $D \in \mathrm{Div}_{\mathbf{V}}$. Then there is $D' \in \mathrm{Div}_{\mathbf{V}}$ such that $D' \sim D$ and $\mathrm{supp}(D') \cap \{x_1, \ldots, x_n\} = \emptyset$.

Proof. We may assume that D is prime, and by induction we may assume that $\mathrm{supp}(D) \cap \{x_1, \ldots, x_{n-1}\} = \emptyset$, thus $D = (x_n)$. Moreover, by choosing an affine open subset containing $\{x_1, \ldots, x_n\}$, we may assume that $\mathbf{V}$ is affine.

Since $\mathbf{V}$ is affine, we may choose a local coordinate $t \in \mathscr{O}_{\mathbf{V}} \subseteq \mathscr{O}_{x_n}$ at x_n. By Hilbert's Nullstellensatz, for $j \in \{1, \ldots, n-1\}$ let $g_j \in \mathscr{O}_{\mathbf{V}}$ such that $g_j(x_j) = 1$, and $g_j(x_i) = 0$ for $j \neq i \in \{1, \ldots, n\}$.

Let $f := t + \sum_{j=1}^{n-1} a_j g_j^2 \in \mathscr{O}_{\mathbf{V}}$, where $a_j \in \overline{\mathbb{F}}$. Then for $i \in \{1, \ldots, n-1\}$ we get $f(x_i) = t(x_i) + \sum_{j=1}^{n-1} a_j g_j(x_i)^2 = t(x_i) + a_i$. Thus choosing $a_i \neq -t(x_i)$ we have $f(x_i) \neq 0$, thus $x_i \notin \mathrm{supp}((f))$.

Moreover, we have $g_j(x_n) = 0$, hence $\nu_{x_n}(\sum_{j=1}^{n-1} a_j g_j^2) \geq 2$, showing that $\nu_{x_n}(f) = \nu_{x_n}(t) = 1$. Thus in conclusion we get $\mathrm{supp}((f)-(x_n)) \cap \{x_1, \ldots, x_n\} = \emptyset$. Hence the divisor $D' := D - (f) \in \mathrm{Div}_{\mathbf{V}}$ is as desired. ♯

(19.5) The Riemann-Roch problem. Let $\mathbf{V}$ be a smooth projective curve, and let $D \in \mathrm{Div}_{\mathbf{V}}$. Then, looking for rational functions which are 'at least as regular' as prescribed by the 'wish-list' or 'bookkeeping device' $(-D)$, that is whose orders of zeroes are bounded below and whose orders of poles are bounded above by $(-D)$, we let

$$\mathscr{L}(D) = \mathscr{L}_{\mathscr{K}_{\mathbf{V}}}(D) := \{f \in \mathscr{K}_{\mathbf{V}}^*; (f) + D \geq 0\} \dot{\cup} \{0\}.$$

By the triangle inequality this an $\overline{\mathbb{F}}$-vector space, called the **(Riemann-Roch) ($\mathscr{L}$-)space of functions over** D. For example, if $D = 0$, then any element of $\mathscr{L}(0)$ is regular on $\mathbf{V}$, hence is constant, so that $\mathscr{L}(0) = \overline{\mathbb{F}}$.

The **Riemann-Roch problem** is to determine $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D))$. This actually only depends on the divisor class $[D] \in \mathrm{Pic}_{\mathbf{V}}$: If $D \sim D'$, that is $D - D' = (g)$ for some $g \in \mathscr{K}_{\mathbf{V}}^*$, then $\mathscr{L}(D) \to \mathscr{L}(D') \colon f \mapsto fg$ is an $\overline{\mathbb{F}}$-isomorphism.

Proposition. i) If $\deg(D) < 0$, then $\mathscr{L}(D) = \{0\}$.

ii) If $\deg(D) \geq 0$, then $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) \leq 1 + \deg(D)$.

Proof. We may assume that there is $0 \neq f \in \mathscr{L}(D)$. Then $\mathscr{L}((f) + D) \cong \mathscr{L}(D) \neq \{0\}$, where $(f) + D \geq 0$ and $\deg((f) + D) = \deg(D)$.

i) We have $(f) + D \geq 0$ and $\deg((f) + D) < 0$, a contradiction.

ii) We may assume that $D = \sum_{i=1}^{r} m_i \cdot (x_i)$ is effective.

We need some preparation first: For $x \in \mathbf{V}$, let $\widehat{\mathscr{O}}_x$ be the $\mathfrak{m}_x$**-adic completion** of $\mathscr{O}_x$. Then $\widehat{\mathscr{O}}_x$ is a graded $\overline{\mathbb{F}}_x$-algebra, such that there is a natural embedding $\mathscr{O}_x \to \widehat{\mathscr{O}}_x$ of $\overline{\mathbb{F}}_x$-algebras, and by **Cohen's Structure Theorem** we have $\widehat{\mathscr{O}}_x \cong \overline{\mathbb{F}}_x[[t_x]]$, the ring of **formal power series** in the indeterminate t_x. (In a certain sense $\widehat{\mathscr{O}}_x$ can be seen as the 'very local ring' of $\mathbf{V}$ at x.)

Moreover, for $m \in \mathbb{Z}$ the $\mathfrak{m}_x$**-adic completion** of $t_x^m \mathscr{O}_x$ is

$$\widehat{t_x^m \mathscr{O}_x} := \prod_{j \geq m} (t_x^j \mathscr{O}_x)/(t_x^{j+1} \mathscr{O}_x) \cong \prod_{j \geq m} \langle t_x^j \rangle_{\overline{\mathbb{F}}_x} \cong t_x^m \widehat{\mathscr{O}}_x \cong t_x^m \cdot \overline{\mathbb{F}}_x[[t_x]],$$

being a **graded $\widehat{\mathscr{O}}_x$-module**, consisting of **formal Laurent series**.

Now we get an $\overline{\mathbb{F}}$-linear map

$$\alpha \colon \mathscr{L}(D) \to \bigoplus_{i=1}^{r} (t_{x_i}^{-m_i} \widehat{\mathscr{O}}_{x_i} / \widehat{\mathscr{O}}_{x_i}) = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \langle t_{x_i}^{-j} \rangle_{\overline{\mathbb{F}}_{x_i}}.$$

Let $f \in \ker(\alpha)$; then f is regular at the points x_i, and since f is regular at all points in $\mathbf{V} \setminus \{x_1, \ldots, x_r\}$ anyway, we conclude that f is regular on $\mathbf{V}$, thus is constant. Hence we have $\dim_{\overline{\mathbb{F}}}(\ker(\alpha)) = 1$.

This yields $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) - 1 \leq \sum_{i=1}^{r} m_i \dim_{\overline{\mathbb{F}}}(\overline{\mathbb{F}}_{x_i}) = \sum_{i=1}^{r} m_i = \deg(D)$. ♯

(19.6) Canonical divisors. a) Let $\mathbf{V}$ be a smooth projective curve, and let $0 \neq \omega \in \Omega_{\mathscr{K}_{\mathbf{V}}}$. Then for any $x \in \mathbf{V}$ and any affine open neighborhood $U \subseteq \mathbf{V}$ of x, we have $\omega|_U \in \Omega_{\mathscr{K}_U}$. Now $\Omega_{\mathscr{K}_U}$ is a $\mathscr{K}_U$-vector space of dimension 1, being generated by $d_U(t_x)$, where $t_x \in \mathfrak{m}_x \setminus \mathfrak{m}_x^2 \subseteq \mathscr{O}_x$ is a local coordinate.

Hence we have $d_x(\omega) = d_x(\omega|_U) = f_{\omega,x} \cdot d_x(t_x)$, for some $f_{\omega,x} \in \mathscr{K}_U^*$; then $f_{\omega,x}$ is called the **derivative** of ω with respect to t_x; we also write $f_{\omega,x} = \frac{d_x(\omega)}{d_x(t_x)}$.

Letting $\nu_x(\omega) := \nu_x(f_{\omega,x}) \in \mathbb{Z}$ be the **valuation** of ω at the point x, using the formal Laurent series expansion $f_{\omega,x} = \sum_{j \geq \nu_x(\omega)} (f_{\omega,x})_j t_x^j \in t_x^{\nu_x(\omega)} \widehat{\mathscr{O}}_x$, we get

$$d_x(\omega) = \sum_{j \geq \nu_x(\omega)} (f_{\omega,x})_j t_x^j \cdot d_x(t_x) \in \widehat{\Omega}_x \cong \Omega_x \otimes_{\mathscr{O}_x} \widehat{\mathscr{O}}_x.$$

Now the **residue** of ω at the point x is defined as $\mathrm{res}_x(\omega) := (f_{\omega,x})_{-1} \in \overline{\mathbb{F}}_x \cong \overline{\mathbb{F}}$. (It can be shown that this is independent of the choice of t_x.)

Theorem: Residue Theorem. We have (note that the following sum is finite)

$$\sum_{x \in \mathbf{V}} \mathrm{res}_x(\omega) = \sum_{x \in \mathbf{V}, \nu_x(\omega) < 0} \mathrm{res}_x(\omega) = 0. \qquad \sharp$$

b) The **divisor** of ω is defined as follows; note that $(\omega) \in \mathrm{Div}_{\mathbf{V}}$ indeed:

$$(\omega) := \sum_{x \in \mathbf{V}} \nu_x(f_{\omega,x}) \cdot (x) = \sum_{x \in \mathbf{V}} \nu_x(\omega) \cdot (x) \in \mathrm{Div}_{\mathbf{V}}.$$

Since any non-empty open subset is dense in $\mathbf{V}$, we conclude that any rational differential form on $\mathbf{V}$ is uniquely determined by considering any open neighborhood of some arbitrarily chosen point, and that the quotient of two non-zero rational differential forms is given by a rational function in $\mathscr{K}_V^*$.

Hence $(\omega) \in \mathrm{Div}_{\mathbf{V}}$ defines a unique divisor class $[(\omega)] \in \mathrm{Pic}_{\mathbf{V}}$, being called the **canonical divisor class**. Any divisor on $\mathbf{V}$ being equivalent to $(\omega) \in \mathrm{Div}_{\mathbf{V}}$ is called a **canonical divisor** (the terminology going back to MUMFORD). Then the **(geometric) genus** of $\mathbf{V}$ is defined as

$$g = g_{\mathbf{V}} := \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega))) \geq 0.$$

(19.7) Special divisors. Let $\mathbf{V}$ be a smooth projective curve, and let $D \in \mathrm{Div}_{\mathbf{V}}$. Then the $\overline{\mathbb{F}}$-vector space

$$\Omega(D) = \Omega_{\mathscr{K}_{\mathbf{V}}}(D) := \{\omega \in \Omega_{\mathscr{K}_{\mathbf{V}}} \setminus \{0\}; (\omega) - D \geq 0\} \dot{\cup} \{0\}$$

is called the **space of differential forms over** D. Moreover, D is called **special** if $\Omega(D) \neq \{0\}$, and $\dim_{\overline{\mathbb{F}}}(\Omega(D)) \in \mathbb{N}_0$ is called the **index of specialty** of D. Indeed, $\Omega(D)$ is finitely generated as $\overline{\mathbb{F}}$-vector space:

Given a fixed $0 \neq \omega \in \Omega_{\mathscr{K}_{\mathbf{V}}}$, for $f \in \mathscr{K}_{\mathbf{V}}^*$ we have $f\omega \in \Omega(D)$, if and only if $(f\omega) = (f) + (\omega) - D \geq 0 \in \mathrm{Div}_{\mathbf{V}}$, which holds if and only if $f \in \mathscr{L}((\omega) - D)$; thus we have an $\overline{\mathbb{F}}$-isomorphism $\mathscr{L}((\omega) - D) \to \Omega(D) \colon f \mapsto f\omega$. $\sharp$

Example. i) For $D = 0$ we get the module of regular differential forms $\Omega_{\mathbf{V}} = \Omega(0)$, so that $\mathscr{L}((\omega)) \cong \Omega_{\mathbf{V}}$ as $\overline{\mathbb{F}}$-vector spaces, thus $g_{\mathbf{V}} = \dim_{\overline{\mathbb{F}}}(\Omega_{\mathbf{V}})$.

ii) For $D = (\omega)$, from $\mathscr{L}(0) \cong \overline{\mathbb{F}}$ we get $\Omega((\omega)) = \langle\omega\rangle_{\overline{\mathbb{F}}}$, thus $\dim_{\overline{\mathbb{F}}}(\Omega((\omega))) = 1$.

iii) If $\deg(D) > \deg((\omega))$, then $\mathscr{L}((\omega) - D) = \{0\}$ implies that D is non-special. (By ii) this bound on $\deg(D)$ is best possible.)

(19.8) Theorem: Riemann–Roch [∼1850, for complex curves].
Let $\mathbf{V}$ be a smooth projective curve having canonical divisor $(\omega) \in \mathrm{Div}_{\mathbf{V}}$, where $0 \neq \omega \in \Omega_{\mathscr{K}_{\mathbf{V}}}$, and genus $g = g_{\mathbf{V}} = \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega))) = \dim_{\overline{\mathbb{F}}}(\Omega_{\mathbf{V}}) \geq 0$. Then for any divisor $D \in \mathrm{Div}_{\mathbf{V}}$ we have

$$\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1 + \deg(D) - \dim_{\overline{\mathbb{F}}}(\Omega_{\mathbf{V}}) + \dim_{\overline{\mathbb{F}}}(\Omega(D)). \qquad \sharp$$

We are not at all able to prove this 'bread-and-butter theorem' here, but in order to underline its importance we rush to present various applications:

Corollary. We have $\deg((\omega)) = 2g - 2$.

Proof. We apply the theorem to the (special) divisor $D = (\omega)$: We get $g = \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega))) = 1 + \deg((\omega)) - g + \dim_{\overline{\mathbb{F}}}(\Omega((\omega))) = \deg((\omega)) + 2 - g$. $\sharp$

(19.9) Plane curves. Let $\mathbf{V} \subseteq \mathbf{P}^2$ be a smooth **plane** projective curve.

Then, by **Krull's principal ideal theorem**, we have $\mathbf{V} = \mathbf{V}(F)$, where $F \in \mathscr{R}_{\mathbf{P}^2} \cong \overline{\mathbb{F}}[X_0, X_1, X_2]$ is irreducible and homogeneous, such that

$$[(\partial_{X_i} F)(x_0, x_1, x_2)]_i \neq [0, 0, 0],$$

for $[x_0 \colon x_1 \colon x_2] \in \mathbf{P}^2$ such that $F(x_0, x_1, x_2) = 0$. Conversely, any such polynomial F defines a smooth projective curve, and given $\mathbf{V}(F)$ then F is unique up to scalar multiples. Then $d := \deg(F) \in \mathbb{N}$ is called the **degree** of $\mathbf{V}$, where $\mathbf{V}$ is called a **conic**, **cubic**, **quartic**, **quintic**, for $d = 2, 3, 4, 5$, respectively.

Theorem. Then $(d-3) \cdot (X_0) \in \mathrm{Div}_{\mathbf{V}}$ is a canonical divisor of $\mathbf{V}$.

Proof. i) We consider the open affine covering of $\mathbf{V}$ given by $U := \mathbf{V} \cap D_{X_0}$ and $V := \mathbf{V} \cap D_{X_1}$ and $W := \mathbf{V} \cap D_{X_2}$. Hence U can be identified with $\mathbf{V}(G)$, where $A_i := \frac{X_i}{X_0}$ and $G(A_1, A_2) := F(1, A_1, A_2) = \frac{1}{X_0^d} \cdot F(X_0, X_1, X_2)$.

Similarly, V can be identified with $\mathbf{V}(H)$, where $B_i := \frac{X_i}{X_1}$ and $H(B_0, B_2) := F(B_0, 1, B_2) = \frac{1}{X_1^d} \cdot F(X_0, X_1, X_2)$; and W can be identified with $\mathbf{V}(I)$, where $C_i := \frac{X_i}{X_2}$ and $I(C_0, C_1) := F(C_0, C_1, 1) = \frac{1}{X_2^d} \cdot F(X_0, X_1, X_2)$.

ii) We first consider U, where we may assume that $U \neq \emptyset$: Let $\partial_i G := \partial_{A_i} G$, and let $D_0 := D_{\partial_1 G} \cap D_{\partial_2 G} \subseteq U$; then by smoothness we have $U = D_{\partial_1 G} \cup D_{\partial_2 G}$. Moreover, $\{d(A_1), d(A_2)\}$ generates Ω_U as an $\mathscr{R}_U$-module, subject to the relation $(\partial_1 G) d(A_1) + (\partial_2 G) d(A_2) = 0 \in \Omega_U$.

Hence on D_0 we have $\frac{1}{\partial_2 G} d(A_1) = -\frac{1}{\partial_1 G} d(A_2) \in \Omega_{D_0}$. Thus letting

$$\omega_0' := -\frac{1}{\partial_1 G} d(A_2) \in \Omega_{D_{\partial_1 G}} \quad \text{and} \quad \omega_0'' := \frac{1}{\partial_2 G} d(A_1) \in \Omega_{D_{\partial_2 G}}$$

defines a regular differential form $\omega_0 \in \Omega_U$, vanishing nowhere on U. Hence we have $\Omega_U = \omega_0 \cdot \mathscr{R}_U$, being a free $\mathscr{R}_U$-module (of rank 1). Let $\omega \in \Omega_{\mathscr{K}_{\mathbf{V}}}$ be the extension of ω_0 to $\mathbf{V}$; hence we have $\mathrm{supp}((\omega)) \cap U = \emptyset$.

Similarly, whenever $V \neq \emptyset$ we have $\Omega_V = \omega_1 \cdot \mathscr{R}_V$, where ω_1 restricts to $\omega_1'' := \frac{1}{\partial_2 H} d(B_0) \in \Omega_{D_{\partial_2 H}}$; and whenever $W \neq \emptyset$ we have $\Omega_W = \omega_2 \cdot \mathscr{R}_W$, where ω_2 restricts to $\omega_2'' := \frac{1}{\partial_1 I} d(C_0) \in \Omega_{D_{\partial_1 I}}$.

iii) If $V \neq \emptyset$, then on $U \cap V$ we have $A_1 = \frac{X_1}{X_0} = \frac{1}{B_0}$ and $A_2 = \frac{X_2}{X_0} = \frac{X_2}{X_1} \cdot \frac{X_1}{X_0} = \frac{B_2}{B_0}$. Thus we get

$$H(B_0, B_2) = \frac{X_0^d}{X_1^d} \cdot G(A_1, A_2) = B_0^d \cdot G(\frac{1}{B_0}, \frac{B_2}{B_0}),$$

entailing that $\partial_2 H = B_0^{d-1} \cdot (\partial_2 G)(\frac{1}{B_0}, \frac{B_2}{B_0}) = B_0^{d-1} \cdot \partial_2 G$. Moreover, we have $d(A_1) = d(\frac{1}{B_0}) = -\frac{1}{B_0^2} d(B_0)$. Hence on $D_{\partial_2 H} \cap U \subseteq D_{\partial_2 G}$ we have

$$\omega|_{U \cap D_{\partial_2 H}} = \frac{1}{\partial_2 G} d(A_1) = -\frac{B_0^{d-3}}{\partial_2 H} d(B_0).$$

Similarly, if $W \neq \emptyset$, then on $U \cap W$ we have $A_1 = \frac{X_1}{X_0} = \frac{X_1}{X_2} \cdot \frac{X_2}{X_0} = \frac{C_1}{C_0}$ and $A_2 = \frac{X_2}{X_0} = \frac{1}{C_0}$. Thus we get

$$I(C_0, C_1) = \frac{X_0^d}{X_2^d} \cdot G(A_1, A_2) = C_0^d \cdot G(\frac{C_1}{C_0}, \frac{1}{C_0}),$$

entailing that $\partial_1 I = C_0^{d-1} \cdot (\partial_1 G)(\frac{C_1}{C_0}, \frac{1}{C_0}) = C_0^{d-1} \cdot \partial_1 G$. Moreover, we have $d(A_2) = d(\frac{1}{C_0}) = -\frac{1}{C_0^2} d(C_0)$. Hence on $D_{\partial_1 I} \cap U \subseteq D_{\partial_1 G}$ we have

$$\omega|_{U \cap D_{\partial_1 I}} = -\frac{1}{\partial_1 G} d(A_2) = \frac{C_0^{d-3}}{\partial_1 I} d(C_0).$$

In conclusion we have the canonical divisor $(\omega) = (d-3) \cdot (X_0) \in \mathrm{Div}_{\mathbf{V}}$. ♯

Corollary: Plücker formula [1834]. We have $g_{\mathbf{V}} = \frac{1}{2}(d-1)(d-2)$.

Proof. Let $0 \neq \varphi = \frac{f}{g} \in \mathscr{K}_{\mathbf{V}}$, where $f, g \in \overline{\mathbb{F}}[X_0, X_1, X_2]$ are coprime and homogeneous such that $k := \deg(f) = \deg(g) \geq 0$. Then we have $\varphi\omega \in \Omega_{\mathbf{V}}$, that is $(\varphi\omega) = (\varphi) + (\omega) \geq 0$, if and only if $\varphi \in \mathscr{L}((\omega)) = \mathscr{L}((d-3) \cdot (X_0))$. The latter is equivalent to $\varphi|_U$ being regular, and $\nu_x(\varphi) \geq -(d-3)$ for $x \in \mathbf{V} \cap \mathbf{V}(X_0)$.

The first condition is equivalent to $\varphi(1, A_1, A_2) = \frac{f(1, A_1, A_2)}{g(1, A_1, A_2)} \in \mathscr{K}_{\mathbf{U}}$ being a polynomial, that is $g = X_0^k$ up to associates, or equivalently $\varphi = \frac{1}{X_0^k} f(X_0, X_1, X_2)$. In this case, by coprimeness we get $\nu_x(\varphi) = -k$, for $x \in \mathbf{V} \cap \mathbf{V}(X_0)$. Hence we have $\varphi = \frac{1}{X_0^k} f(X_0, X_1, X_2) \in \mathscr{L}((\omega))$ if and only if $k \leq d-3$.

Thus $\Omega_{\mathbf{V}} = \{0\}$ if $d \leq 2$; hence we may assume that $d \geq 3$. Then we may write $\varphi = \frac{1}{X_0^{d-3}} f$, where $f \in \overline{\mathbb{F}}[X_0, X_1, X_2]_{d-3}$. Then $\overline{\mathbb{F}}[X_0, X_1, X_2]_{d-3} \to \mathscr{L}((\omega))\colon f \mapsto \frac{1}{X_0^{d-3}} f$ is injective, thus is an $\overline{\mathbb{F}}$-isomorphism:

If $f, f' \in \overline{\mathbb{F}}[X_0, X_1, X_2]_{d-3}$ such that $\frac{1}{X_0^{d-3}} f' = \frac{1}{X_0^{d-3}} f \in \mathscr{K}_{\mathbf{V}}$, then restricting to U implies $(f - f')(1, A_1, A_2) = 0 \in \mathscr{R}_U$, that is $G(A_1, A_2) = F(1, A_1, A_2) \mid (f - f') \in \overline{\mathbb{F}}[A_1, A_2]$; since $U \neq \emptyset$ we have $\deg(G) = d$, implying $f = f'$.

This yields (note that this also holds true for $d \leq 2$)

$$g_{\mathbf{V}} = \dim_{\overline{\mathbb{F}}}(\Omega_{\mathbf{V}}) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega))) = \dim_{\overline{\mathbb{F}}}(\overline{\mathbb{F}}[X_0, X_1, X_2]_{d-3}) = \binom{d-1}{2}. \quad \sharp$$

In view of **Bézout's Theorem**, we observe the following special case: We have $(d-3) \cdot \deg((X_0)) = \deg(\omega) = 2g_{\mathbf{V}} - 2 = d(d-3)$, thus $\deg((X_0)) = d$.

(19.10) The projective line. a) We consider the projective line $\mathbf{P}^1 = \mathbf{A}^1 \dot\cup \{\infty\} = \overline{\mathbb{F}} \dot\cup \{\infty\}$, where $\infty := [0\colon 1] \in \mathbf{P}^1$, and $\mathbf{A}^1 \subseteq \mathbf{P}^1$ is an affine open subset, with respect to the embedding $\mathbf{A}^1 \to \mathbf{P}^1\colon x \mapsto [1\colon x]$. Then we have $\mathscr{K}_{\mathbf{P}^1} \cong \overline{\mathbb{F}}[X_0, X_1]_{(\{0\})} \cong \overline{\mathbb{F}}[X]_{\{0\}} = \overline{\mathbb{F}}(X)$, via $\frac{X_1}{X_0} \mapsto X$.

Then $\mathbf{P}^1$ can be identified with the plane curve $\mathbf{V} := \mathbf{V}(X_2) = \{[x_0\colon x_1\colon 0] \in \mathbf{P}^2; [x_0\colon x_1] \in \mathbf{P}^1\} \subseteq \mathbf{P}^2$. Since $\mathbf{V}$ is defined by the polynomial $F = X_2$ of degree 1, by Plücker's formula we conclude that $\mathbf{P}^1$ has genus $g = 0$.

Alternatively, to find the genus of $\mathbf{P}^1$, we may more directly argue as follows:

We have $(X) = (0) - (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$. For $k \in \mathbb{N}_0$ we have $\{1, X, \ldots, X^k\} \subseteq \mathscr{L}(k \cdot (\infty))$, thus $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(k \cdot (\infty)) \geq k + 1$. Since the divisor $k \cdot (\infty)$ is non-special for $k \geq 2g - 1$, we get $k + 1 \leq \dim_{\overline{\mathbb{F}}}(\mathscr{L}(k \cdot (\infty)) = 1 + k - g$, implying $g = 0$. (Hence we have $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(k \cdot (\infty)) = k + 1$ for all $k \in \mathbb{N}_0$.) $\sharp$

b) Using the notation of (19.9), there is $\omega \in \Omega_{\mathscr{K}_{\mathbf{V}}}$ given piecewise as follows: On $U = D_{X_0} = \mathbf{A}^1$, from $G = F(1, \frac{X_1}{X_0}, \frac{X_2}{X_0}) = \frac{X_2}{X_0} = A_2$ we get $\omega_0 := d(\frac{X_1}{X_0}) \in \Omega_U$. Similarly, on $V = D_{X_1}$, from $H = F(\frac{X_0}{X_1}, 1, \frac{X_2}{X_1}) = \frac{X_2}{X_1} = B_2$ we get $\omega_1 := -(\frac{X_1}{X_0})^2 \cdot d(\frac{X_0}{X_1}) \in \Omega_V$. This yields the canonical divisor $(\omega) = -2 \cdot (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$; thus $\mathrm{res}_x(\omega) = 0$ for $x \in \mathbf{P}^1$.

Moreover, for $\alpha \in \mathbf{A}^1$ the function $\frac{X_0}{X_1 - \alpha X_0} \in \mathscr{K}_{\mathbf{P}^1}$ has the only pole $[1\colon \alpha]$ and the only zero $\infty = [0\colon 1]$, both of order 1; hence $(\frac{X_0}{X_1 - \alpha X_0} \cdot \omega) = -([1\colon \alpha]) - (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$. On U we have $d(\frac{X_1}{X_0}) = d(\frac{X_1 - \alpha X_0}{X_0})$, thus $\mathrm{res}_\alpha(\frac{X_0}{X_1 - \alpha X_0} \cdot \omega) = \mathrm{res}_\alpha(\frac{X_0}{X_1 - \alpha X_0} \cdot d(\frac{X_1 - \alpha X_0}{X_0})) = 1$; we get (as predicted by the Residue Theorem)

$$\mathrm{res}_\infty(\frac{X_0}{X_1 - \alpha X_0} \cdot \omega) = -\mathrm{res}_\infty(\frac{X_1}{X_1 - \alpha X_0} \cdot \frac{X_1}{X_0} \cdot d(\frac{X_0}{X_1})) = -1.$$

Identifying $\mathscr{K}_{\mathbf{P}^1} \cong \overline{\mathbb{F}}(X)$, we get $\omega_0 = d(X)$ and $\omega_1 = -X^2 \cdot d(\frac{1}{X})$, where $\mathrm{res}_\infty(\frac{1}{X - \alpha} \cdot \omega) = -1$ and $\mathrm{res}_\alpha(\frac{1}{X - \alpha} \cdot \omega) = 1$, while $\mathrm{res}_x(\frac{1}{X - \alpha} \cdot \omega) = 0$ elsewise.

c) Actually, there are not too many smooth projective curves of genus 0:

Theorem. Let $\mathbf{V}$ be a smooth projective curve of genus 0. Then $\mathbf{V} \cong \mathbf{P}^1$.

Proof. Let $D \in \mathrm{Div}_{\mathbf{V}}$ have degree 1. Since $\deg(D) \geq 2g-1$, it follows that D is non-special, thus $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1 + \deg(D) - g = 2$. Then, since $\mathscr{L}(D) \neq \{0\}$, there is an effective divisor $D' \sim D$.

We have $\deg(D') = 1$, so that $D' = (x)$, for some $x \in \mathbf{V}$. Since $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D')) = 2$, there is a non-constant rational function $f \in \mathscr{L}(D')$. Thus $(f) \geq -D' = -(x)$, hence the pole divisor of f equals $(f)_\infty = (x)$.

Considering f as a morphism $f\colon \mathbf{V} \to \mathbf{P}^1$, we have $f^*((\infty)) = (f)_\infty$, hence by the degree formula we get $\deg(f) = \deg(f^*((\infty))) = \deg((f)_\infty) = 1$. We conclude that the associated field extension $f^*(\mathscr{K}_{\mathbf{P}^1}) \subseteq \mathscr{K}_{\mathbf{V}}$ has degree 1, that is $\mathscr{K}_{\mathbf{V}} = f^*(\mathscr{K}_{\mathbf{P}^1})$, which implies that f is an isomorphism. ♯

20 Geometric Goppa codes

(20.1) Rationality. a) Let $\mathbb{F}$ be a perfect field, and let $\mathbb{F} \subseteq \overline{\mathbb{F}}$ be an algebraic closure of $\mathbb{F}$. (Later on we will let $\mathbb{F} := \mathbb{F}_q$.)

Then an affine or projective variety $\mathbf{V}$ over $\overline{\mathbb{F}}$ is said to be **defined over** $\mathbb{F}$, if its defining prime ideal in $\overline{\mathbb{F}}[\mathcal{X}]$ is generated by polynomials in $\mathbb{F}[\mathcal{X}]$. Hence we may replace all the algebraic constructions made earlier over $\overline{\mathbb{F}}$ by similar ones over $\mathbb{F}$, where the old ones are recovered by scalar extension from $\mathbb{F}$ to $\overline{\mathbb{F}}$. In particular, there is a separably generated field extension $\mathbb{F} \subseteq \mathscr{K}_{\mathbf{V},\mathbb{F}}$, such that $\mathscr{K}_{\mathbf{V}} \cong \mathscr{K}_{\mathbf{V},\mathbb{F}} \otimes_{\mathbb{F}} \overline{\mathbb{F}}$ is obtained by extending the field of constants.

The Galois group $\Gamma := \mathrm{Aut}_{\mathbb{F}}(\overline{\mathbb{F}})$ acts naturally on $\mathbf{V}$. The fixed points with respect to this action, that is the points $[x_1, \ldots, x_n] \in \mathbf{A}^n$ in the affine case, and $[x_0\colon x_1\colon \ldots\colon x_n] \in \mathbf{P}^n$ in the projective case, having coordinates $x_i \in \mathbb{F}$, are called **$\mathbb{F}$-rational**. Let $\mathbf{V}(\mathbb{F}) \subseteq \mathbf{V}$ be the set of $\mathbb{F}$-rational points of $\mathbf{V}$; if $\mathbb{F} = \mathbb{F}_q$ is a finite field, we also write $\mathbf{V}(q) := \mathbf{V}(\mathbb{F}_q)$, which is a finite set.

In particular, for $x \in \mathbf{V}(\mathbb{F})$ we have $\mathbb{F}_x := \mathscr{O}_{x,\mathbb{F}}/\mathfrak{m}_{x,\mathbb{F}} \cong \mathbb{F}$; thus the $\mathbb{F}$-rational points are also called the points of **degree** $[\mathbb{F}_x\colon \mathbb{F}] = 1$ of $\mathbf{V}(\mathbb{F})$.

b) Let $\mathbf{V}$ be smooth projective curve defined over $\mathbb{F}$. Then the Γ-action on $\mathbf{V}$ induces an action on $\mathrm{Div}_{\mathbf{V}}$, whose fixed points are called **$\mathbb{F}$-rational**; in particular, any divisor supported on $\mathbf{V}(\mathbb{F})$ only is $\mathbb{F}$-rational.

For any $\mathbb{F}$-rational divisor $D \in \mathrm{Div}_{\mathbf{V}}$, the $\overline{\mathbb{F}}$-vector space $\mathscr{L}(D) \leq \mathscr{K}_{\mathbf{V}}$ is Γ-stable, thus carries an $\mathbb{F}$-rational structure, so that there is an $\mathbb{F}$-vector space $\mathscr{L}_{\mathbb{F}}(D) \leq \mathscr{K}_{\mathbf{V},\mathbb{F}}$ such that $\mathscr{L}(D) \cong \mathscr{L}_{\mathbb{F}}(D) \otimes_{\mathbb{F}} \overline{\mathbb{F}}$; see Exercise (24.11).

Since for all $\mathbb{F}$-rational affine open subsets $U \subseteq \mathbf{V}$, for their Kähler differentials we have $\Omega_{\mathscr{K}_{U,\mathbb{F}}} \otimes_{\mathbb{F}} \overline{\mathbb{F}} \cong \Omega_{\mathscr{K}_U}$, we conclude that there is an $\mathbb{F}$-rational differential form $\omega_{\mathbb{F}} \neq 0$ on $\mathbf{V}$; we let $\Omega_{\mathscr{K}_{\mathbf{V},\mathbb{F}}} := \mathscr{K}_{\mathbf{V},\mathbb{F}} \cdot \omega_{\mathbb{F}} \subseteq \Omega_{\mathscr{K}_{\mathbf{V}}}$. Since Γ maps local coordinates to local coordinates, we conclude that $(\omega_{\mathbb{F}}) \in \mathrm{Div}_{\mathbf{V}}$ is

$\mathbb{F}$-rational. Moreover, letting $\Omega_{\mathbb{F}}(D) = \Omega_{\mathscr{K}_{\mathbf{V},\mathbb{F}}}(D) := \Omega(D) \cap \Omega_{\mathscr{K}_{\mathbf{V},\mathbb{F}}}$, we get an $\mathbb{F}$-isomorphism $\mathscr{L}_{\mathbb{F}}((\omega_{\mathbb{F}}) - D) \to \Omega_{\mathbb{F}}(D) \colon f \mapsto f\omega_{\mathbb{F}}$.

(20.2) Geometric Goppa codes [1981]. a) Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$ defined over $\mathbb{F}_q$, let $\{p_1, \ldots, p_n\} \subseteq \mathbf{V}(\mathbb{F}_q)$ be pairwise distinct $\mathbb{F}_q$-rational points, for some $n \in \mathbb{N}$, playing the role of places, and let $P := \sum_{i=1}^{n}(p_i) \in \mathrm{Div}_{\mathbf{V}}$ be the associated ($\mathbb{F}_q$-rational) divisor.

Let $G \in \mathrm{Div}_{\mathbf{V}}$ be an $\mathbb{F}_q$-rational divisor such that $\mathrm{supp}(G) \cap \{p_1, \ldots, p_n\} = \emptyset$. Then the associated **geometric Goppa code** is defined as

$$\mathcal{G}_P(G) := \{[f(p_1), \ldots, f(p_n)] \in \mathbb{F}_q^n; f \in \mathscr{L}_{\mathbb{F}_q}(G)\} \leq \mathbb{F}_q^n,$$

where because of disjointness we have $\nu_{p_i}(f) \geq 0$, so that f is regular p_i, and by $\mathbb{F}_q$-rationality we have $f(p_i) \in (\mathbb{F}_q)_{p_i} \cong \mathbb{F}_q$. Thus we have

$$\overline{\mathcal{G}}_P(G) := \mathcal{G}_P(G) \otimes_{\mathbb{F}_q} \overline{\mathbb{F}} = \{[f(p_1), \ldots, f(p_n)] \in \overline{\mathbb{F}}^n; f \in \mathscr{L}(G)\} \leq \overline{\mathbb{F}}^n,$$

hence $\dim_{\mathbb{F}_q}(\mathcal{G}_P(G)) = \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P(G))$ and $d(\mathcal{G}_P(G)) \geq d(\overline{\mathcal{G}}_P(G))$.

Theorem. Let $k := \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P(G)) \in \mathbb{N}_0$ and $d := d(\overline{\mathcal{G}}_P(G)) \in \mathbb{N} \dot\cup \{\infty\}$.

a) Then we have $k = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) - \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G - P))$ and $d \geq n - \deg(G)$.

Then $\delta := n - \deg(G) \in \mathbb{Z}$ is called the **designed distance** of $\mathcal{G}_P(G)$.

b) If $\deg(G) \leq n - 1$, then we have $k = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) \geq 1 + \deg(G) - g$.

Moreover, if $2g - 1 \leq \deg(G) \leq n - 1$, then we have $k = 1 + \deg(G) - g$.

c) Let $\deg(G) \leq n - 1$ and $k \geq 1$. Then we have $d = \delta$ if and only if there is $P' \in \mathrm{Div}_{\mathbf{V}}$ such that $0 \leq P' < P$, where $\deg(G - P') = 0$ and $\mathscr{L}(G - P') \neq \{0\}$.

Proof. a) We consider the surjective $\overline{\mathbb{F}}$-linear evaluation map

$$\psi_P \colon \mathscr{L}(G) \to \overline{\mathcal{G}}_P(G) \colon f \mapsto [f(p_i)]_i.$$

Then we have $f \in \ker(\psi_P)$ if and only if $\nu_{p_i}(f) > 0$, for $i \in \{1, \ldots, n\}$, which holds if and only if $f \in \mathscr{L}(G - P)$; hence $k = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) - \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G - P))$.

In order to estimate d, we may assume that $k \geq 1$. Then let $f \in \mathscr{L}(G)$ such that $\psi_P(f) \neq 0$, and let $\mathcal{I} := \{1, \ldots, n\} \setminus \mathrm{supp}(\psi_P(f))$; then we have $\mathrm{wt}(\psi_P(f)) = n - |\mathcal{I}|$. We have $0 \neq f \in \mathscr{L}(G - \sum_{i \in \mathcal{I}}(p_i))$, entailing that $0 \leq \deg(G - \sum_{i \in \mathcal{I}}(p_i)) = \deg(G) - n + \mathrm{wt}(\psi_P(f))$, thus $\mathrm{wt}(\psi_P(f)) \geq n - \deg(G) = \delta$.

b) Let $\deg(G) \leq n - 1$. Then from $\deg(G - P) < 0$ we get $\mathscr{L}(G - P) = \{0\}$, so that $k = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G))$, thus by the Riemann-Roch Theorem we get $k = 1 + \deg(G) - g + \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega) - G)) \geq 1 + \deg(G) - g$.

Let $2g - 1 \leq \deg(G) \leq n - 1$. Then $\deg((\omega) - G) < 0$ implies $\mathscr{L}((\omega) - G) = \{0\}$, that is G is non-special. Thus we get $k = 1 + \deg(G) - g$.

c) i) Let $d = \delta$. Then there is $0 \neq f \in \mathscr{L}(G)$ such that $\mathrm{wt}(\psi_P(f)) = \delta = n - \deg(G) > 0$. Let $\mathcal{I} := \{1, \dots, n\} \backslash \mathrm{supp}(\psi_P(f))$, and let $0 \leq P' := \sum_{i \in \mathcal{I}}(p_i) < P$; hence $\deg(P') = |\mathcal{I}| = \deg(G)$. Then we have $f \in \mathscr{L}(G - P') \neq \{0\}$.

ii) Let $P' = \sum_{i \in \mathcal{I}}(p_i) \in \mathrm{Div}_{\mathbf{V}}$ be as asserted, for some $\mathcal{I} \subseteq \{1, \dots, n\}$ such that $|\mathcal{I}| = \deg(G)$, and let $0 \neq f \in \mathscr{L}(G - P')$. Then we have $\mathrm{supp}(\psi_P(f)) \subseteq \{1, \dots, n\} \setminus \mathcal{I}$, entailing $0 < \mathrm{wt}(\psi_P(f)) \leq n - \deg(G) = \delta$. Thus we have $d \leq \delta$, while we have $d \geq \delta$ anyway. ♯

Corollary. Let $\deg(G) \leq n - 1$, let $k \geq 1$ (for example, whenever $\deg(G) \geq g$), and let $d := d(\mathcal{G}_P(D))$ (contrary to the notation used above).

a) If $\{f_1, \dots, f_k\} \subseteq \mathscr{L}_{\mathbb{F}_q}(G)$ is an $\mathbb{F}_q$-basis, then $[f_j(p_i)]_{ji} \in \mathbb{F}_q^{k \times n}$ is a generator matrix of $\mathcal{G}_P(G)$.

b) We have $d - 1 \leq n - k \leq d - 1 + g$; thus $\mathcal{G}_P(G)$ is an MDS code for $g = 0$.

Proof. b) We have $d \geq n - \deg(G)$. Hence we have $n - k \leq d + \deg(G) - (1 + \deg(G) - g) = d - 1 + g$, while $d - 1 \leq n - k$ is the Singleton bound. ♯

(20.3) Remark: Generalized geometric Goppa codes. We can actually discard the disjointness property of $G \in \mathrm{Div}_{\mathbf{V}}$: By (19.4) there is $t \in \mathcal{K}_{\mathbf{V},\mathbb{F}_q}$ such that $\mathrm{supp}(G - (t)) \cap \{p_1, \dots, p_n\} = \emptyset$, and thus we may let

$$\mathcal{G}_P(G - (t)) := \{[f(p_1), \dots, f(p_n)] \in \mathbb{F}_q^n; f \in \mathscr{L}_{\mathbb{F}_q}(G - (t))\} \leq \mathbb{F}_q^n.$$

Then $\mathcal{G}_P(G - (t))$ depends on the choice of t (which is far from being unique), but the equivalence class of codes thus obtained does not:

Let $s \in \mathcal{K}_{\mathbf{V},\mathbb{F}_q}$ such that $\mathrm{supp}(G - (s)) \cap \{p_1, \dots, p_n\} = \emptyset$, then

$$\mathscr{L}_{\mathbb{F}_q}(G - (t)) \to \mathscr{L}_{\mathbb{F}_q}(G - (s)) \colon f \mapsto \frac{s}{t} \cdot f$$

is an $\mathbb{F}_q$-isomorphism. Since $\nu_{p_i}(s) = \nu_{p_i}(t)$, we conclude that $\frac{s}{t}$ is regular at p_i such that $(\frac{s}{t})(p_i) \neq 0$, for $i \in \{1, \dots, n\}$. This entails

$$\mathcal{G}_P(G - (s)) = \mathcal{G}_P(G - (t)) \cdot \mathrm{diag}[(\frac{s}{t})(p_i)]_i.$$ ♯

(20.4) Dual geometric Goppa codes. a) Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$ defined over $\mathbb{F}_q$, let $\{p_1, \dots, p_n\} \subseteq \mathbf{V}(\mathbb{F}_q)$ be pairwise distinct, where $n \in \mathbb{N}$, and let $P := \sum_{i=1}^n (p_i) \in \mathrm{Div}_{\mathbf{V}}$ be the associated divisor.

Let $G \in \mathrm{Div}_{\mathbf{V}}$ be an $\mathbb{F}_q$-rational divisor such that $\mathrm{supp}(G) \cap \{p_1, \dots, p_n\} = \emptyset$. Then the associated **dual geometric Goppa code** is defined as

$$\mathcal{G}_P^*(G) := \{[\mathrm{res}_{p_1}(\omega), \dots, \mathrm{res}_{p_n}(\omega)] \in \mathbb{F}_q^n; \omega \in \Omega_{\mathbb{F}_q}(G - P)\} \leq \mathbb{F}_q^n,$$

where by $\mathbb{F}_q$-rationality we indeed have $\mathrm{res}_{p_i}(\omega) \in (\mathbb{F}_q)_{p_i} \cong \mathbb{F}_q$. Likewise, using the $\mathbb{F}_q$-isomorphism $\mathscr{L}_{\mathbb{F}_q}((\omega_{\mathbb{F}_q}) + P - G) \to \Omega_{\mathbb{F}_q}(G - P)\colon f \mapsto f\omega_{\mathbb{F}_q}$ we get

$$\mathcal{G}_P^*(G) := \{[\mathrm{res}_{p_1}(f\omega_{\mathbb{F}_q}), \ldots, \mathrm{res}_{p_n}(f\omega_{\mathbb{F}_q})] \in \mathbb{F}_q^n; f \in \mathscr{L}_{\mathbb{F}_q}((\omega_{\mathbb{F}_q}) + P - G)\} \leq \mathbb{F}_q^n.$$

Again, for $\overline{\mathcal{G}}_P^*(G) := \mathcal{G}_P^*(G) \otimes_{\mathbb{F}_q} \overline{\mathbb{F}}$ we similarly get

$$\overline{\mathcal{G}}_P^*(G) = \{[\mathrm{res}_{p_1}(\omega), \ldots, \mathrm{res}_{p_n}(\omega)] \in \overline{\mathbb{F}}^n; \omega \in \Omega(G - P)\} \leq \overline{\mathbb{F}}^n,$$

thus we have $\dim_{\mathbb{F}_q}(\mathcal{G}_P^*(G)) = \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P^*(G))$ and $d(\mathcal{G}_P^*(G)) \geq d(\overline{\mathcal{G}}_P^*(G))$.

Theorem. Let $k^* := \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P^*(G))$ and $d^* := d(\overline{\mathcal{G}}_P^*(G))$.

a) Then we have $k^* = \dim_{\overline{\mathbb{F}}}(\Omega(G-P)) - \dim_{\overline{\mathbb{F}}}(\Omega(G))$ and $d^* \geq \deg(G) - (2g-2)$.

Then $\delta^* := \deg(G) - (2g-2) \in \mathbb{Z}$ is called the **designed distance** of $\mathcal{G}_P^*(G)$.

b) If $\deg(G) \geq 2g-1$, then we have $k^* = \dim_{\overline{\mathbb{F}}}(\Omega(G-P)) \geq n - (1 + \deg(G) - g)$.

Moreover, if $2g - 1 \leq \deg(G) \leq n - 1$, then we have $k^* = 1 + \deg(G) - g$.

c) Let $\deg(G) \geq 2g - 1$ and $k^* \geq 1$. Then $d^* = \delta^*$ if and only if there is $P' \in \mathrm{Div}_{\mathbf{V}}$ such that $0 < P' \leq P$, where $\deg(G - P') = 2g - 2$ and $\Omega(G - P') \neq \{0\}$.

Proof. a) We consider the surjective $\overline{\mathbb{F}}$-linear evaluation map

$$\psi_P^*\colon \Omega(G - P) \to \overline{\mathcal{G}}_P(G)\colon \omega \mapsto [\mathrm{res}_{p_i}(\omega)]_i.$$

We show that $\ker(\psi_P^*) = \Omega(G)$, thus $k^* = \dim_{\overline{\mathbb{F}}}(\Omega(G - P)) - \dim_{\overline{\mathbb{F}}}(\Omega(G))$:

Since $\omega \in \Omega(G - P)$ has a pole of order at most 1 at p_i, we have $\mathrm{res}_{p_i}(\omega) = 0$ if and only if ω is regular at p_i, that is $\nu_{p_i}(\omega) \geq 0$. Thus we have $\omega \in \ker(\psi_P^*)$ if and only if $\nu_{p_i}(\omega) \geq 0$, for $i \in \{1, \ldots, n\}$, which holds if and only if $\omega \in \Omega(G)$.

In order to estimate d^*, we may assume that $k^* \geq 1$. Then let $\omega \in \Omega(G - P)$ such that $\psi_P^*(\omega) \neq 0$, and let $\mathcal{I} := \mathrm{supp}(\psi_P^*(\omega))$; then $\mathrm{wt}(\psi_P^*(\omega)) = |\mathcal{I}|$. We have $0 \neq \omega \in \Omega(G - \sum_{i \in \mathcal{I}}(p_i))$, entailing that $\deg(G - \sum_{i \in \mathcal{I}}(p_i)) = \deg(G) - \mathrm{wt}(\psi_P^*(\omega)) \leq 2g - 2$, thus we have $\mathrm{wt}(\psi_P^*(\omega)) \geq \deg(G) - (2g - 2) = \delta^*$.

b) Let $\deg(G) \geq 2g - 1$. Then from $\deg((\omega) - G) = (2g - 2) - \deg(G) < 0$ we get $\mathscr{L}((\omega) - G) = \{0\}$, thus $\Omega(G) = \{0\}$, so that

$$k^* = \dim_{\overline{\mathbb{F}}}(\Omega(G - P)) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega) + P - G)).$$

Using $\deg((\omega)) = 2g - 2$, the Riemann-Roch Theorem yields

$$k^* = 1 - g + \deg((\omega) + P - G) + \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G - P)) \geq g - 1 + n - \deg(G).$$

Finally, if additionally $\deg(G) \leq n-1$, then $\deg(G-P) < 0$ implies $\mathscr{L}(G-P) = \{0\}$, entailing equality $k^* = n + g - 1 - \deg(G)$.

c) i) Let $d^* = \delta^*$. Then there is $0 \neq \omega \in \Omega(G-P)$ such that $\mathrm{wt}(\psi_P^*(\omega)) = \delta^* = \deg(G) - (2g-2) > 0$. Let $\mathcal{I} := \mathrm{supp}(\psi_P^*(\omega))$, and let $0 < P' := \sum_{i\in\mathcal{I}}(p_i) \leq P$; hence $\deg(P') = |\mathcal{I}| = \deg(G) - (2g-2)$. Then we have $\omega \in \Omega(G-P') \neq \{0\}$.

ii) Let $P' = \sum_{i\in\mathcal{I}}(p_i) \in \mathrm{Div}_{\mathbf{V}}$ be as asserted, for some $\mathcal{I} \subseteq \{1,\ldots,n\}$ such that $|\mathcal{I}| = \deg(G) - (2g-2)$, and let $0 \neq \omega \in \Omega(G-P')$. Then we have $\mathrm{supp}(\psi_P^*(\omega)) \subseteq \mathcal{I}$, entailing $0 < \mathrm{wt}(\psi_P^*(\omega)) \leq \deg(G) - (2g-2) = \delta^*$. Thus we have $d^* \leq \delta^*$, while we have $d^* \geq \delta^*$ anyway. ♯

Corollary. Let $\deg(G) \geq 2g-1$, let $k^* \geq 1$, and let $d^* := d(\mathcal{G}_P^*(D))$.

a) If $\{f_1,\ldots,f_k\} \subseteq \mathscr{L}_{\mathbb{F}_q}((\omega_{\mathbb{F}_q}) + P - G)$ is an $\mathbb{F}_q$-basis, then $[\mathrm{res}_{p_i}(f_j\omega_{\mathbb{F}_q})]_{ji} \in \mathbb{F}_q^{k\times n}$ is a generator matrix of $\mathcal{G}_P^*(G)$.

b) We have $d^* - 1 \leq n - k^* \leq d^* - 1 + g$; thus $\mathcal{G}_P^*(G)$ is an MDS code for $g = 0$.

Proof. b) We have $d^* \geq \deg(G) - (2g-2)$, that is $1 + \deg(G) - g \leq d^* - 1 + g$, and $n - k^* \leq 1 + \deg(G) - g$. Hence by the Singleton bound we get $d^* - 1 \leq n - k^* \leq 1 + \deg(G) - g \leq d^* - 1 + g$. ♯

(20.5) Duality of geometric Goppa codes. a) Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$ defined over $\mathbb{F}_q$, let $\{p_1,\ldots,p_n\} \subseteq \mathbf{V}(\mathbb{F}_q)$ be pairwise distinct, where $n \in \mathbb{N}$, let $P := \sum_{i=1}^n (p_i) \in \mathrm{Div}_{\mathbf{V}}$, and let $G \in \mathrm{Div}_{\mathbf{V}}$ be an $\mathbb{F}_q$-rational divisor such that $\mathrm{supp}(G) \cap \{p_1,\ldots,p_n\} = \emptyset$. We show that geometric Goppa codes and dual geometric Goppa codes are indeed duals to each other:

Theorem. We have $\mathcal{G}_P^*(G) = \mathcal{G}_P(G)^\perp$.

Proof. i) We show that $\dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P(G)) + \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P^*(G)) = n$:

We have $k := \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P(G)) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) - \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G-P))$ and $k^* := \dim_{\overline{\mathbb{F}}}(\overline{\mathcal{G}}_P^*(G)) = \dim_{\overline{\mathbb{F}}}(\Omega(G-P)) - \dim_{\overline{\mathbb{F}}}(\Omega(G))$. Moreover, by the Riemann-Roch Theorem we have $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) - \dim_{\overline{\mathbb{F}}}(\Omega(G)) = 1 + \deg(G) - g$, and similarly $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(G-P)) - \dim_{\overline{\mathbb{F}}}(\Omega(G-P)) = 1 + (\deg(G) - n) - g$. This yields $k + k^* = (1 + \deg(G) - g) - (1 + \deg(G) - g - n) = n$. ♯

ii) Thus it suffices to show that $\overline{\mathcal{G}}_P^*(G) \leq \overline{\mathcal{G}}_P(G)^\perp$:

Let $f \in \mathscr{L}(G)$ and $\omega \in \Omega(G-P)$. Then we have to compute

$$S_{f,\omega} := \langle \psi_P(f), \psi_P^*(\omega)\rangle = \langle [f(p_i)]_i, [\mathrm{res}_{p_i}(\omega)]_i\rangle = \sum_{i=1}^n f(p_i)\cdot \mathrm{res}_{p_i}(\omega).$$

Since f is regular at p_i, we have $\frac{d_{p_i}(f\omega)}{d_{p_i}(t_{p_i})} = f(p_i)\cdot\frac{d_{p_i}(\omega)}{d_{p_i}(t_{p_i})}$, from which we infer $\mathrm{res}_{p_i}(f\omega) = f(p_i)\cdot\mathrm{res}_{p_i}(\omega)$. Thus we get $S_{f,\omega} = \sum_{i=1}^n \mathrm{res}_{p_i}(f\omega)$.

From $(f) \geq -G$ and $(\omega) \geq G - P$ we get $(f\omega) \geq -P$. Hence $f\omega \in \Omega_{\mathscr{K}_{\mathbf{V}}}$ is regular on $\mathbf{V} \setminus \{p_1,\ldots,p_n\}$. Hence by the Residue Theorem we get $S_{f,\omega} = 0$. ♯

b) We show that dual geometric Goppa codes are geometric Goppa codes:

Theorem. i) There is $\omega_P \in \Omega_{\mathscr{K}_{\mathbf{V},\mathbb{F}_q}}$ such that $\nu_{p_i}(\omega_P) = -1$ and $\operatorname{res}_{p_i}(\omega_P) = 1$. **ii)** Then we have $\mathcal{G}_P^*(G) = \mathcal{G}_P((\omega_P) + P - G)$.

Proof. i) Let $0 \neq \omega \in \Omega_{\mathscr{K}_{\mathbf{V},\mathbb{F}_q}}$. Then by (19.4) (and its proof) there is $f \in \mathscr{K}_{\mathbf{V},\mathbb{F}_q}$ such that $\nu_{p_i}(f) = -\nu_{p_i}(\omega) - 1$, for $i \in \{1, \ldots, n\}$, entailing $\nu_{p_i}(f\omega) = -1$.

Let $a_i := \operatorname{res}_{p_i}(f\omega) \in \mathbb{F}_q^*$. Then there is $g \in \mathscr{K}_{\mathbf{V},\mathbb{F}_q}$ such that $\nu_{p_i}(g) = 0$ and $g(p_i) = \frac{1}{a_i} \in \mathbb{F}_q^*$, for $i \in \{1, \ldots, n\}$; recall that there is an affine open subset of $\mathbf{V}$ containing $\{p_1, \ldots, p_n\}$, so that we may find g using Hilbert's Nullstellensatz.

Hence we may let $\omega_P := gf\omega \in \Omega_{\mathscr{K}_{\mathbf{V},\mathbb{F}_q}}$.

ii) We have $\operatorname{supp}((\omega_P) + P) \cap \{p_1, \ldots, p_n\} = \emptyset = \operatorname{supp}(G) \cap \{p_1, \ldots, p_n\}$, and $(\omega_P) + P - G$ is $\mathbb{F}_q$-rational, so that $\mathcal{G}_P((\omega_P) + P - G)$ is well-defined.

There is an $\mathbb{F}_q$-isomorphism $\pi\colon \mathscr{L}_{\mathbb{F}_q}((\omega_P) + P - G) \to \Omega_{\mathbb{F}_q}(G - P)\colon f \mapsto f\omega_P$. Then for $f \in \mathscr{L}_{\mathbb{F}_q}((\omega_P) + P - G)$, we have $\nu_{p_i}(f) \geq 0$, that is f is regular at p_i, hence we get $\operatorname{res}_{p_i}(f\omega_P) = f(p_i) \cdot \operatorname{res}_{p_i}(\omega_P) = f(p_i)$, for $i \in \{1, \ldots, n\}$.

Thus, for the evaluation maps $\psi_P\colon \mathscr{L}_{\mathbb{F}_q}((\omega_P) + P - G) \to \mathcal{G}_P((\omega_P) + P - G)$ and $\psi_P^*\colon \Omega_{\mathbb{F}_q}(G - P) \to \mathcal{G}_P^*(G)$ we get $\psi_P = \psi_P^* \circ \pi$, showing the assertion. ♯

Corollary. Let $\omega_P \in \Omega_{\mathscr{K}_{\mathbf{V}}}$ such that $(\omega_P) = 2G - P$ and $\operatorname{res}_{p_i}(\omega_P) = 1$, for $i \in \{1, \ldots, n\}$. Then $\mathcal{G}_P(G)$ is self-dual.

Proof. We observe that $\nu_{p_i}(\omega_P) = \nu_{p_i}(2G - P) = -1$, for $i \in \{1, \ldots, n\}$, hence we have $\mathcal{G}_P(G)^\perp = \mathcal{G}_P^*(G) = \mathcal{G}_P((\omega_P) + P - G) = \mathcal{G}_P(G)$. ♯

21 Rational geometric Goppa codes

(21.1) Rational geometric Goppa codes. Let $\mathbf{V} := \mathbf{P}^1$, which has genus 0, and is defined over any finite field $\mathbb{F}_q \subseteq \overline{\mathbb{F}}$; then we have $\mathscr{K}_{\mathbf{P}^1,\mathbb{F}_q} \cong \mathbb{F}_q(X)$.

Let $\{p_1, \ldots, p_n\} \subseteq \mathbf{P}^1(\mathbb{F}_q) = \mathbb{F}_q \,\dot\cup\, \{\infty\}$ be pairwise distinct, where $n \in \mathbb{N}$; hence we have $n \leq q + 1$. Let $P := \sum_{i=1}^n (p_i) \in \operatorname{Div}_{\mathbf{P}^1}$, and let $G \in \operatorname{Div}_{\mathbf{P}^1}$ be an $\mathbb{F}_q$-rational divisor such that $\operatorname{supp}(G) \cap \{p_1, \ldots, p_n\} = \emptyset$. (Note that if $n = q+1$, then G is supported on non-rational points of $\mathbf{P}^1$ only.) Then $\mathcal{G}_P(G) \leq \mathbb{F}_q^n$ is called a **rational** geometric Goppa code.

Corollary. i) Let $k := \dim_{\mathbb{F}_q}(\mathcal{G}_P(G))$. Then we have $k = 0$, if $\deg(G) < 0$; and $k = 1 + \deg(G)$, if $\deg(G) \in \{0, \ldots, n-1\}$; and $k = n$, if $\deg(G) \geq n$.

ii) If $\deg(G) \in \{0, \ldots, n-1\}$, then we have $d(\mathcal{G}_P(G)) = n - \deg(G)$, that is $\mathcal{G}_P(G)$ is an MDS code.

iii) The dual code $\mathcal{G}_P(G)^\perp$ is a rational geometric Goppa code as well.

Proof. This mostly follows from (20.2) and (20.5); except that for $\deg(G) < 0$ we have $\mathscr{L}(G) = \{0\}$ and thus $\mathcal{G}_P(G)) = \{0\}$; while for $\deg(G) \geq n$, since any divisor of degree $\geq 2g - 1 = -1$ is non-special, we get $k = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) - \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G - P)) = (1 + \deg(G)) - (1 + \deg(G) - n) = n$. ♯

(21.2) Theorem: Rational geometric Goppa codes are GRS codes. We keep the notation of (21.1). Then there are pairwise distinct $\alpha := [\alpha_1, \ldots, \alpha_n] \subseteq \mathbb{F}_q$, and a vector $v := [v_1, \ldots, v_n] \in (\mathbb{F}_q^*)^n$, such that we have; see (17.1):

a) If $n \leq q$, then $\mathcal{G}_P(G) = \mathrm{GRS}_k(\alpha, v)$ (thus having generator matrix $G_k(\alpha, v)$).

b) If $n = q + 1$, then $\mathcal{G}_P(G)$ has generator matrix

$$\left[\begin{array}{c|c} & 0 \\ G_k(\alpha, v) & \vdots \\ & 0 \\ & 1 \end{array}\right] \in \mathbb{F}_q^{k \times (q+1)}.$$

Proof. We may assume that $k = \dim_{\overline{\mathbb{F}}}(\mathcal{G}_P(G)) \notin \{0, n\}$. Thus we have $\deg(G) = k - 1 \geq 0$, and, recalling that any divisor of degree ≥ -1 is non-special, we have $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(G)) = 1 + \deg(G) = k$.

a) Let $n \leq q$, let $p_0 \in \mathbf{P}^1(\mathbb{F}_q) \setminus \{p_1, \ldots, p_n\}$, and let $D := (p_0) - (p_1) \in \mathrm{Div}_{\mathbf{P}^1}$. Thus D is $\mathbb{F}_q$-rational and $\deg(D) = 0$, hence we get $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1$.

Let $0 \neq z \in \mathscr{L}_{\mathbb{F}_q}(D)$. Then we have $(z) + D \geq 0$ and $\deg((z) + D) = 0$, implying $(z) + D = 0$. Thus we have $(p_1) - (p_0) = -D = (z)$, hence $(z)_0 = (p_1)$ and $(z)_\infty = (p_0)$. We conclude that $\deg(z) = \deg((z)_0) = \deg((z)_\infty) = 1$, that is $\mathbb{F}_q(z) = \mathbb{F}_q(X)$, or equivalently $z \colon \mathbf{P}^1 \to \mathbf{P}^1$ is an isomorphism; in particular $\{z^j \in \mathbb{F}_q(X); j \in \mathbb{N}_0\}$ is $\mathbb{F}_q$-linearly independent.

Now we consider $C := G - (k - 1) \cdot (p_0) \in \mathrm{Div}_{\mathbf{P}^1}$. Hence C is $\mathbb{F}_q$-rational and $\deg(C) = 0$, thus $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(C)) = 1$. Let $0 \neq g \in \mathscr{L}_{\mathbb{F}_q}(C)$. We get $(gz^j) = (g) + j \cdot (z) \geq -G + (k-1-j) \cdot (p_0) + j \cdot (p_1) \geq -G \in \mathrm{Div}_{\mathbf{P}^1}$, for $j \in \{0, \ldots, k-1\}$. Hence we have $\{gz^j \in \mathscr{K}_{\mathbf{V}, \mathbb{F}_q}; j \in \{0, \ldots, k-1\}\} \subseteq \mathscr{L}_{\mathbb{F}_q}(G)$, being $\mathbb{F}_q$-linearly independent, and thus being an $\mathbb{F}_q$-basis. This shows that

$$\mathscr{L}_{\mathbb{F}_q}(G) = \{g \cdot f(z) \in \mathscr{K}_{\mathbf{V}, \mathbb{F}_q}; f \in \mathbb{F}_q[X]_{<k}\}.$$

Applying the evaluation map ψ_P yields

$$\mathcal{G}_P(G) = \{[g(p_i) \cdot f(z(p_i))]_i \in \mathbb{F}_q^n; f \in \mathbb{F}_q[X]_{<k}\}.$$

Since z is regular at p_i, we may let $\alpha := [z(p_i)]_i \in \mathbb{F}_q^n$, where since z is bijective the $z(p_i) \in \mathbb{F}_q$ are pairwise distinct. Moreover, since $\mathrm{supp}(G) \cap \{p_1, \ldots, p_n\} = \emptyset$, we infer that g is regular at p_i, so that we may let $v := [g(p_i)]_i \in \mathbb{F}_q^n$.

Assume that $g(p_i) = 0$, for some $i \in \{1, \ldots, n\}$; then we have $g \in \mathscr{L}(C - (p_i))$, but since $\deg(C - (p_i)) = -1$ we have $\mathscr{L}(C - (p_i)) = \{0\}$, a contradiction. Hence $g(p_i) \neq 0$, for $i \in \{1, \ldots, n\}$, so that $\mathcal{G}_P(G) = \mathrm{GRS}(\alpha, v)$.

b) Let $n = q+1$, and let $D := (p_{q+1}) - (p_1) \in \mathrm{Div}_{\mathbf{P}^1}$. (Now p_{q+1} plays the role of p_0 above, but otherwise we proceed similarly.) Hence we have $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1$. Letting $0 \neq z \in \mathscr{L}_{\mathbb{F}_q}(D)$, we get $(z)_0 = (p_1)$ and $(z)_\infty = (p_{q+1})$, so that $z\colon \mathbf{P}^1 \to \mathbf{P}^1$ is an isomorphism, and $\{z^j \in \mathbb{F}_q(X); j \in \mathbb{N}_0\}$ is $\mathbb{F}_q$-linearly independent.

Now we consider $C := G - (k-1)\cdot(p_{q+1}) \in \mathrm{Div}_{\mathbf{P}^1}$. We get $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(C)) = 1$. Letting $0 \neq g \in \mathscr{L}_{\mathbb{F}_q}(C)$, we have $(gz^j) \geq -G + (k-1-j)\cdot(p_{q+1}) + j\cdot(p_1) \geq -G \in \mathrm{Div}_{\mathbf{P}^1}$, for $j \in \{0,\ldots,k-1\}$, thus $\{gz^j \in \mathscr{K}_{\mathbf{V},\mathbb{F}_q}; j \in \{0,\ldots,k-1\}\} \subseteq \mathscr{L}_{\mathbb{F}_q}(G)$ is an $\mathbb{F}_q$-basis. Applying the evaluation map ψ_P yields (slightly differently)

$$\mathcal{G}_P(G) = \{[(g\cdot f(z))(p_i)]_i \in \mathbb{F}_q^n; f \in \mathbb{F}_q[X]_{<k}\}.$$

Since $\mathrm{supp}(G) \cap \{p_1,\ldots,p_n\} = \emptyset$, the rational function $g\cdot f(z)$ is regular at p_i. For $i \in \{1,\ldots,q\}$, both z and g are regular at p_i, thus we may let $\alpha := [z(p_1),\ldots,z(p_q)] \in \mathbb{F}_q^q$ and $v := [g(p_1),\ldots,g(p_q)] \in \mathbb{F}_q^q$, where the $z(p_i)$ are pairwise distinct, and $g(p_i) \neq 0$. For $i = q+1$ we get $\nu_{p_{q+1}}(gz^j) = k-1-j$, implying $(gz^j)(p_{q+1}) = 0$, for $j \in \{0,\ldots,k-2\}$, while $\gamma := (gz^{k-1})(p_{q+1}) \neq 0$. Replacing g by $\frac{1}{\gamma}\cdot g$, thus replacing v by $\frac{1}{\gamma}\cdot v$, yields the claimed matrix. ♯

Corollary: GRS codes are rational geometric Goppa codes. Any generalized Reed-Solomon code can be realized as a rational geometric Goppa code.

Proof. Let $\mathcal{G} = \mathrm{GRS}_k(\alpha, v) \leq \mathbb{F}_q^n$, for some $n \leq q$, where we may assume that $k := \dim_{\mathbb{F}_q}(\mathcal{G}) \notin \{0, n\}$, and where $\alpha = [\alpha_1,\ldots,\alpha_n] \subseteq \mathbb{F}_q$ are pairwise distinct, and $v = [v_1,\ldots,v_n] \in (\mathbb{F}_q^*)^n$.

Then let $P := \sum_{i=1}^n (\alpha_i) \in \mathrm{Div}_{\mathbf{P}^1}$. Letting $z := X \in \mathbb{F}_q[X]$, thus $(z) = (0) - (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$, we get $\alpha = [z(\alpha_i)]_i$. Moreover, let $g \in \mathbb{F}_q(X)$ such that g is regular at α_i and $g(\alpha_i) = v_i$, for $i \in \{1,\ldots,n\}$, thus we have $v = [g(\alpha_i)]_i$. (Note that by Lagrange interpolation we may even choose $g \in \mathbb{F}_q[X]$.) Letting $G := (k-1)\cdot(\infty) - (g) \in \mathrm{Div}_{\mathbf{P}^1}$, the above proof shows that $\mathcal{G} = \mathcal{G}_P(G)$. ♯

(21.3) Rational dual geometric Goppa codes. Let $\{p_1,\ldots,p_n\} \subseteq \mathbb{F}_q \subseteq \mathbf{P}^1(\mathbb{F}_q)$ be pairwise distinct, where $n \in \mathbb{N}$, let $P := \sum_{i=1}^n (p_i) \in \mathrm{Div}_{\mathbf{P}^1}$, let $G \in \mathrm{Div}_{\mathbf{P}^1}$ be $\mathbb{F}_q$-rational such that $\mathrm{supp}(G) \cap \{p_1,\ldots,p_n\} = \emptyset$, and let $\mathcal{G}_P^*(G) \leq \mathbb{F}_q^n$ be the associated **rational** dual geometric Goppa code.

We write $\mathcal{G}_P^*(G)$ as a rational geometric Goppa code:

Let $\omega \in \Omega_{\mathscr{K}_{\mathbf{P}^1},\mathbb{F}_q}$ such that $(\omega) = -2\cdot(\infty) \in \mathrm{Div}_{\mathbf{P}^1}$ and $\mathrm{res}_\infty(\frac{1}{X}\cdot\omega) = -1$; see (19.10). Moreover, let $h := \prod_{i=1}^n (X - p_i) \in \mathbb{F}_q[X]$; then for the **logarithmic derivative** of h we have $\frac{\partial_X h}{h} = \sum_{i=1}^n \frac{1}{X-p_i} \in \mathbb{F}_q(X)$. Finally, let

$$\omega_P := \frac{\partial_X h}{h}\cdot\omega = \Big(\sum_{i=1}^n \frac{1}{X-p_i}\Big)\cdot\omega \in \Omega_{\mathscr{K}_{\mathbf{P}^1},\mathbb{F}_q}.$$

We show that ω_P fulfills the assertions in (20.5):

i) We have $\nu_{p_i}(\frac{1}{X-p_i}) = -1$, and $\nu_{p_i}(\frac{1}{X-p_j}) = 0$ for $j \neq i$. Thus by the strong triangle inequality we get $\nu_{p_i}(\frac{\partial_X h}{h}) = -1$, implying $\nu_{p_i}(\omega_P) = -1$.

ii) We have $\nu_{p_i}(\sum_{j\neq i} \frac{1}{X-p_j}) \geq 0$, hence we get $\mathrm{res}_{p_i}(\omega_P) = \mathrm{res}_{p_i}(\frac{\partial_X h}{h} \cdot \omega) = \mathrm{res}_{p_i}(\frac{1}{X-p_i} \cdot \omega)$. Since $\{p_i, \infty\}$ are the only poles of $\frac{1}{X-p_i} \cdot \omega$, by the Residue Theorem (or by direct calculation as in (19.10)) we get $\mathrm{res}_{p_i}(\frac{1}{X-p_i} \cdot \omega) = 1$. ♯

Thus we have $\mathcal{G}_P^*(G) = \mathcal{G}_P((\omega_P) + P - G)$; note that

$$(\omega_P) + P = (\partial_X h) + (n-2) \cdot (\infty) = (\partial_X h)_0 + (n - 2 - \deg(\partial_X h)) \cdot (\infty) \in \mathrm{Div}_{\mathbf{P}^1}.$$

(21.4) Subfield subcodes of rational geometric Goppa codes. We show that any alternant code over $\mathbb{F}_q$ can be obtained as a subfield subcode of a rational geometric Goppa code defined over a finite extension field of $\mathbb{F}_q$:

Let $\mathbb{F}_q \subseteq \mathbb{F}$ be a finite field extension of degree $f \in \mathbb{N}$, let $n \in \mathbb{N}$, let $k \in \{1, \ldots, n-1\}$, let $\alpha := [\alpha_1, \ldots, \alpha_n]$ be pairwise distinct places of $\mathbb{F}$, let $v := [v_1, \ldots, v_n] \in (\mathbb{F}^*)^n$, and let $\mathcal{A}_k(\alpha, v) := \mathrm{GRS}_k(\alpha, v)^\perp \cap \mathbb{F}_q^n \leq \mathbb{F}_q^n$ be the associated alternant code. (We may safely leave out the cases $k \in \{0, n\}$).

Then by (21.1) we have $\mathrm{GRS}_k(\alpha, v) = \mathcal{G}_P(G) \leq \mathbb{F}^n$, where $P := \sum_{i=1}^n (\alpha_i) \in \mathrm{Div}_{\mathbf{P}^1}$, and $G := (k-1) \cdot (\infty) - (g) \in \mathrm{Div}_{\mathbf{P}^1}$, where $g \in \mathbb{F}(X)$ such that g is regular at α_i and $g(\alpha_i) = v_i$, for $i \in \{1, \ldots, n\}$. Hence we have

$$\mathrm{GRS}_k(\alpha, v)^\perp = \mathcal{G}_P(G)^\perp = \mathcal{G}_P^*(G) = \mathcal{G}_P((\omega_P) + P - (k-1) \cdot (\infty) + (g)) \leq \mathbb{F}^n.$$

In particular, from $d(\mathcal{G}_P^*(G)) = n - \deg((\omega_P) + P - (k-1) \cdot (\infty) + (g)) = k + 1$ we recover the Helgert bound $d(\mathcal{A}_k(\alpha, v)) \geq k + 1$. ♯

i) In particular, let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q, n) = 1$, be a BCH code of designed distance $\delta \in \{2, \ldots, n\}$ (where we again leave out the cases $\delta \in \{1, n+1\}$), and consecutive defining set $\mathcal{V} := \{\zeta_n^a, \ldots, \zeta_n^{a+\delta-2}\} \subseteq \mathbb{F}_q(\zeta_n) =: \mathbb{F}$ of length $\delta - 1$, where $a \in \mathbb{Z}_n$. Then, letting $\alpha := [\zeta_n^{i-1}]_i$ and $v := [\zeta_n^{a(i-1)}]_i$, we have

$$\mathcal{C} = \mathcal{A}_{\delta-1}(\alpha, v) = \mathrm{GRS}_{\delta-1}(\alpha, v)^\perp \cap \mathbb{F}_q^n = \mathcal{G}_P(G^*) \cap \mathbb{F}_q^n,$$

where $P := \sum_{i=1}^n (\zeta_n^{i-1}) \in \mathrm{Div}_{\mathbf{P}^1}$ and $G^* := (\omega_P) + P - (\delta-2) \cdot (\infty) + (g) \in \mathrm{Div}_{\mathbf{P}^1}$, where in turn $g \in \mathbb{F}(X)$ such that $g(\zeta_n^{i-1}) = \zeta_n^{a(i-1)}$, for $i \in \{1, \ldots, n\}$.

Hence we may choose $g := X^a \in \mathbb{F}[X]$, so that $(g) = a \cdot (0) - a \cdot (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$. Moreover, we have $h := \prod_{i=1}^n (X - \zeta_n^{i-1}) = X^n - 1 \in \mathbb{F}[X]$, so that $\partial_X h = nX^{n-1} \neq 0$, entailing $(\omega_P) + P = (n-1) \cdot (0) - (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$. This yields

$$G^* = (n + a - 1) \cdot (0) - (\delta + a - 1) \cdot (\infty) \in \mathrm{Div}_{\mathbf{P}^1}.$$

From $d(\mathcal{G}_P(G^*)) = n - \deg(G^*) = \delta$ we recover the BCH bound $d(\mathcal{C}) \geq \delta$. ♯

ii) Likewise, let $0 \neq g \in \mathbb{F}_q[X]$ having degree $k := \deg(g) \in \mathbb{N}$, let $\mathbb{F}_q \subseteq \mathbb{F}$ be a finite field extension, let $\alpha := [\alpha_1, \ldots, \alpha_n]$ be pairwise distinct places of $\mathbb{F}$, where

$n > k$, such that $g(\alpha_i) \neq 0$, for $i \in \{1, \ldots, n\}$, and let $v := [\frac{1}{g(\alpha_i)}]_i \in (\mathbb{F}^*)^n$. Then the associated (classical) Goppa code is given as $\mathcal{G}(\alpha, g) = \mathcal{A}_k(\alpha, v) = \mathrm{GRS}_k(\alpha, v)^\perp \cap \mathbb{F}_q^n$. (We again leave out the cases $k = 0$ and $k \geq n$.)

We have $\mathrm{GRS}_k(\alpha, v)^\perp = \mathcal{G}_P(G^*) \leq \mathbb{F}^n$, where $P := \sum_{i=1}^n (\alpha_i) \in \mathrm{Div}_{\mathbf{P}^1}$ and $G^* = (\omega_P) + P - (k-1) \cdot (\infty) + (\frac{1}{g}) \in \mathrm{Div}_{\mathbf{P}^1}$. Letting $h := \prod_{i=1}^n (X - \alpha_i) \in \mathbb{F}[X]$ we have $(\omega_P) + P = (\partial_X h) + (n-2) \cdot (\infty)$, thus we get

$$G^* = (\partial_X h) + (n - k - 1) \cdot (\infty) - (g).$$

We recover the Goppa bound $d(\mathcal{G}(\alpha, g)) \geq d(\mathcal{G}_P(G^*)) = n - \deg(G^*) = k + 1$. ♯

22 Non-rational geometric Goppa codes

(22.1) Example: The cubic Fermat curve. a) We consider the plane projective curve $\mathbf{V} = \mathbf{V}(F)$ of degree 3 defined over any field $\mathbb{F}$ such that $\mathrm{char}(\mathbb{F}) \neq 3$ by the irreducible polynomial

$$F := X^3 + Y^3 + Z^3 \in \mathbb{F}[X, Y, Z].$$

(Over $\mathbb{F}_4 = \mathbb{F}_{r^2}$, where $r + 1 = 3$, this is the smallest case of a **Hermitian curve**.) Note that, since F is a symmetric polynomial, $\mathbf{V}$ has $\mathcal{S}_3$ as a group of automorphisms acting by coordinate permutations.

We determine the singular points of $\mathbf{V}$: We have $J_{X,Y,Z}(F) = [3X^2, 3Y^2, 3Z^2]$. Thus for $[x\colon y\colon z] \in \mathbf{P}^2$ we have $(J_{X,Y,Z}(F))(x, y, z) \neq 0$. Hence $\mathbf{V}$ is smooth.

Let $\zeta = \zeta_3 \in \overline{\mathbb{F}}$ be a primitive 3-rd root of unity. By Plücker's formula $\mathbf{V}$ has genus $g_{\mathbf{V}} = 1$; hence is an **elliptic curve**. Thus we have $\dim_{\overline{\mathbb{F}}}(\Omega_{\mathbf{V}}) = 1$, and an element $0 \neq \omega \in \Omega_{\mathbf{V}}$ is given as follows; see (19.9):

On D_Z we have $G(A_1, A_2) := A_1^3 + A_2^3 + 1$, where $A_1 := \frac{X}{Z}$ and $A_2 := \frac{Y}{Z}$, leading to the following forms on D_{ZX}, and by symmetry on D_{XY} and on D_{YZ}:

$$\begin{aligned} \omega_{XY} &:= (\tfrac{X}{Y})^2 \cdot d(\tfrac{Z}{X}) \in \Omega_{D_{XY}}, \\ \omega_{YZ} &:= (\tfrac{Y}{Z})^2 \cdot d(\tfrac{X}{Y}) \in \Omega_{D_{YZ}}, \\ \omega_{ZX} &:= (\tfrac{Z}{X})^2 \cdot d(\tfrac{Y}{Z}) \in \Omega_{D_{ZX}}; \end{aligned}$$

these forms coincide on D_{XYZ}, and $D_{XZ} \cup D_{XY} \cup D_{YZ} = \mathbf{V}$.

Hence we have $(\omega) = 0 \in \mathrm{Div}_{\mathbf{V}}$, but ω is not the differential of a regular (hence constant) function. Moreover, we observe that the set of canonical divisors coincides with the set of principal divisors, by $\mathscr{K}_{\mathbf{V}} \to \Omega_{\mathscr{K}_{\mathbf{V}}} \colon f \mapsto f\omega$.

b) i) For later use, we determine certain principal divisors: Note first that, since the canonical divisor $(X) \in \mathrm{Div}_{\mathbf{V}}$ has degree $d = 3$, for all linear homogeneous $f \in \mathscr{R}_{\mathbf{V}}$ we have $\deg((f)) = \deg((f)_0) = 3$. Now, for $r, s, t \in \mathbb{Z}$ let

$$f_{rst} := \frac{X^r Y^s Z^t}{(Y + Z)^{r+s+t}} \in \mathscr{K}_{\mathbf{V}}.$$

The zeroes of $Y+Z$ are given by $y=-z$, implying $x=0$; hence $(Y+Z)=3\cdot(p_0)$, where $p_0:=[0\colon 1\colon -1]$. Moreover, the zeroes of X are given by $x=0$, implying $y^3\neq -z^3$, that is $y=-\zeta^i z$, for $i\in\{0,\ldots,2\}$; thus $(X)=(p_0)+(p_1)+(p_2)$, where $p_i:=[0\colon 1\colon -\zeta^i]$. By symmetry, we get $(Y)=(p_0')+(p_1')+(p_2')$, where $p_i':=[1\colon 0\colon -\zeta^i]$, and $(Z)=(p_0'')+(p_1'')+(p_2'')$, where $p_i'':=[1\colon -\zeta^i\colon 0]$.

Thus we have $(f_{100})=(p_1)+(p_2)-2\cdot(p_0)$, while $(f_{010})=(p_0')+(p_1')+(p_2')-3\cdot(p_0)$ and $(f_{001})=(p_0'')+(p_1'')+(p_2'')-3\cdot(p_0)$; note that $(f_{rst})=(f_{100}^r f_{010}^s f_{001}^t)=r\cdot(f_{100})+s\cdot(f_{010})+t\cdot(f_{001})\in\mathrm{Div}_{\mathbf{V}}$.

ii) Next, again for later use, we determine the associated residues:

We have

$$\mathrm{res}_{p_i}(f_{rst}\cdot\omega)=\mathrm{res}_{p_i}(\frac{Y^{r+s+2}Z^{t-2}}{(Y+Z)^{r+s+t}}\cdot\frac{X^r}{Y^r}\cdot d(\frac{X}{Y})),$$

hence $\mathrm{res}_{p_i}(f_{rst}\cdot\omega)\neq 0$ if and only if $r=-1$, where $\mathrm{res}_{p_i}(f_{-1,s,t}\cdot\omega)=\zeta^{i(s+2t)}$.

Similarly, we have

$$\mathrm{res}_{p_i'}(f_{rst}\cdot\omega)=\mathrm{res}_{p_i'}(\frac{X^{r-2}Z^{s+t+2}}{(Y+Z)^{r+s+t}}\cdot\frac{Y^s}{Z^s}\cdot d(\frac{Y}{Z})),$$

hence $\mathrm{res}_{p_i'}(f_{rst}\cdot\omega)\neq 0$ if only if $s=-1$, where $\mathrm{res}_{p_i'}(f_{r,-1,t}\cdot\omega)=\zeta^{i(2-r)}$.

Finally, we have

$$\mathrm{res}_{p_i''}(f_{rst}\cdot\omega)=\mathrm{res}_{p_i''}(\frac{X^{r+t+2}Y^{s-2}}{(Y+Z)^{r+s+t}}\cdot\frac{Z^t}{X^t}\cdot d(\frac{Z}{X})),$$

hence $\mathrm{res}_{p_i''}(f_{r,s,-1}\cdot\omega)\neq 0$ if and only if $t=-1$, where $\mathrm{res}_{p_i''}(f_{r,s,-1}\cdot\omega)=\zeta^{i(2-r)}$.

c) We determine the $\mathbb{F}$-rational points of $\mathbf{V}$ for $|\mathbb{F}|\in\{2,4\}$; by symmetry, it suffices to consider points $[x\colon y\colon 1]\in\mathbf{V}$, for which $F(x,y,1)=x^3+y^3+1=0$.

i) Over $\mathbb{F}_2$ we get $x+y=1$, so that $\mathbf{V}(2)=\{p_0,p_0',p_0''\}$.

ii) Over $\mathbb{F}_4=\mathbb{F}_2(\zeta)$, since $\mathbb{F}_2$ is the set of cubes in $\mathbb{F}_4$, we get $\mathbf{V}(4)\setminus\mathbf{V}(2)=\{p_1,p_2,p_1',p_2',p_1'',p_2''\}$; the latter consisting of three $\langle\varphi_2\rangle$-orbits of length 2.

It can shown that $|\mathbf{V}(2^r)|=2^r+1=|\mathbf{P}^1(2^r)|$ if r is odd, while we have $|\mathbf{V}(2^{2s})|-(2^{2s}+1)=(-2)^{s+1}$ if $r=2s$ is even; comparing with the **Hasse-Weil bound**, saying that $\big||\mathbf{V}(q)|-|\mathbf{P}^1(q)|\big|\leq 2g_{\mathbf{V}}\cdot\sqrt{q}$ for any finite field $\mathbb{F}_q$, shows that in the latter cases this bound is actually reached. Here are a few specific cardinalities (which have been computed straightforwardly using GAP):

r	1	2	3	4	5	6	7	8	9	10	11	12
2^r+1	3	5	9	17	33	65	129	257	513	1025	2049	4097
$\lvert\mathbf{V}(2^r)\rvert$		9		9		81		225		1089		3969

We proceed to construct some (dual) geometric Goppa codes over $\mathbb{F}_4$: To do so, let $P:=\sum_{p\in\mathbf{V}(4)\setminus\{p_0\}}(p)\in\mathrm{Div}_{\mathbf{V}}$; hence $n:=\deg(P)=|\mathbf{V}(4)|-1=8$.

Table 11: Goppa codes from the cubic Fermat curve.

k	f	$\nu_{p_0}(f)$	p_1	p_2	p'_0	p'_1	p'_2	p''_0	p''_1	p''_2	d
1	f_{000}	0	1	1	1	1	1	1	1	1	8
2	f_{100}	-2	.	.	1	ζ^2	ζ	1	ζ^2	ζ	6
3	f_{010}	-3	ζ	ζ^2	.	.	.	1	1	1	5
4	f_{200}	-4	.	.	1	ζ	ζ^2	1	ζ	ζ^2	4
5	f_{110}	-5	.	.	.	.	.	1	ζ^2	ζ	3
6	f_{011}	-6	1	1	.	.	.	.	.	.	2
7	f_{210}	-7	.	.	.	.	.	1	ζ	ζ^2	2

We enumerate the points of $\mathbf{V}(4)$ as $[p_0; p_1, p_2, p'_0, p'_1, p'_2, p''_0, p''_1, p''_2]$. Let $G_k := k \cdot (p_0) \in \mathrm{Div}_{\mathbf{V}}$, where $k \in \{1, \ldots, 7\} = \{2g_{\mathbf{V}} - 1, \ldots, n-1\}$; recall that p_0 is $\mathbb{F}_2$-rational. (Geometric and dual geometric Goppa codes associated with a divisor being a multiple of a prime divisor are called **one-point codes**.)

d) Let first $\mathcal{G}_k := \mathcal{G}_P(G_k) \leq \mathbb{F}_4^8$. Then we have $\dim_{\mathbb{F}_4}(\mathcal{G}_k) = \dim_{\mathbb{F}_4}(\mathscr{L}_{\mathbb{F}_4}(G_k)) = 1 + k - g_{\mathbf{V}} = k$. We determine the space of functions $\mathscr{L}(G_k) \subseteq \mathscr{K}_{\mathbf{V}}$:

We have $f_{rst} \in \mathscr{L}(G_k)$ if and only if $r, s, t \geq 0$ and $2r + 3s + 3t \leq k$. Since for any $k \in \{0, 2, \ldots, 7\}$ there are (not necessarily unique) $r, s, t \in \mathbb{N}_0$ such that $2r + 3s + 3t = k$, we get

$$\mathscr{L}(G_k) = \langle f_{rst} \in \mathscr{K}_{\mathbf{V}}; r, s, t \in \mathbb{N}_0, 2r + 3s + 3t \leq k \rangle_{\mathbb{F}}.$$

Our choice of functions is given in Table 11, where evaluation at the various points yields the rows depicted, so that the first k rows constitute a generating matrix of $\mathcal{G}_k$. In particular, $\mathcal{G}_1$ is the repetition code, $\mathcal{G}_2$ is the two-fold repetition of a $[4, 2, 3]$-code, and $\mathcal{G}_7 = \langle 1_8 \rangle_{\mathbb{F}_4}^{\perp}$.

The designed distance of $\mathcal{G}_k$ equals $\delta = n - \deg(G_k) = 8 - k$, where for the true minimum distance $d = d(\mathcal{G}_k)$ we have $d \geq \delta$. By inspection we get equality, except for $k \in \{1, 7\}$; note that $1_8 \in \mathcal{G}_7^{\perp}$, so that $d(\mathcal{G}_7) > 1$. In particular, $\mathcal{G}_k$ is not an MDS code, except for $k \in \{1, 7\}$.

e) Let now $\mathcal{G}_k^* := \mathcal{G}_P^*(G_k) \leq \mathbb{F}_4^8$. We have $\dim_{\mathbb{F}_4}(\mathcal{G}_k^*) = \dim_{\mathbb{F}_4}(\Omega_{\mathbb{F}_4}(G_k - P)) = \dim_{\mathbb{F}_4}(\mathscr{L}_{\mathbb{F}_4}((\omega) + P - G_k)) = \dim_{\mathbb{F}_4}(\mathscr{L}_{\mathbb{F}_4}(P - G_k)) = \deg(P - G_k) = 8 - k = k^*$, where we have the $\mathbb{F}_4$-isomorphism $\mathscr{L}_{\mathbb{F}_4}(P - G_k) \to \Omega_{\mathbb{F}_4}(G_k - P)\colon f \mapsto f\omega$.

We determine the space of functions $\mathscr{L}(P - G_k)$: We have $f_{rst} \in \mathscr{L}(P - G_k)$ if and only if $r, s, t \geq -1$ and $2r + 3s + 3t \leq -k$. For any $l \in \{1, \ldots, 6, 8\}$ there are (not necessarily unique) $r, s, t \geq -1$ such that $2r + 3s + 3t = -l$, hence

$$\mathscr{L}(P - G_k) = \langle f_{r,s,t} \in \mathscr{K}_{\mathbf{V}}; r, s, t \geq -1, 2r + 3s + 3t \leq -k \rangle_{\mathbb{F}}.$$

Our choice of functions is given in Table 12, where taking residues of the associated differential forms at the various points yields the rows depicted, so that

Table 12: Dual Goppa codes from the cubic Fermat curve.

k	f	$\nu_{p_0}(f)$	p_1	p_2	p_0'	p_1'	p_2'	p_0''	p_1''	p_2''	d^*
7	$f_{-1,-1,-1}$	8	1	1	1	1	1	1	1	1	8
6	$f_{0,-1,-1}$	6	.	.	1	ζ^2	ζ	1	ζ^2	ζ	6
5	$f_{-1,-1,0}$	5	ζ^2	ζ	1	1	1	.	.	.	5
4	$f_{1,-1,-1}$	4	.	.	1	ζ	ζ^2	1	ζ	ζ^2	4
3	$f_{0,-1,0}$	3	.	.	1	ζ^2	ζ	.	.	.	3
2	$f_{-1,0,0}$	2	1	1	.	.	.	.	.	.	2
1	$f_{1,-1,0}$	1	.	.	1	ζ	ζ^2	.	.	.	2

the first $8-k$ rows constitute a generating matrix of $\mathcal{G}_k^*$. In particular, $\mathcal{G}_7^*$ is the repetition code, $\mathcal{G}_6^*$ is the two-fold repetition of a $[4,2,3]$-code, and $\mathcal{G}_1^* = \langle 1_8 \rangle_{\mathbb{F}_4}^{\perp}$.

The designed distance of $\mathcal{G}_k^*$ equals $\delta^* = \deg(G_k) = k$, where for the true minimum distance $d^* = d(\mathcal{G}_k^*)$ we have $d^* \geq \delta^*$. By inspection we get equality, except for $k \in \{1,7\}$; note that $1_8 \in (\mathcal{G}_7^*)^{\perp}$, so that $d(\mathcal{G}_7^*) > 1$. In particular, $\mathcal{G}_k^*$ is not an MDS code, except for $k \in \{1,7\}$.

Actually, the similarities between the codes $\mathcal{G}_k^*$ and the codes $\mathcal{G}_k$ turn out to be not surprising, as soon as we discuss aspects of duality; see (20.5) (and in particular the proof of part b)):

i) Firstly, let $G \in \mathbb{F}_4^{7\times 8}$ and $H \in \mathbb{F}_4^{7\times 8}$ be the matrices depicted in Table 11 and Table 12, respectively. Then we get

$$G \cdot H^{\mathrm{tr}} = \begin{bmatrix} . & . & . & . & . & . & . \\ . & . & . & . & . & . & 1 \\ . & . & . & . & . & 1 & . \\ . & . & . & . & 1 & . & . \\ . & . & . & 1 & . & . & . \\ . & . & 1 & . & . & . & . \\ . & 1 & . & . & . & . & . \end{bmatrix} \in \mathbb{F}_4^{7\times 7}.$$

This shows that indeed $\mathcal{G}_k^{\perp} = \mathcal{G}_k^* \leq \mathbb{F}_4^8$, for $k \in \{1,\ldots,7\}$.

ii) Secondly, we may rewrite the dual geometric Goppa code $\mathcal{G}_k^* = \mathcal{G}_P^*(G_k)$ as a geometric Goppa code $\mathcal{G}_P(\cdot)$ with respect to a suitable divisor:

To this end, let $f_P := f_{-1,-1,-1} = \frac{(Y+Z)^3}{XYZ} \in \mathscr{K}_{\mathbf{V}}$. Then we have $(f_P \cdot \omega) = (f_P) = -P + 8 \cdot (p_0) \in \mathrm{Div}_{\mathbf{V}}$, and we have already seen that $\mathrm{res}_p(f_P \cdot \omega) = 1$ for $p \in \mathbf{V}(4) \setminus \{p_0\}$. Thus we conclude that $\mathcal{G}_k^* = \mathcal{G}_P((f_P) + P - G_k) = \mathcal{G}_P((8-k)\cdot(p_0)) = \mathcal{G}_P(G_{8-k}) = \mathcal{G}_{8-k}$.

Hence in combination we actually get $\mathcal{G}_k^{\perp} = \mathcal{G}_k^* = \mathcal{G}_{8-k}$. In particular, $\mathcal{G}_4$ is a self-dual $[8,4,4]$-code over $\mathbb{F}_4$.

(22.2) Example: The Klein quartic. a) We consider the plane projective curve $\mathbf{V}$ of degree $d = 4$ defined over any field $\mathbb{F}$ by the irreducible polynomial

$$F := X^3Y + Y^3Z + Z^3X \in \mathbb{F}[X, Y, Z].$$

Since $F(Y, Z, X) = F(X, Y, Z)$, we conclude that $\mathbf{V}$ has an automorphism of order 3 given by cyclic coordinate permutation $\alpha\colon [x\colon y\colon z] \mapsto [y\colon z\colon x]$. Moreover, letting $\rho = \zeta_7 \in \overline{\mathbb{F}}$ be a primitive 7-th root of unity, chosen to have minimum polynomial $T^3 + T + 1 \in \mathbb{F}_2[T]$, we observe that $\beta\colon [x\colon y\colon z] \mapsto [x\colon \rho y\colon \rho^{-2}z]$ defines an automorphism of $\mathbf{V}$ of order 7: Indeed, we have $F(X, \rho Y, \rho^{-2}Z) = X^3(\rho Y) + (\rho Y)^3(\rho^{-2}Z) + (\rho^{-2}Z)^3X = \rho \cdot F(X, Y, Z)$.

Note that $\alpha \in \mathrm{PSL}_3(\mathbb{F}_2)$ and $\beta \in \mathrm{PGL}_3(\mathbb{F}_8)$. We have $[x\colon y\colon z] \cdot \alpha^{-1}\beta\alpha = [x\colon \rho^{-3}y\colon \rho^{-1}z] = [x\colon y\colon z] \cdot \beta^{-3}$, saying that α normalizes $\langle\beta\rangle$, entailing that $\langle\alpha, \beta\rangle \leq \mathrm{PGL}_3(\mathbb{F}_8)$ has order 21. It can be shown that the automorphism group of $\mathbf{V}$ is isomorphic to $\mathrm{PGL}_3(\mathbb{F}_2)$, which is a simple group of order 168.

We determine the α-fixed points: If $[x\colon y\colon z] = [y\colon z\colon x] \in \mathbf{P}^2$, then we have $xyz \neq 0$, and letting $\gamma := \frac{x}{y} = \frac{y}{z} = \frac{z}{x} \in \overline{\mathbb{F}}$ entails $x = \gamma y = \gamma^2 z = \gamma^3 x$. Hence, if $\mathrm{char}(\mathbb{F}) = 3$ then $[1\colon 1\colon 1] \in \mathbf{V}$ is the only α-fixed point; while if $\mathrm{char}(\mathbb{F}) \neq 3$ then letting $\zeta = \zeta_3 \in \overline{\mathbb{F}}$ be a primitive 3-rd root of unity, we conclude that the α-fixed points are given as $[\zeta^{2i}\colon \zeta^i\colon 1] \in \mathbf{V}$, for $i \in \{0, \ldots, 2\}$.

We determine the β-fixed points: If $[x\colon y\colon z] = [x\colon \rho y\colon \rho^{-2}z] \in \mathbf{P}^2$, then assuming $xy \neq 0$ or $xz \neq 0$ or $yz \neq 0$ yields a contradiction, so that we have $[x\colon y\colon z] \in \{p_0, p_1, p_2\}$, where $p_0 := [0\colon 0\colon 1]$, $p_1 := [1\colon 0\colon 0]$, $p_2 := [0\colon 1\colon 0]$.

b) We determine the singular points of $\mathbf{V}$: We have

$$J := J_{X,Y,Z}(F) = [Z^3 + 3X^2Y, X^3 + 3Y^2Z, Y^3 + 3Z^2X].$$

By applying α it suffices to consider $p := [x, y, 1] \in \mathbf{V} \cap \mathbf{A}^2$: From $F(p) = x^3y + y^3 + x = 0$ we get $x = y = 0$ or $xy \neq 0$. If p is a singular point, we have $J(p) = [1 + 3x^2y, x^3 + 3y^2, y^3 + 3x] = 0$, contradicting the first case, thus we have $xy \neq 0$, and moreover we infer that $\mathrm{char}(\mathbb{F}) \neq 3$.

This entails $y^3 = -3x$ and $x^2y = -\frac{1}{3}$, hence $0 = x^3y + y^3 + x = -\frac{1}{3}x - 3x + x = -\frac{7}{3}x$, thus $\mathrm{char}(\mathbb{F}) = 7$. This shows that $\mathbf{V}$ is smooth whenever $\mathrm{char}(\mathbb{F}) \neq 7$; in this case, by Plücker's formula, $\mathbf{V}$ has genus $g_{\mathbf{V}} = 3$.

If $\mathrm{char}(\mathbb{F}) = 7$, then we have $x = 2y^3$ and $y = \frac{2}{x^2}$, entailing $x = 2 \cdot (\frac{2}{x^2})^3 = \frac{2}{x^6}$, that is $x^7 = 2$, or equivalently $x = 2$, and thus $y = \frac{2}{2^2} = \frac{1}{2} = 4$. Indeed, for $p = [2, 4, 1]$ we have $F(p) = 0$ and $J(p) = 0$, so that $\mathbf{V}$ is not smooth, where $[1\colon 2\colon 4] = [2\colon 4\colon 1] = [4\colon 1\colon 2] \in \mathbf{V}$ is the only singular point.

c) For later use, we determine certain principal divisors: Note first that, since the divisor $(X) \in \mathrm{Div}_{\mathbf{V}}$ has degree $d = 4$, for all linear homogeneous $f \in \mathscr{R}_{\mathbf{V}}$ we have $\deg((f)) = \deg((f)_0) = 4$. Now, for $r, s \in \mathbb{Z}$ let

$$f_{rs} := \frac{X^rY^s}{Z^{r+s}} \in \mathscr{K}_{\mathbf{V}}.$$

In order to find $(f_{rs}) \in \mathrm{Div}_{\mathbf{V}}$, we only have to consider the points $\{p_0, p_1, p_2\}$: We have $J(0,0,1) = [1,0,0]$, implying that the tangent space at $p_0 \in \mathbf{V} \cap \mathbf{A}^2$ is given as the line $D_{[0,0]}(F(X,Y,1)) = X = 0$. Thus Y can be chosen as a local coordinate; see Zariski's Proposition in (18.4). This yields $\nu_{p_0}(Y) = 1$, and from $Y^3 = -X(X^2Y + 1)$ we infer $\nu_{p_0}(X) = 3$, while $\nu_{p_0}(Z) = 0$ anyway.

By applying α we get $\nu_{p_2}(X) = 1$, $\nu_{p_2}(Z) = 3$ and $\nu_{p_2}(Y) = 0$, as well as $\nu_{p_1}(Z) = 1$, $\nu_{p_1}(Y) = 3$ and $\nu_{p_1}(X) = 0$. Thus we get $(f_{10}) = 3 \cdot (p_0) - (p_1) - 2 \cdot (p_2) \in \mathrm{Div}_{\mathbf{V}}$ and $(f_{01}) = (p_0) + 2 \cdot (p_1) - 3 \cdot (p_2) \in \mathrm{Div}_{\mathbf{V}}$.

d) From now on we assume that $\mathrm{char}(\mathbb{F}) = 2$. We determine the $\mathbb{F}$-rational points of $\mathbf{V}$ for $|\mathbb{F}| \in \{2, 4, 8\}$:

i) We first consider $\mathbb{F}_4 = \mathbb{F}_2(\zeta)$: By applying α it suffices to consider $p := [x, y, 1] \in \mathbf{V} \cap \mathbf{A}^2$, for which we have $x^3y + y^3 + x = 0$, and thus $x = y = 0$ or $xy \neq 0$; moreover, in the second case we get $[x, y] = [\zeta^i, \zeta^{2i}]$. Thus we infer that $\mathbf{V}(2) = \{p_0, p_1, p_2\}$ and $\mathbf{V}(4) \setminus \mathbf{V}(2) = \{[\zeta\colon \zeta^2\colon 1], [\zeta^2\colon \zeta\colon 1]\}$; note that the former consists of the β-fixed points in $\mathbf{V}$, and the latter of the α-fixed points.

ii) We now consider $\mathbb{F}_8 = \mathbb{F}_2(\rho)$: For $p := [x\colon y\colon z] \in \mathbf{V}(8) \setminus \mathbf{V}(2)$ we have $xyz \neq 0$, hence we may assume that $x = 1$, and by applying β we may assume that $y = 1$ as well. Then we have $F(p) = 1 + z + z^3 = 0$, thus $z \in \{\rho, \rho^2, \rho^4\}$, being the $\langle\varphi_2\rangle$-orbit of ρ. Letting $p := [1\colon 1\colon \rho]$, we have $p \cdot \beta^4\alpha^2 = [1\colon 1\colon \rho^4]$ and $p \cdot \beta^5\alpha = [1\colon 1\colon \rho^2]$, so that $\mathbf{V}(8) \setminus \mathbf{V}(2) = p \cdot \langle\alpha, \beta\rangle$. Since neither α, being of order 3, nor β, being of order 7, has fixed points in $\mathbf{V}(8) \setminus \mathbf{V}(2)$, we conclude that the latter is a regular $\langle\alpha, \beta\rangle$-orbit, thus we have $|\mathbf{V}(8) \setminus \mathbf{V}(2)| = 21$.

It can shown that $|\mathbf{V}(2^r)| = 2^r + 1 = |\mathbf{P}^1(2^r)|$ if $3 \nmid r$, while $|\mathbf{V}(2^{3s})| - (2^{3s} + 1) = a_s$ if $r = 3s$, where $a_0 = -6$ and $a_1 = 15$, and recursively $8a_s + 5a_{s+1} + a_{s+2} = 0$; note that the **Hasse-Weil-Serre bound** (improving the Hasse-Weil bound) yields $\big||\mathbf{V}(2^r)| - |\mathbf{P}^1(2^r)|\big| \leq 3 \cdot \lfloor 2 \cdot \sqrt{2^r} \rfloor \leq 6 \cdot \sqrt{2^r}$. Here are a few specific cardinalities (which have been computed using GAP); in particular, the Hasse-Weil-Serre bound is reached for $r = 3$:

r	1	2	3	4	5	6	7	8	9	10	11	12
$2^r + 1$	3	5	9	17	33	65	129	257	513	1025	2049	4097
$3 \cdot \lfloor 2 \cdot \sqrt{2^r} \rfloor$			15			48			135			384
$\lvert\mathbf{V}(2^r)\rvert$	3	5	24	17	33	38	129	257	528	1025	2049	4238

e) We proceed to construct some geometric Goppa codes over $\mathbb{F}_8$: To do so, let $P := \sum_{p \in \mathbf{V}(8) \setminus \{p_0\}} (p) \in \mathrm{Div}_{\mathbf{V}}$; hence $n := \deg(P) = |\mathbf{V}(8)| - 1 = 23$; note that $|\mathbf{V}(8)|$ is close to the maximum allowed by the Hasse-Weil bound. Enumerating lexicographically, we get $\mathbf{V}(8) = [p_0; p_1, p_2, [1\colon 1\colon \rho], \ldots, [1\colon \rho^6\colon \rho^6]]$.

Let $G_k := k \cdot (p_0) \in \mathrm{Div}_{\mathbf{V}}$, where $k \in \{5, \ldots, 22\} = \{2g_{\mathbf{V}} - 1, \ldots, n - 1\}$; hence G_k is $\mathbb{F}_2$-rational. Finally, let $\mathcal{G}_k := \mathcal{G}_P(G_k)$; hence $\dim_{\mathbb{F}_8}(\mathcal{G}_k) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(G_k)) = k + 1 - g_{\mathbf{V}} = k - 2$, and $\mathcal{G}_k$ has designed distance $\delta = n - k = 23 - k$. We determine the space of functions $\mathscr{L}(G_k) \subseteq \mathscr{K}_{\mathbf{V}}$:

In view of $(f_{rs}) = (3r + s) \cdot (p_0) + (-r + 2s) \cdot (p_1) + (-2r - 3s) \cdot (p_2) \in \mathrm{Div}_{\mathbf{V}}$

Table 13: Goppa codes from the Klein curve.

k	f	$\mathbf{V}(8) \setminus \{p_0\}$																				
	$f_{0,0}$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	$f_{-1,0}$	0	0	ρ	ρ^2	ρ^4	1	ρ^2	ρ^6	1	ρ^4	ρ^5	ρ^2	ρ^3	ρ^5	1	ρ	ρ^3	ρ	ρ^5	ρ^6	ρ^3	ρ^4	ρ^6
5	$f_{-2,1}$	0	0	ρ	ρ^2	ρ^4	ρ	ρ^3	1	ρ^2	ρ^6	1	ρ^5	ρ^6	ρ	ρ^4	ρ^5	1	ρ^6	ρ^3	ρ^4	ρ^2	ρ^3	ρ^5
6	$f_{-2,0}$	0	0	ρ^2	ρ^4	ρ	1	ρ^4	ρ^5	1	ρ	ρ^3	ρ^4	ρ^6	ρ^3	1	ρ^2	ρ^6	ρ^2	ρ^3	ρ^5	ρ^6	ρ	ρ^5
7	$f_{-2,-1}$	1	0	ρ^3	ρ^6	ρ^5	ρ^6	ρ^5	ρ^3	ρ^5	ρ^3	ρ^6	ρ^3	ρ^6	ρ^5	ρ^3	ρ^6	ρ^5	ρ^5	ρ^3	ρ^6	ρ^3	ρ^6	ρ^5
8	$f_{-3,1}$	0	0	ρ^2	ρ^4	ρ	ρ	ρ^5	ρ^6	ρ^2	ρ^3	ρ^5	1	ρ^2	ρ^6	ρ^4	ρ^6	ρ^3	1	ρ	ρ^3	ρ^5	1	ρ^4
9	$f_{-3,0}$	0	0	ρ^3	ρ^6	ρ^5	1	ρ^6	ρ^4	1	ρ^5	ρ	ρ^6	ρ^2	ρ	1	ρ^3	ρ^2	ρ^3	ρ	ρ^4	ρ^2	ρ^5	ρ^4
10	$f_{-3,-1}$	0	0	ρ^4	ρ	ρ^2	ρ^6	1	ρ^2	ρ^5	1	ρ^4	ρ^5	ρ^2	ρ^3	ρ^3	1	ρ	ρ^6	ρ	ρ^5	ρ^6	ρ^3	ρ^4
11	$f_{-4,1}$	0	0	ρ^3	ρ^6	ρ^5	ρ	1	ρ^5	ρ^2	1	ρ^3	ρ^2	ρ^5	ρ^4	ρ^4	1	ρ^6	ρ	ρ^6	ρ^2	ρ	ρ^4	ρ^3
12	$f_{-4,0}$	0	0	ρ^4	ρ	ρ^2	1	ρ	ρ^3	1	ρ^2	ρ^6	ρ	ρ^5	ρ^6	1	ρ^4	ρ^5	ρ^4	ρ^6	ρ^3	ρ^5	ρ^2	ρ^3
13	$f_{-4,-1}$	0	0	ρ^5	ρ^3	ρ^6	ρ^6	ρ^2	ρ	ρ^5	ρ^4	ρ^2	1	ρ^5	ρ	ρ^3	ρ	ρ^4	1	ρ^6	ρ^4	ρ^2	1	ρ^3
14	$f_{-4,-2}$	1	0	ρ^6	ρ^5	ρ^3	ρ^5	ρ^3	ρ^6	ρ^3	ρ^6	ρ^5	ρ^6	ρ^5	ρ^3	ρ^6	ρ^5	ρ^3	ρ^3	ρ^6	ρ^5	ρ^6	ρ^5	ρ^3
15	$f_{-5,0}$	0	0	ρ^5	ρ^3	ρ^6	1	ρ^3	ρ^2	1	ρ^6	ρ^4	ρ^3	ρ	ρ^4	1	ρ^5	ρ	ρ^5	ρ^4	ρ^2	ρ	ρ^6	ρ^2
16	$f_{-5,-1}$	0	0	ρ^6	ρ^5	ρ^3	ρ^6	ρ^4	1	ρ^5	ρ	1	ρ^2	ρ	ρ^6	ρ^3	ρ^2	1	ρ	ρ^4	ρ^3	ρ^5	ρ^4	ρ^2
17	$f_{-5,-2}$	0	0	1	1	1	ρ^5	ρ^5	ρ^5	ρ^3	ρ^3	ρ^3	ρ	ρ	ρ	ρ^6	ρ^6	ρ^6	ρ^4	ρ^4	ρ^4	ρ^2	ρ^2	ρ^2
18	$f_{-6,0}$	0	0	ρ^6	ρ^5	ρ^3	1	ρ^5	ρ	1	ρ^3	ρ^2	ρ^5	ρ^4	ρ^2	1	ρ^6	ρ^4	ρ^6	ρ^2	ρ	ρ^4	ρ^3	ρ
19	$f_{-6,-1}$	0	0	1	1	1	ρ^6	ρ^6	ρ^6	ρ^5	ρ^5	ρ^5	ρ^4	ρ^4	ρ^4	ρ^3	ρ^3	ρ^3	ρ^2	ρ^2	ρ^2	ρ	ρ	ρ
20	$f_{-6,-2}$	0	0	ρ	ρ^2	ρ^4	ρ^5	1	ρ^4	ρ^3	1	ρ	ρ^3	ρ^4	ρ^6	ρ^6	1	ρ^2	ρ^5	ρ^2	ρ^3	ρ^5	ρ^6	ρ
21	$f_{-6,-3}$	1	0	ρ^2	ρ^4	ρ	ρ^4	ρ	ρ^2	ρ	ρ^2	ρ^4	ρ^2	ρ^4	ρ	ρ^2	ρ^4	ρ	ρ	ρ^2	ρ^4	ρ^2	ρ^4	ρ
22	$f_{-7,-1}$	0	0	ρ	ρ^2	ρ^4	ρ^6	ρ	ρ^5	ρ^5	ρ^2	ρ^3	ρ^6	1	ρ^2	ρ^3	ρ^4	ρ^6	ρ^3	1	ρ	ρ^4	ρ^5	1

we have $(f_{rs}) \in \mathscr{L}(G_k)$ if and only if $3r + s \geq -k$ and $2s \geq r$ and $2r + 3s \leq 0$, which is equivalent to $-\frac{2}{3}r \geq s \geq \max\{\frac{r}{2}, -3r-k\}$. Thus we have $0 \geq r \geq -\frac{3}{7}k$, leaving finitely many cases which can be checked explicitly (using GAP).

It turns out that the 20 functions given in Table 13 belong to $\mathscr{L}_{\mathbb{F}_2}(G_{22})$ and have pole order i at p_0, where $i \in [0, 3, 5, 6, \ldots, 22]$. Hence these functions are $\overline{\mathbb{F}}$-linearly independent, and the first $k-2$ of them are an $\overline{\mathbb{F}}$-basis of $\mathscr{L}(G_k)$.

Evaluating at the points of $\mathbf{V}(8) \setminus \{p_0\}$ yields the explicit vectors given in Table 13 (which is not too enlightening): Evaluating at the points of $\mathbf{V}(8) \setminus \mathbf{V}(2)$ is straightforward (using GAP), while at the points $\{p_1, p_2\}$ both numerator and denominator of f_{rs}, viewed as a rational function in $\mathbb{F}_2(X, Y, Z)$, vanish. But using the divisor of f_{rs}, and noting that by $\mathbb{F}_2$-rationality we have $f_{rs}(p_i) \in \mathbb{F}_2$ anyway, we may read off the latter values directly. (Using the fact that $\mathbb{F}_2 = \{0, 1\}$ is a really dirty trick.)

We determine the minimum distance $d = d(\mathcal{G}_k)$, where the designed distance and the Singleton bound yield $\delta = 23 - k \leq d \leq 23 - (k-2) + 1 = 26 - k$: By inspection (using GAP, looking for codewords of small weight for $k \leq 17$, and taking check matrices into account for $k \geq 17$), we find that $d = \delta = 23 - k$ for $k \in \{5, \ldots, 19, 21\}$, while $d = \delta + 1 = 24 - k$ for $k \in \{20, 22\}$. In particular, $\mathcal{G}_k$ is not an MDS code.

f) Geometric Goppa codes (defined over curves with small genus and 'many' points compared to the Hasse-Weil-Serre bound) have a tendency to provide fairly good codes over prime fields, by using concatenation, see Exercise (24.12):

Let $\mathcal{D}$ be the even-weight $[4, 3, 2]$-code over $\mathbb{F}_2$. Then the concatenation $\widetilde{\mathcal{G}}_k$ of $\mathcal{G}_k$ and $\mathcal{D}$ is a binary code of length $23 \cdot 4 = 92$ and dimension $\widetilde{k} := 3 \cdot (k-2)$.

Moreover, $\widetilde{\mathcal{G}}_k$ has minimum distance $\widetilde{d} \geq 2 \cdot (23 - k + \epsilon)$, where $\epsilon = 1$ if $k \in \{20, 22\}$, and $\epsilon = 0$ otherwise. By inspection (using GAP, again looking for codewords of small weight, and taking check matrices into account), we get equality throughout.

For comparison we also give the maximum minimum distance $\widehat{d}$ of the binary codes with given length and dimension currently known (which does not at all mean that the latter are optimal). As it turns out, for quite a few values of k the code $\widetilde{\mathcal{G}}_k$ is not too far apart from the best codes known, and for $k = 11$ it is even amongst them.

k	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
$\widetilde{k}$	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60
$\widetilde{d}$	36	34	32	30	28	26	24	22	20	18	16	14	12	10	8	8	4	4
$\widehat{d}$	42	40	36	32	30	28	24	24	22	20	20	18	18	14	14	12	12	10

(22.3) Geometric Goppa codes are good. We have already seen that (classical) Goppa codes are good, in the sense that they reach the asymptotic Gilbert-Varshamov bound, see (17.5). It turns out that geometric Goppa codes are even better. In order to give an indication why, we need some preparation:

a) For a fixed prime power q, and $g \geq 0$ let

$$N_q(g) := \max\{|\mathbf{V}(q)| \in \mathbb{N}_0; \mathbf{V} \text{ smooth projective curve over } \mathbb{F}_q \text{ of genus } g\};$$

thus for $g = 0$ we have $N_q(g) = q + 1$. By the Hasse-Weil-Serre bound we have $\big||\mathbf{V}(q)| - (q+1)\big| \leq g \cdot \lfloor 2 \cdot \sqrt{q} \rfloor$, so that $N_q(g) \in \mathbb{N}_0$ is well-defined. We let

$$\tau(q) := \limsup_{g \to \infty} \frac{N_q(g)}{g}.$$

Again by the Hasse-Weil-Serre bound we have $0 \leq \tau(q) \leq \lfloor 2\sqrt{q} \rfloor$.

Actually, it is known that $\tau(q) > 0$, and the **Drinfeld-Vladut bound** says that $\tau(q) \leq \sqrt{q} - 1$. In general, the value of $\tau(q)$ remains obscure, but if q is a square we have the following result (needing the theory of modular curves):

Theorem: Tsfasman–Vladut–Zink [1982]. We have $\tau(q^2) = q - 1$. ♯

Corollary. Let $0 \leq \delta \leq 1 - \frac{1}{q-1}$. Then we have $\kappa_{q^2}(\delta) + \delta \geq 1 - \frac{1}{q-1}$.

Proof. Let $[\mathbf{V}_1, \mathbf{V}_2, \ldots]$ be a series of smooth projective curves over $\mathbb{F}_{q^2}$, having genus g_i and $n_i := |\mathbf{V}_i(q^2)| - 1$ rational points, such that $\lim_{i\to\infty} n_i = \infty$ and $\lim_{i\to\infty} \frac{g_i}{n_i} = \frac{1}{\tau(q^2)} = \frac{1}{q-1}$. Moreover, let $r_i \leq n_i - 1$ such that $\lim_{i\to\infty} \frac{r_i}{n_i} = 1-\delta$.

Now let $p_i \in \mathbf{V}_i(q^2)$, let $P_i := \sum_{p\in\mathbf{V}_i(q^2)\setminus\{p_i\}}(p) \in \mathrm{Div}_{\mathbf{V}_i}$, let $G_i := r_i \cdot (p_i) \in \mathrm{Div}_{\mathbf{V}_i}$, and let $\mathcal{G}_i := \mathcal{G}_{P_i}(G_i) \leq \mathbb{F}_{q^2}^{n_i}$ be the associated geometric Goppa code. Then we have $k_i := \dim_{\mathbb{F}_{q^2}}(\mathcal{G}_i) \geq 1 + r_i - g_i$ and $d_i := d(\mathcal{G}_i) \geq n_i - r_i$.

This entails $\rho_i := \rho(\mathbf{V}_i) = \frac{k_i}{n_i} \geq \frac{r_i+1}{n_i} - \frac{g_i}{n_i}$ and $\delta_i := \delta(\mathbf{V}_i) = \frac{d_i}{n_i} \geq 1 - \frac{r_i}{n_i}$. We may assume that both $\rho := \lim_{i\to\infty} \rho_i$ and $\widetilde{\delta} := \lim_{i\to\infty} \delta_i$ exist. Then we have $\rho \geq (1-\delta) - \frac{1}{q-1}$ and $\widetilde{\delta} \geq \delta$, entailing $\kappa_{q^2}(\delta) \geq \kappa_{q^2}(\widetilde{\delta}) \geq \rho \geq 1 - \delta - \frac{1}{q-1}$. ♯

b) We compare the Tsfasman-Vladut-Zink bound $\kappa_{q^2}(\delta) \geq 1 - \frac{1}{q-1} - \delta$, for $0 \leq \delta \leq 1 - \frac{1}{q-1}$, with the Gilbert-Varshamov bound $\kappa_{q^2}(\delta) \geq 1 - H_{q^2}(\delta)$, where

$$H_{q^2}(\delta) = \delta \log_{q^2}(q^2-1) - \delta \log_{q^2}(\delta) - (1-\delta)\log_{q^2}(1-\delta)$$

is the q-ary entropy function. We aim at finding the even powers q of a prime such that $1 - \frac{1}{q-1} - \delta \geq 1 - H_{q^2}(\delta)$, that is $\delta + \frac{1}{q-1} \leq H_{q^2}(\delta)$, for some δ:

Since $H_q(\delta)$ is strictly concave, we first determine the argument δ_0 such that $H_{q^2}(\delta)$ has a tangent of slope 1, that is $(\frac{\partial}{\partial\delta} H_{q^2})(\delta_0) = 1$: From

$$\frac{\partial}{\partial\delta} H_{q^2}(\delta) = \log_{q^2}(q^2-1) + \log_{q^2}(\frac{1-\delta}{\delta})$$

we get $\log_{q^2}(\frac{q^2}{q^2-1}) = \log_{q^2}(\frac{1-\delta_0}{\delta_0})$, that is $\frac{q^2}{q^2-1} = \frac{1-\delta_0}{\delta_0}$, which is equivalent to $\delta_0 = \frac{q^2-1}{2q^2-1}$. Now we get $1 - \delta_0 = \frac{q^2}{2q^2-1}$, and thus

$$H_{q^2}(\delta_0) = \log_{q^2}(2q^2-1) - \frac{q^2}{2q^2-1} = \log_{q^2}(2q^2-1) + (\delta_0 - 1).$$

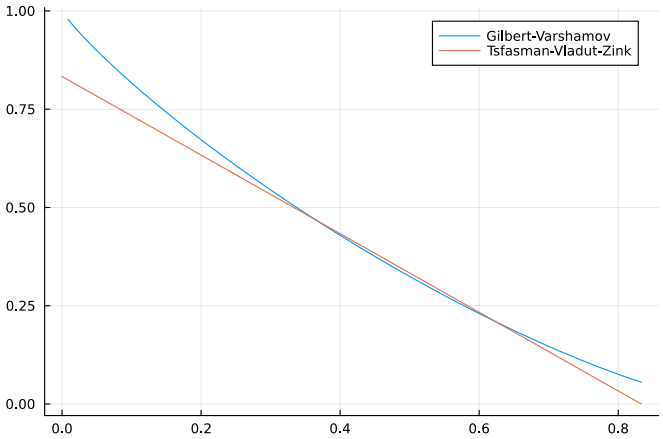
Hence the condition for some δ as desired to exist becomes

$$\delta_0 + \frac{1}{q-1} \leq H_{q^2}(\delta_0) = \log_{q^2}(2q^2-1) + (\delta_0 - 1),$$

that is $\frac{q}{q-1} \leq \log_{q^2}(2q^2-1)$, or equivalently $\frac{1}{q-1} \leq \log_{q^2}(2 - \frac{1}{q^2})$.

Since (deriving with respect to q shows that) the difference of the right and left hand sides is strictly increasing, and for $q_0 := \sqrt{43}$ we have $\frac{1}{q_0-1} < \log_{q_0^2}(2 - \frac{1}{q_0^2})$, we conclude that for all admissible $q \geq 49$ there is some (even a certain open interval of) δ such that $1 - \frac{1}{q-1} - \delta > 1 - H_{q^2}(\delta)$, that is the Tsfasman-Vladut-Zink bound strictly exceeds the Gilbert-Varshamov bound, see Table 14.

Table 14: Asymptotic bounds for $q = 49$.



VII

23 Exercises to Part I

(23.1) Exercise: Arithmetic parity check codes.
Let $\mathcal{C} := \{[x_1, \dots, x_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} i x_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$ be the code used for the ISBN-10. Does $\mathcal{C}$ detect twin errors and jump twin errors?

(23.2) Exercise: Geometric parity check codes.
For $a \in \mathbb{Z}_{11}$ let $\mathcal{C}_a := \{[x_1, \dots, x_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} a^i x_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$. When does $\mathcal{C}_a$ detect single errors, transposition errors, twin errors, jump twin errors?

(23.3) Exercise: Bank account codes.
a) For $x \in \mathbb{N}_0$ let $Q(x) \in \mathbb{N}_0$ be the **cross sum** of x with respect to its decimal representation. Considering $\mathbb{Z}_{10}$ as a subset of $\mathbb{N}_0$, show that $\mathbb{Z}_{10} \to \mathbb{Z}_{10} \colon x \mapsto Q(2x)$ is a bijection.
b) A typical parity check code used for bank account numbers is

$$\mathcal{C} := \{[x_1, \dots, x_{2n}] \in \mathbb{Z}_{10}^{2n}; \sum_{i=1}^{n} (Q(2x_{2i-1}) + x_{2i}) = 0 \in \mathbb{Z}_{10}\} \subseteq \mathbb{Z}_{10}^{2n}.$$

for $n \in \mathbb{N}$. Does $\mathcal{C}$ detect single errors and transposition errors?

(23.4) Exercise: International Bank Account Number.
Assume that in the valid IBAN '`DE`68 3905 0000 0123 4567 89' the leading digit

'0' of the bank account number is replaced by '9', yielding the BBAN

$$\text{'3905 0000 } \mathbf{9}\text{123 4567 89'}.$$

Which further single errors are there such that the resulting BBAN carrying two errors cannot be detected by the parity check symbols '`DE`68'?

(23.5) Exercise: Parity check codes over groups.
We consider a parity check code over a finite group G with respect to bijections $\pi_i\colon G \to G$, for $i \in \{1,\dots,n\}$ and $n \in \mathbb{N}$. Formulate conditions on the π_i such that twin errors and jump twin errors may be detected.

(23.6) Exercise: Parity check codes over dihedral groups.
For $n \geq 2$ let $D_{2n} := \langle a, b; a^n = 1, b^2 = 1, b^{-1}ab = a^{-1}\rangle$ be the dihedral groups of order $2n$, and let $\tau\colon D_{2n} \to D_{2n}$ be given by $a^i \mapsto a^{-i-1}$ and $a^ib \mapsto a^ib$, for $i \in \{0,\dots,n-1\}$. Show that the map τ is well-defined and bijective, and if n is odd then we have $gh^\tau \neq hg^\tau$, for $g \neq h \in D_{2n}$.

(23.7) Exercise: Almost complete maps.
Let $G = \mathbb{Z}_n$, where $n := 2^a$ for some $a \in \mathbb{N}$, and let $\sigma\colon G \to G$ be given by

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

where $m := \frac{n}{2}$. (Thus σ is bijective.) Show that the map $\tau := \sigma + \mathrm{id}_G\colon G \to G$ fulfills $\tau(G) = G \setminus \{m-1\}$ and $\tau^{-1}(\{n-1\}) = \{n-1, n-1-\lceil\frac{n}{4}\rceil\}$.

(23.8) Exercise: Source encoding.
Alice is able to send four messages, which are encoded into two bits as follows:

00: *'Stock prices are stable.'*	11: *'Shall we sell?'*
01: *'Stock prices are falling.'*	10: *'Shall we buy?'*

Bob receives the stream ...001100110011... What does Alice want to tell him?

(23.9) Exercise: Huffman encoding.
The following algorithm can be used for optimal binary source encoding:

Let $\mathcal{X} = \{x_1,\dots,x_q\}$ be an alphabet with probability distribution μ, and let $p_k := \mu(x_k)$ for $k \in \{1,\dots,q\}$. Let first $j \neq i \in \{1,\dots,q\}$ such that $p_i = \min\{p_1,\dots,p_q\}$ and $p_j = \min(\{p_1,\dots,p_q\} \setminus \{p_i\})$, and encode x_i and x_j into 0 and 1, respectively. Now consider the modified alphabet $\mathcal{X}' := (\{x_1,\dots,x_q\} \setminus \{x_i,x_j\}) \dot\cup \{x_{ij}\}$, with probability distribution μ', where $\mu'(x_k) := p_k$ for $k \in \{1,\dots,q\} \setminus \{i,j\}$, and $p_{ij} = \mu'(x_{ij}) := p_i + p_j$, and recurse to find prefixes to the codewords already found.

a) Show that this yields a prefix-free injective encoding $\gamma\colon \mathcal{X} \to (\mathbb{Z}_2)^* \setminus \{\epsilon\}$, being called the **Huffman encoding** of $\mathcal{X}$ with respect to μ.

Show that the Huffman encoding is optimal. (Hence we have $H(\gamma) \leq H(\mathcal{X})+1$.) Do we necessarily have $l(\gamma(x_i)) \leq 1 - \log_2(p_i)$ whenever $p_i > 0$?

b) Now let μ be the uniform distribution, and let $k := \lfloor \log_2(q) \rfloor \in \mathbb{N}$. Show that $\gamma(\mathcal{X}) \subseteq (\mathbb{Z}_2)^k \,\dot\cup\, (\mathbb{Z}_2)^{k+1}$, and that $\gamma(\mathcal{X})$ has average word length $k + 2 - \frac{2^{k+1}}{n}$. How does this compare to $H(\mathcal{X})$? What happens for $q = 2^k$?

c) Consider the alphabet $\mathbb{Z}_4$ with the following probability distribution:

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

Show that the associated Huffman encoding γ is as given above, and determine the average word length of $\gamma(\mathbb{Z}_4)$ and the average information content $H(\mathbb{Z}_4)$.

Moreover, consider the alphabet $\mathbb{Z}_4 \times \mathbb{Z}_4$, with independent entries, and determine its average information content, the associated Huffman encoding, and its average word length. What do you observe?

(23.10) Exercise: Huffman encoding.
a) Write GAP programs computing Huffman encoding and decoding of the Latin alphabet $\mathcal{X} := \{\mathtt{a}, \ldots, \mathtt{z}\} \,\dot\cup\, \{\_\}$ including a blank symbol, with respect to the following relative frequency of letters in English texts:

$\mathcal{X}$	μ
_	0.186
a	0.064
b	0.013
c	0.022
d	0.032
e	0.103
f	0.021
g	0.015
h	0.047

$\mathcal{X}$	μ
i	0.057
j	0.001
k	0.005
l	0.032
m	0.028
n	0.057
o	0.063
p	0.015
q	0.001

$\mathcal{X}$	μ
r	0.048
s	0.051
t	0.080
u	0.023
v	0.008
w	0.018
x	0.001
y	0.016
z	0.001

Determine the average information content of $\mathcal{X}$, and the average word length of the associated Huffman encoding.

b) Apply this to the following text consisting of 377 letters, where commas may

be ignored. How many bits does its Huffman encoding need?

```
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
```

c) Alternatively, apply **adaptive** Huffman encoding by using the relative frequency of letters in the above text as underlying probability distribution. How does this change the average information content of $\mathcal{X}$, the average word length of the Huffman encoding, and the bit length of the encoded text?

(23.11) Exercise: Huffman and Shannon-Fano encoding.
Determine the average information content, the Huffman encoding and the Shannon-Fano encoding, and their respective average word lengths, for the following probability distributions:

$$\frac{1}{100} \cdot [20, 18, 10, 10, 10, 6, 6, 4, 4, 4, 4, 3, 1]$$

$$\frac{1}{238} \cdot [2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41]$$

$$\frac{1}{385} \cdot [1, 4, 9, 16, 25, 36, 49, 64, 81, 100]$$

(23.12) Exercise: Symmetric binary channel.
Determine the maximal capacity of the symmetric binary channel having error probability $\frac{1}{2} \leq p \leq 1$.

(23.13) Exercise: Maximum likelihood decoding.
We consider the symmetric binary channel having error probability $0 \leq p < \frac{1}{2}$; typically we let $p = 10^{-e}$ for $e \in \{1, 2, 3\}$. Applying ML decoding, the channel is used to transmit words in $\mathbb{F}_2^3$, which is assumed to carry the uniform distribution.

a) Determine the error probability $\epsilon(\mathbb{F}_2^3)$, if the words are sent without redundancy, that is with information rate $\rho(\mathbb{F}_2^3) = 1$.

b) Now we use the code $\mathcal{C}_0 \leq \mathbb{F}_2^6$ given by the generator matrix

$$G_0 := \begin{bmatrix} 1 & . & . & 1 & . & . \\ . & 1 & . & . & 1 & . \\ . & . & 1 & . & . & 1 \end{bmatrix} \in \mathbb{F}_2^{3\times 6};$$

hence we have $\rho(\mathcal{C}_0) = \frac{1}{2}$. Determine the error probability $\epsilon(\mathcal{C}_0)$.

c) Finally, we similarly use the code $\mathcal{C} \leq \mathbb{F}_2^6$ given by the generator matrix

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

hence we have $\rho(\mathcal{C}) = \frac{1}{2}$ again. Determine the error probability $\epsilon(\mathcal{C})$.

Hint. Consider the spheres $\mathcal{B}_r(c) \subseteq \mathbb{F}_2^6$, for $c \in \mathcal{C}$ and $r \in \{0, 1, 2\}$.

(23.14) Exercise: Maximum likelihood decoding.
We consider the symmetric binary channel having error probability $0 \leq p < \frac{1}{2}$. Applying ML decoding, the channel is used to transmit words in a $(7, 2^4, 3)$-Code $\mathcal{C}$, which is assumed to carry the uniform distribution. Determine the error probability $\epsilon(\mathcal{C})$.

24 Exercises to Part II

(24.1) Exercise: Hamming distance.
a) Let $n \in \mathbb{N}$, let $v, w \in \mathbb{F}_2^n$ such that $d := d(v, w) \in \mathbb{N}_0$, for $r, s \in \mathbb{N}_0$ let

$$\mathcal{A} := \{u \in \mathbb{F}_2^n; d(v, u) = r,\ d(w, u) = s\},$$

and let $t := \frac{d+r-s}{2}$. Show that $\mathcal{A} = \emptyset$ if $t \notin \mathbb{Z}$, and $|\mathcal{A}| = \binom{d}{t} \cdot \binom{n-d}{r-t}$ if $t \in \mathbb{Z}$.

b) Let $n \in \mathbb{N}$, and let $v, w, x, y \in \mathbb{F}_2^n$ having mutual distance $d \in \mathbb{N}$. Show that d is even, and that there is a unique $u \in \mathbb{F}_2^n$ such that $d(v, u) = d(w, u) = d(x, u) = \frac{d}{2}$. Does $d(y, u) = \frac{d}{2}$ hold as well?

(24.2) Exercise: Semi-linear isometries.
Let $\mathbb{F}_q$ be the finite field with q elements, let $\alpha \in \mathrm{Aut}(\mathbb{F}_q)$ be a field automorphism, and let $n \in \mathbb{N}$. An additive map $\varphi\colon \mathbb{F}_q^n \to \mathbb{F}_q^n$ such that $(va)^\varphi = v^\varphi \cdot a^\alpha$, for $v \in \mathbb{F}_q^n$ and $a \in \mathbb{F}_q$, is called an **α-semi-linear** map. (Hence the $\mathbb{F}_q$-linear maps are precisely the $\mathrm{id}_{\mathbb{F}_q}$-semi-linear ones.)

a) Show that the set of all α-semi-linear isometries of $\mathbb{F}_q^n$ with respect to the Hamming metric is given as the set of pairs $\mathrm{Isom}_\alpha(\mathbb{F}_q^n) = (\mathbb{F}_q^*)^n \times \mathcal{S}_n$, where $\pi \in \mathcal{S}_n$ acts by permuting the components of $v = [x_1, \ldots, x_n] \in \mathbb{F}_q^n$, and $\varphi := [a_1, \ldots, a_n] \in (\mathbb{F}_q^*)^n$ acts component-wise, that is $v^\varphi := [x_1^\alpha \cdot a_1, \ldots, x_n^\alpha \cdot a_n] \in \mathbb{F}_q^n$.

Show that $\mathrm{Isom}_\alpha(\mathbb{F}_q^n)$ fixes $0_n \in \mathbb{F}_q^n$, that is $\mathrm{Isom}_\alpha(\mathbb{F}_q^n) \subseteq \mathrm{Isom}_0(\mathbb{F}_q^n)$. When does $\mathrm{Isom}_\alpha(\mathbb{F}_q^n)$ carry a group structure? What happens if q is a prime?

b) Let $\mathrm{Isom}^*(\mathbb{F}_q^n)$ be the group of all isometries of $\mathbb{F}_q^n$ with respect to the Hamming metric, which map $\mathbb{F}_q$-subspaces to $\mathbb{F}_q$-subspaces. Show that $\mathrm{Isom}^*(\mathbb{F}_q^n)$ preserves minimum distance and minimum weight of $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$.

Show that for $n \geq 3$ as sets we have $\mathrm{Isom}^*(\mathbb{F}_q^n) = \coprod_{\alpha \in \mathrm{Aut}(\mathbb{F}_q)} \mathrm{Isom}_\alpha(\mathbb{F}_q^n)$. Can you describe the group structure of $\mathrm{Isom}^*(\mathbb{F}_q^n)$?

Show that for $n = 0$ we have $\mathrm{Isom}^*(\mathbb{F}_q) = \mathrm{Isom}_0(\mathbb{F}_q)$, while for $n = 1$ we get $\mathrm{Isom}^*(\mathbb{F}_q^2) = \coprod_{\alpha \in \mathrm{Aut}(\mathbb{F}_q^*)} \mathrm{Isom}_\alpha(\mathbb{F}_q^2)$, where $\mathrm{Aut}(\mathbb{F}_q^*)$ is the group of group automorphisms of $\mathbb{F}_q^*$, and $\mathrm{Isom}_\alpha(\cdot)$ is defined analogous to the semi-linear case.

(24.3) Exercise: Linear and non-linear isometries.
Let $\mathbb{F}_4 = \{0, 1, \omega, \omega+1\}$ be the field with 4 elements, where ω is a primitve root. For $\zeta \in \{\omega, \omega^2\}$ we consider the code $\mathcal{C}_\zeta \leq \mathbb{F}_4^8$ given by the generator matrix

$$G_\zeta := \begin{bmatrix} 1 & . & . & . & 1 & 1 & 1 & 1 \\ . & 1 & . & . & . & 1 & 1 & \omega \\ . & . & 1 & . & 1 & . & 1 & \omega \\ . & . & . & 1 & 1 & 1 & . & \zeta \end{bmatrix} \in \mathbb{F}_4^{4\times 8}.$$

a) Show that $\mathcal{C}_\omega$ and $\mathcal{C}_{\omega^2}$ are isometric, hence have the same minimum distance.
b) Show that $\mathcal{C}_\omega$ and $\mathcal{C}_{\omega^2}$ are not linearly isometric.

Hint for a). Show that $\mathcal{C}_\omega$ and $\mathcal{C}_{\omega^2}$ are α-semi-linearly isometric, where α is the Frobenius automorphism of $\mathbb{F}_4$.

(24.4) Exercise: Erasures.
Let $\mathcal{C}$ be a non-trivial block code. An entry of a code word being lost on transmition is called an **erasure**, amounting to an error with known position. For $e, g \in \mathbb{N}_0$ show that it is possible to correct e errors and g erasures simultaneously if and only if $2e + g \leq d(\mathcal{C}) - 1$.

(24.5) Example: Hamming bound.
Show that the parameters $q = 2$ and $n = 90$ and $m = 2^{78}$ and $d = 5$ fulfill the Hamming bound, but there is no binary (n, m, d) code. (Actually, this is the only such example for $q \leq 100$ and $n \leq 1000$ and $e \leq 1000$.)

Hint. Let $\mathcal{C} \subseteq \mathbb{F}_2^n$ be such a code. Consider the sets $\mathcal{A} := \{v = [x_1, \ldots, x_n] \in \mathbb{F}_2^n; x_1 = x_2 = 1, \mathrm{wt}(v) = 3\}$ and $\mathcal{B} := \{w = [y_1, \ldots, y_n] \in \mathcal{C}; y_1 = y_2 = 1, \mathrm{wt}(w) = 5\}$, and determine $|\{[v, w] \in \mathcal{A} \times \mathcal{B}; vw^{\mathrm{tr}} = 1\}|$ by double counting.

(24.6) Exercise: Perfect codes.
a) Show that any perfect code $\mathcal{C}_1 \subseteq \mathbb{F}_q^n$ having minimal distance 3 has the same parameters as a suitable Hamming code.
b) Let $\mathcal{C}_2 \subseteq \mathbb{F}_2^n$ be a perfect code having minimal distance 5. Show that we have $n = 5$, and determine $\mathcal{C}$.
c) Let $\mathcal{C}_3 \subseteq \mathbb{F}_2^n$ be a perfect code having minimal distance 7. Show that we have $n \in \{7, 23\}$. Determine $\mathcal{C}$ for the case $n = 7$.

(24.7) Exercise: Vasiliev codes.
We modify the Hamming code $\mathcal{H} \leq \mathbb{F}_2^7$ as follows: Let $f\colon \mathbb{F}_2^7 \to \mathbb{F}_2$ be given by

$f(0_7) := 0$, and $f(x) := 1$, for $0_7 \neq x = [x_1, \ldots, x_7] \in \mathbb{F}_2^7$. Then let

$$\mathcal{C} := \{[x;\, x+v;\, (\sum_{i=1}^{7} x_i) + f(v)] \in \mathbb{F}_2^{15};\, x \in \mathbb{F}_2^7,\, v \in \mathcal{H}\} \subseteq \mathbb{F}_2^{15}.$$

Determine the parameters of $\mathcal{C}$, and show that it is a perfect code, which is not equivalent to a linear code.

(24.8) Exercise: Check matrices.
Let $\mathcal{C} \leq \mathbb{F}_2^6$ be given by the following check matrix:

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

Compute a generator matrix of $\mathcal{C}$, and determine its minimum distance and its covering radius. Is $\mathcal{C}$ a self-dual code?

(24.9) Exercise: Self-dual codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a linear code having standard check matrix $H = [A \mid E_{n-k}] \in \mathbb{F}_q^{(n-k)\times n}$, where $k = \dim_{\mathbb{F}_q}(\mathcal{C})$ and $A \in \mathbb{F}_q^{(n-k)\times k}$. Show that $\mathcal{C}$ is self-dual if and only if $2k = n$ and $AA^{\mathrm{tr}} = -E_{n-k}$.

(24.10) Exercise: MDS codes.
Let $\mathcal{C} < \mathbb{F}_q^n$ be a non-trivial code. Show that the following are equivalent:
i) $\mathcal{C}$ is an MDS code. **ii)** $\mathcal{C}^\perp$ is an MDS code.
iii) $\mathcal{C}$ is systematic on all k-subsets of components.
iv) For any generator matrix, all k-subsets of columns are linearly independent.
v) For any check matrix, all $(n-k)$-subsets of columns are linearly independent.

(24.11) Exercise: Subfield subcodes.
Let $\mathbb{F}_q \subseteq F$ be a finite field extension, let $\Gamma := \mathrm{Aut}_{\mathbb{F}_q}(F) = \langle \varphi_q \rangle$, and let $\mathrm{Tr}\colon F \to \mathbb{F}_q\colon a \mapsto \sum_{\gamma\in\Gamma} a^\gamma$ be the $\mathbb{F}_q$-linear (surjective) **trace map**. Then by component-wise application we get an $\mathbb{F}_q$-linear automorphism $\varphi_q\colon F^n \to F^n\colon v \mapsto v^q$, and an $\mathbb{F}_q$-linear (surjective) map $\mathrm{Tr}\colon F^n \to \mathbb{F}_q^n$.

a) Given a code $\mathcal{C} \leq F^n$, show that $\dim_{\mathbb{F}_q}(\mathcal{C} \cap \mathbb{F}_q^n) \leq \dim_F(\mathcal{C}) \leq \dim_{\mathbb{F}_q}(\mathrm{Tr}(\mathcal{C}))$.

b) Show **Stichtenoth's Theorem** saying that the following are equivalent:
i) $\mathcal{C} = \mathcal{C}^q$, **ii)** $\dim_{\mathbb{F}_q}(\mathcal{C} \cap \mathbb{F}_q^n) = \dim_F(\mathcal{C})$, **iii)** $\dim_F(\mathcal{C}) = \dim_{\mathbb{F}_q}(\mathrm{Tr}(\mathcal{C}))$.

c) Show **Delsarte's Theorem** saying that $(\mathcal{C} \cap \mathbb{F}_q^n)^\perp = \mathrm{Tr}(\mathcal{C}^\perp) \leq \mathbb{F}_q^n$.

d) Conclude that if $\mathcal{C} = \mathcal{C}^q$, then $\mathcal{C}^\perp = (\mathcal{C}^\perp)^q$ and $(\mathcal{C} \cap \mathbb{F}_q^n)^\perp = \mathcal{C}^\perp \cap \mathbb{F}_q^n$.

(24.12) Exercise: Concatenation of codes.
Let $\mathbb{F}_q \subseteq \mathbb{F}$ be a finite field extension of degree $m = [\mathbb{F}\colon \mathbb{F}_q] := \dim_{\mathbb{F}_q}(\mathbb{F})$, let $\mathcal{C}$ be an $[n,k,d]$-code over $\mathbb{F}$, and let $\mathcal{D}$ be an $[l,m,\delta]$-code over $\mathbb{F}_q$.

Then $\mathcal{C}$ and $\mathbb{F}^n$ can be considered as $\mathbb{F}_q$-vector spaces, and specifying an $\mathbb{F}_q$-basis of $\mathbb{F}$ yields an encoding of $\mathbb{F}$ into $\mathbb{F}_q^l$. Show that this gives rise to an $\mathbb{F}_q$-linear code $\mathcal{D} \circ \mathcal{C} \leq \mathbb{F}_q^{nl}$ of $\mathbb{F}_q$-dimension km, called the **concatenation** of $\mathcal{C}$ and $\mathcal{D}$.

Show that $\mathcal{D} \circ \mathcal{C}$ has minimum distance bounded below by $d\delta$. Given generator matrices of $\mathcal{C}$ and $\mathcal{D}$, how can a generator matrix of $\mathcal{D} \circ \mathcal{C}$ be found?

(24.13) Exercise: Syndrome decoding.
a) Let $\mathcal{H} \leq \mathbb{F}_2^7$ be the Hamming code, having the following generator matrix:

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

Determine syndromes and coset leaders, and decode the following words:

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

b) Let $\mathcal{C} \leq \mathbb{F}_2^7$ be given by following generator matrix:

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

Determine syndromes and coset leaders, and decode the following words:

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

(24.14) Exercise: Unique decodability.
Let $\mathcal{C} \leq \mathbb{F}_2^{10}$ be given by the following generator matrix:

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

Show that all words in $\mathbb{F}_2^{10}$ are uniquely decodable. Determine the minimum distance and the covering radius of $\mathcal{C}$.

(24.15) Exercise: Modifying codes.
Apply the constructions of puncturing, extending, expurgation, augmentation, shortening and lengthening to the binary parity check and repetition codes, and determine the parameters of the resulting codes.

(24.16) Exercise: Extended codes.
Let the ternary codes $\mathcal{C}_1$ and $\mathcal{C}_2$ be given by the following generator matrices:

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

Determine the minimum distance of $\mathcal{C}_1$ and $\mathcal{C}_2$, and of their extensions.

(24.17) Exercise: Extended Hamming codes.
Let $k \geq 2$, and let $\mathcal{C}$ be a binary $[2^k, 2^k - k - 1, 4]$-code. Show that $\mathcal{C}$ is linearly equivalent to the extended Hamming code $\widehat{\mathcal{H}}_k$.

(24.18) Exercise: Hadamard codes.
a) A matrix $H \in \mathbb{R}^{n\times n}$, for $n \in \mathbb{N}$, having entries ± 1 and fulfilling $HH^{\mathrm{tr}} = nE_n$, is called a **Hadamard matrix**; if additionally all the entries in the first row and column are positive, then H is called **normalised**.

Let $H \in \mathbb{R}^{n\times n}$ und $H' \in \mathbb{R}^{n'\times n'}$ be Hadamard matrices. Show that if $n \geq 3$ then $4 \mid n$, and that $H \otimes H' \in \mathbb{R}^{(nn')\times(nn')}$ is a Hadamard matrix again.

b) Replacing the entries 1 and -1 of a Hadamard matrix H by $0 \in \mathbb{F}_2$ and $1 \in \mathbb{F}_2$, respectively, yields the associated **binary Hadamard matrix**. The rows of the binary matrices associated with a normalised Hadamard matrix H and with $-H$ form the associaterd binary **Hadamard code** $\mathcal{A}$. Shortening $\mathcal{A}$ with respect to the first component yields the **shortened** Hadamard code $\mathcal{A}^\circ$.

Show that $\mathcal{A}$ is an $(n, 2n, \frac{n}{2})$-code, and that $\mathcal{A}^\circ$ is an $(n-1, n, \frac{n}{2})$-code. Determine their covering radius. Do they fulfill the Plotkin bound?

c) Let $H_2 = H_2^{\otimes 1} := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ and $H_2^{\otimes(k+1)} := H_2^{\otimes k} \otimes H_2 \in \mathbb{R}^{2^k\times 2^k}$, for $k \in \mathbb{N}$. Show that $H_2^{\otimes k}$ is a normalised Hadamard matrix, also being called **Sylvester matrix**. Moreover, show that the associated binary codes $\mathcal{A}_k$ and $\mathcal{A}_k^\circ$ are linear, so that $\mathcal{A}_k$ is a $[2^k, k+1, 2^{k-1}]$-code and $\mathcal{A}_k^\circ$ is a $[2^k - 1, k, 2^{k-1}]$-code.

For $k \geq 2$ conclude that $\mathcal{A}_k$ linearly equivalent to the Reed-Muller code $\mathcal{R}_k$, and that $\mathcal{A}_k^\circ$ is linearly equivalent to the simplex code $\mathcal{S}_k$.

d) Show that the following matrix is a (normalised) Hadamard matrix:

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

Are the associated binary Hadamard and shortened Hadamard codes linear?

(24.19) Exercise: Sums and products of codes.
Let $\mathcal{C}$ be a non-trivial $[n,k,d]$-code, and let $\mathcal{C}'$ be a non-trivial $[n',k',d']$-code, both over $\mathbb{F}_q$, having generator matrices $G \in \mathbb{F}_q^{k\times n}$ and $G' \in \mathbb{F}_q^{k'\times n'}$, respectively. Show the following assertions, and find out what happens if different generator matrices are chosen:

a) If $k = k'$, then the rows of $[G \mid G'] \in \mathbb{F}_q^{k\times(n+n')}$ generate an $[n+n', k, d'']$-code such that $d'' \geq d + d'$, being called the **glue** of $\mathcal{C}$ and $\mathcal{C}'$.

b) The rows of the matrix $\left[\begin{array}{c|c} G & . \\ \hline . & G' \end{array}\right] \in \mathbb{F}_q^{(k+k')\times(n+n')}$ generate an $[n+n', k+k', \min\{d,d'\}]$-code, being called the **direct sum** of $\mathcal{C}$ and $\mathcal{C}'$.

c) The set of matrices in $\mathbb{F}_q^{n\times n'}$, whose columns and rows are elements of $\mathcal{C}^{\mathrm{tr}}$ and $\mathcal{C}'$, respectively, forms an $[nn', kk', dd']$-code, being called the **direct product** of $\mathcal{C}$ and $\mathcal{C}'$. Upon identifying $\mathbb{F}_q^{n\times n'} \cong \mathbb{F}_q^{nn'}$, the direct product of $\mathcal{C}$ and $\mathcal{C}'$ has generator matrix $G \otimes G' \in \mathbb{F}_q^{(kk')\times(nn')}$.

(24.20) Exercise: Product codes.
Let $\mathcal{C} \leq \mathbb{F}_2^{8\times 16}$ be the direct product of the extended binary Hamming codes $\widehat{\mathcal{H}}_3$ und $\widehat{\mathcal{H}}_4$; hence $\mathcal{C}$ is a $[128, 44, 16]$-code, and thus is 7-error correcting. Actually, $\mathcal{C}$ is much better than that:

Assume that a received word contains precisey 14 errors, in the (randomly chosen, but unknown) positions $[2, 7, 19, 24, 27, 32, 45, 51, 53, 76, 82, 86, 96, 121]$, where the rows of the matrices are enumerated row-wise. Show that the received word is uniquely decodable.

25 Exercises to Part III

(25.1) Exercise: Bounding the code length.
Consider an $(n, q^{n-3}, 3)$-code over an alphabet with q elements, where $n \in \mathbb{N}$. Show that $n \leq q^2 + q + 1$.

(25.2) Exercise: Bounding the code order.
Let $\mathcal{C}$ be a binary code of length $n \in \mathbb{N}$, having minimum distance $d \geq 3$. Show that $|\mathcal{C}| \leq \frac{2^n}{n+1}$, and if n is even then $|\mathcal{C}| \leq \frac{2^n}{n+2}$. What happens for n odd?

Hint. Apply double counting to $\{[v, w] \in \mathcal{C} \times \mathbb{F}_2^n; d(v, w) = 2\}$.

(25.3) Exercise: Bounding the minimum distance.
a) Use the Singleton, Hamming and Plotkin bounds to show that for any binary $(6, 9, d)$-code we have $d \leq 3$.
b) Show that there is a binary $(6, 9, 2)$-code, but there is no binary $(6, 9, 3)$-code.

(25.4) Exercise: Weight sum.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a code such that $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \geq 1$, having a generator matrix which does not contain a zero column. Show that $\sum_{v \in \mathcal{C}} \mathrm{wt}(v) = n(q-1)q^{k-1}$.

(25.5) Exercise: Equidistant codes.
Let $\mathcal{C} \subseteq \mathbb{F}_2^{16}$ be a binary code such that $\mathrm{wt}(v) = 6$ and $d(v, w) = 8$, for $v \neq w \in \mathcal{C}$. Show that $|\mathcal{C}| \leq 16$. Is there such a code fulfilling $|\mathcal{C}| = 16$?

(25.6) Exercise: Griesmer bound.
a) Show that the parameters $q := 3$, $n := 14$, $k := 4$ and $d := 9$ fulfill the Griesmer bound, but there is no ternary $[14, 4, 9]$.code.
b) Show that the parameters $q := 2$, $n := 15$, $k := 8$ and $d := 5$ fulfill the Griesmer bound, but there is no binary $[15, 8, 5]$-code.

(25.7) Exercise: Optimal binary codes.
For $n, d \in \mathbb{N}$ let $K_2(n, d) := \max\{k \in \mathbb{N}$; there is a binary $[n, k, d]$-code$\}$. Show that $K_2(n, 2d-1) = K_2(n+1, 2d)$ and $K_2(n+1, d) \leq K_2(n, d) + 1$.

(25.8) Exercise: Punctured simplex code.
Beginning with the binary simplex $[31, 5, 16]$-code, use iterated puncturing to construct a $[21, 5, 10]$-code. Does the latter fulfill the Griesmer bound?

(25.9) Exercise: Shortened Hamming codes.
Let $k \geq 2$. Show that the shortened binary Hamming code $(\mathcal{H}_k)^\circ$ is optimal.

(25.10) Exercise: Optimal codes.
a) Determine $K_2(10, 5)$.
b) Determine bounds for $K_2(17, 8)$.

(25.11) Exercise: Best code [1978].
We provide an example showing that non-linear codes might outperform optimal linnear codes, but that it might be difficult to find them:

a) Show first that $K_2(10,4) = 5$.

b) Now let $\mathcal{C}_0 \leq \mathbb{F}_2^{10}$ be given by the generator matrix $[G \mid G] \in \mathbb{F}_2^{3\times 10}$, where

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

Letting $v := [1,0,0,0,0;\ 0,0,1,0,0] \in \mathbb{F}_2^{10}$ and $\pi := (1,2,3,4,5)(6,7,8,9,10) \in \mathcal{S}_{10}$, let $\mathcal{C} := (v+\mathcal{C}_0)^{\langle\pi\rangle} \subseteq \mathbb{F}_2^{10}$, that is the $\langle\pi\rangle$-orbit of the coset $v+\mathcal{C}_0 \subseteq \mathbb{F}_2^{10}$, be the (non-linear) **Best code**.

Show that $\mathcal{C}$ is a $(10,40,4)$-code. (It can be shown that $\mathcal{C}$ is optimal.)

(25.12) Exercise: Even weight.
Assume there is a binary (n,m,d)-code such that $d \in \mathbb{N}$ is even. Show that there is a binary (n,m,d)-code such that all its words have even weight.

(25.13) Exercise: Non-linear codes.
a) Use suitable bounds to show that any binary $(6,9,d)$-code fulfills $d \leq 3$.
b) Provide a binary $(6,9,2)$-code, but show that there is no binary $(6,9,3)$-code.

(25.14) Exercise: Weight distributions.
Let $\mathcal{C} \leq \mathbb{F}_2^n$ having weight enumerator $A_\mathcal{C} \in \mathbb{C}[X,Y]$.
a) Determine the weight enumerator of the extended code $\widehat{\mathcal{C}}$, the expurgated code $\mathcal{C}'$ and the augmented code $\widetilde{\mathcal{C}}$.
b) For a polynomial $0 \neq f \in \mathbb{C}[X]$ let $f^* := X^{\deg(f)} \cdot f(X^{-1}) \in \mathbb{C}[X]$ be the **reversed** polynomial. Give a necessary and sufficient condition for $A_\mathcal{C}(X,1) \in \mathbb{C}[X]$ being self-reversed.

(25.15) Exercise: MacWilliams identities.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a code such that $k := \dim_{\mathbb{F}_q}(\mathcal{C})$, let $[A_0,\ldots,A_n]$ be the weight distrubution of $\mathcal{C}$, and let $[A_0^\perp,\ldots,A_n^\perp]$ be the weight distrubution of $\mathcal{C}^\perp \leq \mathbb{F}_q^n$. For $r \in \{0,\ldots,n\}$ show that

$$\sum_{i=0}^{n-r} \binom{n-i}{r} A_i = q^{k-r} \cdot \sum_{j=0}^{r} \binom{n-j}{r-j} A_j^\perp.$$

(25.16) Exercise: Weight distribution of MDS codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be an MDS code such that $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \geq 1$; hence $d := d(\mathcal{C}) = n-k+1$. Moreover, let $[A_0,\ldots,A_n]$ be the weight distrubution of $\mathcal{C}$.

a) Show that $A_i = 0$ for $i \in \{1, \ldots, n-k\}$, and that for $i \in \{1, \ldots, k\}$ we have

$$A_{n-k+i} = \binom{n}{k-i} \cdot \sum_{j=0}^{i-1} (-1)^j \binom{n-k+i}{j} (q^{i-j} - 1);$$

this says that the weight distrubution of $\mathcal{C}$ is determined by n and k.

b) Conclude (again) that $k \leq 1$ or $d \leq q$, and that $d \leq 2$ or $k \leq q-1$.

Hint for a). Let $\mathcal{C}_{\mathcal{I}} := \{[x_1, \ldots, x_n] \in \mathcal{C}; x_i = 0 \text{ for } i \in \mathcal{I}\}$, for $\mathcal{I} \subseteq \{1, \ldots, n\}$. For $j \in \{0, \ldots, k-1\}$ apply double counting to $\{[\mathcal{I}, v]; |\mathcal{I}| = j, v \in \mathcal{C}_{\mathcal{I}} \setminus \{0\}\}$.

(25.17) Exercise: Weight distribution of Hamming codes.
a) For $k \geq 2$ and $n := 2^k - 1$ let $\mathcal{H}_k$ be the associated binary Hamming code, having weight distribution $[A_0, \ldots, A_n]$. Show that for $i \geq 2$ the latter fulfills the recursion $iA_i + A_{i-1} + (n-i+2)A_{i-2} = \binom{n}{i-1}$, where $A_0 = 1$ und $A_1 = 0$.
b) Determine the weight distribution of $\widehat{\mathcal{H}}_k$ and $\mathcal{H}'_k$.
c) Determine the weight distribution of the Reed-Muller codes $\mathcal{R}_k$ and $\mathcal{R}_k^{\bullet}$.

(25.18) Exercise: Weight distribution of extended codes.
Let $\mathcal{C} \leq \mathbb{F}_2^n$ be a binary code having weight distribution $[A_0, \ldots, A_n]$, let $\widehat{\mathcal{C}} \leq \mathbb{F}_2^{n+1}$ be the associated extended code having weight distribution $[\widehat{A}_0, \ldots, \widehat{A}_{n+1}]$, and assume that $\mathrm{Aut}_{\mathbb{F}_q}(\widehat{\mathcal{C}})$ acts transitively on the components.

Show **Prange's Theorem**, saying that for $i \in \{1, \ldots, \frac{n+1}{2}\}$ we have

$$A_{2i-1} = \frac{2i}{n+1-2i} \cdot A_{2i} = \frac{2i}{n+1} \cdot \widehat{A}_{2i}.$$

Conclude that $\mathcal{C}$ has odd minimum distance.

(25.19) Exercise: Divisible codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a code such that $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \in \mathbb{N}$, having no zero component.

a) Show that $\sum_{v \in \mathcal{C}} \mathrm{wt}(v) = nq^{k-1}(q-1)$. How is this statement related to the MacWilliams identities?

b) Let now $\mathcal{C}$ be r-divisible, for some $r \geq 2$, and let $s := \frac{r}{\gcd(r, q^{k-1})}$. Show **Ward's Theorem**, saying that $\mathcal{C}$ is an s-fold **repeated** code, that is of shape $\mathcal{C} = \{[v, \ldots, v] \in \mathbb{F}_q^n; v \in \mathcal{C}_0\}$, for some code $\mathcal{C}_0 \leq (\mathbb{F}_q)^{\frac{n}{s}}$.

(25.20) Exercise: Divisible self-dual codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be an r-divisible self-dual $[n, \frac{n}{2}, \frac{n}{2}]$-code, for $r \geq 2$ and $n \geq 4$ even.
a) Show that $q = 2$ and $n = 8$, or $q = 3$ and $n = 12$, or $q = 4$ and $n = 8$.
b) Show that $\mathcal{C}$ is linearly equivalent to $\widehat{\mathcal{H}}_3$, or $\mathcal{G}_{12}$, or $\widehat{\mathcal{H}}_3 \otimes_{\mathbb{F}_2} \mathbb{F}_4$, respectively.

(25.21) Exercise: Self-dual codes.
Let p be a prime. Provide self-dual $\mathbb{F}_p$-linear codes of length $n \in \{2, 4, 6, 8\}$.

Hint. Distinguish the cases $p = 2$ and $p \equiv \pm 1 \pmod 4$.

(25.22) Exercise: Weakly self-dual binary codes.
Let $\mathcal{C} \leq \mathcal{C}^\perp \leq \mathbb{F}_2^n$ be a weakly self-dual code, and let n be odd.
a) Show that $\mathcal{C}$ is 4-divisible.
b) Assume that $\dim_{\mathbb{F}_2}(\mathcal{C}) = \frac{n-1}{2}$. Show that $\mathcal{C}^\perp = \widetilde{\mathcal{C}}$.

(25.23) Exercise: Self-dual binary codes.
a) Show that a self-dual binary code of length 10 has minimum distance at most 4, and determine the possible weight enumerators. Show that there is no self-dual binary $[10, 5, 4]$-code, but provide a self-dual binary $[10, 5, 2]$-code. Is the latter extremal?

b) Show that a self-dual binary code of length 16 has minimum distance at most 6, and determine the possible weight enumerators. Provide self-dual binary $[16, 8, 2]$- and $[16, 8, 4]$-codes. Is there a self-dual binary $[16, 8, 6]$-code? Would the latter be extremal?

c) Show that a self-dual binary code of length 32 has minimum distance at most 10. Show that there is no self-dual binary $[32, 16, 10]$-code. Determine the weight enumerator of a putative self-dual binary $[32, 16, 8]$-code. Would the latter be extremal?

d) Show that a self-dual binary code of length 48 has minimum distance at most 14. Show that there is no self-dual binary $[48, 24, 14]$-code. Determine the weight enumerator of a putative doubly-even self-dual binary $[48, 24, 12]$-code. In which sense would the latter be extremal?

Hint. Use suitable bounds.

(25.24) Exercise: Quaternary codes.
In general, a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is called **formally self-dual** if $A_\mathcal{C} = A_{\mathcal{C}^\perp}$. Now let $\mathcal{C} \leq \mathbb{F}_4^n$ be a 2-divisible quaternary code of even length, such that $\dim_{\mathbb{F}_4}(\mathcal{C}) = \frac{n}{2}$.

a) Show that $\mathcal{C}$ is formally self-dual. Moreover, if $\mathcal{C}$ is self-dual, then there is a subfield subcode $\mathcal{C}_0 \leq \mathbb{F}_2^n$ such that $\mathcal{C} = \mathcal{C}_0 \otimes_{\mathbb{F}_2} \mathbb{F}_4$.

b) Provide a formally self-dual 2-divisible quaternary $[6, 3, 4]$-code, and determine its weight enumerator. But show that there is no extremal 2-divisible self-dual quaternary code of length 6.

(25.25) Exercise: Extremal binary codes.
For $n \in \mathbb{N}$ let $\mathcal{C} \leq \mathbb{F}_2^n$ be a (putative) extremal (even-weight or doubly-even) self-dual binary code. How do its parameters compare to the asymptotic Gilbert-Varshamov, Hamming, Elias and McEliece bounds, for $n \gg 0$?

(25.26) Exercise: Counting self-dual binary codes.
a) Let $\mathcal{C} \leq \mathbb{F}_2^n$ be a weakly self-dual binary code of dimension $k \geq 1$. Show that the number of self-dual binary codes containing $\mathcal{C}$ equals $\prod_{i=1}^{\frac{n}{2}-k}(2^i+1)$.
b) Conclude that the number of self-dual binary codes equals $\prod_{i=1}^{\frac{n}{2}-1}(2^i+1)$.
c) Let $v \in \mathbb{F}_2^n$ have even weight, where $v \notin \{0_n, 1_n\}$. Show that the number of self-dual binary codes containing v equals $\prod_{i=1}^{\frac{n}{2}-2}(2^i+1)$.

26 Exercises to Part IV

(26.1) Exercise: Modifying cyclic codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a cyclic code having generator polynomial $g \in \mathbb{F}_q[X]$. Which of the constructions puncturing, extension, expurgation, augmentation, shortening and lengthening of $\mathcal{C}$ are cyclic again? Provide suitable generator polynomials.

(26.2) Exercise: Constructing cyclic codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ and $\mathcal{C}' \leq \mathbb{F}_q^{n'}$ be non-trivial cyclic codes having generator polynomials $g \in \mathbb{F}_q[X]$ and $g' \in \mathbb{F}_q[X]$, respectively.
a) Let $\gcd(n, n') = 1$. Show that the direct product of $\mathcal{C}$ und $\mathcal{C}'$ is a cyclic code as well. Provide a suitable generator polynomial.
b) Let $q := 2$, let $n = n'$ be odd, and assume that $g \mid g'$. Show that the Plotkin sum of $\mathcal{C}$ and $\mathcal{C}'$ is linearly equivalent to a cyclic code having generator polynomial $gg' \in \mathbb{F}_2[X]$.

(26.3) Exercise: Generalized cyclic codes.
Let $n \in \mathbb{N}$ and $0 \neq a \in \mathbb{F}_q$. A code $\mathcal{C} \leq \mathbb{F}_q^n$ is called **a-cyclic**, if for any $[x_0, \ldots, x_{n-1}] \in \mathcal{C}$ we have $[ax_{n-1}, x_0, \ldots, x_{n-2}] \in \mathcal{C}$ as well.

Provide a bijective correspondence between the set of a-cyclic codes of length n and the set of ideals of a suitable quotient of the polynomial ring $\mathbb{F}_q[X]$.

Moreover, show that the set of a-cyclic codes is closed under taking sums and intersections, and describe the associated generator polynomials.

(26.4) Exercise: CRC codes.
a) Write GAP programs for CRC encoding and decoding. Which input data is needed? What is the output? Which consistency checks should be made?

b) Apply the programs to the CRC code $\mathcal{H} \leq \mathbb{F}_2^7$ having generator polynomial $X^3 + X + 1 \in \mathbb{F}_2[X]$: Compute the encoding of the following words:

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

and decide which of the following vectors belong to $\mathcal{H}$:

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

(26.5) Exercise: RWTH-ID.
Write GAP programs to generate and verify RWTH-IDs, applying the programs from Exercise (26.4) with respect to the generator polynomial $X^5 + X^2 + 1 \in \mathbb{F}_2[X]$, and using the source coding from Table 7.

(26.6) Exercise: Cyclotomic polynomials.
Let $\mathbb{F}_q$ be the field with q elements, and let $n \in \mathbb{N}$.
a) For $m \in \mathbb{N}$ show that $X^m - 1$ divides $X^n - 1$ if and only if m divides n.
b) Let $p := \mathrm{char}(\mathbb{F}_q)$, and let $\overline{\mathbb{F}}$ be an algebraic closure of $\mathbb{F}_q$. How are the factorizations of $X^{p^n} - 1 \in \mathbb{F}_q[X]$ and $X^n - 1 \in \mathbb{F}_q[X]$ related? How are the sets of zeroes $\mathcal{V}(X^{p^n} - 1) \subseteq \overline{\mathbb{F}}$ and $\mathcal{V}(X^n - 1) \subseteq \overline{\mathbb{F}}$ related?

(26.7) Exercise: Classification of cyclic codes.
a) Determine all cyclic binary codes of length $n \in \{1, \ldots, 32\}$, their sets of zeroes, and their generator polynomials. How are these codes related with respect to inclusion? How do these codes behave with respect to duality?
b) For $n \leq 10$ determine the minimum distance of the above codes.

(26.8) Exercise: Reversible codes.
A cyclic code $\mathcal{C} \leq \mathbb{F}_q^n$ having generator polynomial $g \in \mathbb{F}_q[X]$ is called **reversible**, if for any $[x_0, \ldots, x_{n-1}] \in \mathcal{C}$ we have $[x_{n-1}, \ldots, x_0] \in \mathcal{C}$ as well.

a) Show that the following assertions are equivalent:
i) The code $\mathcal{C}$ is reversible.
ii) We have $g^* = ag \in \mathbb{F}_q[X]$ for some $a \in \mathbb{F}_q^*$.
iii) The set of zeroes $\mathcal{V}(\mathcal{C}) = \mathcal{V}(g)$ is closed under taking inverses.

b) Assume that $\gcd(q, n) = 1$, and that $-1 \in \mathbb{Z}_n^*$ is a q-power. Show that any cyclic code $\mathcal{C} \leq \mathbb{F}_q^n$ is reversible.

(26.9) Exercise: Cyclic binary codes of even length.
For $k \geq 3$ let $\mathcal{H}_k \leq \mathbb{F}_2^n$ the associated Hamming code, where $n := 2^k - 1$. Show that the expurgated shortened code $(\mathcal{H}_k^\circ)' \leq \mathbb{F}_2^{n-1}$ is linearly equivalent to a cyclic $[n-1, n-k-2, 4]$-code.

Hint. Use the Plotkin sum of $\mathcal{H}'_{k-1}$ and an even-weight code.

(26.10) Exercise: Binary simplex codes.
For $k \geq 2$ let $\mathcal{S}_k \leq \mathbb{F}_2^n$ be the associated simplex code, where $n := 2^k - 1$. Show that $\mathcal{S}_k$ is linearly equivalent to a cyclic code, determine its set of zeroes $\mathcal{V}(\mathcal{S}_k)$, and use this to recover $\dim_{\mathbb{F}_2}(\mathcal{S}_k) = k$ und $d(\mathcal{S}_k) = 2^{k-1}$.

(26.11) Exercise: BCH codes.
Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a BCH code defined by $\{\zeta_n^a, \ldots, \zeta_n^{a+\delta-2}\}$, where $a \in \mathbb{Z}_n$ and $\delta \in \{1, \ldots, n+1\}$. Is the dual code $\mathcal{C}^\perp$ necessarily a BCH code as well?

(26.12) Exercise: Binary BCH codes.
a) Let $\mathcal{C} \leq \mathbb{F}_2^n$ be a narrow sense BCH code having designed distance $\delta \in \{1, \ldots, n+1\}$. Show that $\dim_{\mathbb{F}_2}(\mathcal{C}) \geq n - \lfloor \frac{\delta}{2} \rfloor \cdot |\mathrm{Aut}_{\mathbb{F}_2}(\mathbb{F}_2(\zeta_n))|$.

b) Now assume that $\mathcal{C}$ is primitive, where $n := 2^k - 1$ for some $k \geq 1$, such that $\lfloor \frac{\delta}{2} \rfloor \leq 2^{\lceil \frac{k}{2} \rceil - 1}$. Show that $\dim_{\mathbb{F}_2}(\mathcal{C}) = n - \lfloor \frac{\delta}{2} \rfloor \cdot k$.

c) Determine the dimension of the primitive narrow sense BCH codes for $n = 31$.

(26.13) Exercise: Ternary BCH codes.
Determine the maximum dimension of a primitive ternary BCH code of length 26 and designed distance 5.

(26.14) Exercise: Reed-Solomon codes.
a) Let $\mathcal{C} \leq \mathbb{F}_q^{q-1}$ be a Reed-Solomon code such that $1 \notin \mathcal{V}(\mathcal{C})$. Show that the extended code $\widehat{\mathcal{C}} \leq \mathbb{F}_q^q$ is an MDS-Code.
b) Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a primitive BCH code. Show that there is a finite field $\mathbb{F}_q \subseteq F$ and a Reed-Solomon code $\mathcal{D} \leq F^n$ such that $\mathcal{C} = \mathcal{D} \cap \mathbb{F}_q^n$.

(26.15) Exercise: Roos bound.
Let $n := 2^k - 1$, where $k \geq 3$, and let $\mathcal{C} \leq \mathbb{F}_2^n$ be cyclic.
a) Let $\{\zeta_n, \zeta_n^5\} \subseteq \mathcal{V}(\mathcal{C})$. Show that $\mathcal{C}$ has minimum distance ≥ 4.
b) Let $\{\zeta_n, \zeta_n^{-1}\} \subseteq \mathcal{V}(\mathcal{C})$. Show that $\mathcal{C}$ has minimum distance ≥ 5, and that the expurgated code $\mathcal{C}' \leq \mathbb{F}_2^n$ has minimum distance ≥ 6.

(26.16) Exercise: van-Lint–Wilson bound.
a) Let $n := 2^{2k} + 1$, where$k \geq 1$, and let $\mathcal{C} \leq \mathbb{F}_2^n$ be the cyclic code associated with $\{\zeta_n\}$. Show that $\mathcal{C}$ is reversible and has minimum distance ≥ 5.
b) Let $\mathcal{C} \leq \mathbb{F}_2^{31}$ be the cyclic code associated with $\{\zeta_{31}, \zeta_{31}^5, \zeta_{31}^7\}$. Show that $\mathcal{C}$ has minimum distance ≥ 7. Does equality hold?

27 Exercises to Part V

(27.1) Example: QR codes.
a) Determine the minimum distance of the binary QR code of length 47.
b) Determine all perfect 1-error correcting QR codes.

(27.2) Example: Extended QR codes.
Provide a self-dual binary $[32, 16, 8]$-code and a self-dual binary $[48, 24, 12]$-code.

(27.3) Exercise: Golay codes.
Determine the residual code of the extended ternary Golay code $\mathcal{G}_{12}$ and of the extended binary Golay code $\mathcal{G}_{24}$.

(27.4) Exercise: Weight distribution of Golay codes.
Determine the weight distribution of the binary Golay codes $\mathcal{G}_{23}$ and $\mathcal{G}'_{23}$, and of the ternary Golay codes $\mathcal{G}_{11}$ and $\mathcal{G}'_{11}$.

(27.5) Exercise: Steiner systems.
A **Steiner system** $S(t,k,v)$, where $t,k,v \in \mathbb{N}$ such that $t \leq k \leq v$, is a set $\mathcal{P}$ of **points** of cardinality v, together with a set $\mathcal{B}$ of k-subsets of $\mathcal{P}$, called **blocks**, such that any t-subset of $\mathcal{P}$ is contained in precisely one block.

a) For $s \in \{0,\dots,t\}$ show that any s-subset of points is contained in precisely $\lambda_s \in \mathbb{N}$ blocks, where $\lambda_s \cdot \binom{k-s}{t-s} = \binom{v-s}{t-s}$.

Conclude that there are precisely $b := |\mathcal{B}| \in \mathbb{N}$ blocks, where $b \cdot \binom{k}{t} = \binom{v}{t}$; that each point belongs to precisely $r \in \mathbb{N}$ blocks, where $bk = vr$; and that for $t = 2$ we have $r(k-1) = v-1$.

b) Given a Steiner system $S(t,k,v)$, where $t \geq 2$, show that there is a Steiner system $S(t-1,k-1,v-1)$.

Steiner systems for $t = 1$ or $t = k$ are not too interesting (why?), and are fairly easy to find for $t = 2$, which we will see soon. But it is hard to find any for $t \geq 3$, where conjecturally there are none for $k > t \geq 6$; we will see below that there are sporadic ones for $t = 4$ und $t = 5$.

c) Let $\mathbf{A}^2(\mathbb{F}_q) := \mathbb{F}_q^2$ be the **affine plane** over $\mathbb{F}_q$, where the subsets $w + \langle v \rangle_{\mathbb{F}_q} \subseteq \mathbb{F}_q^2$, for $v, w \in \mathbb{F}_q^2$ such that $v \neq 0$, are called **affine lines**. Show that the affine plane, together with the affine lines, forms a Steiner system $S(2,q,q^2)$, and determine the parameters λ_s, b und r.

d) Let $\mathbf{P}^2(\mathbb{F}_q) := \{\langle v \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^3; 0 \neq v \in \mathbb{F}_q^3\}$ be the **projective plane** over $\mathbb{F}_q$, where the subspaces $\langle v, w \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^3$ such that $\langle v \rangle_{\mathbb{F}_q} \neq \langle w \rangle_{\mathbb{F}_q} \in \mathbf{P}^2(\mathbb{F}_q)$, are called **projective lines**. Show that the projective plane, together with the projective lines, forms a Steiner system $S(2, q+1, q^2+q+1)$, and determine the parameters λ_s, b und r. Finally, draw a picture of the **Fano plane** $\mathbf{P}^2(\mathbb{F}_2)$.

(27.6) Exercise: Codes and Steiner systems.
We relate codes with Steiner systems as follows: Let $\mathcal{C} \leq \mathbb{F}_2^n$ be a non-trivial $[n,k,d]$-code, and for $\mathcal{P} := \{1,\dots,n\}$ let $\mathcal{B} := \{\mathrm{supp}(v) \subseteq \mathcal{P}; v \in \mathcal{C}, \mathrm{wt}(v) = d\}$.

a) Show that $\mathcal{C}$ is perfect such that $d = 2e+1$ if and only if $\mathcal{B}$ is the set of blocks of a Steiner system $S(e+1, 2e+1, n)$. Moreover, in this case the extended code $\widehat{\mathcal{C}}$ yields a Steiner system $S(e+2, 2e+2, n+1)$. How are these systems related?

b) Conclude that the Hamming code $\mathcal{H}_k$ and the extended Hamming code $\widehat{\mathcal{H}}_k$, for $k \geq 2$, yield Steiner systems $S(2,3,2^k-1)$ and $S(3,4,2^k)$, respectively.

c) Similarly, the Golay code $\mathcal{G}_{23}$ and the extended Golay code $\mathcal{G}_{24}$ yield Steiner systems $S(4,7,23)$ and $S(5,8,24)$, respectively, also being called **Witt systems**. Use this to determine the weight distribution of $\mathcal{G}_{24}$ (again).

(27.7) Exercise: A weighing problem.
a) We have $n \geq 3$ identical coins, amongst which there is at most one fake coin having a different weight. How many weighings using a beam balance are needed to identify the fake coin? How many weighings are needed if their order is fixed right from the beginning?

b) How many weighings are needed for 11 coins, amongst which there are up to two fake coins having the same weight? How many are needed for fixed order?

Hint for a). Consider the cases $n = \frac{3^k-1}{2}$, for $k \geq 2$.

28 Exercises to Part VI

(28.1) Exercise: Generalized Reed-Solomon codes.
a) Let $\mathrm{GRS}_k(\alpha, v) \leq \mathbb{F}_q^n$ be a generalized Reed-Solomon code, and let $\beta := c \cdot \alpha + d \cdot 1_n$, where $c \in \mathbb{F}_q^*$ and $d \in \mathbb{F}_q$. Show that $\mathrm{GRS}_k(\alpha, v) = \mathrm{GRS}_k(\beta, v)$.
b) Let $\mathrm{GRS}_k(\alpha, w) \leq \mathbb{F}_q^n$ also be a generalized Reed-Solomon code. When are $\mathrm{GRS}_k(\alpha, v)$ and $\mathrm{GRS}_k(\alpha, w)$ linearly equivalent?

(28.2) Exercise: Goppa codes.
Let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_q^n$ be a Goppa code, where $\alpha = [\alpha_1, \ldots, \alpha_n]$. Show that it can be written as subfield subcode $\mathcal{G}(\alpha, g) = \mathrm{GRS}_n(\alpha, [v_1, \ldots, v_n]) \cap \mathbb{F}_q^n$, where

$$v_i = \frac{g(\alpha_i)}{\prod_{j \neq i}(\alpha_i - \alpha_j)} \quad \text{for} \quad i \in \{1, \ldots, n\}.$$

(28.3) Exercise: Goppa codes in characteristic 2.
Let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_q^n$ be a separable Goppa code, where q is even. Show that its minimum distance fulfills $d(\mathcal{G}(\alpha, g)) \geq 2\deg(g) + 1$.

Hint. Consider the polynomial $f_\alpha := \prod_{i=1}^n (X - \alpha_i)$ and its derivative.

(28.4) Exercise: Cyclic Goppa codes.
Let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_2^n$ be a binary Goppa code, where $\alpha = \mathcal{V}_n$. Assume that $\mathcal{G}(\alpha, g)$ is cyclic. Show that the Goppa polynomial can be chosen as $g = X^k$, for some $k \in \{0, \ldots, n\}$. Conclude that $\mathcal{G}(\alpha, g)$ is a narrow sense BCH code.

(28.5) Exercise: Goppa codes and BCH codes.
a) Let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_2^{15}$ be the (non-separable) binary Goppa code given by $\alpha := \mathbb{F}_{16}^*$ and $g := Z^2 + 1$. Determine its $\mathbb{F}_2$-dimension and its minimum distance.

b) Let $\mathcal{C} \leq \mathbb{F}_2^{15}$ be the cyclic binary code having generator polynomial X^2+X+1. Show that $\mathcal{C}$ is a BCH code, but is not a Goppa code.

(28.6) Exercise: Irreducible Goppa codes.
a) Let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_2^8$ be the irreducible binary Goppa code given by $\alpha := \mathbb{F}_8$ and $g := X^2 + X + 1$. Determine its $\mathbb{F}_2$-dimension and its minimum distance.

b) Let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_2^{32}$ be the irreducible binary Goppa code given by $\alpha := \mathbb{F}_{32}$ and $g := X^3 + X + 1$. Determine its $\mathbb{F}_2$-dimension and its minimum distance.

c) Let $g := X^2 + X + \zeta_5 \in \mathbb{F}_{16}[X]$. Show that g is irreducible.

Hence let $\mathcal{G}(\alpha, g) \leq \mathbb{F}_2^{16}$ be the associated irreducible binary Goppa code given by $\alpha := \mathbb{F}_{16}$. Determine its $\mathbb{F}_2$-dimension and its minimum distance.

(28.7) Exercise: Divisors.
a) Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$, let $D \in \mathrm{Div}_{\mathbf{V}}$, and let $0 \neq f \in \mathscr{L}(D)$. Show that $f \notin \mathscr{L}(D - (x))$ for almost all $x \in \mathbf{V}$.
b) Conclude that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D - (x))) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) - 1$ for almost all $x \in \mathbf{V}$.

(28.8) Exercise: Effective divisors.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$, and let $D \in \mathrm{Div}_{\mathbf{V}}$ be effective. Show that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1 + \deg(D)$ if and only if $D = 0$ or $g = 0$.

(28.9) Exercise: Divisors of degree 0.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$, and let $D \in \mathrm{Div}_{\mathbf{V}}$.
a) Assume that $\deg(D) = 0$ and $\mathscr{L}(D) \neq \{0\}$. Give a description of $\mathscr{L}(D)$.
b) Let $g = 1$ and let D be canonical. Show that $D \sim 0$.

(28.10) Exercise: Canonical divisors.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$, and let $K \in \mathrm{Div}_{\mathbf{V}}$.

a) Show that K is canonical if and only if $\deg(K) = 2g-2$ and $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(K)) \geq g$.

b) Show that there is an effective canonical divisor if and only if $g \geq 1$.

c) Let $k \in \mathbb{Z}$, and assume that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1+\deg(D)-k+\dim_{\overline{\mathbb{F}}}(\mathscr{L}(K-D))$, for all $D \in \mathrm{Div}_{\mathbf{V}}$. Show that $k = g$ and that K is canonical.

(28.11) Exercise: Special divisors.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$.

a) For $D \leq D' \in \mathrm{Div}_{\mathbf{V}}$ show $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D')) - \deg(D') \leq \dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) - \deg(D)$. Conclude that if D is non-special, then D' is non-special as well.

b) Show that any $D \in \mathrm{Div}_{\mathbf{V}}$ such that $\mathscr{L}(D) \neq \{0\}$ and $\deg(D) < g$ is special.

c) Let $\{x_1, \ldots, x_r\} \subseteq \mathbf{V}$ be pairwise distinct points, where $r \geq g$. Show that there is a non-special effective divisor of shape $D = \sum_{i=1}^{r} n_i \cdot (x_i) \in \mathrm{Div}_{\mathbf{V}}$, where $n_i \geq 0$, having degree $\deg(D) = g$.

Hint for c). Show that whenever D is effective such that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = 1$ and $\deg(D) < g$, that there is $i \in \{1, \ldots, r\}$ such that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D + (x_i))) = 1$.

(28.12) Exercise: Clifford's Theorem.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$.
a) Let $D, D' \in \mathrm{Div}_{\mathbf{V}}$ such that $\mathscr{L}(D) \neq \{0\} \neq \mathscr{L}(D')$. Show that

$$\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) + \dim_{\overline{\mathbb{F}}}(\mathscr{L}(D')) \leq 1 + \dim_{\overline{\mathbb{F}}}(\mathscr{L}(D + D')).$$

b) Let $D \in \mathrm{Div}_{\mathbf{V}}$ such that $\deg(D) \in \{0, \ldots, 2g-2\}$. Show that

$$\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) \leq 1 + \frac{1}{2} \cdot \deg(D).$$

(28.13) Exercise: Brill-Noether reciprocity.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$, let $D, D' \in \mathrm{Div}_{\mathbf{V}}$ such that $D + D'$ is canonical. Show that

$$\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) - \frac{1}{2} \cdot \deg(D) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(D')) - \frac{1}{2} \cdot \deg(D').$$

(28.14) Exercise: (Max) Noether's reduction theorem.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$, let $D \in \mathrm{Div}_{\mathbf{V}}$ such that $\mathscr{L}(D) \neq \{0\}$, let $(\omega) \in \mathrm{Div}_{\mathbf{V}}$ be canonical, and let $x \in \mathbf{V}$.

Show that we have $\dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega) - D - (x))) \neq \dim_{\overline{\mathbb{F}}}(\mathscr{L}((\omega) - D))$ if and only if $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D + (x))) = \dim_{\overline{\mathbb{F}}}(\mathscr{L}(D))$.

(28.15) Exercise: Bases of Riemann-Roch spaces.
Let $\mathbf{V}$ be a smooth projective curve, let $p \in \mathbf{V}$, let $n \in \mathbb{N}$, and let $\{f_1, \ldots, f_k\} \subseteq \mathscr{L}(n \cdot (p))$, for some $k \in \mathbb{N}$, such that the values $\{\nu_p(f_1), \ldots, \nu_p(f_k)\}$ are pairwise distinct, and for any $f \in \mathscr{L}(n \cdot (p))$ we have $\nu_p(f) \in \{\nu_p(f_1), \ldots, \nu_p(f_k)\}$.

Show that $\{f_1, \ldots, f_k\}$ is an $\overline{\mathbb{F}}$-basis of $\mathscr{L}(n \cdot (p))$.

(28.16) Exercise: Pole divisors.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$.
a) Let $x \in \mathbf{V}$. Show that for any $k \geq 2g$ there is $f \in \mathscr{K}_{\mathbf{V}}^*$ having pole divisor $(f)_\infty = k \cdot (x) \in \mathrm{Div}_{\mathbf{V}}$.
b) Let $\{x_1, \ldots, x_r\} \subseteq \mathbf{V}$ be pairwise distinct. Show that there is $f \in \mathscr{K}_{\mathbf{V}} \setminus \overline{\mathbb{F}}$ having pole divisor of shape $(f)_\infty = \sum_{i=1}^{r} n_i \cdot (x_i) \in \mathrm{Div}_{\mathbf{V}}$, where $n_i > 0$.

(28.17) Exercise: Morphisms of curves.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$. Show that there is a non-constant morphism $f\colon \mathbf{V} \to \mathbf{P}^1$ of degree $\deg(f) \leq g + 1$.

(28.18) Exercise: Canonical morphisms.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 2$, and let $D \in \mathrm{Div}_{\mathbf{V}}$ be canonical; then we have $\dim_{\overline{\mathbb{F}}}(\mathscr{L}(D)) = g$.

a) Let $\{f_1, \ldots, f_g\} \subseteq \mathscr{L}(D)$ be an $\overline{\mathbb{F}}$-basis. Show that there is a morphism

$$\varphi\colon \mathbf{V} \to \mathbf{P}^{g-1}\colon x \mapsto [f_1(x)\colon \ldots \colon f_g(x)].$$

Let $\{f_1', \ldots, f_g'\} \subseteq \mathscr{L}(D)$ be an $\overline{\mathbb{F}}$-basis, and $\varphi'\colon \mathbf{V} \to \mathbf{P}^{g-1}$ be the associated morphism. Show that there is an automorphism α of $\mathbf{P}^{g-1}$ such that $\varphi' = \alpha \circ \varphi$.

b) Any map φ as above is called a **canonical morphism**. We distinguish the cases whether φ is injective or not:

If φ is non-injective, show that there is a non-constant morphism $\mathbf{V} \to \mathbf{P}^1$ of degree 2; hence in this case $\mathbf{V}$ is a **hyperelliptic curve**.

If φ is injective, show that φ is a **closed immersion**, that is $\varphi(\mathbf{V}) \subseteq \mathbf{P}^{g-1}$ is closed, and $\varphi\colon \mathbf{V} \to \varphi(\mathbf{V})$ is an isomorphism.

(28.19) Exercise: Pure Picard groups.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$. Show that the pure Picard group $\mathrm{Pic}^0(\mathbf{V})$ is trivial if and only if $g = 0$, that is $\mathbf{V} \cong \mathbf{P}^1$.

(28.20) Exercise: Curves of genus 0.
Let $\mathbf{V}$ be a smooth projective curve of genus $g = g_{\mathbf{V}} \geq 0$.

a) If $g \geq 1$, show that for any $x, y \in \mathbf{V}$ we have $(x) \sim (y)$ if and only if $x = y$. Does this also hold for $g = 0$?

b) For $g = 0$, show that any two divisors on $\mathbf{V}$ of degree 1 are linearly equivalent.

c) Show that the following are equivalent: **i)** We have $g = 0$. **ii)** There is $f \in \mathscr{K}_{\mathbf{V}}^*$ such that $\deg((f)_0) = 1$. **iii)** There is $x \in \mathbf{V}$ such that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}((x))) \geq 2$.

(28.21) Exercise: Projective lines.
We consider the projective line $\mathbf{P}^1 = \mathbf{A}^1 \dot\cup \{\infty\}$, whose field of rational functions $\mathscr{K}_{\mathbf{P}^1}$ is naturally identified with $\overline{\mathbb{F}}(X)$.

a) Show that there is a unique $\omega \in \Omega_{\mathscr{K}_{\mathbf{P}^1}}$ such that $(\omega) = -2 \cdot (\infty) \in \mathrm{Div}_{\mathbf{P}^1}$ and $\mathrm{res}_\infty(\frac{1}{X} \cdot \omega) = -1$. Moreover, for $a \in \overline{\mathbb{F}}$ and $k \in \mathbb{Z}$ determine $\mathrm{res}_\infty((X-a)^k \cdot \omega)$ and $\mathrm{res}_a((X-a)^k \cdot \omega)$. Finally, determine $\Omega(-2 \cdot (\infty))$.

b) Show that X^{-1} is not the derivative of any element of $\overline{\mathbb{F}}(X)$; if $\mathrm{char}(\overline{\mathbb{F}}) = p > 0$ show that X^{p-1} is not the derivative of any element of $\overline{\mathbb{F}}(X)$. Conclude that there is a differential form which is not of the form df, for any $f \in \overline{\mathbb{F}}(X)$. Moreover, for $f \in \overline{\mathbb{F}}[X]$ irreducible determine the divisors (df) and $(f dX)$.

(28.22) Exercise: Cubic curves.
a) For $\lambda \in \overline{\mathbb{F}}$ let $\mathbf{V}_\lambda := \mathbf{V}(f_\lambda) \subseteq \mathbf{P}^2$ be the plane projective curve given by

$$f_\lambda := Y^2Z - X(X-Z)(X-\lambda Z) \in \overline{\mathbb{F}}[X, Y, Z].$$

For which $\lambda \in \overline{\mathbb{F}}$ is $\mathbf{V}_\lambda$ smooth? (Then, by Plücker's formula, $\mathbf{V}_\lambda$ has genus 1.)

b) Let $\mathbf{V}_\lambda$ be smooth. Then in affine coordinates we have $\mathscr{K}_{\mathbf{V}} \cong \mathrm{Q}(\mathscr{A})$, where

$$\mathscr{A} := \overline{\mathbb{F}}[X,Y]/\langle Y^2 - X(X-1)(X-\lambda)\rangle).$$

Determine the principal divisors (X) and (Y).

c) Show that $\mathscr{L} := \mathscr{L}(k \cdot (X)_\infty) \subseteq \mathscr{A}$ and that $\dim_{\overline{\mathbb{F}}}(\mathscr{L}) = 2k$, for $k \in \mathbb{N}$.

d) Ler $\omega := \frac{1}{Y} \cdot dX \in \Omega_{\mathscr{K}_{\mathbf{V}}}$. Show that $(\omega) = 0 \in \mathrm{Div}_{\mathbf{V}}$.

(28.23) Exercise: A quintic curve.
Let $\mathbf{V} = \mathbf{V}(f)$ be the plane projective curve given by

$$f := X^4Y + Y^4Z + Z^4X \in \mathbb{F}_2[X,Y,Z].$$

Show that $\mathbf{V}$ is smooth, determine its genus, and compute the principal divisor associated with $\frac{X}{Y} \in \mathscr{K}_{\mathbf{V}}$.

(28.24) Exercise: Goppa codes.
Let $\mathbf{V}$ be a smooth projective curve defined over $\mathbb{F}_q$, let $\{p_1, \ldots, p_n\} \subseteq \mathbf{V}(\mathbb{F}_q)$ be pairwise distinct, where $n \in \mathbb{N}$, and let $P := \sum_{i=1}^n (p_i) \in \mathrm{Div}_{\mathbf{V}}$; note that we do not specify the order of the places given.

a) Let $G \sim G' \in \mathrm{Div}_{\mathbf{V}}$ be $\mathbb{F}_q$-rational divisors whose support is disjoint to $\{p_1, \ldots, p_n\}$. Show that $\mathcal{G}_P(G)$ and $\mathcal{G}_P(G')$ are $\mathbb{F}_q$-linearly equivalent, and that so are $\mathcal{G}_P^*(G)$ and $\mathcal{G}_P^*(G')$.

b) Conversely, let G be an $\mathbb{F}_q$-rational divisor whose support is disjoint to $\{p_1, \ldots, p_n\}$, and let $\mathcal{C} \leq \mathbb{F}_q^n$ be $\mathbb{F}_q$-linearly equivalent to $\mathcal{G}_P(G)$. Show that there is an $\mathbb{F}_q$-rational divisor G' whose support is disjoint to $\{p_1, \ldots, p_n\}$, such that $G' \sim G$ and $\mathcal{C} = \mathcal{G}_P(G')$. Show the analogous assertion for $\mathcal{G}_P^*(G)$.

(28.25) Exercise: Goppa codes from elliptic curves.
Let $\mathbb{F}_4 = \mathbb{F}_2(\zeta)$, where ζ is a primitive 3-rd root of unity, and let $\mathbf{V} = \mathbf{V}(f)$ be the plane projective curve given by

$$f := X^2Y + \zeta Y^2Z + \zeta^2 Z^2X \in \mathbb{F}_4[X,Y,Z].$$

a) Show that $\mathbf{V}$ is smooth, determine its genus, and compute $\mathbf{V}(4)$.

b) Let $q := [\zeta\colon 1\colon 1] \in \mathbf{V}$ and $q' := [1\colon \zeta\colon 1] \in \mathbf{V}$ and $q'' := [1\colon 1\colon \zeta] \in \mathbf{V}$. Show that the divisor $G := 2(q) + (q') \in \mathrm{Div}_{\mathbf{V}}$ is $\mathbb{F}_4$-rational, and compute an $\mathbb{F}_4$-basis of $\mathscr{L}_{\mathbb{F}_4}(G) \subseteq \mathscr{K}_{\mathbf{V}}$.

c) Let $P := \sum_{p \in \mathbf{V}(4) \setminus \{q,q',q''\}} (p) \in \mathrm{Div}_{\mathbf{V}}$. Show that the geometric Goppa code $\mathcal{G}_P(G)$ over $\mathbb{F}_4$ has length 6, dimension 3, designed distance 3, but actually minimum distance 4. (Thus $\mathcal{G}_P(G)$ is an MDS code, being called the **hexacode**.)

(28.26) Exercise: Goppa codes from Hermitian curves.
Let $\mathbf{V} = \mathbf{V}(f)$ be the plane projective curve given by

$$f := X^{q+1} + Y^{q+1} + Z^{q+1} \in \mathbb{F}_{q^2}[X, Y, Z].$$

a) Show that $\mathbf{V}$ is smooth, and that $g_{\mathbf{V}} = \frac{1}{2}(q^2 - q)$. Moreover, show that $|\mathbf{V}(q^2)| = q^3 + 1$. How does this compare to the Hasse-Weil bound?

b) Now let q be even, let $p_0 := [0\colon 1\colon 1] \in \mathbf{V}$, let $P := \sum_{p \in \mathbf{V}(q^2) \setminus \{p_0\}}(p) \in \mathrm{Div}_{\mathbf{V}}$, and let $G := m \cdot (p_0) \in \mathrm{Div}_{\mathbf{V}}$, where $m \in \{q^2 - q - 1, \ldots, q^3\}$. Determine the length, the dimension and the designed distance of the geometric Goppa code $\mathcal{G} := \mathcal{G}_P(G)$ over $\mathbb{F}_{q^2}$.

c) Finally, let $q := 4$, and choose m such that $\mathcal{G}$ has rate $\rho(\mathcal{G}) = \frac{1}{2}$. Compute a generating matrix of $\mathcal{G}$ and determine its minimum distance. Is $\mathcal{G}$ an MDS code? Is $\mathcal{G}$ self-dual?

d) Suppose $\mathcal{G}$ is used to transit data through a symmetric binary channel with error probability $0 < \epsilon < 1$, by writing the elements of $\mathbb{F}_{16}$ as words of length 4 over $\mathbb{F}_2$. Determine the error probability of $\mathcal{G}$, by treating non-uniquely decodable words as erasures. Compare with the error probability of a Reed-Solomon code $\mathcal{R}$ over $\mathbb{F}_{q^2}$ being chosen such that $\rho(\mathcal{R}) \sim \frac{1}{2}$.

(28.27) Exercise: Goppa codes from the Fermat curve.
Let $\mathbf{V} = \mathbf{V}(f)$ be the smooth plane projective curve given by

$$f := X^3 + Y^3 + Z^3 \in \mathbb{F}_2[X, Y, Z].$$

a) Show that $|\mathbf{V}(8) \setminus \mathbf{V}(2)| = 6$, consisting of two $\langle\varphi_2\rangle$-orbits of length 3.

b) Let $P := \sum_{p \in \mathbf{V}(4)}(p) \in \mathrm{Div}_{\mathbf{V}}$, and let $G := (q) + (q^{\varphi_2}) + (q^{\varphi_2^2}) \in \mathrm{Div}_{\mathbf{V}}$, where $q \in \mathbf{V}(8) \setminus \mathbf{V}(2)$. Show that the geometric Goppa code $\mathcal{G}_P(G)$ and the dual geometric Goppa code $\mathcal{G}_P^*(G)$ over $\mathbb{F}_4$ are well-defined, and determine their dimension and their designed distance.

c) Compute $\mathbb{F}_4$-bases of $\mathscr{L}_{\mathbb{F}_4}(G)$ and $\mathscr{L}_{\mathbb{F}_4}((\omega_{\mathbb{F}_4}) + P - G)$, and determine the minimum distance of $\mathcal{G}_P(G)$ and $\mathcal{G}_P^*(G)$. Are $\mathcal{G}_P(G)$ or $\mathcal{G}_P^*(G)$ MDS codes?

(28.28) Exercise: Goppa codes from the Klein quartic.
Let $\mathbf{V} = \mathbf{V}(f)$ be the plane projective curve given by

$$f := X^3Y + Y^3Z + Z^3X \in \mathbb{F}_2[X, Y, Z].$$

a) Let $p_0 := [0\colon 0\colon 1] \in \mathbf{V}(2)$, and let $G_k := k \cdot (p_0) \in \mathrm{Div}_{\mathbf{V}}$ for $k \in \{0, \ldots, 18\}$. Compute an $\mathbb{F}_2$-basis of the function space $\mathscr{L}(G_k) \subseteq \mathscr{K}_{\mathbf{V}}$.

b) Let $P := \sum_{p \in \mathbf{V}(8) \setminus \{p_0\}}(p) \in \mathrm{Div}_{\mathbf{V}}$. Compute an $\mathbb{F}_8$-basis of the geometric Goppa code $\mathcal{G}_P(G_k)$, and determine its designed and minimum distances.

c) Let $\widetilde{P} := \sum_{p \in \mathbf{V}(16) \setminus \{p_0\}}(p) \in \mathrm{Div}_{\mathbf{V}}$. Determine $\mathbf{V}(16)$, compute an $\mathbb{F}_{16}$-basis of $\mathcal{G}_{\widetilde{P}}(G_k)$, and determine its designed and minimum distances.

(28.29) Exercise: Goppa codes from the Klein quartic.
Let $\mathbf{V} = \mathbf{V}(f)$ be the plane projective curve given by

$$f := X^3Y + Y^3Z + Z^3X \in \mathbb{F}_2[X, Y, Z].$$

a) Compute a differential form $\omega \in \Omega_{\mathbf{V}}$ and a canonical divisor $(\omega_{\mathbb{F}_2}) \in \mathrm{Div}_{\mathbf{V}}$.

b) Let $q_i := [\zeta_3^i \colon \zeta_3^{2i} \colon 1] \in \mathbf{V}(4)$ for $i \in \{1, 2\}$, where $\zeta_3 \in \mathbb{F}_4$ is a primitive 3-rd root of unity, let $G_k := k \cdot ((q_1) + (q_2)) \in \mathrm{Div}_{\mathbf{V}}$ for $k \in \{0, \ldots, 11\}$, and let $P := \sum_{p \in \mathbf{V}(8)} (p) \in \mathrm{Div}_{\mathbf{V}}$. Compute $\mathbb{F}_2$-bases of $\mathscr{L}_{\mathbb{F}_2}(G_k)$ and $\mathscr{L}_{\mathbb{F}_2}((\omega_{\mathbb{F}_2}) + P - G_k)$.

c) Show that the geometric Goppa code $\mathcal{G}_P(G_k)$ and the dual geometric Goppa code $\mathcal{G}_P^*(G_k)$ over $\mathbb{F}_8$ are well-defined, and determine their dimension and their designed and minimum distances.

---

29 References

[1] E. Assmus, J. Key: Designs and their codes, Cambridge Tracts in Mathematics 103, Cambridge University Press, 1992.

[2] E. Assmus, H. Mattson, R. Turyn: Research to develop the algebraic theory of codes, Air Force Cambridge Research Laboratories AFCRL-67-0365, 1967.

[3] A. Betten, H. Fripertinger, A. Kerber, A. Wassermann, K. Zimmermann: Codierungstheorie, Springer, 1998.

[4] A. Betten, M. Braun, H. Fripertinger, A. Kerber, A. Kohnert, A. Wassermann: Error-correcting linear codes, Classification by isometry and applications, Algorithms and Computation in Mathematics 18, Springer, 2006.

[5] E. Berlekamp: Algebraic coding theory, Mac-Graw Hill, 1968.

[6] W. Ebeling: Lattice and codes, 3rd edition, Advanced Lectures in Mathematics, Springer, 2013.

[7] D. Knuth: The art of computer programming, vol. 1: Fundamental algorithms, 3rd edition, Addison-Wesley, 1997.

[8] D. Knuth: The art of computer programming, vol. 3: Sorting and searching, 2nd edition, Addison-Wesley, 1998.

[9] F. MacWilliams, N. Sloane: The theory of error-correcting codes, North-Holland Mathematical Library 16, North-Holland, 1986.

[10] G. Nebe, E. Rains, N. Sloane: Self-Dual codes and invariant theory, Algorithms and Computation in Mathematics 17, Springer, 2006.

[11] O. Pretzel: Codes and algebraic curves, Oxford Lecture Series in Mathematics and Its Applications 8, Oxford Science Publications, 1998.

[12] E. Rains, N. Sloane: Self-dual codes, in: W. Huffman (ed.): Handbook of coding theory, vol. 1: Algebraic coding, Elsevier, 1998.

[13] C. Shannon: A mathematical theory of communication, The Bell System Technical Journal 27, 379–423, 623–656, 1948.

[14] H. Stichtenoth: Algebraic function fields and codes, Graduate Texts in Mathematics 254, Springer, 2009.

[15] D. Stinson: Cryptography, theory and practice, 3rd edition, CRC Press Series on Discrete Mathematics and its Applications 36, 2006.

[16] G. van der Geer, J. van Lint: Introduction to coding theory and algebraic geometry, DMV Seminar 12, Birkhäuser, 1988.

[17] J. van Lint: Introduction to coding theory, 2nd edition, Graduate Texts in Mathematics 86, Springer, 1991.

[18] W. Willems: Codierungstheorie, de Gruyter, 1999.