Teorie kódování

verze: 6. května 2024

Obsah

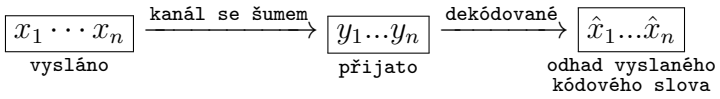
Kapitola 1

Úvod do samoopravných kódů

Pomocí nějakého zařízení chceme přenášet zprávy. Zařízením lze přenášet 0 a 1, ale toto zařízení není dokonalé - tzv. kanál se šumem. Občas dochází k záměně 0 na 1 a naopak. Předpokládáme, že při přenosu nastává záměna 0 za 1 a záměna 1 za 0 se stejnou pravděpodobností p (tzv. symetrický kanál). Bez újmy na obecnosti předpokládáme, že $p < \frac{1}{2}$, protože v případě $p > \frac{1}{2}$ lze interpretovat nulu jako jedničku a naopak. V případě $p = \frac{1}{2}$, nemá smysl vysílat zprávy.

Definice: Podmnožinu $\mathcal{C}$ množiny $\{0,1\}^n$ nazýváme kódem délky n. Prvek $x \in \mathcal{C}$ nazýváme kódovým slovem.

Zprávě, kterou chceme vyslat, přiřadíme prostým zobrazením kódové slovo x z $\mathcal{C}$ a tuto n-tici x vyšleme kanálem.



Na množině všech n-tic definujeme přirozenou metriku

Definice:Pro každé $x, y \in \{0,1\}^n,\ x = x_1 \cdots x_n,\ y = y_1 \cdots y_n$ klademe

$$\mathrm{dist}(x,y) = \#\{i : x_i \neq y_i\}\,.$$

Snadno lze ověřit, že dist je metrika na $\{0,1\}^n$, speciálně tedy platí trojúhelníková nerovnost $\mathrm{dist}(x,y) \leq \mathrm{dist}(x,z) + \mathrm{dist}(z,y)$.

Definice: Nechť $\mathcal{C} \subset \{0,1\}^n$ je kód a $M = \#\mathcal{C}$ počet jeho slov. Číslo

$$d = \min\{\mathrm{dist}(x,y) : x, y \in \mathcal{C}, x \neq y\}$$

nazveme minimální vzdálenost kódu $\mathcal{C}$. Říkáme, že $\mathcal{C}$ je (M,n,d)-kód.

Lemma: Nechť $x \in \mathcal{C} \subset \{0,1\}^n$ a $y \in \{0,1\}^n$ takové, že $\mathrm{dist}(x,y) = i$. Pak pravděpodobnost, že při přenosu kanálem bylo vysláno x a přijato y je rovna $p^i(1-p)^{n-i}$.

Poznámka: Při pevném $p < \frac{1}{2}$ a $n \in \mathbb{N}$ platí

$$i < j \implies p^i(1-p)^{n-i} > p^j(1-p)^{n-j}\,.$$

Tedy při přijatém y má největší pravděpodobnost vyslaného kódového slova takové $x \in \mathcal{C}$, kde je $\text{dist}(y, x)$ nejmenší.

Strategie dekódování: Nechť $y \in \{0,1\}^n$ je přijatá n-tice. Pak odhad vyslaného slova je

$$\hat{x} = \underset{x \in \mathcal{C}}{\text{argmin}}\, \text{dist}(x, y).$$

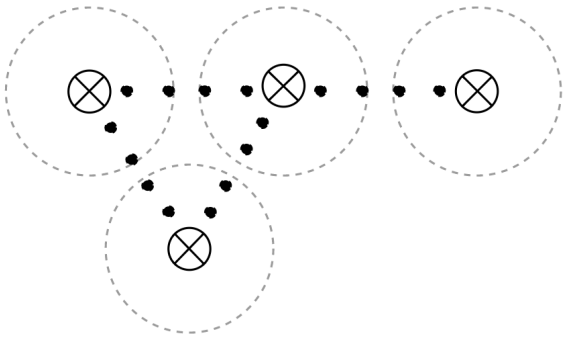
Poznámka: Pokud se minima nabývá na více $x \in \mathcal{C}$, zvolíme náhodně jedno z nich.

Věta: Kód s minimální vzdáleností d opravuje $\left\lfloor \frac{d-1}{2} \right\rfloor$ chyb, t.j. pokud při vysílání slova $x \in \mathcal{C}$ nastane chyba na i místech a $i \leq \left\lfloor \frac{d-1}{2} \right\rfloor$, pak $\hat{x} = x$.

Důkaz. Označme y přijatou n-tici, pak $i = \text{dist}(x, y)$. Z trojúhelníkové nerovnosti plyne pro každé $z \in \mathcal{C},\ z \neq x$

$$d \leq \text{dist}(x, z) \leq \underbrace{\text{dist}(x, y)}_{i} + \text{dist}(y, z) \leq \tfrac{d-1}{2} + \text{dist}(y, z).$$

Proto $\text{dist}(y, z) \geq \frac{d+1}{2} > \left\lfloor \frac{d-1}{2} \right\rfloor \geq \text{dist}(y, x)$. Proto naše dekódovací strategie vybere $\hat{x} = x$, viz Obr. 1.1. □



Obrázek 1.1:
$\otimes$ označuje prvky kódu $\mathcal{C}$
$\bullet$ označuje prvky $\{0,1\} \setminus \mathcal{C}$
$d = 5$

Věta (Hammingova) Nechť $\mathcal{C}$ je $(M, n, d)-$kód. Položme $t = \left\lfloor \frac{d-1}{2} \right\rfloor$. Pak platí

$$M\left(1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t}\right) \leq 2^n.$$

Důkaz. Uvažujme kouli o poloměru t se středem $x \in \mathcal{C}$

$$B(x, t) = \{y \in \{0,1\}^n \ :\ \text{dist}(x, y) \leq t\}.$$

1. Tvrzení: $B(x, t) \cap B(x', t) = \emptyset$ pro $x, x' \in \mathcal{C},\ x \neq x'$. Kdyby ne, pak by existovalo v průniku z a platilo by $d \leq \text{dist}(x, x') \leq \underbrace{\text{dist}(z, x')}_{\leq t} + \underbrace{\text{dist}(z, x)}_{\leq t} \leq 2t < d$, spor.

2. Tvrzení: Koule $B(x,t)$ má $1+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{t}$ prvků. Platí totiž, že počet n-tic z $\{0,1\}^n$, které se liší od x na i místech je roven:

$$\begin{aligned}&1, \ \texttt{pokud}\ i=0\\ &n, \ \texttt{pokud}\ i=1\\ &\binom{n}{2}, \ \texttt{pokud}\ i=2\\ &\texttt{atd.}\end{aligned}$$

□

V případě, že přenosovým kanálem lze vysílat q symbolů - jejich množinu budeme značit T, pak předchozí nerovnost má tvar

$$M\left(1+\binom{n}{1}(q-1)+\binom{n}{2}(q-1)^2+...+\binom{n}{t}(q-1)^t\right)\leq q^n.$$

Tato nerovnost se nazývá *Hammingova mez.*

Definice: Pokud v Hammingově mezi nastává rovnost, kód $\mathcal{C}$ se nazývá *perfektní.*

Příklad: Kód kontroly parity

$$\mathcal{C}=\left\{x_1x_2\cdots x_n\in\{0,1\}^n \ : \ \sum_{i=1}^{n}x_i=0 \mod 2\right\}.$$

Zřejmě $M=\#\mathcal{C}=2^{n-1}$ a $d=2$. Tento kód neopravuje žádnou chybu, umí detekovat chybu na 1 místě.

Příklad: Koktavý kód $\mathcal{C}=\{\underbrace{00\cdots0}_{n\text{-krát}},\underbrace{11\cdots1}_{n\text{-krát}}\}$. Zřejmě $M=2,\ d=n$, a tedy koktavý kód opravuje $\left\lfloor\frac{n-1}{2}\right\rfloor$ chyb. Pokud je n liché, $n=2l+1$, pak je koktavý kód perfektní. Platí totiž

$$2\left(1+\binom{n}{1}+\cdots+\binom{n}{l}\right)=\sum_{i=0}^{n}\binom{n}{i}=2^n.$$

Což je rovnost v Hammingově mezi.

Kapitola 2

Lineární kódy

2.1 Definice a základní pojmy

V této kapitole T označuje konečné těleso, nejčastěji to bude těleso $\mathbb{Z}_2$. T^n označuje prostor všech n-tic se složkami z T. Zřejmě

$$T^n = \{\alpha_1\alpha_2\cdots\alpha_n \ : \ \alpha_i \in T \ \texttt{pro každé} \ i = 1, 2, \ldots, n\}$$

s obvykle definovanými operacemi sčítání po složkách a násobením prvkem $c \in T$ taky po složkách je vektorovým prostorem dimenze n. Když $q = \#T$, pak T^n má q^n prvků.

Definice: Množina $\mathcal{C} \subset T^n$ se nazývá lineárním kódem délky n a dimenze k, pokud $\mathcal{C}$ je vektorovým prostorem dimenze $k \geq 1$.
Nechť

$$\begin{aligned}
\alpha^{(1)} &= \alpha_1^{(1)}\alpha_2^{(1)}\ldots\alpha_n^{(1)},\\
\alpha^{(2)} &= \alpha_1^{(2)}\alpha_2^{(2)}\ldots\alpha_n^{(2)}\\
&\vdots\\
\alpha^{(k)} &= \alpha_1^{(k)}\alpha_2^{(k)}\ldots\alpha_n^{(k)}
\end{aligned}$$

je báze $\mathcal{C}$. Matice

$$G = \begin{pmatrix}
\alpha_1^{(1)}\alpha_2^{(1)}\ldots\alpha_n^{(1)}\\
\alpha_1^{(2)}\alpha_2^{(2)}\ldots\alpha_n^{(2)}\\
\vdots\\
\alpha_1^{(k)}\alpha_2^{(k)}\ldots\alpha_n^{(k)}
\end{pmatrix} \in T^{k\times n}$$

se nazývá *generující matice* kódu $\mathcal{C}$.

Poznámka: Pokud těleso T má q prvků a lin. kód $\mathcal{C} \subset T^n$ má dimenzi k, pak $\#\mathcal{C} = q^k$.

Věta: Nechť G je generující matice lineárního kódu $\mathcal{C} \subset T^n$ dimenze k. Pak

$$\alpha_1\alpha_2\ldots\alpha_n \in \mathcal{C} \iff \exists\, \beta_1\beta_2\ldots\beta_k \in T^k \text{ takové že } \alpha_1\ldots\alpha_n = (\beta_1\ldots\beta_k)\,G.$$

Důkaz. Každý prvek vektorového prostoru je lineární kombinací prvků báze. □

Poznámka: Protože báze není daná jednoznačně, ani generující matice není dána jednoznačně.

Ověříme, že koktavý kód a kód kontroly parity jsou lineární kódy.

Příklad: *Koktavý kód*: $\mathcal{C} = \{00\ldots0, 11\ldots1\} \in T^n$, těleso $T = \mathbb{Z}_2$, $\dim\mathcal{C} = 1$, generující matice $G = (11\ldots1) \in T^{1\times n}$.

Kód kontroly parity: $\mathcal{C} = \{\alpha_1\ldots\alpha_n \in T^n \ : \ \sum_{i=1}^{n}\alpha_i = 0\}$ má $\dim\mathcal{C} = n-1$ a jeho generující matice je $G = \begin{pmatrix} 1100\ldots00 \\ 0110\ldots00 \\ 0011\ldots00 \\ \vdots \\ 0000\ldots11 \end{pmatrix} \in T^{(n-1)\times n}$.

Má-li kód $\mathcal{C}$ dimenzi k (neboli hodnost$(G) = k$), má jádro $\ker G \subset T^n$ dimenzi $n-k$. Uvážením libovolné báze podprosotoru $\ker G$ dostáváme následující tvrzení:

Věta: Nechť $\mathcal{C} \subset T^n$ je lineární kód dimenze k. Pak existuje matice $H \in T^{(n-k)\times n}$ taková, že $\mathcal{C}$ je řešením homogenní soustavy lineárních rovnic s maticí H. Tuto matici nazýváme *kontrolní matice* kódu $\mathcal{C}$. Dále platí, že

$$HG^\top = \Theta,$$

kde Θ označuje $(n-k)\times k$ matici složenou ze samých nul.

Příklad: Opakovací kód $\mathcal{C} = \{00\ldots0, 11\ldots1\}$ má kontrolní matici $H = \begin{pmatrix} 1100\ldots00 \\ 0110\ldots00 \\ 0011\ldots00 \\ \vdots \\ 0000\ldots11 \end{pmatrix}$.

Kód kontroly parity má kontrolní matici $H = (111\ldots11)$.

Příklad: 2.1: Vytvořme $H \in T^{3\times7}$ tak, že každá nenulové trojice $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \in T^3$ bude sloupcem H, tedy např.

$$H = \begin{pmatrix} 1110100 \\ 1101010 \\ 1011001 \end{pmatrix}.$$

Kód s kontrolní maticí H má dimenzi 4. Bázi tvoří např. 1000111, 0100110, 0010101, 0001011.

$G = \begin{pmatrix} 1000111 \\ 0100110 \\ 0010101 \\ 0001011 \end{pmatrix}$ tzv. *Hammingův binární kód* délky 7.

Označme wt(x) = počet nenulových složek v n-tici $x \in \mathcal{C} \subset T^n$. Číslo wt$(x)$ se nazývá váha slova x.

Věta: Nechť $\mathcal{C} \subset T^n$ je lineární kód. Pro minimální vzdálenost kódu $\mathcal{C}$ platí

$$d = \min\{\mathrm{wt}(z) \ : \ z \in \mathcal{C},\ z \neq 0\}.$$

Důkaz. Zřejmě $\mathrm{dist}(x, y) = \mathrm{wt}(x - y)$. Protože $\mathcal{C}$ je lineární kód, je $z = x - y \in \mathcal{C}$ a $z \neq 0$ pro každé $x, y \in \mathcal{C},\ x \neq y$. □

Věta: Nechť $\mathcal{C} \subset T^n$ je lineární kód s minimální vzdáleností d a H jeho kontrolní matice. Potom existuje d sloupců matice H, které jsou lineárně závislé (LZ), a naopak každých $d-1$ sloupců matice H je lineárně nezávislých (LN). Jinými slovy,

$$d = \min\left\{k \geq 1 : \exists k \text{ lineárně závislých sloupců } H\right\}.$$

Důkaz. Na základě předchozí věty si stačí uvědomit následující dva fakty.

1. Existuje $x = x_1 x_2 \cdots x_n \in T^n$ takové, že $\mathrm{wt}(x) = d$ a $Hx^\top = \vec{0}$. To znamená, že příslušných d sloupců matice H jejichž indexy i splňují $x_i \neq 0$ jsou LZ.
2. Pro každé nenulové $y \in T^n$ s $\mathrm{wt}(y) \leq d-1$ platí, že $y \notin \mathcal{C}$ a tudíž $Hy^\top \neq \vec{0}$. Jinými slovy, každých $d-1$ sloupců H je LN.

□

Důsledek: Je-li H kontrolní maticí kódu se vzdáleností d, tak $\mathrm{hodnost}(H) \geq d - 1$.

Příklad: Hammingův kód z příkladu 2.1 s kontrolní maticí $H \in T^{3\times 7}$ má všechny sloupce různé, nenulové, t.j. žádné dva sloupce matice H nejsou LZ. Součet 1.,4. a 5. sloupce je roven $\ominus$, tedy 1.,4. a 5. sloupec jsou LZ. Tedy tento Hammingův kód má min. vzdálenost $d = 3$ a opravuje 1 chybu.
Délka kódu je 7, dimenze 4. Proto $\#\mathcal{C} = 2^4 = 16$. V Hammingově mezi nastává rovnost $2^4\Big(\underbrace{1 + \binom{7}{1}}_{2^3}\Big) \leq 2^7$. Je to perfektní kód.

Definice: Nechť T je těleso o q prvcích, $m \in \mathbb{N}$. Nechť H je matice tvořena všemi nenulovými sloupci $\in T^m$, které mají navíc vlastnost, že první nenulová složka ve sloupci je rovna 1. Kód, jehož kontrolní matice je H, se nazývá $(q, m)-$Hammingův kód.

Příklad: Hammingův kód z příkladu 2.1 je $(2, 3)-$Hammingův kód.

Příklad: $(3, 3)-$Hammingův kód je vektorový podprostor nad tělesem $T = \mathbb{Z}_3$.

$$H = \begin{pmatrix} 0000111111111 \\ 0111000111222 \\ 1012012012012 \end{pmatrix}.$$

Tvrzení: Pro $m \in \mathbb{N}$ a $\#T = q$ má kontrolní matice $(q, m)-$Hammingova kódu $n = \frac{q^m-1}{q-1}$ sloupců. Kód má dimenzi $k = \frac{q^m-1}{q-1} - m$ a minimální vzdálenost kódových slov $d = 3$.

Důkaz. Z výše popsané konstrukce okamžitě plynou deklarované hodnoty n a k. Zbývá tedy ověřit, že každé dva sloupce kontrolní matice H jsou lineárně nezávislé.

Předpokládejme, že H obsahuje dva sloupce $s_1 \neq s_2$ takové, že $\alpha s_1 + \beta s_2 = 0$ pro nějaké $\alpha, \beta \in T$. Naším cílem je ukázat, že $\alpha = \beta = 0$. Označme i_1 resp. i_2 indexy první nenulové souřadnice v s_1 resp. s_2, a bez újmy na obecnosti předpokládejme, že $i_1 \geq i_2$. Je-li $i_1 > i_2$, tak potom i_1-ní souřadnice $\alpha s_1 + \beta s_2$ je rovna α, a proto nutně $\alpha = 0$. Protože sloupce H jsou nenulové, musí i $\beta = 0$. Zbývá tedy uvážit $i_1 = i_2$. V tom případě však nulový součet na i_1-ní souřadnici dává $\alpha + \beta = 0$, neboli $\alpha = -\beta$. Získáváme tak, že $\alpha(s_1 - s_2) = 0$, což je možné jen pokud $\alpha = 0$. □

Věta: Nechť $m \in \mathbb{N}$, $\#T = q$. Pak $(q, m)-$Hammingův kód je perfektní.

Důkaz. Dosadíme do Hammingovy nerovnosti pro $t = \left\lfloor \frac{d-1}{2} \right\rfloor = 1$ vidíme, že v

$$\underbrace{q^k}_{q^{n-m}} \left(1 + \underbrace{\binom{n}{1}(q-1)}_{q^m - 1} \right) \leq q^n$$

nastává rovnost. □

Definice: Lineární kód $\mathcal{C} \subset T^n$ dimenze k se nazývá *systematický*, pokud pro každé $\alpha_1 \dots \alpha_k \in T^k$ existuje právě jeden $\alpha_{k+1} \dots \alpha_n \in T^{n-k}$ tak, že $\alpha_1 \dots \alpha_n \in \mathcal{C}$.

Příklad: Koktavý kód je systematický, kód kontroly parity je systematický, zmiňovaný Hammingův kód délky 7 je systematický.

Věta: Nechť $\mathcal{C} \subset T^n$ je systematický lineární kód dimenze k. Pak generující a kontrolní matice lze najít ve tvaru:

$$G = (I_k | F) \quad \text{a} \quad H = (-F^\top | I_{n-k}), \quad \text{kde } F \in T^{k \times (n-k)}.$$

Důkaz. Protože $\overbrace{0 \dots 0 \underbrace{1}_{i\text{-tá pozice}} 0 \dots 0}^{\text{délka } k}$ lze jednoznačně doplnit na $n-$tici tak, aby byla v $\mathcal{C}$, tvoří k takto získaných prvků z kódu $\mathcal{C}$ bázi, která napsaná do řádků pod sebe tvoří matici G uvedenou ve větě.
Stačí ověřit, že H v tvrzení věty splňuje vztah

$$H \cdot G^\top = (-F^\top | I_{n-k}) \begin{pmatrix} I_k \\ F^\top \end{pmatrix} = -F^\top + F^\top = \Theta\,.$$

□

Definice: Nechť $\mathcal{C}_1, \mathcal{C}_2 \subset T^n$ jsou lineární kódy. Řekneme, že jsou *ekvivalentní*, pokud existuje permutace indexů $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$ taková, že

$$\alpha_1 \alpha_2 \dots \alpha_n \in \mathcal{C}_1 \quad \Longleftrightarrow \quad \alpha_{\pi(1)} \alpha_{\pi(2)} \dots \alpha_{\pi(n)} \in \mathcal{C}_2.$$

Věta: Každý lineární kód je ekvivalentní nějakému systematickému kódu.

Důkaz. Z lineární algebry víme: když $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(k)}$ je LN soubor vektorů, pak tento soubor zůstane LN po libovolné z těchto operací:

a) změníme libovolné pořadí vektorů v souboru

b) libovolný $\alpha^{(i)}$ vynásobíme číslem $c \in T,\ c \neq 0$

c) $\alpha^{(i)}$ nahradíme vektorem $\alpha^{(i)} + c \cdot \alpha^{(j)}$

To znamená, že pokud $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(k)}$ je báze kódu $\mathcal{C}$ a generující matice $G = \begin{pmatrix} \alpha^{(1)} \to \\ \alpha^{(2)} \to \\ \vdots \\ \alpha^{(k)} \to \end{pmatrix}$, pak provázením úprav typu a) b) c) na řádkové vektory dostaneme jinou generující matici stejného kódu. Úpravy a) b) c) provádíme tak dlouho až dostaneme generující matici v horním stupňovitém tvaru. Označme $i_1, i_2, \ldots, i_k$ indexy hlavních sloupců a definujme permutaci

$$\pi : 1 \to i_1, 2 \to i_2, \ldots, k \to i_k$$

a $\pi(k+1), \ldots, \pi(n)$ libovolně tak, aby π byla permutace. Položme

$$\mathcal{C}' = \{\alpha_{\pi(1)}\alpha_{\pi(2)} \ldots \alpha_{\pi(n)} \ :\ \alpha_1\alpha_2 \ldots \alpha_n \in \mathcal{C}\}.$$

Pak generující matici kódu $\mathcal{C}'$ získáme permutací sloupců původní generující matice G. Dostaneme G' ve tvaru $G' = (U|\texttt{něco})$, kde $U \in T^{k\times k}$ je čtvercova horní trojúhelníková matice s nenulovými prvky na diagonále. Tato matice je generující maticí kódu $\mathcal{C}'$. Abychom ukázali, že $\mathcal{C}'$ je systematický, převedeme opět řádkovými úpravami a), b), c) matici G' do tvaru $G'' = (I_k|\texttt{něco})$. Tato matice je taky generující maticí kódu $\mathcal{C}'$, a proto podle předchozí věty je $\mathcal{C}'$ systematický. □

2.2 Vztahy mezi parametry lineárního kódu

Hammingova mez svazuje počet kódových slov M, délku kódu n a minimální vzdálenost kódu d pro libovovlný kód $\mathcal{C} \subset T^n$. V případě, že $\mathcal{C}$ je lineární kód, lze pro parametry n, d a $M = q^k$, kde $k = \dim \mathcal{C}$, nalézt přesnější meze.

Věta (Singletonova nerovnost): Nechť $\mathcal{C} \subset T^n$ je lineární kód dimenze k s minimální vzdáleností d. Pak

$$n + 1 \geq k + d.$$

Důkaz. Z Frobeniovy věty platí, že $\dim \mathcal{C} = k = n - \text{hodnost}(H)$. Dále z předchozí kapitoly víme, že každých $d-1$ sloupců H je lineárně nezávislých. Speciálně tedy

$$d - 1 \leq \text{hodnost}(H) = n - k.$$

□

Věta (Gilbert-Varshamova mez pro binární kódy): Buď $T = \mathbb{Z}_2$, a nechť $r, n, d \in \mathbb{N}$ splňují

$$1 + \binom{n-1}{1} + \binom{n-1}{2} + \cdots + \binom{n-1}{d-2} < 2^r \,. \tag{2.1}$$

Pak existuje kontrolní matice $H \in T^{r\times n}$ taková, že lineární binární kód s kontrolní maticí H má minimální vzdálenost $\geq d$.

Důkaz. Již víme, že k důkazu stačí nalézt matici H takovou, že každých $d-1$ sloupců H je lineárně nezávislých nad $\mathbb{Z}_2$. To uděláme tak, že budeme po jednom vybírat celkem n vektorů z $\mathbb{Z}_2^r$, které budou tvořit sloupce hledané matice H, tak, aby každých $d-1$ z nich bylo LN.

Jako první vektor vybereme libovolnou nenulovou r-tici z 0 a 1. Nyní předpokládejme, že již máme vybráno i vektorů, pro nějaké $i \leq n-1$, takových, že každých $d-1$ z nich je LN. Naším cílem je vybrat $(i+1)$-ní vektor a stále zachovat vlastnost, že libovolných $d-1$ vektorů je LN. Nemůžeme vybrat vektor $\vec{0}$, ani žádný z již vybraných i vektorů, ani součet dvou již vybraných, ..., a ani součet $d-2$ již vybraných vektorů. Avšak, to jsou jediná omezení pro $(i+1)$-ní vektor, a tudíž ze zbývajících vektorů v $\mathbb{Z}_2^r$, zbyly-li nějaké, můžeme vybrat libovolný z nich.

Celkový počet zakázaných možností pro výběr $(i+1)$-ního vektoru je nejvýše

$$1 + \binom{i}{1} + \binom{i}{2} + \cdots + \binom{i}{d-2} \leq 1 + \binom{n-1}{1} + \binom{n-1}{2} + \cdots + \binom{n-1}{d-2}.$$

Protože je pravá strana dle (2.1) ostře menší než 2^r, existuje požadovaný $(i+1)$-ní vektor. □

Při dané dimenzi kódu k a počtu chyb, které chceme opravovat, je žádoucí minimalizovat délku kódových slov n.

Nechť $k, d \in \mathbb{N}$. Označme

$$N(k,d) := \min\{n \in \mathbb{N} \,:\, \texttt{existuje lineární binární kód délky } n, \\ \texttt{dimenze } k \texttt{ a minimální vzdálenosti } d\}.$$

Lze si snadno rozmyslet, že $N(k,1) = k$ a $N(1,d) = d$ pro každé $k, d \in \mathbb{N}$. Též je zřejmé, že funkce $N(k,d)$ je neklesající v paramteru d, tj. $N(k,d_1) \leq N(k,d_2)$, pro každé $k \in \mathbb{N}$ a $d_1 \leq d_2$. Podobně platí, že $N(k_1,d) \leq N(k_2,d)$, pro všechna $d \in \mathbb{N}$ a $k_1 \leq k_2$.

Lineární kód $\mathcal{C}$, pro který se nabývá minimální délky $N(k,d)$, není dán jednoznačně, a např. každý kód vzniklý z $\mathcal{C}$ libovolnou permutací jeho složek je opět kódem, kde se nabývá minima délky.

Věta: Pro $k, d \in \mathbb{N}, k \geq 2$ platí $N(k,d) \geq d + N(k-1, \lceil \frac{d}{2} \rceil)$.

Důkaz. Označme $n = N(k,d)$ a nechť $\mathcal{C}$ je linenární binární kód délky n, dimenze k a minimální vzdálenosti d. V tomto kódu existuje slovo s váhou d. Bez újmy na

obecnosti pozice jsou zpermutovány tak, že toto slovo je tvaru $\underbrace{00\ldots0}_{n-d}\underbrace{11\ldots1}_{d}$. Toto slovo dáme do báze $\mathcal{C}$ a doplníme dalšími $k-1$ vektory kódu a vytvoříme generující matici

$$G = \left(\begin{array}{c|c} 00\ldots0 & 11\ldots1 \\ \hline \underbrace{G_1}_{n-d} & \texttt{něco} \end{array}\right)\Bigg\}k = \dim\mathcal{C}.$$

Uvažujme lineární kód s generující maticí G_1.

- Hodnost G_1 je $k-1$: jinak bychom mohli nakombinovat řádky tak, že první řádek G_1 je roven $\underbrace{0\ldots0}_{n-d}$ a jeho prodloužení na rozměr n je kódové slovo původního kódu $\mathcal{C}$. Nemůže to být nulové slovo délky n (to by hodnost$(G) < k$). Tedy má tvar $\underbrace{0\ldots0}_{n-d}\,\underbrace{\texttt{něco nenulové}}_{d} \in \mathcal{C}$. To ale implikuje, že vzdálenost od 1. řádku celé matice G je $< d$ - spor.
- Označme d_1 minimální vzdálenost kódu s generující maticí G_1. V tomto kódu existuje slovo $u \in \{0,1\}^{n-d}$ s váhou $\mathrm{wt}(u) = d_1$. Prodloužení u v původním kódu je tvaru $uv \in \mathcal{C}$. Protože $\underbrace{0\ldots0}_{n-d}\underbrace{1\ldots1}_{d} \in \mathcal{C}$ a $\mathcal{C}$ je lineární, platí

$$\underbrace{u}_{n-d}\,\underbrace{v}_{d} + \underbrace{00\ldots0}_{n-d}\underbrace{11\ldots1}_{d} = u\overline{v} \in \mathcal{C},$$

  kde $\overline{v}$ značí slovo vzniklé ze slova v záměnou $0 \leftrightarrow 1$. Proto

$$\left.\begin{array}{rcrcl} \mathrm{wt}(u) & + & \mathrm{wt}(v) & \geq & d \\ \mathrm{wt}(u) & + & d - \mathrm{wt}(v) & \geq & d \\ \hline 2\mathrm{wt}(u) & & & \geq & d \end{array}\right\} \quad\Longrightarrow\quad d_1 = \mathrm{wt}(u) \geq \frac{d}{2}.$$

Kód s generující matici G_1 má dimenzi $k-1$, minimální vzdálenost $d_1 \geq \left\lceil \frac{d}{2} \right\rceil$ a délku $n-d$, tudíž

$$N(k,d) - d = n - d \geq N(k-1, d_1) \geq N\left(k-1, \left\lceil \frac{d}{2} \right\rceil\right).$$

□

Důsledek (Griesmerova mez). $N(k,d) \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{2^i} \right\rceil$.

Důkaz. Matematickou indukcí na k dokazujeme výrok:

$$\text{Pro každé } d \in \mathbb{N} \text{ platí } N(k,d) \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{2^i} \right\rceil.$$

Jak už bylo zmíněno, $N(1,d) = d = \sum_{i=0}^{0} \left\lceil \frac{d}{2^i} \right\rceil$. Tvrzení tedy platí pro $k = 1$.

Nechť $k \geq 2$ a předpokládejme, že tvrzení je pravdivé pro všechny hodnoty menší než k. Pak podle předchozí věty a indukčního předpokladu

$$N(k,d) \geq d + N(k-1, \lceil \tfrac{d}{2} \rceil) \geq d + \sum_{i=0}^{k-2} \left\lceil \tfrac{1}{2^i} \left\lceil \tfrac{d}{2} \right\rceil \right\rceil \geq \left\lceil \tfrac{d}{2^0} \right\rceil + \sum_{i=0}^{k-2} \left\lceil \tfrac{d}{2^{i+1}} \right\rceil = \sum_{i=0}^{k-1} \left\lceil \tfrac{d}{2^i} \right\rceil .$$

V poslední nerovnosti jsme využili zřejmý vztah

$$\lceil \alpha \cdot \lceil y \rceil \rceil \geq \lceil \alpha \cdot y \rceil \quad \texttt{pro každé } y \in \mathbb{R} \texttt{ a } \alpha \geq 0.$$

(Lze nahlédnout, že v případě, kdy $\frac{1}{\alpha}$ je celé číslo, platí dokonce $\lceil \alpha \cdot \lceil y \rceil \rceil = \lceil \alpha \cdot y \rceil$.) □

2.3 Standardní tabulka pro dekódování

V celé této podkapitole se věnujeme pouze binárním kódům.

Chybový vektor. Když při přenosu kódového slova x délky n nastanou chyby, přijmeme vektor $y = x + e$, kde $e = e_1 \dots e_n$ je tzv. *chybový vektor*, pro který platí $e_i = 1$, pokud na $i-$tém místě nastala chyba a $e_i = 0$, pokud chyba nenastala.
Jeho váha $\mathrm{wt}(e)$ je tedy počet chybných pozic při přenosu.

Nechť $\mathcal{C}$ je lineární binární kód délky n a dimenze k. Vytvořme tzv. **standardní tabulku** následujícím způsobem:
Do nultého řádku N_0 tabulky napíšeme postupně všechny prvky kódu $\mathcal{C}$, t.j. 2^k slov. Pak vybereme $n-$tici a_1, která neleží v N_0 a další řádek tabulky N_1 bude tvořen $n-$ticemi $a_1 + x$, kde $x \in \mathcal{C}$. Řádek tabulky N_2 vznikne tak, že vybereme $a_2 \notin N_0 \cup N_1$ a do N_2 napíšeme všechny $n-$tice typu $a_2 + x,\ x \in \mathcal{C}$. Pro tvorbu N_3 opět vybereme $a_3 \notin N_0 \cup N_1 \cup N_2$, atd. Postupujeme tak dlouho, až každá $n-$tice leží v některém řádku tabulky. Máme tedy 2^{n-k} řádků.

Pozorování: Když vyšleme kódové slovo $x \in \mathcal{C}$ a nastane chyba e, pak přijatý vektor $y = x + e$ a chyba e leží ve stejném řádku tabulky.

Důkaz. Nechť y leží v $i-$tém řádku a e leží v $j-$tém. Pak

$$y = a_i + x_1, \texttt{ kde } x_1 \in \mathcal{C}$$
$$e = a_j + x_2, \texttt{ kde } x_2 \in \mathcal{C}$$

Odtud $a_i = y - x_1 = x + e - x_1 = a_j + \underbrace{x_2 + x - x_1}_{\in \mathcal{C}}$ a tedy a_i leží v $j-$tém řádku, proto z konstrukce plyne, že $a_i = a_j$. □

Pozorování: Nechť kód $\mathcal{C}$ má kontrolní matici H. Pak pro každé dvě n-tice y a z platí $Hy^\top = Hz^\top \quad \Leftrightarrow \quad y$ a z leží ve stejném řádku tabulky.

Důkaz. y a z leží ve stejném řádku $\quad \Leftrightarrow \quad y - z \in \mathcal{C} \quad \Leftrightarrow \quad H(y-z)^\top = \vec{0} \quad \Leftrightarrow$ $Hy^\top = Hz^\top$. □

Definice: $Hy^\top$ nazýváme *syndrom* vektoru y.

Standardní tabulku upravíme tak, že na první místo každého řádku umístíme ten vektor, který má nejmenší váhu a na konec řádku umístíme jeho syndrom.

Dekódování: Protože přijatý vektor y je roven $y = x + e$, musíme od přijatého y odečíst neznámé e, abychom získali skutečně vyslané slovo x. Víme, že e leží ve stejném řádku jako y. V tomto řádku, je 2^k kandidátů na možné e. Současně víme, že nejpravděpodobnější je chyba na nejmenším počtu míst, a proto za e prohlásíme vektor s nejmenším počtem jedniček, t.j. při naší organizaci tabulky vektor na prvním místě v řádku, v kterém leží y. Jak ale při přijetí y hledat řádek, ve kterém leží? Snadno pomocí $Hy^\top$. Z celé tabulky nás tedy zajímá pouze první sloupec (tam jsou vektory s nejmenší vahou v řádku) a poslední sloupec (tam jsou syndromy jednotlivých řádků). Zbytek tabulky se ve skutečnosti může vymazat.

Příklad: V tomto příkladě sestavíme standardní tabulku pro Hammingův kód s kontrolní maticí $H = \begin{pmatrix} 1001101 \\ 0101011 \\ 0010111 \end{pmatrix}$. Tedy délka $n = 7$ a dimenze $k = 4$. Proto každý řádek má $2^4 = 16$ slov a celá tabulka má $2^{7-4} = 8$ řádků.

0000000	1110001	0110010	1000011	1010100	0100101	1100110	0010111	000
1101000	0011001	1011010	0101011	0111100	1001101	0001110	1111111	
1000000	0110001	1110010	0000011	0010100	1100101	0100110	1010111	100
0101000	1011001	0011010	1101011	1111100	0001101	1001110	0111111	
0100000	1010001	0010010	1100011	1110100	0000101	1000110	0110111	010
1001000	0111001	1111010	0001011	0011100	1101101	0101110	1011111	
0010000	1100001	0100010	1010011	1000100	0110101	1110110	0000111	001
1111000	0001001	1001010	0111011	0101100	1011101	0011110	1101111	
0001000	1111001	0111010	1001011	1011100	0101101	1101110	0011111	110
1100000	0010001	1010010	0100011	0110100	1000101	0000110	1110111	
0000100	1110101	0110110	1000111	1010000	0100001	1100010	0010011	101
1101100	0011101	1011110	0101111	0111000	1001001	0001010	1111011	
0000010	1110011	0110000	1000001	1010110	0100111	1100100	0010101	011
1101010	0011011	1011000	0101001	0111110	1001111	0001100	1111101	
0000001	1110000	0110011	1000010	1010101	0100100	1100111	0010110	111
1101001	0011000	1011011	0101010	0111101	1001100	0001111	1111110	

Vyšleme-li kódové slovo $x = 1010100$ a nastane-li jedna chyba na 4. místě, přijmeme $y = 1011100$. (tj. chybový vektor $e = 000100$) Syndrom $Hy^\top = 110$, který odpovídá 4. řádku tabulky (tabulka začíná nultým řádkem). V tomto řádku skutečně y najdeme. Protože první vektor ve 4. řádku je 0001000, odečteme tento vektor od y a

dostaneme $\hat{x}$, které je v našem případě rovné skutečně vyslanému x. Tedy Hammingův kód tuto jednu chybu na 4. místě správně opravil (to jsme ale čekali, protože víme, že Hammingův kód má minimální vzdálenost $d = 3$, a tedy opravuje jednu chybu). Vyšleme-li stejné x a chyba nastane na 2. a 4. místě, přijmeme $y = 1111100$. Syndrom $Hy^\top$ je v tomto případě 100, což odpovídá 1. řádku, a tedy od y odečítáme předpokládaná chybový vektor 0000001. Dostaneme odhadované $\hat{x} = 1111101$, které není rovno vyslanému x, tj. Hammingův kód tuto chybu na dvou místech už neumí opravit.

Jak vidíme, ve skutečnosti z tabulky potřebujeme pouze poslední sloupec syndromů a první sloupec vektorů s minimální vahou.

Pozorování: Nechť binární kód $\mathcal{C}$ má minimální vzdálenost d. Označme $t = \left\lfloor \frac{d-1}{2} \right\rfloor$. Pak každý vektor e s váhou $\mathrm{wt}(e) \leq t$ je na prvním místě některého řádku standardní tabulky.

Důkaz. Sporem: kdyby e byl v řádku, kde na prvním místě sedí e'. Pak z toho, že $He = He' \Rightarrow H(e - e') = \vec{0} \Rightarrow e - e' \in \mathcal{C}$. Zřejmě, $\mathrm{wt}(e - e') \leq \mathrm{wt}(e) + \mathrm{wt}(e') \leq 2t \leq 2\left\lfloor \frac{d-1}{2} \right\rfloor \leq 2\frac{d-1}{2} = d - 1$. Tedy v $\mathcal{C}$ existuje slovo s váhou $< d$. Spor s definicí minimální vzdáleností. □

Pravděpodobnost chybného dekódování

Předpokládejme, že všechna kódová slova jsou vysílaná se stejnou hustotou a že dekódujeme podle stand. tabulky.
Pak pravděpodobnost chybného dekódování je zřejmě

$$P_{err} = Prob\{\texttt{chybový vektor } e \texttt{ se nerovná žádnému vektoru, který je v prvním sloupci standardní tabulky}\}.$$

Nechť $p < \frac{1}{2}$ je pravděpodobnost záměny $0 \leftrightarrow 1$ a nechť γ_i = počet vektorů v prvním sloupci tabulky, které mají váhu i. Pravděpodobnost, že se chybový vektor e shoduje s vektorem váhy i se rovná $p^i(1-p)^{n-i}$. Proto pravděpodobnost, že se e trefí do některého vektoru z 1. sloupce tabulky je $\sum_{i=0}^{n} \gamma_i p^i (1-p)^{n-i}$, a tedy

$$P_{err} = 1 - \sum_{i=0}^{n} \gamma_i p^i (1-p)^{n-i}\,.$$

Příklad: Binární Hammingovy kódy jsou perfektní, jejich délka je $n = 2^m - 1$, dimenze $k = 2^m - 1 - m$, mají minimální vzdálenost $d = 3$, odtud $t = 1$. Počet řádků ve standardní tabulce je roven 2^{n-k}. V našem případě to je $2^{n-k} = 2^m = n + 1$. Podle předchozího pozorování se každá n-tice s váhou $\leq t = 1$ vyskytne v prvním sloupci standardní tabulky. Tedy $\gamma_0 = 1$ (jediný vektor váhy 0) a $\gamma_1 = n$ (n vektorů s váhou 1). Těchto $n+1$ vektorů už vystačí na $n+1$ řádků standardní tabulky, tj. $\gamma_i = 0$,když $i \leq 2$. Pro Hammingovy kódy jsme tedy dovodili, že

$$P_{err} = 1 - (1-p)^n - np(1-p)^{n-1}\,.$$

Např. Hammingův kód $n = 7$ a dimenze $k = 4$ má při pravděpodobnosti záměny $0 \leftrightarrow 1$ rovné $p = \frac{1}{10}$.

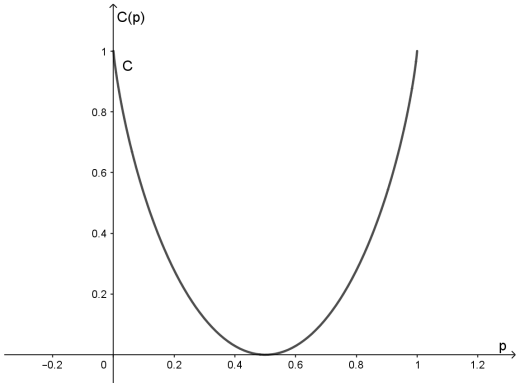
$$P_{err} = 1 - \left(\frac{9}{10}\right)^7 - 7\left(\frac{9}{10}\right)^6 \cdot \frac{1}{10} \doteq 0.1497$$

Kdyby se místo sedmic vysílaly původné čtveřice symbolů (kód má dimnezi 4), tak

$$P_{err} = 1 - (1-p)^4 \doteq 0.3439\,.$$

Shannon ukázal, že i při špatném přenosovém zařízení (kromě $p = \frac{1}{2}$) lze zkonstruovat lineární kód s docela dobrým poměrem mezi dimenzi a délkou, a přitom pravděpodobnost chybného dekódování je menší než předem zvolené libovolně malé ϵ. Vágní pojem "docela dobrý" nahradíme kapacitou binárního symetrického kanálu. Pro pravděpodobnost $p \in (0, 1)$ položme

$$C(p) = 1 + p\log_2(p) + (1-p)\log_2(1-p)\,.$$



Obrázek 2.1: Graf $C(p)$.

Následující věta je obsahem kurzu Teorie informace, a proto ji uvádíme bez důkazu.

Věta: (Shannonova) Pro každé $\epsilon > 0$ a každé $R < C(p)$ existuje lineární kód délky n a dimenze k takový, že $\frac{k}{n} \geq R$ a $P_{err} < \epsilon$.

Kapitola 3

Cyklické kódy

3.1 Generující a kontrolní polynomy

Opět vše omezeno na binární abecedu.

Definice: Lineární kód $\mathcal{C}$ délky n se nazývá *cyklický*, jestliže pro každé slovo $v = v_0v_1\dots v_{n-1} \in \mathcal{C}$ je také $v_{n-1}v_0v_1\dots v_{n-2} \in \mathcal{C}$. Kódová slova budeme zapisovat ve tvaru polynomů stupně $< n$, takže místo $v_0v_1\dots v_{n-1}$ budeme psát $v(z) = v_0 + v_1z + v_2z^2 + \dots + v_{n-1}z^{n-1}$.
Podmínku cykličnosti pak můžeme napsat ve tvaru

$$v(z) \in \mathcal{C} \;\Leftrightarrow\; z \odot v(z) \in \mathcal{C}, \tag{3.1}$$

kde operace $\odot$ znamená, že násobíme v okruhu polynomů $\mathbb{Z}_2[x]/x^n-1$, t.j. $z^n-1 = 0$.

Poznámka: Všimněme si, že z ekvivalence (3.1) plyne, že

$$v(z) \in \mathcal{C} \;\Leftrightarrow\; a(z) \odot v(z) \in \mathcal{C}, \texttt{ pro lib. polynom } a(z). \tag{3.2}$$

Úmluva: Budeme-li násobit a sčítat v okruhu $\mathbb{Z}_2[x]/x^n - 1$, budeme u polynomů používat proměnnou z. Budeme-li mít na mysli operace násobení a sčítání polynomů v obvyklém smyslu, budeme používat proměnnou x.

Příklad: Kód kontroly parity délky 4 je tvořen slovy:

$$0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111.$$

Je to cyklický kód. Pomocí polynomů jej lze zapsat

$$\mathcal{C} = \{0, z^2 + z^3, z + z^3, z + z^2, 1 + z^3, 1 + z^2, 1 + z, 1 + z + z^2 + z^3\}.$$

Všimněme si, že všechny prvky tohoto kódu jsou násobky polynomu $1+z$ polynomem stupně ≤ 2. (Násobíme v okruhu $\mathbb{Z}_2[x]/x^4 - 1$).

$$\begin{aligned}
0 &= 0 \cdot (z+1) & 1+z^3 &= (1 + z + z^2)(1+z)\\
z^2 + z^3 &= z^2(1+z) & 1+z^2 &= (1+z)(1+z)\\
z + z^3 &= (z + z^2)(1+z) & 1+z &= 1 \cdot (1+z)\\
z + z^2 &= z(1+z) & 1 + z + z^2 + z^3 &= (1+z^2)(1+z)
\end{aligned}$$

Věta: Každý cyklický kód $\mathcal{C}$ délky n a dimenze k obsahuje polynom $g(z)$ stupně $n-k$ s těmito vlastnostmi:

1. Kód $\mathcal{C}$ sestává ze všech násobků polynomu $g(z)$ v $\mathbb{Z}_2[x]/_{x^n-1}$, t.j.

$$\mathcal{C} = \{a(z)g(z) \ : \ a(z) \in \mathbb{Z}_2[x]/x^n - 1\}.$$

2. Polynomy $g(z), z\,g(z), \ldots, z^{k-1}\,g(z)$ tvoří bázi $\mathcal{C}$.

3. Polynom $g(x)$ dělí polynom $x^n - 1$.

Důkaz. 1. a 2. Zvolme v $\mathcal{C}$ nenulový polynom co nejmenšího stupně. Označme jej $g(z)$ a jeho stupeň nechť je s.
Nechť $v(z) \in \mathcal{C}$. Pak $s \leq \operatorname{st} v \leq n-1$. Dělíme $v(x)$ polynomem $g(x)$

$$v(x) = a(x)g(x) + r(x),$$

kde $r(x)$ je zbytek po dělení, tj. $r(x) \equiv 0$ nebo $\operatorname{st} r < s$. Po dosazení z za proměnnou dostaneme rovnost

$$r(z) = \underbrace{v(z)}_{\in\mathcal{C}} - \underbrace{a(z)g(z)}_{\in\mathcal{C} \text{ podle (3.2)}} \in \mathcal{C} \qquad (\texttt{z linearity } \mathcal{C})$$

Protože $r(z)$ má $\operatorname{st} r < s$, plyne z minimality s, že $r(z) \equiv 0$, a tedy $v(z) = a(z)g(z)$, kde $a(z)$ je polynom stupně $\leq n-1-s$. Označme $a(z) = a_0+a_1z+\cdots+a_{n-1-s}z^{n-1-s}$. Odtud

$$v(z) = a_0g(z) + a_1zg(z) + a_2z^2g(z) + \cdots + a_{n-1-s}z^{n-1-s}g(z)\,.$$

Ukázali jsme, že každé kódové slovo $v(z)$ lze nakombinovat pomocí koeficientů $a_0, \ldots, a_{n-s-1}$ z polynomů $g(z), zg(z), \ldots, z^{n-s-1}g(z)$. Tyto polynomy jsou lineárně nezávislé, a tak tvoří bázi kódu $\mathcal{C}$. Proto $\dim\mathcal{C} = k = n - s$ a stupeň polynomu $g(z)$ je $s = n - k$.
3. Nechť $x^n - 1 = a(x)g(x) + r(x)$, kde $\operatorname{st} r < st\,g$. Po dosazení z a počítání v okruhu $\mathbb{Z}_2[x]/x^n - 1$ dostaneme $r(z) = -a(z)g(z)$. Což znamená, že $r(z)$ je prvek $\mathcal{C}$. Protože $\operatorname{st} r < \operatorname{st} g$, je r nulový polynom. □

Definice: Polynom $g(x)$ z předchozí věty nazýváme *generující polynom* kódu $\mathcal{C}$ a polynom $h(x) = \frac{x^n-1}{g(x)}$ se nazývá *kontrolní*.

Příklad: Jak vypadají cyklické kódy délky 5? Víme

$$x^5 - 1 = (x+1)(x^4 + x^3 + x^2 + x + 1)$$

a polynomy napravo jsou už ireducibilní, nelze je dál rozložit na součin. Proto existují dva cyklické kódy:

$\mathcal{C}_1$ s generujícím polynomem $x+1$ a kontrolním $x^4+x^3+x^2+x+1$. Má bázi tvořenou polynomy: $z+1, z(z+1), z^2(z+1), z^3(z+1)$ neboli kódovými slovy $11000, 01100, 00110, 00011$, tedy generující matice je

$$G = \begin{pmatrix} 11000 \\ 01100 \\ 00110 \\ 00011 \end{pmatrix} \quad \texttt{kód kontroly parity}$$

Abychom dostali kontrolní matici H, hledáme ortogonální doplněk (jelikož má platit $HG^\top = \Theta$). Nutně tedy $H = (11111)$.

$\mathcal{C}_2$ s generujícím polynomem $x^4+x^3+x^2+x+1$. Tedy bázi tvoří jediný polynom $z^4+z^3+z^2+z+1$. Jedná se o opakovací kód $\mathcal{C}_2 = \{11111, 00000\}$.

3.2 Minimální polynomy prvků tělesa $\mathbb{Z}_2[x]/q(x)$

Pro konstrukci cyklických kódů délky n potřebujeme hledat generující polynomy, jejichž vlastností je, že dělí polynom x^n-1. Proto je naším cílem (podobně jako v příkladu x^5-1 z minulé sekce) nalézt rozklad x^n-1 na ireducibilní polynomy

$$x^n-1 = M_1(x) \cdot M_2(x) \cdot \cdots \cdot M_l(x)\,.$$

Z nich pak budeme volit generující polynomy pro cyklický kód $\mathcal{C}$ tak, aby měl dobré parametry. K tomu nám pomůže práce s polynomy z tělesa $\mathbb{Z}_2[x]/q(x)$, kde $q(x)$ je ireducibilní polynom.

Definice: Nechť $b \in \mathbb{Z}_2[x]/q(x)$, kde $q(x)$ je ireducibinlní polynom. Minimálním polynomem prvku b nazýváme nenulový polynom $f(x)$ s koeficienty v $\mathbb{Z}_2$ minimálního stupně takový, že $f(b) = 0$.

Příklad: V tělese $\mathbb{Z}_2[x]/x^3+x+1$ mají jednotlivé prvky tyto minimální polynomy.

0	1	z	$z+1$	z^2	z^2+1	z^2+z	z^2+z+1
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
x	$x+1$	x^3+x+1	x^3+x^2+1	x^3+x+1	x^3+x^2+1	x^3+x+1	x^3+x^2+1

Všimněme si, že máme 4 různé minimální polynomy a pro jejich součin platí

$$x(x^7-1) = x(x+1)(x^3+x+1)(x^3+x^2+1).$$

Vzniká přirozená otázka, zda musí ke každému prvku existovat minimální polynom? Kladná odpověď plyne z následujícího faktu. Víme, že prvky konečného tělesa bez 0 tvoří cyklickou grupu vzhledem k násobení, jinými slovy existuje $\alpha \in \mathbb{Z}_2[x]/q(x)$ takové, že

$$\mathbb{Z}_2[x]/q(x) \setminus \{0\} = \{\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{2^m-2}, \alpha^{2^m-1} = 1\}\,.$$

Připomeňme, že 2^m-1 je počet nenulových prvků tělesa $\mathbb{Z}_2[x]/q(x)$, kde $m = \operatorname{st} q$. Tedy každé nenulové $b \in \mathbb{Z}_2[x]/q(x)$ se dá napsat ve tvaru $b = \alpha^l$ pro nějaké l, kde

$0 \leq l \leq 2^m-2$. Ze vztahu $\alpha^{2^m-1} = 1$ plyne $(\alpha^l)^{2^m-1} = 1$. Tedy každý nenulový prvek tělesa je kořenem polynomu $x^{2^m-1}-1$. A proto každý prvek (včetně 0) je kořenem polynomu

$$x^{2^m} - x = x(x^{2^m-1}-1).$$

Vlastnosti minimálních polynomů

1. Ke každému prvku existuje jediný minimální polynom.

   Důkaz. Kdyby pro $b \in \mathbb{Z}_2[x]/q(x)$ existovaly 2 minimální polynomy f_1 a f_2, pak by měly stejný stupeň řekněme m. Koeficient u x^m u obou polynomů je 1. Proto polynom $f_1(x) - f_2(x)$ je stupně $< m$ a b je jeho kořenem - spor s minimalitou stupně m. □

2. Minimální polynom je ireducibilní nad $\mathbb{Z}_2$.

   Důkaz. Kdyby $b \in \mathbb{Z}_2[x]/q(x)$ měl reducibilní minimální polynom $f(x) = f_1(x) \cdot f_2(x)$, kde f_1 a f_2 mají stupeň alespoň 1, pak b je kořenem buď f_1 nebo f_2 - spor s minimalitou stupně f. □

3. Minimální polynom prvku b dělí každý polynom s koeficienty v $\mathbb{Z}_2$, který má b za kořen.

   Důkaz. Nechť f je minimální polynom prvku b a p je polynom, který má kořen b. Po dělení

   $$p(x) = a(x)f(x) + \underbrace{r(x)}_{\text{st}\, r < \text{st}\, f}$$

   Po dosazení b dostaneme $0 = 0 + r(b)$. Tedy b je kořenem polynomu r. Jelikož f je nenulový polynom minimálního stupně s kořenem b, je nutně $r(x)$ nulový polynom. □

4. Každý minimální polynom prvku $b \in \mathbb{Z}_2[x]/q(x)$ dělí polynom $x^{2^m} + x$.

   Důkaz. Důsledek bodu 3. □

5. Je-li $f(x)$ minimální polynom prvku b, pak $f(x)$ je minimálním polynomem prvku b^2.

   Důkaz. Nechť $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Využijeme toho, že v $\mathbb{Z}_2$ platí $a_i^2 = a_i$ a $1 + 1 = 0$. Proto

   $$(f(x))^2 = a_0 + a_1x^2 + a_2x^4 + \cdots + a_nx^{2n} = f(x^2)\,.$$

   Tedy $f(b^2) = (f(b))^2 = 0$. Kdyby b^2 bylo kořenem polynomu stupně menšího než st$f$, pak by podle bodu 3. minimální polynom prvku b^2 dělil polynom f a to je nemožné, protože podle 2. je f ireducibilní.

   □

Následující věta nám dá návod, jak pro speciální hodnoty n řešit úlohu rozložit $x^n - 1$ na ireducibilní polynomy.

Věta: Nechť $q(x) \in \mathbb{Z}_2[x]$ je ireducibilní polynom stupně m, pak

$$x^{2^m} - x = \texttt{součin všech různých minimálních polynomů prvků ze } \mathbb{Z}_2[x]/q(x).$$

Důkaz. Označme $F(x)$ polynom rovný součinu všech různých minimálních polynomů prvků ze $\mathbb{Z}_2[x]/q(x)$. Jak víme, každý prvek tělesa $\mathbb{Z}_2[x]/q(x)$ je kořenem polynomu $x^{2^m} - x$. Uvažme nyní rozklad $x^{2^m} - x$ na polynomy ireducibilní nad $\mathbb{Z}_2$. Z 2. a 3. vlastnosti minimálních polynomů vyplývá, že každý minimální polynom se nutně musí vyskytnout v rozkladu $x^{2^m} - x$, a proto $F(x)$ dělí $x^{2^m} - x$. Dále platí, že libovolný polynom stupně k má nejvýše k kořenů, a proto $\operatorname{st} F(x) \geq 2^m$ (těleso $\mathbb{Z}_2[x]/q(x)$ má celkem 2^m prvků). Dohormady s předchozím tedy vyplývá, že $F(x)$ je přímo roven $x^{2^m} - x$. □

Zbývá vyřešit otázku jak hledat minimální polynomy k prvkům tělesa?

Věta: Minimální polynom prvku $b \in \mathbb{Z}_2[x]/q(x)$ je

$$f(x) = (x - b)(x - b^2)(x - b^4) \cdots (x - b^{2^k}),$$

kde k je nejmenší číslo takové, že $b^{2^{k+1}} = b$.

Poznámka: Připomeňme, že každý prvek tělesa $\mathbb{Z}_2[x]/q(x)$, které má 2^m prvků, splňuje $b^{2^m} = b$. Hodnota k v předchozí větě je proto nanejvýš $m-1$. Jinými slovy minimální polynom je nanejvýš stupně m.

Důkaz. Z 5. vlastnosti víme, že minimální polynomy k prvku b má kromě b taky kořen b^2, a ze stejných důvodů b^4, atd. Proto minimální polynom prvku b nemůže mít menší stupeň, než je stupeň $f(x)$ ve znění věty. Zřejmě b je kořenem f. Stačí tedy ukázat, že koeficienty polynomu f jsou v $\mathbb{Z}_2$. S využitím toho, jak je definováno číslo k, dostaneme

$$(f(x))^2 = (x^2 - b^2)(x^2 - b^4) \cdots (x^2 - \underbrace{b^{2^{k+1}}}_{=b}) = (x^2 - b)(x^2 - b^2) \cdots (x^2 - b^{2^k}) = f(x^2)$$

Rozepišme $f(x)$ definované ve větě do tvaru

$$f(x) = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1} + x^k, \ \texttt{kde} \ a_i \in \mathbb{Z}_2[x]/q(x)\,.$$

$$(f(x))^2 = a_0^2 + a_1^2 x^2 + \cdots + a_{k-1}^2 x^{2(k-1)} + x^{2k} = f(x^2) = a_0 + a_1 x^2 + \cdots + a_{k-1} x^{2(k-1)} + x^{2k}.$$

Porovnáním koeficientů u stejných mocnin x dostaneme $a_i = a_i^2$ pro každý koeficient. Tedy každý koeficient je kořenem polynomu $x^2 - x$. Tento polynom má pouze dva kořeny 0 a 1. Proto $a_i \in \{0, 1\}$, a tedy všechny koeficienty polynomu $f(x)$ jsou v $\mathbb{Z}_2$.

□

Příklad: V tělese $\mathbb{Z}_2[x]/x^3+x+1$ spočítejme pomocí předchozí věty minimální polynom k prvku $b = z{+}1$. Jelikož

$$b^2 = z^2{+}1, \quad b^4 = z^4{+}1 = z^2+z+1, \quad b^8 = z^4+z^2+1 = z{+}1 = b,$$

dostaneme

$$f(x) = (x{-}b)(x{-}b^2)(x{-}b^4) = x^3{+}\underbrace{(b+b^2+b^4)}_{1}x^2{+}\underbrace{(b^3+b^5+b^6)}_{0}x{+}b^7 = x^3{+}x^2{+}1$$

3.3 Generující kořeny cyklických kódů

Pro konstrukci cyklických kódů délky n potřebujeme hledat generující polynomy $g(x)$ takové, že $x^n{-}1 = g(x)h(x)$ pro nějaký polynom $h(x)$. Když rozložíme $x^n{-}1$ na součin ireducibilních polynomů s koeficienty v $\mathbb{Z}_2$, řekněme $x^n{-}1 = f_1(x)f_2(x)\cdots f_s(x)$, pak zřejmě, všechny možné generující polynomy $g(x)$ cyklických kódů délky n získáme součinem libovolného výběru z polynomů f_1 až $f_s(x)$.

Víme už, jak rozkládat na ireducibilní faktory s koeficienty v $\mathbb{Z}_2$ polynomy tvaru $x^{2^m-1}{-}1$. Ty se rovnají součinu minimálních polynomů nenulových prvků tělesa $\mathbb{Z}_2[x]/q(x)$, kde m je $=\mathrm{st}\, q$. Ale málokdy je n tvaru $n = 2^m{-}1$. Jak ale postupovat, pokud není?

- Pro zadané n nalezneme m tak, aby n dělilo $2^m{-}1$, tj. existuje r takové, že $nr = 2^m{-}1$. To vynucuje n liché - proto od této chvíle **vždy** n **liché**.
- Nalezneme ireducibilní polynom $q(x)$ stupně m a zkonstruujeme těleso $\mathbb{Z}_2[x]/q(x)$.
- Nalezneme $\beta \in \mathbb{Z}_2[x]/q(x)$ tak, že $\beta, \beta^2, \ldots, \beta^{2^m-1} = 1$ obsahuje všechny nenulové prvky tělesa.
- Prvky $\alpha_i = \beta^{ir}$, kde $i = 1, 2, \ldots, n$ jsou všechny navzájem různé kořeny rovnice $x^n = 1$. Vskutku,

  $$\alpha_i^n = \beta^{irn} = \big(\beta^{rn}\big))^i = \big(\beta^{2^m-1}\big)^i = 1^i = 1\,.$$

  Navíc prvek $\alpha_1 = \beta^r$ má řád n, tj. $\alpha_i^k \neq 1$, pokud $0 < k < n$. Řád každého prvku β^{ir} je dělitelem čísla n. Z vlastností minimálního polynomu k prvku b nějakého tělesa víme, že minimální polynom dělí každý polynom, který má kořen b. Odtud získáváme

  $x^n{-}1 =$ `součin min. polynomů prvků` $b \in \mathbb{Z}_2[x]/q(x)$`, kde řád` b `dělí` n.

Odvodili jsme následující větu.

Věta: Pro generující polynom $g(x)$ cyklického kódu $\mathcal{C}$ liché délky n platí

$$g(x) = \texttt{součin min. polynomů k prvkům}\ \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_s}$$

kde α je prvek řádu n v nějakém tělese $\mathbb{Z}_2[x]/q(x)$.

Definice: Prvky $\alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_s}$ z předchozí věty nazýváme generující kořeny cyklického kódu $\mathcal{C}$.

Protože každý prvek cyklického kódu v polynomiálním zápise je násobkem generujícího polynomu, zřejmě platí:

$$v_0 v_1 \ldots v_{n-1} \in \mathcal{C} \quad \Leftrightarrow \quad v(\alpha^{i_j}) = 0 \ \texttt{pro každé} \ j = 1, 2, \ldots, s,$$

kde $v(z)$ je polynom $v_0 + v_1 z + \cdots + v_{n-1} z^{n-1}$.

Příklad: $n = 7$. Pak stačí vzít $m = 3$, protože $7|2^3{-}1$. Za $q(x)$ zvolíme ireducibilní polynom $q(x) = x^3 + x + 1$ a pracujeme tedy s tělesem $\mathbb{Z}_2[x]/x^3 + x + 1$. Prvek $\alpha = z{+}1$ má řád $n = 7$ a lze ověřit (viz. předchozí podkapitola), že jeho minimální polynom je $x^3 + x^2 + 1$. Můžeme proto zvolit $g(x) = x^3 + x^2 + 1$. Odtud

$$v_0 + v_1 z + \ldots v_6 z^6 = v(z) \in \mathcal{C} \Leftrightarrow v(\alpha) = 0, \ v_0, \ldots, v_6 \in \{0, 1\}.$$

Neboli $v_0 + v_1\alpha + v_2\alpha^2 + \cdots + v_6\alpha^6 = 0$ v tělese $\mathbb{Z}_2[x]/x^3 + x + 1$ dosadíme za $\alpha, \alpha^2, \ldots, \alpha^6$.

$$v_0 + v_1(z{+}1) + v_2 \underbrace{(z{+}1)^2}_{z^2+1} + \cdots + v_6 \underbrace{(z{+}1)^6}_{z^2+z} = 0$$

$$v_0 + v_1(z{+}1) + v_2(z^2{+}1) + v_3(z^2) + v_4(z^2 + z + 1) + v_5(z) + v_6(z^2 + z) = 0$$

Porovnáním koeficientu u stejných mocnin z dostaneme podmínky

$$\begin{array}{ccccccccccccccc} v_0 & + & v_1 & + & v_2 & & & + & v_4 & & & & & = & 0 \\ & & v_1 & & & & & + & v_4 & + & v_5 & + & v_6 & = & 0 \\ & & & & v_2 & + & v_3 & + & v_4 & & & + & v_6 & = & 0 \end{array}$$

Maticově

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}}_{H} \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

kde H je kontrolní matice cyklického kódu s jedním generujícím kořenem α řádu 7. H obsahuje všechny nenulové trojice jako sloupce, t.j. kód je Hammingův. Obecně platí

Fakt: Každý binární Hammingův kód je ekvivalentní cyklickému kódu. Má jeden generující kořen α řádu $n = 2^m{-}1$ a minimální vzdálenost je $d = 3$.

Kapitola 4

BCH kódy

Je to nejdůležitější třída bezpečnostních kódů. Je pojmenována podle iniciál objevitelů, kterými byli Bose, Chaudhuri, Hocquenghem. Jsou to cyklické kódy liché délky n, ve kterých lze volbou generujících kořenů získat požadovanou minimální vzdálenost d a rychle dekódovat (daleko rychleji než se standardní tabulkou).

4.1 BCH kódy pro opravy dvojnásobných chyb

Definice: Binárním BCH kódem pro dvojnásobné opravy chyb se nazývá kód liché délky $n \geq 5$ s generujícími kořeny α a α^3, kde α je prvek řádu n v některém tělese $\mathbb{Z}_2[x]/q(x)$. Generujícím polynomem tohoto kódu je součin minimálních polynomů prvků α a α^3.

Příklad A: Konstruujeme cyklický kód délky $n = 15$ s generujícími kořeny α a α^3. Najdeme m tak, aby $n|2^m-1$. Tedy $m = 4$. Jako ireducibilní polynom stupně 4 zvolíme $q(x) = x^4 + x + 1$ a pracujeme v tělese $\mathbb{Z}_2[x]/q(x)$.

$$\mathbb{Z}_2[x]/q(x) = \{a_0 + a_1 z + a_2 z^2 + a_3 z^3 \ : \ a_0, a_1, a_2, a_3 \in \{0, 1\}\}.$$

Víme, že $\mathbb{Z}_2[x]/q(x) \setminus \{0\} = \{z, z^2, z^3, \ldots, z^{15} = 1\}$, t.j. $\alpha = z$. Využijeme zápisu prvků tělesa ve tvaru mocniny generátoru α i znalost minimálních polynomů, viz

následující tabulka.

prvek $\mathbb{Z}_2[x]/q(x) \setminus \{0\}$	ve tvaru z^i	minimální polynom prvku
1	$z^{15} = z^0$	$x{+}1$
z	z^1	$x^4{+}x{+}1$
z^2	z^2	$x^4{+}x{+}1$
z^3	z^3	$x^4{+}x^3{+}x^2{+}x{+}1$
$1{+}z$	z^4	
$z{+}z^2$	z^5	
$z^2{+}z^3$	z^6	
$1{+}z{+}z^3$	z^7	
$1{+}z^2$	z^8	
$z{+}z^3$	z^9	
$1{+}z{+}z^2$	z^{10}	
$z{+}z^2{+}z^3$	z^{11}	
$1{+}z{+}z^2{+}z^3$	z^{12}	
$1{+}z^2{+}z^3$	z^{13}	
$1{+}z^3$	z^{14}	

Uvažujme cyklický kód s generujícími kořeny $\alpha = z$ a $\alpha^3 = z^3$. Generující polynom kódu je $g(x) = (x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1)$. Generující polynom má stupeň 8, proto dimenze kódu $k = n{-}8 = 15 - 8 = 7$.

Polynom $v(z)$ stupně ≤ 14 je kódovým slovem právě tehdy, když $v(\alpha) = 0$ a $v(\alpha^3) = 0$, formálně

$$v_0 v_1 \dots v_{14} \in \mathcal{C} \quad \Longleftrightarrow \quad \sum_{i=0}^{14} v_i \alpha^i = 0 \;\; \texttt{a} \;\; \sum_{i=0}^{14} v_i \alpha^{3i} = 0.$$

Připomeňme, že každý prvek tělesa je polynom stupně ≤ 3, tj. je reprezentován 4 koeficienty u mocnin $z^0, \dots, z^3$. Proto rovnost nule u dvou předchozích sum dává dohromady 8 rovnic pro keficienty polynomů.

I když kontrolní matici formálně zapisujeme takto

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 & \alpha^8 & \alpha^9 & \alpha^{10} & \alpha^{11} & \alpha^{12} & \alpha^{13} & \alpha^{14} \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} \end{pmatrix},$$

její skutečný tvar je

$$H = \left(\begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & \dots \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & \dots \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & \dots \\ \hline 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & \dots \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & \dots \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & \dots \end{array}\right) \in \{0,1\}^{8\times 15}$$

4.2 Dekódování BCH kódů pro opravy dvojnásobných chyb

Dekódování BCH kódu s generujícími kořeny α, α^3

- Vyslali jsme slovo v a předpokládejme, že došlo k dvěma chybám na pozici $i-$té a $j-$té. Chybový polynom je tedy tvaru $e(z) = z^i + z^j$. Vypočteme syndrom $H \cdot w = \begin{pmatrix} w(\alpha) \\ w(\alpha^3) \end{pmatrix} = \begin{pmatrix} s_1 \\ s_3 \end{pmatrix} = \begin{pmatrix} \alpha^i + \alpha^j \\ \alpha^{3i} + \alpha^{3j} \end{pmatrix}$. Ze znalosti s_1 a s_3 chceme zjistit pozici i a j. Řešíme soustavu

$$s_1 = \alpha^i + \alpha^j$$
$$s_3 = \alpha^{3i} + \alpha^{3j}$$

  Protože $i \neq j$ (chyba na dvou různých místech) je $s_1 \neq 0$. Z první rovnice

$$s_1^3 = \alpha^{3i} + \alpha^{2i}\alpha^j + \alpha^i\alpha^{2j} + \alpha^{3j} = \underbrace{\alpha^{3i} + \alpha^{3j}}_{s_3} + \alpha^i\alpha^j \underbrace{(\alpha^i + \alpha^j)}_{s_1},$$

  a odtud $\alpha^i\alpha^j = s_1^{-1}(s_1^3 + s_3) = s_1^2 + s_1^{-1}s_3$. Uvažujme kvadratickou rovnici

$$(\lambda + \alpha^i)(\lambda + \alpha^j) = \lambda^2 + (\alpha^i + \alpha^j)\lambda + \alpha^i\alpha^j = \lambda^2 + s_1\lambda + s_1^2 + s_1^{-1}s_3.$$

  Když nalezneme kořeny této rovnice, zjistíme pozici i a j.

- Předpokládejme, že došlo k jediné chybě na pozici $i-$té. Pak $s_1 = \alpha^i$ a $s_3 = \alpha^{3i}$. Neboli $s_1^3 = s_3$ a předchozí rovnice má tvar $\lambda^2 + s_1\lambda$. Tato rovnice má kořeny $s_1 = \alpha^i$ a 0.

Algoritmus pro dekódování:
Spočítej syndrom $\begin{pmatrix} s_1 \\ s_3 \end{pmatrix}$.
Pokud $s_1 = s_3 = 0$, pak za vyslané slovo prohlásíme přijaté.
Jinak hledáme kořeny rovnice $\lambda^2 + s_1\lambda + s_1^2 + s_3s_1^{-1}$.
Buď jsou kořeny 0 a α^i, pak nastala jedna chyba a opravíme $i-$tou pozici.
Nebo jsou kořeny α^i a α^j a opravíme $i-$tou a $j-$tou pozici.

Poznámka: Cyklické kódy s kořeny α a α^3 jsme apriori nazvali "kódy opravující dvě chyby". Vidíme, že tento název byl opodstatněný.
Dekódování, které jsme právě popsali, je založené na řešení kvadratické rovnice. To probereme v následující sekci. Toto dekódování opraví nejvíce dvě chyby. Pro dekódování lze použít i standardní tabulku, která opraví i některé další chyby, ale je mnohem pomalejší.

4.3 Řešení kvadratických rovnic v tělese $\mathbb{Z}_2[x]/q(x)$

Řešení kvadratických rovnic v konečných tělesech se podstatně liší od metody, kterou známe pro těleso $\mathbb{R}$ nebo $\mathbb{C}$. Je založeno na existenci tzv. normální báze. Konečné těleso $\mathbb{Z}_2[x]/q(x)$, kde $q(x)$ je ireducibilní polynom, je vektorovým prostorem dimenze m nad tělesem $\mathbb{Z}_2$.

Věta(Davenport 1968): V tělese $GF(p^m)$ existuje primitivní prvek β takový, že

$$\beta, \beta^p, \beta^{p^2}, \ldots, \beta^{p^{m-1}}$$

tvoří bázi $GF(p^m)$ jako vektorového prostoru nad $\mathbb{Z}_p$. tzv. *normální báze*. Jinými slovy, každý prvek $d \in GF(p^m)$ lze vyjádřit ve tvaru

$$d = \sum_{i=0}^{m-1} a_i \beta^{p^i}\,, \quad \texttt{kde} \quad a_0, a_1, \ldots, a_{m-1} \in \mathbb{Z}_p\,.$$

Číslo $\mathrm{Tr}(d) := a_0 + a_1 + \cdots + a_{m-1}$ nazýváme stopa prvku d.

Příklad: V tělese $GF(2^4) = \mathbb{Z}_2[x]/q(x)$, kde $q(x) = x^4+x+1$ (uvedené u BCH kódů) tvoří

$$\beta = z^3, \quad \beta^2 = z^6 = z^2+z^3, \quad \beta^4 = z^{12} = 1+z+z^2+z^3, \quad \beta^8 = z^{24} = z^9 = z+z^3$$

normální bázi. Platí totiž, že každý polynom původní báze $1, z, z^2, z^3$ v ní lze vyjádřit

$$\begin{array}{rcccccccc} 1 & = & 1\cdot\beta & + & 1\cdot\beta^2 & + & 1\cdot\beta^4 & + & 1\cdot\beta^8 \\ z & = & 1\cdot\beta & & & + & & & 1\cdot\beta^8 \\ z^2 & = & 1\cdot\beta & + & 1\cdot\beta^2 & & & & \\ z^3 & = & 1\cdot\beta & & & & & & \end{array}$$

Volba $\beta = z$, $\beta^2 = z^2$, $\beta^4 = z^4 = z+1$, $\beta^8 = z^2+1$ netvoří bázi.

Lemma: Nechť $a\lambda^2 + b\lambda + c = 0$ je kvadratická rovnice pro neznámou λ s koeficienty $a, b, c \in \mathbb{Z}_2[x]/q(x)$, kde $a, b \neq 0$. Pak substituce $\lambda = \frac{b}{a}\mu$ převede rovnici na tvar $\mu^2+\mu+\frac{ca}{b^2} = 0$.

Důkaz. Plyne z tvaru substituce. □

Důsledek: Stačí se omezit na řešení kvadratických rovnic ve tvaru

$$\mu^2 + \mu + d = 0 \quad \texttt{a} \quad \mu^2 + d = 0\,.$$

Věta 1: Rovnice $\mu^2 + d = 0$ má vždy řešení. Pokud $\beta, \beta^2, \beta^{2^2}, \ldots, \beta^{2^{m-1}}$ je normální báze a d v ní má tvar

$$d = d_0\beta + d_1\beta^2 + d_2\beta^{2^2} + \cdots + d_{m-1}\beta^{2^{m-1}}, \quad \texttt{kde} \quad d_0, d_1, \ldots, d_{m-1} \in \mathbb{Z}_2\,,$$

pak řešením je $\mu = d_1\beta + d_2\beta^2 + d_3\beta^{2^2} + \cdots + d_0\beta^{2^{m-1}}$.

Důkaz. Dosazením navrhovaného řešení do kvadratické rovnice s využitím rovnosti $\beta^{2^m} = \beta$.

□

Věta 2: Rovnice $\mu^2 + \mu + d = 0$ má řešení právě tehdy, když $\mathrm{Tr}(d) = 0$. V případě, že $\mathrm{Tr}(d) = 0$, pak existují dvě řešení

$$\mu^{(1)} = 1 \cdot \beta + (d_1{+}1)\beta^{2} + (d_1{+}d_2{+}1)\beta^{2^2} + \cdots + (d_1{+}d_2 + \ldots {+}d_{m-1}{+}1)\beta^{2^{m-1}}$$
$$\mu^{(0)} = d_1\beta^{2} + (d_1{+}d_2)\beta^{2^2} + (d_1{+}d_2{+}d_3)\beta^{2^3} + \cdots + (d_1{+}d_2 + \ldots {+}d_{m-1})\beta^{2^{m-1}}$$

Důkaz. Hledejme řešení μ zapsané v normální bázi

$$\mu = \mu_0\beta + \mu_1\beta^{2} + \mu_2\beta^{2^2} + \cdots + \mu_{m-1}\beta^{2^{m-1}}.$$

Pak $\mu^2 = \mu_{m-1}\beta + \mu_0\beta^{2} + \mu_1\beta^{2^2} + \cdots + \mu_{m-2}\beta^{2^{m-1}}$.
Odtud

$$\mathrm{Tr}(\mu^2{+}\mu) = \mu_0{+}\mu_{m-1}{+}\mu_1{+}\mu_0{+}\ldots{+}\mu_{m-1}{+}\mu_{m-2} = 2(\mu_0{+}\mu_1{+}\ldots{+}\mu_{m-1}) = 0 = \mathrm{Tr}(d).$$

Z druhé strany ověřme, že $\mu^{(1)}$ a $\mu^{(0)}$ jsou řešení za předpokladu, že $\mathrm{Tr}(d) = 0$. Vskutku

$$\begin{aligned}\left(\mu^{(0)}\right)^2 + \mu^{(0)} =& d_1\beta^{2^2} + (d_1{+}d_2)\beta^{2^3} + (d_1{+}d_2{+}d_3)\beta^{2^4} + \cdots + (d_1{+}d_2{+}\ldots{+}d_{m-2})\beta^{2^{m-1}} \\ & (d_1{+}d_2{+}\ldots{+}d_{m-1})\beta + d_1\beta^{2} + (d_1{+}d_2)\beta^{2^2} + (d_1{+}d_2{+}d_3)\beta^{2^3} + \\ & + \cdots + (d_1{+}d_2{+}\cdots{+}d_{m-1})\beta^{2^{m-1}} = \\ =& \underbrace{(d_1{+}d_2{+}\cdots{+}d_{m-1})}_{d_0}\beta + d_1\beta^{2} + d_2\beta^{2^2} + = d_3\beta^{2^3} + \cdots + d_{m-1}\beta^{2^{m-1}} \\ =& d\end{aligned}$$

□

Příklad: V tělese V tělese $GF(2^4) = \mathbb{Z}_2[x]/q(x)$, kde $q(x) = x^4{+}x{+}1$ (uvedené u BCH kódů) máme rozhodnout o řešení rovnice

$$\lambda^2 + \lambda + \underbrace{z^2{+}z{+}1}_{d} = 0.$$

V normální bázi z příkladu za Davenportovou větou

$$d = 1 \cdot \beta + 0 \cdot \beta^2 + 1 \cdot \beta^4 + 0 \cdot \beta^8 \quad \Rightarrow \quad \mathrm{Tr}(d) = 1{+}0{+}1{+}0 = 0.$$

Máme dvě řešení

$$\lambda^{(0)} = 0 \cdot \beta^2 + 1 \cdot \beta^4 + 1 \cdot \beta^8 = 1{+}z^2,$$
$$\lambda^{(1)} = 1 \cdot \beta + 1 \cdot \beta^2 = z^2.$$

Ověřme

$$\underbrace{(1{+}z^2)^2}_{(\lambda^{(0)})^2} + \underbrace{1{+}z^2}_{\lambda^{(0)}} + \underbrace{z^2{+}z{+}1}_{d} = 1 + \underbrace{z^4}_{z+1} + 1{+}z^2{+}z^2{+}z{+}1 = 0$$
$$\underbrace{(z^2)^2}_{\lambda^{(1)}} + z^2{+}z^2{+}z{+}1 = \underbrace{z{+}1}_{z^4} + z^2{+}z^2{+}z{+}1 = 0$$

4.4 BCH kódy pro opravy t násobných chyb

Definice: BCH kód s plánovanou vzdáleností $d = 2, 3, 4, \ldots$ je kód liché délky $n \geq d$ s generujícími kořeny $\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{d-1}$, kde α je prvek $n-$tého řádu v nějakém tělese $\mathbb{Z}_2[x]/q(x)$. Kontrolní matice tohoto kódu je

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \ldots & \alpha^{n-1} \\ 1 & \alpha^2 & (\alpha^2)^2 & (\alpha^2)^3 & \ldots & (\alpha^2)^{n-1} \\ 1 & \alpha^3 & (\alpha^3)^2 & (\alpha^3)^3 & \ldots & (\alpha^3)^{n-1} \\ \vdots & & & & & \\ 1 & \alpha^{d-1} & (\alpha^{d-1})^2 & (\alpha^{d-1})^3 & \ldots & (\alpha^{d-1})^{n-1} \end{pmatrix}$$

Poznámka: Sudé mocniny mezi generujícími kořeny jsou zbytečné. Polynom $v(z)$ má totiž s každým kořenem α^i taky kořen $\alpha^{2i}, \alpha^{4i}, \ldots$ Odtud plyne, že BCH kódy s plánovanou vzdáleností $d = 2t$ a $d = 2t+1$ jsou totožné a mají generující kořeny $\alpha, \alpha^3, \alpha^5, \ldots, \alpha^{2t-1}$. Proto budeme předpokládat, že plánovaná vzdálenost d je liché číslo a kontrolní matice je tvaru

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \ldots & \alpha^{n-1} \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \ldots & \alpha^{3(n-1)} \\ 1 & \alpha^5 & \alpha^{10} & \alpha^{15} & \ldots & \alpha^{5(n-1)} \\ \vdots & & & & & \\ 1 & \alpha^{d-2} & \alpha^{2d-4} & \alpha^{3d-6} & \ldots & \alpha^{(d-2)(n-1)} \end{pmatrix}$$

Příklad: BCH kód délky 15 s plánovanou vzdáleností 7 má kořeny $\alpha, \alpha^3, \alpha^5$, a tedy generující polynom $g(x)$ bude součinem minimálních polynomů k těmto kořenům.

V tělese $\mathbb{Z}_p[x]/q(x)$ z příkladu **A** s polynomem $q(x) = x^4+x+1$ a z tabulky ze stejného příkladu nalezneme minimální polynom k prvku $\alpha = z$, t.j. $M_\alpha(x) = x^4 + x + 1$ a k $\alpha = z^3$, t.j. $M_{\alpha^3}(x) = x^4 + x^3 + x^2 + x + 1$.
Minimální polynom k α^5 spočítáme podle věty o hledání minimální polynomů

$$M_{\alpha^5}(x) = (x - \alpha^5)(x - (\alpha^5)^2), \quad \text{jelikož už} \quad (\alpha^5)^{2^2} = \alpha^{20} = \alpha^5 .$$

Tedy

$$M_{\alpha^5}(x) = (x - \alpha^5)(x - \alpha^{10}) = x^2 - (\alpha^5 + \alpha^{10})x + \alpha^{15} = x^2 + x + 1.$$

Při úpravách jsme opět využili tabulku z **A**, podle ní $\alpha^5 + \alpha^{10} = \alpha+\alpha^2+1+\alpha+\alpha^2 = 1$.

Generující polynom BCH kódu s kořeny $\alpha, \alpha^3, \alpha^5$ je

$$g(x) = M_\alpha(x)M_{\alpha^3}(x)M_{\alpha^5}(x) = (x^4+x+1)(x^4+x^3+x^2+x+1)(x^2+x+1).$$

Protože g má stupeň 10, je dimenze kódu $\mathcal{C}$ rovna $15 - 10 = 5$.

Naším cílem dále bude ukázat, že kód s plánovanou vzdáleností d má minimální vzdálenost alespoň d. K tomu budeme potřebovat určit, kdy je nenulový determinant

Vandermondovy matice rozměru $d \times d$ s parametry $a_1, a_2, \ldots, a_d$

$$V(a_1, a_2, \ldots, a_d) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ a_1 & a_2 & & a_d \\ a_1^2 & a_2^2 & & a_d^2 \\ a_1^3 & a_2^3 & & a_d^3 \\ \vdots & \vdots & & \vdots \\ a_1^{d-1} & a_2^{d-1} & & a_d^{d-1} \end{pmatrix}$$

Lemma: $V(a_1, a_2, \ldots, a_d) = V(a_1, a_2, \ldots, a_{d-1}) \cdot \prod_{i=1}^{d-1} (a_d - a_i)$.

Důkaz. Uvažujme polynom proměnné t (poslední sloupec matice jsme nahradili mocninami t)

$$p(t) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ a_1 & a_2 & & t \\ a_1^2 & a_2^2 & & t^2 \\ a_1^3 & a_2^3 & & t^3 \\ \vdots & \vdots & & \vdots \\ a_1^{d-1} & a_2^{d-1} & & t^{d-1} \end{pmatrix}.$$

Počítáme-li determinant rozvojem podle posledního sloupce, dostaneme, že koeficient u nejvyšší mocniny t^{d-1} je $V(a_1, a_2, \ldots, a_{d-1})$. Pokud za proměnnou t dosadíme a_i, pak ma matice dva stejne sloupce a determinant je 0. Proto $a_1, a_2, \ldots, a_{d-1}$ jsou kořeny $p(t)$. Polynom stupně $d-1$, u kterého známe $d-1$ kořenů i koeficient u největší mocniny, má tvar

$$p(t) = V(a_1, a_2, \ldots, a_{d-1})(t - a_1)(t - a_2) \cdots (t - a_{d-1}).$$

Stačí za t dosadit a_d a uvědomit si, že $p(a_d) = V(a_1, a_2, \ldots, a_d)$.

□

Důsledek: $V(a_1, a_2, \ldots, a_d) = \prod_{d \geq j > i \geq 1} (a_j - a_i)$.

Věta: BCH kód s plánovanou vzdálenosti d má minimální vzdálenost $\geq d$.

Důkaz. Lineární kód má minimální vzdálenost $\geq d$ pokud každých $d-1$ sloupců kontrolní matice H je lineárně nezávislých. Pro důkaz sporem předpokládejme, že sloupce $i_1, i_2, \ldots, i_{d-1}$ matice

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \ldots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \ldots & \alpha^{2(n-1)} \\ \vdots & & & & & \\ 1 & \alpha^{d-1} & & & \ldots & \alpha^{(d-1)(n-1)} \end{pmatrix}$$

jsou LZ. Pak determinant matice utvořené z těchto sloupců je $= 0$. Ale

$$\det \begin{pmatrix} \alpha^{i_1} & \alpha^{i_2} & \dots & \alpha^{i_{d-1}} \\ \alpha^{2i_1} & \alpha^{2i_2} & \dots & \alpha^{2i_{d-1}} \\ \vdots & \vdots & & \vdots \\ \alpha^{(d-1)i_1} & \alpha^{(d-1)i_2} & \dots & \alpha^{(d-1)i_{d-1}} \end{pmatrix} = \alpha^{i_1}\alpha^{i_2}\dots\alpha^{i_{d-1}} V(\alpha^{i_1}, \alpha^{i_2}, \dots, \alpha^{i_{d-1}})$$
$$= \alpha^{i_1}\alpha^{i_2}\dots\alpha^{i_{d-1}} \prod_{d-1 \geq k > j \geq 1} (\alpha^{i_k} - \alpha^{i_j}) \neq 0$$

protože $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ jsou navzájem různá čísla. Spor. □

Dekódování BCH kódu

Nechť plánovaná vzdálenost $d = 2t+1$. Předpokládejme, že víme, že došlo k p chybám, $p \leq t$. Existují tedy souřadnice $i_1, i_2, \dots, i_p$, kde má chybový vektor e složku 1.

$$e = (0, \dots, 0, \underset{i_1}{1}, 0, \dots, 0, \underset{i_2}{1}, 0, \dots, 0, \underset{i_p}{1}, 0, \dots, 0).$$

Přejděme opět k polynomiálnímu zápisu slov. Vyslali jsme $v(z) \in \mathcal{C}$, nastala chyba $e(z) = \sum_{k=1}^{p} z^{i_k}$ a obdrželi jsme na příjmu slovo $w(z) = v(z) + e(z)$.
Protože $v(z) \in \mathcal{C}$ je násobek generujícího polynomu, který má kořeny $\alpha, \alpha^2, \dots, \alpha^{d-1}$, platí $v(\alpha^k) = 0$ pro $k = 1, \dots, d-1$.
Proto $H \cdot w(\alpha) = H \cdot e(\alpha)$

$$H \cdot e(\alpha) = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(n-1)} \\ \vdots & & & & \\ 1 & \alpha^{d-1} & & \dots & \alpha^{(d-1)(n-1)} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_{d-1} \end{pmatrix} = \texttt{syndrom},\ d-1 = 2t,$$

kde

$$\begin{aligned} s_1 &= \alpha^{i_1} + \alpha^{i_2} + \dots + \alpha^{i_p} \\ s_2 &= (\alpha^{i_1})^2 + (\alpha^{i_2})^2 + \dots + (\alpha^{i_p})^2 \\ &\vdots \\ s_{d-1} &= (\alpha^{i_1})^{d-1} + (\alpha^{i_2})^{d-1} + \dots + (\alpha^{i_p})^{d-1}. \end{aligned}$$

Úkolem je ze znalosti $s_1, s_2, \dots, s_{d-1}$ (to jsou složky syndromu) vypočítat neznámé $a_1 = \alpha^{i_1}, a_2 = \alpha^{i_2}, \dots, a_p = \alpha^{i_p}$. Když budeme znát $a_1, \dots, a_p$ budeme vědět, na kterých p pozicích nastaly chyby, a ty opravíme.

Definice: Lokátorem p násobné chyby rozumíme polynom

$$f(x) = (1 - a_1 x)(1 - a_2 x)\dots(1 - a_p x).$$

1. Ukážeme, jak ze znalosti $s_1 = a_1 + a_2 + \cdots + a_p, s_2 = a_1^2 + a_2^2 + \cdots + a_p^2, \ldots,$ $s_{d-1} = s_{2t} = a_1^{2t} + a_2^{2t} + \cdots + a_p^{2t}$ určit koeficienty $f_1, f_2, \ldots, f_p$ lokátoru chyb $f(x) = 1 + f_1 x + f_2 x^2 + \cdots + f_p x^p$.

2. Pak nalezneme kořeny $r_1, r_2, \ldots, r_p$ polynomu f. A potřebné a_i získáme ze vztahu $a_1 = r_1^{-1}, a_2 = r_2^{-1}, \ldots, a_p = r_p^{-1}$. Teď už víme, které pozice $i_1, \ldots, i_p$ opravit.

Určení koeficientů $f_1, \ldots, f_p$ v lokátoru chyb.

$$1 + f_1 x + f_2 x^2 + \cdots + f_p x^p = (1 - a_1 x)(1 - a_2 x) \ldots (1 - a_p x) = f(x)$$

Formálně derivujeme:

$$\begin{aligned}\frac{f'(x)}{f(x)} &= \frac{-a_1}{1 - a_1 x} + \frac{-a_2}{1 - a_2 x} + \cdots + \frac{-a_p}{1 - a_p x} = -\sum_{j=1}^{\infty} a_1^j \cdot x^{j-1} - \ldots - \sum_{j=1}^{\infty} a_p^j \cdot x^{j-1} \\ &= -\sum_{j=1}^{\infty} \big(\underbrace{a_1^j + a_2^j + \cdots + a_p^j}_{s_j \text{ pro } j=1,2,\ldots,2t} \big) \cdot x^{j-1}\end{aligned}$$

Získáváme tak

$$-f'(x) = f(x) \sum_{j=1}^{\infty} (a_1^j + a_2^j + \cdots + a_p^j) x^{j-1} \,. \tag{4.1}$$

Na druhou stranu víme, že

$$\begin{aligned}f'(x) &= f_1 + 2f_2 x + 3f_3 x^2 + 4f_4 x^3 + 5f_5 x^4 + \cdots + p f_p x^{p-1} \\ &= f_1 + f_3 x^2 + f_5 x^4 + \ldots\end{aligned}$$

protože $2 \equiv 0$ v tělese charakteristiky 2. Porovnáním koeficientů v (4.1) u stejných mocnin x dostaneme

$$\begin{aligned}x^0 :&\quad f_1 = s_1 f_0 = s_1 \\ x^2 :&\quad f_3 = s_3 + f_1 s_2 + f_2 s_1 \\ x^4 :&\quad f_5 = s_5 + s_4 f_1 + s_3 f_2 + s_2 f_3 + s_1 f_4 \\ \vdots :&\quad \vdots \\ x^{2p-2} :&\quad 0 = s_{2p-1} + s_{2p-2} f_1 + s_{2p-3} f_2 \cdots + s_{p-1} f_p\end{aligned}$$

Řešíme soustavu pro neznámé $f_1, \ldots, f_p$ se známými koeficienty $s_1, \ldots, s_{2p-1}$. Maticově dostaneme

$$\begin{pmatrix} s_1 \\ s_3 \\ s_5 \\ \vdots \\ s_{2p-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ s_2 & s_1 & 1 & 0 & 0 & & 0 & 0 \\ s_4 & s_3 & s_2 & s_1 & 1 & & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \\ s_{2p-2} & s_{2p-3} & s_{2p-4} & s_{2p-5} & s_{2p-6} & & s_p & s_{p-1} \end{pmatrix}}_{\text{Označme matici } M_p} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_p \end{pmatrix}$$

Problémem pro sestavení těchto rovnic je to, že nevíme hodnotu p, t.j. skutečný počet chyb. Dále též apriori nevíme, že příslušná soustava nemá více než jedno řešení. Oba tyto problémy řeší následující věta, kterou uvedeme bez důkazu.

Věta: Pro každé $p = 1, 2, \ldots, t{+}1$ a faktický počet chyb τ platí

$$\begin{aligned} \det M_\tau &= 0 \ \texttt{pro} \ \tau > E{+}1, \\ \det M_\tau &\neq 0, \\ \det M_{\tau+1} &\neq 0. \end{aligned}$$

Závěr: Za p postupně dosazujeme $t{+}1, t, t{-}1, t{-}2, \ldots$ tak dlouho, až narazíme na $\det M_p \neq 0$. Pro první takové p prohlásíme, že došlo k $\tau = p{-}1$ chybám, a vyřešením výše uvedené soustavy s M_τ dostaneme koeficienty lokátoru chyb.

Kapitola 5

Binární kódy s velkou minimální vzdáleností

V této kapitole se budeme zabývat obecnými binárními kódy, nemusejí být lineární. Pro lineární kódy jsme při zkoumání vztahu mezi parametry zavedli $N(k, d)$ jako minimální délku lineárního binárního kódu, který má dimenzi k a minimální vzdálenost d. Binární lineární kód dimenze k má 2^k slov. Tedy jsme se pro fixovaný počet slov $M = 2^k$ a fixovanou minimální zdálenost d snažili hledat co nejkratší délku kódových slov n.

Nyní budeme uvažovat duální úlohu. K předepsané délce kódového slova n a minimální vzdálenosti alespoň d budeme hledat kód s co největším počtem kódových slov M. O kódu $\mathcal{C}$ řekneme, že je to (n, M, d)-kód, pokud platí $\mathcal{C} \subset \{0,1\}^n$, $M = \#\mathcal{C}$ a $d \geq \min\{\mathrm{dist}(x, y) \ : \ x, y \in \mathcal{C},\ x \neq y\}$.

Definujme

$$A(n, d) = \max\{M : \texttt{existuje } (n, M, d)\texttt{-kód}\}$$

V případě, že d je dostatečně velké vzhledem k n, jsou známé přesné hodnoty $A(n, d)$. V následující sekci odvodíme horní odhady na $A(n, d)$ a později ukážeme, že se jich za jistých okolností nabývá.

5.1 Plotkinova mez

Nejdříve odvodíme některé vlastnosti hodnot $A(n, d)$.

Lemma 1: Pro každé $n, r \in \mathbb{N}$ platí $A(n, 2r{-}1) = A(n{+}1, 2r)$.

Důkaz. Nechť $\mathcal{C}$ je $(n, M, 2r{-}1)$-kód, kde $M = A(n, 2r{-}1)$. Prodlužme každé kódové slovo $x = x_1 \cdots x_n \in \mathcal{C}$ symbolem $x_{n+1} \in \{0, 1\}$ tak, aby $x_1 + \cdots + x_n + x_{n+1} = 0 \mod 2$. Když se $x = x_1 \cdots x_n \in \mathcal{C}$ a $y = y_1 \ldots y_n \in \mathcal{C}$ lišily na lichém počtu míst, doplněné symboly jsou různé, tj. $x_{n+1} \neq y_{n+1}$ a rozšířené slova se tak nově liší o jednu pozici navíc. Tedy rozšířené slova z $\mathcal{C}$ tvoří $(n{+}1, M, 2r)$-kód. Tím jsme ukázali, že $A(n, 2r{-}1) \leq A(n{+}1, 2r)$.

Na druhé straně nechť $\tilde{\mathcal{C}}$ je $(n{+}1, \tilde{M}, 2r)$-kód, kde $\tilde{M} = A(n{+}1, 2r)$. Umažme z každé $(n{+}1)$-tice $\tilde{x} \in \tilde{\mathcal{C}}$ poslední souřadnici. Dostaneme $(n, \tilde{M}, 2r{-}1)$-kód. Proto platí i obrácená nerovnost $A(n, 2r{-}1) \geq A(n{+}1, 2r)$. □

Lemma 2: Pro každé $n, d \in \mathbb{N}, n > 1$, platí $A(n, d) \leq 2A(n{-}1, d)$.

Důkaz. Nechť $\mathcal{C}$ je (n, M, d)-kód, kde $M = A(n, d)$. Označme

$$\mathcal{C}_1 = \{x \in \mathcal{C} \ : \ \texttt{první složka } x \texttt{ je } 1\}$$
$$\mathcal{C}_0 = \{x \in \mathcal{C} \ : \ \texttt{první složka } x \texttt{ je } 0\}.$$

Bez újmy na obecnosti předpokládejme, že $\#\mathcal{C}_1 \geq \#\mathcal{C}_0$. Odtud $\#\mathcal{C}_1 \geq \frac{M}{2}$. Odstraněním první složky (které je u všech slov z $\mathcal{C}_1$ rovna 1) dostaneme $(n{-}1, M_1, d)$-kód, kde $M_1 = \#\mathcal{C}_1$. Proto platí také $A(n{-}1, d) \geq M_1 \geq \frac{M}{2} = \frac{1}{2}A(n, d)$. □

Lemma 3: Nechť $2d > n$. Pak $A(n, d) \leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor$.

Důkaz. Nechť $\mathcal{C}$ je (n, M, d)-kód, kde $M = A(n, d)$. Napišme si všechna slova kódu $\mathcal{C}$ pod sebe. Dostaneme tak tabulku o M řádcích a n sloupcích. Označme $\alpha_i =$ počet 0 v $i-$tém sloupečku tabulky. Zřejmě $M - \alpha_i =$ počet 1 v $i-$tém sloupečku.

$$dM(M-1) \leq \sum_{x \in C} \sum_{y \in C} \mathrm{dist}(x, y) = 2 \sum_{i=1}^{n} \alpha_i (M - \alpha_i).$$

Protože funkce $x(M-x)$ nabývá maxima v bodě

$$x = \frac{M}{2} \ \texttt{pokud je } M \texttt{ sudé, odhadneme} \quad \alpha_i(M-\alpha_i) \leq \frac{M^2}{4};$$
$$x = \frac{M-1}{2} \ \texttt{pokud je } M \texttt{ liché, odhadneme} \quad \alpha_i(M-\alpha_i) \leq \frac{M^2-1}{4}.$$

Pro M **sudé** tak odvodíme:

$$dM(M-1) \leq \frac{nM^2}{2} \quad \Rightarrow \quad 2d(M-1) \leq nM \quad \Rightarrow \quad M(2d-n) \leq 2d$$
$$\text{a odtud} \quad \underbrace{\frac{M}{2}}_{\in \mathbb{N}} \leq \frac{d}{2d-n} \quad \Rightarrow \quad \frac{M}{2} \leq \left\lfloor \frac{d}{2d-n} \right\rfloor.$$

Pro M **liché**:

$$dM(M-1) \leq n\frac{M^2-1}{2} \quad \Rightarrow \quad 2dM \leq n(M+1) \quad \Rightarrow \quad M \leq \frac{n}{2d-n} = \frac{2d}{2d-n} - 1$$
$$\text{a proto z celočíslenosti } M \text{ vyplývá } \quad M \leq \left\lfloor \frac{2d}{2d-n} \right\rfloor - 1 \leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor.$$

Poznamenjme, že v poslední nerovnosti jsme použili fakt $\lfloor 2t \rfloor \leq 2\lfloor t \rfloor + 1$ pro každé $t \geq 0$. □

Věta 4 (Plotkinova mez)

1. `Pokud` d `je sudé a` $2d > n$`, pak` $A(n,d) \leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor$.
2. `Pokud` d `je sudé, pak` $A(2d,d) \leq 4d$.
3. `Pokud` d `je liché a` $2d+1 > n$`, pak` $A(n,d) \leq 2\left\lfloor \frac{d+1}{2d+1-n} \right\rfloor$.
4. `Pokud` d `je liché, pak` $A(2d+1,d) \leq 4d+4$.

Důkaz:

1. To je tvrzení lemmatu 3.
2. $A(2d,d) \overset{\text{Lemma 2}}{\leq} 2A(2d-1,d) \overset{\text{Lemma 3}}{\leq} 2 \cdot 2 \cdot \left\lfloor \frac{d}{2d-(2d-1)} \right\rfloor = 4d.$
3. $A(n,d) \overset{\text{Lemma 1}}{=} A(n+1,d+1) \overset{\text{Lemma 3}}{\leq} 2\left\lfloor \frac{d+1}{2d+1-n} \right\rfloor.$
4. $A(2d+1,d) \overset{\text{Lemma 1}}{=} A(2d+2,d+1) \overset{\text{Věta 4 (2)}}{\leq} 4(d+1).$

Naším cílem bude ukázat, že v Plotkinových mezích platí všude rovnosti. K tomu budeme využívat kódy konstruované pomocí Hadamardových matic.

Definice: Čtvercová matice H řádu n s prvky $H_{ij} \in \{1,-1\}$ pro každé $i,j \in \{1,2,\ldots,n\}$, taková, že

$$HH^\top = nI.$$

se nazývá *Hadamardova matice.*

Poznámka: Jelikož $\frac{1}{n}HH^\top = I$, je matice $\frac{1}{n}H$ inverzní k maticí $H^\top$. Proto také $H^\top(\frac{1}{n}H) = I$, tj.

$$H^\top H = HH^\top = nI\,.$$

Tedy řádky i sloupce matice H jsou navzájem kolmé a mají velikost $\sqrt{n}$. Libovolná permutace řádků resp. sloupců neporuší jejich vzájemnou kolmost ani jejich velikost, proto neporuší ani hadamardovskost matice.

Poznámka: Vynásobíme-li v Hadamardově matici libovolný sloupec nebo řádek číslem -1, opět neporušíme vzájemnou kolmost ani velikost sloupců resp. řádků, tj. dostaneme opět Hadamardovu matici. Tedy existuje-li Hadamardova matice řádu n, existuje i Hadamardova matice stejného řádu, která má všechny prvky v 1. řádku a v 1. sloupci rovny 1. Nazýváme ji *standardní Hadamardova matice.*

Úmluva: Řád Hadamardovy matice občas vyznačujeme dolním indexem, píšeme H_n. Kvůli přehlednosti budeme prvky -1 v Hadamardově matice zapisovat jako $\overline{1}$.

Příklad: $H_1 = (1)\,,\quad H_2 = \begin{pmatrix} 1 & 1 \\ 1 & \overline{1} \end{pmatrix}\,,\quad H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \overline{1} & 1 & \overline{1} \\ 1 & 1 & \overline{1} & \overline{1} \\ 1 & \overline{1} & \overline{1} & 1 \end{pmatrix}\,.$

Pokusíme-li se najít Hadamardovou matici řádu 3, nepodaří se nám to. Fakt, že nejaký řádek má být kolmý na řádek ze samých jedniček totiž vynucuje, že v něm musí být polovina jedniček a polovina minus jedniček. Tedy řád matice musí být nutně sudý. Ani to však nestačí, jak říká následující věta.

Věta: Nechť existuje Hadamardova matice řádu n. Pak $n = 1$ nebo $n = 2$ nebo n je dělitelné čtyřkou.

Důkaz: Uvažujme $n \geq 3$ a bez újmy na obecnosti nechť H je standardní Hadamardova matice řádu n. Zpermutujeme sloupce tak, aby 1.,2. a 3. řádek měl tvar

$$\underbrace{\begin{array}{c|c|c|c} 111 & 11111 & 111 & 11 \\ 111 & 11111 & \overline{1}\overline{1}\overline{1} & \overline{1}\overline{1} \\ \underbrace{111}_{a\text{-krát}} & \underbrace{\overline{1}\overline{1}\overline{1}\overline{1}\overline{1}}_{b\text{-krát}} & \underbrace{111}_{c\text{-krát}} & \underbrace{\overline{1}\overline{1}}_{d\text{-krát}} \end{array}}_{n}$$

Ze vzájemné kolmosti řádků dostaneme

$$\begin{array}{rr} 1. \perp 2. \Rightarrow a + b - c - d & = 0 \\ 1. \perp 3. \Rightarrow a - b + c - d & = 0 \\ 2. \perp 3. \Rightarrow a - b - c + d & = 0 \\ a + b + c + d & = n \\ \hline \sum \Rightarrow \qquad 4a & = n \end{array}$$

Konstrukci Hadamardových matic se budeme věnovat ve zvláštní sekci. Dodnes se však neví, zda nutná podmínka z předchozí věty je i postačující. Všeobecně se usuzuje, že platí

Domněnka: Když n je dělitelné čtyřmi, pak existují Hadamardova matice řádu n.

V roce 2014 se o celkem 12 číslech n, jež jsou dělitelná čtyřmi a zároveň jsou menší než 2000, nevědělo, zda existuje Hadamardova matice řádu n. Jsou to tato čísla: 668, 716, 892, 1132, 1244, 1388, 1436, 1676, 1772, 1916, 1948 a 1964. Zajímavostí je, ze všech těchto 12 čísel n splňuje, že $n/4$ je prvočíslo kongruentní s 3 modulo 4.

5.2 Levenshteinova věta

Cílem této sekce je dokázat, že se v Plotkinových mezích nabývá rovnosti. To je obsahem Levenshteinovy věty. K jejímu důkazu budeme potřebovat dvě ingredience: tzv. Hadamardovy kódy a skládání kódů. Budeme využívat operaci E, která ve slovech z $\{0,1\}^n$ zaměňuje 0 a 1. Příklad: $E(10011) = 01100$.

Hadamardovy kódy

Nechť H_n je standardní Hadamardova matice řádu n.

A_n **:** V matici H_n zaměníme $1 \to 0$ a $\overline{1} \to 1$. Řádky takto vzniklé matice prohlásíme za slova kódu A_n.

Příklad: A_4 obsahuje slova $0000, 0101, 0011, 0110$.

Tvrzení: Kód A_n je $(n, n, \frac{n}{2})-$kód.

Důkaz: Každé dva různé řádky Hadamardovy matice, označme je
$$x = x_1x_2\cdots x_n \in \{1, \overline{1}\}^n \quad \text{a} \quad y = y_1y_2\cdots y_n \in \{1, \overline{1}\}^n,$$
jsou na sebe kolmé. To jest $x_1y_1 + x_2y_2 + \cdots + x_ny_n = 0$. Proto pro polovinu indexů i platí, že $x_iy_i = 1$ (a tedy $x_i = y_i$) a pro další polovinu platí $x_iy_i = \overline{1}$ (a tedy $x_i \neq y_i$). Tedy každé dvě kódová slova se liší na $\frac{n}{2}$ místech.

$\mathcal{A}_n$ **:** Vznikne tak, že z každého slova kódu A_n ubereme 1. složku.

Tvrzení: Kód $\mathcal{A}_n$ je $(n-1, n, \frac{n}{2})-$kód.

Důkaz: Protože H_n je standardní matice, každý řádek matice začíná jedničkou, a tedy každý prvek kódu A_n začíná nulou. Proto po umaznání první složky se nezmění počet míst, kde se různé řádky liší.

$\mathcal{A}'_n$ **:** Vznikne z $\mathcal{A}_n$ tak, že vybereme z $\mathcal{A}_n$ slova, která začínají na 0, a tu škrtneme.

Tvrzení: Kód $\mathcal{A}'_n$ je $(n-2, \frac{n}{2}, \frac{n}{2})-$kód.

Důkaz: Protože druhý sloupec matice H_n je kolmý na první sloupec, který je ze samých jedniček, obsahuje druhý sloupec matice H_n polovinu jedniček a polovinu minus jedniček. Tj. u kódu $\mathcal{A}_n$ začíná polovina slova nulou a polovina jedničkou.

B_n **:** Pokládáme $B_n = A_n \cup EA_n$.

Příklad: B_4 obsahuje slova

$$0000, 0101, 0011, 0110$$
$$1111, 1010, 1100, 1001$$

Tvrzení: Kód B_n je $(n, 2n, \frac{n}{2})-$kód.

Důkaz: Každá dvě slova kódu A_n se liší na $\frac{n}{2}$ místech a shodují se $\frac{n}{2}$ místech. Stačí si tedy uvědomit, že pro operaci E na slovech libovolné délky k platí: $\mathrm{dist}(x, E(y)) = k-\mathrm{dist}(x, y)$, $\mathrm{dist}(x, E(x)) = k$ a $\mathrm{dist}(x, y) = \mathrm{dist}(E(x), E(y))$.

Skládání kódů

Definujeme dvě operace s kódy a určíme jejich nové parametry.

$\mathcal{C}_1 \oplus \mathcal{C}_2$ **:** Nechť $\mathcal{C}_1$ je (n_1, M_1, d_1)-kód a $\mathcal{C}_2$ je (n_2, M_2, d_2)-kód. Označme $M = \min\{M_1, M_2\}$. Do nového kódu $\mathcal{C}_1 \oplus \mathcal{C}_2$ dáváme slova délky $n_1 + n_2$, která vzniknou tak, že první slovo kódu $\mathcal{C}_1$ zřetězíme s prvním slovem kódu $\mathcal{C}_2$, druhé slovo kódu $\mathcal{C}_1$ zřetězíme s druhým slovem kódu $\mathcal{C}_2$, atd. až M-té slovo kódu $\mathcal{C}_1$ zřetězíme s M-tým slovem kódu $\mathcal{C}_2$. Zřejmě platí

Tvrzení: Kód $\mathcal{C}_1 \oplus \mathcal{C}_2$ je $\big(n_1 + n_2, \min\{M_1, M_2\}, d_1 + d_2\big)$-kód.

$b \times \mathcal{C}$ **:** Nechť $\mathcal{C}$ je (n, M, d)-kód a $b \in \mathbb{N}$. Klademe $b \times \mathcal{C} = \underbrace{\mathcal{C} \oplus \mathcal{C} \oplus \cdots \oplus \mathcal{C}}_{b\text{-krát}}$.

Tvrzení: Kód $b \times \mathcal{C}$ je $\big(bn, M, bd\big)$-kód.

Věta (Levenshtein): Pokud existují příslušné Hadamardovy matice, pak ve všech nerovnostech věty s názvem Plotkinova mez platí rovnost.

Důkaz: Nejdříve ukážeme, že ve 2. a 4. nerovnosti Plotkinovy meze na $A(n, d)$ platí rovnost.

2. *d je sudé a $n = 2d$*: Zřejmě n je dělitelné 4, a existuje-li Hadamardova matice řádu n, tak B_n je $(2d, 4d, d)-$kód, tedy má $4d$ slov. Proto $A(2d, d) = 4d$.

4. *d je liché a $n = 2d + 1$*: Podle Lemmatu 1 platí

$$A(2d + 1, d) = A(2d+2, \underbrace{d+1}_{\text{sudé}}) = 4(d + 1)$$

díky právě dokázanému tvrzení pro sudou minimální vzdálenost.

K důkazu rovnosti v 1. bodě Plotkinovy meze využijme výše popsaných kódů $\mathcal{A}$ a $\mathcal{A}'$.

1. *d je sudé a $2d > n$*: Hledáme kód $\mathcal{C}$ délky n s minimální vzdáleností alespoň d, který má $2k$ slov pro $k = \left\lfloor \frac{d}{2d-n} \right\rfloor$. Protože $k \leq \frac{d}{2d-n} < k + 1$, existují $a, b \in \mathbb{N}_0$, že

$$k = \frac{d - b}{2d - n} \quad \text{a} \quad k + 1 = \frac{d + a}{2d - n}. \tag{5.1}$$

Odečtením předchozích rovností a vynásobením jmenovatelem získáme $2d-n = a+b$. Po dosazení do jmenovatele prvního vztahu v (5.1) dostaneme

$$(a + b)k = d - b\,, \tag{5.2}$$

a odtud

$$(a + b)2k + b - a = 2(d - b) + b - a = 2d - b - a = n\,. \tag{5.3}$$

Konstrukce hledaného kódu $\mathcal{C}$ bude záviset na paritě parametru k.

- Pokud je k sudé, pak z (5.1) plyne, že $d - b$ je sudé, a tedy i b sudé. Položme

$$\mathcal{C} = a \times \mathcal{A}_{2k} \oplus \frac{b}{2} \times \mathcal{A}'_{4k+4}.$$

Parametry tohoto kódu spočteme pomocí pravidel pro skládání kódů a z parametrů Hadamardových kódů $\mathcal{A}$ a $\mathcal{A}'$. Předně kódová slova mají délku

$$a(2k - 1) + \frac{b}{2}(4k + 2) = (a + b)2k + b - a = n,$$

kde poslední rovnost plyne z (5.3). Dále minimální vzdálenost je alespoň

$$ak + \frac{b}{2}(2k + 2) = (a + b)k + b = d,$$

kde poslední rovnost využívá (5.2). Konečně $|\mathcal{C}| = \min\{2k, \frac{4k+4}{2}\} = 2k$.

- Je-li k liché, tak k+1 a $d+a$ jsou sudé, a díky tomu je i a je sudé. Položme

$$\mathcal{C} = \frac{a}{2} \times \mathcal{A}'_{4k} \oplus b \times \mathcal{A}_{2k+2}.$$

Ověření, že se skutečně jedná o $(n, 2k, d)-$kód, přenecháme čtenáři.

Zbývá dokázat rovnost v 3. bodu Plotkinovy meze. Předpokládáme $2d+1 > n$.

3. *d je liché*: Protože $2(d+1) = 2d+2 > n+1$, opět můžeme pomocí Lemmatu 1 převést tento problém na již dokázaný 1. bod Plotkinovy meze. Platí tedy, že

$$A(n,d) = A(n{+}1, d{+}1) = 2 \left\lfloor \frac{d{+}1}{2d+1-n} \right\rfloor.$$

Kapitola 6

Konstrukce Hamadardových matic

V minulé kapitole jsme viděli, že v jistém smyslu nejlepší kódy se získávají pomocí Hadamardových matic. Teď si popíšem dvě konstrukce Hadamardových matic. První z nich je založena na tenzorovém součinu matic, druhá vyžaduje znalosti z teorie čísel o kvadratických reziduích.

6.1 Sylvestrova kontrukce

Definice: Nechť $A \in \mathbb{R}^{n\times n}$ a $B \in \mathbb{R}^{s\times s}$ jsou čtvercové matice. Jejich tenzorový součin definujeme po blocích velikosti $s \times s$ takto

$$A \otimes B = \left(\begin{array}{c|c|c|c} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & & & \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{array}\right) \in \mathbb{R}^{(ns)\times(ns)}$$

Snadno nahlédneme přímo z definice, že pro počítání s tenzorovým součinem platí tato pravidla

1. $(A \otimes B)^\top = A^\top \otimes B^\top$
2. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, pokud AC a BD má smysl.

Věta: (Sylvestrova konstrukce) Nechť H_n a H_m jsou Hadamardovy matice. Pak $H_n \otimes H_m$ je Hadamardova matice.

Důkaz: Z definice tenzorového součinu je zřejmé, že prvky matice $H_n \otimes H_m$ jsou pouze 1 a -1.

$$\begin{aligned}(H_n \otimes H_m)(H_n \otimes H_m)^\top = (H_n \otimes H_m)(H_n^\top \otimes H_m^\top) = (H_n H_n^\top) \otimes (H_m H_m^\top) = \\ = (n \cdot I_n) \otimes (m \cdot I_m) = nm \cdot (I_n \otimes I_m) = nm \cdot I_{nm}.\end{aligned}$$

Když do tenzorového součinu dosadíme $H_2 = \begin{pmatrix} 1 & 1 \\ 1 & \overline{1} \end{pmatrix}$, předchozí věta dává

Důsledek: Je-li $H_n \in \mathbb{R}^{n\times n}$ Hadamardova matice, pak $\begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix} \in \mathbb{R}^{2n\times 2n}$ je taky Hadamardova matice.

6.2 Kvadratická rezidua

Definice: Nechť p je prvočíslo. Řekneme, že $x \in \mathbb{Z}_p \setminus \{0\}$ je kvadratické reziduum mod p, pokud $x = k^2 \mod p$ pro nějaké $k \in \mathbb{Z}_p$. Množinu kvadratických reziduí mod p značíme $\mathcal{R}_p$.

Příklad:

$$\begin{aligned} p = 7: \quad & 1 \equiv 1^2 \mod 7,\ 4 \equiv 2^2 \mod 7,\ 2 \equiv 3^2 \mod 7, \\ & 2 \equiv 4^2 \mod 7,\ 4 \equiv 5^2 \mod 7,\ 1 \equiv 6^2 \mod 7. \end{aligned}$$

Tedy $\mathcal{R}_7 = \{1, 2, 4\}$ a čísla 3,5,6 nejsou kvadratická rezidua mod 7. Říkáme jim kvadratická nerezidua.

Věta: Nechť p je prvočíslo, $p > 2$. Pak v množině $\mathbb{Z}_p \setminus \{0\}$ přesně polovina prvků je kvadratickým reziduem $\mod p$.

Důkaz. Uvažujme $1^2, 2^2, \ldots, (p-1)^2 \mod p$. Pro $a = 1, 2, \ldots, p-1$ zřejmě platí rovnost

$$(p-a)^2 = p^2 - 2pa + a^2 = a^2 \mod p\,.$$

Proto reziduí je maximálně $\frac{p-1}{2}$ a pro jejich získání stačí probrat a^2, kde $a = 1, 2, \ldots, \frac{p-1}{2}$.

Ještě ukážeme, že k různým a z této množiny dostaneme různá rezidua. Až teď využijeme toho, že p je prvočíslo, a tedy $\mathbb{Z}_p$ je těleso. Nechť $a_1, a_2 \in \{1, 2, \ldots, \frac{p-1}{2}\}$. Kdyby $a_1^2 = a_2^2 \mod p$, pak

$$a_1^2 - a_2^2 = (a_1 - a_2)\underbrace{(a_1 + a_2)}_{1\le \quad \le p-1} = 0 \mod p.$$

Protože $\mathbb{Z}_p$ je těleso, součin dvou čísel je 0, pouze když alespoň jedno číslo je 0 v $\mathbb{Z}_p$. A to musí být $a_1 - a_2 = 0$. Tedy rovnost $a_1^2 \equiv a_2^2 \mod p$ implikuje $a_1 = a_2$.

□

Lemma: Nechť p je prvočíslo a $a, b \in \mathbb{Z}_p \setminus \{0\}$.

1. Když $a, b \in \mathcal{R}_p$, pak $a.b \in \mathcal{R}_p$.
2. Když $a \in \mathcal{R}_p$ a $b \notin \mathcal{R}_p$, pak $a.b \notin \mathcal{R}_p$.
3. Když $a, b \notin \mathcal{R}_p$, pak $a.b \in \mathcal{R}_p$.

Důkaz: Přenecháno čtenáři.

Uvědomme si, že $1 = 1^2$ je kvadratické reziduum pro libovolné p, tj. $1 \in \mathcal{R}_p$. V dalším povídání bude důležité vědět, kdy $-1 = p-1 \mod p$ je kvadratickým reziduem. Na příkladu $p = 7$ jsme viděli, že $-1 = 6 \notin \mathcal{R}_7$.

Věta: Nechť p je prvočíslo, $p > 2$. Pak $-1 \in \mathcal{R}_p$ právě tehdy, když $p - 1$ je dělitelné číslem 4.

Důkaz: Na množině $\mathbb{Z}_p \setminus \{0\}$ definujeme ekvivalenci takto: $x \sim y$, pokud $x = \pm y$ nebo $x = \pm y^{-1}$. Snadno lze ověřit, že to je skutečně ekvivalence.

Třída ekvivalence je tvořena prvky $x, -x, x^{-1}, -x^{-1}$ a je tedy nanejvýš čtyřprvková. Ale některé prvky v ní mohou splynout. Určitě nesplynou x a $-x$. Kdyby totiž $x = -x$, pak $2x = 0$, a součin dvou nenulových prvků v tělese by dával 0, což je spor. Ze stejných důvodů $x^{-1} \neq -x^{-1}$.

Třída, která obsahuje 1, je určitě dvouprvková $\{1, -1\}$. To splynuly $x = x^{-1}$, neboli kořeny rovnice $x^2 = 1$, a to jsou $\{1, -1\}$.

Jestli by existovala další třída s méně než 4 prvky, tak v ní by musely splynout $x = -x^{-1}$, a pak už nutně i $-x = x^{-1}$. Tedy by tato třída byla dvouprvková a obsahovala by kořeny rovnice $x^2 = -1 \mod p$. Jinými slovy, existence další dvouprvkové třídy znamená, že -1 je kvadratické reziduum $\mod p$. Naše ekvivalence je definovaná na množině $\mathbb{Z}_p \setminus \{0\}$, která má $p-1$ prvků. Všechny třídy ekvivalence kromě jedné nebo dvou tříd jsou čtyřprvkové. Proto dvě dvouprvkové třídy existují právě tehdy, když číslo 4 dělí $p-1$.

Definice: Nechť p je prvočíslo. Pro prvky tělesa $\mathbb{Z}_p$ definujeme Legendreův symbol $\chi : \mathbb{Z}_p \to \{0, 1, -1\}$ takto:

$$\begin{array}{rcll} \chi(0) & = & 0\,, & \\ \chi(i) & = & 1, & \text{když } i \in \mathcal{R}_p\,, \\ \chi(i) & = & -1, & \text{když } i \notin \mathcal{R}_p\,. \end{array}$$

Z vlastnosti kvadratických reziduí popsaných v lemmatu plyne

$$\chi(a \cdot b) = \chi(a)\chi(b) \quad \text{pro každé } a, b \in \mathbb{Z}_p.$$

Věta: Nechť p je prvočíslo. Pak platí

$$\underbrace{\sum_{b=0}^{p-1} \chi(b)\chi(\overbrace{b+c}^{\text{součet v } \mathbb{Z}_p})}_{\text{součet v } \mathbb{R}} = \begin{cases} -1, & \texttt{když } c \in \mathbb{Z}_p \setminus \{0\}, \\ p-1, & \texttt{když } c = 0. \end{cases} \tag{6.1}$$

Důkaz: Nejdříve uvažujme pevné $c \neq 0$ a zobrazení $x \in \mathbb{Z}_p \setminus \{0\} \to \mathbb{Z}_p$ dané předpisem $f(x) = x^{-1}(x + c) = 1 + x^{-1}c$. Ukažme, že to je prosté zobrazení:

$$f(x_1) = f(x_2) \iff 1 + x_1^{-1}c = 1 + x_2^{-1}c \iff x_1^{-1} = x_2^{-1} \iff x_1 = x_2.$$

Protože zobrazení je prosté, má obor hodnot stejný počet prvků jako definiční obor, tedy $p - 1$. Jelikož $x^{-1}c \neq 0$, číslo 1 není v oboru hodnot, a proto $f(x) = x^{-1}(x + c)$ probíhá prvky $0, 2, 3, \ldots, p-1$ (obraz množiny $\mathbb{Z}_p \setminus \{0\}$).

V následujících úpravách využijme, že $xf(x) = x + c$ a dvou vlastnosti Legendreaova symbolu, a to $\chi(x \cdot f(x)) = \chi(x)\chi(f(x))$ a $\chi(0) = 0$.

$$\sum_{b=0}^{p-1} \chi(b)\chi(b+c) = \sum_{b=1}^{p-1} \chi(b)\chi(\underbrace{b+c}_{bf(b)}) = \sum_{b=1}^{p-1} \underbrace{(\chi(b))^2}_{1}\chi(f(b)) =$$

$$= \sum_{k\in\mathbb{Z}_p\setminus\{1\}} \chi(k) = \sum_{k\in\mathbb{Z}_p\setminus\{0,1\}} \chi(k) = -1$$

Poslední rovnost plyne z toho, že reziduí a nereziduí je stejně v $\mathbb{Z}_p \setminus \{0\}$ a 1 je reziduum.
Ukázat platnost věty pro $c = 0$ je už jednoduché:

$$\sum_{b=0}^{p-1} \chi(b)\chi(b+c) = \sum_{b=1}^{p-1} \chi(b)\chi(b) = \sum_{b=1}^{p-1} 1 = p-1.$$

6.3 Paleyova konstrukce

Tato konstrukce umožňuje nalézt Hadamardovy matice řádu $p+1$, pokud p je prvočíslo a $p = 3 \mod 4$ a taky řádu $2(p+1)$, pokud p je prvočíslo a $p = 1 \mod 4$. Konstrukce využívají Jacobstahlovy matice.

Definice: Nechť $p \neq 2$, p prvočíslo. Jacobstahlovou matici Q rozměru $p \times p$, jejíž řádky a sloupce jsou číslovány indexy $0, 1, \ldots, p-1$ definujeme takto:

$$Q_{ij} = \chi(j - i).$$

Příklad: Nechť $p = 7$. Protože $\mathcal{R}_7 = \{1, 2, 4\}$ a nerezidua jsou $3, 5, 6$, je 0-tý řádek matice Q roven $0, 1, 1, \overline{1}, 1, \overline{1}, \overline{1}$. Celá matice pak vznikne cyklickými posuny nultého řádku.

$$Q = \begin{pmatrix} 0 & 1 & 1 & \overline{1} & 1 & \overline{1} & \overline{1} \\ \overline{1} & 0 & 1 & 1 & \overline{1} & 1 & \overline{1} \\ \overline{1} & \overline{1} & 0 & 1 & 1 & \overline{1} & 1 \\ 1 & \overline{1} & \overline{1} & 0 & 1 & 1 & \overline{1} \\ \overline{1} & 1 & \overline{1} & \overline{1} & 0 & 1 & 1 \\ 1 & \overline{1} & 1 & \overline{1} & \overline{1} & 0 & 1 \\ 1 & 1 & \overline{1} & 1 & \overline{1} & \overline{1} & 0 \end{pmatrix}$$

Všimneme si, že matice Q je antisymetrická.

Když $p = 5$, pak $\mathcal{R}_5 = \{1, 4\}$ a Jacobstahlova matice má tvar

$$Q = \begin{pmatrix} 0 & 1 & \overline{1} & \overline{1} & 1 \\ 1 & 0 & 1 & \overline{1} & \overline{1} \\ \overline{1} & 1 & 0 & 1 & \overline{1} \\ \overline{1} & \overline{1} & 1 & 0 & 1 \\ 1 & \overline{1} & \overline{1} & 1 & 0 \end{pmatrix}$$

Tato matice je symetrická.

Lemma: Nechť p je prvočíslo, $p > 2$, a Q je Jacobstahlova matice řádu p.

- Pokud $p = 3 \mod 4$, pak $Q^\top = -Q$.
- Pokud $p = 1 \mod 4$, pak $Q^\top = Q$.

Důkaz: Připomeňme, že -1 je reziduum právě tehdy, když $p = 1 \mod 4$. V Legendreově symbolice: $\chi(-1) = -1$, když $p = 3 \mod 4$ a $\chi(-1) = 1$, když $p = 1 \mod 4$. Odtud

$$Q_{ij} = \chi(j-i) = \chi((-1)\cdot(i-j)) = \chi(-1)\cdot Q_{ji}.$$

Lemma: Nechť Q, I a J jsou čtvercové matice řádu p, po řadě Jacobstahlova, jednotková a matice, která má všechny prvky rovny 1. Pak

$$QQ^\top = pI - J \quad \text{a} \quad QJ = JQ = \Theta\,.$$

Důkaz: Diagonální prvky matice $QQ^\top$ jsou $\left(QQ^\top\right)_{ii} = \sum_{k=0}^{p-1} q_{ik}^2 = \sum_{\substack{k=0\\k\neq i}}^{p-1} 1 = p-1$.

Pro výpočet mimodiagonálních prvků $\left(QQ^\top\right)_{ij}$, kde $i \neq j$, použijeme lemma

$$\left(QQ^\top\right)_{ij} = \sum_{k=0}^{p-1} q_{ik}q_{jk} = \sum_{k=0}^{p-1} \chi(k-i)\chi(k-j) \underbrace{=}_{\substack{\texttt{Označme } b=k-i\\ c=i-j\neq 0\\ \texttt{a změníme sčítací index}}} \sum_{b=0}^{p-1} \chi(b)\chi(b+c) = -1.$$

Tím je dokázáno první tvrzení.

Abychom ukázali $QJ = JQ = \Theta$, stačí si uvědomit, že $Q\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$ = součet všech sloupců matice Q. V každém řádku matice Q je jedna 0 a stejný počet jedniček a minus jedniček, protože reziduí je stejně jako nereziduí. Proto je součet prvků v každém rádku matice Q roven 0.

Teď už můžeme popsat dvě Paleyovy konstrukce Hadamardových matic.

Věta (Paleyova konstrukce I): Nechť $n = p+1$, kde p je prvočíslo, $p = 3 \mod 4$. Nechť Q je Jacobstahlova matice řádu p. Položme

$$H = \left(\begin{array}{c|c} 1 & 11\ldots1 \\ \hline 1 & \\ 1 & Q - I_p \\ \vdots & \\ 1 & \end{array}\right) \in \mathbb{R}^{n\times n}.$$

Pak H je Hadamardova matice.

Důkaz: Máme ověřit, že $HH^\top = nI$. K tomu účelu označíme vektor $j = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{p\times 1}$.

Uvědomme si, že $j\,j^\top = J \in \mathbb{R}^{p\times p}$ a $j^\top j = p \in \mathbb{R}$.

Matice $HH^\top$ je zřejmě symetrická a v blokovém zápisu má tvar $HH^\top = \begin{pmatrix} A & B^\top \\ B & C \end{pmatrix}$.

K ověření hadamardovskosti matice H máme ukázat, že $A = n$, $B = 0 \in \mathbb{R}^{p\times 1}$ a $C = nI \in \mathbb{R}^{p\times p}$. Samotné H má tvar $H = \begin{pmatrix} 1 & j^\top \\ j & Q - I_p \end{pmatrix}$. Vynásobením $HH\top$ zjistíme, čemu se matice A, B, C rovnají.

$$A = 1 + j^\top j = 1 + p = n, \qquad B = j + (Q - I)j = Qj = 0 \in \mathbb{R}^{p\times 1}$$

$$C = j\,j^\top + (Q - I)(Q^\top - I) = J + \underbrace{QQ^\top}_{=pI-J} - Q - Q^\top + I = (p+1)I - Q - Q^\top = nI.$$

Poslední úprava využila antisymetrie matice Q.

Příklad: Pomocí Paleyovy konstrukce získáme matici H_8. To lze, protože prvočíslo $p = 7 = 3 \mod 4$. Využijeme Jacobstahlovou matici 7×7, kterou jsem našli v předchozím příkladě.

$$H_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \overline{1} & 1 & 1 & \overline{1} & 1 & \overline{1} & \overline{1} \\ 1 & \overline{1} & \overline{1} & 1 & 1 & \overline{1} & 1 & \overline{1} \\ 1 & \overline{1} & \overline{1} & \overline{1} & 1 & 1 & \overline{1} & 1 \\ 1 & 1 & \overline{1} & \overline{1} & \overline{1} & 1 & 1 & \overline{1} \\ 1 & \overline{1} & 1 & \overline{1} & \overline{1} & \overline{1} & 1 & 1 \\ 1 & 1 & \overline{1} & 1 & \overline{1} & \overline{1} & \overline{1} & 1 \\ 1 & 1 & 1 & \overline{1} & 1 & \overline{1} & \overline{1} & \overline{1} \end{pmatrix}$$

Jinou Hadamardovou matici rozměru 8×8 lze získat Sylvestrovou konstrukcí pomocí tenzorových součinů matice $H_2 = \begin{pmatrix} 1 & 1 \\ 1 & \overline{1} \end{pmatrix}$.

$$H_8 = H_2 \otimes H_2 \otimes H_2 = H_2 \otimes \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \overline{1} & 1 & \overline{1} \\ 1 & 1 & \overline{1} & \overline{1} \\ 1 & \overline{1} & \overline{1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \overline{1} & 1 & \overline{1} & 1 & \overline{1} & 1 & \overline{1} \\ 1 & 1 & \overline{1} & \overline{1} & 1 & 1 & \overline{1} & \overline{1} \\ 1 & \overline{1} & \overline{1} & 1 & 1 & \overline{1} & \overline{1} & 1 \\ 1 & 1 & 1 & 1 & \overline{1} & \overline{1} & \overline{1} & \overline{1} \\ 1 & \overline{1} & 1 & \overline{1} & \overline{1} & 1 & \overline{1} & 1 \\ 1 & 1 & \overline{1} & \overline{1} & \overline{1} & \overline{1} & 1 & 1 \\ 1 & \overline{1} & \overline{1} & 1 & \overline{1} & 1 & 1 & \overline{1} \end{pmatrix}$$

Proto je Paleyova kontrukce užitečnější při určení H_{12}. To přenecháme čtenáři.

Věta (Paley konstrukce II): Nechť $n = 2(p+1)$, kde p je prvočíslo, $p = 1 \mod 4$.

Položme

$$H = I_{p+1} \otimes A + \begin{pmatrix} 0 & 11\cdots 11 \\ 1 & \\ \vdots & Q \\ 1 & \end{pmatrix} \otimes B, \quad \text{kde } A = \begin{pmatrix} 1 & \overline{1} \\ \overline{1} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & \overline{1} \end{pmatrix}$$

a Q je Jacobstahlova matice řádu p. Pak H je Hadamardova matice řádu $n = 2(p+1)$.

Důkaz Protože Jacobstahlova matice Q je symetrická a A, B, I_{p+1} jsou symetrické je H symetrická. Při ověřování hadamardovskosti použijeme

$$A^2 = B^2 = 2I_2 \qquad \text{a} \qquad AB + BA = \Theta\,,$$

vlastnosti násobení tenzorového součinu, jak jsme je uvedli na začátku kapitoly, a faktu, že teď je Q symetrická. Detaily důkazu přenecháme čtenáři.

Kapitola 7

Dodatek Konečná tělesa

Cílem této kapitoly je připomenout některé znalosti z algebry o konečných tělesech. O tyto znalosti se opírají metody konstrukce lineárních kódů. U tvrzení většinou chybějí důkazy - ty lze najít v každé učebnici obecné algebry. Tvrzení jsou však doplněna příklady a neřešenými úlohami, to aby si čtenář mohl hned ověřit, zda zavedenému pojmu nebo tvrzení dobře rozumí.

Definice: Nechť M je neprázdná množina a $\circ : M \times M \to M$ binární operace na M s vlastnostmi:

1. $a \circ b = b \circ a$ pro každé $a, b, \in M$ (komutativita)
2. $a \circ (b \circ c) = (a \circ b) \circ c$ pro každé $a, b, c \in M$ (asociativita)
3. existuje $e \in M$ takové, že $e \circ a = a$ pro každé $a \in M$ (e nazýváme neutrální prvek)
4. pro každé $a \in M$ existuje $b \in M$ takové, že $a \circ b = e$ (b nazýváme inverzní prvek k prvku a).

Množinu M spolu s operací $\circ$ značíme $(M, \circ)$ a nazývýme *abelovská grupa*, dále jen grupa.

Příklad V tomto příkladě uvažujeme podmnožiny reálných čísel a sčítání $+$ a násobení $\cdot$ jak jsou obvykle zavedeny na $\mathbb{R}$.

$(\mathbb{Z}, +)$ - je grupa (Co je neutrální prvek e ? Co je inverzní prvek k číslu 5 ?)

$(\mathbb{N}, +)$ - není grupa (Proč?)

$(\mathbb{R}, +)$ - je grupa

$(\mathbb{R}, \cdot)$ - není grupa (Proč?)

$(\mathbb{R} \setminus \{0\}, \cdot)$ - je grupa

$(\mathbb{Z} \setminus \{0\}, \cdot)$ - není grupa (Proč?)

Kromě operace + a $\cdot$ na $\mathbb{R}$ máme i operaci odčítání a dělení. Odčítání je vlastně přičítání inverzního prvku (nazývaného opačným prvkem) v grupě $(\mathbb{R}, +)$ a dělení je násobení inverzním prvkem (nazývaným převráceným prvkem) v grupě $(\mathbb{R} \setminus \{0\}, \cdot)$.

Definice: Nechť $p \in \mathbb{N}$. Na množině $\{0, 1, \ldots, p-1\}$ definujme operace sčítání $\oplus$ a násobení $\otimes$ takto

$$a \oplus b = \text{zbytek čísla } a + b \text{ po dělení číslem } p;$$

$$a \otimes b = \text{zbytek čísla } a \cdot b \text{ po dělení číslem } p.$$

Množinu $\{0, 1, \ldots, p-1\}$ spolu s operacemi $\oplus$ a $\otimes$ značíme $\mathbb{Z}_p$.

Příklad V $\mathbb{Z}_7$ platí: $4 \oplus 5 = 2$, $4 \otimes 5 = 6$

$3 \oplus 4 = 0$, tj. 4 je inverzní (opačný) prvek k prvku 3 při operaci sčítání $\oplus$

$3 \otimes 5 = 1$, tj. 5 je inverzní (převrácený) prvek k prvku 3 při operaci násobení $\otimes$

V $\mathbb{Z}_6$ například platí $4 \oplus 5 = 3$, $4 \otimes 5 = 2$. Opačný prvek při operaci sčítání k číslu 2 je 4, protože $2 \oplus 4 = 0$. Kdežto inverzní prvek (převrácený) k číslu 2 vzhledem na nasobení neexistuje, protože

$$2 \otimes 1 = 2, \quad 2 \otimes 2 = 4, \quad 2 \otimes 3 = 0, \quad 2 \otimes 4 = 2, \quad 2 \otimes 5 = 4.$$

Tedy výsledek násobení prvku 2 s žádným jiným prvkem není roven 1.

Tvrzení Množina $\{0, 1, \ldots, p-1\}$ s operací $\oplus$ je grupa pro každé $p \in \mathbb{N}$. Množina $\{1, \ldots, p-1\}$ s operací $\otimes$ je grupa právě tehdy, když p je prvočíslo.

Věta: Nechť G je konečná grupa (s operací násobení) a nechť $a \in G$. Pak existuje přirozené j tak, že $a^j = 1$. Minimální takové j nazýváme *řád prvku* a v grupě G.

Důkaz Když $a = 1$, tak zřejmě $j = 1$. Uvažujme $a \neq 1$. Protože G je konečná, mezi prvky $a^0 = 1, a^1, a^2, a^3, \ldots$ existuje index $j > 0$ takový, že $a^j = a^i$ pro nějaký menší index i, tj. $0 \leq i < j$. Vezmime minimálné takové j, tj. prvně, kdy se mocnina trefí do nějaké mocniny před ní. Kdyby $i > 0$, pak po vynásobení rovnosti $a^i = a^j$ prvkem inverzním k a dostaneme $a^{i-1} = a^{j-1}$, a to by byl spor s minimalitou j. Tedy $a^j = 1$ a toto j je řádem prvku a.

Příklad: V grupě $\{1, \ldots, 6\}$ s operací násobení $\mod 7$ má prvek $a = 3$ řád $j = 6$, protože

$$3^1 = 3, \quad 3^2 = 2, \quad 3^3 = 6, \quad 3^4 = 4, \quad 3^5 = 5, \quad 3^6 = 1 \tag{7.1}$$

Prvek $a = 6$ má řád $j = 2$, protože $6^2 = 1 \mod 7$. Snadno se přesvědčíme, že žádný prvek nemá řád 4 nebo 5.

Řád prvku v grupě může nabývat jenom některé hodnoty.

Věta: Nechť G je konečná grupa s N prvky. Pak řád prvku v G je dělitelem čísla N.

Definice: Pokud v grupě G existuje prvek $b \in G$ takový, že $G = \{b^i : i = 1, 2, 3, \ldots\}$, pak se grupa nazývá *cyklická* a prvek b se nazývá *generátor* grupy G.

Všechny prvky cyklické grupy získáme mocněním jediného prvku grupy.

Příklad: Grupa $\{1, \ldots, 6\}$ s operací násobení $\mod 7$ je cyklická a jak dokládá (7.1),

číslo $b = 3$ je jejím generátorem.

Důsledek: Pokud je G cyklická grupa s N prvky, tak každý dělitel čísla N je řádem některého prvku grupy G.

Důkaz: Nechť b je genrátorem grypy G. Pak nutně $b^N = 1$ a $G = \{b^1, b^2, \ldots, b^N\}$. Nechť d je dělitelem čísla N, tj. $N = d.r$ pro nějaké přirozené r. Potom prvek $a = b^r$ má řád d. Snadno se ověří, že $a^d = b^{rd} = 1$.

Jak jsme už poznamenali, $(\mathbb{R}, +)$ a $(\mathbb{R} \setminus \{0\}, \cdot)$ jsou grupy. V reálných číslech jsou operace násobení a sčítání navzájem provázány a kromě asociativního a komutativního zákona pro $+$ a $\cdot$ platí navíc distributivní zákon. Vlastnosti reálných čísel jsou abstrahovány do definice tělesa.

Definice: Nechť M je alespoň dvouprvková množina a $\oplus$ a $\otimes$ jsou binární operace na M takové, že

1. $(M, \oplus)$ tvoří grupu - její neutrální prvek označíme 0;
2. $(M \setminus \{0\}, \otimes)$ tvoří grupu - její neutrální prvek označíme 1;
3. pro každé $a, b, c \in M$ platí $(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$.

Pak trojici $(M, \oplus, \otimes)$ nazveme *těleso*.

Nechceme-li zavádět pojem grupa, můžeme použít **alternativní definici tělesa**.

Definice: Těleso je alespoň dvouprvková množina M s binárními operacemi $\oplus$ a $\otimes$ na M takovými, že platí

1. asociativní zákony: $a \oplus (b \oplus c) = (a \oplus b) \oplus c$ a $a \otimes (b \otimes c) = (a \otimes b) \otimes c$ pro každé $a, b, c \in M$;
2. komutativní zákony: $a \oplus b = b \oplus a$ a $a \otimes b = b \otimes a$ pro každé $a, b \in M$;
3. distributivní zákon: $(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$ pro každé $a, b, c \in M$;
4. existuje neutrální prvek 0 vzhledem k operaci $\oplus$ a neutrální prvek $1 \neq 0$ vzhledem k operaci $\otimes$, tj. prvky takové, že pro každé $a \in M$ platí $a \oplus 0 = a$ a $a \otimes 1 = a$;
5. ke každému prvku $a \in M$ existuje opačný prvek vzhledem k operaci $\oplus$, tj. prvek, značíme jej $-a$, pro který platí $a \oplus (-a) = 0$;
6. ke každému prvku $a \in M \setminus \{0\}$ existuje převrácený prvek vzhledem k operaci $\otimes$, tj. prvek, značíme jej a^{-1}, pro který platí $a \otimes \left(a^{-1}\right) = 1$.

Příklad: $\mathbb{Z}_p$ splňuje podmínky 1. - 5. pro každé přirozené p. Podmínka 6. je splněná pouze pokud p je prvočíslo.

Důsledek: $\mathbb{Z}_p$ je těleso právě tehdy, když p je prvočíslo.

Definice: Alespoň dvouprvková množina M s binárními operacemi $\oplus$ a $\otimes$ na M taková, že platí podmínky 1.- 5. z alternativní definice tělesa, se nazývá *okruh*.

Příklad: $\mathbb{Z}_p$ je okruh pro každé přirozené p.

V okruhu lze sčítat a odčítat, násobit, ale obecně nelze dělit! Příkladem takové struktury jsou polynomy s reálnými koeficienty s operacemi, jak je známe. Výsledkem součtu, rozdílu i násobení dvou polynomů je polynom. Výsledkem podílu dvou polynomů je polynom pouze tehdy, když zbytek po dělení polynomu polynomem je nula. To je ale málo kdy. Většina podílů dvou polynomů, už polynomem není. Proto tato množina je okruh, ale ne těleso.

Příklad a úmluva: Nechť $\mathbb{T}$ je těleso. Pak množinu všech polynomů s proměnnou x s koeficienty z tělesa $\mathbb{T}$ značíme $\mathbb{T}[x]$. Operaci sčítání polynomů a násobení polynomů definujeme obvyklým způsobem, přičemž operaci s koeficienty provádíme v tělese $\mathbb{T}$. Snadno nahlédneme, že $\mathbb{T}[x]$ je okruh.

Např. pro polynomy $f(x) = x^2 + x + 1$ a $g(x) = x + 1$ v okruhu $\mathbb{Z}_2[x]$ platí

$$f(x)+g(x) = x^2+x+1+x+1 = x^2 \quad \text{a} \quad f(x)\cdot g(x) = x^3+x^2+x+x^2+x+1 = x^3+1\,,$$

protože v $\mathbb{Z}_2$ je $1 + 1 = 0$.

V okruhu $\mathbb{T}[x]$ lze provádět **dělení polynomu polynomem**, jak jsme zvyklí u polynomů s koeficienty v $\mathbb{R}$, jenom teď operace s koeficienty provádíme v tělese $\mathbb{T}$.

Příklad: Vydělme $f(x) = x^3 + x + 1$ polynomem $g(x) = x + 1$ v okruhu $\mathbb{Z}_2[x]$. Opět využíváme, že $1 + 1 = 0$, a tedy $-1 = 1$.

v prvním kroku

$$(x^3 + x + 1) : (x + 1) = x^2$$
$$\underline{-(x^3 + x^2)}$$
$$x^2 + x + 1$$

v druhém kroku

$$(x^3 + x + 1) : (x + 1) = x^2 + x$$
$$\underline{-(x^3 + x^2)}$$
$$x^2 + x + 1$$
$$\underline{-(x^2 + x)}$$
$$1$$

Tedy zbytek je 1 a můžeme napsat $f(x) = (x^2 + x) \cdot g(x) + 1$.

Definice: Nechť $q(x) \in \mathbb{T}[x]$ je pevně daný polynom stupně $m \in \mathbb{N}$. Pro dvojici polynomů $f(x), g(x) \in \mathbb{T}[x]$, které mají stupeň ostře menší než m,

- operaci $\oplus$ definujeme tak, že sčítáme koeficienty jako v tělese $\mathbb{T}$ u stejných mocnin proměnné x;

• operaci $\otimes$ definujeme takto:
$f(x) \otimes g(x) =$ zbytek po dělení součinu $f(x) \cdot g(x)$ polynomem $q(x)$ v $\mathbb{T}[x]$.

Množinu $\{f(x) \in \mathbb{T}[x] \ : \text{st}\, f < m\}$ s operacemi $\oplus$ a $\otimes$ značíme $\mathbb{T}[x]/q(x)$.

Věta: $\mathbb{T}[x]/q(x)$ je okruh.

V teorii cyklických kódů hraje důležitou roli okruh s $\mathbb{Z}_2[x]/q(x)$ s polynomem $q(x) = x^n - 1$. Protože koeficienty polynomu jsou ze $\mathbb{Z}_2$, platí $q(x) = x^n - 1 = x^n + 1$.

Příklad: Uvažujme okruh $\mathbb{Z}_2[x]/x^3+1$. Je tvořeny polynomy stupně ≤ 2 a koeficienty u mocnin x^i jsou 0 nebo 1. Okruh má proto 8 prvků

$$\mathbb{Z}_2[x]/x^3 + 1 = \{0,\ 1,\ x,\ 1+x,\ x^2,\ x^2+1,\ x^2+x,\ x^2+x+1\}$$

Spočtěme výsledek násobení $(x^2+1) \otimes (x^2+x)$. Pro obyčejné násobení polynomů platí
$$(x^2+1) \cdot (x^2+x) = x^4 + x^2 + x^3 + x = x(x^3+1) + (x^3+1) + x + 1.$$
Teď bychom měli výsledek dělit polynomem x^3+1 a zjistit zbytek. Ale z úpravy už je zřejmé, že zbytek je $x+1$, tj. $(x^2+1) \otimes (x^2+x) = x+1$. Neboli z pohledu nového násobení $\otimes$ lze prohlásit $x^3 + 1 = 0$.

Okruh $\mathbb{Z}_2[x]/x^3+1$ není tělesem. Součin dvou nenulových prvků $(x^2+x+1) \otimes (x+1) = 0$, a to by se v tělese nestalo.

Příklad: Uvažujme okruh $\mathbb{R}[x]/x^2 + 1$. Zřejmě $\mathbb{R}[x]/x^2 + 1 = \{a + bx \ : \ a, b \in \mathbb{R}\}$. Pro obyčejný součine dvou polynomů platí

$$(a + bx) \cdot (c + dx) = bdx^2 + ac + xcb + xad = bd(x^2 + 1) + ac - bd + x(cb + ad).$$

Z poslední úpravy je vidět zbytek po dělení polynomem $x^2 + 1$. Proto
$$(a + bx) \otimes (c + dx) = ac - bd + x(cb + ad)$$

Všimněme si, že z výsledku obyčejného součinu dvou polynomů získame výsledek nového násobení $\otimes$ tak, že $x^2 + 1$ prohlásíme za 0. Pokud proměnnou označíme písmenkem i místo písmenkem x, pak za 0 prohlašujeme výraz $i^2 + 1$. Prvky okruhu jsou tvaru $a + ib$ a násobí se podle pravidla $(a+bi) \otimes (c+di) = ac - bd + i(cb + ad)$. Tedy se jedná o komplexní čísla.

U komplexních čísel se vžilo pravidlo, že pokud proměnnou označujeme i, tak speciální značky $\oplus$ a $\otimes$ pro sčítání a násobení můžeme nahradit obvyklými znaky $+$ a $\cdot$, přičemž při operacích používame rovnost $i^2 + 1 = 0$. Tuto konvenci přeneseme i na okruhy nad konečnými tělesy.

Úmluva: Nechť p je prvočíslo. Pro proměnnou u prvků okruhu $\mathbb{Z}_p[x]/q(x)$, budeme používat písmeno z, pro násobení a sčítání budeme používat obvyklé značky $+$ a $\cdot$ a při těchto operacích využíváme identity $q(z) = 0$.

Jak jsme zmínili, v okruhu obecně nelze dělit. Jinými slovy okruh $\mathbb{T}[x]/q(x)$ nemusí, ale může být tělesem. Ve dvou předchozích příkladech jsme mohli vidět obě varianty: $\mathbb{Z}_2[x]/x^3 + 1$ není tělesem a $\mathbb{C} = \mathbb{R}[x]/x^2 + 1$ tělesem je. O tom, který z případu nastane, rozhoduje jistá vlastnost polynomu $q(x)$.

Definice: Nechť $q(x)$ je polynom s koeficienty v tělese $\mathbb{T}$, tj. $q(x) \in \mathbb{T}[x]$. Řekneme, že $q(x)$ je *reducibilní* nad $\mathbb{T}$ pokud existují polynomy $h(x), g(x) \in \mathbb{T}[x]$ stupně alespoň

1 takové, že $q(x) = h(x) \cdot g(x)$. V opačném případě je $q(x)$ *ireducibilní* nad $\mathbb{T}$.

Příklad: Polynom $x^2 + 1$ je ireducibilní nad $\mathbb{R}$. Kdyby totiž bylo možné jej zapsat jako součin dvou polynomů s reálnými koeficienty stupně alespoň jedna, měly by oba polynomy nutně stupeň 1. Kazdý polynom stupně 1 má kořen v $\mathbb{R}$, a proto by i polynom $x^2 + 1$ měl reálný kořen - spor.

Polynom $x^2 + 1$ je reducibilní nad $\mathbb{Z}_2$, protože $x^2 + 1 = (x + 1)(x + 1)$.

Polynom $x^2 + x + 1$ je ireducibilní nad $\mathbb{Z}_2$. Důvodem je fakt, že tento polynom nemá kořen v tělese $\mathbb{Z}_2$. Pokud by byl $x^2 + x + 1$ napsatelný jako součin dvou polynomů stupně 1, pak by kořen v $\mathbb{Z}_2$ měl.

Věta: Okruh $\mathbb{T}[x]/q(x)$ je tělesem právě tehdy, když $q(x)$ je ireducibilní nad $\mathbb{T}$.

Příklad: Zvolme polynom $q(x) = x^2+x+1$, který je ireducibilní nad $\mathbb{Z}_2$. Popišme explicitně těleso $\mathbb{Z}_2[x]/q(x)$. Jeho prvky jsou polynomy z $\mathbb{Z}_2[x]$ se stupněm $< \operatorname{st} q(x) = 2$. Proto

$$\mathbb{Z}_2[x]/x^2 + x + 1 = \{0, 1, z, 1 + z\}.$$

V tabulkách uvedeme výsledky operací sčítání i násobení. Při úpravách používáme vztah $z^2 + z + 1 = 0$.

součet	0	1	z	$1+z$
0	0	1	z	$1+z$
1	1	0	$1+z$	z
z	z	$1+z$	0	1
$1+z$	$1+z$	z	1	0

součin	0	1	z	$1+z$
0	0	0	0	0
1	0	1	z	$1+z$
z	0	z	$1+z$	1
$1+z$	0	$1+z$	1	z

Fakt: Nechť $q(x)$ je ireducibilní polynom nad $\mathbb{Z}_2$ a nechť $\operatorname{st} q = m$. Pak těleso $\mathbb{Z}_2[x]/q(x)$ má 2^m prvků.

Důkaz: Do tělesa patří všechny polynomy tvaru $a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1}$, kde pro každý z m koeficientů a_i máme dvě volby 0 nebo 1.

Věta: V konečném tělese $\mathbb{F}$ existuje takový prvek b, že

$$\mathbb{F} \setminus \{0\} = \{b, b^2, b^3, \ldots, \underbrace{b^{M-1}}_{=1}\}, \quad \text{kde } M \text{ je počet prvků v } \mathbb{F}\,.$$

Jinými slovy, $\mathbb{F} \setminus \{0\}$ je vzhledem k operaci násobení cyklická grupa.

Příklad V tělese $\mathbb{F} = \mathbb{Z}_2[x]/x^2 + x + 1$ je takovým prvkem (generátorem) např. $b = z$, protože $z^2 = z + 1$ a $z^3 = 1$. Prvek b z předchozí věty není dán jednoznačně. Jako generátor lze vzít i $b = z + 1$, protože $(z + 1)^2 = z$ a $(z + 1)^3 = 1$. Můžeme napsat

$$\mathbb{F} \setminus \{0\} = \{z, z^2, z^3\} = \{1 + z, (1 + z)^2, (1 + z)^3\}\,.$$

Úkol: K ireducibilnímu polynomu $q(x) = x^3 + x + 1 \in \mathbb{Z}_2[x]$ vypište všechny prvky tělesa $\mathbb{F} = \mathbb{Z}_2[x]/x^3 + x + 1$ a nalezněte generátor jeho cyklické grupy $\mathbb{F} \setminus \{0\}$.

Příklad V $\mathbb{Z}_2[x]$ existují dva ireducibilní polynomy stupně 3. Jsou to $x^3 + x + 1$ a $x^3 + x^2 + 1$. Můžeme tedy zkonstruovat dvě tělesa: $\mathbb{F}_1 = \mathbb{Z}_2[x]/x^3 + x + 1$ a

$\mathbb{F}_2 = \mathbb{Z}_2[x]/x^3 + x^2 + 1$. Každé z nich bude mít $8 = 2^3$ prvků, množiny prvků budou stejné, ale v $\mathbb{F}_1$ platí $z^3 + z + 1 = 0$, zatímco v $\mathbb{F}_2$ platí $z^3 + z^2 + 1 = 0$. Nicméně se tato tělesa moc neliší.

Definice: Dvě konečná tělesa jsou navzájem *izomorfní*, pokud existuje bijektivní zobrazení $\psi : \mathbb{F}_1 \to \mathbb{F}_2$ takové, že pro každé $a, b \in \mathbb{F}_1$ platí

$$\psi(a + b) = \psi(a) + \psi(b) \quad \text{a} \quad \psi(a \cdot b) = \psi(a) \cdot \psi(b).$$

Jinými slovy prvkům v tělese $\mathbb{F}_1$ lze dát nová jména (podle matrikáře ψ) tak, že jejich chování je nerozlišitelné od chování prvků v tělese $\mathbb{F}_2$.

(Oblíbená matematická formulace zní: Bůh je jen jeden, až na izomorfizmus!)

Tento přehled zakončíme důležitou větou o tom, jak vypadají konečná tělesa.

Věta: (Galoisova)

1. Ke každému konečném tělesu $\mathbb{F}$ existuje prvočíslo p a přirozené číslo m tak, že počet prvků v tělese $\mathbb{F}$ je p^m.
2. Těleso, které má p^m prvků, je izomorfní s tělesem $\mathbb{Z}_p[x]/q(x)$, kde $q(x) \in \mathbb{Z}_p[x]$ je ireducibilní polynom stupně m.
3. Dvě konečná tělesa se stejným počtem prvků jsou izomorfní.

Abychom mohli zkonstruovat libovolné konečné těleso, potřebujeme pro každé prvočíslo p a přirozené číslo m znát alepoň jeden ireducibilní polynom $q(x) \in \mathbb{Z}_p[x]$ stupně m.

Že takový polynom existuje, plyne z Gaussovy věty:

Věta: Nechť p je prvočíslo. Pro každé $m \in \mathbb{N}$ označme a_m počet ireducibilních polynomů stupně m z okruhu $\mathbb{Z}_p[x]$. Pak pro každé m platí

$$p^m = \sum_{d|m} d a_d \qquad \text{(součet přes všechny dělitele } d \text{ čísla } m\text{)}.$$

Příklad: V kódování nás nejvíce zajímají tělesa tvaru $\mathbb{Z}_2[x]/q(x)$. Proto pro malé hodnoty m určíme počty ireducibilních polynomů s koeficienty v $\mathbb{Z}_2$ a příslušné polynomy uvedeme v tabulce.

Zřejmě $a_1 = 2$, protože x a $x + 1$ jsou jediné polynomy stupně 1, a jsou tedy nutně ireducibilní.

Je-li m prvočíslo (jediní jeho dělitelé jsou $d = 1$ a $d = m$), pak Gaussova věta říká $2^m = a_1 + m a_m$. Odtud $a_m = \frac{1}{m}(2^m - 2)$.

Proto $a_2 = 1$, $a_3 = 2$ a $a_5 = 6$.

Pro $m = 4$ je podle Gaussovy věty: $2^4 = a_1 + 2a_2 + 4a_4$. Odtud $a_4 = 3$.

Pro $m = 6$ je podle Gaussovy věty: $2^6 = a_1 + 2a_2 + 3a_3 + 6a_6$. Odtud $a_6 = 9$.

stupeň m	počet a_m	ireducibilní polynomy
1	2	$x,\ \ x+1$
2	1	x^2+x+1
3	2	$x^3+x+1,\ \ x^3+x^2+1$
4	3	$x^4+x^3+x^2+x+1,$ $x^4+x^3+1,\ \ x^4+x+1$
5	6	$x^5+x^3+x^2+x+1,\ \ x^5+x^4+x^2+x+1,$ $x^5+x^4+x^3+x+1,\ \ x^5+x^4+x^3+x^2+1,$ $x^5+x^2+1,\ \ x^5+x^3+1,$
6	9	$x^6+x^5+x^4+x^2+1,\ \ x^6+x^5+x^4+x+1,$ $x^6+x^5+x^3+x^2+1,\ \ x^6+x^4+x^3+x+1,$ $x^6+x^4+x^2+x+1,\ \ x^6+x^5+x^2+x+1,$ $x^6+x^5+1,\ \ x^6+x+1,\ \ x^6+x^3+1$