Inhaltsverzeichnis

Kapitel 1

Codierungstheorie

1.1 Grundbegriffe und einfache Beispiele

1.1.1 Codierung

(Kanalcodierung)
Sicherung von Daten/Nachrichten gegen zufällig auftretenden Fehler bei Speicherung/Übertragung.

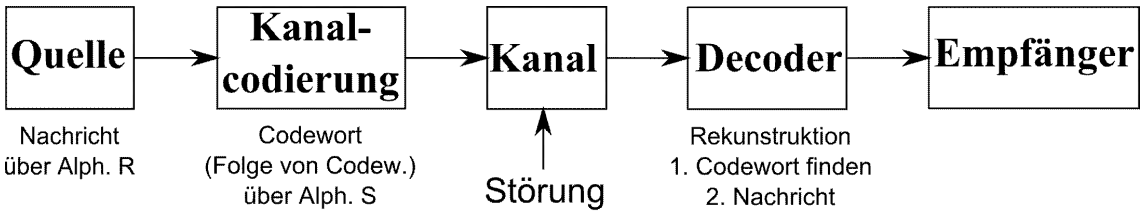

Abbildung 1.1: Schaubild der Codierung

1.1.2 Ziele

- Möglichst viele Fehler erkennen und gegebenenfalls korrigieren.
- Aufwand für Codierung und Decodierung möglichst gering.

1.2 Grundprinzip

Hinzufügen von Redundanz
Es gibt zwei Typen um Redundanz zu erzeugen.

1.2.1 FEC-Verfahren (Forward Error Correction)

Aufgetretene Fehler sollen erkannt und korrigiert werden.
Vorteil: keine Verzögerung der Übertragung aber ggf. große Redundanz notwendig.

1.2.2 ARQ-Verfahren (Automatic Repeat Request)

Aufgetretene Fehler sollen erkannt werden, werden nicht korrigiert. Stattdessen wiederholt die Übertragung beim Sender anfordern.

Vorteil: geringe Redundanz, aber Verzögerung.

Beispiele

1.2.3 Parity-Check-Codes

z.B. Nachrichten: 00, 01, 10, 11

Codierung:

$$00 \rightarrow 000$$
$$01 \rightarrow 011$$
$$10 \rightarrow 101$$
$$11 \rightarrow 110$$

(gerade Anzahl von Einsen in den Codewörtern)

1 Fehler wird erkannt, nicht korrigiert.
2 Fehler werden nicht erkannt.

1.2.4 Wiederholungscode

Nachrichten wie in 1.

Codierung:

$$00 \rightarrow 000000$$
$$01 \rightarrow 010101$$
$$10 \rightarrow 101010$$
$$11 \rightarrow 111111$$

(3-Fache Wiederholung)

1 Fehler wird erkannt und korrigiert.
$\underline{01}0101\underline{01} \to 010101 \to 01$
Nachrichten wie in 1. Codierung:

$$00 \to 00000$$
$$01 \to 01101$$
$$10 \to 10110$$
$$11 \to 11011$$

Je zwei Codewörter unterscheiden sich an mindestens 3 Positionen.
Angenommen 1 Fehler tritt bei Übertragung auf. Dann gibt es genau ein Codewort, dass sich vom empfangenen Wort an genau einer Stelle unterscheidet; in das wird decodiert.

Muss immer Ungerade unterschiede in Codewörtern sein. Bei 5 diffs sind 2 Fehler korrigierbar.

1.2.5 (ehmaliger) ISBN-Code

International Standard Book Number

10-Stelliger Code
Erste 9 Ziffern haben inhaltliche Bedingung ($\hat{=}$ Nachricht)
10. Ziffer: Prüfziffer

Beispiel: 3-540-26121-? (Land - Verlag - Buchnummer - Prüfziffer)

Uncodierte Wörter sind gebildet über $R = \{0, \ldots, 9\}$
Codierte Wörter sind gebildet über $S = \{0, \ldots, 9, X\}$

ISBN-Wort $C_{10}C_9 \ldots C_2C_1$
$C_{10} \ldots C_2$ inhaltliche Bedingung, C_1 wird so gewählt, dass

$$\sum_{k=1}^{10} k \cdot C_k \equiv 0 (\text{mod } 11)$$
$$10 \cdot C_{10} + \ldots + 2 \cdot C_2 + C_1 \equiv 0 (\text{mod } 11)$$

falls $C_1 = 10$ so setzte $C_1 = X$
C_1 vom Beispiel ausrechnen.

$$10 \cdot 3 + 9 \cdot 5 + 8 \cdot 4 + 7 \cdot 0 + 6 \cdot 2 + 5 \cdot 6 + 4 \cdot 1 + 3 \cdot 2 + 2 \cdot 1 + C_1 \equiv 0 (\text{mod } 11)$$
$$161 + C_1 \equiv 0 (\text{mod } 11) \Rightarrow C_1 = 4$$

Ändern einer Ziffer wird erkannt:
$C_{10}C_9 \ldots C_2C_1 \to C_i$ wird $X_i \neq C_i$ ersetzt

$C_{10} \dots C_{i+1} X_i C_{i-1} \dots C_1$

$$\sum_{k=1,k\neq i}^{10} k \cdot C_k + i \cdot x_i = \underbrace{\sum_{k=1,k\neq i}^{10} k \cdot C_k}_{\equiv 0 (\text{mod } 11)} \overbrace{\underset{\neq 0 (\text{mod } 11)}{\underset{\uparrow}{i}} \cdot (\underbrace{x_i - c_i}_{\neq 0 (\text{mod } 11)})}^{\neq 0 (\text{mod } 11)} \not\equiv 0 (\text{mod } 11)$$

Fehler wird erkannt, Korrektur nicht möglich.

$3 - 540 - 26121 - 4 \equiv 0 (\text{mod } 11)$

$\left. \begin{array}{l} 3 - 540 - 26121 - \mathbf{6} \\ 3 - 540 - 2612\mathbf{2} - 4 \end{array} \right\}$ Prüfsumme 2.

Vertauschung von Zwei Ziffern wird erkannt.
C_i und C_j vertauscht.
O.B.d.A $C_i \neq C_j$
$C_{10} \dots \underset{\overset{\uparrow}{i}}{C_j} \dots \underset{\overset{\uparrow}{j}}{C_i} \dots C_1$

$$\begin{aligned} \sum_{k=1,k\neq i,j}^{10} k \cdot C_k + i \cdot C_j + j \cdot C_i &= \sum_{k=1}^{10} k \cdot C_k + i(C_j - C_i) + j(C_i - C_j) \\ &= \underbrace{\sum_{k=1}^{10} k \cdot C_k}_{\equiv 0 (\text{mod } 11)} + \underbrace{(C_j - C_i)}_{\neq 0 (\text{mod } 11)} \underbrace{(i - j)}_{\neq 0 (\text{mod } 11)} \not\equiv 0 (\text{mod } 11) \end{aligned}$$

Vertauschung wird durch gewichtete Quersummen erkannt.

1.2.6 EAN-13-Code

European Article Number
13-Stelliger Code, erste 12 Ziffer sind inhaltlich festgelegt.
13. Ziffer ist Prüfziffer.
$R = S = \{0, \dots, 9\}$
$C_1 \dots C_{12} C_{13}$

$C_1 \dots C_{12}$ inhaltliche Angabe (in der Regel):
$C_1 C_2$ Herstellerland (40-43 Deutschland)
$C_6 \dots C_7$ Hersteller $C_8 \dots C_{12}$ interne Produktions Nummer

C_{13} so gewählt, dass

$$C_1 + 3 \cdot C_2 + C_3 + 3 \cdot C_4 + \dots + 3 \cdot C_{12} + C_{13} \equiv 0 (\text{mod } 10)$$

$x \to 3x$ Permutation auf $\mathbb{Z}_{10} (\text{mod } 10)$, da $ggT(3, 10) = 1$, 1 Fehler wird erkannt.
Vertauschung in der Regel nicht erkannt.

Übersetzung in Barcode:

$$C_1C_2 \dots C_7C_8 \dots C_{13}$$

Jede der Ziffern $C_2, \dots, C_{13}$ wird durch einen 0-1-String der Länge 7 binär codiert.
$0 \hat{=}$ weißer Balken, $1 \hat{=}$ schwarzer Balken.
Codierung sorgt dafür, dass nie mehr als 4 weiße oder schwarze Balken nebeneinander stehen.

Abbildung 1.2: EAN-13 Barcode

Schmalen Balken in Mitte und am Rand, sind nur Abtrennzeichen, die nichts mit EAN zu tun haben und nur beim einscannen helfen.

5 zu 0110001_2

$C_2, \dots, C_7$ werden nach Code A oder Code B codiert. C_1 bestimmt welcher dieser beiden Codes verwendet wird.
$C_8, \dots, C_{13}$ werden nach Code C codiert.
C_1 ergibt sich aus der Art der Codierung von $C_2, \dots, C_7$

	Ziffern $C_2 - C_7$		**Ziffern $C_8 - C_{13}$**	**bestimmt durch C_1**
Zeichen	**Code A**	**Code B**	**Code C**	**Code D**
0	0001101	0100111	1110010	AAAAAA
1	0011001	0110011	1100110	AABABB
2	0010011	0011011	1101100	AABBAB
3	0111101	0100001	1000010	AABBBA
4	0100011	0011101	1011100	ABAABB
5	0110001	0111001	1001110	ABBAAB
6	0101111	0000101	1010000	ABBBAA
7	0111011	0010001	1000100	ABABAB
8	0110111	0001001	1001000	ABABBA
9	0001011	0010111	1110100	ABBABA

Codewörter von Code A,B oder C kommen nur einmal vor. Daher treten nie mehr als 4 gleiche Balken nebeneinander auf.

1.3 Blockcodes

$$
\begin{aligned}
00 &\rightarrow 00000\\
01 &\rightarrow 01101\\
10 &\rightarrow 10110\\
11 &\rightarrow 11011
\end{aligned}
$$

1.3.1 Definition

S endl. Menge (=Alphabet), $n \in \mathbb{N}$.
Ein Blockcode C der (Block-)Länge n über S ist Teilmenge von $S^n = \underset{\longleftarrow}{S} \times \underset{n}{\ldots} \times \underset{\longrightarrow}{S}$

Elemente von C heißen **Codewörter**.
Ist $|S| = 2$ (i.d.R. $S = \{0, 1\}$, so **binär** Code.
$|C| = m$, so ist $m \leq |S|^n$.

Dann lassen sich n Informationssymbole (oder Strings von Informationssymbolen) codieren (Codierungsfunktion). Folge von Informationssymbolen (oder Strings) werden dann in Folge von Codewörtern codiert.

1.3.2 Definition: Hamming-Abstand

S endl. Alphabet, $n \in \mathbb{N}$.
$a, b \in S^n$ $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$
$d(a, b) = \sharp\{i : a_i \neq b_i\}$
Hamming-Abstand von a und b (Anzahl der unterschiedlichen Stellen).
(Richard W. Hamming, 1915-1998, Begründer der Codierungstheorie)

Eigenschaften

a) $d(a, b) = 0 \Leftrightarrow a = b$

b) $d(a, b) = d(b, a)$

c) $d(a, b) \leq d(a, c) + d(c, b)$ (Dreiecksungleichung)
$(a_i \neq b_i \Rightarrow a_i \neq c_i$ oder $b_i \neq c_i)$

d) Wenn $(S, +)$ komm. Gruppe, dann auch S^n
$[(a_1, \ldots a_n) + (b_1, \ldots b_n) = (a_1 + b_1, \ldots, a_n + b_n)]$
$d(a, b) = d(a + c, b + c)$ (Translationsinvarianz)

Also: Wird $x \in C$ gesendet und $y \in S^n$ wird empfangen und $d(x, y) = k$, so sind k Fehler aufgetreten.

1.3.3 Definition

a) Hamming-Decodierung

für Blockcode $C \subseteq S^n$
Wird $y \in S^n$ empfangen, so wird y zu einem Codewort $x' \in C$ decodiert, das unter allen Codewörtern minimalen Hamming-Abstand zu y hat.

$$d(x', y) = min\ d(x, y), x \in C$$

(x' muss nicht eindeutig bestimmt sein)
z.B. $C = \{(0000), (1111)\}$
Empfangen: 0011 x' nicht eindeutig in diesem Fall.

($|S| = 2$: Hamming-Decodierung ist bestmöglich, falls jedes Symbol in einem Codewort mit der gleichen Wahrscheinlichkeit $p < \frac{1}{2}$ verändert wird und wenn jedes Codewort gleich wahrscheinlich ist.)

b) Minimalabstand

C Blockcode in S^n, Minimalabstand von C:

$$d(C) = min\ d(x, x'),\ \ x, x' \in C, x \neq x'$$

(Ist $|C| = 1$, so $d(C) = n$)
$[Bsp : C = \{(00000), (01101), (10110), (11011)\}, d(C) = 3]$

c)

Ein Blockcode C ist **t-Felder-korrigierend**, falls $d(C) \geq 2t + 1$, und er heißt **t-Fehler-erkennend**, falls $d(C) \geq t + 1$.

Begründung für die Bezeichnung in c)
„Kugel" vom Radius t um $x \in C : K_t(x) = \{y \in S^n : d(x, y) \leq t\}$

Ist $d(C) \geq 2t + 1$, so sind Kugelm vom Radius t um Codewörter disjunkt.

Angenommen es existiert $y \in S^n$ mit $y \in K_t(x) \cap K_t(x')$, $x, x' \in C, x \neq x'$. Dann $d(x, x') \leq d(x, y) + d(y, x') \leq t + t = 2t$. Widerspruch
$x \in C$ gesendet, y wird empfangen, und angenommen maximal t-Fehler sind aufgetreten, dann $y \in K_t(x)$ und Abstand zu jedem anderem Codewort ist $> t$
$\Rightarrow$ Hamming-Decodierung ist korrekt.
$d(C) \geq t + 1$ und es treten maximal t minimal 1 Fehler auf, so ist y kein Codewort.

Bsp:

a) n-fach Wiederholungscode

$$\begin{aligned} S_n &\to \underbrace{S_1 S_1 \ldots S_1}_{n} \\ &\vdots \\ S_k &\to \underbrace{S_k S_k \ldots S_k}_{n} \end{aligned}$$

$C = \{(s, s, \ldots, s) : s \in S\} \subseteq S^n$
$d(C) = n$

$\left\lfloor \frac{n-1}{2} \right\rfloor$-Fehler-korr.

b) ISBN, EAN-Codes, $d(C) = 2$, 1-Fehler-erkennend.

1.4 Titel???

$d(C) \geq 2 \cdot t + 1,\ C \subseteq R^N$
$K_t(x) \cap K_t(x') = \emptyset$
$x, x' \in C,\ x \neq x'$

y empfangen:

- falls y in $K_t(x)$ liegt für einen $x \in C$, so wird y nach x decodiert (Korrekt, falls max. t Fehler aufgetreten sind)
- falls y in keiner $K_t(x)$ liegt, so kann es mehrere Codewörter geben mit gleichem min. Abstand zu y. (Dann keine eindeutige Decodierung)

1.4.1 Definition: Perfekter Code

Code $C \subseteq R^n$ heißt perfekt, falls es ein $t \in \mathbb{N}_0$ gibt, mit der Eigenschaft:

$$R^n = \bigcup_{x \in C} K_t(x) \quad \text{und} \quad K_t(x) \cap K_t(x') = \emptyset \quad \text{für} \quad x, x' \in C,\ x \neq x'$$

Dann ist $d(C) = 2 \cdot t + 1$, falls $|C| > 1$:
Ang. $d(C) \leq 2 \cdot t$. Wähle $x, x' \in C,\ x \neq x'$, mit $d(x, x') = d(C) \leq 2 \cdot t$.
Wähle $y \in R^n$ mit $d(x, y) = t,\ d(y, x') \leq t$
$y \in K_t(x) \cap K_t(x')$ Widerspruch
$d(C) \leq 2 \cdot t + 1$

Wähle $x \in C$, wähle $y \in R^n$ mit $d(x, y) = t + 1$. Nach Vorraussetzung existiert $x' \in C$ mit $y \in K_t(x')$.

$$\begin{aligned} d(x, x') &\leq d(x, y) + d(y, x') \leq t + 1 + t = 2 \cdot t + 1 \\ d(C) &\leq 2 \cdot t + 1 \end{aligned}$$

1.4.2 Gibt es perfekte Codes?

Trivial Beispiele:

- einelementige Codes (t=n)
- $C = R^n$ (t=0)
  (Jedes Element ist ein Codewort)
- n-fache Wiederholungscode über Z_2
  $n = 2 \cdot t + 1$
  $C = \{(\underbrace{0, \dots, 0}_{\longleftarrow n \longrightarrow}), (\underbrace{1, \dots, 1}_{\longleftarrow n \longrightarrow})\}$

1.4.3 Lemma

$|R| = q,\ x \in R^n,\ t \in \mathbb{N}$
Dann ist $|K_t(x)| = \sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i$
$\left(\binom{n}{i} = \frac{n}{i(n-i)}\right)$

Beweis

Abstand 0 zu x: 1 Word (nämlich x): $\binom{n}{0} \cdot (q-1)^0 = 1$
Abstand $i > 0$ zu x:
Anzahl der Auswahl von i Positionen aus n Positionen: $\binom{n}{i}$
An jeder Position $q - 1$ Änderungsmöglichkeiten.
$\rightarrow$ insgesamt $(q-1)^i$ Möglichkeiten,
Anzahl der Wörter vom Abstand i von x: $\binom{n}{i} \cdot (q-1)^i$

Satz

Sei C ein Code der Länge n über R, $|C| > 1, |R| = q$. Sei $t \in \mathbb{N}_0$ maximal mit $d(C) \geq 2 \cdot t + 1,\ t = \lfloor \frac{d(C)-1}{2} \rfloor$.

a) (Kugelpackungsschranke)
   $|C| \leq \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i}$

b) C ist perfekt $\Leftrightarrow$ in a) gilt Gleichheit, d.h.
   $|C| = \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i}$

Beweis

a)

$$d(C) \geq 2 \cdot t + 1, \text{ daher } K_t(x) \cap K_t(x') = \emptyset, \ x \neq x', \ x, x' \in C$$
$$R^n \geq \bigcup_{x \in C} K_t(x)$$
$$q^n = \left|R^n\right|$$
$$\left|\bigcup_{x \in C} K_t(x)\right| = \sum_{x \in C} |K_t(x)| \underset{Lemma}{=} |C| \cdot \sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i$$

b)

$$\Rightarrow: d(C) = 2 \cdot t + 1$$
$$R^n = \bigcup_{x \in C} K_t(x) \Rightarrow \text{Gleichheit in a)}$$
$$\Leftarrow: \text{Gleichheit} \Rightarrow R^n = \bigcup K_t(x) \Rightarrow C \text{ perfekt.}$$

1.4.4 Bsp: Binärer Hamming-Code der Länge 7

$R = \mathbb{Z}_2 = \{0, 1\}$ C perfekt, $d(C) = 3, \ |C| = 16$
1-Fehler-Korrigierend

$$C = \{(C_1, \ldots, C_7) : C_i \in \mathbb{Z}_2, C_1 + C_4 + C_6 + C_7 = 0,$$
$$C_2 + C_4 + C_5 + C_7 = 0,$$
$$C_3 + C_5 + C_6 + C_7 = 0$$
$$\} \subseteq \mathbb{Z}_2^7$$

C ist Unterraum von $\mathbb{Z}_2^7$
$(C_1, \ldots, C_7) \in C, \ (C_1', \ldots, C_7') \in C$
$(C_1 + C_1', \ldots, C_7 + C_7') \ \ (C_1 + C_1') + (C_4 + C_4') + (C_6 + C_6') + (C_7 + C_7') = 0$
$dim(C) = 4, \ C_4, C_5, C_6, C_7$ frei wählbar $\rightsquigarrow C_1, C_2, C_3$ festgelegt
Basis:

$$(\ldots 1000) \to (1101000)$$
$$(\ldots 0100) \to (0110100) \quad |C| = 2^4 = 16$$
$$(\ldots 0010) \to (1010010)$$
$$(\ldots 0001) \to (1110001)$$

$d(C) = 3$:
Ang. $d(C) = d$. Wähle $x, x' \in C$ mit $d(x, x') = d$

Translationsinvarianz der Metrik:

$$\begin{aligned} d = d(x, x') &= d(x + x, x + x') = d(0, x + x') \\ wt(x) &= \text{ Anzahl der Einsen in } x \\ &= d(0, x) \\ d(C) &= \min wt(x), \quad x \in C, \quad x \neq \mathcal{V} \end{aligned}$$

Zeige: Jeder Vektor $\neq v$ in C enthält mind. 3 Einsen.
$= 3$ weist man nach durch überprüfen aller 15 von $\mathcal{V}$ verschiedenen Codewörtern oder durch Analyse der Gleichung.

$$\begin{aligned} &(C_1, \ldots, C_7) \in C \text{ Ang. } C_7 = 1 \\ \Rightarrow &C_1 + C_4 + C_6 = 1. \text{ Wenn alle Eins } \checkmark \\ &C_1 = 1, C_4 = C_6 = 0 \\ &C_4 = 1, C_1 = C_6 = 0 \\ &C_6 = 1, C_1 = C_4 = 0 \\ &C_1, C_2 \text{ oder } C_3 = 1 \end{aligned}$$

2. Fall: $C_7 = 1, C_4 = 1, C_1 = 0, C_2$ oder $C_3 = 1$
3. Fall: analog zu Fall 2.
1. Fall: $C_1 = 1, C_4 = C_6 = 0, C_7 = 1$, o.B.d.A. $C_2 = C_3 = 0 \Rightarrow C_5 = 1$
$d(C) \leq 3,\ d(C) = 3 = 2 \cdot 1 + 1$

Prüfe nach, ob bei Kugelpackungsschranke Gleichheit gilt:

$$\begin{aligned} &|C| = 16 \\ &|C| \leq \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} \cdot (q-1)^i} \quad (q = 2, t = 1, n = 7) \\ &= \frac{2^7}{1 + \binom{7}{1}} = \frac{2^7}{2^3} = 2^4 = 16 \\ &C \text{ perfekt!} \end{aligned}$$

1.5 Lineare Codes

1.5.1 Definition: linearer Code

Sei K ein endlicher Körper, $n \in \mathbb{N}$. Ein linearer Code C der Länge n ist ein Unterraum von K^n. (Zeilenvektoren) [Alphabet = k]

Ist $dim(C) = k$, so heißt C $[n, k]$-Code.
Ist $d(C) = d$, so $[n, k, d]$-Code.

Beachte: $|K| = q \Rightarrow |C| = q^k$.

1.5.2 Definition: Informationsrate

Informationsrate (Rate) von C : $\frac{k}{n}$.

1.5.3 Bemerkung über endliche Körper

a) p Primzahl, $\mathbb{Z}_p$ ist Körper der Ordnung p

b) K endlicher Körper $\Rightarrow |K| = p^m$, p Primzahl, $m \in \mathbb{N}$.

c) Zu jeder Primzahlpotenz p^m existiert (bis auf Isomorphie) genau ein Körper der Ordnung p^m.

d) f sei irreduzibles Polynom vom Grad m über $\mathbb{Z}_p$.
$K = \{g \in \mathbb{Z}_p[x] : Grad(g) \leq m-1\}, \quad |K| = p^m$
K wird Körper:
Addition = übliche Addition von Polynomen
Multiplikation = normale Multiplikation + Reduktion mod f
(AES : $|K| = 2^8$)

1.5.4 Bsp

a) n-facher Wiederholungscode über $\mathbb{Z}_p$
$C = \{(\underbrace{0, \ldots, 0}_{\longleftarrow n \longrightarrow}), (1, \ldots, 1), \ldots, (p-1, \ldots, p-1)\}$
C ist linerer Code, $C =< (1, \ldots, 1) >$
$[n, 1, n]$-Code

b) Hamming-Code ist linearer [7,4,3]-Code über $\mathbb{Z}_2$

c) $C = \{(c_1, \ldots, c_n) : c_i \in \mathbb{Z}_p, \sum_{i=1}^{n} c_i = 0\}$
($p = 2$: Parity Check Code), linear $[n, n-1, 2]$-Code über $\mathbb{Z}_p$
Basis von C : $(1, 0, \ldots, 0, p-1), (0, 1, 0, \ldots, 0, p-1), \ldots, (0, \ldots, 0, 1, p-1)$

1.5.5 Definition: Gewicht und Minimalgewicht

K endl. Körper

a) $x \in K^n$, so Gewicht von x, $wt(x)$, definiert durch $wt(x) = \sharp\{i : x_i \neq 0\}$

b) $\{0\} \neq C \subseteq K^n$, so ist das Minimalgewicht von C definiert durch $wt(C) = \min\limits_{x \in C, x \neq 0} wt(x)$

1.5.6 Satz

Ist $C \neq \{0\}$ ein linearer Code, so ist $d(C) = wt(C)$. (Beweis wie beim [7,4,3]-Hamming Code)

1.5.7 Definition: Erzeugermatrix

Sei C ein $[n,k]$-Code über K, sei $g_1 = (g_{11}, \ldots, g_{1n}), \ldots, (g_{k1}, \ldots, g_{kn}) = (g_{k1}, \ldots, g_{kn})$ eine Basis von C.

Dann heißt die $k \times n$ -Matix $G = \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix} = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \cdots & g_{kn} \end{pmatrix}$ Erzeugermatrix von C

1.5.8 Satz

Sei G ein Erzeugermatrix von C.
Dann ist $C = \{ \underset{1\times k}{u} \cdot \underset{k\times x}{G} : u \in K^k\}$
Beweis:
$u = (u_1, \ldots, u_k), u_i \in K$
$uG = (u_1, \ldots, u_k) \cdot (g_1, \ldots, g_k)^t = u_1 g_1 + \ldots + u_k g_k \in C$

1.5.9 Bemerkung

a) Die Abb $\begin{cases} K^k & \to C \\ u & \mapsto uG \end{cases}$ ist bijektiv.
$u \in K^k$ Informationswörter
Codert in Codewörter durch uG.

b) Elementare Zeilenumformungen an Erzeugermatrix liefern Erzeugermatrix.

1.5.10 Beispiel: Hamming-$[7,4]$-Code über $\mathbb{Z}_7$

$$\begin{aligned} C = \{(C_1, \ldots, C_7) : C_i \in \mathbb{Z}_2, C_1 + C_4 + C_6 + C_7 = 0, \\ C_2 + C_4 + C_5 + C_7 = 0, \\ C_3 + C_5 + C_6 + C_7 = 0 \\ \} \subseteq \mathbb{Z}_2^7 \end{aligned}$$

Erzeugermatrix:

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

Cod. eines Informationswort (u_1, u_2, u_3, u_4) mit G

$(u_1, u_2, u_3, u_4) \to (u_1, u_2, u_3, u_4) \cdot G = (\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, u_1{+}u_3{+}u_4, u_2{+}u_3{+}u_4, u_1{+}u_2{+}u_3)$

1.5.11 Definition: Standardform

C $[n, k]$-Code Erzeugermatix C ist in Standardform, falls sie folgende Gestalt hat.

$$G = \overbrace{\begin{pmatrix} 1 & & 0 & & \\ & \ddots & & & * \\ 0 & & 1 & & \end{pmatrix}}^{\begin{matrix} \leftarrow & k & \rightarrow & \leftarrow (n-k) \rightarrow \end{matrix}} \Big\} k$$

Cod. $(u_1, \ldots, u_k) \cdot G = (u_1, \ldots, u_k, *, \ldots, *)$

1.5.12 Satz

Sei C ein $[n, k]$-Code über K. Dann existiert $(n - k) \times n$-Matrix H über K mit folgenden Eigenschaften:
Sei $y \in K^n$. Dann: $y \in C \Leftrightarrow H \cdot y^t = \vec{0}$
H heißt Kontrollmatrix von C ($\Leftrightarrow y \cdot H^t = \vec{0}$)
Es ist $rg(H) = n - k$ (Dann ist $H \cdot G^t = 0$)

1.5.13 Beweis

Sei $g_1, \ldots, g_k$ Basis von C, $G = \begin{pmatrix} g_1 \\ \vdots \\ g_2 \end{pmatrix}$

$g_i = (g_{i1}, \ldots, g_{in})$
Betrachte LGS:

$$\begin{aligned} g_{11}x_1 + \ldots + g_{1n}x_n &= 0 \\ &\vdots \\ g_{k1}x_1 + \ldots + g_{kn}x_n &= 0 \end{aligned}$$

d.h. $G \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$. Koeffizientmatrix G hat Rang k.
Dimension des Löungsraums dieses LGS $= n - k$
Sei $h_1, \ldots, h_{n-k} \in K^n$ Basis des Lösungsraums dieses LGS.

$$H = \begin{pmatrix} m \\ \vdots \\ h_{n-k} \end{pmatrix}, \quad H \cdot g_i^t = \begin{pmatrix} h_1 g_i^t \\ \vdots \\ h_{n-k} g_i^t \end{pmatrix} = 0, i = 1, \ldots, k$$

$Hy^t = 0$ für alle $y \in C$.
$rg(H) = n - k \Rightarrow dim\ Kern(H) = k = dim(C)$
$C = Kern(H)$

1.5.14 Bermerkung

- Kontrollmatrix kann zur Fehlererkennung verwendet werden.
- Beweis liefert Verfahren: Erzeugermatrix → Kontrollmatrix
- Umgekehrt: Kontrollmatrix → Erzeugermatrix (Bilde Basis des Lösungsraums von $Hy^t = 0$)

1.5.15 Beispiel

a) Parity-Check-Code über $\mathbb{Z}_p$
$C = \{(c_1, \ldots, c_n) : \sum_{i=1}^{n} c_i = 0\}$
$H = (1, 1, \ldots, 1)$
$$H \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = 0 \Leftrightarrow c_1 + \ldots c_n = 0 \Leftrightarrow (c_1, \ldots, c_n) \in C$$

b) $[7, 4]$-Hamming-Code

$$\begin{aligned} C = \{(C_1, \ldots, C_7) : C_i \in \mathbb{Z}_2,\ & C_1 + C_4 + C_6 + C_7 = 0, \\ & C_2 + C_4 + C_5 + C_7 = 0, \\ & C_3 + C_5 + C_6 + C_7 = 0 \\ & \} \subseteq \mathbb{Z}_2^7 \end{aligned}$$

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

c) C Code mit Erzeugermatrix
$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ $[4, 2]$-Code über $\mathbb{Z}_2$
$$G \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_4 \end{pmatrix} = 0$$

$$\begin{aligned} x_1 + x_2 + x_4 &= 0 \\ x_2 + x_4 &= 0 \end{aligned}$$

x_5, x_4 frei wählen, x_1, x_2 fesgelegt.

Basis $(0010), (0101)$

$$\text{Kontrollmatrix } H = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

$C\{(c_1, \ldots, c_4) : c_3 = 0, c_2 + c_4 = 0\}$

1.5.16 Satz

C $[n, k]$-Code, $C \neq \{\vec{0}\}$, K^n, Kontrollmatrix H.

$$\begin{aligned} d(C) = wt(C) = min \quad & \{r : \text{ in H gibt es r linear abhänige Spalten}\} \\ = max \quad & \{r : \text{ je r-1 Spalten linear unabhängig}\} \end{aligned}$$

Beweis

$s_1, \ldots, s_n$ Spalten von H, Länge $n - k$.
$C \neq \{\vec{0}\}, k \geq 1, n - k < n \Rightarrow s_1, \ldots, s_n$ lin. abhängig.
Sei $min\{r : \ldots\} = w$. $s_{i_1}, \ldots, s_{i_w}$ lin. abhängig.
Existiert $c_{i_1}, \ldots, c_{i_w} \in K$, nicht alle $= 0$, $c_{i_1} s_{i_1} + \ldots + c_{i_w} s_{i_w} = 0$
$w \quad min \Rightarrow$ alle $c_{i_1}, \ldots, c_{i_w} \neq 0$.
Def. $c = (c_1, \ldots, c_n)$ mit den c_{i_j} an den Stellen i_j, übrige $c_i = 0$

$$\sum_{i=1}^{n} c_i s_i = c_{i_1} s_{i_1} + \ldots + c_{i_w} s_{i_w} = 0$$

$$\sum_{i=1}^{n} c_i s_i^t = 0$$

$$Hc^t = 0 \quad c \in C$$

$wt(c) = w$, Min. Gewicht von $C \leq wt(c) = w$
Ang. es ex. $0 \neq c' \in C, wt(c') = w' < w$. $Hc'^t = 0$
$c' = (c'_1, \ldots, c'_n) \quad \sum c'_i s_{i=1}^n = 0 \Rightarrow w'$ der Spalten $c_1, \ldots, c_n$ sind linear abhänging.
Widerspruch!
$wt(c) = w$

1.5.17 Beispiel: $[7, 4]$-Hamming-Code über $\mathbb{Z}_2$

$H = \begin{pmatrix} 1 & & & 0 & 1 & 0 & 1 \\ & \ddots & & 0 & 1 & 1 & \\ 0 & & 1 & 1 & 1 & 1 & \end{pmatrix}$ Kontrollmatrix.

Keine Spalte ist Nullspalte, keine zwei Spalten sind gleich. 1.,2.,4. Spalte sind linear abhänging.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad d(C) = 3$$

1.5.18 Korollar: (Singleton-Schranke)

Ist C ein linearer $[n,k]$-Code, $d(C) = d$, so gilt:

$$d \leq n - k + 1$$

Beweis

2. Gleichheit: $d \leq rg(H) + 1 = n - k + 1$ (Zeilen von H sind lin. unabhängig)

1.5.19 Bemerkung: (Nebenklassen von Unterräumen in Vektorräumen)

C ein Umterraum von Vektorraum V. Für jedes $v \in V$:

$$v + C = \{v + x : x \in C\}$$

Nebenklasse von C zu v.

a) $v_1, v_2 \in V$. Dann:
$v_1 + C = v_2 + C$ oder $(v_1 + C) \cap (v_2 + C) = \emptyset$

b) $v_1 + C = v_2 + C \Leftrightarrow v_1 - v_2 \in C$
$(v + C = C(= \vec{0} + C) \Leftrightarrow v \in C)$

c) Wähle aus jeder Nebenklasse einen Vektor v_i:

$$V = \dot{\bigcup}(v_i + C)$$

d) V Vektorraum über endl. Körper: $|v + C| = |C|$

e) C $[n,k]$-Code $(V = K^n, dim(C) = k, |C| = q^k, \text{ falls } |K| = q)$
Anzahl der Nebenklassen ist q^{n-k}

1.6 Syndrom-Decodierung linearer Code

C $[n,k]$-Code über K, $|K| = q$, Kontrollmatrix H, $(n-k) \times n$-Matrix.
Ist $y \in K^n$, so heißt $Hy^t \in K^{n-k}$ **Syndrom** von y.

a) $x \in C \Leftrightarrow Hx^t = 0$ (x hat Syndrom 0)

b) $y_1, y_2 \in K^n$. y_1, y_2 liegen in der gleichen Nebenklasse zu C (d.h. $y_1 + C = y_2 + C$)
$\Leftrightarrow y_1, y_2$ haben gleiches Syndrom
(d.h. $Hy_1^t = Hy_2^t$)

$[y_1 + C = y_2 + C \Leftrightarrow y_1 - y_2 \in C \Leftrightarrow 0 = H(y_1 - y_2)^t = Hy_1^t - Hy_2^t \Leftrightarrow Hy_1^t = Hy_2^t]$

c) Jedes $z \in K^{n-k}$ tritt als Syndrom auf.

Ang. $x \in C$ wird gesendet, $y = x + f$, wird empfangen.
f "Fehlervektor".
$y + C = f + C$, y und f haben das gleiche Syndrom, nämlich Hy^t.
Bestimmt in der Nebenklasse von y ein e mit kleinstmögliche Gewicht (**Nebenklassenführer**)
Decodierung: $y \to y - e \in C$ (Hamming-Decodierung)

Ordne die Nebenklassenführer nach der lexikogr. ihrer Syndome.
Speicherbedarf: q^{n-k} Nebenklassenführer, jeder hat Länge n
(Besser als Durchforsten der Liste aller Codewörter (q^k), falls $k \geq \frac{n}{2}$)
C $[70, 50]$-Code über $\mathbb{Z}_2$. 2^{20} Nebenklassenführer, je 70 BitLänge.
Speicher: $70 \cdot 2^{20} Bit \approx 8,75$ MegaByte
Speicher für Codewörter: $70 \cdot 2^{50}$ Bit = 9 PetaByte

1.6.1 Beispiel

C $[5, 2]$-Code über $\mathbb{Z}_2$, Kontrollmatrix

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

$d(C) = 3$

$$\begin{aligned} (x_1, \ldots, x_5) \in C \Leftrightarrow & x_1 + x_5 = 0 \\ & x_2 + x_3 = 0 \\ & x_2 + x_4 + x_5 = 0 \end{aligned}$$

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

Nebenklassen von C.

C =(000000) + C = {(00000), (11101), (01110), (10011)}

(10000) + C = {(10000), (01101), (11110), (00011)}

Nebenklassenführer:(10000)

(01000) + C = {(01000), (10101), (00110), (11011)}

Nebenklassenführer:(01000)

(00100) + C = {(00100), (11001), (01010), (10111)}

Nebenklassenführer:(00100)

(00010) + C = {(00010), (11111), (01100), (10001)}

Nebenklassenführer:(00010)

(00001) + C = {(00001), (11100), (01111), (10010)}

Nebenklassenführer:(00001)

(00111) + C = {(00111), (11010), (01001), (10100)}

Mögliche Nebenklassenführer:(01001), (10100)

(00101) + C = {(00101), (11000), (01011), (10110)}

Mögliche Nebenklassenführer:(00101), (11000)

Angenommen als Nebenklassenführer werden gewählt:
$f_0 = (00000), f_1 = (10000), \ldots, f_5 = (00001), f_6 = (01001), f_7 = (00101)$

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

Syndrome:

$$Hf_0^t = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, Hf_1^t = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, Hf_2^t = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, Hf_3^t = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, Hf_4^t = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$Hf_5^t = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, Hf_6^t = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, Hf_7^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(Ordnung: $f_0, f_4, f_3, f_2, f_1, f_5, f_6, f_7$)

Empfangen: $y = (10110)$

$$Hy^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Decodierung: $y \to y + f_7 = (10011) \in C$

(Hätte man für die Nebenklasse $f_7 + C$ als Nebenklassenführer (11000) gewählt, so wäre decodiert worden in $y + (11000) = (01110) \in C$)

1.7 Beispiel guter linear Codes

1.7.1 Hamming-Codes

Sei q ein Primzahlpotenz, K Körper mit $|K| = q$
Sei $l \in \mathbb{N}$. $n = \frac{q^l - 1}{q-1}, k = n - l$
Denn ex. perfekter $[n, k]$-Code C über K, $d(C) = 3$. Hamming-Code.

Konstruktion

$\left|K^l \backslash \{\vec{0}\}\right| = q^l - 1$, je $q - 1$ von 0 versch. Vektoren erzeugen den gleichen 1-dim. Unterräume in K^l, d.h.

$$n = \frac{q^l - 1}{q^l - 1} \quad \text{1-dim Unterraum}$$

Bilde $l \times n$-Matrix H: Wähle aus jedem der 1-dim. Unterraum von K^l einen Vektor $\neq 0$ aus und schreibe ihn als Spalte in H
$C = \{x \in K^n : Hx^t = 0\}$ $rg(H) = l$, denn H enthält l lin. unabhängige Spalten.
$dim(C) = n - l = k$, $|C| = q^k$
$d(C) = 3$
Nach Konstruktion von H sind je zwei Spalten linear unabhänging. Es gibt drei linear abhängige Spalten:

$$\begin{pmatrix} a \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ b \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} c \\ c \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad a, b, c \neq 0$$

$$\frac{c}{a}\begin{pmatrix} a \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \frac{c}{b}\begin{pmatrix} 0 \\ b \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} c \\ c \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Kugelpackungsbed.:

$\sum_{j=0}^{1} \binom{n}{j}(q-1)^i = 1 + n \cdot (q-1) = 1 + \frac{q^l - 1}{q-1} = q^l$
$\frac{q^n}{q^l} = q^{n-l} = q^k = |C|$
C perfekt.