Anton Malevich

Einführung in die Kodierungstheorie

Skript zu einer
im Februar 2013
gehaltenen Kurzvorlesung

Fakultät für Mechanik und Mathematik
Belorussische Staatliche Universität

Institut für Algebra und Geometrie
Otto-von-Guericke-Universität Magdeburg

1 Grundlagen der Kodierungstheorie

Ein *Sender* möchte einem *Empfänger* über einen *Kanal* gewisse *Daten* übermitteln. Dabei werden die Nachrichten durch Interferenzen zufällig gestört, d.h. es passieren Fehler. Wie kann der Qualitätsverlust effizient (z.B. kostengünstig) minimiert werden?

Beispiel: TV, radio, Musik-CD, Datenübertragung von Raumsonden (space probes), ISBN (Bücher-Code), etc.

Erinnerung: Zu jeder Primzahl p gibt es einen Körper $\mathbb{F}_p$ mit genau p Elementen (Operationen mod p; $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$). Wir beschränken uns der Einfachheit halber in folgenden auf den Fall $p = 2$, der zu sogenannten *binären Codes* führt.

Der Körper $\mathbb{F}_2 = \{0, 1\}$ ist durch die Verknüpfungstafeln gegeben

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Wir fixieren ein $n \in \mathbb{N}$ und schreiben $V = \mathbb{F}_2^n$ für den n-dimensionalen Standardvektorraum über $\mathbb{F}_2$. (Die Elemente sind Zeilenvektoren.)

Nachricht: stets ein binäres n-Tupel, also ein Element von V.

Fehlervektor: Wird eine Nachricht $x \in V$ über den Kanal versandt, so erhält der Empfänger die gestörte Nachricht $y = x + e$, wobei $e \in V$ ein "zufälliger" Fehlervektor ist.

Solch ein y soll dann möglichst als x dekodiert werden. Dafür benutzen wir als Nachrichten nur Elemente c eines sorgfältig gewählten Unterraumes C von V. Solch ein C heißt dann *linearer Code*.

1.1 Definition (Hamming-Abstand)**.** Der *(Hamming-)Abstand* von $v, w \in V$ mit $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$ ist die Anzahl der Stellen, an denen sich v und w unterscheiden:

$$\mathrm{d}_H(v, w) = \#\{\, i \in \{1, 2, \ldots, n\} \mid v_i \neq w_i \}\,.$$

1.2 Definition. Ist allgemein X eine beliebige Menge und $d : X \times X \to \mathbb{R}_{\geq 0}$, $(x, y) \mapsto d(x, y)$ eine Abbildung mit den Eigenschaften

(M1) $d(x, y) = 0$ genau dann, wenn $x = y$

(M2) $d(x, y) = d(y, x)$ (Symmetrie)

(M3) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

für alle $x, y, z \in X$, so heißt d eine *Abstandsfunktion* oder *Metrik* auf X und (X, d) ein metrischer Raum.

1.3 Beispiel. 1) $X = \mathbb{R}$ mit dem gewöhnlichen Abstand $d(x, y) = |x - y|$.

2) $X = \mathbb{R}^n$ mit dem euklidischen Abstand

$$d\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}.$$

3) Französische Eisenbahnmetrik: $X = \mathbb{R}^2$, $O = (0, 0)$. Der Abstand wird berechnet je nachdem, ob eine direkte Strecke zwischen x und y existiert.

$$d(x, y) = \begin{cases} \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, & \text{falls } x \text{ — } y \\ \sqrt{x_1^2 + x_2^2} + \sqrt{y_1^2 + y_2^2}, & \text{falls } x,\ O,\ y \end{cases}$$

1.4 Satz. *Der Hamming-Abstand* $\mathrm{d} = \mathrm{d}_H$ *ist eine Metrik auf* V.

Beweis. Wir überprüfen die Eigenschaften aus der Definition 1.2. Es seien $u, v, w \in V$ mit $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$.

(M1) $\mathrm{d}(v, w) = 0$ genau dann, wenn $v_i = w_i$ für alle i gilt, d.h. $v = w$.

(M2) $\mathrm{d}(v, w) = \#\{i \mid v_i \neq w_i\} = \#\{i \mid w_i \neq v_i\} = \mathrm{d}(w, v)$.

(M3) Für jedes i gilt $u_i \neq w_i \;\Rightarrow\; v_i \neq u_i$ oder $v_i \neq w_i$ Dies ergibt

$$\begin{aligned} \mathrm{d}(u, w) &= \#\{i \mid u_i \neq w_i\} = \#\{i \mid u_i \neq v_i\} + \#\{i \mid v_i \neq w_i\} \\ &= \mathrm{d}(u, v) + \mathrm{d}(v, w). \end{aligned}$$

□

1.5 Definition (abgeschlossene Hamming-Kugel)**.** Für $v \in V$ und $r \in \mathbb{R}_{\geq 0}$ heißt

$$S_r = \{w \in V \mid \mathrm{d}(v, w) \leq r\}$$

die abgeschlossene Hamming-Kugel von Radius r um v.

1.6 Definition (Code, Minimaldistanz)**.** Ein *Code* der Länge n über $\mathbb{F}_2$ ist eine nicht leere Teilmenge $C \subseteq V = \mathbb{F}_2^n$. Elemente eines Codes werden oft *Codewörter* genannt.

Die Minimaldistanz eines Codes C ist die Invariante

$$\mathrm{d}(C) = \begin{cases} 0, & \text{falls } |C| = 1, \\ \min\{\mathrm{d}(v, w) \mid v, w \in C\}, & \text{sonst.} \end{cases}$$

Ist $\mathrm{d}(C) \geq 2t + 1$ für ein $t \in \mathbb{N}$, so heißt C ein t-fehlerkorrigirender Code.

1.7 Lemma. *Sei $C \subseteq V$ ein t-fehlerkorrigierender Code. Dann gelten:*

(a) Für jedes $v \in V$ gibt es höchstens ein $c \in C$ mit $\mathrm{d}(v,c) \leq t$.

(b) Die Hammingkugeln $S_t(c)$ von Radius t um Codewörter $c \in C$ sind paarweise disjunkt.

Beweis. (a) Es sei $v \in V$ und seien $c, \widetilde{c} \in C$ mit $\mathrm{d}(v,c),\ \mathrm{d}(v,\widetilde{c}) \leq t$. Es folgt dann $\mathrm{d}(c,\widetilde{c}) \leq \mathrm{d}(c,v) + \mathrm{d}(v,\widetilde{c} \leq 2t$. Aus $\mathrm{d}(C) \geq 2t+1$ ergibt sich nun $c = \widetilde{c}$.

(b) Es seien $c, \widetilde{c} \in C$ mit $S_t(c) \cap S_t(\widetilde{c}) \neq \varnothing$. Für $v \in S_t(c) \cap S_t(\widetilde{c})$ gilt dann $\mathrm{d}(v,c),\ \mathrm{d}(v,\widetilde{c}) \leq t$. Aus (a) folgt nun $c = \widetilde{c}$.

□

Interpretation: Angenommen, bei der Übertragung eines Codeworts c eines t-fehlerkorrigierenden Codes $C \subseteq V$ treten höchstens Fehler in bis zu t Koordinatenstellen auf. Die veränderte Nachricht $y \in V$ kann dann wie folgt korrekt dekodiert werden: wähle das eindeutige Codewort $\widetilde{c} \in C$ mit $\mathrm{d}(y,\widetilde{c}) \leq t$.

Was will man haben?

- $C \subseteq V$ soll möglichst groß sein
  (hohe Informationsrate, $|C| \to \max$).
- $\mathrm{d}(C)$ soll möglichst groß sein
  (hohe Fehlerkorrekturmöglichkeit, $\mathrm{d}(C) \to \max$).
- der Dekodieralgorithmus soll möglichst effizient sein.

1.8 Beispiel.

(i) Wiederholungscode: $C = \{(0,0,\ldots,0),(1,1,\ldots,1)\} \subseteq V$, $\mathrm{d}(C) = n \rightsquigarrow \left\lfloor \frac{n-1}{2} \right\rfloor$-fehlerkorrigierend, aber sehr geringe Informationsrate.

(ii) $C = \{(0,0,0,0,0),(0,1,1,0,1),(1,0,1,1,0),(1,1,0,1,1)\} \subseteq \mathbb{F}_2^5$, $\mathrm{d}(C) = 3 \rightsquigarrow$ 1-fehlerkorrigierend, aber sehr geringe Informationsrate.

Bemerkung: Im $\mathbb{F}_2^4$ gibt es keinen Code C mit $|C| = 4$ und $\mathrm{d}(C) \geq 3$.

1.9 Definition (linearer Code, Gewicht, Minimalgewicht)**.** Ein *linearer Code* der Länge n über $\mathbb{F}_2$ ist ein Unterraum C von $V = \mathbb{F}_2^n$ (wir schreiben $C \leq V$).

Das *Gewicht* eines Vektors $v = (v_1,\ldots,v_n) \in V$ ist die Anzahl der Koordinatenstellen ungleich Null:

$$\mathrm{wt}(v) = \mathrm{d}(v,0) = \#\{\, i \in \{1,2,\ldots,n\} \mid v_i \neq 0\}\,.$$

Das *Minimalgewicht* eines Codes $C \subseteq V$ ist das kleinste Gewicht unter allen Codewörter ungleich Null:

$$\mathrm{wt}(C) = \begin{cases} 0, & \text{falls } C = \{0\}\,, \\ \min\{\,\mathrm{wt}(c) \mid c \in C \text{ mit } c \neq 0\}\,, & \text{sonst.} \end{cases}$$

1.10 Lemma. *Es sein $C \leq V$ ein linearer Code. Dann gilt $\mathrm{d}(C) = \mathrm{wt}(C)$, d.h. "Minimaldistanz = Minimalgewicht".*

Beweis. Da C ein Unterraum ist, haben wir $0 \in C$ und somit $\mathrm{d}(C) \leq \mathrm{wt}(C)$.

Ist $C = \{0\}$, so sind $\mathrm{d}(C) = 0 = \mathrm{wt}(C)$.

Sei nun $C \neq \{0\}$. Wähle $v \neq w \in C$ mit $\mathrm{d}(C) = \mathrm{d}(v, w)$. Wir behaupten, dass $\mathrm{wt}(v - w) = \mathrm{d}(v, w)$ ist $(*)$. Daraus folgt wegen $0 \neq v - w \in C$ dann

$$\mathrm{d}(C) = \mathrm{d}(v, w) = \mathrm{wt}(v - w) \leq \mathrm{wt}(C),$$

und somit $\mathrm{d}(C) = \mathrm{wt}(C)$.

Zu $(*)$: sind $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$, so gilt

$$\mathrm{d}(v, w) = \#\{i \mid v_i \neq w_i\} = \#\{i \mid v_i - w_i \neq 0\} = \mathrm{d}(v - w, 0) = \mathrm{wt}(v - w).$$

□

1.11 Definition (Erzeugendenmatrix)**.** Es sei $C \leq V$ ein linearer Code und $k = \dim C$. Eine *Erzeugendenmatrix* G für C ist eine $k \times n$ Matrix über $\mathbb{F}_2$, deren Zeilen eine Basis für C liefern.

1.12 Bemerkung. Die Hamming-Distanz und die abgeliteten Invarianten (Minimaldistanz, Gewicht, Minimalgewicht) werden von Koordinatenvertauschungen (d.h. Umsortierungen der Standardbasis $(e_1, \ldots, e_n)$ von V) nicht berührt.

Es sei $C \leq V$ ein linearer Code und $\dim C = k$. Erlaubt man sich gegebenenfalls Koordinatenvertauschungen vorzunehmen, so besitzt C eine Erzeugendenmatrix der Form $G = (E_k|A)$, wobei E_k die $k \times k$ Einheitsmatrix, und A eine $k \times (n - k)$-Matrix ist.

Hat C eine Erzeugendenmatrix G von der obigen Form, so kann man die ersten k Koordinaten eines Codeworts als *Informationsanteil* und die übrigen $(n - k)$ Koordinaten als *Prüfstellen* betrachten: Klar ist jedes $c \in C$ eindeutig durch seine ersten k Koordinaten bestimmt.

2 Standardskalarprodukt und Orthogonalräume

Wir betrachten in diesem Abschnitt dem n-dimensionalen Standardvektorraum $V = K^n$ über einem Körper K.

2.1 Definition (Standardskalarprodukt)**.** Die Verknüpfung $\langle ., . \rangle : V \times V \to K$ mit

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{i=1}^{n} x_i y_i$$

heißt *Standardskalarprodukt* auf V.

2.2 Lemma. *Für alle $v, v', w \in V$ und $a_1, a_2 \in K$ gilt*

(a) $\langle v, w\rangle = \langle w, v\rangle$ *(symmetrisch),*

(b) $\langle a_1 v + a_2 v', w\rangle = a_1 \langle v, w\rangle + a_2 \langle v', w\rangle$ *(linear).*

2.3 Definition (Orthogonalraum)**.** Es sei $U \leq V$. Dann heißt

$$U^\perp = \{ w \in V \mid \langle u, w\rangle = 0 \text{ für alle } u \in U\}$$

der zu U *orthogonale Raum* in V.

2.4 Bemerkung. Es sei U ein Unterraum von V.

(a) $U^\perp$ ist auch ein Unterraum von V.

(b) Ist $U = \langle u_1, \ldots, u_k\rangle$ (Erzeugnis, lineare Hülle), so gilt

$$U^\perp = \{ w \in V \mid \langle u_i, w\rangle = 0 \text{ für alle } i \in \{1, \ldots, k\}\}.$$

(c) (Dimensionsformel für Orthogonalräume) Es gilt

$$\dim U + \dim U^\perp = \dim V.$$

Achtung! In allgemein gilt nicht $U \cap U^\perp = \{0\}$ und somit auch nicht $U \oplus U^\perp = V$. Für $V = \mathbb{R}^n$ ist das jedoch richtig, da $\langle v, v\rangle \neq 0$ für $v \neq 0$ ist.

2.5 Folgerung. *Für jeden Unterraum U von V gilt $\left(U^\perp\right)^\perp = U$.*

Beweis. Es sei $u \in U$. Dann ist $\langle u, w\rangle = 0$ für alle $w \in U^\perp$, also ist $\left(U^\perp\right)^\perp$, d.h. $U \subseteq \left(U^\perp\right)^\perp$. Mit Bemerkung 2.4 gilt nun

$$\dim U = \dim V - \dim U^\perp = \dim\left(U^\perp\right)^\perp.$$

□

3 Dualer Code und Kontrollmatrix

Es sei nun wieder $V = \mathbb{F}_2^n$ und betrachte lineare Codes $C \leq V$.

3.1 Definition (dualer Code)**.** Der Orthogonalraum $C^\perp \leq V$ heißt der zu C *duale Code*.

3.2 Satz. *C und $C^\perp$ sind dual in dem Sinne, dass es $\left(C^\perp\right)^\perp = C$ gilt*

Beweis. Siehe Folgerung 2.5. □

3.3 Definition (Kontrollmatrix)**.** Es sei $C \leq V$ ein linearer Code und $k = \dim C$. Eine *Kontrollmatrix H* für C ist eine $(n-k) \times n$ Matrix über $\mathbb{F}_2$, deren Zeilen eine Basis für $C^\perp$ liefern (d.h. eine Erzeugendenmatrix für $C^\perp$).

3.4 Lemma. *Es sei $C \leq V$ ein linearer Code mit $k = \dim C$ und Erzeugendenmatrix G der Form $G = (E_k|A)$, wobei E_k die $k \times k$ Einheitsmatrix, und A eine $k \times (n-k)$ Matrix ist. Dann kann man für die Kontrollmatrix von C die Matrix $H = (-A^\mathsf{T}|E_{n-k})$ wählen.*

Beweis. Es sei $(e_1, \ldots, e_n)$ die Standardbasis von V. Seien $(u_1, \ldots, u_k)$ die Zeilen von G, also Basis von C, wobei

$$u_i = e_i + \sum_{j=k+1}^{n} a_{ij} e_j \text{ für } 1 \leq i \leq k,\ a_{ij} \in K.$$

Zu zeigen ist, dass mit Basis $(w_1, \ldots, w_{n-k})$ eine Basis von $C^\perp$ ist, wobei

$$w_r = -\sum_{j=k+1}^{n} a_{s,k+r} e_s + e_{k+r} \text{ für } 1 \leq r \leq n-k.$$

Für $1 \leq i \leq k$, $1 \leq r \leq n-k$ gilt

$$\langle u_i, w_r \rangle = \left\langle e_i + \sum_{j=k+1}^{n} a_{ij} e_j, -\sum_{j=k+1}^{n} a_{s,k+r} e_s + e_{k+r} \right\rangle = -a_{i,k+r} + a_{i,k+r} = 0,$$

also sind $w_1, \ldots, w_{n-k} \in U^\perp$ nach Bemerkung 2.4 (b). Die Vektoren $w_1, \ldots, w_{n-k}$ sind wegen ihrer Form auch linear unabhängig. Aus der Dimensionsformel aus Bemerkung 2.4 (c) folgt nun, dass $(w_1, \ldots, w_{n-k})$ eine Basis von $C^\perp$ ist. □

3.5 Definition (Syndrom)**.** Das *Syndrom* eines Vektors $v \in V$ bezüglich eines Vektorsystems $(w_1, \ldots, w_m)$ in V ist der Vektor

$$\left(\langle v, w_1 \rangle, \langle v, w_2 \rangle, \ldots, \langle v, w_m \rangle\right) \in \mathbb{F}_2^m.$$

Ist $C \leq V$ ein linearer Code der Dimension k und H eine Kontrollmatrix von C, so bezeichnet

$$s_H(v) = \left(\langle v, w_1 \rangle, \ldots, \langle v, w_{n-k} \rangle\right) = v \cdot H^\mathsf{T} \in \mathbb{F}_2^{n-k}$$

das *Syndrom* von v bezüglich der Basisvektoren $(w_1, \ldots, w_{n-k})$ von $C^\perp$, die den Zeilen von H entsprechen.

3.6 Lemma. *Es sei $C \leq V$ ein linearer Code mit einer Kontrollmatrix H. Für alle $v \in V$ gilt*

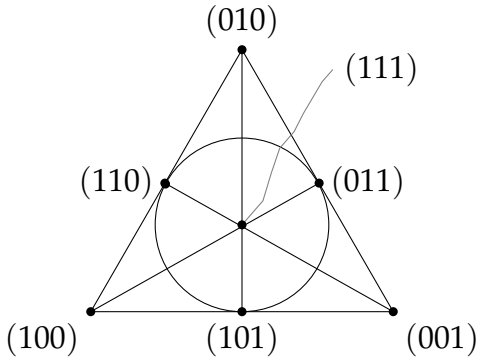
$$s_H(v) = 0 \ \Leftrightarrow\ v \in C.$$

Beweis. Es seien $w_1, \ldots, w_{n-k}$ die Basisvektoren von $C^\perp$, die den Zeilen von H entsprechen. Es gilt

$$\begin{aligned} s_H(v) = v \cdot H^\mathsf{T} &= \left(\langle v, w_1 \rangle, \ldots, \langle v, w_{n-k} \rangle\right) = 0 \\ &\Leftrightarrow\ \langle v, w_i \rangle = 0 \text{ für alle } i \in \{1, \ldots, n-k\} \\ &\overset{2.4\,(a)}{\Leftrightarrow}\ v \in \left(C^\perp\right)^\perp = C. \end{aligned}$$

□

3.7 Beispiel (Hamming-Code der Länge 7)**.** Die projektive Ebene über $\mathbb{F}_2$ ist:



Die 7 Punkte entsprechen den Vektoren in $\mathbb{F}_2^3 \setminus \{(0,0,0)\}$ (und damit den 1-dimensionalen Unterräumen des $\mathbb{F}_2^3$). Die Geraden sind die Teilmengen $\{v_1, v_2, v_3\}$ von $\mathbb{F}_2^3 \setminus \{(0,0,0)\}$ mit $v_1 + v_2 + v_3 = 0$ (und entsprechen den 2-dimensionalen Unterräumen des $\mathbb{F}_2^3$.

Ein *Hamming-Code* der Länge 7 hat die *Kontrollmatrix* H, deren Spalten den Punkten der projektiven Ebene über $\mathbb{F}_2$ entsprechen, zum Beispiel

$$H = \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

(Eine andere Reihenfolge der Spalten führt zu Codes mit ähnlichen Eigenschaften, z.B. gleicher Dimension und gleicher Minimaldistanz.)

Gemäß Lemma 3.4 erhält man die *Erzeugendenmatrix*

$$G = \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right).$$

- Offenbar haben C und $C^\perp$ die Länge 7.
- $\dim C = 4$ und $\dim C^\perp = 3$.
- Die *Minimaldistanzen* sind $\mathrm{d}(C) = \mathrm{wt}(C) = 3$ und $\mathrm{d}(C^\perp) = \mathrm{wt}(C^\perp) = 4$. (Wegen Lemma 1.10 ist $\mathrm{d}(C) = \mathrm{wt}(C)$.)

  Beweis. Übungsaufgabe! □

Da $\mathrm{d}(C) = 3$ ist, ist C 1-fehlerkorrigierend. Obwohl $\mathrm{d}(C^\perp) = 4 > \mathrm{d}(C)$ ist, ist $C^\perp$ auch "nur" 1-fehlerkorrigierend. Wegen $\dim(C) = 4 > \dim(C^\perp) = 3$ ist C besser für die Praxis geeignet (höhere Informationsrate).

Für die Dekodierung sind die Syndrome von Vektoren $v \in \mathbb{F}_2^7$ bezüglich H hilfreich. Man erhält die folgende Tabelle:

$s_H(v)$	v
$(0,0,0)$	$v \in C$
$(1,1,1)$	$v = (1,0,0,0,0,0,0) + c$ mit $c \in C$
$(1,1,0)$	$v = (0,1,0,0,0,0,0) + c$ mit $c \in C$
$(1,0,1)$	
$(0,1,1)$	...
$(1,0,0)$	
$(0,1,0)$	
$(0,0,1)$	$v = (0,0,0,0,0,0,1) + c$ mit $c \in C$

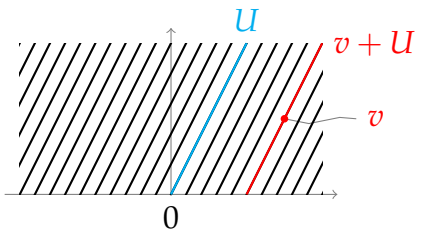
Die möglichen Syndrome sind die 2^3 Vektoren von $\mathbb{F}_2^3$. Zu jedem Syndrom haben wir genau 2^4 Vektoren aus V, die diesen Syndrom haben. Diese bilden die sogenannten *Nebenklassen* von C in V, oder *affine Unterräume* parallel zu C in V.

4 Nebenklassen. Syndrom-Dekodierung

4.1 Definition. Es sei V ein Vektorraum über einem Körper K und U ein Unterraum von V. Für $v \in V$ heißt $v + U = \{v + u \mid u \in U\}$ die *Nebenklasse* von U in V bezüglich v. (Die Elemente einer Nebenklasse werden oft *Vertreter* genannt.)

Die *Menge aller Nebenklassen* von U in V wird mit $V/U = \{v + U \mid v \in V\}$ bezeichnet.

4.2 Beispiel. Es sei $V = \mathbb{R}^2$ und $U = \langle(1,2)\rangle = \{(x,y) \in \mathbb{R}^2 \mid y = 2x\}$. Die Menge V/U aller Nebenklassen von U in V besteht alles allen Geraden, die parallel zu U sind ($y = 2x + b$, $b \in \mathbb{R}$).



4.3 Satz. *Es sei V ein Vektorraum über einem Körper K, $U \leq V$. Dann wird durch*

$$v \sim_U w \;\Leftrightarrow\; v - w \in U \quad (v, w \in V)$$

eine Äquivalenzrelation auf V definiert.

Die $\sim_U$-Äquivalenzklassen sind genau die Nebenklassen von U in V und somit gilt $V/\sim_U \;=\; V/U$. ($V/\sim_U$ bezeichnet die Menge der Äquivalenzklassen bezüglich der Relation $\sim_U$.)

4.4 Folgerung. *V/U liefert eine Partition von V, d.h. eine Zerlegung in paarweise disjunkte Teilmengen: Jedes $v \in V$ liegt in genau einer Nebenklasse von U in V, nähmlich $v + U$.*

4.5 Folgerung. *Sind $v + U,\ w + U \in V/U$, so gilt:*

$$v + U = w + U \;\Leftrightarrow\; v + U \cap w + U \neq \varnothing \;\Leftrightarrow\; v - w \in U.$$

Beweis von Satz 4.3. Wir zeigen zunächst, dass $\sim_U$ eine Äquivalenzrelation auf V ist. Es seien $v,\ v_1,\ v_2,\ v_3 \in V$. Dann gelten

(i) $v - v = 0 \in U$, also $v \sim_U v$ [reflexiv].

(ii) $v_1 \sim_U v_2 \;\Leftrightarrow\; v_1 - v_2 \in U \;\Leftrightarrow\; -(v_1 - v_2) = v_2 - v_1 \in U \;\Leftrightarrow\; v_2 \sim_U v_1$ [symmetrisch].

(iii) $v_1 \sim_U v_2$ und $v_2 \sim_U v_3 \;\Rightarrow\; v_1 - v_2,\ v_2 - v_3 \in U$
$\Rightarrow\; v_1 - v_3 = (v_1 - v_2) + (v_2 - v_3) \in U \;\Rightarrow\; v_1 \sim_U v_3$ [transitiv].

Für die Äquivalenzklasse von $v \in V$ gilt:

$$\begin{aligned}[v]_{\sim_U} &= \{\, w \in V \mid w \sim_U v \,\} = \{\, w \in V \mid \exists u \in U \colon w - v = u \,\} \\ &= \{\, w \in V \mid \exists u \in U \colon w = v + u \,\} = v + U.\end{aligned}$$

□

4.6 Bemerkung. V/U hat eine Vektorraumstruktur und wird der *Faktorraum* von V nach U genannt. Es gilt

$$\dim V/U = \dim V - \dim U.$$

Es sei im folgenden $V = \mathbb{F}_2^n$ und $C \leq V$ ein linearer Code mit $\dim C = k$.

Hauptbeispiel 3.7: Hamming-Code der Länge 7. In diesem Beispiel ist $\dim C = 4$ und $\dim C^\perp = 3$, und C ist 1-fehlerkorrigierend.

4.7 Lemma. *Es sei H eine Kontrollmatrix für C. Dann gilt für die Syndrome von $v,\ w \in V$ bezüglich H*

$$s_H(v) = s_H(w) \;\Leftrightarrow\; v + C = w + C$$

Beweis. Es seien $v,\ w \in V$. Es gilt

$$\begin{aligned}s_H(v) = s_H(w) \;\Leftrightarrow\; v \cdot H^\mathsf{T} = w \cdot H^\mathsf{T} \;\Leftrightarrow\; (v - w) \cdot H^\mathsf{T} = 0 \\ \overset{3.6}{\Leftrightarrow}\; v - w \in C \;\overset{4.5}{\Leftrightarrow}\; v + C = w + C.\end{aligned}$$

□

4.8 Definition (Nebenklassenanführer)**.** Ein *Anführer* einer Nebenklasse $y + C \in V/U$ ist ein Vertreter $e \in y + C$, der unter allen Vektoren dieser Nebenklasse minimales Gewicht hat:

$$\mathrm{wt}(e) = \min\{\mathrm{wt}(v) \mid v \in y + C\}.$$

4.9 Satz (Eindeutigkeit der Nebenklassenanführer)**.** *Es sei $C \leq V$ ein linearer t-fehlerkorrigierender Code. Dann gilt:*

(a) Jeder Vektor $v \in V$ mit Gewicht $\mathrm{wt}(v) \leq t$ ist ein Anführer seiner Nebenklasse $v + C$.

(b) Sind $v, \widetilde{v} \in C$ Anführer derselben Nebenklasse $v + C = \widetilde{v} + C$ und ist $\mathrm{wt}(v) = \mathrm{wt}(\widetilde{v}) \leq t$, so gilt $v = \widetilde{v}$.

(In Worten: Vektoren von Gewicht $\leq t$ sind eindeutigen Anführer ihrer Nebenklassen.)

Beweis. Es sei $v \in V$ mit $\mathrm{wt}(v) \leq t$.

Ist $C = \{0\}$, so ist $v + C = \{v\}$ und offenbar ist v der eindeutige Anführer von $v + C$.

Es sei nun $C \neq \{0\}$ und $\widetilde{v} \in v + C$ mit $\widetilde{v} \neq v$. Zu zeigen: $\mathrm{wt}(\widetilde{v}) > t$.

Wegen $v + C = \widetilde{v} + C$ und $\widetilde{v} \neq v$ ist $v - \widetilde{v} \in C \setminus \{0\}$. Aus Definitionen 1.6 und 1.9 und Lemma 1.10 folgt

$$\mathrm{wt}(v - \widetilde{v}) \geq \mathrm{wt}(C) \geq \mathrm{d}(C) \geq 2t + 1.$$

Mit der Dreiecksungleichung für den Hamming-Abstand ergibt dies

$$\begin{aligned} 2t + 1 \leq wt(v - \widetilde{v}) = d(v, \widetilde{v}) &\leq d(v, 0) + d(0, \widetilde{v}) \\ &= \mathrm{wt}(v) + \mathrm{wt}(\widetilde{v}) \leq t + \mathrm{wt}(\widetilde{w}), \end{aligned}$$

also ist $\mathrm{wt}(\widetilde{v}) \geq t + 1 > t$. □

4.10 Algorithmus (Syndrom-Dekodierung)**.** Es sei $C \leq V$ ein t-fehlerkorrigierender Code der Länge $n = \dim V$ und der Dimension $k = \dim C$. Es sei ferner H eine Kontrollmatrix für C. Zu jeder Nebenklasse $v + C \in V/C$ wähle man einen Anführer (nach Satz 4.9 teilweise eindeutig bestimmt) $e \in v + C$ und schreibe mit seinem Syndrom $s_H(e)$ in einer Liste.

Merke: Insgesamt gibt es $|V| = 2^n$ Vektoren in V, die sich nach Folgerung 4.4 auf die $|V/C| = 2^{n-k}$ verschiedenen Nebenklassen aufteilen. Die zu erstellende Liste hat somit 2^{n-k} Einträge $e, s_H(e)$.

Die Dekodierung erfolgt nun anhand der Liste:

- Es wird der Vektor $y \in V$ empfangen. Angenommen, es sind höchstens t Fehler passiert. Dann ist $y = c + e$, wobei $c \in C$ das zu ermittelnde Codewort und $e \in V$ der unbekannte Fehlervektor mit $\mathrm{wt}(e) \leq t$ ist.

- Berechne das Syndrom $s_H(y) = y \cdot H^{\mathsf{T}}$. Wegen $y = c + e$ gilt $y + C = e + C$ und nach Lemma 4.7: $s_H(y) = s_H(e)$.
- Stelle anhand der Anführer-Syndrom-Liste den Nebenklassenanführer von $y + C = e + C$ fest. Wegen $\mathrm{wt}(e) \leq t$ ist dies nach Satz 4.9 gerade der gesuchte Fehlervektor e!
- Dekodiere y zu $y - e = c \in C$.

Dieses Verfahren läßt sich in Praxis auch effizient durchführen.

4.11 Beispiel. In Beispiel 3.7 haben wir bereits eine Anführer-Syndrom-Liste für den Hamming-Code der Länge 7 aufgestellt. Angenommen, es ist

$$c = (1, 0, 0, 1, 1, 0, 0)$$

übertragen worden, und dabei ist ein Fehler in der letzten Ziffer entstanden, sodass der empfangene Vektor

$$y = (1, 0, 0, 1, 1, 0, 1)$$

lautet. Dann ist

$$s_H(y) = (0, 0, 1),$$

und der Anführer der Nebenklasse mit diesem Syndrom ist tatsächlich

$$e = (0, 0, 0, 0, 0, 0, 1).$$

y wird daher richtig zu $c = y - e$ dekodiert.

Index