Zusammenfassung Codierungstheorie

$$\text{Datenquelle} \xrightarrow{\text{senden}} \text{Kanal} \xrightarrow{\text{empfangen}} \text{Senke}$$

Die Daten liegen bereits digitalisiert vor. Mit dem Problem wie Daten wie bspw. natürliche Sprache möglichst effizient codiert werden, befasst sich die Informationstheorie. In dieser Vorlesung soll es darum gehen, Daten mit einer Kanalcodierung so zu übersetzen, dass Fehler, die bei einer Übertragung über einen fehlerhaften Kanal, korrigiert oder zumindest bemerkt werden.

$$\text{Datenquelle} \xrightarrow[E]{\text{codieren}} \text{Code} \xrightarrow{\text{senden}} \text{Kanal} \xrightarrow{\text{empfangen}} \square$$
$$\xrightarrow[D]{\text{decodieren}} \text{Code} \xrightarrow[E^{-1}]{} \text{Senke}$$

Def. Ein **Alphabet** ist eine Menge Q mit $q > 1$ Elementen, typischerweise $\{0, 1, \ldots, q-1\} \cong \mathbb{Z}_q$.

Bem. $\mathbb{Z}_q$ trägt die Struktur eines Ringes. Falls q eine Primzahlpotenz ist, so gibt es einen Körper $\mathbb{F}_q$ mit q Elementen.

Def. Sei $n \geq 1$. Eine nichtleere Menge $C \subseteq Q^n$ mit $q = |Q|$ heißt **Blockcode** der Länge n über Q oder **q-närer Code** der Länge n. Jedes $c = (c_1, \ldots, c_n) \in C$ heißt ein **Codewort**. Falls $M = |C|$, so nennt man C einen **(n, M)-Code** über Q.

Def. Die **Informationsrate** von C ist dann $R(C) := \log_n(M)/n$. Falls $|M| = q^k$, dann ist $R(C) = k/n$.

Bem. Ist $Q \cong \mathbb{F}_q$, dann ist Q^n ein $\mathbb{F}_q$-VR. Falls C ein Unterraum von Q^n ist, so ist $R(C) = \dim_{\mathbb{F}_q}(C)/n$.

Def. Der **Hamming-Abstand** von $u, v \in Q^n$ ist

$$d(u, v) := |\{i = 1, \ldots, n \,|\, u_i \neq v_i\}|.$$

Lem. Der Hamming-Abstand ist eine Metrik auf Q^n.

Notation. Es sei $C \subseteq Q^n$ ein Code und $y \in Q^n$. Wenn y empfangen wurde, so geht man davon aus, dass das gesendete Wort dasjenige des Codes mit den wenigsten Unterschieden zu y ist, also ein Wort, welches den **Hamming-Abstand** $d(y, C) := \min_{c \in C} d(y, c)$ von y zu C realisiert. Es existiert i. A. kein eindeutiges solches Element, sondern eine Menge

$$N_c(y) := \{\bar{c} \,|\, d(y, C) = d(y, \bar{c})\}.$$

Def.
- Man nennt einen Kanal einen **q-nären symmetrischen Kanal**, falls ein $p \in \mathbb{R}$ mit $0 < p < (q-1)/q$ existiert, sodass

  $$\mathbb{P}(\beta \text{ empfangen} \mid \alpha \text{ gesendet}) = p/q-1$$

  für alle $\beta \neq \alpha \in Q$, also $\mathbb{P}(\alpha \text{ empfangen} \mid \alpha \text{ gesendet}) = 1 - p$.
- Man nennt einen Kanal **gedächtnislos**, falls

  $$\mathbb{P}(y \text{ empfangen} \mid c \text{ gesendet}) = \prod_{i=1}^{n} \mathbb{P}(y_i \text{ empfangen} \mid c_i \text{ gesendet})$$

  für alle Wörter $x, y \in Q^n$ gilt.

Def (**Maximum-Likelihood-Prinzip**). Gegeben sei ein Code $C \subseteq Q^n$ und $y \in Q^n$. Gesucht ist $\hat{c} = \arg\max_{c \in C} \mathbb{P}(y \mid c)$.

Satz. Es seien ein q-närer symm, gedächtnisloser Kanal und ein Code $C \subseteq Q^n$ gegeben. Sei $y \in Q^n$ und $\hat{c} \in C$. Dann sind äquivalent:
- $\mathbb{P}(y \mid \hat{c}) = \max_{c \in C} \mathbb{P}(y \mid c)$
- $\hat{c} \in N_c(y)$

Def. $D : Q^n \to C$ heißt **vollständige Decodierabbildung**, falls

$$\forall\, y \in Q^n \,:\, D(y) \in N_C(y).$$

Def. Die **Kanalkapazität** eines q-nären symmetrischen Kanal ist

$$\kappa(q, p) := \log_2(q) + p \cdot \log_2(p/q-1) + (1-p) \cdot \log_2(1-p).$$

Sie ist ein Maß für die maximale Information, die über den Kanal übertragen werden kann. Die **Entropiefunktion** ist

$$H(q, p) := 1 - \kappa(q, p).$$

Def. Sei C ein Code und D sei eine zugehörige (vollständige) Decodierabbildung. Die **Restfehlerwahrscheinlichkeit** zu (C, D):

$$\mathbb{P}_{\text{err}}(C) := \max_{y \in Q^n, c \in C} \mathbb{P}(D(y) \neq c \mid c \text{ gesendet}, y \text{ empfangen})$$

Satz (**Shannon**). Sei $0 < R < \kappa(q, p)$. Dann gibt es eine Folge $(C_n)_{n \in \mathbb{N}}$ von Codes und zugehörigen Decodierabbildungen D_n mit:
- C_n ist ein (n, M_n)-Code mit Informationsrate $R \leq R(C_n) < \kappa(q, p)$
- $\lim_{n \to \infty} (\mathbb{P}_{\text{err}}(C_n)) = 0$

Def. Der **Minimalabstand** eines (n, M)-Codes C über Q ist

$$d := d(C) := \min_{c, c' \in C, c \neq c'} d(c, c').$$

Man sagt dann, C ist ein q-närer (n, M, d)-Code.

Notation. Für $u \in Q^n$, $l \in \mathbb{N}$ sei $B_l(u) := \{x \in Q^n \,|\, d(x, x) \leq l\}$.

Def.
- Ein Code C heißt **l-fehlerkorrigierend**, falls $B_l(c) \cap B_{l'}(c') = \emptyset$ für alle $c, c' \in C$ mit $c \neq c'$.
- C heißt **m-fehlererkennend**, wenn $B_m(c) \cap C = \{c\}$ f. a. $c \in C$.
- C heißt **genau l-fehlerkorrigierend/-erkennend**, falls C m-fehlerkorr./-erkennend für $m := l$ aber nicht $m := l + 1$ ist.

Satz. Jeder (n, M, d)-Code C ist genau
- $(d-1)$-fehlererkennend und
- $(t := \lfloor d-1/2 \rfloor)$-fehlerkorrigierend.

Bsp. $C = \{000, 111\}$ ist ein binärer $(3, 2, 3)$-Code.

Problem. Gegeben: q, Länge n, Minimalabstand d. Gesucht:

$$A_q(n, d) := \max\{M \,|\, \exists\, (n, M, d)\text{-Code}\}$$

Def. Ein (n, M, d)-Code heißt **optimal**, falls $M = A_q(n, d)$.

Lem. Seien $q, n \in \mathbb{N}$, $q \geq 2$, $n \geq 1$.
- $A_q(n, 1) = q^n$, realisiert durch $C = Q^n$.
- $A_q(n, n) = q$, realisiert durch $C = \{(a, \ldots, a) \,|\, a \in Q\} \subseteq Q^n$
- $d \leq d' \implies A_q(n, d) \geq A_q(n, d')$
- Sei $n \geq 2$ und $d \geq 2$. Dann gilt $A_q(n, d) \leq A_q(n-1, d-1)$.

Kor (**Singletonschranke**). $A_q(n, d) \leq q^{n-d+1}$

Def. Ein Code, der die Singletonschranke mit Gleichheit erfüllt, heißt ein **MDS-Code** (MDS = maximum distance separable).

Bem. Sei $C \subseteq Q^n$ ein (n, M, d)-Code, $T = \{1 \leq t_1 < \ldots < t_{|T|} \leq n\}$ und $\pi_T : C \to Q^{|T|}$, $c \mapsto (c_{t_1}, \ldots, c_{t_{|T|}})$. Ist C ein MDS-Code, so ist π_T bijektiv für alle T mit $|T| = n - d + 1$.

Satz. $A_q(n, 2) = q^{n-1}$, realisiert durch einen Code mit Prüfziffer

Def. Sei $(G, +, 0)$ eine kommutative Gruppe.
Das **Hamming-Gewicht** von $x \in G^n$ ist

$$\mathrm{wt}(x) := |\mathrm{supp}(x)|, \quad \text{wobei} \quad \mathrm{supp}(x) := \{i \,|\, x_i \neq 0\}.$$

Lem. Sei G wie oben, $x, y \in G^n$. Dann $\mathrm{wt}(x - y) = d(x, y)$.

Satz. $A_q(n, 2) = q^{n-1}$ für alle $q \geq 2$ und alle $n \geq 2$.

Beweis. Wir konstruieren einen $(n, q^{n-1}, 2)$-Code. Sei R ein kommutativer Ring mit q Elementen, $\lambda_1, \ldots, \lambda_{n-1} \in R$ Einheiten und $\lambda_n := -1$. Wir betrachten die Kontrollgleichung

$$\kappa : R^n \to R, \quad z \mapsto \lambda_1 z_1 + \ldots + \lambda_n z_n.$$

Dann ist $C := \ker(\kappa)$ ein 1-fehlererkennender Code. $\square$

Lem. Falls $\lambda_2 - \lambda_1, \ldots, \lambda_n - \lambda_{n-1}$ ebenfalls Einheiten sind, so sind Nachbarvertauschungen als Fehler erkennbar.

Bspe.
- Für $q = 2, R = \mathbb{Z}_2, \lambda_1 = \ldots = \lambda_{n-1} = 1$ heißt $C := \ker(\kappa)$ **Parity-Check-Erweiterung**.
- Beim ISBN-Code ist $R = \mathbb{Z}_{11}$, $\lambda_1 = 1, \ldots, \lambda_9 = 9$, also $\kappa(z) = \sum_{i=1}^{10} i z_i$.

Bem. Es gilt $A_q(4, 3) = q^2 \iff$ es gibt ein Paar orthogonaler lateinischer Quadrate der Größe $q \iff q \neq 2$ oder $q \neq 6$.

Lem. Für $x, y \in \mathbb{Z}_2^n$ gilt $d(x, y) = \mathrm{wt}(x) + \mathrm{wt}(y) - 2 \cdot \mathrm{wt}(x \cdot y)$.

Satz. Für alle $n \geq 1$ und d ungerade gilt $A_2(n, d) = A_2(n+1, d+1)$, realisiert durch die Parity-Check-Erweiterung.

Def. Zwei (n, M)-Codes C, C' über Q heißen **äquivalent**, falls gilt: Es gibt eine Permutation γ auf $\{1, \ldots, n\}$ und Permutationen $\sigma_1, \ldots, \sigma_n$ auf Q, sodass

$$\alpha : Q^n \to Q^n, \quad (x_1, \ldots, x_n) \mapsto (\sigma_1(x_{\gamma(1)}), \ldots, \sigma_n(x_{\gamma(n)}))$$

den Code C auf C' abbildet.

Bsp. $A_2(5, 3) = 4$ realisiert durch $\{00000, 11100, 00111, 11011\}$

Lem. Sei Q ein Alphabet, $u \in Q^n$. Dann gilt

$$|B_l(u)| = \sum_{j=0}^{l} \binom{n}{j} (|Q| - 1)^j.$$

Satz (**Kugelpackungsschranke** (KPS))**.** Sei $q \geq 2$, $n \geq 2$, $1 \leq d \leq n$, $t \coloneqq \lfloor \frac{d-1}{2} \rfloor$. Dann ist

$$A_q(n,d) \leq q^n / \sum_{j=0}^{t} \binom{n}{j} (q-1)^j.$$

Def. Ein q-närer (n,M,d)-Code C heißt **perfekt**, falls M gleich der Kugelpackungsschranke ist.

Bem. Die KPS kann zur **Johnsen-Schranke** verbessert werden. Zusammen mit dem letzten Beispiel liefert diese $A_2(6,3) = 8$.

Bsp. Für q=2, n=7, d=3 liefert die KGS genau $A_2(7,3) \leq 16$.

Bem. Zu jeder Primzahlpotenz $q = p^m \geq 2$ gibt es bis auf Isomorphie genau einen Körper $\mathbb{GF}_q = \mathbb{F}_q$ mit q Elementen. Die Charakteristik dieses Körpers ist p. Ist q keine Primzahlpotenz, so gibt es auch keinen Körper mit q Elementen.

Konstr. Sei $q = p^m$, p prim. Dann gibt es ein irreduzibles Polynom $g(x)$ in $\mathbb{Z}_p$ mit $\deg(g) = m$. Dann ist $\mathbb{F}_q \coloneqq \mathbb{Z}_p[x]/(g(x))$.

Def. Ein **$\mathbb{F}_q$-linearer Code** der Länge n ist ein $\mathbb{F}_q$-Teilraum $\mathbb{F}_q^n$.

Notation. Sei C ein $\mathbb{R}_q$-linearer Code. Sei $k \coloneqq \dim(C)$. Dann ist $|C| = q^k$, also C ein (n, q^k)-Code. Man sagt, C ist ein $[n,k]$-Code. Ist d der Minimalabstand von C, so: C ist ein $[n,k,d]$-Code.

Def. Sei C ein $\mathbb{F}_q$-linearer Code mit $\dim(C) \geq 1$. Das **Minimalgewicht** von C ist $\min\{\mathrm{wt}(c) \mid c \in C, c \neq 0\}$.

Lem. Sei C ein $\mathbb{F}_q$-linearer Code mit $\dim(C) \geq 1$. Dann:

$$\text{Minimalgewicht von } C = \text{Minimalabstand von } C.$$

Bsp. Folgender Code ist ein bin. $(6,8,3)$-Code bzw. $[6,3,3]$-Code:

$$\left\{\begin{matrix} 000000, 100101, 010110, 001111, \\ 110011, 101010, 011001, 111100 \end{matrix}\right\} = \mathrm{span}\{100101, 010110, 001111\}$$

Problem. Gegeben sei $\mathbb{F}_q$, die Länge n und der Minimalabstand d. Gesucht ist $A_q^{\mathrm{lin}}(n,d)$, die bestmögliche Anzahl Wörter eines Codes mit diesen Parametern.

Bem. Klar ist $A_q^{\mathrm{lin}}(n,d) \leq A_q(n,d)$.

Lem.
- $A_q^{\mathrm{lin}}(n,1) = q^n = A_q(n,1)$
- $A_q^{\mathrm{lin}}(n,n) = q = A_q(n,n)$
- $d \leq d' \implies A_q^{\mathrm{lin}}(n,d) \geq A_q^{\mathrm{lin}}(n,d')$
- Für $n \geq 2$, $d \geq 2$ ist $A_q^{\mathrm{lin}}(n,d) \leq A_q^{\mathrm{lin}}(n-1,d-1)$.

Bem. Da die Parity-Check-Erweiterung durch eine lineare Abbildung geschieht, gilt:

Satz. $A_1^{\mathrm{lin}}(n,2) = q^{n-1} = A_q(n,2)$

Satz. Falls d ungerade, so ist $A_2^{\mathrm{lin}}(n,d) = A_2^{\mathrm{lin}}(n+1,d+1)$

Def. Sei C ein $[n,k]$-Code über $\mathbb{F}_q$, d. h. es gibt eine injektive Codierabbildung $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ mit $\mathrm{im}(E) = C$. Dann heißt für jede Basis $g^1, \dots, g^k \in C$ von C die Matrix

$$G \coloneqq \begin{pmatrix} g^1 \\ \vdots \\ g^k \end{pmatrix} \in \mathbb{F}_q^{k \times n} \qquad \text{eine } \textbf{Generatormatrix} \text{ von } C.$$

Bem. Dann ist $E(u) = uG = \sum_{j=1}^{k} u_j g^j \in C$

Def. Zwei $[n,k]$-Codes $C, C' \subseteq \mathbb{F}_q^n$ heißen **linear äquivalent**, falls es $\gamma \in S_n$ und $\lambda_1, \dots, \lambda_n \in \mathbb{F}_q^\times$ gibt, sodass die monomiale Transf.

$$\alpha : \mathbb{F}_q^n \to \mathbb{F}_q^n, \quad (x_1, \dots, x_n) \mapsto (\lambda_1 x_{\gamma(1)}, \dots, \lambda_n x_{\gamma(n)})$$

den Code C in C' überführt.

Def. Ein $[n,k]$-Code heißt **systematisch**, falls die ersten k Spalten seiner Generatormatrix die Standardbasisvektoren sind.

Notation. Sei $C \subset \mathbb{F}_q^n$ ein UVR. Für $x, y \in \mathbb{F}_q^n$ schreiben wir

$$x \equiv y \pmod{C} \quad :\iff \quad x - y \in C.$$

Die zu $x \in V$ gehörende Kongruenzklasse modulo C ist $x + C$.

Def. Ein Repräsentantensystem $\mathcal{R}$ dieser Klassen heißt **gewichtsminimal**, falls $\mathrm{wt}(r) = \min_{c \in C} \mathrm{wt}(r+c)$ für alle $r \in \mathcal{R}$.

Satz. Sei C ein $[n,k]$-Code über $\mathbb{F}_q$, $\mathcal{R}$ ein gewichtsmin. Repräsentantensystem mod C. Zu $y \in \mathbb{F}_q^n$ sei $\mathcal{R}(y) \in \mathbb{R}$ mit $\mathcal{R}(y) + C = y + C$. Dann ist $D : \mathbb{F}_q^n \to C$, $y \mapsto y - \mathcal{R}(y)$ eine Decodierabbildung.

Bem. Sei $\mathbb{F}$ ein Körper, $n \in \mathbb{N}^*$. Das Standard-Skalarprodukt

$$(-|-) : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}, \quad (x,y) \mapsto \sum_{i=1}^{n} x_i y_i$$

ist eine nicht-ausgeartete, symmetrische Bilinearform.

Achtung. Es ist $\dim(U^\perp) = n - k$, im Allgemeinen gilt aber $U \cap U^\perp \neq 0$, z. B. ist $11011 \in \mathbb{F}_2^5$ senkrecht zu sich selbst.

Def. Sei C ein $[n,k]$-Code über $\mathbb{F}_q$. Dann heißt $C^\perp$ der zu C gehörende **duale Code**.

Def. Die Generatormatrix H von $C^\perp$ heißt **Kontrollmatrix** zu C.

Lem. $x \in C \iff Hx^T = 0$

Algorithmus. Sei C ein $[n,k]$-Code, $H \in \mathbb{F}_q^{n-k \times n}$ die Kontrollmatrix. Dann ist

$$\psi_H : \mathbb{F}_q^n \to \mathbb{F}_q^{n-k}, \quad x \mapsto Hx^T$$

eine surjektive lineare Abbildung mit $\ker(\psi_H) = C$.

- Sei $c \in C$ gesendet, $y \in \mathbb{F}_q^n$ empfangen, etwa $y = c + e$. Wir als Empfänger kennen jedoch c und e nicht, nur y. Trotzdem können wir das **Syndrom** $s \coloneqq \phi_H(y) = Hc^T + He^T = He^T \in \mathbb{F}_q^{n-k}$ berechnen.
- Wahrscheinlich ist e ein gewichtsminimaler Repräsentant von y. Sei also $\mathcal{R}$ ein minimales Repräsentantensystem.
  Dann ist $\psi \coloneqq \psi_H|_{\mathcal{R}} : \mathcal{R} \to \mathbb{F}_q^{n-k}$ bijektiv.
  Dann definiert $D(y) \coloneqq y - \psi^{-1}(s)$ eine Decodierabbildung.

Satz. Sei C ein linearer $[n,k,d]$-Code über $\mathbb{F}_q$, H eine Kontrollmatrix zu C. Dann gilt:

$$\begin{aligned} d &= 1 + \max\{a \in \mathbb{N}^* \mid \text{je } a \text{ Spalten von } H \text{ sind lineare unabhängig}\} \\ &= \min\{m \in \mathbb{N}^* \mid \text{es gibt } m \text{ linear abhängige Spalten in } H\} \end{aligned}$$

Def. Sei C ein linearer Code der Länge n über $\mathbb{F}_q$.
Die **Gewichtsverteilung von C** ist $A = A_C \in \mathbb{N}^{\{0,1,\dots,n\}}$ mit

$$A(i) \coloneqq \{w \in C \mid \mathrm{wt}(w) = i\}, \quad 0 \leq i \leq n.$$

Bem. Es gilt $A_0 = 1$ und $A_1 = A_2 = \ldots = A_{d-1} = 0$, falls d das Minimalgewicht von C ist.

Def. $A_C(Z) \coloneqq \sum_{i=0}^{k} A_i Z^k \in \mathbb{C}[Z]$ heißt **Gewichtszählpolynom**,

$$A_C^{\mathrm{hom}}(X,Y) \coloneqq \sum_{i=0}^{n} A_i X^{n-i} \cdot Y^i \in \mathbb{C}[X,Y]$$

heißt **homogenes Gewichtszählpolynom**.

Bem.
- $A_C(Z) = A_C^{\mathrm{hom}}(1,Z)$ • $A_C^{\mathrm{hom}}(X,Y) = X^n \cdot A_C(\frac{Y}{X})$
- Aus $A_C(X,Y)$ erhält man durch die sogenannte Mac-Williams-Transformation $A_{C^\perp}(X,Y)$

Lem. Sei C ein perfekter (n,M,d)-Code. Dann ist d ungerade.

Bem. Wir betrachten nun perfekte Codes C mit $t = 1$, also $d = 3$. Es gilt dann $|C| = q^n/1+n(q-1)$, es ist also $1 + n(q-1)$ ein Teiler von q^n. Beispielsweise ist für $q \geq 2$ und $n = q+1$ die Zahl $1 + n(q-1) = q^2$ ein Teiler von q^n. Diese Teilbarkeit ist eine notwendige, aber nicht hinreichende Bedingung für die Existenz eines perfekten $(n,M,3)$-Codes über Q mit $q = |Q|$.

Lem. Seien $p, u, v \in \mathbb{N}$, $p \geq 2$. Dann gilt $u|v \iff p^u - 1 | p^v - 1$.

Prop. Sei C perfekt mit $t = 1$ über Q, wobei $|Q| = q$ eine Primzahlpotenz ist. Dann ist $|C|$ eine q-Potenz.

Bem. Sei nun $q \geq 2$ eine Primzahlpotenz, C ein q-närer perfekter $(n,M,3)$-Code. Dann ist

$$q^k = |C| = M = q^n/1+n(q-1) \iff n = (q^{n-k}-1)/q-1$$

Wie viele Lösungspaare (n,k) gibt es bei festem q? Wir setzen $m \coloneqq n - k$. Dann ist $k(m) \coloneqq n - m$ und $n(m) \coloneqq \frac{q^m-1}{q-1}$. Die Lösungspaare hängen damit nur noch vom Parameter m ab.

Satz. Zu jedem $m \geq 2$ und zu jeder Primzahlpotenz $q \geq 2$ gibt es einen linearen perfekten $[\frac{q^m-1}{q-1}, \frac{q^m-1}{q-1} - m, 3]$-Code über $\mathbb{F}_q$.

Kor. Ist $q \geq 2$ eine Primzahlpotenz, so gilt

$$A_q^{\mathrm{lin}}\left(\tfrac{q^m-1}{q-1}, 3\right) = A_q\left(\tfrac{q^m-1}{q-1}, 3\right) = q^{q^0+\ldots+q^{m-1}-m} \quad \forall m \geq 2, m \in \mathbb{N}$$

Konstr. Ein bin. **Hamming-Code** $\mathrm{Ham}_2(m)$ (ein $[n, n-m, 3]$-Code mit $n \coloneqq 2^m - 1$) ist geg. durch die Kontrollmatrix $H \in \mathbb{F}_2^{m \times n}$, welche jeden Vektor aus $\mathbb{F}_2^m \setminus \{0\}$ in genau einer Spalte stehen hat.

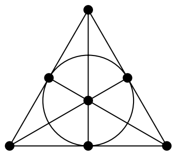
Algorithmus (Decodierung von binären Hamming-Codes)**.** Angenommen, die Spalten der Kontrollmatrix H codieren die Zahlen $1, \dots, 2^m - 1$ im Binärsystem und sind geordnet. Sei $y \in \mathbb{F}_2^n$ empfangen worden. Falls $Hy = 0$, so wurde wsl. y gesendet. Falls das Syndrom Hy ungleich null ist, so ist vermutlich das j-te Bit gekippt, wobei j die Zahl ist, deren Binärcodierung Hy ist.

Prop. Sei $m \geq 2$, $n = 2^m - 1$ und $A \in \mathbb{N}^{0,1,\dots,\mathbb{N}}$ die Gewichtsverteilung des $[n, n-m, 3]$-Hamming-Codes. Dann gilt $A_{n-j} = A_j$ für alle $j = 0, 1, \dots, 2^{m-1} - 1$.

Satz. Die Gewichtsverteilung des binären $[7,4]$-Hamming-Codes ist

$$A = (1, 0, 0, 7, 7, 0, 0, 1).$$

Bem. Sei $C = \mathrm{Ham}_2(3)$, $C_3 := \{c \in C \mid \mathrm{wt}(c) = 3\}$. Für $c \in C_3$ seien $P(c) := \{i = 1, \dots, 7 \mid c_i = 1\}$ die Positionen der in c gesetzten Bits. Falls $i \in P(c)$, so sagen wir, dass i auf der Geraden c liege.
Dies definiert die folgende geometrische Struktur:



~~Die Heiligtümer des Todes~~
Fano-Ebene
$S_1(2, 3, 7)$-Blockplan
Steinersystem
Projektive Ebene $PG(2, \mathbb{F}_2)$

Wir bemerken, dass jede Gerade drei Punkte enthält, jeder Punkt auf drei Geraden liegt, durch je zwei verschiedene Punkte genau eine Gerade verläuft und jedes Paar von Geraden sich in genau einem Punkt schneidet. Die Vierecke in der Fano-Ebene sind die Komplemente von Geraden. Sie entsprechen den Codeworten mit Hamming-Gewicht 4.

Satz. Die Parity-Check-Erweiterung des $[7,4]$-Hamming-Codes ist ein binärer $[8,4,4]$-Code. Dieser ist selbst-dual und optimal. Sein homogenes Gewichtszählpolynom ist $X^8 + 14X^4Y^4 + Y^8$.

Konstr. Wir definieren auf $A := \mathbb{F}_q^n \setminus \{0\}$ eine Äq'relation durch

$$u \sim v \;\; :\Longleftrightarrow \; \exists \lambda \in \mathbb{F}_q \, : \, u = \lambda v.$$

Wir setzen $\mathbb{P} := PG(m-1, \mathbb{F}_q) := A/\!\sim$. Es gilt $|\mathbb{P}| = {}^{q^m - 1}\!/_{q-1} = n$.
Sei $v_1, \dots, v_n$ ein Representantensystem der Äquivalenzklassen. Dann definiert die Kontrollmatrix $H_q(m) := (v_1 \cdots v_n)^T \in \mathbb{F}_q^{m \times n}$ den **q-nären Hamming-Code** $\mathrm{Ham}_q(n)$.

Bem. Wir wählen das Representantensystem wie folgt:

$$\left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ * \end{pmatrix} \right\} \cup \ldots \cup \left\{ \begin{pmatrix} 1 \\ * \\ \vdots \\ * \\ * \end{pmatrix} \right\},$$

also so, dass der erste Eintrag jedes Vektors eine 1 ist.

Algorithmus (Decodieren des q-nären Hamming-Codes)**.** Sei y empfangen mit höchstens einem Fehler. Berechne das Syndrom $s = H_q(m)y^T$. Falls $s = 0$, so ist $D(y) := y$. Angenommen, $s_i \neq 0$. Sei i minimal mit $s_i \neq 0$. Dann ist ${}^{s}\!/_{s_i}$ eine Spalte von $H_q(m)$, etwa die l-te Spalte. Decodiere $D(y) := y - s_i \cdot e_l$.

Def. Sei $q \geq 2$ eine Primzahlpotenz und $m \geq 2$. Der Code $\mathrm{Sim}_q(m) := \mathrm{Ham}_q(m)^\perp$ heißt **Simplex-Code**.

Bem. $\mathrm{Sim}_q(m)$ ist ein $[n, m]$-Code.

Satz. $\mathrm{Sim}_q(m)$ ist **gewichtskonstant**, d. h. jedes vom Nullwort verschiedene Codewort hat Gewicht q^{m-1}.

Bem. Also ist $A_{\mathrm{Sim}_q(m)}(z) = 1 + (q^m - 1) \cdot z^{q^{m-1}}$.

Satz. $A_{\mathrm{Ham}_2(m)}(z) = \frac{1}{n+1}(1+z)^n + \frac{n}{n+1}(1+z)^{\frac{n-1}{2}} \cdot (1-z)^{\frac{n+1}{2}}$, wobei $n = 2^m - 1$. Es gilt die Rekursionsgleichung

$$\binom{n}{l-1} = (n - l + 2) \cdot A_{l-2} + A_{l-1} + l \cdot A_l.$$

Prop. Ist $n \geq 3$ ungerade, so ist der binäre n-Wiederholungscode ein perfekter Code.

Bem. Es sei $d \geq 5$, q eine Primzahlpotenz. Für $n \leq 1000$, $\log_q(M) \leq 1000$ und $q \leq 1000$ könnte es nur Codes mit folgenden Parametern geben:

- $q = 2$, $n = 23$, $d = 7$, $M = 2^{12} = 4096$
- $q = 2$, $n = 90$, $d = 5$, $M = 2^{78}$
- $q = 3$, $n = 11$, $d = 5$, $M = 3^6 = 729$

Satz. Es gibt keinen binären $(90, 2^{78}, 5)$-Code.

Bem. Sei q eine Primzahlpotenz, C ein perfekter q-närer $(n, M, 2t+1)$-Code. Dann hat das **Lloyd-Polynom**

$$L_t(X) := \sum_{j=0}^{t} (-1)^j \cdot (q-1)^{t-j} \cdot \binom{X-1}{j} \binom{n-1-X}{t-j}$$

mindestens t verschiedene Nullstellen in $\{1, \dots, n\}$.

Bsp. Für $n = 90$, $q = 2$, $t = 2$ ist $L_2(X) = 2(X^2 - 90 + 2003)$. Dessen Diskriminante ist 88, also keine Quadratzahl. Somit besitzt $L_2(X)$ keine natürlichen Nullstellen.

Prop. Sei C ein binärer selbst-dualer Code (insb. linear). Dann gilt:

- Jedes Codewort hat ein gerades Gewicht.
- $\forall c \in C \, : \, 4 | \, \mathrm{wt}(c) \iff C$ hat eine Basis B mit $\forall b \in B \, : \, 4 | \, \mathrm{wt}(c)$

Prop. Für jeden ternären selbstdualen Code C gilt $\forall c \in C : 3 | \, \mathrm{wt}(c)$.

Konstr. Beginne mit dem $[7,4,3]$-Hamming-Code C_1 mit Generatormatrix

$$G_1 = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Sei $\overline{C}_1$ die Parity-Check-Erweiterung von C_1 mit Generatormatrix

$$\overline{G}_1 = \left(\begin{array}{c|c} & 1 \\ G_1 & 1 \\ & 1 \\ & 1 \end{array} \right)$$

Dann ist $\overline{C}_1$ eine selbstdualer $[8,4,4]$-Code über $\mathbb{F}_2$.
Sei G_2 die Matrix G_1 mit Spalten in umgekehrter Reihenfolge,

$$G_2 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

der Code C_2 von G_2 erzeugt, $\overline{C}_2$ die Parity-Check-Erw. von C_2. Dann ist $\overline{C}_2$ selbstdual.

Satz. Sei $\Gamma \subset \mathbb{F}_2^{24}$ definiert durch

$$\Gamma := \{(a + f, b + f, a + b + f) \mid a, b \in \overline{C}_1, f \in \overline{C}_2\}$$

Dann ist Γ ein selbst-dualer binärer $[24, 12, 8]$-Code.

Def. Γ wird **erweiterter binärer Golay-Code $\mathcal{G}(24)$** genannt.

Satz. Es gibt einen perfekten binären $[23, 12, 7]$-Code, den **Golay-Code $\mathcal{G}(23)$**. Diesen erhält man aus $\mathcal{G}(24)$ durch Streichen einer Koordinate.

Bem. Umgekehrt ist $\mathcal{G}(24)$ eine Parity-Check-Erw. von $\mathcal{G}(23)$.

Satz.
- $A_{\mathcal{G}(24)}(z) = 1(z^0 + z^{24}) + 759(z^8 + z^{16}) + 2576z^{12}$
- $A_{\mathcal{G}(23)}(z) = 1(z^0 + z^{23}) + 253(z^7 + z^{16}) + 506(z^8 + z^{15}) + 1288(z^{11} + z^{12})$

Bem. $\mathcal{G}(24)$ hat eine Generatormatrizen der Form $G_1 = [E|M]$ und $G_2 = [M|E]$, wobei M symmetrisch ist. Beide Matrizen sind gleichzeitig auch Kontrollmatrizen. Die Matrix M hat dabei besondere Eigenschaften, die zum Decodieren ausnutzen kann.

Satz. $\mathcal{G}(12) := \Omega \subset \mathbb{F}_3^{12}$ sei der Code mit Generatormatrix $G = [E_6|M] \in \mathbb{F}_3^{6 \times 12}$, wobei

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 2 & 1 & 2 \\ 1 & 0 & 1 & 2 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 & 2 \\ 2 & 2 & 1 & 0 & 1 & 2 \\ 1 & 2 & 2 & 1 & 0 & 2 \end{pmatrix}$$

Dann ist Ω ein $[12, 6, 6]$-Code über $\mathbb{F}_3$, selbstdual.

Satz. Es gibt einen $[11, 6, 5]$-Code über $\mathbb{F}_3$. Dieser ist perfekt und heißt **ternärer Golay-Code $\mathcal{G}(11)$**.

Konstr. Streichen der letzten Koordinate von $\mathcal{G}(12)$.

Def. Eine **Inzidenzstruktur** ist ein Tupel $\mathcal{D} = (V, \mathcal{B}, I)$ bestehend aus

- einer (nichtleere) Menge von **Punkten** (oder *Knoten*) V,
- einer (nichtleere) Menge von **Blöcken** (oder *Geraden*) $\mathcal{B}$ und
- einer **Inzidenzrelation** $I \subseteq V \times \mathcal{B}$

Notation. $pIB \iff (p, B) \in I$

Bem. Wir können die Inzidenzrelation durch eine Matrix $M \in \mathbb{C}^{|V| \times |\mathcal{B}|}$ darstellen, welche Einträge in $\{0, 1\}$ besitzt.

Notation. $\sigma(p) \coloneqq \{B \in \mathcal{B} \mid pIB\}$
$\Sigma(B) \coloneqq \{p \in V \mid pIB\}$

Def. $\mathcal{D}$ heißt **einfach**, falls Σ injektiv ist.

Notation. Falls $\mathcal{D}$ einfach: $p \in B \;:\!\iff pIB$

Notation. $v \coloneqq |V|$, $b \coloneqq |\mathcal{B}|$

Def. Eine endliche Inzidenzstruktur $\mathcal{D} = (V, \mathcal{B}, I)$ heißt **linearer Raum**, falls gilt:

- $|\Sigma(B)| \geq 2$ für alle $B \in \mathcal{B}$
- $|\sigma x \cap \sigma y| = 1$ für alle $x \neq y \in V$.

Satz. $\mathcal{D} = (V, \mathcal{B}, \in)$ sei ein linearer Raum mit $b \leq 2$. Dann gilt $b \geq v$.

Bsp. Die Inzidenzstruktur

$$V = \{x_1, \dots, x_{v-1}, x_v\}, \quad \mathcal{B} = \{L_1, \dots, L_{v-1}, B\}$$
$$\Sigma(L_i) = \{x_i, x_v\}, \quad \Sigma(B) = \{x_1, \dots, x_{v-1}\}$$

ist ein linearer Raum mit $b = v$. Dieser heißt **entartet**.

Def. Ein nicht-entarteter linearer Raum mit $v = b$ heißt eine (endliche) **projektive Ebene**.

Satz. Sei $\pi = (V, G, \in)$ eine projektive Ebene. Dann gilt:

- Je zwei verschiedene Geraden schneiden sich in genau einem Pkt.
- Es gibt ein $n \in \mathbb{N}$ mit $n \geq 2$, sodass:
  - Jede Gerade enthält $n + 1$ Punkte.
  - Jeder Punkt liegt auf genau $n + 1$ Geraden.
  - $b = v = n^2 + n + 1$

Def. n heißt **Ordnung** von π.

Bsp. Die Fano-Ebene ist die projektive Ebene der Ordnung 2.