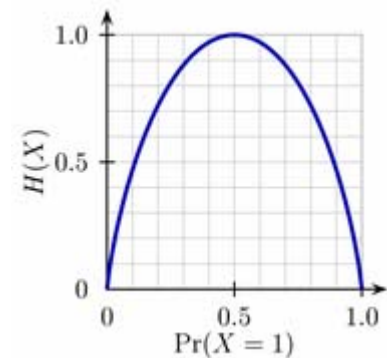


# Information und Kommunikation

Hartmut Klauck  
Universität Frankfurt  
SS 07  
18.5.



# Codes

- Parameter von Codes:
  - $q=|A|$  Größe des Alphabets
  - $n$ : Länge der Codeworte
  - $k=\log_q |C|$  : Länge der Nachrichten
  - $d=\Delta(C)$ : Distanz
- Wir sprechen von  $(n,k,d)_q$ -Codes
- Die Rate eines Codes ist  $k/n$
- Wir wollen also zeigen:
- **Theorem 10.1**
  - $\{bG \mid b \in \{0,1\}^4\}$  ist ein  $(7,4,3)_2$ -Code

# Beweis

- Angenommen  $d_H(bG, b'G) < 3$  für  $b \neq b'$
- Dann ist  $d_H((b-b')G, 0^7) < 3$   
[ $b-b'$  ist hier dasselbe wie  $b \oplus b'$ , die bitweise Paritätsoperation]
- Es gibt also ein  $b'' \neq 0000$ , so dass  $y = b''G$  < 3 Positionen ungleich 0 hat
- $y$  kann nicht  $= 0^7$  sein, da 0000 dieses Codewort hat

# Beweis

- Wir betrachten folgende Matrix:

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

- Behauptung:  $\{bG \mid b \in \{0,1\}^4\} = \{y : Hy=0\}$ 
  - Dies ist leicht nachzuprüfen
- Wir unterscheiden 2 Fälle:
  - $|y|=2$ :  $Hy=h_i+h_j=0$  ( $h_i$  sind die Spalten von  $H$ ): unmöglich
  - $|y|=1$ :  $Hy=h_i=0$ : unmöglich
- Damit gilt, dass das minimale Hamming Gewicht von Codeworten (außer  $0^7$ ) 3 ist

# Bemerkungen

- Hamming Codes sind lineare Codes, d.h. Codes bei denen für Codeworte  $x, y$  auch  $x+y$  ein Codewort ist
- Lineare Codes werden durch eine Generatormatrix  $G$  oder durch eine Parity-Check Matrix  $H$  definiert wie oben

# Hamming Codes

- Wir können die obige Konstruktion verallgemeinern:
  - Sei  $H$  die  $l \times 2^l - 1$  Matrix, die alle binären Strings der Länge  $l$  bis auf den 0-string enthält
  - Die Codeworte sind durch  $\{y: Hy=0\}$  gegeben
  - Die Blocklänge ist damit  $n=2^l-1$

# Hamming Codes

- Was ist die Distanz?
- Wie zuvor müssen wir nur das kleinste Gewicht eines Codewortes  $\neq 0$  bestimmen
  - Es gibt keine Nullspalte, also kann kein Codewort Gewicht 1 haben
  - Jede Summe zweier Spalten ist ungleich 0, denn nur zwei gleiche Spalten addieren sich zu 0.
- D.h. die Distanz ist 3

# Hamming Codes

- Wir brauchen jetzt noch den Wert von  $k$
- Dies ist  $\log_2 |\{y:Hy=0\}|$
- $H$  hat einen Rang von  $l$
- Damit hat der Kern der Abbildung/Matrix  $H$  eine Dimension von  $2^l-l-1$
- D.h.  $k=2^l-l-1$
- Die Rate der Codes ist damit  $1-l/(2^l-1)$  und die Codes korrigieren einen Fehler

# Generatormatrix

- Wie erhalten wir  $G$ ?
- $G$  soll  $\{0,1\}^{2^l-1-l}$  auf den Kern von  $H$  abbilden
- Da der Kern ein Unterraum ist, existiert eine solche lineare Abbildung  $G$

# Fehlerkorrektur

- Wir kehren zum  $(7,4,3)_2$  Code zurück
- Wir wissen, die Distanz ist 3. Wie können wir dekodieren?
- Sei  $e_i$  der Vektor mit 1 an Stelle  $i$  und 0 sonst
- Wenn  $c$  ein Codewort ist, ist  $r=c\oplus e_i$  ein korrumpiertes Codewort mit 1 Fehler
- Ziel ist es,  $i$  zu bestimmen

# Fehlerkorrektur

- Es gilt nun  $Hr = H(c + e_i) = 0 + He_i = h_i$ , die  $i$ -te Spalte von  $H$
- D.h.  $Hr$  gibt uns an, wo der Fehler aufgetreten ist, und wir können korrigieren
- $h_i$  ist die Binärdarstellung von  $i$

# Kodierung

- $H$  beschreibt die Menge der Codeworte
- Man kann leicht sehen, dass die ersten 4 Zeichen der Codeworte über alle Werte von  $\{0,1\}^4$  laufen.
- Wir können daher diese als die Nachricht auffassen, und die restlichen 3 Bits als parity check Bits
- Solche Codes nennt man *systematisch*

# Größere Distanz

- Angenommen wir wollen Distanz 4 erreichen
- Dazu hängen wir einfach ein parity check bit an (die Parität der 7 Bits des Codewortes)
- **Lemma 10.2**
  - Gegeben einen Code  $C$  mit Distanz  $2t-1$ , erhalten wir einen Code mit Distanz  $2t$  durch Anhängen eines Parity Check Bits.
- **Beweis:**
  - Klar ist, dass für  $x, y \in C$  die Distanz nicht sinkt
  - Wenn  $d_H(x, y) = 2t-1$ , dann stimmen  $x$  und  $y$  an einer ungeraden Anzahl von Stellen überein, d.h. sie haben eine unterschiedliche Parität, d.h. ihre Distanz steigt auf  $2t$
- **Beispiele:**
  - Hamming Code mit Paritätsbit hat Distanz 4
  - Ein Paritätsbit allein ergibt einen Code mit Distanz 2

# Lineare Code

- Das Alphabet  $A$  sei ein endlicher Körper
- Wir sagen ein Code  $C$  ist linear, wenn er ein Unterraum von  $A^n$  ist
  - $x, y \in C \Rightarrow x + y \in C$
  - $a \in A, x \in C \Rightarrow ax \in C$
- Unterräume können immer als Kern einer linearen Abbildung/Matrix angesehen werden
- D.h. wir erhalten immer eine „parity check matrix“ und können Fehler entdecken (nicht notwendigerweise einfach beheben)
- Umgekehrt können wir immer mittels einer Generatormatrix kodieren, d.h. durch Matrix-Vektor Multiplikation

# Duale Codes

- Zu einem linearen Code  $C$  gibt es einen dualen Code:
  - Wenn  $C$  Generatormatrix  $G$  und Parity Check Matrix  $H$  hat, hat der duale Code Generatormatrix  $H$
  - Der duale Code des dualen Codes von  $C$  ist  $C$

# Duale Codes

- Was ist der duale Code zum Hamming Code?
- Wir betrachten den Hamming Code mit Blocklänge  $n=2^l-1$
- $H$  ist die  $l \times 2^l-1$  Matrix, die alle Strings  $\neq 0$  als Spalten hat
- $\{bH: b \in \{0,1\}^l\}$  ist die Menge der Codeworte
- Codeworte haben eine Länge von  $n=2^l-1$
- Die Kodierung ist also  $\langle b, x \rangle_{x \in \{0,1\}^l - \{0\}}$
- D.h. die Parität jeder nichtleeren Teilmenge der Bits von  $b$
- $\langle y, x \rangle = \sum_i y_i x_i \pmod 2$

# Hadamard Codes

- Wir haben also einen  $(n, l, ?)_2$  Code
- Was ist die Distanz dieses (sehr langen) Codes?
- Es gilt:
  - Wenn  $x \neq y$  und  $x, y \in \{0, 1\}^l$ , dann ist mit Wahrscheinlichkeit  $1/2$  über alle  $z \in \{0, 1\}^l$ :  
 $\langle x, z \rangle \neq \langle y, z \rangle$
  - Beweis Übung
- D.h. Die Distanz ist  $2^{l-1}$ , also extrem groß
- Der Duale Code zu Hamming Codes ist also ein  $(2^l - 1, l, 2^{l-1})_2$  Code
- Die Rate ist damit  $l / (2^l - 1)$ , aber es können  $n/4$  Fehler korrigiert werden

# Hadamard Codes

- Die Matrix der Codeworte ist eine  $2^l - 1 \times 2^l$  Matrix
- Wenn wir überall eine Null anhängen, und 0 nach -1 ändern erhalten wir eine Hadamard Matrix

# Hadamard Codes

- Wir betrachten noch eine interessante Eigenschaft von Hadamard Codes:
  - Lokale Dekodierbarkeit
- Angenommen wir erhalten zu einem Hadamard Codewort  $x$  (für string  $a$ ) der Länge  $n$  ein Wort  $z$  mit  $\delta n$  Fehlern
- Wir wollen  $x_u$  wissen, dabei sei  $u \in \{0,1\}^l$
- Es kann sein, dass  $x_u$  in  $z$  verfälscht ist
- Wir wollen aber nur wenige Positionen von  $z$  anschauen, um das richtige  $x_u$  zu bestimmen
- Folgende Prozedur funktioniert:
  - Ziehe  $v$  zufällig aus  $\{0,1\}^l$
  - Setze  $w = u \oplus v$
  - Gebe  $z_v \oplus z_w$  aus

# Hadamard Codes

- Behauptung:
  - Die Ausgabe ist mit Wahrscheinlichkeit  $1-2\delta$  korrekt, d.h.  $=x_u$
- Beweis:
  - $v, w$  sind uniform zufällig
  - Mit Wahrscheinlichkeit  $\delta$  gilt  $x_v \neq z_v$  und analog für  $w$
  - Mit Wahrscheinlichkeit  $1-2\delta$  gilt daher  $z_v = x_v$  und  $z_w = x_w$
  - Mit Wahrscheinlichkeit  $1-2\delta$  ist die Ausgabe  $x_w \oplus x_v = \langle a, w \rangle \oplus \langle a, v \rangle = \langle a, v \oplus w \rangle = \langle a, u \rangle = x_u$