Codierungstheorie

Ausblick zur Linearen Algebra A

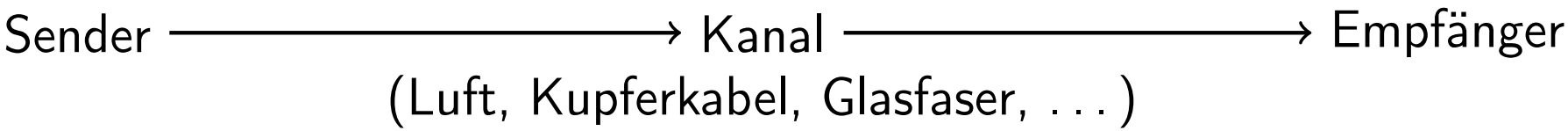
Benjamin Sambale

Leibniz Universität Hannover

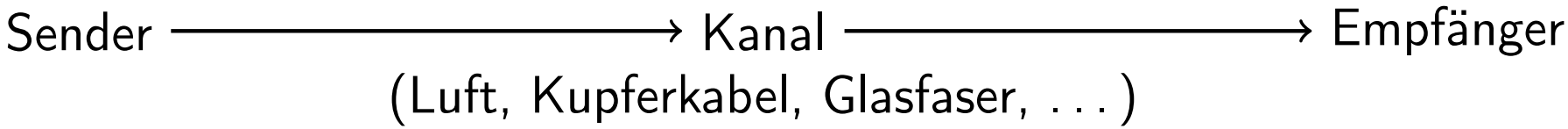
29. 01. 2021

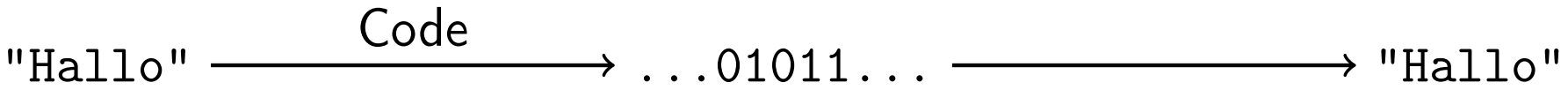
programmiert mit LaTeX + Ti*k*Z + BEAMER

Situation

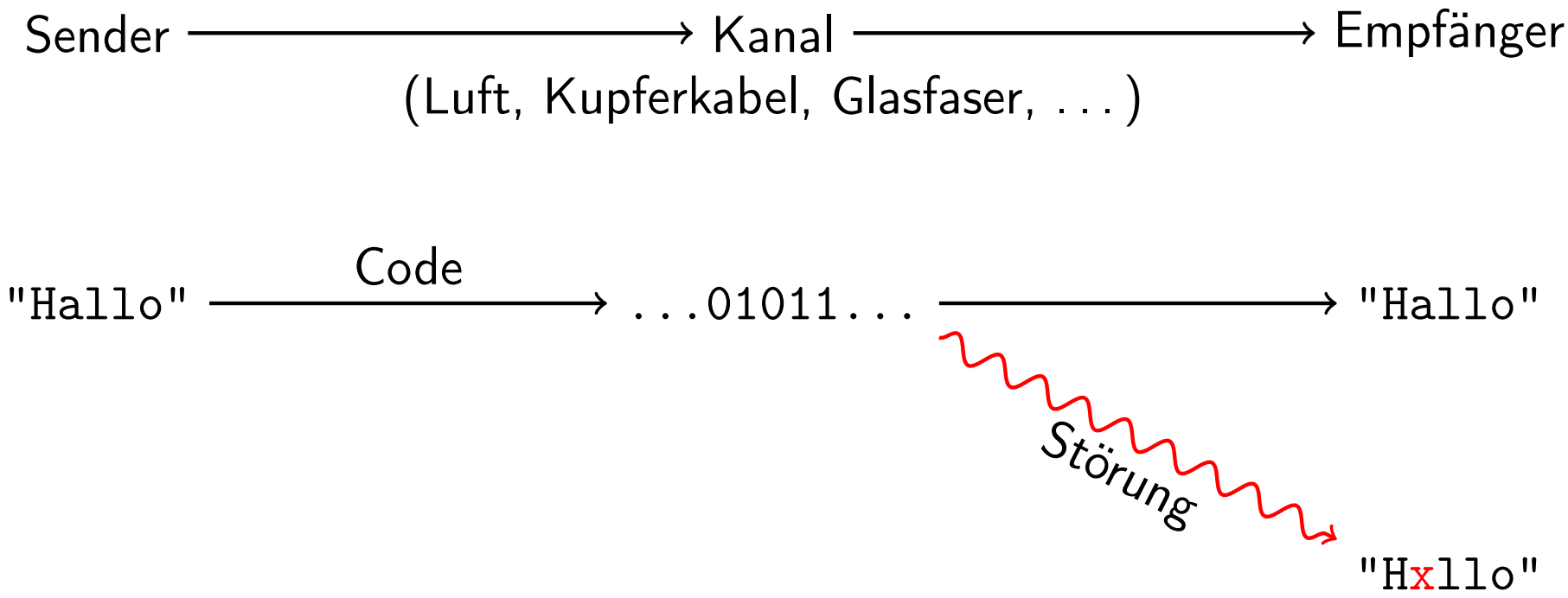


Situation


Sender
Kanal
Empfänger
(Luft, Kupferkabel, Glasfaser, . . .)


"Hallo"
Code
...01011...
"Hallo"

Situation


Sender
Kanal
Empfänger
(Luft, Kupferkabel, Glasfaser, . . .)
"Hallo"
Code
...01011...
"Hallo"
Störung
"Hxllo"

Warum werden Daten codiert?

1. Speicherung/Verarbeitung akustischer oder visueller Signale
   $\longrightarrow$ Signalverarbeitung $\longrightarrow$ Nachrichtentechnik

Warum werden Daten codiert?

1. Speicherung/Verarbeitung akustischer oder visueller Signale $\longrightarrow$ Signalverarbeitung $\longrightarrow$ Nachrichtentechnik
2. Geringe Bandbreite/Geschwindigkeit $\longrightarrow$ Datenkompression (Bsp. zip, mp3, jpeg) $\longrightarrow$ Informationstheorie

Warum werden Daten codiert?

1. Speicherung/Verarbeitung akustischer oder visueller Signale $\longrightarrow$ Signalverarbeitung $\longrightarrow$ Nachrichtentechnik
2. Geringe Bandbreite/Geschwindigkeit $\longrightarrow$ Datenkompression (Bsp. zip, mp3, jpeg) $\longrightarrow$ Informationstheorie
3. Geheimhaltung $\longrightarrow$ Kryptographie (Bsp. AES, DSA, RSA) $\longrightarrow$ Zahlentheorie

Warum werden Daten codiert?

1. Speicherung/Verarbeitung akustischer oder visueller Signale $\longrightarrow$ Signalverarbeitung $\longrightarrow$ Nachrichtentechnik
2. Geringe Bandbreite/Geschwindigkeit $\longrightarrow$ Datenkompression (Bsp. zip, mp3, jpeg) $\longrightarrow$ Informationstheorie
3. Geheimhaltung $\longrightarrow$ Kryptographie (Bsp. AES, DSA, RSA) $\longrightarrow$ Zahlentheorie
4. Störungsresistenz $\longrightarrow$ Codierungstheorie $\longrightarrow$ Lineare Algebra

Warum werden Daten codiert?

1. Speicherung/Verarbeitung akustischer oder visueller Signale $\longrightarrow$ Signalverarbeitung $\longrightarrow$ Nachrichtentechnik
2. Geringe Bandbreite/Geschwindigkeit $\longrightarrow$ Datenkompression (Bsp. zip, mp3, jpeg) $\longrightarrow$ Informationstheorie
3. Geheimhaltung $\longrightarrow$ Kryptographie (Bsp. AES, DSA, RSA) $\longrightarrow$ Zahlentheorie
4. Störungsresistenz $\longrightarrow$ Codierungstheorie $\longrightarrow$ Lineare Algebra

In der Praxis kombiniert man diese Verfahren.

Wie erkennt man Übertragungsfehler?

- Idee: Sende jedes Symbol zweimal:

$$\texttt{"Test"} \longrightarrow \texttt{"TTeesstt"} \longrightarrow \texttt{"TTeesxtt"}$$

Wie erkennt man Übertragungsfehler?

- Idee: Sende jedes Symbol zweimal:

  `"Test"` $\longrightarrow$ `"TTeesstt"` $\longrightarrow$ `"TTeesxtt"`

- Damit lässt sich ein Fehler erkennen, aber nicht korrigieren (`s` oder `x` ist falsch).

Wie erkennt man Übertragungsfehler?

- Idee: Sende jedes Symbol zweimal:

  $$\texttt{"Test"} \longrightarrow \texttt{"TTeesstt"} \longrightarrow \texttt{"TTeesxtt"}$$

- Damit lässt sich ein Fehler erkennen, aber nicht korrigieren (`s` oder `x` ist falsch).
- Besser: Sende jedes Symbol dreimal (`s` $\rightarrow$ `sss` $\rightarrow$ `sxs`).
  Dann lässt sich ein Fehler erkennen und korrigieren, aber nicht zwei.

Wie erkennt man Übertragungsfehler?

- Idee: Sende jedes Symbol zweimal:

  $$\texttt{"Test"} \longrightarrow \texttt{"TTeesstt"} \longrightarrow \texttt{"TTeesxtt"}$$

- Damit lässt sich ein Fehler erkennen, aber nicht korrigieren (`s` oder `x` ist falsch).
- Besser: Sende jedes Symbol dreimal (`s` $\rightarrow$ `sss` $\rightarrow$ `sxs`).
  Dann lässt sich ein Fehler erkennen und korrigieren, aber nicht zwei.
- Allgemein: Mit n Wiederholungen lassen sich $\frac{n-1}{2}$ Fehler korrigieren.
  Nachteil: Datenmenge erhöht sich.

Coderaum

Definition

Ein Code der Länge n ist eine nichtleere Teilmenge $C \subsetneq \mathbb{F}_2^n$. Die Elemente von C heißen Codeworte.

Coderaum

Definition

Ein Code der Länge n ist eine nichtleere Teilmenge $C \subsetneq \mathbb{F}_2^n$. Die Elemente von C heißen Codeworte. Für $x, y \in \mathbb{F}_2^n$ sei

$$d(x, y) := |\{i : x_i \neq y_i\}|$$

der Abstand von x und y (Metrik im Sinne der Analysis).

Coderaum

Definition

Ein Code der Länge n ist eine nichtleere Teilmenge $C \subsetneq \mathbb{F}_2^n$. Die Elemente von C heißen Codeworte. Für $x, y \in \mathbb{F}_2^n$ sei

$$d(x, y) := |\{i : x_i \neq y_i\}|$$

der Abstand von x und y (Metrik im Sinne der Analysis).

- Sei S eine Menge von zu sendenden Daten (Bsp. lateinisches Alphabet).

Coderaum

Definition

Ein Code der Länge n ist eine nichtleere Teilmenge $C \subsetneq \mathbb{F}_2^n$. Die Elemente von C heißen Codeworte. Für $x, y \in \mathbb{F}_2^n$ sei

$$d(x, y) := |\{i : x_i \neq y_i\}|$$

der Abstand von x und y (Metrik im Sinne der Analysis).

- Sei S eine Menge von zu sendenden Daten (Bsp. lateinisches Alphabet).
- Codierung ist eine bijektive Abbildung $\gamma : S \to C$.

Coderaum

Definition

Ein Code der Länge n ist eine nichtleere Teilmenge $C \subsetneq \mathbb{F}_2^n$. Die Elemente von C heißen Codeworte. Für $x, y \in \mathbb{F}_2^n$ sei

$$d(x, y) := |\{i : x_i \neq y_i\}|$$

der Abstand von x und y (Metrik im Sinne der Analysis).

- Sei S eine Menge von zu sendenden Daten (Bsp. lateinisches Alphabet).
- Codierung ist eine bijektive Abbildung $\gamma : S \to C$.
- Decodierung ist eine Abbildung $\gamma' : \mathbb{F}_2^n \to S$ mit $\gamma' \circ \gamma = \mathrm{id}_S$.

Fehler erkennen und korrigieren

- Übertragungsfehler werden erkannt, falls $x := \gamma(s) \notin C$.

Fehler erkennen und korrigieren

- Übertragungsfehler werden erkannt, falls $x := \gamma(s) \notin C$.
- Fehler werden korrigiert, falls genau ein Codewort $c \in C$ mit minimalem Abstand zu x existiert. Setze $\gamma'(x) := \gamma^{-1}(c)$.

Fehler erkennen und korrigieren

- Übertragungsfehler werden erkannt, falls $x := \gamma(s) \notin C$.
- Fehler werden korrigiert, falls genau ein Codewort $c \in C$ mit minimalem Abstand zu x existiert. Setze $\gamma'(x) := \gamma^{-1}(c)$.
- Diese Heuristik setzt voraus, dass Fehler „selten“ und zufällig auftreten.

Fehler erkennen und korrigieren

- Übertragungsfehler werden erkannt, falls $x := \gamma(s) \notin C$.
- Fehler werden korrigiert, falls genau ein Codewort $c \in C$ mit minimalem Abstand zu x existiert. Setze $\gamma'(x) := \gamma^{-1}(c)$.
- Diese Heuristik setzt voraus, dass Fehler „selten“ und zufällig auftreten.
- Bei zerkratzten CDs ist dies beispielsweise nicht erfüllt.

Beispiel Wiederholungscode

$$
\begin{aligned}
S &:= \mathbb{F}_2, \\
C &:= \{(0,\ldots,0),(1,\ldots,1)\} \subseteq \mathbb{F}_2^n, \\
\gamma(x) &:= (x,\ldots,x).
\end{aligned}
$$

Beispiel Wiederholungscode

$$
\begin{aligned}
S &:= \mathbb{F}_2, \\
C &:= \{(0,\ldots,0),(1,\ldots,1)\} \subseteq \mathbb{F}_2^n, \\
\gamma(x) &:= (x,\ldots,x).
\end{aligned}
$$

- Durch eine Störung wird 1 auf $x := (1,1,1,0,1,0) \notin C$ abgebildet.

Beispiel Wiederholungscode

$$
\begin{aligned}
S &:= \mathbb{F}_2, \\
C &:= \{(0,\ldots,0),(1,\ldots,1)\} \subseteq \mathbb{F}_2^n, \\
\gamma(x) &:= (x,\ldots,x).
\end{aligned}
$$

- Durch eine Störung wird 1 auf $x := (1,1,1,0,1,0) \notin C$ abgebildet.
- Offenbar ist $c := (1,1,1,1,1,1) \in C$ das einzige Codewort mit minimalem Abstand $d(x,c) = 2$.

Beispiel Wiederholungscode

$$
\begin{aligned}
S &:= \mathbb{F}_2, \\
C &:= \{(0,\ldots,0),(1,\ldots,1)\} \subseteq \mathbb{F}_2^n, \\
\gamma(x) &:= (x,\ldots,x).
\end{aligned}
$$

- Durch eine Störung wird 1 auf $x := (1,1,1,0,1,0) \notin C$ abgebildet.
- Offenbar ist $c := (1,1,1,1,1,1) \in C$ das einzige Codewort mit minimalem Abstand $d(x,c) = 2$.
- Somit kann x zu c korrigiert werden.

Beispiel Wiederholungscode

$$
\begin{aligned}
S &:= \mathbb{F}_2, \\
C &:= \{(0,\ldots,0),(1,\ldots,1)\} \subseteq \mathbb{F}_2^n, \\
\gamma(x) &:= (x,\ldots,x).
\end{aligned}
$$

- Durch eine Störung wird 1 auf $x := (1,1,1,0,1,0) \notin C$ abgebildet.
- Offenbar ist $c := (1,1,1,1,1,1) \in C$ das einzige Codewort mit minimalem Abstand $d(x,c) = 2$.
- Somit kann x zu c korrigiert werden.
- Andererseits kann $x := (1,0,1,0,1,0)$ nicht korrigiert werden, denn $d(x,(0,\ldots,0)) = 3 = d(x,(1,\ldots,1))$.

Beispiel Wiederholungscode

$$
\begin{aligned}
S &:= \mathbb{F}_2, \\
C &:= \{(0,\ldots,0),(1,\ldots,1)\} \subseteq \mathbb{F}_2^n, \\
\gamma(x) &:= (x,\ldots,x).
\end{aligned}
$$

- Durch eine Störung wird 1 auf $x := (1,1,1,0,1,0) \notin C$ abgebildet.
- Offenbar ist $c := (1,1,1,1,1,1) \in C$ das einzige Codewort mit minimalem Abstand $d(x,c) = 2$.
- Somit kann x zu c korrigiert werden.
- Andererseits kann $x := (1,0,1,0,1,0)$ nicht korrigiert werden, denn $d(x,(0,\ldots,0)) = 3 = d(x,(1,\ldots,1))$.
- Bei mehr als drei Fehlern wird $x := (1,0,0,0,1,0)$ sogar falsch „korrigiert“ zu $c = (0,\ldots,0)$.

ASCII-Code

- $S = \{s_0, \ldots, s_{127}\}$ häufig verwendete Symbole:

`	´	ˆ	˜	¨	˝	°	ˇ	˘	¯	·	¸	˛	,	‹	›
“	”	„	«	»	–	—		o	ı	j	ff	fi	fl	ffi	ffl
␣	!	"	#	$	%	&	’	(	)	*	+	,	-	.	/
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	ˆ	_
‘	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
p	q	r	s	t	u	v	w	x	y	z	{	|	}	˜	-

ASCII-Code

- $S = \{s_0, \ldots, s_{127}\}$ häufig verwendete Symbole:

`	´	ˆ	˜	¨	˝	°	ˇ	˘	¯	·	¸	˛	,	‹	›
“	”	„	«	»	–	—		o	ı	j	ff	fi	fl	ffi	ffl
␣	!	"	#	$	%	&	’	(	)	*	+	,	-	.	/
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
‘	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	-

- Binärcode von k liefert Zuordnung $s_k \longleftrightarrow (a_1, \ldots, a_7) \in \mathbb{F}_2^7$ (beachte: $2^7 = 128$).

ASCII-Code

- $S = \{s_0, \ldots, s_{127}\}$ häufig verwendete Symbole:

`	´	ˆ	˜	¨	˝	°	ˇ	˘	¯	·	¸	˛	,	‹	›
“	”	„	«	»	–	—		o	ı	j	ff	fi	fl	ffi	ffl
␣	!	"	#	$	%	&	’	(	)	*	+	,	-	.	/
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	ˆ	_
‘	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
p	q	r	s	t	u	v	w	x	y	z	{	|	}	˜	-

- Binärcode von k liefert Zuordnung $s_k \longleftrightarrow (a_1, \ldots, a_7) \in \mathbb{F}_2^7$ (beachte: $2^7 = 128$).
- Ergänze Prüfbit $a_8 := a_1 + \ldots + a_7 \in \mathbb{F}_2$. Danach: 8 Bits = 1 Byte.

ASCII-Code

- $S = \{s_0, \ldots, s_{127}\}$ häufig verwendete Symbole:

`	´	ˆ	˜	¨	˝	°	ˇ	˘	¯	·	¸	˛	,	‹	›
“	”	„	«	»	–	—		o	ı	j	ff	fi	fl	ffi	ffl
␣	!	"	#	$	%	&	’	(	)	*	+	,	-	.	/
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	ˆ	_
‘	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
p	q	r	s	t	u	v	w	x	y	z	{	|	}	˜	-

- Binärcode von k liefert Zuordnung $s_k \longleftrightarrow (a_1, \ldots, a_7) \in \mathbb{F}_2^7$ (beachte: $2^7 = 128$).
- Ergänze Prüfbit $a_8 := a_1 + \ldots + a_7 \in \mathbb{F}_2$. Danach: 8 Bits = 1 Byte.
- Nun ist

$$C := \left\{(x_1, \ldots, x_8) \in \mathbb{F}_2^8 : x_1 + \ldots + x_8 = 0\right\} \subseteq \mathbb{F}_2^8$$

ein Code der Länge 8 und Codierung $\gamma : S \to C$, $s_k \mapsto (a_1, \ldots, a_8)$.

ASCII-Code

- Beispiel: $s_{98} = b \longrightarrow (1, 1, 0, 0, 0, 1, 0$

ASCII-Code

- Beispiel: $s_{98} = b \longrightarrow (1, 1, 0, 0, 0, 1, 0, 1)$

ASCII-Code

- Beispiel: $s_{98} = b \longrightarrow (1,1,0,0,0,1,0,1) =: \gamma(b)$

ASCII-Code

- Beispiel: $s_{98} = b \longrightarrow (1,1,0,0,0,1,0,1) =: \gamma(b)$
- ASCII-Code erkennt einen Fehler, aber kann nicht korrigieren, denn für $x \in \mathbb{F}_2^8 \setminus C$ gibt es 8 Codewörter $c \in C$ mit $d(x,c) = 1$.

ASCII-Code

- Beispiel: $s_{98} = b \longrightarrow (1,1,0,0,0,1,0,1) =: \gamma(b)$
- ASCII-Code erkennt einen Fehler, aber kann nicht korrigieren, denn für $x \in \mathbb{F}_2^8 \setminus C$ gibt es 8 Codewörter $c \in C$ mit $d(x,c) = 1$.
- Ähnlich funktionieren

GTIN: 

IBAN und ISBN (erkennt auch Vertauschung von Ziffern).

Lineare Codes

Definition

Ein Code $C \subsetneq \mathbb{F}_2^n$ heißt linear, falls C ein nicht-trivialer Unterraum von $\mathbb{F}_2^n$ ist.

Lineare Codes

Definition

Ein Code $C \subsetneq \mathbb{F}_2^n$ heißt linear, falls C ein nicht-trivialer Unterraum von $\mathbb{F}_2^n$ ist. Gegebenenfalls heißt

$$w(C) := \min\{d(0,c) : c \in C \setminus \{0\}\}$$

Gewicht von C.

Lineare Codes

Definition

Ein Code $C \subsetneq \mathbb{F}_2^n$ heißt linear, falls C ein nicht-trivialer Unterraum von $\mathbb{F}_2^n$ ist. Gegebenenfalls heißt

$$w(C) := \min\{d(0,c) : c \in C \setminus \{0\}\}$$

Gewicht von C.

- Man nennt C einen (n,k,w)-Code, falls $k = \dim C$ und $w = w(C)$.

Lineare Codes

Definition

Ein Code $C \subsetneq \mathbb{F}_2^n$ heißt linear, falls C ein nicht-trivialer Unterraum von $\mathbb{F}_2^n$ ist. Gegebenenfalls heißt

$$w(C) := \min\{d(0,c) : c \in C \setminus \{0\}\}$$

Gewicht von C.

- Man nennt C einen (n,k,w)-Code, falls $k = \dim C$ und $w = w(C)$.
- Die Rate $\frac{k}{n} \leq 1$ von C beschreibt das Verhältnis von Informationsgehalt zu Speicherbedarf.

Fehlerschranken

Satz

Für jeden linearen Code C gilt

1. *C erkennt e Fehler* $\iff$ $w(C) > e$.
2. *C korrigiert e Fehler* $\iff$ $w(C) > 2e$.

Fehlerschranken

Satz

Für jeden linearen Code C gilt

1. *C erkennt e Fehler* $\iff$ $w(C) > e$.
2. *C korrigiert e Fehler* $\iff$ $w(C) > 2e$.

Beweis.

1. $C \leq \mathbb{F}_2^n$ erkennt genau dann e Fehler, falls

$$1 \leq d(c, x) \leq e \implies x \notin C \qquad (\forall c \in C,\ x \in \mathbb{F}_2^n)$$

Fehlerschranken

Satz

Für jeden linearen Code C gilt

1. *C erkennt e Fehler* $\iff$ $w(C) > e$.
2. *C korrigiert e Fehler* $\iff$ $w(C) > 2e$.

Beweis.

1. $C \leq \mathbb{F}_2^n$ erkennt genau dann e Fehler, falls

$$1 \leq d(c, x) \leq e \implies x \notin C \qquad (\forall c \in C,\ x \in \mathbb{F}_2^n)$$
$$d(c, c') \leq e \implies c = c' \qquad (\forall c, c' \in C)$$

Fehlerschranken

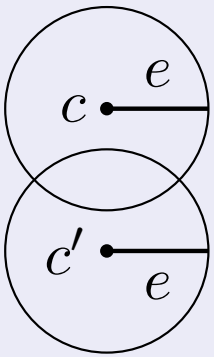
Satz

Für jeden linearen Code C gilt

1. C *erkennt* e *Fehler* $\iff w(C) > e.$
2. C *korrigiert* e *Fehler* $\iff w(C) > 2e.$

Beweis.

1. $C \leq \mathbb{F}_2^n$ erkennt genau dann e Fehler, falls



$$1 \leq d(c,x) \leq e \implies x \notin C \qquad (\forall c \in C,\ x \in \mathbb{F}_2^n)$$
$$d(c,c') \leq e \implies c = c' \qquad (\forall c, c' \in C)$$
$$d(c-c',0) \leq e \implies c - c' = 0 \qquad (\forall c, c' \in C)$$

Fehlerschranken

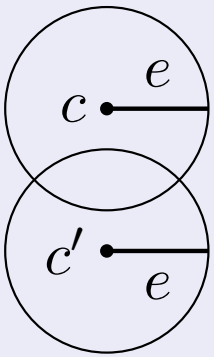
Satz

Für jeden linearen Code C gilt

1. C *erkennt* e *Fehler* $\iff w(C) > e.$
2. C *korrigiert* e *Fehler* $\iff w(C) > 2e.$

Beweis.

1. $C \leq \mathbb{F}_2^n$ erkennt genau dann e Fehler, falls



$$
\begin{aligned}
1 \leq d(c,x) \leq e &\implies x \notin C \qquad (\forall c \in C,\ x \in \mathbb{F}_2^n)\\
d(c,c') \leq e &\implies c = c' \qquad (\forall c, c' \in C)\\
d(c-c',0) \leq e &\implies c - c' = 0 \qquad (\forall c, c' \in C)\\
w(C) &> e.
\end{aligned}
$$

Fehlerschranken

Beweis.

2. $C \leq \mathbb{F}_2^n$ korrigiert genau dann e Fehler, falls

$$d(c, x) \leq e \implies c = x \qquad (\forall c \in C,\ x \in \mathbb{F}_2^n)$$

Fehlerschranken

Beweis.

2. $C \leq \mathbb{F}_2^n$ korrigiert genau dann e Fehler, falls

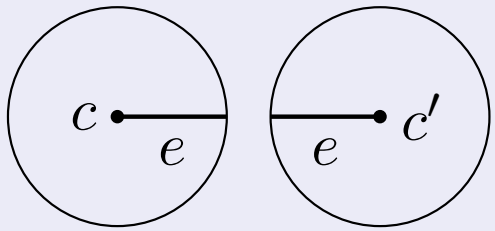
$$d(c,x) \leq e \implies c = x \qquad (\forall c \in C,\ x \in \mathbb{F}_2^n)$$
$$d(c,x), d(c',x) \leq e \implies c = x = c' \qquad (\forall c, c' \in C,\ x \in \mathbb{F}_2^n)$$

Fehlerschranken

Beweis.

2. $C \leq \mathbb{F}_2^n$ korrigiert genau dann e Fehler, falls

$$
\begin{aligned}
d(c,x) \leq e &\implies c = x && (\forall c \in C,\ x \in \mathbb{F}_2^n) \\
d(c,x), d(c',x) \leq e &\implies c = x = c' && (\forall c, c' \in C,\ x \in \mathbb{F}_2^n) \\
d(c,c') \leq 2e &\implies c = c' && (\forall c, c' \in C)
\end{aligned}
$$

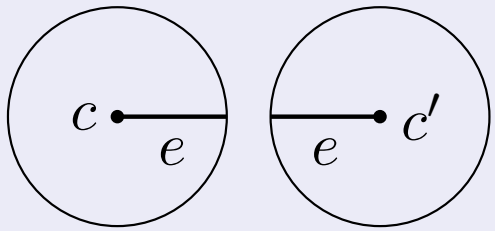


Fehlerschranken

Beweis.

2. $C \leq \mathbb{F}_2^n$ korrigiert genau dann e Fehler, falls

$$\begin{aligned} d(c,x) \leq e &\implies c = x && (\forall c \in C,\ x \in \mathbb{F}_2^n) \\ d(c,x), d(c',x) \leq e &\implies c = x = c' && (\forall c, c' \in C,\ x \in \mathbb{F}_2^n) \\ d(c,c') \leq 2e &\implies c = c' && (\forall c, c' \in C) \\ w(C) &> 2e. \end{aligned}$$



Ziele der Codierungstheorie

- Konstruiere (n, k, w)-Codes mit möglichst großen $\frac{k}{n}$ und w!
- Entwickle effizienten Decodier-Algorithmus!

Ziele der Codierungstheorie

- Konstruiere (n, k, w)-Codes mit möglichst großen $\frac{k}{n}$ und w!
- Entwickle effizienten Decodier-Algorithmus!

Satz (Shannons Hauptsatz)

Es gibt Codes mit „großer“ Rate, sodass die Wahrscheinlichkeit einer fehlerhaften Decodierung beliebig klein ist.

Ziele der Codierungstheorie

- Konstruiere (n, k, w)-Codes mit möglichst großen $\frac{k}{n}$ und w!
- Entwickle effizienten Decodier-Algorithmus!

Satz (Shannons Hauptsatz)

Es gibt Codes mit „großer“ Rate, sodass die Wahrscheinlichkeit einer fehlerhaften Decodierung beliebig klein ist.

Beweis benutzt Statistik und ist nicht konstruktiv ☹

Beispiele

- Wiederholungscode: $(n, k, w) = (n, 1, n)$, großes Gewicht, aber kleine Rate $\frac{1}{n}$.

Beispiele

- Wiederholungscode: $(n, k, w) = (n, 1, n)$, großes Gewicht, aber kleine Rate $\frac{1}{n}$.
- ASCII-Code: $(n, k, w) = (n, n-1, 2)$, große Rate $\frac{n-1}{n}$, aber kleines Gewicht.

Beispiele

- Wiederholungscode: $(n, k, w) = (n, 1, n)$, großes Gewicht, aber kleine Rate $\frac{1}{n}$.
- ASCII-Code: $(n, k, w) = (n, n-1, 2)$, große Rate $\frac{n-1}{n}$, aber kleines Gewicht.
- Mittelweg?

Matrizen

- Jeder lineare Code $C \leq \mathbb{F}_2^n$ lässt sich durch eine Erzeugermatrix $G \in \mathbb{F}_2^{k \times n}$ mit

$$C = \{xG : x \in \mathbb{F}_2^k\}$$

beschreiben.

Matrizen

- Jeder lineare Code $C \leq \mathbb{F}_2^n$ lässt sich durch eine Erzeugermatrix $G \in \mathbb{F}_2^{k \times n}$ mit

$$C = \{xG : x \in \mathbb{F}_2^k\}$$

  beschreiben.
- Setzt man $S := \mathbb{F}_2^k$, so ist die Codierung nun die lineare Abbildung $\gamma : S \to C$, $x \mapsto x \cdot G$.

Matrizen

- Jeder lineare Code $C \leq \mathbb{F}_2^n$ lässt sich durch eine Erzeugermatrix $G \in \mathbb{F}_2^{k \times n}$ mit

$$C = \{xG : x \in \mathbb{F}_2^k\}$$

beschreiben.

- Setzt man $S := \mathbb{F}_2^k$, so ist die Codierung nun die lineare Abbildung $\gamma : S \to C$, $x \mapsto x \cdot G$.
- Ebenso lässt sich C durch eine Kontrollmatrix $H \in \mathbb{F}_2^{(n-k) \times n}$ mit

$$C = \{x \in \mathbb{F}_2^n : Hx = 0\}$$

beschreiben.

Matrizen

- Jeder lineare Code $C \leq \mathbb{F}_2^n$ lässt sich durch eine Erzeugermatrix $G \in \mathbb{F}_2^{k \times n}$ mit

$$C = \{xG : x \in \mathbb{F}_2^k\}$$

beschreiben.

- Setzt man $S := \mathbb{F}_2^k$, so ist die Codierung nun die lineare Abbildung $\gamma : S \to C$, $x \mapsto x \cdot G$.
- Ebenso lässt sich C durch eine Kontrollmatrix $H \in \mathbb{F}_2^{(n-k) \times n}$ mit

$$C = \{x \in \mathbb{F}_2^n : Hx = 0\}$$

beschreiben.

- Dann ist $x \in C \iff Hx = 0$ und $w(C)$ die minimale Anzahl linear abhängiger Spalten von H.

Hamming-Codes

Definition

Sei $m \geq 2$, $n := 2^m - 1$ und $\mathbb{F}_2^m \setminus \{0\} = \{v_1, \ldots, v_n\}$. Der Code H_n mit Kontrollmatrix $H = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \in \mathbb{F}_2^{m \times n}$ heißt Hamming-Code der Länge n.

Hamming-Codes

Definition

Sei $m \geq 2$, $n := 2^m - 1$ und $\mathbb{F}_2^m \setminus \{0\} = \{v_1, \ldots, v_n\}$. Der Code H_n mit Kontrollmatrix $H = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \in \mathbb{F}_2^{m \times n}$ heißt Hamming-Code der Länge n.

Satz

H_n ist ein $(n, n-m, 3)$-Code.

Hamming-Codes

Beweis.

Nach Konstruktion ist H_n die Lösungsmenge des homogenen Gleichungssystems $Hx = 0$.

Hamming-Codes

Beweis.

Nach Konstruktion ist H_n die Lösungsmenge des homogenen Gleichungssystems $Hx = 0$. Da die Spalten von H die Standardbasis von $\mathbb{F}_2^m$ umfassen, gilt

$$\dim C = n - \text{Rang}(H) = n - m \qquad \text{(Satz 6.6)}.$$

Hamming-Codes

Beweis.

Nach Konstruktion ist H_n die Lösungsmenge des homogenen Gleichungssystems $Hx = 0$. Da die Spalten von H die Standardbasis von $\mathbb{F}_2^m$ umfassen, gilt

$$\dim C = n - \text{Rang}(H) = n - m \qquad \text{(Satz 6.6)}.$$

Je zwei Spalten von H sind als verschiedene Vektoren über $\mathbb{F}_2$ automatisch linear unabhängig.

Hamming-Codes

Beweis.

Nach Konstruktion ist H_n die Lösungsmenge des homogenen Gleichungssystems $Hx = 0$. Da die Spalten von H die Standardbasis von $\mathbb{F}_2^m$ umfassen, gilt

$$\dim C = n - \text{Rang}(H) = n - m \qquad \text{(Satz 6.6)}.$$

Je zwei Spalten von H sind als verschiedene Vektoren über $\mathbb{F}_2$ automatisch linear unabhängig. Dies zeigt $w(H_n) \geq 3$.

Hamming-Codes

Beweis.

Nach Konstruktion ist H_n die Lösungsmenge des homogenen Gleichungssystems $Hx = 0$. Da die Spalten von H die Standardbasis von $\mathbb{F}_2^m$ umfassen, gilt

$$\dim C = n - \text{Rang}(H) = n - m \qquad \text{(Satz 6.6)}.$$

Je zwei Spalten von H sind als verschiedene Vektoren über $\mathbb{F}_2$ automatisch linear unabhängig. Dies zeigt $w(H_n) \geq 3$. Andererseits besitzt H die drei linear abhängigen Spalten

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hamming-Codes

Beweis.

Nach Konstruktion ist H_n die Lösungsmenge des homogenen Gleichungssystems $Hx = 0$. Da die Spalten von H die Standardbasis von $\mathbb{F}_2^m$ umfassen, gilt

$$\dim C = n - \text{Rang}(H) = n - m \qquad \text{(Satz 6.6)}.$$

Je zwei Spalten von H sind als verschiedene Vektoren über $\mathbb{F}_2$ automatisch linear unabhängig. Dies zeigt $w(H_n) \geq 3$. Andererseits besitzt H die drei linear abhängigen Spalten

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Es folgt $w(H_n) \leq 3$. □

Beispiel H_7

Für $m = 3$ ist $n = 2^3 - 1 = 7$ und

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

Beispiel H_7

Für $m = 3$ ist $n = 2^3 - 1 = 7$ und

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

Nach Satz 6.15 ist

$$H_7 = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Vergleich für $n = 7$

Code	(n, k, w)
Wiederholungscode	$(7, 1, 7)$
Hamming-Code	$(7, 4, 3)$
ASCII-Code	$(7, 6, 2)$

Vergleich für $n = 7$

Code	(n, k, w)
Wiederholungscode	$(7, 1, 7)$
Hamming-Code	$(7, 4, 3)$
ASCII-Code	$(7, 6, 2)$

Geht es besser?

Parameter-Schranken

Satz

Für jeden linearen (n, k, w)-Code C mit $w > 2e$ gilt

(1) $$w \leq n - k + 1$$ (*Singleton-Schranke*).

(2) $$\sum_{i=1}^{e} \binom{n}{i} \leq 2^{n-k}$$ (*Hamming-Schranke*).

Parameter-Schranken

Satz

Für jeden linearen (n, k, w)-Code C mit $w > 2e$ gilt

(1) $$w \leq n - k + 1$$ (*Singleton-Schranke*).

(2) $$\sum_{i=1}^{e} \binom{n}{i} \leq 2^{n-k}$$ (*Hamming-Schranke*).

Beweis.

1. Die lineare Abbildung (Projektion)

$$\rho : C \to \mathbb{F}_2^{n-w+1},$$
$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{n-w+1})$$

ist injektiv, denn für $x \in \mathrm{Ker}(\rho)$ gilt $d(x, 0) \leq w - 1$, also $x = 0$.

Parameter-Schranken

Satz

Für jeden linearen (n, k, w)-Code C mit $w > 2e$ gilt

(1) $$w \leq n - k + 1$$ (*Singleton-Schranke*).

(2) $$\sum_{i=1}^{e} \binom{n}{i} \leq 2^{n-k}$$ (*Hamming-Schranke*).

Beweis.

1. Die lineare Abbildung (Projektion)

$$\rho : C \to \mathbb{F}_2^{n-w+1},$$
$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{n-w+1})$$

ist injektiv, denn für $x \in \mathrm{Ker}(\rho)$ gilt $d(x, 0) \leq w - 1$, also $x = 0$.
Dies zeigt $2^k = |C| = |\rho(C)| \leq 2^{n-w+1}$ und $k \leq n - w + 1$.

Parameter-Schranken

Beweis.

2. Für $c \in C$ sei $K_e(c) := \{x \in \mathbb{F}_2^n : d(c,x) \leq e\}$ die „Kugel“ mit Radius e und Mittelpunkt c.

Parameter-Schranken

Beweis.

2. Für $c \in C$ sei $K_e(c) := \{x \in \mathbb{F}_2^n : d(c,x) \leq e\}$ die „Kugel“ mit Radius e und Mittelpunkt c. Für $d(c,x) = i$ unterscheidet sich x von c an genau i Positionen.

Parameter-Schranken

Beweis.

2. Für $c \in C$ sei $K_e(c) := \{x \in \mathbb{F}_2^n : d(c,x) \leq e\}$ die „Kugel“ mit Radius e und Mittelpunkt c. Für $d(c,x) = i$ unterscheidet sich x von c an genau i Positionen. Es gibt $\binom{n}{i}$ Möglichkeiten, diese Positionen zu wählen.

Parameter-Schranken

Beweis.

2. Für $c \in C$ sei $K_e(c) := \{x \in \mathbb{F}_2^n : d(c,x) \leq e\}$ die „Kugel“ mit Radius e und Mittelpunkt c. Für $d(c,x) = i$ unterscheidet sich x von c an genau i Positionen. Es gibt $\binom{n}{i}$ Möglichkeiten, diese Positionen zu wählen. Dies zeigt

$$|K_e(c)| = \sum_{i=0}^{e} \binom{n}{i}.$$

Parameter-Schranken

Beweis.

2. Für $c \in C$ sei $K_e(c) := \{x \in \mathbb{F}_2^n : d(c,x) \leq e\}$ die „Kugel“ mit Radius e und Mittelpunkt c. Für $d(c,x) = i$ unterscheidet sich x von c an genau i Positionen. Es gibt $\binom{n}{i}$ Möglichkeiten, diese Positionen zu wählen. Dies zeigt

$$|K_e(c)| = \sum_{i=0}^{e} \binom{n}{i}.$$

Aus $w(C) > 2e$ folgt $K_e(c) \cap K_e(c') = \varnothing$, falls $c \neq c'$.

Parameter-Schranken

Beweis.

2. Für $c \in C$ sei $K_e(c) := \{x \in \mathbb{F}_2^n : d(c,x) \leq e\}$ die „Kugel“ mit Radius e und Mittelpunkt c. Für $d(c,x) = i$ unterscheidet sich x von c an genau i Positionen. Es gibt $\binom{n}{i}$ Möglichkeiten, diese Positionen zu wählen. Dies zeigt

$$|K_e(c)| = \sum_{i=0}^{e} \binom{n}{i}.$$

Aus $w(C) > 2e$ folgt $K_e(c) \cap K_e(c') = \varnothing$, falls $c \neq c'$. Daher ist

$$2^k \sum_{i=0}^{e} \binom{n}{i} = |C| \sum_{i=0}^{e} \binom{n}{i} = \sum_{c \in C} |K_c(e)| = \left| \bigcup_{c \in C} K_e(c) \right| \leq |\mathbb{F}_2^n| = 2^n.$$

Perfekte Codes

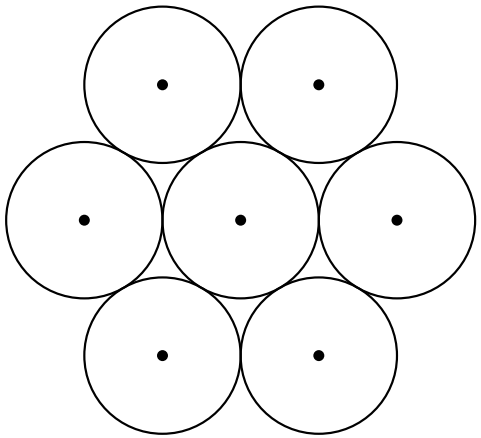
- Gilt Gleichheit in der Hamming-Schranke, so heißt C perfekt.

Perfekte Codes

- Gilt Gleichheit in der Hamming-Schranke, so heißt C perfekt.
- Dann ist $\mathbb{F}_2^n$ die disjunkte Vereinigung von Kugeln mit Radius e um Codewörter.

Perfekte Codes

- Gilt Gleichheit in der Hamming-Schranke, so heißt C perfekt.
- Dann ist $\mathbb{F}_2^n$ die disjunkte Vereinigung von Kugeln mit Radius e um Codewörter.
- Somit lässt sich jedes $x \in \mathbb{F}_2^n$ eindeutig decodieren.



Beispiele perfekter Codes

- Die Hamming-Codes sind perfekt mit $e = 1$, denn

$$\sum_{i=0}^{1} \binom{n}{i} = 1 + n = 2^m = 2^{n-k}.$$

Beispiele perfekter Codes

- Die Hamming-Codes sind perfekt mit $e = 1$, denn

$$\sum_{i=0}^{1} \binom{n}{i} = 1 + n = 2^m = 2^{n-k}.$$

- Unter allen Codes, die einen Fehler korrigieren können, haben die Hamming-Codes die größte Rate.

Beispiele perfekter Codes

- Die Hamming-Codes sind perfekt mit $e = 1$, denn

$$\sum_{i=0}^{1} \binom{n}{i} = 1 + n = 2^m = 2^{n-k}.$$

- Unter allen Codes, die einen Fehler korrigieren können, haben die Hamming-Codes die größte Rate.
- Nachteil: Existieren nur für $n = 2^m - 1$.

Beispiele perfekter Codes

- Die Hamming-Codes sind perfekt mit $e = 1$, denn

$$\sum_{i=0}^{1} \binom{n}{i} = 1 + n = 2^m = 2^{n-k}.$$

- Unter allen Codes, die einen Fehler korrigieren können, haben die Hamming-Codes die größte Rate.
- Nachteil: Existieren nur für $n = 2^m - 1$.
- Wiederholungscode mit $n = 2e + 1 = w$ ist ebenfalls perfekt, denn jedes $x \in \mathbb{F}_2^n$ unterscheidet sich an höchstens e Positionen von genau einem der Codewörter $(0, \ldots, 0)$ oder $(1, \ldots, 1)$.

Beispiele perfekter Codes

- Die Hamming-Codes sind perfekt mit $e = 1$, denn

$$\sum_{i=0}^{1} \binom{n}{i} = 1 + n = 2^m = 2^{n-k}.$$

- Unter allen Codes, die einen Fehler korrigieren können, haben die Hamming-Codes die größte Rate.
- Nachteil: Existieren nur für $n = 2^m - 1$.
- Wiederholungscode mit $n = 2e + 1 = w$ ist ebenfalls perfekt, denn jedes $x \in \mathbb{F}_2^n$ unterscheidet sich an höchstens e Positionen von genau einem der Codewörter $(0, \ldots, 0)$ oder $(1, \ldots, 1)$.
- Außer diesen gibt es im Wesentlichen nur einen weiteren linearen perfekten Code (über $\mathbb{F}_2$):

Golay-Code

Der $(23, 12, 7)$-Golay-Code ist perfekt mit $e = 3$ und Erzeugermatrix:

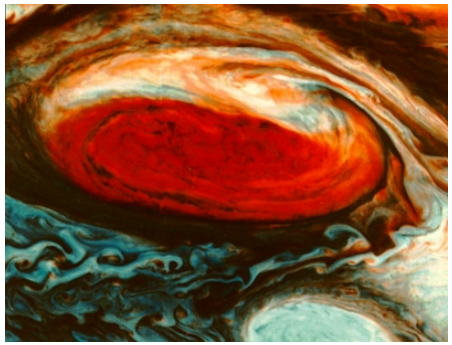
$$G = \begin{pmatrix}
1 & . & 1 & . & 1 & 1 & 1 & . & . & . & 1 & 1 & . & . & . & . & . & . & . & . & . & . & . \\
. & 1 & . & 1 & . & 1 & 1 & 1 & . & . & . & 1 & 1 & . & . & . & . & . & . & . & . & . & . \\
. & . & 1 & . & 1 & . & 1 & 1 & 1 & . & . & . & 1 & 1 & . & . & . & . & . & . & . & . & . \\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

Golay-Code

Der $(23, 12, 7)$-Golay-Code ist perfekt mit $e = 3$ und Erzeugermatrix:

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

Damit haben die Voyager-Sonden Bilder vom Jupiter zur Erde gesendet:



MDS-Codes

- Gilt Gleichheit in der Singleton-Schranke, so heißt C MDS-Code (maximum distance separable).

MDS-Codes

- Gilt Gleichheit in der Singleton-Schranke, so heißt C MDS-Code (maximum distance separable).
- Wiederholungscodes sind MDS-Codes.

MDS-Codes

- Gilt Gleichheit in der Singleton-Schranke, so heißt C MDS-Code (maximum distance separable).
- Wiederholungscodes sind MDS-Codes.
- Ein weiteres Beispiel ist der Reed-Solomon-Code, der unter anderen bei QR-Codes eingesetzt wird:

MDS-Codes

- Gilt Gleichheit in der Singleton-Schranke, so heißt C MDS-Code (maximum distance separable).
- Wiederholungscodes sind MDS-Codes.
- Ein weiteres Beispiel ist der Reed-Solomon-Code, der unter anderen bei QR-Codes eingesetzt wird:

- Datenbank von optimalen linearen Codes:

`http://www.codetables.de/`