3. Grundlagen der Quanteninformationsübertragung

3.1. Das klass. Maß der Information, klass. Datenkompression

Zunächst müssen wir uns verständigen, was wir unter Information verstehen wollen.

Beispiel: Würfel
- Wenn ein Würfelspieler Ihnen sagt, er habe eine "5" gewürfelt, so hat diese Aussage für Sie nur Information, wenn Sie nicht beim Würfeln zugesehen haben
- Wenn der Würfelspieler sagt er habe eine Zahl zwischen 1 und 6 gewürfelt hat das für Sie überhaupt keine Information

> Informationsgehalt einer Aussage = Maß der a priori Ignoranz

Sei A ein zufälliges Ereignis mit Wahrscheinlichkeit $p(A)$. $\bar{A}$ heißt das Gegenteil von A mit $p(\bar{A}) = 1 - p(A)$. Eine Menge $\{A_1, \dots A_n\}$ von Ereignissen heißt vollständig falls $\sum_{i=1}^{n} p(A_i) = 1$.

Def: sei $\{A_1, \dots, A_n\}$ eine vollständige Menge von Ereignissen. Die A_i heißen dann Elementarereignisse falls

(1) $A_i \Rightarrow \bar{A}_j \quad \forall\, j \neq i$

Die Menge alla Elementarereignisse $\Omega = \{A_1, \dots, A_n\}$ heißt Ereignisraum. Jedes Ereignis A ist Teilmenge von Ω.

Prinzip der maximalen Unbestimmtheit

Wenn über ein Ereignis A nichts bekannt ist wird jedem Elementarereignis die a priori Wahrscheinlichkeit

(2) $$p_i = 1/n$$

zugeordnet

Nun wollen wir ein quantitatives Maß S der Unbestimmtheit konstruieren. Es soll folgenden Kriterien genügen

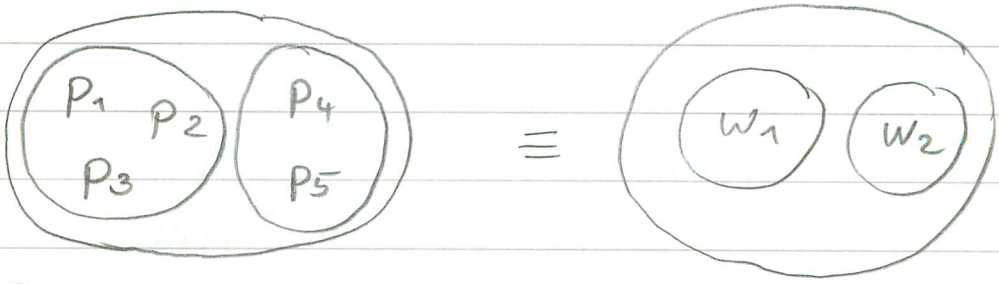
Annahme: n diskrete Ereignisse

(i) S ist Funktion der Wahrscheinlichkeiten p_i
$S = S(p_1, \dots p_n)$ und stetig in p_i

(ii) $S = 0$ für sicheres Ereignis (ein $p_i = 1$)

(iii) $S = \max$ für maximale Unbestimmtheit ($p_i = 1/n$) und monoton wachsend in n
(Je größer der Ereignisraum über den ich nichts weiß umso mehr Information steckt in der Zufallsvariablen)

(iv) H soll folgende Kompositionsregel erfüllen



Gruppe aus 3 Ereignissen — Gruppe aus 2 Ereignissen

$w_1 = p_1 + p_2 + p_3$

$w_2 = p_4 + p_5$

(3) $$S(p_1,\dots,p_5) = S(w_1,w_2) + w_1 S\left(\frac{p_1}{w_1},\frac{p_2}{w_1},\frac{p_3}{w_1}\right) + w_2 S\left(\frac{p_4}{w_2},\frac{p_5}{w_2}\right)$$

Gesamtinformation = Information "w_1 oder w_2"
+ gewichtete Information
"falls w_1: p_1 oder p_2 oder p_3"
"falls w_2: p_4 oder p_5"

Die Forderungen (i) – (iv) legen H eindeutig fest. Man findet die sog. Shannon Entropie

(4) $$\boxed{S(p_1,\dots,p_n) = -a \sum_{i=1}^{n} p_i \ln p_i} \qquad a>0$$

Beweis: (i) ✓

(ii) ✓ da aus $p_i = 1$ $p_j = 0$ $\forall j \neq i$
$\Rightarrow S = 0$

(iii) Um (iii) zu zeigen beweisen wir zunächst folgende Aussage: Seien $(p_1, \dots, p_n)$ und $(\bar{p}_1, \dots, \bar{p}_n)$ zwei Wahrscheinlichkeitsverteilungen für n Elementarereignisse. Dann gilt

$$\sum_{i=1}^{n} \bar{p}_i \ln \frac{\bar{p}_i}{p_i} \geq 0 \tag{5}$$

Gleichheit gilt für $\bar{p}_i \equiv p_i$

Bew: $\ln x \leq x - 1 \quad \forall x \geq 0$

$$\Rightarrow \ln \tfrac{1}{x} = -\ln x \geq 1 - x$$

$$\Rightarrow \ln \frac{\bar{p}_i}{p_i} \geq 1 - \frac{p_i}{\bar{p}_i}$$

$$\sum_i \bar{p}_i \ln \frac{\bar{p}_i}{p_i} \geq \sum_i \bar{p}_i - \sum_i p_i = 0 \qquad \square$$

Aus (5) folgt mit $p_i \equiv \frac{1}{n}$

$$\sum_{i=1}^{n} \bar{p}_i (\ln \bar{p}_i - \ln p_i) \geq 0$$

$$S(\bar{p}_1, \bar{p}_2 \cdots \bar{p}_n) = -a \sum_{i=1}^{n} \bar{p}_i \ln \bar{p}_i \leq -a \sum_{i=1}^{n} \bar{p}_i \ln \frac{1}{n}$$

$$= a \ln n \qquad \square$$

(iv) Es sei $\omega_1 = p_1 + \cdots + p_m$

$\omega_2 = p_{m+1} + \cdots + p_n$

$$S(w_1, w_2) = -a\, w_1 \ln w_1 - a\, w_2 \ln w_2$$

$$w_1\, S\left(\frac{p_1}{w_1}, \dots, \frac{p_m}{w_m}\right) = -a\, w_1 \sum_{i=1}^{m} \frac{p_i}{w_1} \ln\left(\frac{p_i}{w_i}\right)$$

$$= -a \sum_{i=1}^{m} p_i \ln p_i + a\, w_1 \ln w_1$$

$$\Rightarrow S(w_1, w_2) + w_1\, S\left(\frac{p_1}{w_1}, \dots, \frac{p_m}{w_1}\right) + w_2\, S\left(\frac{p_{m+1}}{w_2}, \dots, \frac{p_n}{w_2}\right)$$

$$= -a\, w_1 \ln w_1 - a\, w_2 \ln w_2$$

$$-a \sum_{i=1}^{m} p_i \ln p_i + a\, w_1 \ln w_1$$

$$-a \sum_{i=m+1}^{n} p_i \ln p_i + a\, w_2 \ln w_2$$

$$= -a \sum_{i=1}^{n} p_i \ln p_i = S(p_1, \dots, p_n) \quad \square$$

Beweis Ende

Bem: Der Vorfaktor a kann nach beliebig gewählt werden

Der klassische nichttriviale Ereignisraum hat zwei Alternativen

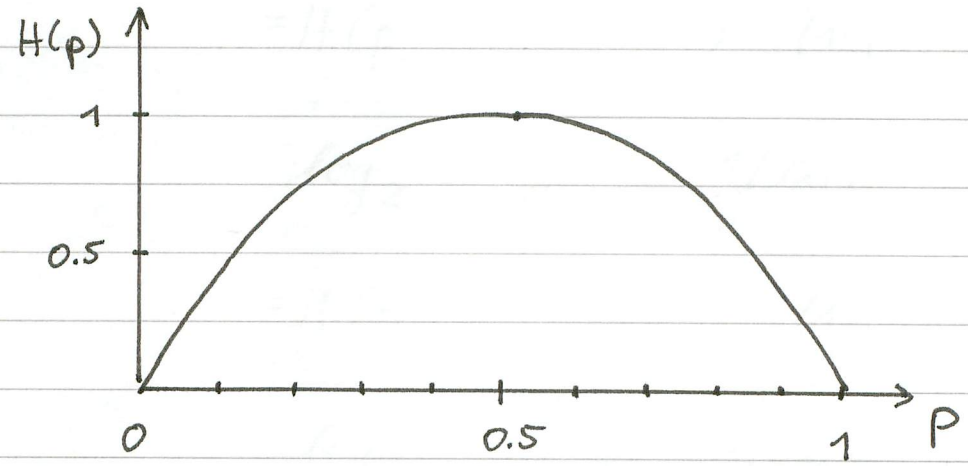
$$\{A, \bar{A}\} \hat{=} \{0, 1\}$$

Diesen Ereignisraum wollen klassisches bit oder cbit nennen.

Für ein cbit ist die Shannon Entropie gegeben durch

$$(6) \qquad \boxed{S(p, 1-p) = H(p) = -p \log p - (1-p) \log (1-p)}$$

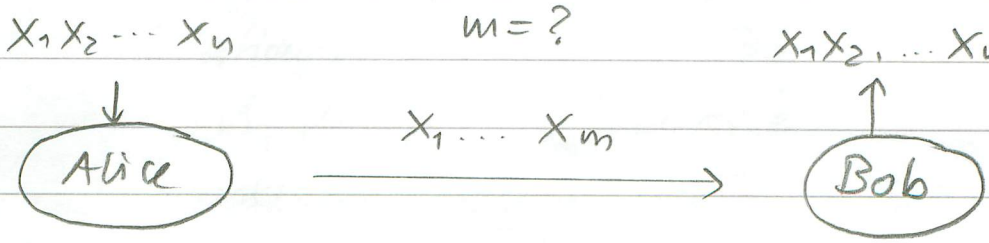
wobei $\log = \log_2$ und $a = 1/\ln 2$ gewählt wurde



Klassische Information kann durch Reihen von cbits, sog. <u>bit-Wörter</u> übertragen werden

$$(7) \qquad x_1 x_2 \cdots x_n \qquad x_i \in \{0,1\}$$

Die Bedeutung der Shannon Entropie als Information was? Stärkeres wir nun an einem zentralen Problem der Informationsübertragung: Gegeben ein bitWort der Länge n mit $n \to \infty$. Wieviel cbits werden benötigt um das bitWort von A nach B zu übertragen



Die Antwort wird durch

Shannons noiseless coding Theorem

Zur Übertragung eines n bit Wortes genügen für $n \to \infty$ n $\cdot$ H(p) bits, wobei p die Wahrscheinlichkeit des Auftretens von "0" oder "1" in dem n bit Wort ist

Darüberhinaus gilt: Es gibt immer einen eindeutig decodierbaren Code aus $1 + n H(p)$ bits. Shannons Theorem gibt aber keine Konstruktionsvorschrift.
Falls "0" und "1" gleichwahrscheinlich $H(p) = \log 2$.

Bspl:

0000 0001 0100 0101 0000 0110 0100 1000 0000 0011
1000 0101 1100 0001 0100 0010 0000 0101 0001 0110
0110 0001 0100 1100 0000 0001 0100 0101 0000 0101
1001 0100 0010 0000 0101 0001 0000 0011 0001 0100
0100 1100 0001 0000 0011 0001 1100 0100 0100 0010
0100 0101 0000 0101 0100 0010 0000 0101 0001 0100
1000 0101 1100 1101 0100 0010 0000 0101 0001 0100
0000 0101 0001 0100 1000 0101 1100 0001 1100 1000

$$p(0) = \frac{225}{320} = 0{,}703 > p(1) = \frac{95}{320} = 0{,}297$$

Original	Code
0000	10
0001	000
0010	001
0100	010
1000	011
0011	1000
0101	1001
1001	1011
0110	1010
1010	1100
1100	1101
0111	11000
1011	11111
1101	11110
1110	11101
1111	111001

häufige Wörter erhalten kurzen Code

codierter Text:

10 000 010 1001 10 1010 010 011 10 1000

011 1001 1101 000 010 001 10 1001 000 1010

1010 000 010 1101 10 000 010 1001 10 1001

1011 010 001 10 1001 000 10 1000 000 010

010 1101 000 10 1000 000 1101 010 010 001

010 1001 10 1001 010 001 10 1001 000 010

011 1001 1101 11110 010 001 10 1001 000 010

10 1001 000 010 011 1001 1101 000 1101 011

Die Verkürzung ist deutlich zu erkennen.

Der in diesem Beispiel verwendete code ist illustrativ aber ein bißchen gemogelt da Leerzeichen im Text beibehalten wurden. Ein code der für Wörter ohne Leerzeichen anwendbar ist, ist der sog. Huffmann code

0000	10
0001	000
0010	001
0100	010
1000	011
0011	11000
0101	11001
0110	11010
1001	11011
1010	11100
1100	11101
0111	1111000
1011	111111
1101	111110
1110	111101
1111	1111001

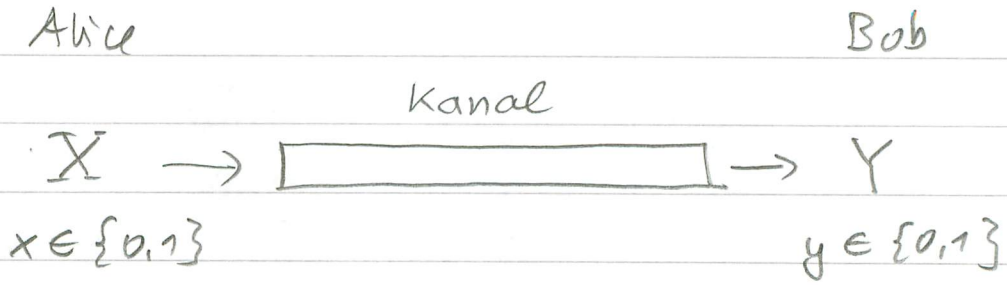
für $p(0) = 3/4$
$p(1) = 1/4$

Bem.: Huffmann code praktisch nicht verwendet.

3.2. Kommunikationskanäle

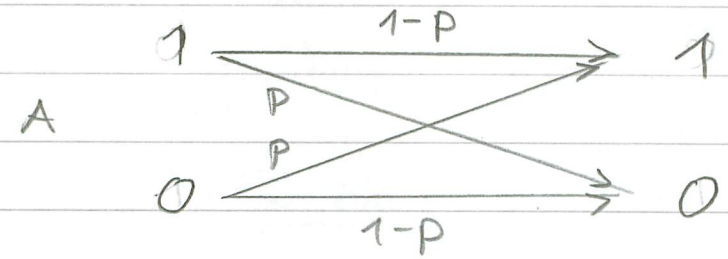
Shannons noiseless coding Theorem macht eine Aussage über die Informationsübertragung durch ideale d.h. verlustfreie Kanäle. Kommunikationskanäle sind aber i.A. verlustbehaftet und führen zu Rauschen.

Betrachten binären Übertragungskanal (Alice kann an Bob "0" oder "1" senden)



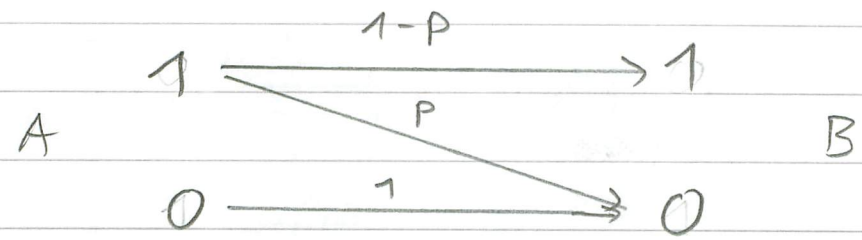
Eine wichtige Klasse von Kanälen sind gedächtnisfreie Kanäle. Bei ihnen hängt die Wahrscheinlichkeit p eines Übertragungsfehlers nicht von der Vorgeschichte ab.

Symmetrischer binärer Kanal



gleiche Fehlerwahrscheinlichkeit
0 → 1
1 → 0

binärer Zerfallskanal



Den Kanälen kann eine Übertragungsmatrix P_{ij} zugeordnet werden.

symmetrisch binärer Kanal Zufallskanal

(8) $\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$ $\begin{pmatrix} 1 & p \\ 0 & 1-p \end{pmatrix}$

Eine wichtige Größe zur Charakterisierung eines Kanals ist die gegenseitige Information

(9) $$I(X:Y) \equiv S(X) - S(X|Y)$$
$$= S(Y) - S(Y|X)$$

wobei

(10) $$S(Y|X) = -\sum_i p(x_i) \sum_j p(y_j|x_i) \log p(y_j|x_i)$$
$$= \sum_{ij} p(x_i, y_j) \log p(y_j|x_i)$$

die bedingte Entropie ist. Hierbei sind

$p(x_i, y_j)$ die Verbundwahrscheinlichkeit für das Auftreten von x_i und y_j

$p(y_j|x_i)$ die konditionelle Wahrscheinlichkeit des Auftretens von y_j falls x_i eingetreten ist

Es gilt

(11) $$p(x,y) = p(x)\, p(y|x)$$

$$S(Y|X) = ? = -\sum_x p(x) \sum_y p(y|x) \log p(y|x)$$

simetrični binarni kanal

$$p(y=0|x=0) = p(y=1|x=1) = 1-p_e$$

$$p(y=0|x=1) = p(y=1|x=0) = p_e$$

$$S(Y|X) = -p(x=0)\left[p(y=0|x=0)\log p(y=0|x=0) + p(y=1|x=0)\log p(y=1|x=0)\right]$$

$$- p(x=1)\left[p(y=0|x=1)\log p(y=0|x=1) + p(y=1|x=1)\log p(y=1|x=1)\right]$$

$$= -p(x=0)\left[(1-p_e)\log(1-p_e) + p_e \log p_e\right]$$

$$-p(x=1)\left[p_e \log p_e + (1-p_e)\log(1-p_e)\right]$$

$$= -p_e \log p_e - (1-p_e)\log(1-p_e) = H(p_e)$$

Für unabhängige Ereignisse ist

$$p(y|x) = p(y) \Rightarrow \quad p(x,y) = p(x)\,p(y) \tag{12}$$

Die gegenseitige Information kann geschrieben werden als

$$I(X:Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)\,p(y)} \tag{13}$$

Falls X und Y unabhängig folgt aus (10)

$$S(X|Y) = S(X) \Rightarrow I(X:Y) = 0 \tag{14}$$

und somit

Für einen binären symmetrischen Kanal gilt

$$\boxed{I(X:Y) = H(p(y)) - H(P_{error})} \tag{15}$$

Wir definieren jetzt die Kanalkapazität

$$C \equiv \max_{\{p(x)\}} I(X:Y) \tag{16}$$

Für den binären symmetrischen Kanal erhalten wir somit

$$\boxed{C = 1 - H(P_{error})} \tag{17}$$

Nun wollen wir uns die Frage stellen wieviel cbits benötigt werden um ein n bit Wort über einen verrauschten (fehlerhaften) Kanal zu übertragen. Dazu führen wir zunächst die Übertragungsrate R ein:

$$R = \frac{\#\text{ Nachrichtenbits}}{\#\text{ Benutzung des Kanals}} \qquad 0 \leq R \leq 1 \tag{18}$$

Nehmen wir an der Kanal hat eine Fehlerwahrscheinlichkeit $p_b \leq 1$. Senden wir die Nachricht mit $R=1$, ist die Fehlerwahrscheinlichkeit im erhaltenen Signal $e = p_b$.
Um die Fehlerwahrscheinlichkeit zu verringern benutzen wir die Kodierung

$$0 \to 000 \qquad 1 \to 111 \qquad \text{(Repitionscode)} \tag{19}$$

D.h. die Übertragungsrate ist $R = \frac{1}{3}$. Wenn nur eines der 3 cbits falsch übertragen wird, kann die Nachricht immer noch exakt rekonstruiert werden. D.h. nur wenn mindestens 2 cbits falsch sind tritt ein Übertragungsfehler auf. Die Wahrscheinlichkeit dafür ist

$$e = p_b^3 + 3p_b^2(1-p_b) < p_b \tag{20}$$

m-facher Repitionscode

- $e \xrightarrow[m\to\infty]{} 0$ 
- $R \xrightarrow[m\to\infty]{} 0$ 

Läßt sich Information über einen verrauschten Kanal mit beliebig kleiner Fehlerwahrscheinlichkeit nur mit $R \to 0$ realisieren? Nein!

Shannons noisy coding theorem (1948)

Für einen Kanal der Kapazität C existiert ein Code der eine Datenübertragung mit beliebig kleiner Fehlerwahrscheinlichkeit und $0 < R < C$ erlaubt.

Fehlerfreie Datenübertragung über verrauschte Kanäle mit nichtverschwindender Übertragungsrate ist möglich!

3.3. Klassische Fehlerkorrektur

In den letzten 50 Jahren ist ein ganzes Theoriegebäude fehlerkorrigierender (klassischer) Codes entwickelt worden. Ein wichtiger Vertreter sind binäre lineare Codes, die jetzt kurz diskutiert werden sollen.

binärer Code: k Bits sind in $n > k$ Bits kodiert

Anzahl der Codewörter 2^k
Anzahl der Wörter insges. 2^n

linearer binärer Code:

Menge G von Codewörtern die bezüglich bitweiser Addition modulo 2 ein Galoisfeld (binärer lin. Raum)

Bspl: $\{(0,1,1), (1,0,1), (1,1,0), (0,0,0)\}$

$(0,1,1) \oplus (1,0,1) = (1,1,0)$ $\quad n = 3$

$(0,1,1) \oplus (1,1,0) = (1,0,1)$ $\quad k = 2$

$(1,0,1) \oplus (1,1,0) = (0,1,1)$ $\quad$ Basis z.B $(0,1,1)$, $(1,0,1)$

Basis von G: $\{v_1, v_2, \dots v_k\}$ $\quad v_i$ Codewörter aus $n \geq k$ bit

Jedes Codewort läßt sich dann als Linearkombination der $\underline{v_i}$ schreiben

$$(21) \qquad \underline{v} = \bigoplus_{i=1}^{k} \alpha_i \underline{v_i} \qquad \alpha_i \in \{0,1\}$$

und $\oplus$ bedeutet bitweise Addition modulo 2.
$\underline{v}$ kodiert die k-bit Nachricht $(\alpha_1, \alpha_2, \dots, \alpha_k)$.
Die k Basisvektoren bilden die $k \times n$ Erzeugende

$$(22) \qquad \underline{\underline{G}} = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$$

Im obigen Beispiel gibt es zwei Basisvektoren z.B. $(0,1,1)$ und $(1,0,1)$ und somit wäre

$$\underline{\underline{G}} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \tag{23}$$

Dann kann Gl. (21) in der Form

$$\underline{v} = \underline{\alpha}\,\underline{\underline{G}} \qquad \underline{\alpha} = (\alpha_1, \alpha_2, \ldots \alpha_k) \tag{24}$$

geschrieben werden.

Alternativ zur Charakterisierung des Codes mittels der Matrix $\underline{\underline{G}}$ kann man auch die $n-k$ linearen Bedingungen

$$\underline{\underline{H}}\,\underline{v}^T = 0 \tag{25}$$

benutzen, wobei $\underline{\underline{H}}$ eine $(n-k) \times n$ Matrix ist. H heißt parity-check Matrix. Die Zeilen von $\underline{\underline{H}}$ sind $(n-k)$ linear unabhängige Vektoren die auf allen Codewörtern senkrecht stehen. Im obigen Beispiel:

Beispiel

$$\underline{\underline{H}} = (1,1,1) \tag{26}$$

Basis: $\{(0,1,1), (1,0,1)\}$

Codewörter

$\vec{\alpha} = (0,0) : (0,0,0)$ $\quad \vec{\alpha} = (1,0) : (0,1,1)$

$\vec{\alpha} = (0,1) : (1,0,1)$ $\quad \vec{\alpha} = (1,1) : (1,1,0)$

Im Falle klassischer Kanäle ist der einzig mögliche Fehler ein bit-flip. Ein bit-flip Fehler entspricht der Addition eines Fehlervektors

(27) $$\underline{v} \longrightarrow \underline{v} \oplus \underline{e}$$

wobei $\underline{e}$ Einträge "1" an den Stellen hat an denen ein bit flip stattgefunden hat, z.B.

(28) $$\underline{e} = (1, 0, 1, 0, \ldots, 0)$$

entspricht einem flip des ersten und dritten Bits.
In einem linearen binären Code können Fehler einfach durch Anwenden der parity check Matrix $\underline{\underline{H}}$ detektiert werden:

(29) $$\underline{\underline{H}}(\underline{v} \oplus \underline{e}) = \underline{\underline{H}}\,\underline{v} \oplus \underline{\underline{H}}\,\underline{e} = \underline{\underline{H}}\,\underline{e}$$

$\underline{\underline{H}}\,\underline{e}$ ist das Fehlersyndrom.

Die Menge $\mathcal{E}$ aller Fehler $\{\underline{e}_i\}$ die unterschiedliche Fehlersyndrome haben sind mittels des verwendeten Codes korrigierbar. Nach eindeutiger Detektion des Fehlers wird dieser dazu einfach noch einmal addiert

(30) $$(\underline{v} \oplus \underline{e}_i) \oplus \underline{e}_i = \underline{v}$$

Falls aber $\underline{\underline{H}}\,\underline{e}_1 = \underline{\underline{H}}\,\underline{e}_2$ könnten wir den

uis interpretieren und

$$\underline{v} \oplus \underline{e}_1 \longrightarrow (\underline{v} \oplus \underline{e}_1) \oplus \underline{e}_2 \neq \underline{v} \tag{31}$$

Der Hamming abstand eines Codes ist die minimalanzahl von bits die geändert werden müssen um ein Element des Codes in ein anderes zu überführen.

> Ein linearer Code mit Abstand $d = 2t+1$ kann t bit-flip Fehler korrigieren

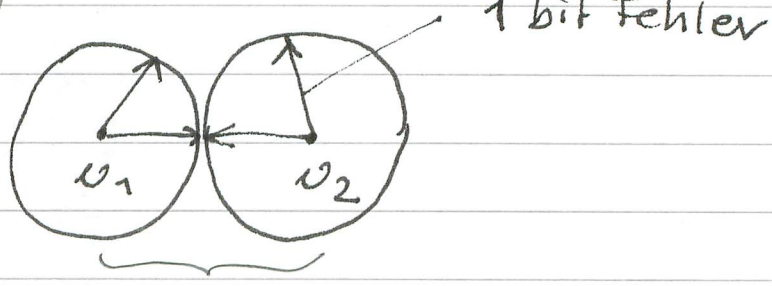
Der oben eingeführte Code hat $d=2$ und kann daher keinen Fehler korrigieren, denn

$$v_1 = (0,1,1) \oplus e_1 = (1,0,0)$$

$$\| \tag{32}$$

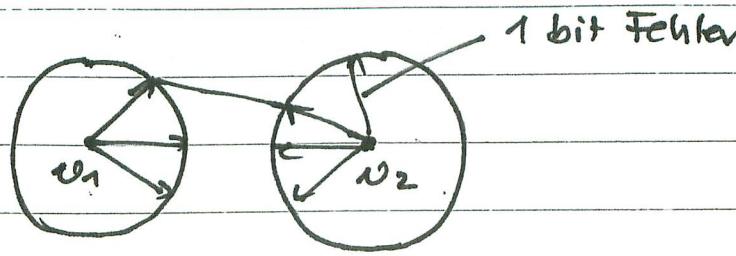
$$v_2 = (1,0,1) \oplus e_2 = (0,1,0)$$

Hier hat man



Hamming abst. $d=2$

Für $d=3$ hat man



$d = 3$

Beispiel für $[n=7, k=4, d=3]$ Code

Message	Huffman	Hamming
0000	10	0000000
0001	000	1010101
0010	001	0110011
0011	11000	1100110
0100	010	0001111
0101	11001	1011010
0110	11010	0111100
0111	1111000	1101001
1000	011	1111111
1001	11011	0101010
1010	11100	1001100
1011	111111	0011001
1100	11101	1110000
1101	111110	0100101
1110	111101	1000011
1111	1111001	0010110

Der Hamming [7,4,3] Code erlaubt Korrektur aller 1-bit Fehler.

3.4. Quanteninformation

In Abschnitt 3.1. hatten wir klassische Zufallsgrößen X betrachtet, die mit Wahrscheinlichkeit p_i die Werte (oder klassischen Zustände) x_i annehmen können. Wir hatten dann die von Shannon Entropie

$$S_{sh} = -\sum_i p_i \ln p_i \tag{33}$$

als Maß für unser Unwissen über den tatsächlichen Zustand x_i eingeführt.

Nun wollen wir die x_i durch quantenmechanische Zustände ersetzen. Zunächst wollen wir verschiedenen Werten x_i orthogonale quantenmechanische Zustände zuordnen

$$x_i \longleftrightarrow |x_i\rangle\langle x_i| \quad \text{mit} \quad \langle x_i | x_j \rangle = \delta_{ij} \tag{34}$$

die mit Wahrscheinlichkeit p_i auftreten sollen. Quantenmechanisch entspricht dies einem Dichteoperator

$$\varrho = \sum_i p_i |x_i\rangle\langle x_i| \tag{35}$$

und (33) kann geschrieben werden als

$$S_{sh} = -\mathrm{Tr}\{\varrho \ln \varrho\} = S_N \tag{36}$$

was gerade der von-Neumann Entropie entspricht.

Als quantitatives Maß für die Unbestimmtheit eines Quantenzustandes $\hat{\varrho}$ wollen wir daher die von-Neumann Entropie

(37) $$\boxed{S_N = -\mathrm{Tr}\{\hat{\varrho}\log\hat{\varrho}\}}$$

verwenden.

Eigenschaften von S_N

(i) reiner Zustand $\hat{\varrho} = |\psi\rangle\langle\psi|$ $\Rightarrow$ $\boxed{S_N = 0}$ (38)

(ii) maximaler gemischter Zustand $\hat{\varrho} = \frac{1}{N}\mathbb{1}_N$
$N = \dim(\mathcal{H})$

$\Rightarrow$ $\boxed{S_N = \log N}$ (39)

(iv) allg. $$\boxed{0 \le S_N \le \log(\dim(\mathcal{H}))}$$ (40)

(v) S_N ist konkav, d.h. $\lambda_1, \lambda_2, \dots, \lambda_N \ge 0$
und $\lambda_1 + \lambda_2 + \dots + \lambda_N = 1$

(41) $$\boxed{S_N(\lambda_1\hat{\varrho}_1 + \dots + \lambda_n\hat{\varrho}_n) \ge \lambda_1 S_N(\hat{\varrho}_1) + \dots + \lambda_n S_N(\hat{\varrho}_n)}$$

Die Eigenschaft (v) bringt folgenden Sachverhalt zum Ausdruck: Wenn wir wissen, daß ein Quanten-

Zustand $\varrho = \lambda_1 \varrho_1 + \cdots + \lambda_N \varrho_N$ aus einem Gemisch der Zustände $\varrho_1, \ldots, \varrho_N$ mit Gewichten (Wahrscheinlichkeiten) λ_i konstruiert wurde besitzen wir i.A. mehr apriori Information über das System als wenn uns nur ϱ bekannt ist. Das liegt daran, daß dasselbe ϱ auf verschiedene Arten zusammengesetzt worden kann.

(vi) unitäre Evolutionen lassen S_N unverändert

$$S_N(U\varrho U^{-1}) = S_N(\varrho) \tag{42}$$

(vii) es gilt die Subadditivität

$$S_N(\varrho_{AB}) \leq S_N(\varrho_A) + S_N(\varrho_B) \tag{43}$$

wobei $\varrho_A = \mathrm{Tr}_B\{\varrho_{AB}\}$ und $\varrho_B = \mathrm{Tr}_A\{\varrho_{AB}\}$

Additivität
$$S_N(\varrho_{AB}) = S_N(\varrho_A) + S_N(\varrho_B) \tag{44}$$

falls $\varrho_{AB} = \varrho_A \otimes \varrho_B$

Diese Eigenschaft ist analog zur Shannon Entropie

$$S_{sh}(X,Y) = S_{sh}(X) + S_{sh}(Y) - S_{sh}(X|Y) \tag{45}$$

d.h. $S_{sh}(X,Y) \leq S_{sh}(X) + S_{sh}(Y)$

und $S_{sh}(X|Y) = 0$ für unkorrelierte Variable X, Y

(viii) es gilt weiterhin die starke Subadditivität

(46) $$\boxed{S_N(\varrho_{ABC}) + S_N(\varrho_B) \leq S_N(\varrho_{AB}) + S_N(\varrho_{AC})}$$

Diese für die von Neumann Entropie nichttriviale Bzhg. ist die Grundlage vieler formaler Beweise in der Q.-Inf., (die wir aber in dieser Vorlesung i.A. nicht behandeln werden).

(ix) Araki-Lieb Ungleichung

(47) $$\boxed{S_N(\varrho_{AB}) \geq |S(\varrho_A) - S(\varrho_B)|}$$

Diese Ungleichung ist in scharfem Kontrast zur klassischen Ungleichung!

$$S_{sh}(X,Y) \geq S_{sh}(X), S_{sh}(Y)$$

D.h. insbesondere kann $S_N(\varrho_{AB}) < S(\varrho_A), S(\varrho_B)$.

Bspl: reiner Zwei-Parteien-Zustand

$$|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \qquad \varrho_A = \frac{1}{2}\mathbb{1} = \varrho_B$$

$$S(\varrho_A) = S(\varrho_B) = 1 \quad \text{aber} \quad S(\varrho_{AB}) = 0$$

In einem verschränkten reinen Zstd. steckt alle Infor-

mation in nichtlokalen Korrelationen.

Aus Gleichung (40) sehen wir, daß der kleinste Hilbertraum für den $S_N \neq 0$ möglich ist, der eines Spin 1/2 Teilchens ist. In Analogie zum klassischen Bit als Einheit der Information wollen wir

(48) $$\boxed{\mathcal{H} = (|0\rangle, |1\rangle) \quad \text{qubit}}$$

nennen.

Der Quantenzustand eines qubits kann stets in der Form

(49) $$\boxed{\rho = \frac{1}{2}\left(\mathbb{1} + \vec{s}\cdot\vec{\sigma}\right)}$$

wobei $\vec{s} \in \mathbb{R}^3$, $|\vec{s}| \leq 1$ der Polarisationsvektor des qubits bedeutet und $\vec{\sigma} = \sigma_x \vec{e}_x + \sigma_y \vec{e}_y + \sigma_z \vec{e}_z$ der Paulivektoroperator ist,

Es gilt

(50) $$s_i = \mathrm{Tr}\{\sigma_i \rho\}$$

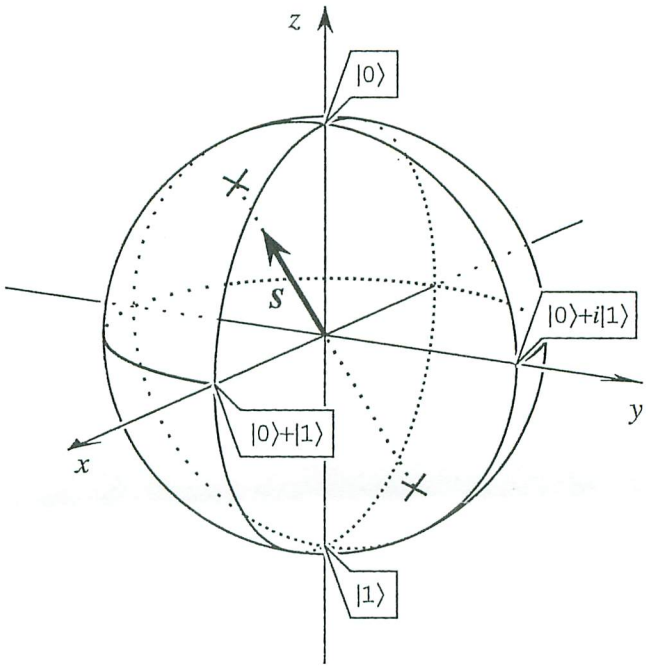
(49) kann auf Diagonalform gebracht werden

(51) $$\boxed{\rho = \rho_+ |\uparrow_s\rangle\langle\uparrow_s| + \rho_- |\downarrow_s\rangle\langle\downarrow_s|}$$

$$\boxed{\rho_\pm = \frac{1}{2}(1 \pm |s|)}$$

wobei $|\uparrow_s\rangle$, $|\downarrow_s\rangle$ Eigenzustände von $\vec{s}\cdot\vec{e} = \vec{e}_s$ sind. Die Darstellung ist eindeutig falls $|\vec{s}| \neq 0$.

Der Quantenzustand eines qubits kann mittels der <u>Bloch Kugel</u> dargestellt werden:



$|\vec{s}| = 1$ entspricht den reinen Zuständen. $|\vec{s}| = 0$ ist der maximal gemischte Zustand.

3.5. Quanteninformation vs. Information aus der Erzeugung bzw. der Messung von Quantenzuständen

In der klassischen Inf.-theorie ist die Information die in einer Zufallsvariable X steckt identisch zu der die wir durch eine Messung von X erhalten können bzw. die wir durch Kenntnis der Erzeugungsprozedur von X erlangen können. Quantenmechanisch ist die

nicht mehr der Fall:

Gegeben ein Quantenzustand ϱ. Nun messen wir die Observable

(52) $$\hat{A} = \sum_y |a_y\rangle\, a_y\, \langle a_y|$$

was uns die Ergebnisse a_y mit Wahrscheinlichkeit $p_y = \langle a_y|\varrho|a_y\rangle$ liefert. Die Messergebnisse $\{a_y, p_y\}$ sind jetzt eine klassische Zufallsvariable der wir eine Shannon Entropie zuordnen können. Diese wird <u>entropy of measurement</u> genannt. Es gilt:

(53) $$\boxed{S_{sh}(a_y) \geq S_N(\varrho)}$$

Gleichheit gilt falls $[\hat{A}, \varrho] = 0$, d.h. wenn die $|a_y\rangle$ Eigenzustände von ϱ sind.

Umgekehrt kann man die Erzeugung eines Quantenzustandes ϱ aus Mischung reiner Zustände $|\varphi_x\rangle$ mit Wahrscheinlichkeiten p_x betrachten

(54) $$\varrho = \sum_x p_x\, |\varphi_x\rangle\langle\varphi_x|$$

Wenn die $|\varphi_x\rangle$ jetzt nicht notwendig mehr orthogonal sind gilt für die Shannon Entropie der Zufallsvariable $\{x, p_x\}$:

(55) $$\boxed{S_{sh}(\varphi_x) \geq S_N(\varrho)}$$

Gleichheit gilt falls alle $|\varphi_x\rangle$ orthogonal (so hatten wir ja in Abschnitt 3.4. die von Neumann Entropie eingeführt). $S_{sh}(\varphi_x)$ wird <u>entropy of preparation</u> genannt.

<u>Bspl:</u> Quelle die qubits in $|\uparrow_{\vec{a}}\rangle$ mit Wahrscheinlichkeit p_a und in $|\uparrow_{\vec{b}}\rangle$ mit p_b erzeugt. Sei $p_a = p_b = 0.5$. Falls $\vec{a}$ und $\vec{b}$ verschieden sind, seind sie klassisch unterscheidbar !!!

- Shannon entropy of preparation

(56) $$H_{prep} = -p_a \log p_a - p_b \log p_b = 1$$

- Q.-Zustand

(57) $$\varrho = p_a |\uparrow_{\vec{a}}\rangle\langle\uparrow_{\vec{a}}| + p_b |\uparrow_{\vec{b}}\rangle\langle\uparrow_{\vec{b}}|$$

Da jeder qubit Zustand sich in der Form (51) schreiben läßt, gilt

(58) $$\varrho = \varrho_+ |\uparrow_{\vec{e}}\rangle\langle\uparrow_{\vec{e}}| + \varrho_- |\downarrow_{\vec{e}}\rangle\langle\downarrow_{\vec{e}}|$$

mit $$\varrho_\pm = \frac{1}{2}\left(1 \pm |\langle\uparrow_{\vec{a}}|\uparrow_{\vec{b}}\rangle|\right)$$

$\Rightarrow$

(59) $$H_N = -S_+ \log S_+ - S_- \log S_- \leq 1$$

- Messung: z.B. Richtg $\vec{a}$

(60) $$H_{Mess} = -p_{a+} \log p_{a+} - p_{a-} \log p_{a-} \leq 1$$

$$p_{a+} = \langle \uparrow_{\vec{a}} | S | \uparrow_{\vec{a}} \rangle = p_a + p_b |\langle \uparrow_{\vec{a}} | \uparrow_{\vec{b}} \rangle|^2$$

$$p_{a-} = p_b |\langle \downarrow_{\vec{a}} | \uparrow_{\vec{b}} \rangle|^2$$

(61) $\Rightarrow$ $$\boxed{0 \leq H_N \leq H_{Mess} \leq H_{prep} \leq 1}$$

Übertragung klassischer Information über einen Quantenkanal, d.h. mittels qubits ist nur dann ohne Verlust von Information möglich, falls die

(i) klassische Information in orthogonalen Zuständen kodiert wird

$$H_N = H_{prep}$$

(ii) bei der Messung die dabei verwendete Basis bekannt ist

$$H_N = H_{Mess}$$

Diese Besonderheit der klass. Informationsübertragung mittels Quantenkanälen werden wir uns in der Quantenkryptographie zu Nutze machen.

3.6. Quantendatenkompression

Im der klass. Inf.-theorie hatten wir Shannons noiseless coding theorem als nachträgliche Rechtfertigung dafür kennengelernt, daß die Shannon Entropie ein vernünftiges Informationsmaß ist. Jetzt wollen wir uns die Fragestellung: Was ist das Quantenanalogon von Shannons Theorem? Dazu müssen wir zwei Fälle unterscheiden

(a) Übertragung von reinen Zuständen

(b) Übertragung von gemischten Zuständen

(a) Quantendatenkompression für reine Zustände

Nehmen wir an Alice benutzt ein Alphabet aus reinen Zuständen $\{|\varphi_x\rangle\langle\varphi_x|\}$ die nicht notwendig orthogonal sein müssen. In einem (langen) Wort taucht $|\varphi_x\rangle$ mit Wahrscheinlichkeit p_x auf, d.h. wir haben

(62) $$\varrho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x|$$

Welche Dimension müsste ein Hilbertraum haben,

um ϱ effektiv zu speichern? Genauer:
Wie groß muß die Dimension des Hilbertraums sein um n Kopien von ϱ zu speichern?

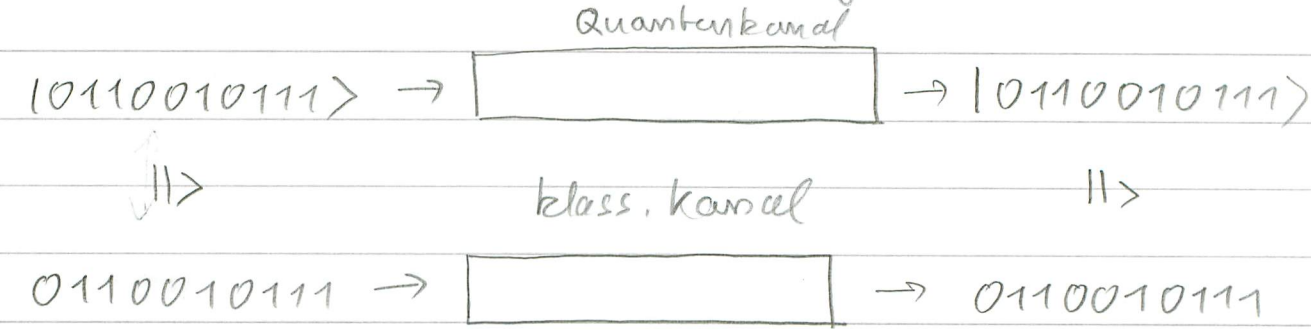
Schumacher $n \to \infty$ reine Zustände
(63) $\log(\dim \mathcal{H}) = n S_N(\varrho)$

D.h. um n qubits über einen verlustfreien Kanal zu schicken benötigt man im Mittel nur $n H_N(\varrho)$ qubits.

Der Beweis des Theorems von Schumacher ist einfach wenn die Zustände des Alphabets orthogonal sind. In diesem Fall können wir einfach die 1:1 Abbildung auf eine klassische Zufallsvariable und $S_N(\varrho) = S_{Sh}(X)$ verwenden

$$\{|\varphi_x\rangle, p(\varphi_x)\} \iff \{x, p_x\}$$

sowie Shannons noiseless coding Theorem



Wenn die Zustände $|\varphi_x\rangle$ nicht orthogonal sind, ist Schumachers Theorem nicht ganz so einsichtig.

Beispiel: Alice Alphabet

(64) $$|\uparrow_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |\uparrow_x\rangle = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Beide treten mit $p_x = p_z = 1/2$ auf. Die entropy of preparation H_{prep} ist offensichtlich gleich 1. Um die von Neumann Entropie zu berechnen diagonalisieren wir den Dichteoperator

$$\varrho = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x|$$

(65) $$= \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

ϱ hat die Eigenvektoren und Eigenwerte

(66) $$|0'\rangle = \begin{pmatrix} \cos\frac{\pi}{8} \\ \sin\frac{\pi}{8} \end{pmatrix} \qquad \varrho_{0'} = \cos^2\frac{\pi}{8}$$

(67) $$|1'\rangle = \begin{pmatrix} -\sin\frac{\pi}{8} \\ \cos\frac{\pi}{8} \end{pmatrix} \qquad \varrho_{1'} = \sin^2\frac{\pi}{8}$$

Somit ist

(68) $$H_N(\varrho) = -\mathrm{Tr}\{\varrho\log\varrho\} = -\varrho_{0'}\log\varrho_{0'} - \varrho_{1'}\log\varrho_{1'}$$

$= 0,600876\ldots \quad < 1$

Der Überlapp von $|0'\rangle$ mit $|\uparrow_x\rangle$ und $|\uparrow_z\rangle$ ist deutlich größer als der von $|1'\rangle$

$$|\langle 0'|\uparrow_x\rangle|^2 = |\langle 0'|\uparrow_z\rangle|^2 = \cos^2\frac{\pi}{8} = 0,8535$$

(69)

$$|\langle 1'|\uparrow_x\rangle|^2 = |\langle 1'|\uparrow_z\rangle|^2 = \sin^2\frac{\pi}{8} = 0,1465$$

Wenn Alice ein qubit Wort der Länge 3 an Bob senden will

(70) $$|\psi\rangle = |\psi_1\rangle|\psi_2\rangle|\psi_3\rangle$$

wobei jedes der qubits $|\psi_i\rangle$ <u>entweder</u> $|\uparrow_x\rangle$ <u>oder</u> $|\uparrow_z\rangle$ darstellt kann sie dies am effektivsten tun indem sie nur die 4 Zustände

(71) $$\{|0',0',0'\rangle,\ |1',0',0'\rangle,\ |0',1',0'\rangle,\ |0',0',1'\rangle\}$$

verwendet anstelle der $2^3 = 8$ möglichen 3er kombinationen aus $|\uparrow_x\rangle$ und $|\uparrow_z\rangle$. Man rechnet leicht nach, daß gilt

$$|\langle 0',0',0'|\psi\rangle|^2 = \left[\cos^2\left(\frac{\pi}{8}\right)\right]^3 = 0,6219$$

(72)

$$|\langle 1',0',0'|\psi\rangle|^2 = |\langle 0',1',0'|\psi\rangle|^2 = |\langle 0',0',1'|\psi\rangle|^2 = \left[\cos^2\left(\frac{\pi}{8}\right)\right]^2 \sin^2\frac{\pi}{8}$$

$$= 0,1067$$

D.h. die Wahrscheinlichkeit ein beliebiges 3 qubit Wort in dem durch (71) definierten Unterraum zu finden ist

(73) $\quad \text{Perfolg} = 0{,}6219 + 3 \cdot 0{,}1067 = 0{,}9419$

Im asymptotischen Grenzfall von n-qubit Wörtern mit $n \to \infty$ kann man einen Quantencode aus Zuständen $|0'\rangle$ und $|1'\rangle$ finden (ähnlich zu (71)) mit $n \cdot H(\cos^2 \frac{\pi}{8}) \simeq 0{,}600876\, n$ qubits und

(74) $\quad \text{Perfolg} \xrightarrow[n \to \infty]{} 1$

(b) Quantendatenkompression für gemischte Zustände

Jetzt wollen wir den Fall betrachten, daß das Alphabet von Zuständen das Alice benutzt aus gemischten Zuständen besteht. Gilt dann ebenfalls Schumachers Theorem (63)? Man kann leicht sehen daß das nicht der Fall ist:

Bsp!: Alice verwendet nur einen "Buchstaben" (qubit Zustand) ϱ_0 mit $S_N(\varrho_0) > 0$. Alice Nachricht hat also stets die Form

(75) $\quad \varrho_0 \otimes \varrho_0 \otimes \varrho_0 \otimes \varrho_0 \ldots$

In dieser Nachricht steckt gar keine Information und es wird überhaupt kein Quantenzustand zur Übertragung benötigt, obwohl $S_N(\rho_0) > 0$.

Um eine Verallgemeinerung von Schumachers Theorem für gemischte Zustände zu finden, erinnern wir uns daran, daß für reine Zustände galt

$$S_N(\{|\varphi_i\rangle\}) = S_{sh}(p(\varphi_i)) \tag{76}$$

falls die Zustände $|\varphi_i\rangle$ die mit Wahrscheinlichkeit $p(\varphi_i)$ im Alphabet auftreten (d.h. $\rho = \sum_i p(\varphi_i)|\varphi_i\rangle\langle\varphi_i|$) orthogonal sind. Besteht das Alphabet nun aus orthogonalen gemischten Zuständen ρ_i, d.h. ist

$$Tr\{\rho_i \rho_j\} = 0 \quad \text{für } i \neq j \tag{77}$$

so sollte die mögliche Kompression ebenfalls durch die Shannon Entropie $S_{sh}(p(\rho_i))$ gegeben sein. Ist

$$\rho = \sum_i p_i \rho_i \tag{78}$$

mit orthogonalen ρ_i gilt

$$\begin{aligned} S_N(\rho) &= -Tr\{\rho \log \rho\} \\ &= -\sum_i Tr\{p_i \rho_i \log(p_i \rho_i)\} \end{aligned} \tag{79}$$

$$= -\sum_i p_i \log p_i - \sum_i p_i \, \mathrm{Tr}\{\varsigma_i \log \varsigma_i\}$$

$$= S_{Sh}(p(\varsigma_i)) + \sum_i p_i S_N(\varsigma_i)$$

Die Größe

(80) $$\boxed{S_{Sh}(p(\varsigma_i)) = S_N(\varsigma) - \sum_i p_i S_N(\varsigma_i) \equiv \chi(\mathcal{E})}$$

heißt <u>Holevo Information</u> und stellt eine Verallgemeinerung der von Neumann Entropie dar. $\chi(\mathcal{E})$ hängt offensichtlich nicht nur vom Quantenzustand ς sondern auch von der Art wie ς durch ein Ensemble von gemischten Zuständen zusammengesetzt ist, d.h. $\mathcal{E} = \{p_i, \hat{\varsigma}_i\}$. Für ein Alphabet aus reinen Zuständen ist $\chi(\mathcal{E}) = S_N(\varsigma)$.

<u>Holevo Vermutung</u>

Zur Übertragung eines n qubit Wortes aus gem. Zuständen ς_i die mit Wahrscheinlichkeit p_i in dem qubit Wort auftreten sind im Limes $n \to \infty$

$$n\,\chi(\{p_i, \varsigma_i\}) \leq n\, S_N(\varsigma = \sum_i p_i \varsigma_i)$$

qubits notwendig