Zusammenfassung der 8. Vorlesung (07.06.2010)

2.4 *Dense Coding*:

Der im Abschnitt 2.3 bewiesene Satz hat neben der Teleportation noch eine weitere Anwendung. Teilen sich Alice und Bob ein bipartites System in einem maximal verschränken Zustand $|\Omega >< \Omega|$ in $\mathcal{H} \otimes \mathcal{H}$ wie bei der Teleportation, dann dann gibt es in $\mathcal{H}$ eine Familie $\{U_\nu\}$, unitärer Operatoren, $\nu = 0, 1, 2, \ldots, (M^2 - 1)$, $M = \dim \mathcal{H}$, mit den im Satz geforderten Eigenschaften. Alice kann nun mit der lokalen Operation $\mathcal{J}_\nu : \sigma \mapsto U_\nu \sigma U_\nu^+$ an ihrem Atom den nZustand des Gesamtsystems $|\Omega >< \Omega|$ in den Zustand $(\mathcal{J}_\nu \otimes \mathbf{1})(|\Omega >< \Omega|) = |\Phi_\nu >< \Phi_\nu|$ überführen´. Wenn Alice nun ihr Teilchen klassisch, etwa durch einen Boten, an Bob sendet, so dass d ieser nun auf beiden Teilchen operieren kann, dann ist Bob in der Lage, ν durch Idealmessung von $A = \sum_\mu \mu |\Phi_\mu >< \Phi_\mu|$ fast sicher, d.h. mit Wahrscheinlichkeit 1, festzustellen.

Als Träger klassischer Information speichert das transportierte Teilchen $2 \log_2 M$ Bit, obwohl es nur M Anregungszustände hat. Deshalb heisst dieses Phänomes "dichte Kodierung" (*dense coding*). Die Verschränktheit des Gesamtzustands mit dem anderen ermöglicht dem transportierten Teilchen, die Kapazität zu speichern, die ohne Verschränktheit das aus beiden Teilchen bestehende System hätte. Ein weiterer Effekt ist, dass nur Bob, der im Besitz des anderen Teilchens ist, die Nachricht fast sicher entschlüsseln kann.

3. Shannon und v. Neumann Entropie

3.1 *Informationsmaße:*

Treten K sich paarweise gegensetig ausschließende, zufällige Ereignisse mit den Wahrscheinlichkeiten p_k,$k1, 2, 3, \ldots, K$, $0 \leq p_k$, $\sum_k p_k = 1$ auf, dann ist die Shannon Entropie

$$0 \leq H := -\sum_{k=1}^{K} p_k \log p_k \leq \log K, \quad = \log 0 = 0$$

ein Maß für die in dieser Wahrscheinlichkeitsverteilung enthaltenen Information. Duch Bahl der Basis des Logarithmus kann man den Maximalwert des

Maßes festlegen. Deses Maß wurde im Abschnitt 2.2 unter dem Namen als Verschränktheitsentropie für reine, bipartite Zustände erwähnt,

$$E(\Psi) = -\sum_{i=0}^{M-1} c_i^2 \log c_i^2$$

weil es den Grad der Verschränktheit quantifiziert. H wurde con E. Shannon bereits 1948 zur Diskussion der gestörten Informationsübertragung in der Arbeit eingeführt. die als Ursprung der klassischen Informationstheorie gilt. Heute geht die Bedeutung dieses Informationsmaßes weit über die ursprungliche hinaus.

Das Verschränktheitsmaß für reine bipartite Zustände ist auch die v. Neumann Entropie der partiellen Spuren von $|\Psi><\Psi|$. Diese wurde von Johann v. Neumann schon 1932 [Mathematische Grundlagen der Quantentheorie. Berlin: Julius Springer. (1932)] beid er Diskussion des quantenmechanischen Messprozesses eingeführt

3.2 *Die Informationsfunktion*:

Man betrachte ein Alphabet $\mathcal{A} = \{a_i\}_{i=0,1,2,\ldots,(M-1)}$ und eine damit gebildete Zeichenreihe $(x_1, x_2, \ldots, x_N) \in \mathcal{A}^N$, die wir auch Wort nennen. Ist die Zeichenreihe unbekannt, dann liefert das zufällige Erkennen von Zeichen aus dieser Reihe Information über das Wort, die durch eine zu bestimmende Informationsfunktion I quantifiziert werden soll. Wie schon im Abschnitt 1.1 bemerkt wurde, kümmert dich die Informationstheorie nicht um die semiotische Bedeutung des Wortes, die Reihenfolge der Zeichen ist deshalb nicht erheblich. Die Erkennung beschränkt sich darauf, die Häufigkeiten festzustellen, mit der einzelne Zeichen in dem Wort vorkommen. Kommt das Zeichen a_i in dem Wort $n(a_i)$ mal vor, und ist die Länge des Wortes N, dann lieferrt das zufällige Erkennen des Zeichens a_i Information über die die Wahrscheinlichkeit $p(a_i) = \frac{n(a_i)}{N}$.

Um zu verdeutlichen. was unter zufälligem Erkennen von Zeichen verstanden wird, denke man sich die M Buchstaben des Alphabets auf Spielkarten geruckt. Jedes der möglichen M^N Wörter kann dann mit Hilfe von N solcher Karten zusammengestellt werden. Mischt man die Karten und legt sie verdeckt auf, entspricht dies der Unkenntnis des Wortes. Das Aufdecken einer willkürlich gewählten Karte führt zum zufälligen Erkennen eines Ze-

ichens. Nachdem diese Karte zurückgelegt und erneut gemischt wurde, kann dieses Spiel zum zufälligen Erkennen eines weiteren Zeichens wiederholt werden. Dies kann schrittweise beliebig oft fortgesetzt werden, wobei die Information über die Wahrscheinlichkeitsverteilung der Zeichen in dem Wort immer größer wird. Wird nach $\tilde{N}$ Schritten $\tilde{n}(a_i)$ mal das das Zeichen a_i erkannt, dann gilt $(\tilde{n}(a_i)/\tilde{N}) \to p_i$.

Seien nun in $\tilde{N}$ Schritten die Zeichen y_ν, $y_\nu \in \mathcal{A}$, $\nu = 1, 2, 3, \ldots, \tilde{N}$ erkannt Da die Wahrscheinlichkeiten für das Erkennen der einzelnen Zeichen nach dem oben beschriebenen Verfahren paarweise voneinander unabhängig sind, tritt dieser Fall mit der Wahrscheinlichkeit

$$p = (y_1)p(y_2)p(y_3)\ldots p(y_{\tilde{N}})$$

ein. Wie groß ist die gewonnwnw Information? Dazu kann man die folgenden drei Forderungen für die Informationsfunktion I stellen:

Axiom 1: I ist nicht negativ und nur von der Wahrscheinlichkeit p abhängig, d.h. $I : [0, 1] \cap \mathbf{Q} \to [0, \infty]$.

Hierbei wurde berücksichtigt, dass p_i eine rationale Zahl ist. Das zweite Axiom legt das Anwachsen der Information fest, wenn das schrittweise Erkennen eines oder mehrerer Zeichen in Stufen erfolgt, etwa mit $\tilde{N}$ Schritten und einem Ergebnis, das mit der Wahrscheinlichkeit p auftritt, in der ersten Stufe, und mit $\tilde{M}$ Schritten und einem Ergebnis, das mit der Wahrscheinlichkeit q auftritt, in der zweiten. Beachte, dass Wahrscheinlichkeit für die Kombination dieser Ergebnisse pq ist.

Axiom 2: I ist additiv, d.h. $I(p \cdot q) = I(p) + I(q)$

Das dritte Axiom legt die Maßeinheit fest, in der die Informationsfunktion die Information angibt. Wir wählen das Bit als Einheit.

Axiom 3: $I(1/2) = 1$.

Man folgert unmittelbar, dass diese Axiome von

$$I(p) = -\log_2 p = \log_2 \frac{1}{p}$$

erfüllt werden. Weniger trivial ist es, dass diese Funktion die einzige Funktion ist, die diese Axiom erfüllt. Der folgende Beweis wurde von Martin Schneeweiß und Matthias Singer erdacht:

Wegen $0^2 = 0$ und $1^2 = 1$ sind $I(0)$ und $I(1)$ Lösungen der Gleichung $x = 2x$, also sind diese Funktionswerte 0 oder ∞. $I(p)$ ist monoton fallend, denn aus $p < q$ folgt $I(p) = I(\frac{p}{q}q) = I(q) + I(\frac{p}{q}) \geq I(q)$. Damit ist gezeigt:

$$p < q \Rightarrow I(p) \geq I(q), \quad I(0) = \infty, \quad I(1) = 0.$$

Für $p \in (0,1)$, $\alpha, \beta \in \mathbf{N}$, $(\beta \neq 0)$, $q = \frac{\alpha}{\beta}$, gilt $I(p^q) = qI(p)$, denn $I(p) = I((p^{\frac{1}{\alpha}})^\alpha) = \alpha I(p^{\frac{1}{\alpha}})$, also $I(p^{\frac{1}{\alpha}}) = \frac{1}{\alpha}I(p)$. Betrachte nun eine Folge rationaler Zahlen $0 < p_n \leq 1$ mit $p_n \to 1$. Diese enthält eine Teilfolge q_n mit $q_n < q_{n+1}$und $q_n \to 1$ und diese wiederum eine solche r_n mit $q_1^{\frac{1}{n}} < r_n$. Nun ist $\frac{1}{n}I(q_1) = I(q_1^{\frac{1}{n}}) \geq I(r_n) \geq 0$ und damit $0 = \lim_{n\to\infty} \frac{1}{n}I(q_1) \geq \lim_{n\to\infty} I(r_n) = \lim_{n\to\infty} I(q_n) = 0$. Damit ist gezeigt.

$$p_n \in \mathbf{Q}, \quad 0 < p_n \leq 1, \quad p_n \to 1 \Rightarrow I(p_n) \to I(1)$$

Sei nun $p \in (0,1)$, rational, und p_n eine Folge rationaler Zahlen mit $p \leq p_n < 1$ und $p_n \to p$, dann gilt $I(p) = I((\frac{p}{p_n})p_n) = I(p_n) + I(\frac{p}{p_n})$. Für $n \to \infty$ gilt $I(p) = \lim_{n\to\infty} I(p_n) + \lim_{n\to\infty} I(\frac{p}{p_n}) = \lim_{n\to\infty} I(p_n) + 0$, weil $\frac{p}{p_n} \to 1$. Analog zeigt man $I(p) = \lim_{n\to\infty} I(p_n)$ für Folgen rationaler Zahlen, die von links gegen p streben. $I(p) = \lim_{n\to\infty} I(p_n)$ gilt deshalb für alle Folgen rationaler Zahlen die gegen p streben.

$$p_n, p \in \mathbf{Q}, \quad 0 < p_n < 1, \quad p_n \to p \Rightarrow I(p_n) \to I(p)$$

Ist der Grenzwwert p irrational, dann ist $I(p_n)$ eine Cauchyfolge, und dies folgt so: p_n ist eine Cauchyfolge, also gibt es für alle $\epsilon > 0$ eine Zahl $N(\epsilon)$ mit $|p_m - p_n| < \epsilon$ für $m, n \geq N(\epsilon)$. Sei o.B.d.A. $p_m > p_n$, dann ist für alle $0 < \tilde{\epsilon} = \epsilon/p_m$ auch $|1 - (p_n/p_n)| < \tilde{\epsilon}$, wenn nur $m, n \geq M(\tilde{\epsilon}) := N(\tilde{\epsilon}/p_m)$ ist. Sei nun $\epsilon_1 > 0$, dann gibt es eine Zahl $N_1(\epsilon_1)$, so dass

$$|I(p_m) - I(p_n)| = |I(\frac{p_n}{p_m})| = |I(\frac{p_n}{p_m}) - I(1)| < \epsilon_1$$

ist, wenn nur $m, n \geq N_1(\epsilon_1)$ ist, denn wir hatten weiter oben gezeigt, dass dass für Folgen rationaler Zahlen, die von links gegen 1 konvergieren, die Folge der Funktionswerte von I gegen $I(1) = 0$ konvergiert. Man kann nun für irrationale p den Wert $I(p)$ als den Grenzwert der Cauchyfolge $I(p_n)$ definieren. Die stetige Fortsetzung von I auf $(0,1]$ ist damit eindeutig durch die Werte auf den rationalen Zahlen bestimmt.

Die eindeutige stetige Fortsetzung der Informationsfunktion auf $(0,1)$ ist auch differenzierbar: Für $p \in (0,1)$ sei $1 > p_n > p_n + 1 > p$ und $p_n \to p$, dann gilt mit $\kappa_n := 1 - \frac{p}{p_n}$

$$\begin{aligned} \frac{T(p_n) - I(p)}{p_n - p} &= \frac{1}{p_n - p} I(\frac{p}{p_n}) = -I((\frac{p}{p_n})^{\frac{1}{p_n - p}}) = -\frac{1}{p_n} I((\frac{p}{p_n})^{\frac{p_n}{p_n - p}}) \\ &= -\frac{1}{p_n} I((1-\kappa_n)^{\frac{1}{\kappa_n}}) \qquad \longrightarrow \qquad -\frac{1}{p} I(\frac{1}{e}). \end{aligned}$$

Analog folgt dies für die linksseitige Ableitung. Damit ist. Damit ist $I'(p) = -\frac{1}{p} I(\frac{1}{e})$ und, wegen $I(1) = 0$ ist $I(p) = -I(\frac{1}{e}) \ln p$ die die eindeutig bestimmte Lösung der Differentialgleichung. Hierbei kann noch über $I(\frac{1}{e})$ verfügt werden. Aus $e^{\alpha x} = 2^x = y$ folgen $x = \log_2 y$, $\alpha x = \ln y$ und $\alpha = \ln 2$, und damit $log_2 y = \frac{1}{\ln 2} \ln p$. Setzt man also $I(\frac{1}{e}) = \frac{1}{\ln 2}$, dann ist

$$I(p) = -\log_2 p$$

und $I(\frac{1}{2}) = 1$, wie im Axion 3 festgelegt wurde. Damit ist folgender Satz bewiesen :

Satz: Die Axiome 1 - 3 legen die Informationsfunktion eindeutig als die Einschränkung von $I(p) = -\log_2 p$ auf rationale Argumente fest.

Die Informationsfunktion hängt nur von den Häufigkeiten ab. mit denen einzelne Zeichen in einer Zeichenreihe vorkommen, nicht aber von deren Anordnung. $I(p) = -\log_2 p$ wird auch für beliebige Verteilungen mit nichtrationalen Wahrscheinlichkeiten $p_i \geq 0$, $\sum_i p_i = 1$, die idealisiert bei unendlichen Zeichenreihen auftreten können, verwendet.

3.3 *Die Shannon Entropie*:

Die Shannon Entropie ist der Erwartungswert der Informationsfunktion

$$H = \sum_{i=0}^{M-1} p_i I(p_i) = -\sum_{i=0}^{M-1} p_i \log_2 p_i, \quad \text{wobei} \quad 0 \log_2 0 = 0$$

gerechnet wird. Besteht das Wort aus paarweise gleichen Zeichen, dann ist $H = 0$. Sind alle Zeichen gleichverteilt, dann ist $H = \log_2 M$.

$$-1 \log_2 1 = 0 \quad \leq \quad H \quad \leq \quad -\sum_{i=0}^{M-1} \frac{1}{M} \log_2 \frac{1}{M} = \log_2 M.$$

Letzterer Wert ist stationär, denn mit $0 = d1 = d(\sum_{i=0}^{M-1} p_i) = \sum_{i=0}^{M-1} dp_i$ ist $p_i = (1/M)$ Lösung von $dH = -\sum_{i=0}^{M-1}(\log_2 p_i + \frac{1}{\ln 2})dp_i = 0$. Nun ist $(\partial^2 H/\partial p_i^2) = -\frac{1}{p_i \ln 2} < 0$ und $(\partial^2 H/\partial p_i \partial p_k) = 0$ für $i \neq k$, so dass die quadratische Form $\sum_{i,k=0}^{M-1}(\partial^2 H/\partial p_i \partial p_k)dp_i dp_k$ strikt negativ ist. H ist deshalb strikt konkav und nimmt bei $p_i = (1/M)$ ihren maximalen Wert, $\log_2 M$, an.