Vorlesung am 19.05.06

6.V

IC SS-06

1.8.1: Absolute und relative Entropien (Forts.)

Die Absolute Entropie kontinuierlicher Quelle ist unendlich:

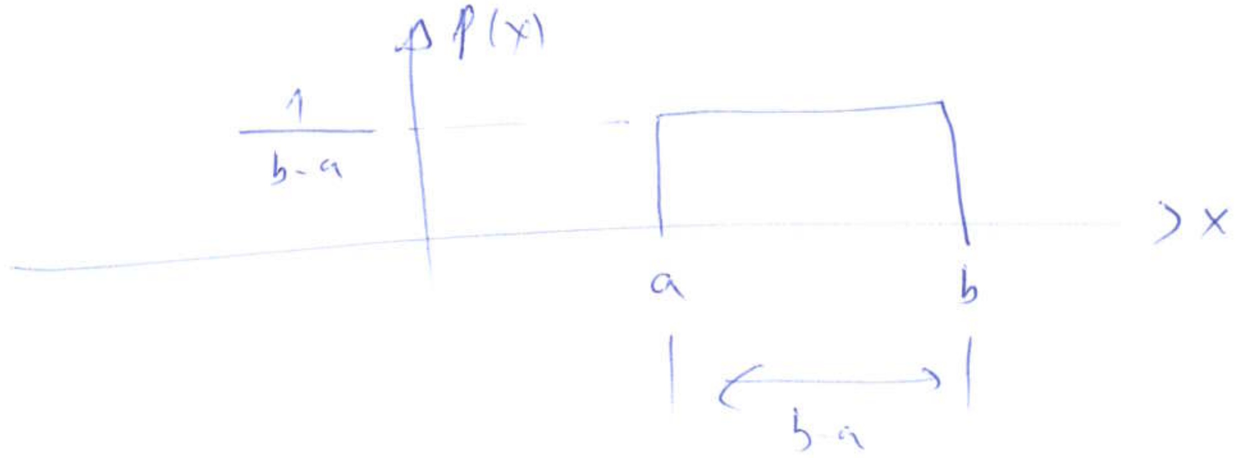
$$H_a(X) = \lim_{\Delta x \to 0} \left[\log_2 \left(\frac{1}{\Delta x} \right) \right] + H(X)$$

$$H(x) : \text{relative Entropie} = \int_a^b p(x) \log_2 \frac{1}{p(x)} \, dx$$

1.8.2: Die maximale relative Entropie kontinuierlicher Quelle

Falls a = endlich, b = endlich, erreicht die relative Entropie ihr Maximum bei $p(x) = \text{konst} = \frac{1}{b-a}$

$$\left| H(x) \right|_{max} = \int_a^b \left(\frac{1}{b-a} \right) \log_2 \left(\frac{1}{\frac{1}{(b-a)}} \right) dx = \log_2 (b-a)$$



Falls $a \to -\infty$ oder $b \to +\infty$ ist die konstante Wahrscheinlichkeitsverteilung nicht möglich, da $\frac{1}{b-a} \to 0$.

In diesem Fall müssen wir

$$H(x) = \int_{-\infty}^{\infty} p(x) \log_2 \frac{1}{p(x)} \, dx$$

für eine bestimmte Klasse von $p(x)$ maximieren.

- Beispiel : Maximale relative Entropie unter allen $p(x)$ mit $\sigma^2 = \int_{-\infty}^{\infty} x^2 p(x)\, dx = \text{konst.}$

Für diesen Fall maximieren wir $H(x) = \frac{-1}{\ln 2} \int_{-\infty}^{\infty} p(x) \cdot \ln[p(x)]\, dx$ für die Klasse von $p(x)$, die folgende Einschränkungen erfüllen:

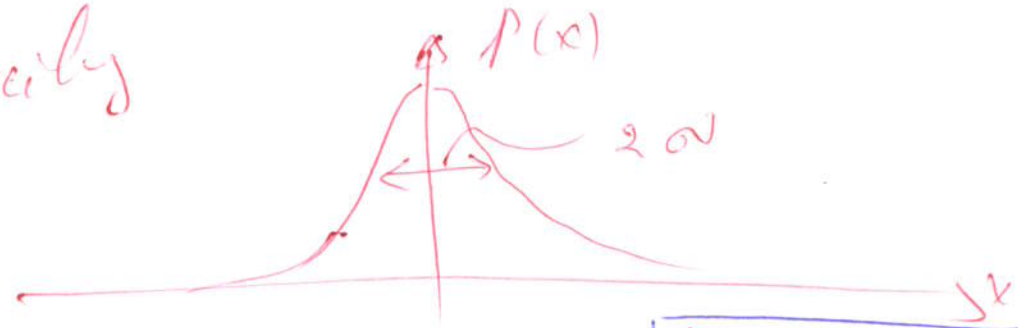
1-) $\int_{-\infty}^{\infty} p(x)\, dx = 1$

2-) $\int_{-\infty}^{\infty} x^2 p(x)\, dx = \sigma^2$ (vorgegeben)

Unter Verwendung des Lagrangischen Verfahrens, erhalten wir:

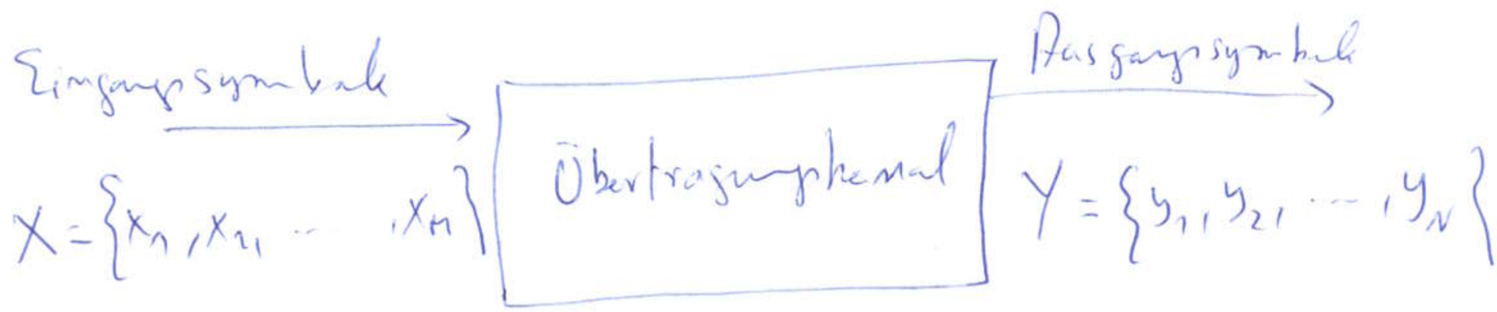
$$\boxed{p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{x^2}{2\sigma^2}}}$$

Gauß'sche Verteilung

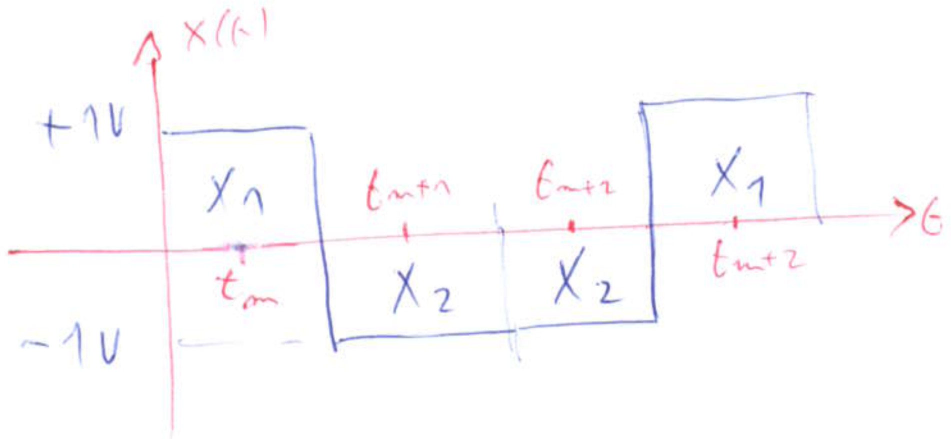


Die entsprechende, maximale Entropie beträgt: $\boxed{\hat{H}(x) = \frac{1}{2}\log_2(2\pi e \sigma^2)}$

1.9 : Diskrete Übertragungskanäle



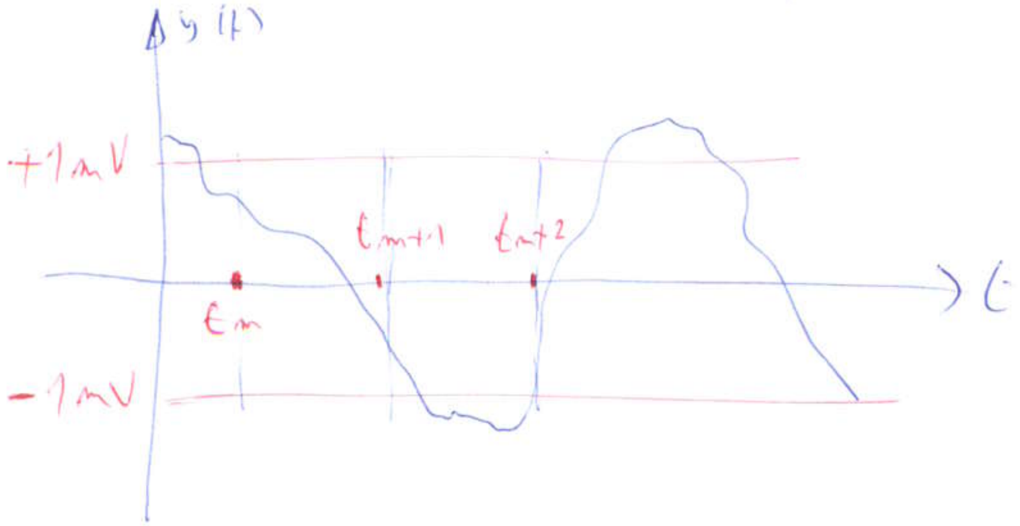
Beispiel: binäre Übertragung



$X = \{x_1, x_2\}$

$x_1 : x(t_m) = +1V$

$x_2 : x(t_m) = -1V$



$Y = \{y_1, y_2, y_3\}$

$y_1 : y(t_m) > +1\,mV$

$y_2 : y(t_m) < -1\,mV$

$y_3 : -1\,mV < y(t_m) < +1\,mV$

Das Rauschen und die begrenzte Bandbreite verfälschen das Signal. Der Effekt der begrenzten Bandbreite (Verzerrung + Abschwächung) ist deterministisch und kann berücksichtigt werden. Die Auswirkung des Rauschens ist nur stochastisch zu ermitteln:

1.9.1: ~~Verbund~~ bedingte Übertragungswahrscheinlichkeiten

$P(x_i, y_j)$ = Wahrscheinlichkeit, daß x_i gesendet und y_j empfangen wurde

$P(x_i | y_j)$ = Wahrscheinlichkeit, daß x_i gesendet wurde, wenn es bekannt ist, daß y_j empfangen wurde

$P(y_j|x_i)$ = Wahrscheinlichkeit, daß y_j empfangen wurde, wenn es bekannt ist, daß x_i gesendet wurde

Zusammenhänge:

$$P(x_i) = \sum_{j=1}^{N} P(x_i, y_j)$$

$$P(y_j) = \sum_{i=1}^{M} P(x_i, y_j)$$

$$P(x_i, y_j) = P(x_i|y_j) \cdot P(y_j) = P(y_j|x_i) \cdot P(x_i)$$

Die zwei Informationsquellen $X = \{x_1, x_2, \ldots, x_M\}$ und $Y = \{y_1, y_2, \ldots, y_N\}$ sind statistisch unabhängig, wenn

$$P(x_i|y_j) = P(x_i)$$

$$P(y_j|x_i) = P(y_j)$$

Hier gilt dann $P(x_i, y_j) = P(x_i) \cdot P(y_j)$

1.9.2: Kanalentropien und die Transinformation

$H(X|y_j)$ = bedingte Entropie der Eingangsquelle, wenn es bekannt ist, daß y_j empfangen wurde

$$H(X|y_j) = \sum_{i=1}^{M} P(x_i|y_j) \log_2 \frac{1}{P(x_i|y_j)}$$

$H(Y|x_i)$ = bedingte Entropie der Ausgangsquelle, wenn es bekannt ist, daß x_i gesendet wurde

$$H(Y|x_i) = \sum_{j=1}^{N} P(y_j|x_i) \log_2 \frac{1}{P(y_j|x_i)}$$

Rauschfreie Kanäle (selbst mit Verzerrung durch die begrenzte Bandbreite) bilden jedes Symbol x_i auf ein Symbol y_i eindeutig. In diesem Fall gilt $N = M$,

$$P(x_i|y_j) = \delta_{ij}$$
$$P(y_j|x_i) = \delta_{ij}$$

In diesem Fall verschwinden $H(X|y_j)$ für alle j's sowie $H(Y/x_i)$ für alle i's.

Extrem rauschende Kanäle führen dazu, daß die Bekanntgabe von y_j keine Hilfe leistet, um x_i zu erfahren. Dies bedeutet $P(x_i|y_j) = P(x_i)$

Oder die Bekanntgabe von x_i keine Hilfe leistet, um y_j zu erfahren. $\Rightarrow P(y_j|x_i) = P(y_j)$

In diesen Fällen gilt $H(X|y_j) = H(X)$

sowie $H(Y/x_i) = H(Y)$.

Nun definieren zwei Entropien:

$$H(X|Y) = \text{Mittelwert von } H(X|y_j)$$

$$= \sum_{j=1}^{N} H(X|y_j)\, \underbrace{P(y_j)}_{}$$

$P(x_i, y_j)$

$$= \sum_{j=1}^{N} P(y_j) \sum_{i=1}^{M} P(x_i|y_j) \log_2 \frac{1}{P(x_i|y_j)}$$

$$H(X|Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log_2 \frac{1}{P(x_i|y_j)}$$

= Entropie der Äquivokation (Zweifel oder Mehrdeutigkeit).

$H(X|Y)$ ist ein Maß für den Zweifel über das gesendete Symbol x_i, selbst, wenn das empfangene Signal bekannt ist.

Ähnlich

$H(Y/X)$ = Mittelwert von $H(Y|x_i)$

$$= \sum_{i=1}^{M} P(x_i)\, H(Y/x_i)$$

$$= \sum_{i=1}^{M} P(x_i) \sum_{j=1}^{N} P(y_j|x_i) \log_2 \frac{1}{P(y_j|x_i)}$$

$P(x_i, y_j)$

$$H(Y|X) = \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log_2 \frac{1}{P(y_j|x_i)}$$

$H(Y|X)$ = Entropie der Irrelevanz, da die Bekanntgabe von x_i zu keiner Entropie der Ausgangsquelle führen darf, falls die Übertragung fehlerfrei wäre. Dass die Ausgangsquelle die Entropie $H(Y|X)$ noch hat, trotz der Tatsache, daß x_i bekannt ist, bedeutet daß $H(Y|X)$ irrelevant ist. Oder, die Bekanntgabe von x_i führt zum Wissen von y_j mit 100%-Sicherheit falls der Kanal fehlerfrei wäre. Dass die Ausgangsquelle trotzdem eine Entropie $H(Y|X)$ noch besitzt, bedeutet, daß diese Entropie durch das Rauschen und nicht durch die Eingangsquelle produziert wurde. Dies bedeutet, daß $H(Y|X)$ irrelevante Info darstellt.

Nun, wegen $P(x_i, y_j) = P(x_i|y_j) \cdot P(y_j)$

$$= P(y_j|x_i) \cdot P(x_i)$$

$$\left.\begin{aligned} \Rightarrow\ & \log_2 \frac{1}{P(x_i|y_j)} + \log_2 \frac{1}{P(y_j)} \\ = & \log_2 \frac{1}{P(y_j|x_i)} + \log_2 \frac{1}{P(x_i)} \end{aligned}\right\} \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j)$$

$$\Rightarrow \quad H(X|Y) + H(Y) =$$

$$H(Y|X) + H(X)$$

Oder

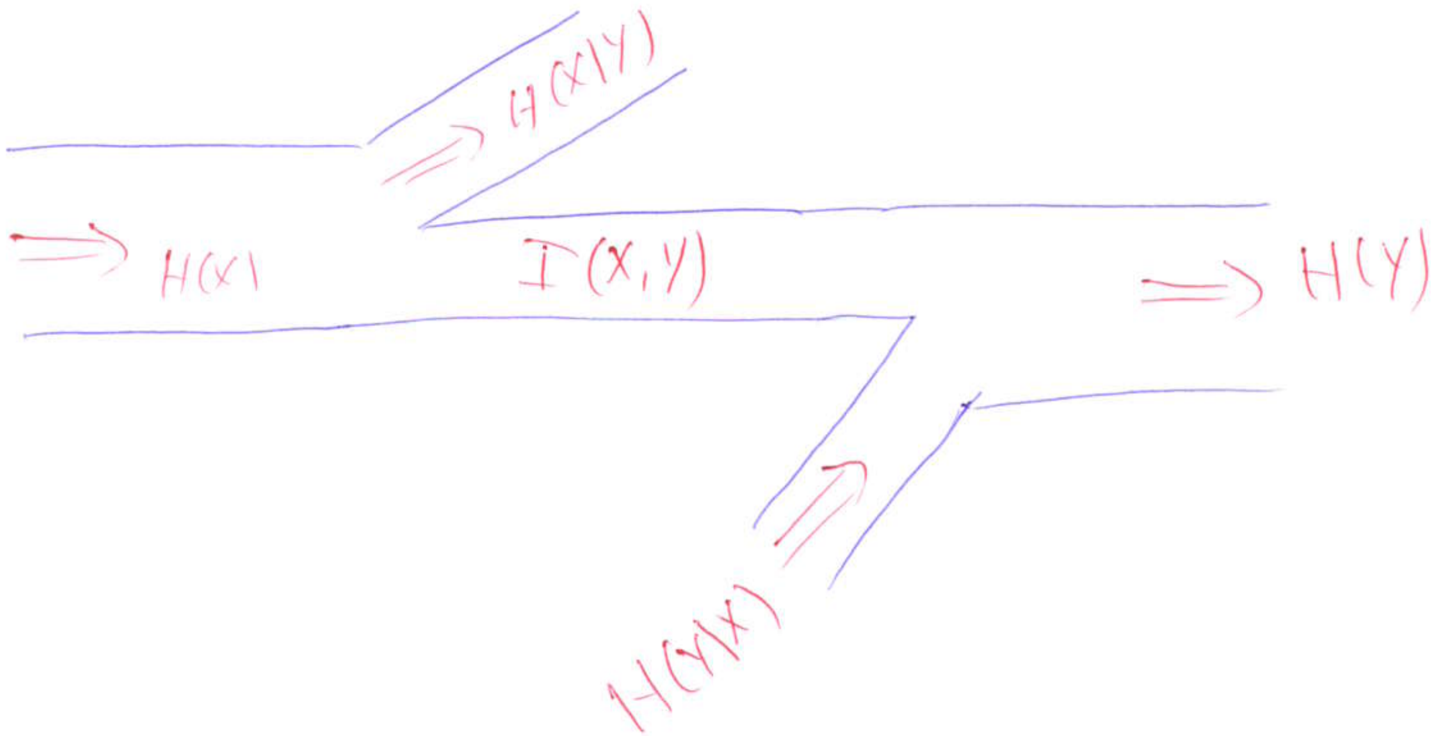
$$H(X) - H(X|Y) = H(Y) - H(Y|X)$$

$$= \underbrace{I(X,Y)}_{\text{Transinformation}}$$

Die Informationsmenge $I(X,Y)$, die über den

Kanal übertragen wurde ist die Differenz zwischen der Eingangsentropie H(X) und der durch den Zweifel verlorene Entropie H(X|Y). Oder die Differenz zwischen der empfangenen Entropie H(Y) und der durch das Rauschen erzeugte Irrelevanz Entropie H(Y|X)

Graphisch



Sonderfälle.

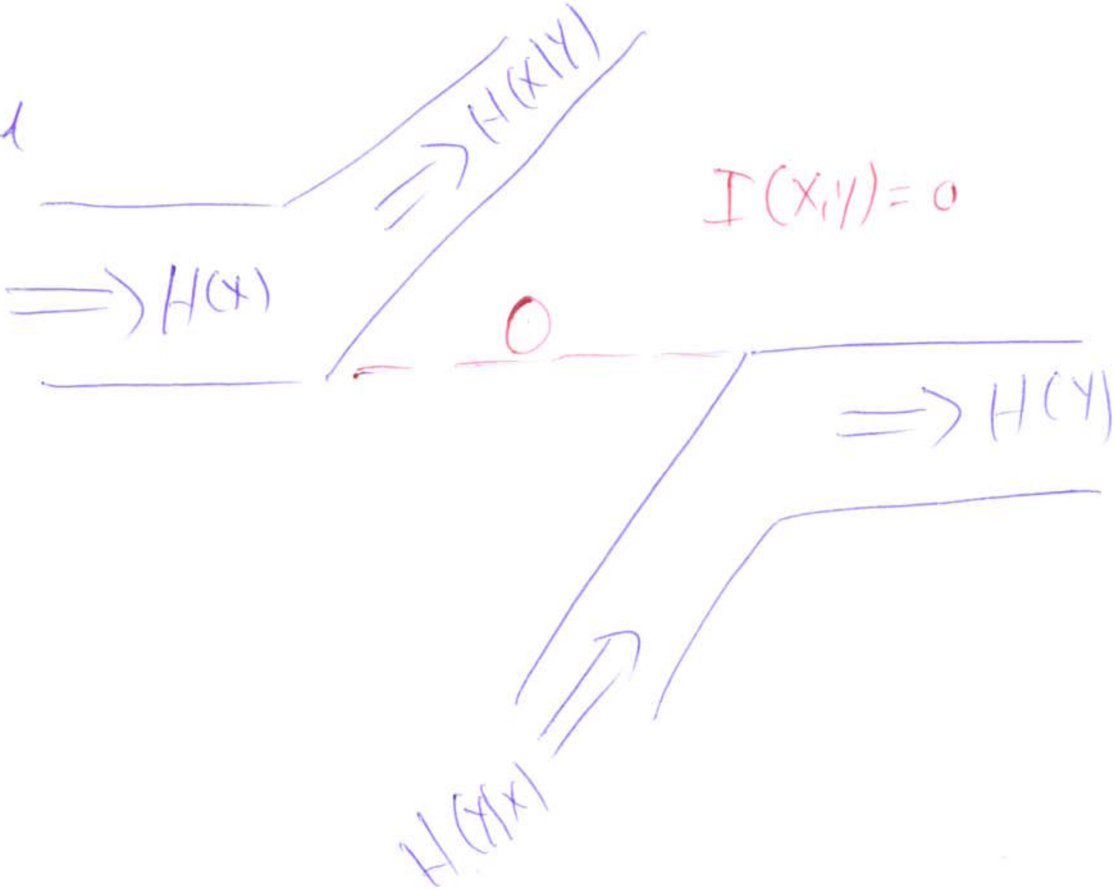
Rauschfrei

$I(X,Y) = H(X) = H(Y)$

$H(X) \Longrightarrow I(X,Y) \Longrightarrow H(Y)$

Extrem rauschend

$I(X,Y) = 0$



Symmetrischer Binärkanal ✓