

Vorlesungsmaterial
von
Prof. Dr. Horst Völz

Shannon: Informationstheorie

Dieses Material beruht weitgehend auf dem Buch

Völz, H.: Wissen - Erkennen - Information. Allgemeine Grundlagen für
Naturwissenschaft, Technik und Medizin. Shaker Verlag, Aachen 2001

Es ist u.a. auch vollständig auf der CD enthalten

Völz, H.: Wissen - Erkennen - Information. Datenspeicher von der Steinzeit bis
ins 21. Jahrhundert. Digitale Bibliothek Bd. 159, Berlin 2007

Der Inhalt wurde verkürzt, ergänzt, auf den neuesten Stand gebracht und die
Bilder farbig umgesetzt

Bei Angabe der Quelle ist das Material zum privaten Gebrauch voll nutzbar
Bei kommerzieller Nutzung bzw. in Publikationen usw. ist eine Abstimmung mit mir notwendig
Bilder sind in höherer Qualität ca. 2000×3000 Pixel oder *.cdr Version 12 verfügbar

Dieses Material wurde heruntergeladen von
aes.cs.tu-berlin.de/voelz bzw. <http://www.medienwissenschaft.hu-berlin.de/vlz>
Email: [hvoelz\(at\).fpk.cs.tu-berlin.de](mailto:hvoelz(at).fpk.cs.tu-berlin.de) bzw. [h.voelz\(at\)infocity.de](mailto:h.voelz(at)infocity.de)
Prof. Dr. Horst Völz, Koppenstr. 59, 10243 Berlin, Tel./Fax 030 296 31 07

Ziele einer Informationstheorie

- Einheitliche Beschreibung aller Informationsprozesse (z.Z. kaum erfüllbar)
- Auffinden der wesentlichen Parameter, Kennwerte, Eigenschaften
- Entwickeln einer Maß-Theorie
- Gezielte Konstruktion von Informationsprozessen

Leistungen der (Informations-) Theorie der Nachrichtenübertragung

Beschreibung: Quelle; Kanal, Störung, Empfänger; Quelle = Output: System = Input + Output

Analytisch-messtechnisch: Quelle → Entropie; Durchlassfähigkeit Kanals → Kapazität

Konstruktiv: Modulation und Codierung der Signale

a) Redundanz der Quelle b) Anpassung Kanal ↔ Quelle

Mittelbar zur Theorie: Erfassung von Fehlern; Fehlererkennung; Fehlerkorrektur;
Distortiontheorie; Kryptographie

Was die Theorie nicht leisten kann: Aussagen und Berücksichtigung von Inhalt auch
syntaktisch z.B. Bilder mehr noch Semantik/Pragmatik

Cloude Elwood **SHANNON**

30.04.1916 - 24.2.2001

1936 - 1938	Magister (PhD) „Boole“sche Algebra bei Relais
1940	Dissertation „Algebra für theoretische Genetik“
1941	Fellow of the IAS, Princeton
1940 - 1945	Grundlagen zur Kryptographie (Geheimhaltung USA)
1941 - 1972	Mathematics Department der Bell-Laboratorien
1948 (1940?!)	Mathematical Theory of Communication Informationstheorie
1958	Professor am MIT (Massachusetts Institute of Technology)
1960	Kybernetik, Shannon Maus, Labyrinth usw.

Shannon-Literatur

A Mathematical Theory of Communication) Bell Systems Technical Journal 27 (Juli 1948) S. 379-423 und (Oktober 1948) S.623 - 656. (eingereicht 24.3.1940). (Ebenfalls in: University Illinois Press 1949). Teil 2 auch: Communication in the Presence of Noise. Proc. IRE 37 (1949) pp. 10 - 20, (eingereicht am 24.3.1940), Übersetzt in: Mathematische Grundlagen der Informationstheorie. R. Oldenbourg, München - Wien, 1976; ISBN 3-486-39851-2 (WEAVER sollte nicht erwähnt werden, da völlig unwissenschaftlich u. z. T. sogar falsch)

Prediction and Entropie of printed English Bell Sys. Techn. J. 30 (1951) 1, 50f

A symbolic analysis of relay and switsching circuits (Eine symbolische Analyse von Relaisschaltkreisen), Transactions American Institute of Electrical Engineers 57; 1938, S. 713 - 723 (Eingang 1.3.38, revidiert 27.5.38; entspricht seiner Master Thesis v. 10.8.37)

Die mathematische Kommunikationstheorie der Chiffriersysteme. Bell Systems Technical Journal 28 (1949) S. 656 - 715 (Ursprünglich 1.9.45)

Datenglättung und Vorhersage in Feuerleitsystemen (mit) Technical Report Bd. 1; Gunfire Control (1946) S. 71 - 159, 166f. (Nur Auszug vorhanden)

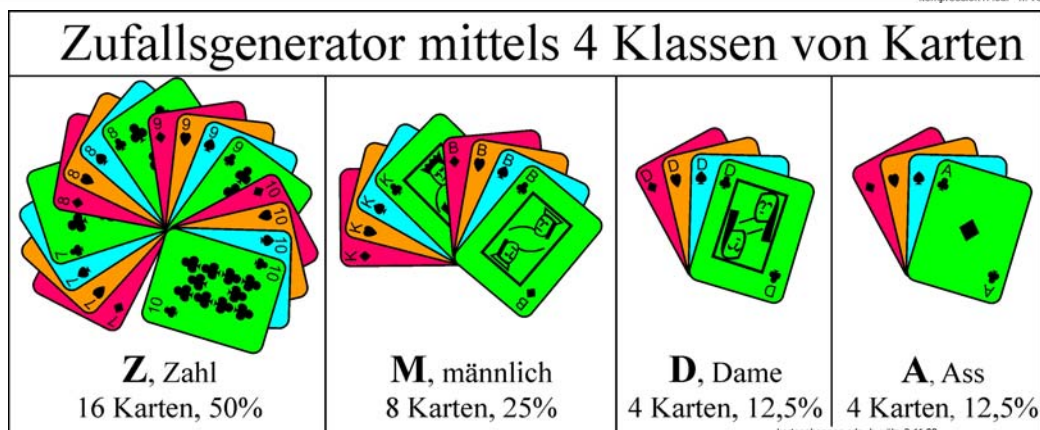
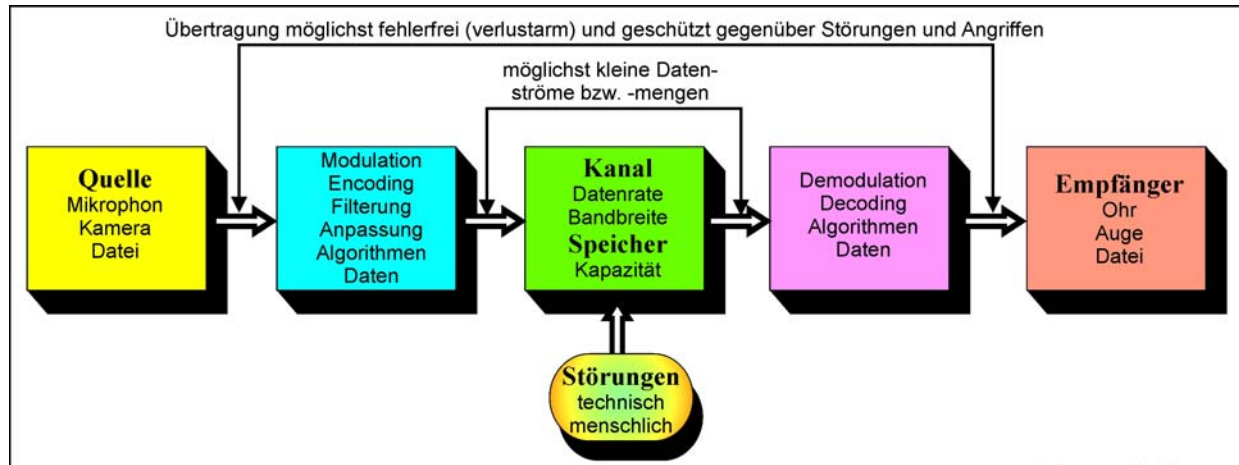
Die Philosophie der PCM Proc. IRE 36 (1948) S. 1324 - 1331 (Eingang 24.5.48)

Vorführung einer Maschine zur Lösung des Labyrinthproblems. Transactions 8th Cybernetics Conference 15. - 16.3.1951 in New York. Josiah Macy Jr. Foundation, (1952) S. 169 - 181

Eine Maschine, die beim Entwurf von Schaltkreisen behilflich ist (Proc. IRE 41 (1953) S. 1348 - 1351, (Eingang 28.5.53, revidiert 29.6.53)

Eine gedankenlesende (?) Maschine. Bell Systems Memorandum 18.3.53 (Typoskript 4 Seiten)
Vorhersage und Entropie der gedruckten englischen Sprache. (Prediction and Entropie of printed English) Bell Systems Technical Journal 30 (1951) S. 50 - 64 (Eingang 15.9.1950)

Ein/Aus, Ausgewählte Schriften zur Kommunikations- und Nachrichtentheorie. Verlag Brinkmann + Bose, Berlin 2000.



Spielablauf

- Häufig wird wiederholt: Mischen → Karte ziehen, nicht zeigen
- Über mit Ja/Nein beantwortbare Fragen jeweilige Klasse erraten

Ziel

Im statischen Mittel sind mit möglichst wenig Fragen sind möglichst viele Karten bestimmen
Dies erfordert eine Spielstrategie

Bezüge zur Informationstheorie

Kartenauswahl	↔	Signalquelle
Raten	↔	Signalerkennung
Strategie	↔	Codebaum

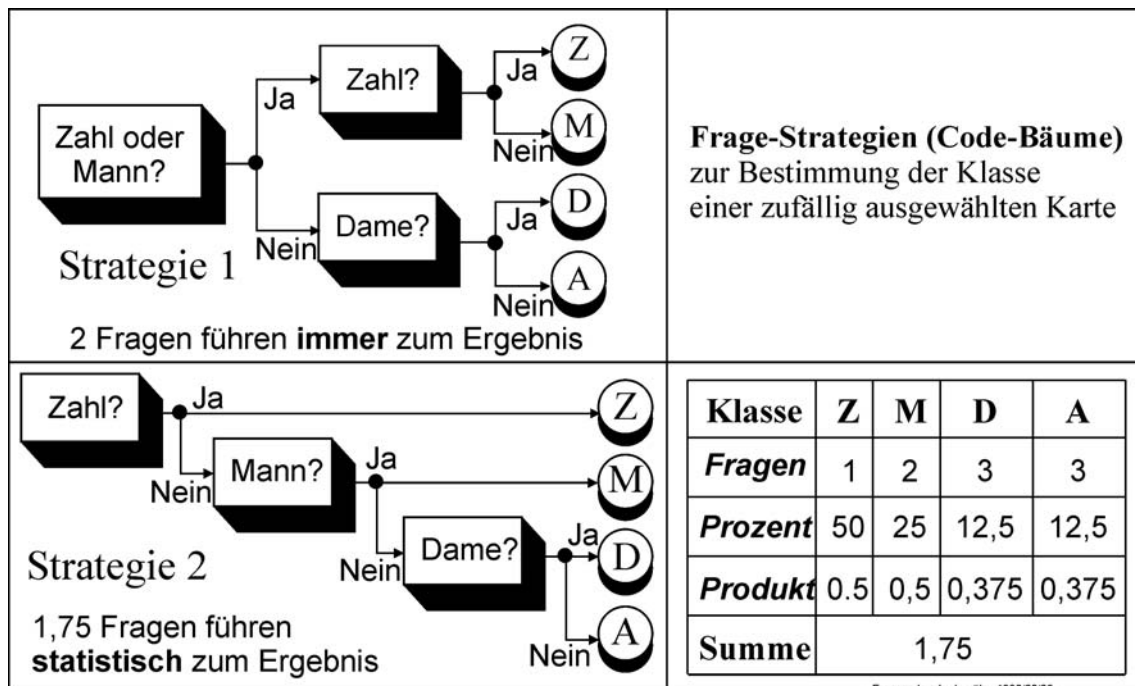
Bit

Entspricht einer Ja/Nein-Frage = Antwort beseitigte Unsicherheit

1,75 Bit/Karte als Entropie =

theoretisch minimaler Aufwand um ein Zeichen zu codieren, zu erkennen

Fragestrategie ist Codierung = konstruktiver Teil der Informationstheorie:



Ableitung der Entropie-Formel

Für 4 Speicherplätze werden unterschieden:

binäre Speicherplätze Symbol-Bits, Signal-Länge	genutzten Zeichen, mögliche Zustände Symbol-Realisierungen, konkrete Signale															
$n = \text{ld}(m)$	$m = 2^n$															
Bit 1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Bit 2	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
Bit 3	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
Bit 4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

Ein mögliches Signal ist die Bit-Folge 0101

Folglich sind zu unterscheiden:

- Die **Informationsmenge (n)**, entspricht der Anzahl von Speicherplätzen
- Die **möglichen Signale (m)** als spezielle **Varianten** der Information

$$H = n = \text{ld}(m)$$

Die Informationsmenge je Symbol **H** heißt Entropie

Es wird der binäre Logarithmus $\text{ld} = {}_2\log$ genutzt

Für **gleichwahrscheinliche Signale** gilt $p = 1/m$ (Wahrscheinlichkeit)

$$H = 1/\text{ld}(p) = -\text{ld}(p)$$

Signale treten jedoch meist mit Wahrscheinlichkeiten p_i auf.

Daher **statistische Mittelwertbildung** erforderlich

Sinnvoll ist zusätzlich eine **Bewertung** mit Gewicht p_i

$$H = - \sum_{i=1}^m p_i \cdot \text{ld}(p_i)$$

Informationen werden **addiert**

Abhängige **Wahrscheinlichkeiten** werden **multipliziert**

Voraussetzungen für Shannon-Entropie

- Existenz genau definiertes Alphabet aus n Symbolen
(dürfen nicht nachträglich geändert werden)
- Jedes Symbol hat **feste Wahrscheinlichkeit** p_i

Bemerkungen

1. Wahrscheinlichkeit *a priori*

Meist physikalisch bedingt

z. B.: Würfel, Urne, Kartenspiel, Los, Quantentheorie, thermisches Rauschen

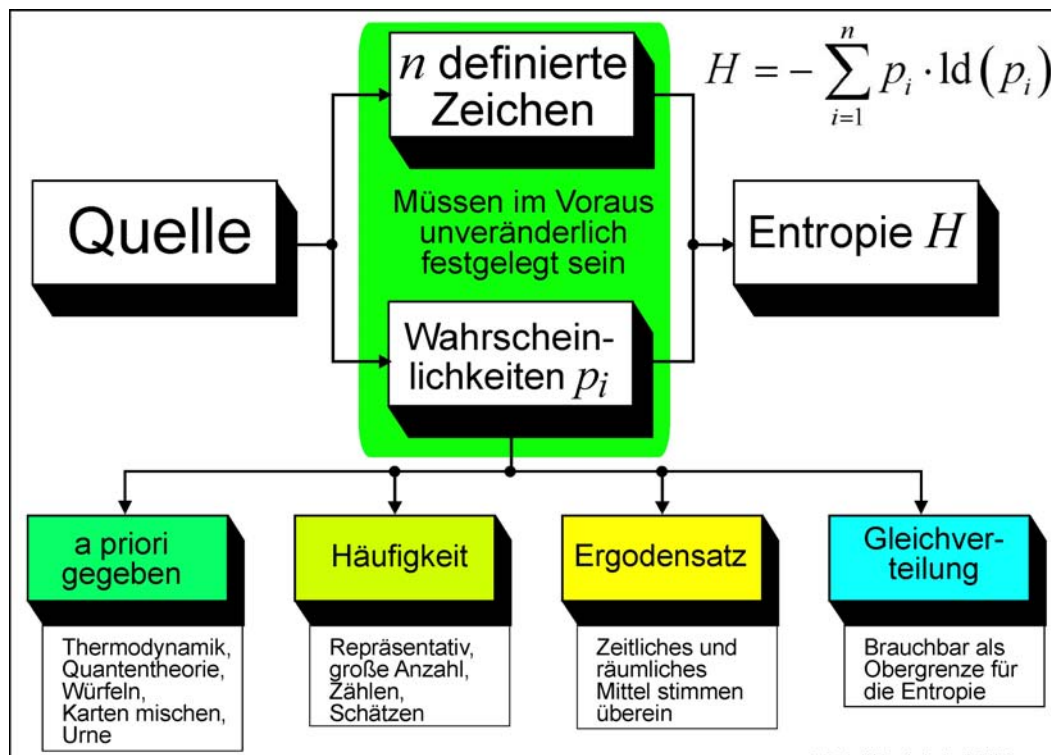
2. Häufigkeit, *a posteriori*

Abzählen, repräsentative Stichproben, Grenzwertsatz

3. Gleichverteilung

= oberer Grenzwert für Entropie, Nutzung zur Abschätzung

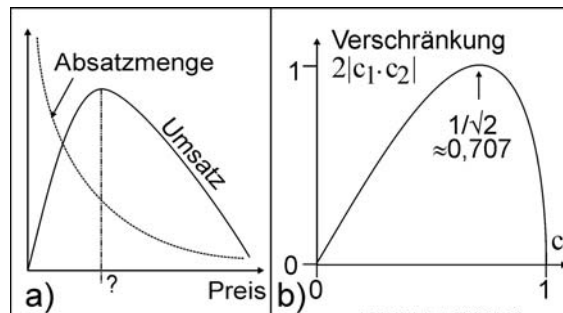
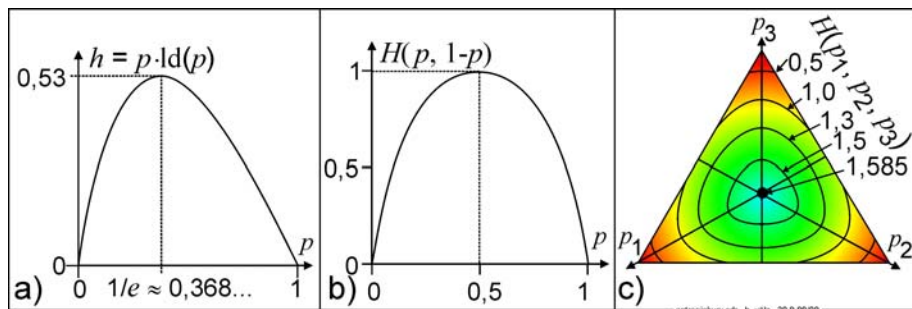
4. Ergodensatz beachten



gleichwertig serielle und parallele Ereignisse

Ergodizität

- Nicht ergodische Prozesse sind z. B. Sprachen, deren Statistik in der Zeit ändert
- Zeitlicher Mittelwert muss mit Ensemble-Mittelwert übereinstimmen
- Bei Markow-Prozessen darf
 1. der Graph darf nicht so in Teile zerfallen, dass man nicht zu allen Knoten zurück gelangen kann
 2. muss der größte gemeinsame Teiler aller Zykluslängen (der Zahl Kanten in einer geschlossenen Schlaufe) gleich 1 sein.
 3. Nichtergodische Quellen, laufen zu Stellen, die Endknoten sind, aus denen der Prozess nicht mehr heraus kann. Problem der Berechnung der Übergangswahrscheinlichkeiten.



Axiomatische Ableitung der Entropie-Formel nach Feinstein

1. $H(p, 1-p)$ ist **stetig** in p mit $0 \leq p \leq 1$ 2 Zeichen
2. $H(0,5, 0,5) = 1$ **Normierung**
3. $H(p_1, p_2, \dots, p_n)$ ist unabhängig von der **Reihenfolge** n Zeichen
4. Verfeinerungstheorem: Zerlegung $p_n = q_1 + q_2$ $n \rightarrow n+1$ Zeichen
dann: $H(p_1, p_2, \dots, q_1+q_2) = H(p_1, p_2, \dots, p_n) + p_n \cdot H(q_1 / p_n + q_2 / p_n)$

Hinweise für die Ableitung

- $H(1) = 0$: **Gewissheit**
- n **gleichwahrscheinliche** Symbole: $p_i = 1/n$: $H(n_1) > H(n_2)$ wenn $n_1 > n_2$
- zwei **unabhängige** Urnenversuche: $p(A; B) = p(A) \cdot p(B) \rightarrow H(A; B) = H(A) + H(B)$
erfordert den **Logarithmus**, dessen **Basis** noch unbestimmt ist
- **ein Versuch** mit $p = 1/n \rightarrow h = c \cdot {}_x\log(1/n)$ folglich bei Summierung Faktor p notwendig
 $h = p \cdot {}_x\log(1/n) = -p \cdot {}_x\log(p)$
- Durch **Summierung** über alle Versuche folgt also $H = - \sum_{i=1}^n p_i \cdot {}_x\log(p_i)$
 - Wegen **Normierung** (oben 2.) kommt nur noch den **binäre Logarithmus** in Betracht

Beispiel für das Verfeinerungstheorem

Fall K_0

Klasse	Karten	p	$\text{ld}(p)$	$-p \cdot \text{ld}(p)$
Zahl	7, 8, 9, 10	0,5	-1	0,5
Männlich	Buben, Könige	0,25	-2	0,5
Sonstige	Damen, Asse	0,25	-2	0,5

Entropie = 1,5 Bit/Karte

Nun wird „Sonstige“ zerlegt in **Weiblich** + **Aß**

Klasse	Karten	p	$\text{ld}(p)$	$-p \cdot \text{ld}(p)$
Weiblich	Damen	0,5	-1	0,5
Aß	Asse	0,5	-1	0,5

Zwischen-Ergebnis = 1 ($\times 0,25$)

Für die verfeinerte Klasse folgt deshalb wieder:

Klasse	Karten	p	$\text{ld}(p)$	$-p \cdot \text{ld}(p)$
Zahl	7, 8, 9, 10	0,5	-1	0,5
Männlich	Buben, Könige	0,25	-2	0,5
Weiblich	Damen	0,125	-3	0,375
Sonstige	Aß	0,125	-3	0,375

Ergebnis = 1,75 Bit/Karte

Codierung von n Zeichen

➤ Shannon- Code

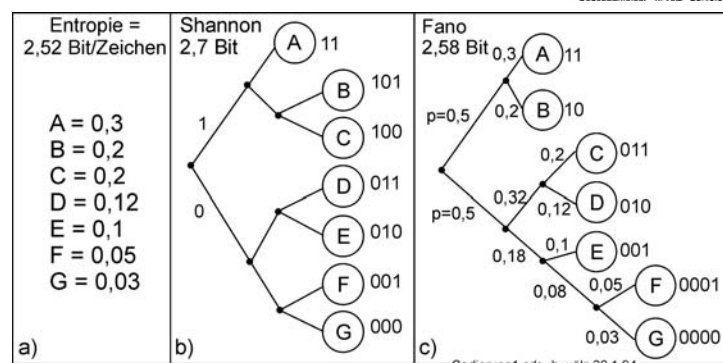
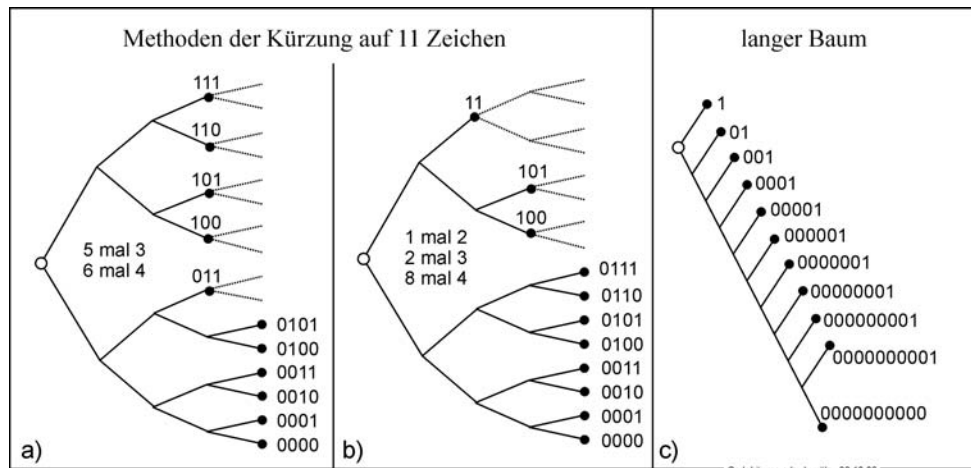
1. Codebaumtiefe k so wählen, dass $2^k \geq n$
2. Überzählige Endknoten durch Zusammenfassen von je zwei Endknoten schrittweise entfernen
3. Zusammengefassten Endknoten mit Symbolen der größten Wahrscheinlichkeit belegen

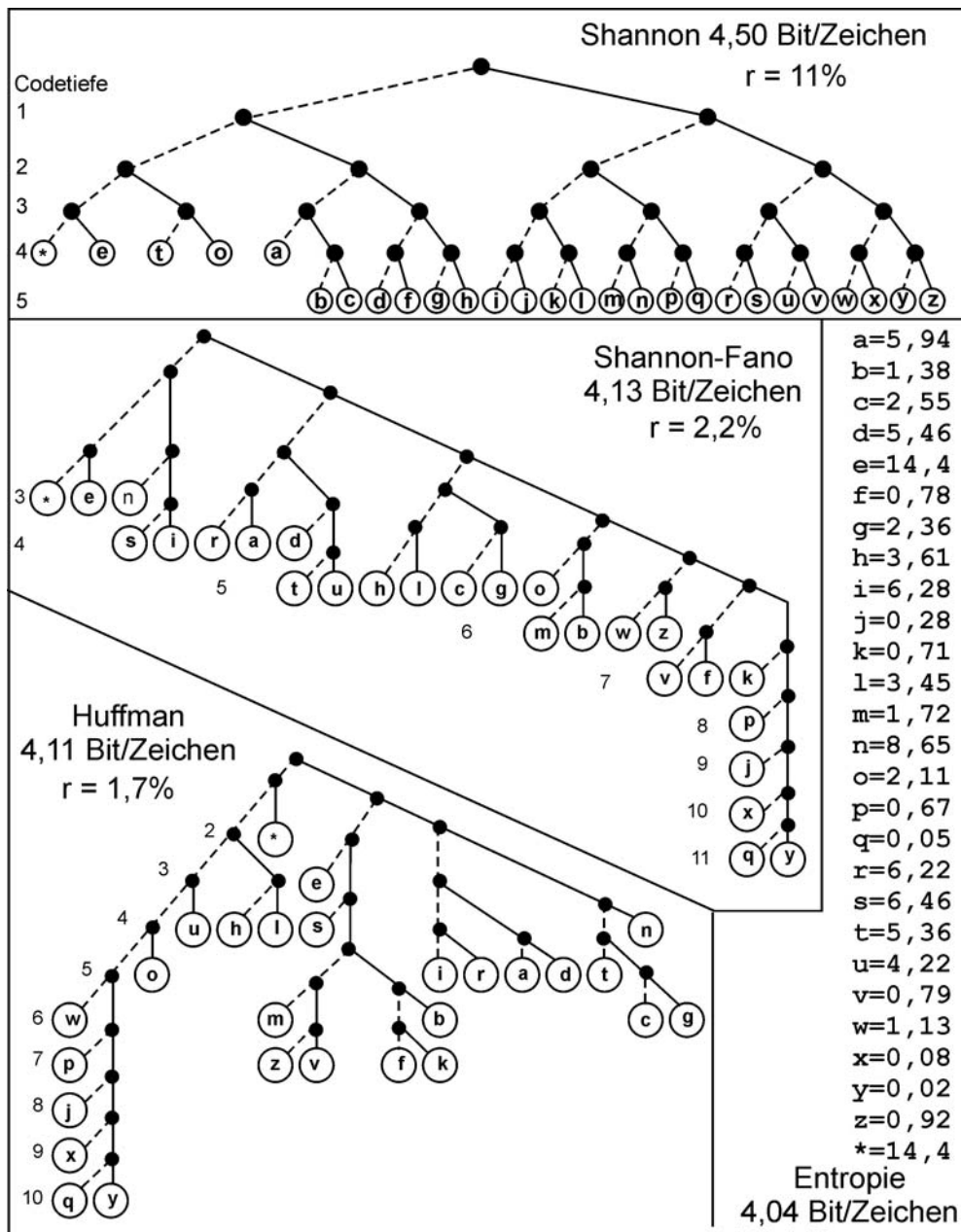
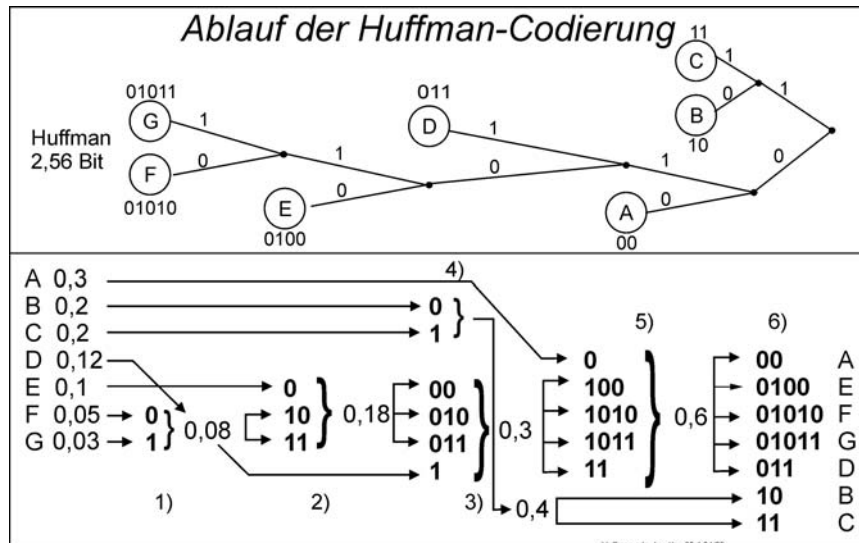
➤ Fano-Shannon-Code

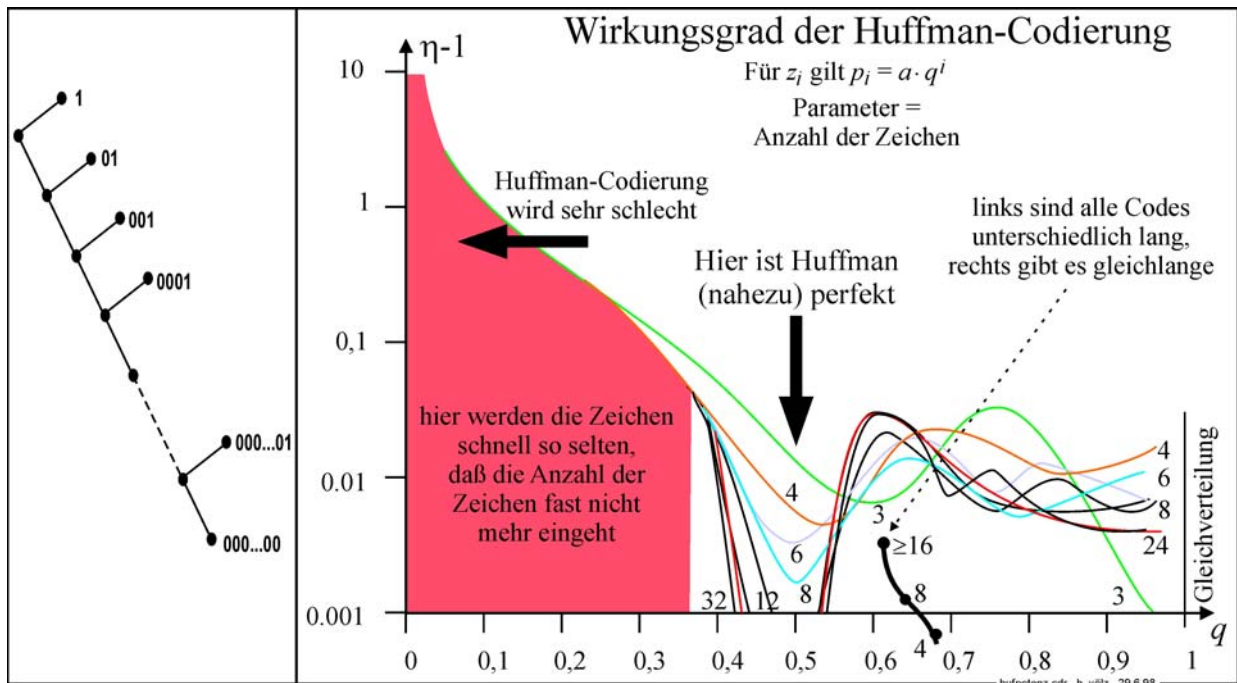
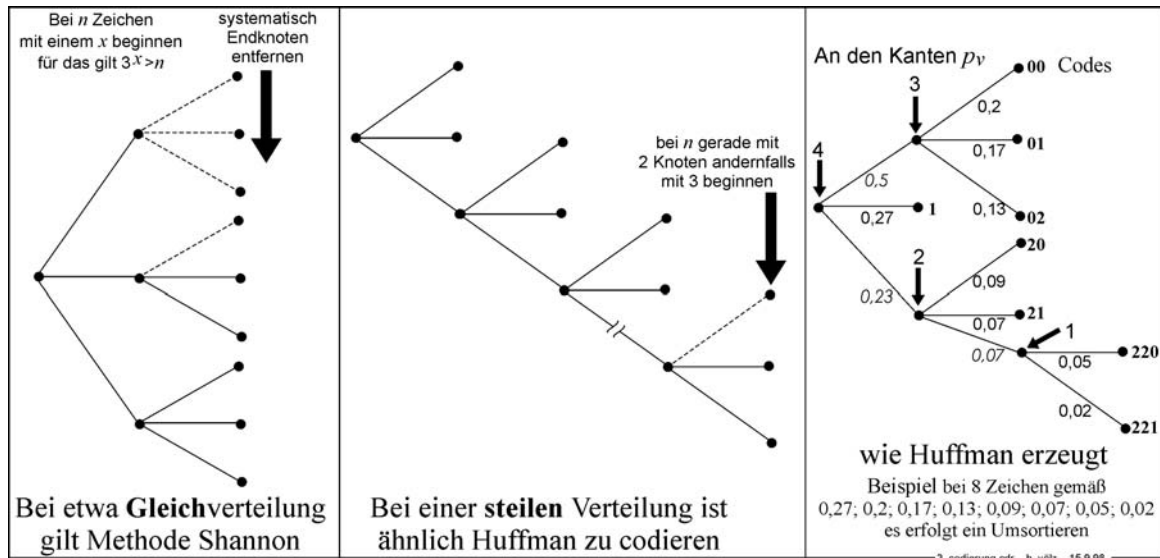
1. Symbole nach fallender Wahrscheinlichkeit sortieren
2. Zwei Gruppen mit möglichst gleicher Wahrscheinlichkeit bilden
3. Obere Teilgruppe mit „1“ codieren, untere mit „0“ (*Code an vorhanden anfügen*)
4. Mit jeder Teilgruppe nach 2. solange fortfahren, bis jeweils einzelne Symbole erreicht

➤ Huffman-Code

1. Symbole nach fallender Wahrscheinlichkeit sortieren
2. Die beiden Symbole mit kleinster Wahrscheinlichkeit mit „0“ bzw. „1“ codieren (*Code vor vorhandenen einfügen*)
3. Beide Symbole aus dem Symbolvorrat entfernen und als ein neues Hilfssymbol zusammenfassen, und mit der addierten Wahrscheinlichkeiten versehen
4. Bei 1. solange fortfahren, bis nur zwei Hilfssymbole existieren







Ergebnis der Codierung

Von n Symbolen mit Wahrscheinlichkeit p_i

Theorie → Entropie H	Anwendung → Codeaufwand A
jedes Symbol liefert Anteil $-\text{ld}(p_i) = 1/\text{ld}(p_i)$	Zu jedem Symbol gehört Code-Länge l_i
$H = - \sum_{i=1}^n p_i \cdot \text{ld}(p_i)$	$A = \sum_{i=1}^n p_i \cdot l_i$

Es gilt immer $H \leq A$

Die Differenz $R = A - H$ heißt (Shannon-) **Redundanz**.

Für eine ideale Codierung ist sie 0

Oft wird auch die **relative Redundanz** $r = \frac{A-H}{H}$ benutzt

Übergang $A \rightarrow H$ bzw. $R \rightarrow 0$

Es gilt allgemein $A \xrightarrow{\text{Kombinationslänge} \rightarrow \infty} H$ bzw. $R \rightarrow 0$

Dafür hat schon Shannon Code mit Gedächtnis eingeführt und bewiesen, dass bei Gedächtnis $\rightarrow \infty$ der erreichbare Codeaufwand $\rightarrow$ Entropie
 Leider nicht praktikabel: Speicherkapazität und Zeitverzug $\rightarrow \infty$

Zeichen	Wahrscheinlichkeit	Code	Aufwand, Redundanz
A	0,75	0	1
B	0,25	1	23,26%
Zweierzeichen			
AA	0,5625	0	0,84375×2 4%
AB	0,1875	11	
BA	0,1875	100	
BB	0,0625	101	
dreifach			
AAA	0,421875	1	0,82292×3 1,435%
AAB	0,140625	001	
ABA	0,140625	010	
BAA	0,140625	011	
BBA	0,046875	00000	
BAB	0,046875	00001	
ABB	0,046875	00010	
BBB	0,015625	00011	

Gültigkeit des Huffman-Codes

notwendige Voraussetzungen:

- Zeichensatz (Alphabet) steht fest
- Wahrscheinlichkeit der Zeichen steht fest
- Encoder und Decoder haben kein Gedächtnis
- Es wird Präfix-Code verwendet (binär $\leftrightarrow$ trinär?!)
- Synchronisiert sich selbst

Präfix-Code = irreduzibel, kommafrei oder natürlich. Kein Code-Wort darf der Anfang eines anderen sein. Shannon, Fano, Huffman: Kodebaum nur Endknoten.

Kein Codewort darf der Anfang eines anderen sein.

wenn *Haus* dann nicht zulässig *Haustür*, *Haushalt* usw.

für Menschen schwer lesbar

IMPRAKTISCHENLEBENISTMIRKEINBEISPIELHIERFÜRBEBKANNTBEREITSAL
 LEZUSAMMENGESETZTENWÖRTERWIDERSPRECHENVOLLDIESEMPRINZIP

impraktischenlebenistmirkeinbeispielhierfurbekanntbereitsallezusammengesetztenwörterwiderr
 prechenvolldiesemprinzip

Decodierbarkeit

1. **Präfix-Code** = SHANNON, FANO, HUFFMAN: Kodebaum nur Endknoten

2. **Gleichmäßiger Code** alle Codewörter gleich lang: Z.B.

- ASCII-Code = 8 Bit
- Postleitzahlen
- Genetischer Code 3 Sequenzen
- BLZ Bankleitzahl
- ISBN für Bücher

Dennoch eventuell Startzeichen notwendig.

3. **Sondertrennzeichen:**

- Allgemein spezielle Zeichen zur Trennung z.B. asynchrone seriellles Protokoll (V24 usw.) „Mark“ und Stoppbit
- Synchronübertragung spezielles, meist langes Codewort
- Sprache: „Zitat“ und „Ende des Zitats“
- Interpunktionszeichen der Schrift

NS-Zeit im KZ Künstler sprachliche Laute: „,“ bum, „,“ peng, „,“ bum-peng, „!“ sching-bum, „?“ tütelit-bum, „-“ tatütata. Goethes „Erkönig“.

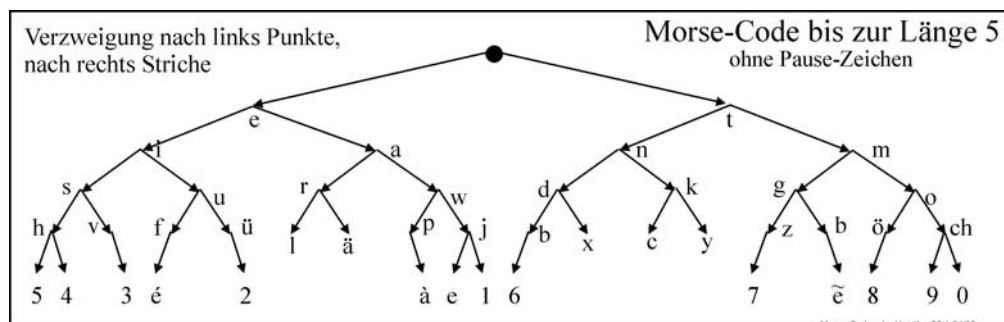
Beispiel Morse-Code

a	d	e	h	i	l	n	r	s	w
.-	-..	.		..	.-.	-.	.-.	...	.-.

Zeichenfolge:-.....-..... ergibt u.a.:
seinen adel herr viel ies nah d

benötigt zusätzlich Pausen als Sondertrennzeichen

also zwei Symbole: Punkt und Strich + Hierarchie von Trennzeichen



normal nach Punkt oder Strich (eventuell verschieden) + nach Zeichen + nach Wort

Sonderfälle mit Decodierbarkeit

A. Römische Zahlen

Es muss altes und neues System unterschieden werden (Regeln hier unvollständig)

1. Es gibt ein Alphabet aus 7 Zeichen: **M** 1000; **D** = 500; **C** = 100; **L** 50; **X** = 10; **V** = 5; **I** = 1
2. Zeichen müssen in dieser Reihenfolge erscheinen, dann erfolgt Addition
3. Mit Ausnahme von **M** maximal 3-fache Wiederholung
4. Es ist möglich genau ein niederwertiges Zeichen voran zu stellen. Das bedeutet Subtraktion.

Später Festlegung von genau vier Subtraktionszeichen

900 = CM; 400 = CD; 90 = XC; 40 = XL; 9 = IX; 4 = IV

B Trennzeichen = echtes Zeichen

A = 0 Jedes Zeichen beginnt mit 0,

B = 01 trotzdem 0 kein Trennzeichen

Erweiterung durch Anfügen von 1: C = 011; D = 0111 usw.

Entropie kontinuierlicher Signale

1. Grenzeinflüsse

- Übliche Ableitung gilt für **Gauß**-Verteilung beim Signal und Rauschen
- **Amplitudenbegrenzte** Signale mit $x_1 \leq x \leq x_2 \rightarrow p(x) = 1/(x_2 - x_1)$; $h(x) = \text{ld}(x_2 - x_1)$
- **Leistungsbegrenzt** $x_{m2} = \int_{-\infty}^{+\infty} x^2 \cdot p(x) \cdot dx$ mit Verteilung $p(x) = \frac{1}{x_m \cdot \sqrt{2 \cdot \pi} \cdot e} \cdot e^{-(x/x_m)^2/2}$
ergibt $h(x) = \text{ld}(x_m \cdot \sqrt{2 \cdot \pi} \cdot e \approx \text{ld}(4,133 \cdot x_m)$

2. Verteilungen

- Üblich ist Gaus-Verteilung
- Außerdem berechnet: Rechteck, Dreieck, $1/x^2$, e^{-x/x_0}

Wegen schwerer Randwertaufgabe sind keine weiteren Verteilungen bekannt

Wichtige Begriffe

Analog, kontinuierlich, stetig, diskret, Quantisieren, digital, dual usw.

Analog

griechisch **logos** Vernunft + lateinische **ana** auf, wieder, aufwärts, nach oben; lateinisch **analogia** mit der Vernunft übereinstimmend, Gleichmäßigkeit

Substantiv **Analogie** etwa: Entsprechung, Ähnlichkeit, Gleichwertigkeit, Übereinstimmung.

Anwendungen u. a.

- **Technik**: unterschiedliche Systeme mit gleichartigem Verhalten, z. B. Analogrechner und elektromechanische, akustische oder wärmeelektrische Analogien; Analoguhr drehende Zeiger
- **Biologie/Medizin**: analoge Organe bzgl. Morphologie/Struktur, z. B. Auge Wirbeltier, Tintenfisch und Insekt
- **Rechtsprechung**: juristisch nicht geregelter Tatbestand auf etwa wesensgleichen bezogen; in Deutschland unzulässig.
- **Literatur** $\approx$ Fabeln, Parabeln, Märchen, Gleichnis
- **Psychologie**: Denken, MAXWELL (Wasserströmung $\Leftrightarrow$ elektrische Felder)
KEKULE (Affens $\Leftrightarrow$ Benzol-Ring)
- **Kybernetik** betrachtet Kommunikations- und Steuerungssysteme bei lebenden Organismen in Analogie zu Maschinen. Achtung!: Umkehr von Technik zu Mensch (harte KI) führt zu Fehlschlüssen
- **Bionik**: Umsetzung des Körperbaus von Delphinen zur Optimierung von Schiffsrümpfen, Lotos-Effekt usw.
- **Platon**: menschliche Seele dreigeteilt (Vernunft, Wille, Begierden). Gerechter Mensch kontrolliert Begierden durch Vernunft, mit Unterstützung des Willens. **Analogie** zum Dreiständeaufbau des Staates: der erleuchtete Philosoph | König regiert Gesellschaft mittels der Krieger.
- **Logik**: induktive Beweisführung: Wenn Größen in einer oder mehrerer Hinsicht einander ähnlich sind, dann könnten sie auch in anderen Hinsichten einander ähnlich sein

Gegensatz zu **analog** kein Begriff (**nicht-analog**)

kontinuierlich

lateinisch continens, continuus: zusammenhängend, angrenzend an, unmittelbar folgend, ununterbrochen, jemand zunächst stehend

Continuare: aneinanderfügen, verbinden, fortsetzen verlängern, gleich darauf, ohne weiteres
contingere: berühren, kosten, streuen, jemandem nahe sein, beeinflussen.

Mathematik: Kontinuum der reellen Zahlen zwischen zwei Zahlen gibt es immer eine weitere

Physik: Kontinuumsmechanik berücksichtigt nicht die Mikrostruktur der Materie.

Signale: Möglichkeit beliebiger Zwischenwerten, bezüglich *Zeit und Amplitude* (Energie)

Umgangssprache: nicht vorhanden

Verwandt mit stetig

Deutsch: von stehen verwandt mit stet, stets. das Gegenteil von unstet.

Mathematik: spezieller als Grenzwert $\lim_{x \rightarrow 0} f(x)$

Umgangssprache: beharrlich, ununterbrochen, ständig

diskret

lateinisch discretio: Unterscheidungsvermögen, Urteil und Entscheid, **discretus**: abgesondert, getrennt, **discernere**: scheiden, trennen, unterscheiden, beurteilen, entscheiden

16. Jh. in deutsche Sprache

Physik Größen, die sich nur in endlichen Schrittweiten ändern.

Mathematik: etwas zerfällt oder gliedert sich in einzelne Punkte oder Elemente

Signal: nur endlich viele, meist genau definierte Werte

Umgangssprache: taktvoll, rücksichtsvoll, zurückhaltend, unauffällig, unaufdringlich, vertrauensvoll, geheim, verschwiegen.

Quantisieren

lateinisch quantitas Größe, Anzahl bzw. **quantum** wie viel, so viel wie, inwieweit, irgendwie. Also Menge

Physik Quant durch Max Planck eingeführt. Gegenteil **Qualität** etwa Güte

Philosophie: Zusammenhang von Quantität ($\approx$ Menge) $\leftrightarrow$ Qualität ($\approx$ Güte)

Signal: Erzeugen diskreter Amplituden und Zeitpunkte (Takte)

digital

lateinisch digitus: Finger

Inhaltlich: zählen, ziffernmäßig, in Zahleneinheiten angeben.

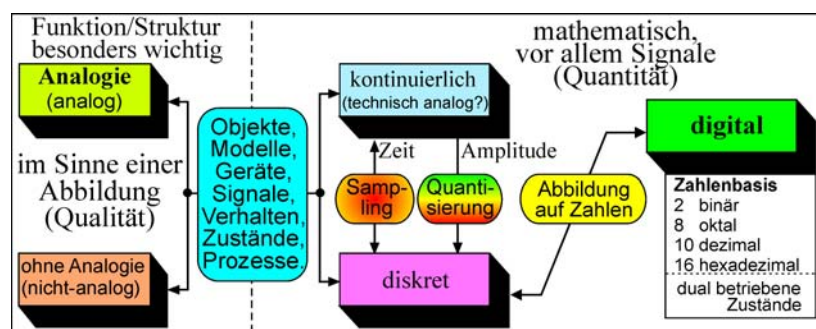
Biologie: digitalis Pflanze Fingerhut

Englisch: alte Maßeinheit Fingerbreite = 18,5 mm

Signal: Zahlenbasis erforderlich: *binär 2, oktal 8, dezimal 10, hexadezimal 16*.

Zustände können abweichen, zum Signal viele Codierungen möglich. z. B. dual bei zwei physikalische Zuständen

BCD =binär codierte Dezimalzahl.

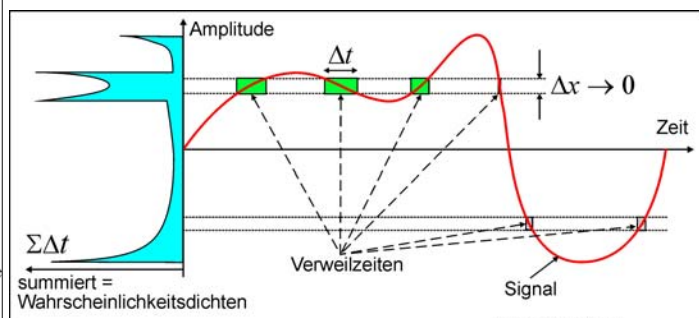
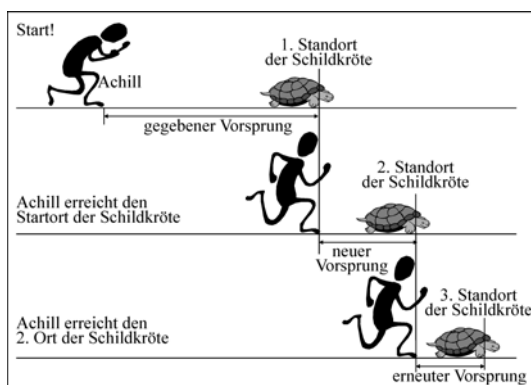


Vor-, Nachteile kontinuierliche, digitale Signale

kontinuierlich (analog)	diskret - digital
Stärken, Vorteile	
• es sind beliebig kleine (Verstärkung) und große Signale möglich • fast nur von technischen Sensoren, Aktoren und menschlichen Sinnen benutzt • Bei Übersteuerung setzen Verzerrungen ein	• Signale leicht fehlerfrei zu regenerieren • Verlustfreies Übertragen, und Kopieren (Vervielfältigen) • Verschachtelung mehrerer, auch recht unterschiedlicher Signale • Fehlererkennung und -korrektur • Verlustfreie Komprimierung • Datenschutz (Verschlüsselung)
Schwächen, Nachteile	
• Bei jeder Übertragung, Speicherung und Vervielfältigung kommen Störungen, zumindest Rauschen hinzu	• Signale müssen eng begrenzte Pegel besitzen, insbesondere hinreichend groß sein • Der Takt darf nicht verloren gehen

Beschreibungen von Welt

Typ	kontinuierlich	diskret
Griechen	ARISTOTELES	DEMOKRIT
Mathematik	unendlich	endlich
Antinomien	Hilbert-Hotel, grot = grün-rot	XENON (Pfeil, Wettlauf)
Paradoxien	Alle Raben sind schwarz	Kretaer Lügner, GÖDEL- Unentscheidbarkeit
Methoden	Limes Differential	Zählen, Messen
Beschreibung	Differentialgleichungen	Algorithmen
Geräte	Analogrechner	Digital-Computer
Entsprechung	Welle	Korpuskel, Quant
Übergang	Bohr'sches Korrespondenz -Prinzip HEISENBERG-Unbestimmtheit	



Entropie kontinuierlicher Signale

Einführung der Wahrscheinlichkeitsdichte

Es werden n Intervalle Δx gewählt, die eine Wahrscheinlichkeit $p_n(\Delta x)$ besitzen

Für die Entropie gilt dann

$$H(n, \Delta x) = - \lim_{\substack{\Delta x \rightarrow 0 \\ n \rightarrow \infty}} \sum_{i=0}^n \underbrace{p_i(\Delta x) \cdot \Delta x}_{\text{Wahrscheinlichkeit}} \cdot \text{ld}(p_i(\Delta x) \cdot \Delta x)$$

Grenzübergang zum Integral (Produkt unter dem Logarithmus in die Summe zerlegen)

$$H(x) = - \int_{-\infty}^{+\infty} p(x) \cdot \text{ld}(p(x)) \cdot dx - \lim_{\substack{\Delta x \rightarrow 0 \\ n \rightarrow \infty}} \sum_{i=0}^n p_i(\Delta x) \cdot \text{ld}(p_i(\Delta x)) \cdot \Delta x$$

Der zweite Teil divergiert wegen

$$\lim_{\Delta x \rightarrow 0} \log(\Delta x) \rightarrow -\infty$$

Daher wird der erste Term als relative Entropie bezeichnet

$$h(x) = - \int_{-\infty}^{+\infty} p(x) \cdot \text{ld}(p(x)) \cdot dx$$

In der Praxis gibt es neben Nutzsignal immer eine Störung. Für sie gilt ebenfalls der Grenzübergang. Differenzbildung unterdrückt dann 2mal die divergierenden Terme (!!)

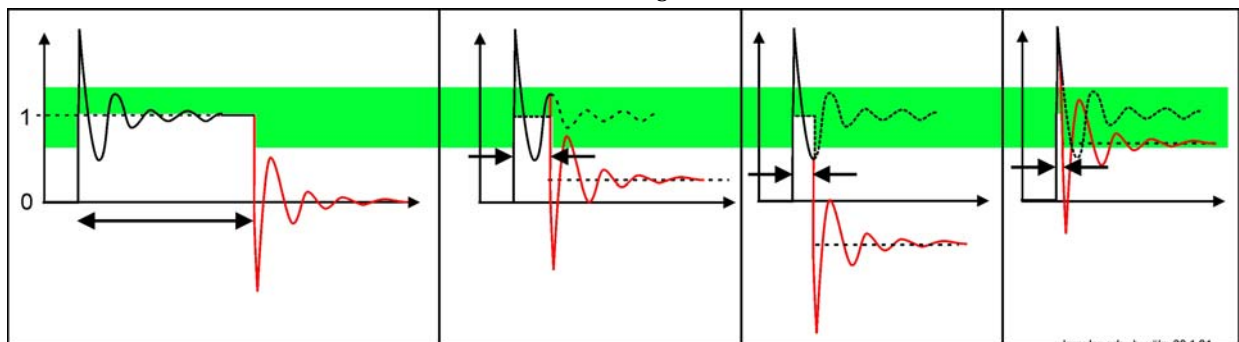
$$H(x) = h_{\text{Nutz}}(x) - h_{\text{Stör}}(x)$$

Gaußsche Verteilungen bei Signal und Störung ergibt dann

$$H = \text{ld} \left(\frac{P_N + P_S}{P_S} \right)$$

Mögliche Interpretation: Es gibt m Pegel mit unterscheidbaren Amplitudenstufen

$$m = \frac{P_N + P_S}{P_S}$$



Sampling-Theorem

englisch *sample* Probe

Karl KÜPFMÜLLER (1897- 1977) Einschwingen bei Systemen etwa 1924

Harry NYQUIST (1889 - 1976) Pulsmodulation etwa 1930

Vladimir Alexandrowitsch KOTELNIKOW, (1908 - 2005) Abtastung 1933

Claude SHANNON (1916 - 2001) exakt abgeleitet 1940

Rekonstruktion kontinuierlicher Signale mit Bandbreite B Probenabstand

$$\Delta T \leq \frac{1}{2 \cdot B}$$

Kernstück ist Whittaker-Funktion (Tonfilm, Magnetband als Spaltfunktion bekannt)

$$x = \frac{\sin(\alpha)}{\alpha}.$$

Nullstellen im Abstand $n/(2B)$ mit $n = \pm 1, \pm 2, \pm 3$ usw.

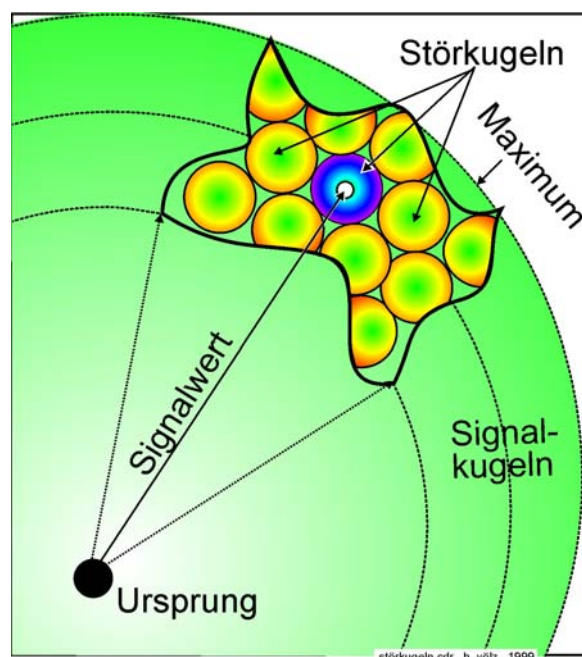
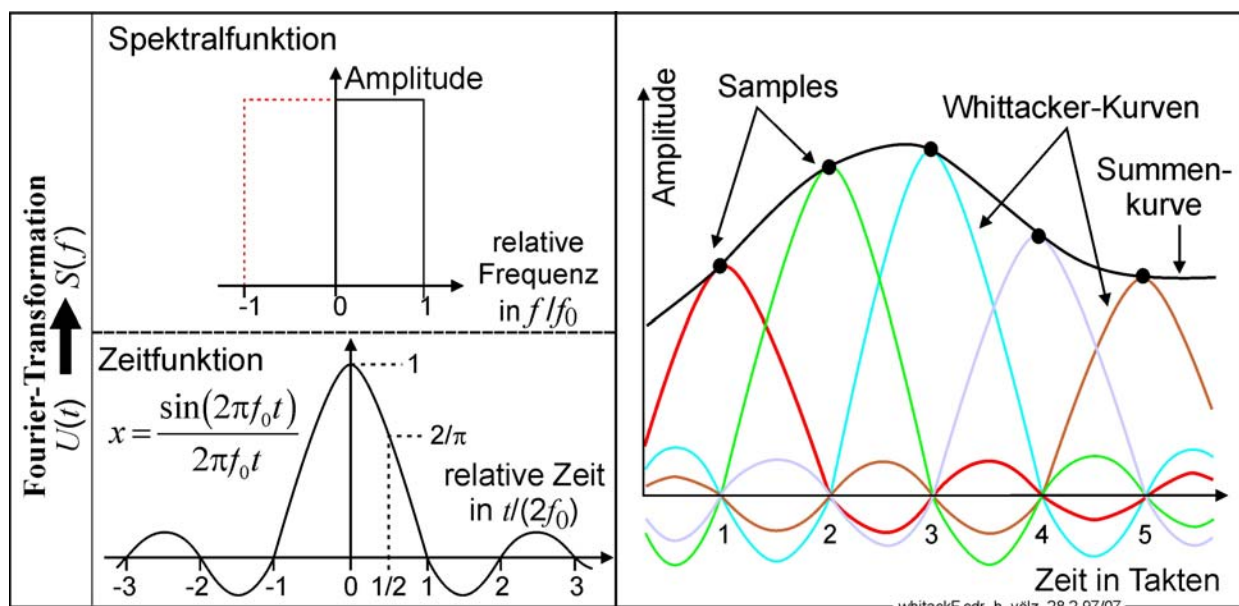
Analogie Heisenberg-Unschärfe-Relation, für konjugierten Größen z.B. Zeit Δt und Energie ΔE mit der Planck-Konstante h

$$\Delta t \cdot \Delta E = h/2.$$

gilt speziell für Photon (Lichtquant) der Frequenz ν

$$\Delta E = h \cdot \nu.$$

oben eingesetzt folgt Sampling-Theorem



Kanalkapazität im n-dimensionalen Raum

Signal der Zeitdauer T und Bandbreite B benötigt wegen Sampling-Theorem $n = 2 \cdot B \cdot T$ Proben.

Alle Werte spannen einen $2 \cdot B \cdot T$ -dimensionalen Raum auf

Durch die Nutzleistung P_n und Störleistung P_s sind mit einer Konstanten a Radien bestimmt

$$r_n = a \cdot \sqrt{P_n} \text{ und } r_s = a \cdot \sqrt{P_s}$$

Jedes Eingangssignal ist die Addition beider

$$r_e = r_n + r_s$$

Das Volumen einer Hyperkugel berechnet sich gemäß

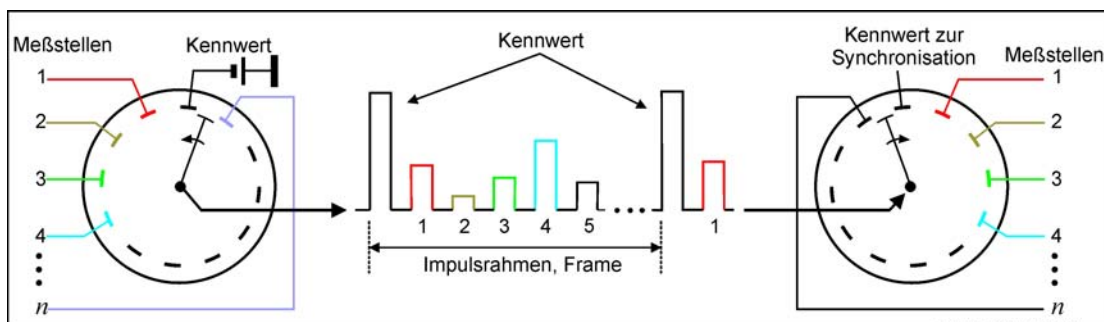
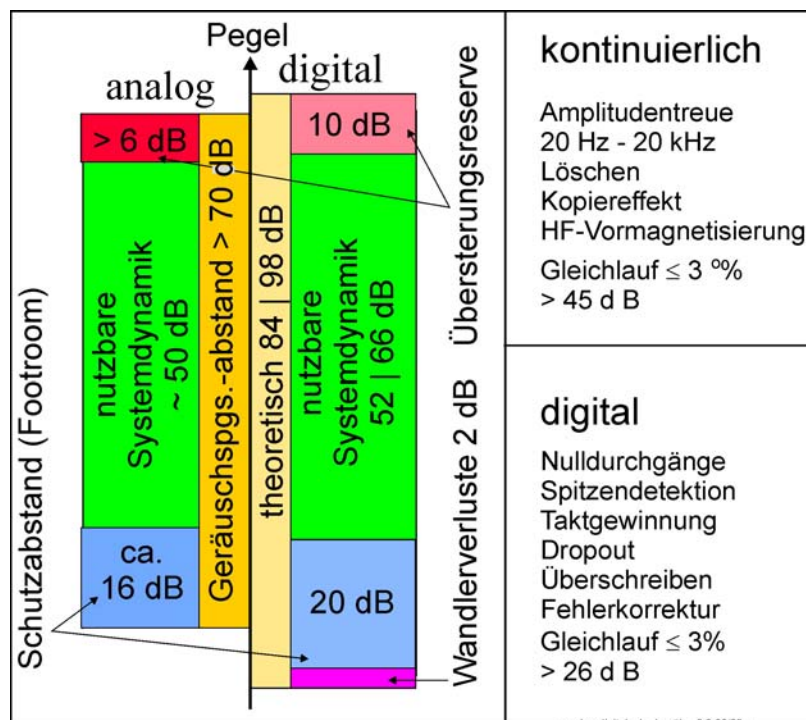
$$V = \frac{\Pi^{n/2}}{\Gamma(1+n/2)}$$

In der Hyperkugel mit Radius r_e befinden sich maximal M unterscheidbare Kugeln mit r_s

$$M \leq \frac{V_e}{V_s} = \left(1 + \frac{P_n}{P_s}\right)^{B \cdot T}$$

M sind quasi Amplitudenstufen. Die Kanalkapazität muss nur auf die Zeit normiert werden

$$C = \frac{\text{ld}(M)}{T} = B \cdot \text{ld} \left(1 + \frac{P_n}{P_s}\right)$$

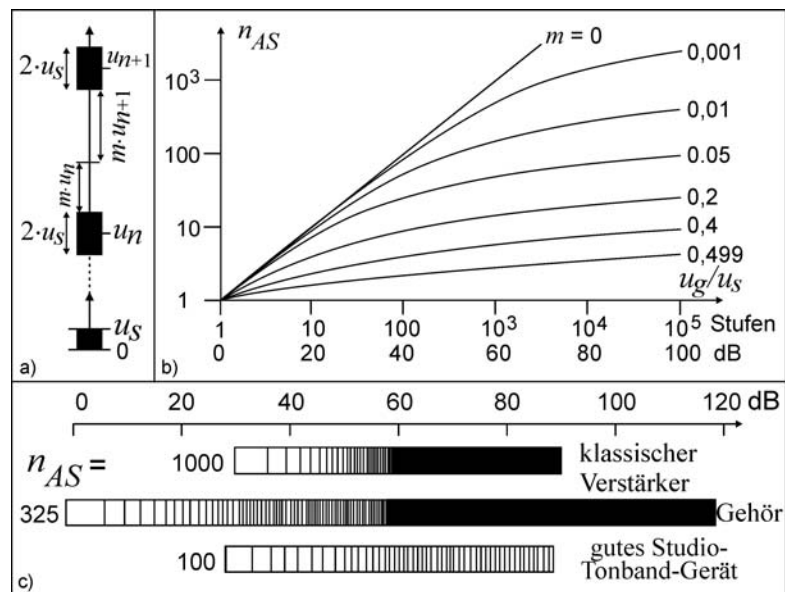
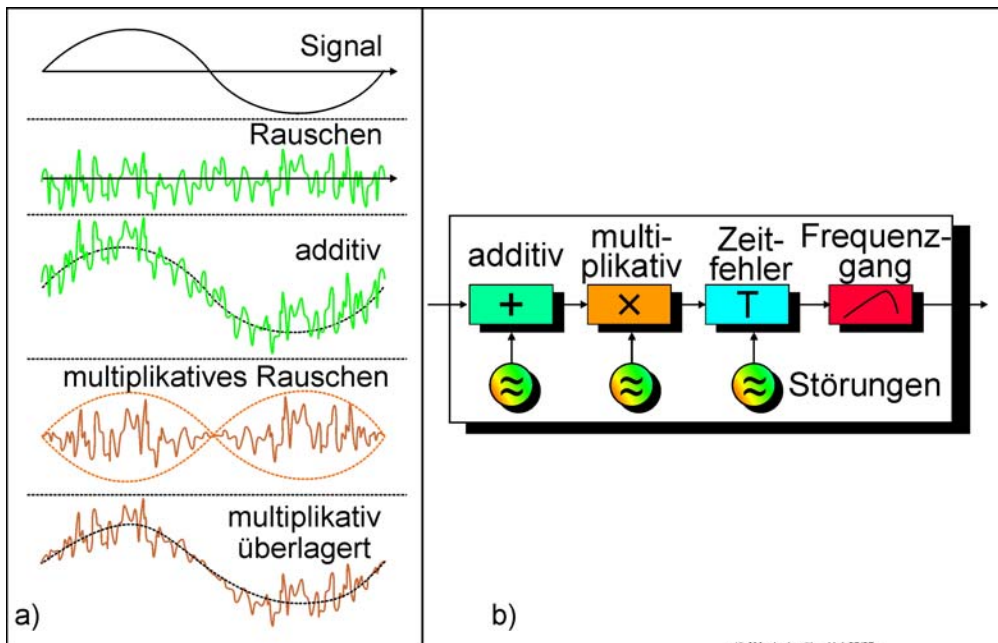


Wichtige Kenngröße SHANNON

Entropie	Kanalkapazität
diskret $H = \sum p_i \cdot \lg(p_i)$	$C = 2 \cdot B \cdot H$
kontinuierlich $H = \lg \frac{P_{\text{nutz}} + P_{\text{st}}}{P_{\text{st}}}$	$= 2 \cdot B \cdot \lg \frac{P_{\text{nutz}} + P_{\text{st}}}{P_{\text{st}}}$

Vergleich Wirkungsgrade

SHANNON	CARNOT
Redundanz gegenüber Entropie bzw. Kanalkapazität	Verlust gegen idealen Kreisprozess



Störende Amplitudenmodulation

Abstand Amplitudenstufen unterschiedlich

- additives Grundrauschen u_s

Mit Modulationsgrad m .

- Modulationsrauschen u_M , etwa proportional der Signalamplitude u_n

Für die gesamte Störung gilt

$$u_{St} = u_s + m \cdot u_n.$$

Grenze zwischen der n -ten und der $(n+1)$ ten Amplitudenstufe

$$u_n \cdot (1+2 \cdot m) + u_s = u_{n+1} \cdot (1-2 \cdot m) - u_s.$$

Es wird eine Substitution benutzt

$$u_n^+ = u_n + \frac{u_s}{2 \cdot m} \quad \text{daraus folgt} \quad \frac{u_{n+1}^+}{u_n^+} = \frac{1+2m}{1-2m} \quad \text{Daraus folgt}$$

Für n Stufen gilt weiter

$$\frac{u_{n+1}^+}{u_1^+} = \left(\frac{1+2m}{1-2m} \right)^n$$

Rücknahme der Substitution und u_{n+1} als größter Wert u_g ergibt Anzahl der möglichen unterscheidbaren Amplitudenstufen

$$n_{AS} = \frac{2 \cdot \log \frac{2 \cdot m \cdot \frac{u_g}{u_s} + 1}{2 \cdot m + 1}}{\log \frac{1+2 \cdot m}{1-2 \cdot m}} + 1 \approx \frac{\ln \left(2 \cdot m \cdot \frac{u_g}{u_k} + 1 \right) - \ln(1+2 \cdot m)}{2 \cdot m} + 1$$

Superzeichen

von FELIX CUBE bestimmt

Text enthält k Wörter mit je n Zeichen,

also gesamt $m = k \cdot n$ Zeichen

Jedes Wort besitzt eine Einzel-Entropie
(da gleichverteilt angenommen $\approx$)

$$H_1 \approx n \cdot \text{ld}(n)$$

Alle Wörter liefern folglich

$$H_2 \approx m \cdot \text{ld}(n)$$

Wörter können auch zu einem Text
zusammengefügt werden. Dafür gilt

$$H_3 \approx k \cdot \text{ld}(k)$$

Beides zusammen ergibt wegen $k = m/n$

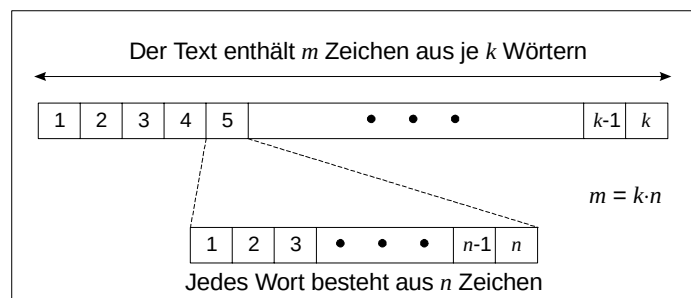
$$H_{total} \approx m \cdot \text{ld}(n) + \frac{m}{n} \cdot \text{ld} \left(\frac{m}{n} \right)$$

Dieser Wert besitzt ein Minimum für

$$m = n \cdot e^{n-1}$$

Hieraus leitet sich eine optimale Gruppenbildung im Sinne von Superzeichen nach der Tabelle ab.

Wortlänge	5,44 $\approx$ 6	22,2 $\approx$ 23	80,3 $\approx$ 81	273	891	2824	8773	26829	81031
Gruppengröße	2	3	4	5	6	7	8	9	10



Entropie von Theorien

Gute Theorie soll möglichst viele experimentelle Fakten exakt beschreiben und weitere voraussagen

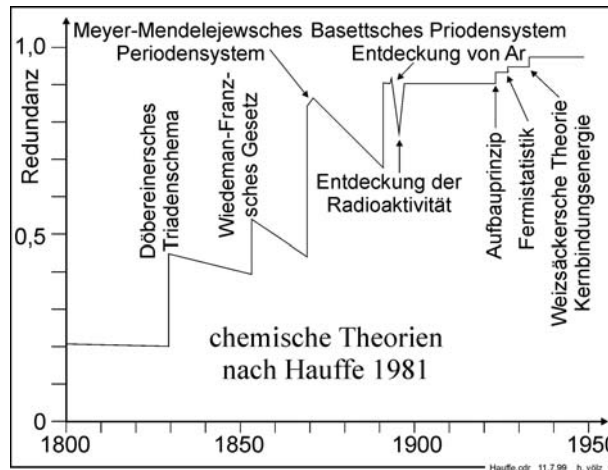
Weiterentwicklung wird Theorie immer dichter und macht alte Aussagen redundant

Chemischen Aspekt von HAUFFE am Beispiel Periodensystem untersucht:

Formal wird CARNAP-Entropie benutzt, es genügt aber die SHANNON-Entropieformel.

Benutzt werden die Eigenschaftspaare:

schwer	↔	leicht
tiefschmelzend	↔	hochschmelzend
wärmeleitend	↔	wärmeisolierend
elektrisch leitend	↔	elektrisch isolierend
leicht ionisierbar	↔	schwer ionisierbar
stabil	↔	instabil.



Information austauschbare Größen

Durch Gleichlauffehler G sind die Zeiten um $1-G$ unsicher. Dann

$$C = 2 \cdot B \cdot (1 - G) \cdot \lg(n_{AS})$$

Frequenzgang des Störabstandes berücksichtigen einzelnen Frequenzstufen Δf_i zu ersetzen:

$$C = 2 \cdot \sum_i \Delta f_i \cdot \lg(n_{AS,i})$$

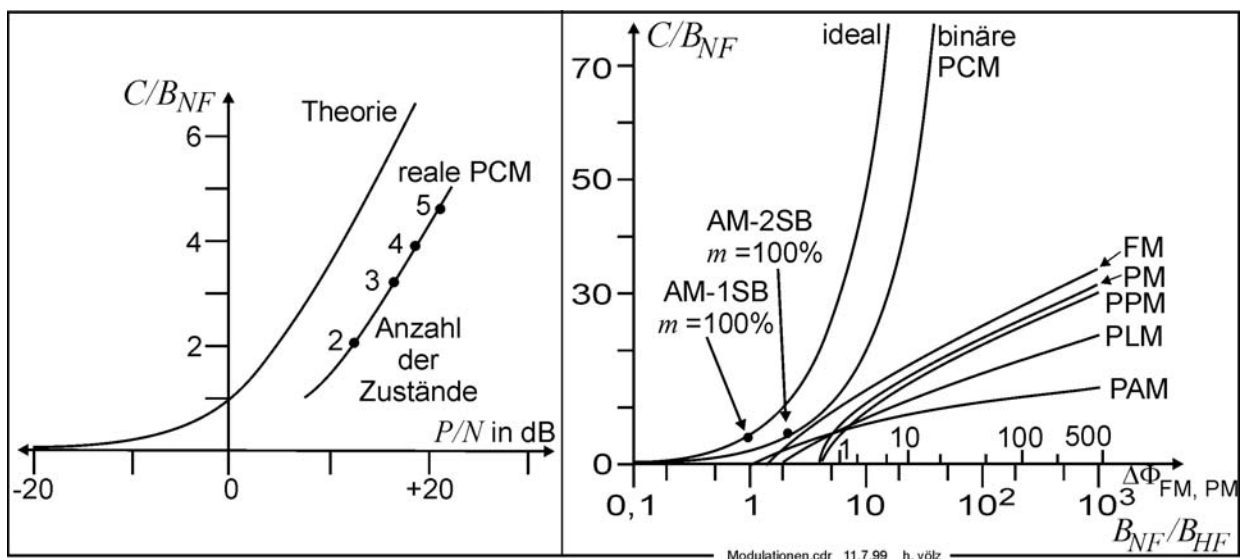
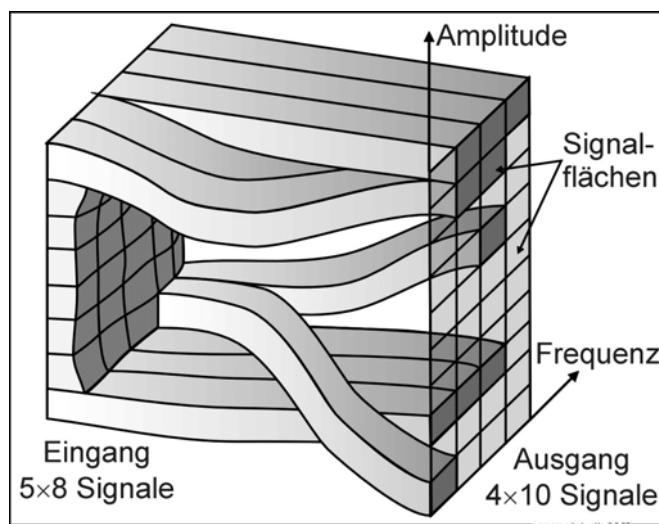
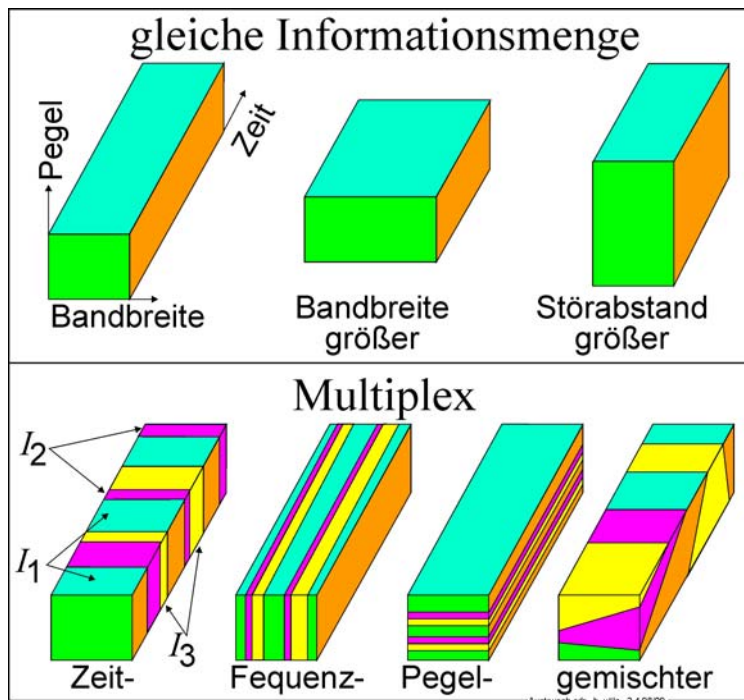
Informationsmenge I ; bestimmten Anzahl Bits oder Bytes, kontinuierlich Bandbreite B , Dynamik (Störabstand) P_N/P_S und Dauer T

$$I = T \cdot B \cdot \lg\left(1 + \frac{P_N}{P_S}\right)$$

Bei einer Zeitdauer $t_{\bar{U}}$ wird die Informationsmenge I übertragen (Störabstand $s = u_g/u_k$)

$$I = C \cdot t_{\bar{U}} = f(B, G, m, s) \cdot t_{\bar{U}}$$

Anpassungen der Signalfächen durch Dynamikregelungen, Modulationen, Codierungen usw. ermöglichen mehr oder weniger effektiven Austausch zwischen den Größen $t_{\bar{U}}$, B , G , m und s .



Energie je Bit

Nach Shannon gilt für die Kanalkapazität $C = B \cdot \lg \left(\frac{P_n + P_s}{P_n} \right)$

B Bandbreite des Kanals; P_s Leistung des Signals;

P_n Leistung des Störungen (noise), bei rein thermischen Rauschen $P_n = k \cdot B \cdot T$.

k BOLTZMANN-Konstante mit $1,381 \cdot 10^{-23}$ J/K; T absolute Temperatur

Die Signalleistung sei das z -fache der Störleistung: $P_s = z \cdot P_n \rightarrow C = B \cdot \lg(1+z)$

Für Verhältnis von Signalleistung zur Kanalkapazität gilt deshalb

$$\frac{P_s}{C} = k \cdot T \cdot \frac{z}{\lg(1+z)} \ln \frac{\text{J}}{\text{Bit}} \text{ bzw. } \frac{\text{W}}{\text{Bit/s}}$$

Der Ausdruck von z kann nun in eine Reihe entwickelt werden:

$$\frac{z}{\ln(1+z)} = \frac{1}{1 - \frac{z}{2} + \frac{z^2}{3} - \frac{z^3}{4} \pm \dots}$$

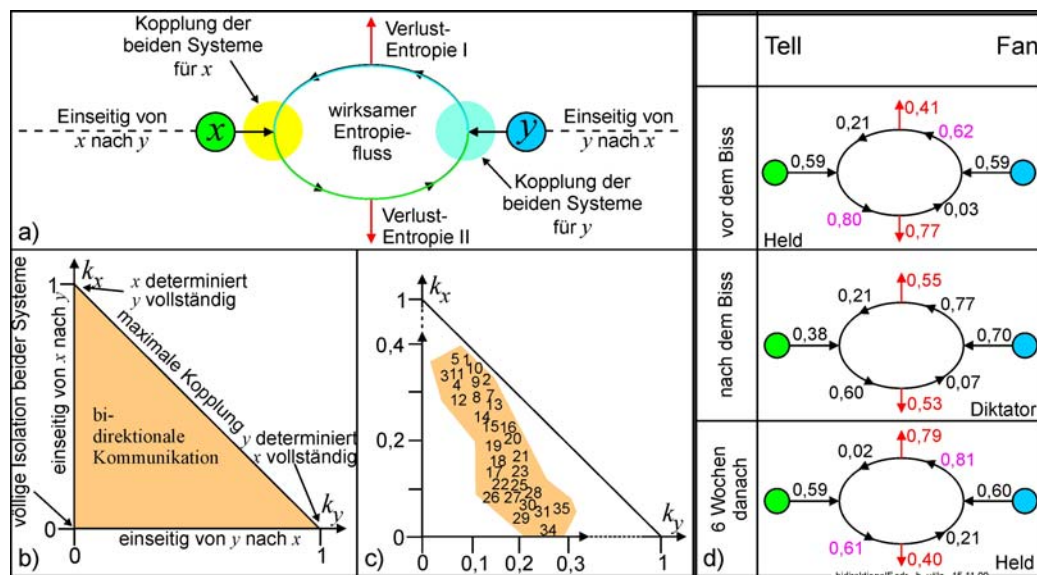
Hierfür gelten die Grenzen

$$1 < \frac{z}{\ln(1+z)} \rightarrow 1 \quad \text{für } z \rightarrow 0$$

Also gilt

$$E/\text{Bit} \geq k \cdot T \cdot \ln(2)$$

Bei 300 K ($\approx$ Zimmertemperatur) folgt $E/\text{Bit} = 3 \cdot 10^{-21}$ J $\cong 5 \cdot 10^{11}$ Hz $\cong 5 \cdot 10^{-22}$ cal $\cong 26$ mV



α -Entropie von Renyi

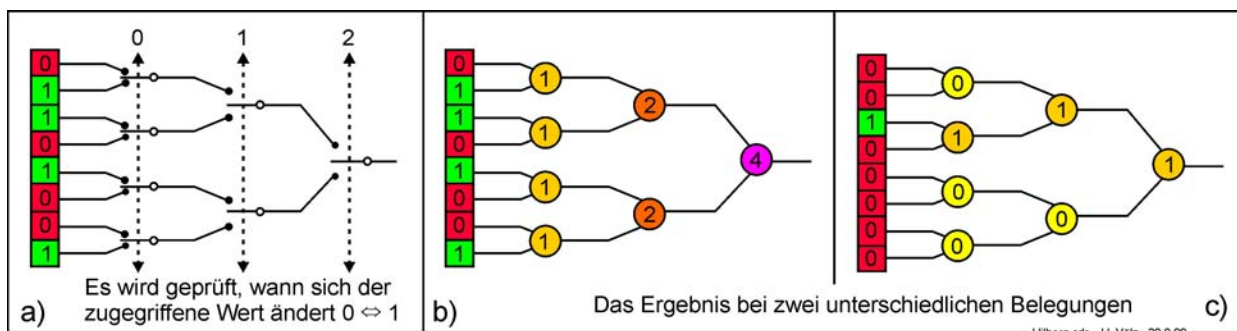
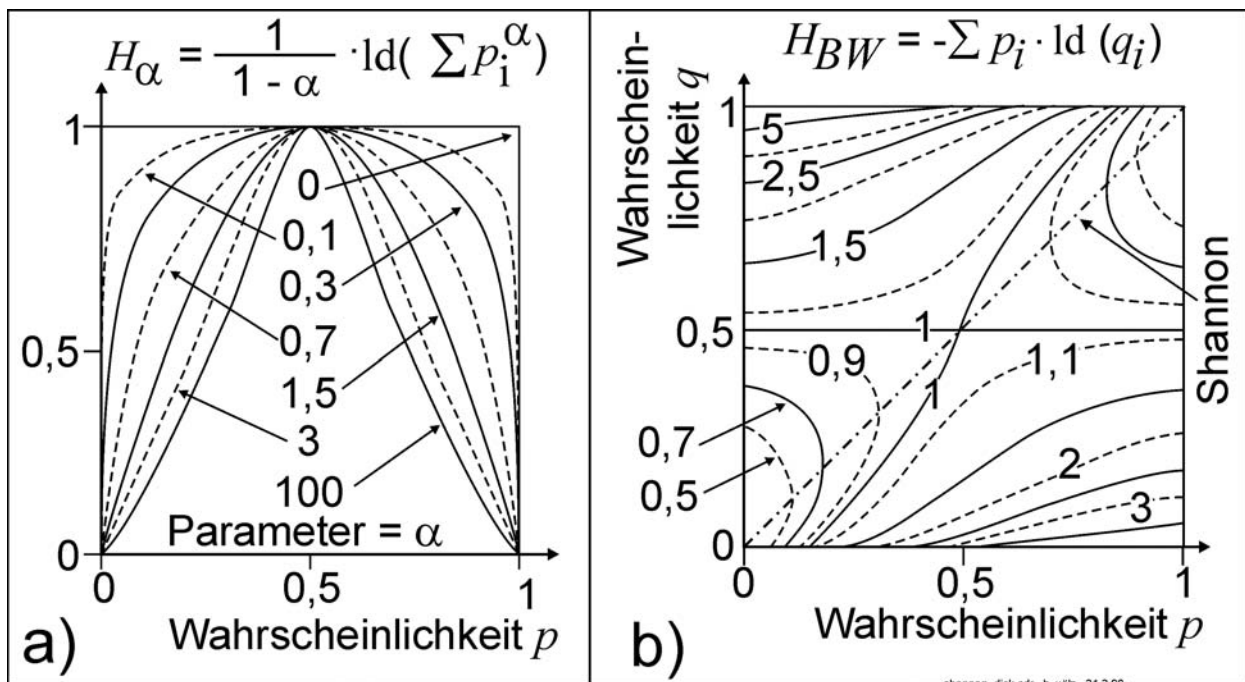
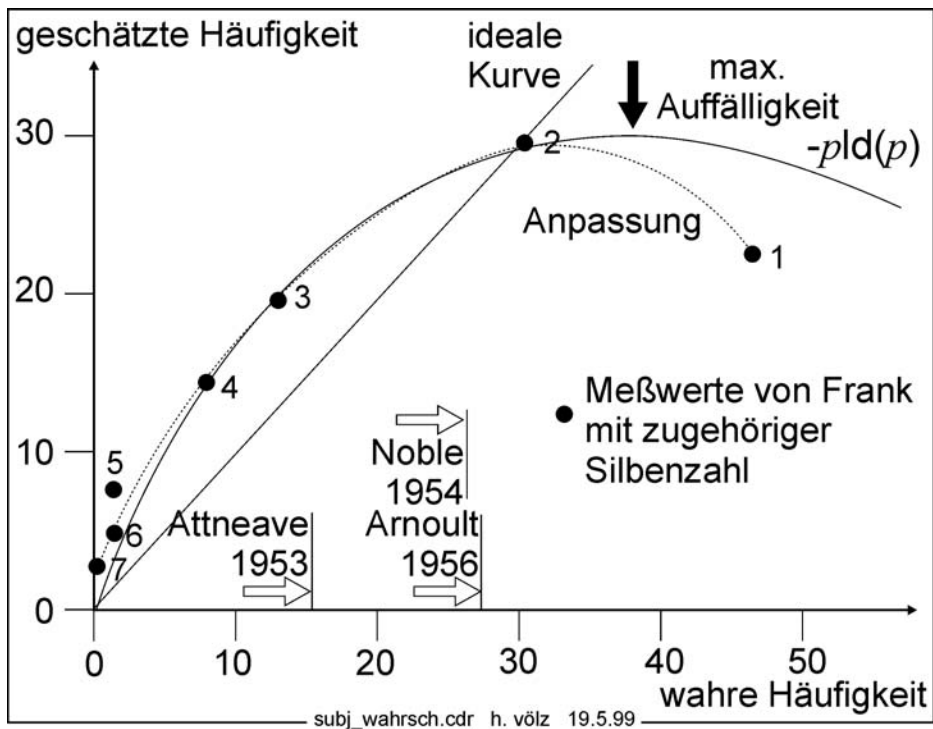
$$H_\alpha = \frac{1}{1-\alpha} \cdot \lg \left(\sum_{i=1}^n p_i^\alpha \right)$$

$\alpha = 1$ Shannon-Entropie

Bongard-Weiß-Entropie

$$H_{BW} = - \sum_{i=1}^n p_i \cdot \lg(q_i)$$

objektive Wahrscheinlichkeit p und die subjektive q



Kolmogoroff-Entropie K (1959)

K ist ein Maß für den Grad des Chaos. K gibt die Wahrscheinlichkeit dafür an, welche benachbarten Orte im nächsten Schritt erreicht werden. Es können so unterschieden werden:

- $K = 0$ deterministische Bewegung
- $K = \infty$ vollkommen stochastische Bewegung
- $K = \text{positiv und endlich}$; deterministisches Chaos; es gibt einen seltsamen Attraktor

Leider ist keine Normierung bekannt. Mit K ist aber eine Aussage zur möglichen Vorhersagezeit T gegeben:

$$T \sim \frac{1}{K} \cdot \log\left(\frac{1}{l}\right)$$

Darin ist l die geforderte Genauigkeit.

Auch für Objekte möglich: so groß, wie kleinstmöglich Beschreibung

Thermodynamische Entropie

Wahrscheinlichkeit W für die Zustände, k Boltzmann-Konstante

$$S = k \cdot \ln(W)$$

zweifach entartetes System.

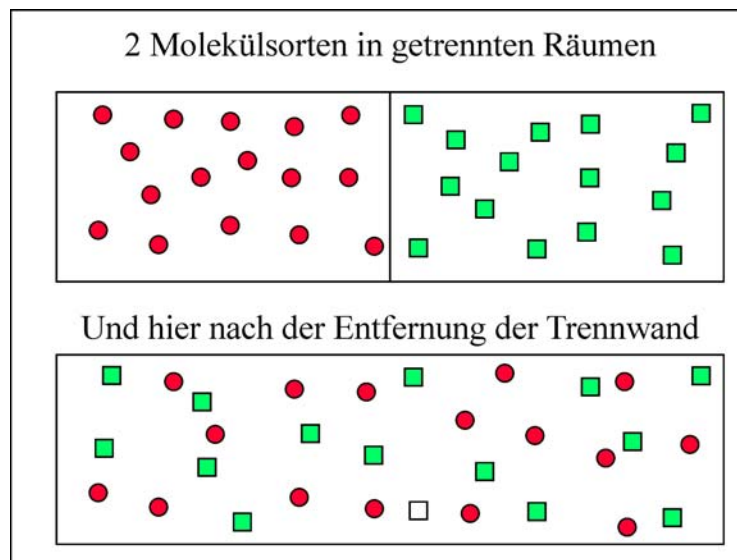
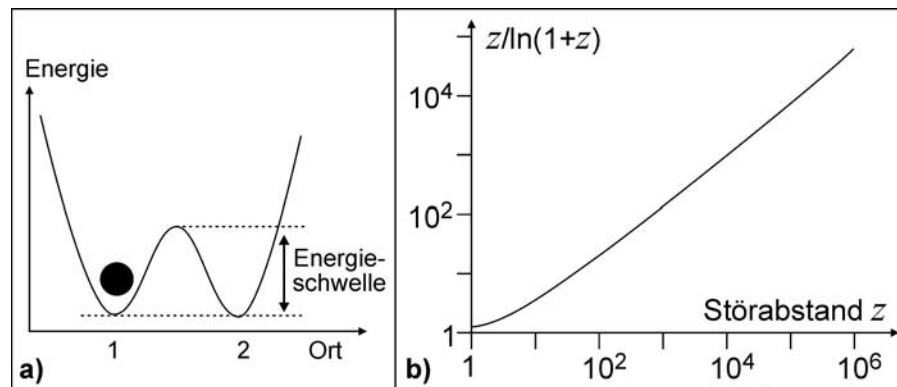
Vor Messung nicht bekannt welcher Zustand (Kugel in Energiemulde)

Wahrscheinlichkeit je $\frac{1}{2}$

Nach Messung ist Wahrscheinlichkeit 1

Für die Messung ist Energie erforderlich, errechnet sich aus der Differenz ΔS der beiden Entropien unter Multiplikation mit der absoluten Temperatur:

$$\Delta E = \Delta S \cdot T = k \cdot T \cdot [\ln(1) - \ln(0,5)] = k \cdot T \cdot \ln(2)$$



In eine Urne mit zwei Vertiefungen werden zwei Kugeln geworfen. Es sei gleichwahrscheinlich in welche Vertiefung eine Kugel beim Wurf fällt. Dann sind 3 Zustände mit den Wahrscheinlichkeiten $W_V = p_V$ möglich.



Boltzmann-Entropie

gilt für jeden Zustand und besagt, dass B am wahrscheinlichsten ist:

$$S_V = k \cdot \ln(W_V)$$

Er wird bei freier Entwicklung schließlich angenommen. Die Geschwindigkeit dorthin hängt von der absoluten Temperatur ab.

Shannon-Entropie

gibt die minimale Anzahl binärer Fragen an, die bei häufigen Versuchen erforderlich sind, um den jeweils aktuellen Zustand zu bestimmen:

$$H = - \sum_{v=1}^n p_v \cdot \lg(p_v)$$

Der Versuch ergibt einen einzigen Wert $H=1,5$, anstatt der verschiedenen S_V .

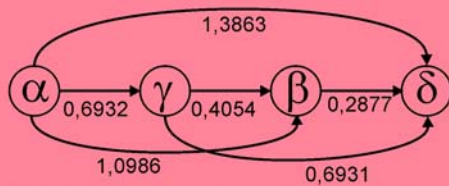
Ein anderes statistisches System kann 4 Zustände mit unterschiedlichen Wahrscheinlichkeiten annehmen:

α : 0,4 β : 0,3

γ : 0,2 δ : 0,1

$$-S_A/k = \ln(0,4) \approx 0,9163 \quad -S_B/k = \ln(0,3) \approx 1,2040$$

$$-S_C/k = \ln(0,2) \approx 1,6094 \quad -S_D/k = \ln(0,1) \approx 2,3026$$



Mögliche Übergänge und relative Entropieänderungen.

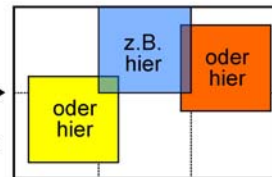
$$H = 1,8644 \text{ Bit /Zustand}$$

Entropie_thermodynamisch.odt h. völk 19.11.01/21.7.06

50 Atome



10²⁸ Jahre warten bis
sich alle Atome
in 1/6 Volumen
befinden

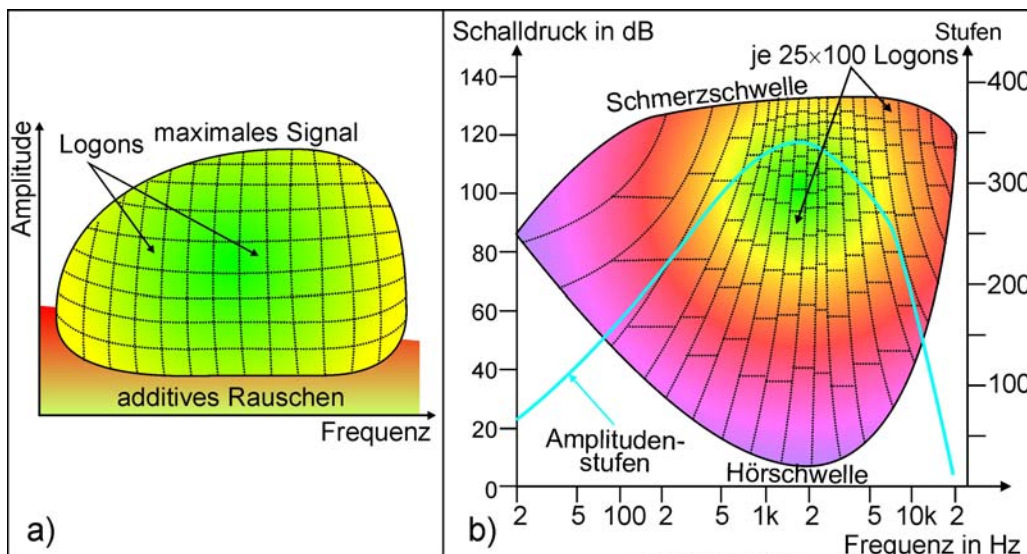


Bezüglich der Welt gilt

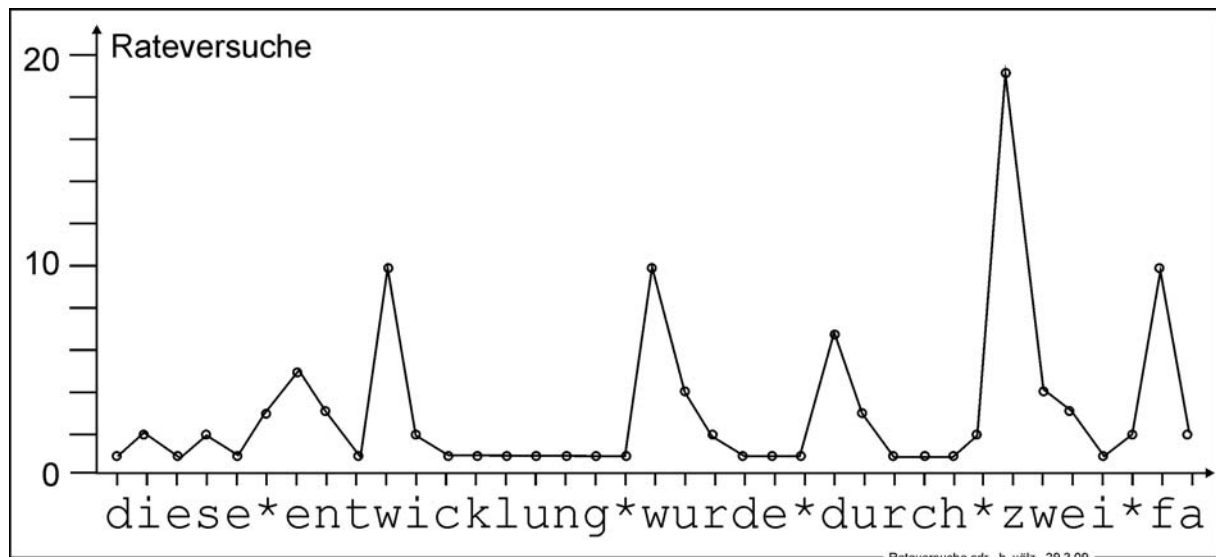
Urknall vor 1,5 · 10 ¹⁰ Jahren	heute	totale Gleichverteilung irgendwann sehr viel später
relative Entropie = 1	10 ⁸⁸ -Fach	10 ⁸⁸ · 10 ¹²³ = 10 ²⁰³ -Fach

$$S = k \cdot \ln(W) \quad \text{Wartezeit proportional} \Rightarrow W = e^{S/k}$$

Entropie_Werte.odt h. völk 5.5.06

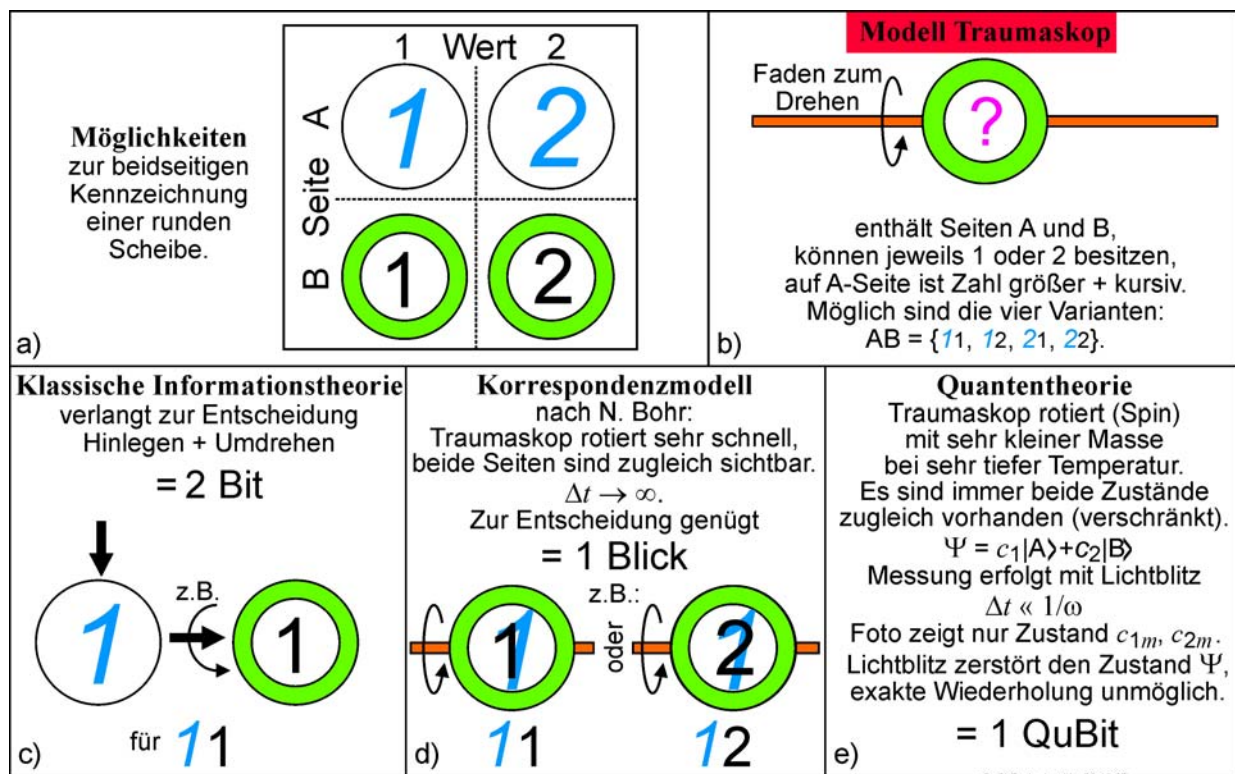
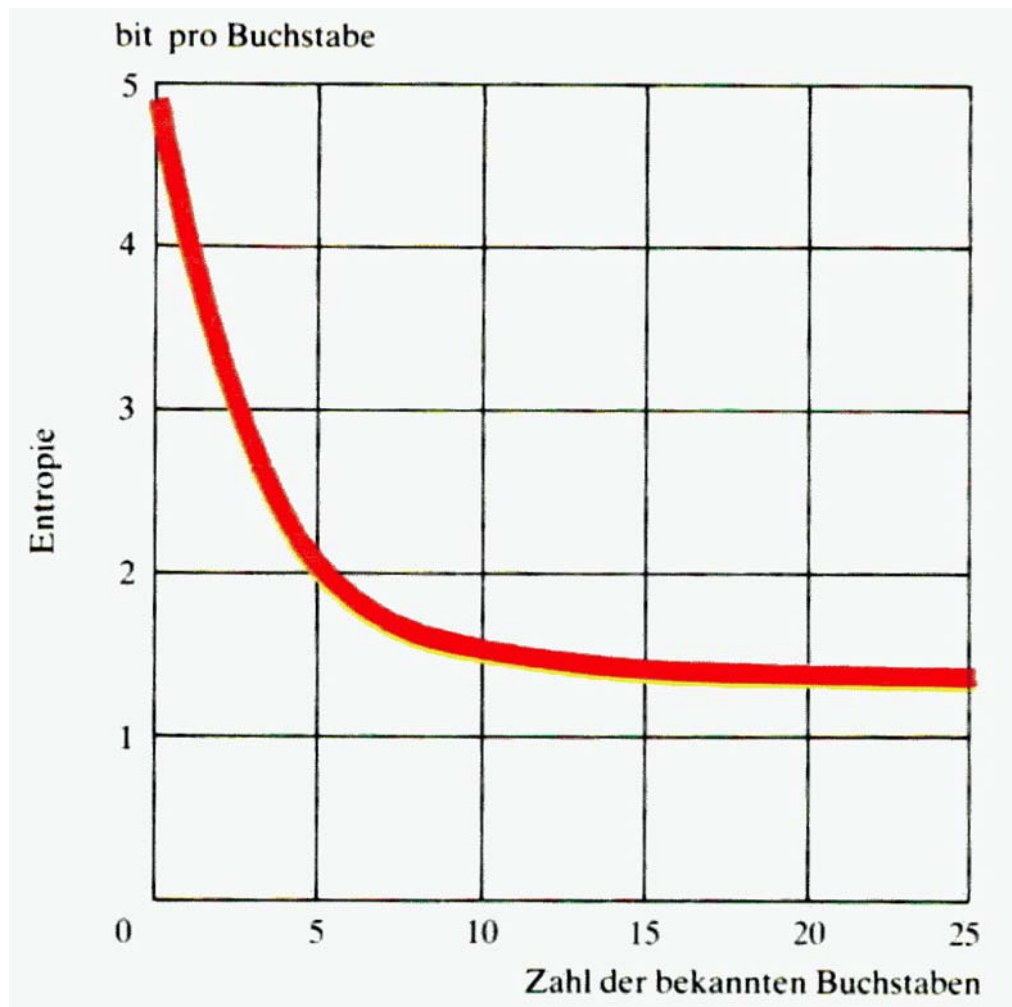


Logons.odt h. völk 11.1.95/07



	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	□	+	SUM	
A	.	.	.	.	●	.	.	.	.	.	.	●	.	●	.	.	.	.	●	●	●	●	.	.	.	.	.	.	.	538
B	.	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	.	.	.	.	.	.	.	149
C	.	.	.	.	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	230
D	●	.	.	.	●	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	397
E	.	.	.	.	.	.	.	.	●	.	.	.	●	.	.	.	.	●	●	●	●	.	.	.	.	.	●	.	.	1552
F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	.	155
G	.	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	214
H	●	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	366
I	.	.	●	.	●	.	.	.	.	.	.	.	.	●	.	.	.	.	●	●	●	●	.	.	.	.	.	.	.	679
J	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	12
K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	119
L	●	.	.	.	●	.	.	.	●	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	307
M	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	195
N	●	.	.	●	●	.	●	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	●	●	.	950
O	.	.	.	.	.	.	.	.	.	.	.	.	.	●	.	.	.	.	●	●	●	●	.	.	.	.	.	.	.	213
P	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	.	70
Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
R	●	.	.	.	●	.	.	.	●	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	●	.	656
S	.	.	●	.	●	.	.	.	●	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	.	●	.	555
T	.	.	.	.	●	.	.	.	●	.	.	.	.	.	.	.	.	.	.	●	●	●	●	.	.	.	.	●	.	521
U	.	.	.	.	●	.	.	.	.	.	.	.	.	●	.	.	.	.	●	●	●	●	.	.	.	.	.	.	.	375
V	.	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	80
W	.	.	.	.	●	.	.	.	●	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	116
X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4
Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
Z	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	110
□	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	1342	
+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	●	●	●	●	●	●	●	●	●	●	76

Σ = 9984





Geschichte zur Information

1824	CARNOT: Kreisprozess
1854	Entropie-Begriff; 2. Thermodynamischer Hauptsatz
1857	BOLTZMANN: $S = k \cdot \ln(W)$
1924	KÜPFMÜLLER: Systemtheorie
1928	R. V. L. HARTLEY: Logarithmus zwischen Symbolen und Information
1929	LEO SZILARD: Physik Zusammenhang von Information und Entropie
1933	Kotelnikow: Informationstheorie + Abtasttheorem
1940	24.3. Eingangsdatum der <i>Shannon</i> -Arbeit im JIRE
1946	GABOR Arbeiten zur Informationstheorie, logons
1948	J. W. TUKEY benennt kleinste Nachrichteneinheit „bit“ (binary digit)
1949	SHANNON: „Communication in the Presence of Noise“
1951	SHANNON Ratestrategie für Texte
1954	CARNAP-Entropie
1957	FUCKS: Analysen zu Musik, Literatur usw.
1960	LANDAUER: IBM nicht Messen sondern Löschen von Information benötigt Energie
1961	BRENNER, JACOB, MESSELSOHN Wirkung der m-RNS
1962	RENYI: α -Entropie
1963	BONGARD-WEISS-Entropie
1965	MARKO: bidirektionale Information
1987	HILBERG deterministische Informationstheorie