

Herbert Heyer

## ÜBER ASYMPTOTISCH FEHLERFREIE ÜBERTRAGBARKEIT VON INFORMATION

Der vorliegende Beitrag möge als erweiterte Synopse einer Vorlesung Aufnahme finden, welche der Autor an der Universität Tübingen für Studenten der Mathematik und mathematikbezogener Nebenfächer gehalten hat, wobei der methodische Zugang elementar, also ohne Heranziehung maßtheoretischer Hilfsmittel, gewählt wurde. Wie bekannt sind jedoch elementare Methoden nicht notwendigerweise gedanklich anspruchslos. Diese Erfahrung machen Hörer und Dozent auch am Gegenstand der Informationstheorie, sobald mit geringem mathematischen Aufwand außerhalb der Mathematik entstandene Begriffe wie Entropie oder Kapazität in ihrer Komplexität in einer für Praktiker wie Theoretiker befriedigenden Weise präzise eingeführt werden sollen. Ist die Einführung der Begriffe gelungen, stellt sich als nächstes die Aufgabe, auch die Beweise der die Theorie konstituierenden Ergebnisse elementar zu führen. Daß dies ehrgeizige Ziel erreicht werden konnte, verdankt der Verfasser den ideenreichen zeitlich dicht aufeinander folgenden Vorlagen von K. Winkelbauer ([6], [7], [8]), W. Böge ([2]) und R. Ahlswede ([1]).

Bei den so beschriebenen Anstrengungen dürfte verständlich sein, daß der Autor gelegentlich davon träumt, im Hörsaal neben wohl motivierten Studenten auch hochverehrte Kolleginnen und Kollegen anzutreffen, die in ihrer von der Mathematik eher entfernten Forschungstätigkeit der Informationstheorie begegnen und das Bedürfnis haben, ihre Vorstellungen von den Grundbegriffen vertiefen zu wollen. Eine in diesem Sinne ideale Hörerin wäre Elisabeth Walther-Bense, mit der der Autor seit mehr als zwei Dekaden über Probleme der theoretischen Ästhetik, auch über Informationsästhetik, im Gespräch ist. Und wenn sie schon keine Hörerin sein kann, so soll sie wenigstens Leserin werden... dieser Synopse, die ich ihr in freundschaftlicher Verbundenheit zueigne.

1. Skizze des Zugangs. Die Informationstheorie hat ihren Ursprung in der Nachrichtentechnik. Sie hat die mathematische Behandlung von Fragen zum Gegenstand, welche bei der Speicherung, Umformung und

Übermittlung von Nachrichten (Information) auftreten. Das Hauptproblem der Informationstheorie läßt sich wie folgt beschreiben: Gegeben sei eine aus Sender und Empfänger (Mund-Ohr, Mikrophon-Lautsprecher) zusammengesetzte Vorrichtung zur Übermittlung von Nachrichten, ein sogenannter Kanal. In dieser Vorrichtung treten Störungen statistischer Natur auf (Rauschen), die es unmöglich machen, am Ende des Kanals exakt zu bestimmen, welches die am Anfang des Kanals eingegebene Nachricht war. In der Tat wird in diesem Modell stets vorausgesetzt, daß jeder derartige Identifikationsversuch mit Fehlerwahrscheinlichkeiten belastet ist. Die sich sodann stellende Aufgabe besteht im Bemühen, die hier auftretenden Fehlerwahrscheinlichkeiten zu minimieren. Etwas genauer läßt sich diese Aufgabe so formulieren: Es geht um die Übertragung von Nachrichten, welche die Gestalt von Worten über einem (endlichen) Alphabet besitzen. Dabei wird für jede festgewählte Wortlänge eine Wahrscheinlichkeitsverteilung der Worte dieser Länge vorgegeben, die sogenannte Quellverteilung. Im Kanal entstehen aus den eingegebenen Worten wiederum Worte über einem möglicherweise weiteren (endlichen) Alphabet, die mit gewissen Übergangswahrscheinlichkeiten ausgegeben werden. Es interessiert nun die Frage, ob zu einer beliebig gewählten Zahl  $\epsilon > 0$  durch Wahl einer hinreichend großen Wortlänge und durch geeignete Verschlüsselung (Kodierung) der Nachrichten vor Eingabe in den Kanal sowie durch geeignete Entschlüsselung (Dekodierung) der Nachrichten nach Ausgabe am Kanalende erreicht werden kann, daß die Nachricht nach Entschlüsselung höchstens mit der Wahrscheinlichkeit  $\leq \epsilon$  von der Nachricht vor Verschlüsselung abweicht. Das bahnbrechende Resultat von C.E. Shannon aus dem Jahre 1948 (Siehe [ 5]) besagt, daß diese Frage für gewisse Klassen von Quellverteilungen (ergodische) und für gewisse Typen von Kanälen (stationäre ohne Vorgriff) positiv beantwortet werden kann, sofern die mittlere Entropie der Quelle echt kleiner ist als die Kapazität des Kanals, und daß die Antwort negativ ist, falls die umgekehrte Ungleichung vorliegt. Zu Beginn der 60er Jahre zeigte nun K. Winkelbauer, daß die Aussage des Satzes von Shannon für periodische, also insbesondere für stationäre Quellen richtig bleibt, und W. Böge wies später nach, daß sogar periodische und zugleich abklingende Kanäle zugelassen werden können. Allerdings erforderten diese Verallgemeinerungen die Einführung neuer Qualitätsmaßstäbe für Quellverteilungen und Kanäle. An die Stelle der mittleren

Entropie trat die allgemeiner gefaßte Sperrigkeit und an die Stelle der Kapazität das ebenfalls allgemeiner gültige Trennvermögen. Die bedeutende Leistung Shannons, die in dem ihm zugeschriebenen Hauptsatz der Informationstheorie zum Ausdruck kommt, ist die zum Zeitpunkt der Entdeckung überraschende Einsicht, daß eine einzige reelle Skala, also eine total geordnete Menge, genügt, um auf ihr vermöge Sperrigkeits- und Trennrates die Übertragbarkeit bzw. Nichtübertragbarkeit von Nachrichten ablesen zu können.

2. Der mathematische Rahmen. Die für das Verständnis der nachfolgenden Ausführungen notwendigen mathematischen Hilfsmittel beziehen sich im wesentlichen auf die Begriffe der Wahrscheinlichkeitsverteilung auf einer nichtleeren endlichen Menge  $X$  sowie des stochastischen (Markoff) Kerns von  $X$  nach einer weiteren endlichen Menge  $Y$ . Beide Begriffe sind grundlegend für jede Einführung in die Elemente der (diskreten) Wahrscheinlichkeitstheorie, wovon man sich etwa in den handlichen Schriften [3] und [4] überzeugen kann. Für die Gesamtheit der Wahrscheinlichkeitsverteilungen auf  $X$  benützt man die Abkürzung  $M(X)$ . Ein Kanal  $Q$  von  $X$  nach  $Y$  (in Zeichen  $Q: X \rightarrow Y$ ) ist eine Abbildung  $x \mapsto Q_x$  von  $X$  in  $M(X)$ . Im Modell der Informationstheorie repräsentieren die Mengen  $X$  und  $Y$  das Eingangs- bzw. Ausgangsalphabet, und das Rauschen im Kanal drückt sich darin aus, daß für jede Menge  $A \subset Y$  die Zahl  $Q_x(A)$  gerade die Wahrscheinlichkeit dafür angibt, daß der empfangene Buchstabe in  $A$  liegt, wenn der Buchstabe  $x \in X$  gesendet wurde. Man spricht dann zusammenfassend von der (stochastischen) Matrix  $(Q_x(\{y\}))_{x \in X, y \in Y}$  der Übergangswahrscheinlichkeiten  $Q_x(\{y\})$ , zumal per definitionem

$$\sum_{y \in Y} Q_x(\{y\}) = 1$$

gilt. Als nächstes wird zu gegebener Quellverteilung  $P \in M(X)$  eine neue Wahrscheinlichkeitsverteilung  $P * Q$  auf  $X \times Y$  (also ein Element aus  $M(X \times Y)$ ) durch die Festsetzung

$$P * Q(\{(x, y)\}) := P(\{x\}) Q_x(\{y\})$$

für alle  $(x, y) \in X \times Y$  erklärt. Ihr Wert an der Stelle  $(x, y)$  repräsentiert die Wahrscheinlichkeit dafür, daß  $x \in X$  bei gegebener Quellverteilung  $P$  gesendet und  $y \in Y$  empfangen wird.

Kanäle  $Q: X \rightarrow Y$  und  $R: Y \rightarrow Z$  (für weitere nichtleere endliche Alphabete  $Y, Z$ ) lassen sich komponieren zu einem Kanal  $QR: X \rightarrow Z$  gemäß

$$(QR)_x(\{z\}) := \sum_{y \in Y} Q_x(\{y\}) R_y(\{z\}),$$

wenn immer  $x \in X$  und  $z \in Z$ . Ein für die Anwendungen besonders wichtiger Typ von Kanälen sind die *deterministischen Kanäle*, die von (Punkt) Abbildungen erzeugt werden. Genauer gesagt ist der von einer Abbildung  $f: X \rightarrow Y$  erzeugte Kanal  $[f]: X \rightarrow Y$  definiert durch die Festlegung

$$[f]_x(\{y\}) := \begin{cases} 1, & \text{falls } y=f(x) \\ 0, & \text{falls } y \neq f(x). \end{cases}$$

Man nennt in dieser Situation  $f$  den (deterministischen) Erzeuger des Kanals  $[f]$ . Das Vorschalten bzw. Nachschalten deterministischer Kanäle ermöglicht, wie später auszuführen sein wird, die Kodierung bzw. Dekodierung von Nachrichten.

3. Fehlerwahrscheinlichkeit. Zu gegebener Quellverteilung  $P \in M(X)$  und gegebenem Kanal  $Q: X \rightarrow Y$  erklärt man die *Fehlerwahrscheinlichkeit* (bzgl.  $P$  und  $Q$ ) als die Zahl

$$\delta(P, Q) := (P * Q) (\{(x, y) \in X \times Y: x \neq y\}).$$

Ist speziell  $Y \subset X$ , so gibt  $\delta(P, Q)$  die Wahrscheinlichkeit dafür an, daß der empfangene Buchstabe vom gesendeten verschieden ist und somit ein Fehler vorliegt.

Sind nun weitere Mengen  $Z$  und  $U$  vorgelegt, so zeigt man

3.1 daß deterministische Kanäle  $[f]: X \rightarrow Y$  und  $[g]: Z \rightarrow U$  mit Erzeugern  $f: X \rightarrow Y$  bzw.  $g: Z \rightarrow U$  existieren, so daß für alle Kanäle  $Q': X \rightarrow Y$  und  $Q'': Z \rightarrow U$  die Ungleichung

$$\delta(P, [f]Q[g]) \leq \delta(P, Q'Q'')$$

erfüllt ist. Hieraus folgert man

3.2 daß deterministische Kanäle bzgl. der Fehlerwahrscheinlichkeit optimal sind und daß das über alle Kanäle  $Q': X \rightarrow Y$  und  $Q'': Z \rightarrow U$  erstreckte Infimum

$$\varepsilon(P, Q) := \inf \delta(P, Q'QQ')$$

für deterministische Kanäle angenommen wird.

Wird also  $U:=X$  gesetzt, so existieren Abbildungen  $f:X \rightarrow Y$  und  $g:Z \rightarrow X$  derart, daß

$$\varepsilon(P, Q) = \delta([f]Q[g]).$$

Mit den Bezeichnungen *Verschlüsselung* (Kodierung) und *Entschlüsselung* (Dekodierung) für die Abbildungen  $f$  bzw.  $g$  besagt diese Gleichheit, daß zu den Vorgaben einer Quellverteilung  $P \in \mathcal{M}(X)$  und eines Kanals  $Q: X \rightarrow Y$  Kodierung und Dekodierung (deterministisch) so gewählt werden können, daß die Wahrscheinlichkeit für Übertragungsfehler minimal wird.

4. Sperrigkeit einer Quellverteilung. Es sei  $P \in \mathcal{M}(X)$  eine Quellverteilung. Für jede Wahl von  $\lambda \in [0, 1]$  heißt die Zahl

$$Sp_\lambda(P) := \min \{ |A| : P(A) \geq 1 - \lambda \},$$

wobei das Minimum über alle Teilmengen  $A$  von  $X$  erstreckt wird, die  $\lambda$ -Sperrigkeit ( $\lambda$ -Kardinalität) von  $P$ .

Zunächst ist die Abbildung  $\lambda \mapsto Sp_\lambda(P)$  von  $[0, 1]$  in die Menge  $\mathbb{Z}_+$  der nichtnegativen ganzen Zahlen monoton fallend, und es gilt

$$4.1 \quad Sp_1(P) = 0 \quad \text{sowie}$$

$$4.2 \quad Sp_\lambda(P) \geq 1 \quad \text{für alle } \lambda \in [0, 1[.$$

Offenbar ist  $Sp_\lambda(P)$  ein Maß für die Unbestimmtheit, mit der die Erzeugung von Buchstaben aus  $X$  mittels der Quellverteilung  $P$  belastet ist. Über das Resultat einer solchen Erzeugung läßt sich dann am wenigsten voraussagen, wenn  $P$  die Gleichverteilung  $P_1$  auf  $X$  ist, welche durch

$$P(\{x\}) := \frac{1}{|X|}$$

für alle  $x \in X$  erklärt wird. Andererseits ist der Ausgang vollständig bestimmt, wenn  $P$  eine Dirac-Verteilung  $P_2$  auf  $X$  ist, d.h. wenn es ein  $x_0 \in X$  gibt, so daß

$$P(\{x\}) = \begin{cases} 1 & \text{für } x=x_0 \\ 0 & \text{für } x \neq x_0 \end{cases}$$

gilt. Man erhält somit,

$$Sp_\lambda(P_2) \leq Sp_\lambda(P) \leq Sp_\lambda(P_1),$$

und es ist leicht einzusehen, daß  $Sp_\lambda(P)$  gleichmäßig maximal für  $\lambda \in [0,1]$  genau dann ist, wenn  $P=P_1$  gilt, und gleichmäßig minimal für  $\lambda \in [0,1]$  genau dann wird, wenn  $P=P_2$  vorliegt. Diese letzten beiden Äquivalenzen lassen ohne Schwierigkeit den Zusammenhang zwischen der Sperrigkeit und dem in der Informationstheorie historisch etablierten Begriff der Entropie erkennen, welcher als Maß für die Unordnung in der Quelle bzw. für die Ausbreitung der Quellverteilung Verwendung findet. Dabei wird für jede Quellverteilung  $PEM(X)$  die Entropie von  $P$  als die Zahl

$$H(P) := - \sum_{x \in X} P(\{x\}) \log P(\{x\})$$

eingeführt. Es gilt stets  $0 \leq H(P) \leq \log |X|$ , und den oben formulierten Äquivalenzen entsprechen die Aussagen, daß  $H(P)$  genau dann maximal (also  $= \log |X|$ ) ist, wenn  $P$  mit  $P_1$  übereinstimmt, und genau dann minimal (also  $= 0$ ) ist, wenn  $P=P_2$  vorliegt.

5. Trennvermögen eines Kanals. Nun sei  $Q: X \rightarrow Y$  ein Kanal. Für jede Wahl von  $\lambda \in [0,1]$  heißt die Zahl

$$Tr_\lambda(Q) := \max\{|M| : Q_{Y_m}(Z_m) \geq 1-\lambda \text{ für alle } m \in M\},$$

wobei die Ungleichung in Klammern an eine (als existent geforderte) Folge  $(y_m)_{m \in M}$  von Punkten in  $Y$  und eine (ebenfalls als existent geforderte) disjunkte Folge  $(Z_m)_{m \in M}$  von Mengen in  $Z$  geknüpft ist, das  $\lambda$ -Trennvermögen ( $\lambda$ -Diszernibilität) von  $Q$ .

Die Abbildung  $\lambda \mapsto Tr_\lambda(Q)$  von  $[0,1]$  in die Menge  $\overline{\mathbb{N}}$  der natürlichen Zahlen einschließlich des Symbols  $+\infty$  ist offensichtlich monoton steigend, und es gilt

5.1  $Tr_1(Q) = +\infty$  sowie

5.2  $1 \leq Tr_\lambda(Q) < +\infty$ , wenn immer  $\lambda \in [0,1[$ .

Das  $\lambda$ -Trennvermögen nimmt ab, wenn man von einem Kanal  $Q:Y \rightsquigarrow Z$  zur Linkskomposition  $Q'Q$  mit einem Kanal  $Q':X \rightsquigarrow Y$  übergeht. Beim Übergang von einem Kanal  $Q':X \rightsquigarrow Y$  zur Rechtskomposition  $Q'Q$  mit einem Kanal  $Q:Y \rightsquigarrow Z$  geht diese Aussage im allgemeinen verloren. Im Spezialfall deterministischer Kanäle mit Erzeugern  $f:X \rightarrow Y$  bzw.  $g:Z \rightarrow U$  gilt sie jedoch in der Form

$$\text{Tr}_\lambda(Q) \geq \text{Tr}_\lambda([f]Q[g]).$$

In dieser Ungleichung tritt Gleichheit auf, falls  $f$  als surjektiv und zugleich  $g$  als injektiv vorausgesetzt wird.

Der Zusammenhang zwischen dem hier betrachteten Trennvermögen und der in der traditionellen Darstellung der Informationstheorie grundlegend verwendeten Kapazität kann hier nur unvollständig beschrieben werden. Zunächst benötigt man den Begriff der *bedingten Entropie*  $H(\phi|\psi;P)$ , welche für Quellverteilungen  $P \in M(X)$  und Abbildungen (Messungen)  $\phi: X \rightarrow Y$ ,  $\psi: X \rightarrow Z$  definiert wird.  $H(\phi|\psi;P)$  beschreibt die Unbestimmtheit der Messung  $\phi$  für den Fall, daß die Messung  $\psi$  bekannt ist. Im Falle, daß  $\psi$  gleich der konstanten Abbildung auf  $X$  ist, schreibt man  $H(\phi;P)$  anstelle von  $H(\phi|\psi;P)$ . Für diese beiden bedingten Entropien gilt die Shannonsche Ungleichung

$$H(\phi|\psi;P) \leq H(\phi;P),$$

so daß sich sogleich die Frage nach Shannonscher Gleichheit stellt. Zunächst wird die Zahl

$$R(\phi|\psi;P) := H(\phi;P) - H(\phi|\psi;P)$$

die (Größe der) *Information über  $\phi$  bei Messung von  $\psi$*  genannt. Durch diese Differenzbildung wird die der Messung  $\phi$  anhaftende Unbestimmtheit  $H(\phi;P)$  um  $H(\phi|\psi;P)$  verringert. Verringerung der Unbestimmtheit ist aber Informationsgewinn. Also ist  $R(\phi|\psi;P)$  als *entropische Information über  $\phi$  durch  $\psi$*  motiviert. Übrigens verschwindet diese entropische Information genau dann, wenn die Messungen  $\phi$  und  $\psi$  (stochastisch) unabhängig sind.

Im weiteren seien  $P$  eine Quellverteilung auf  $X$  und  $Q$  ein Kanal  $X \rightsquigarrow Y$ . Mit  $\phi$  bzw.  $\psi$  mögen die Projektionen  $X \times Y \rightarrow X$  bzw.  $X \times Y \rightarrow Y$  gemeint

sein. Dann wird die *Kapazität* von  $Q$  als die Zahl

$$C(Q) := \sup\{R(\phi|\psi; P*Q) : PEM(X)\}$$

erklärt, wobei  $R(\phi|\psi; P*Q)$  in diesem Fall diejenige Information beschreibt, welche durch den Kanal  $Q$  im Mittel (pro Buchstabe) vom Quellalphabet  $X$  zum Alphabet  $Z$  übertragen wird. Analog zum Trennvermögen nimmt die Kapazität ab, wenn man zu Links- oder Rechtskompositionen übergeht. Unter Heranziehung der Kompaktheit der Menge  $M(X)$  und der Stetigkeit der Abbildung  $P \mapsto R(\phi|\psi; P*Q)$  wird das Supremum in der Definition von  $C(Q)$  tatsächlich angenommen, dh. es existiert (zu  $Q$ ) ein  $P_0 \in M(X)$  mit der Eigenschaft

$$C(Q) = R(\phi|\psi; P_0*Q).$$

Schließlich können wir ein Resultat formulieren, das den angekündigten Zusammenhang zwischen  $Tr_\lambda(Q)$  und  $C(Q)$  erkennen läßt: Mit der Abkürzung  $h$  für die Funktion  $x \mapsto -x \log x - (1-x) \log(1-x)$  auf  $]0,1[$ , welche ihr Maximum ( $=\log 2$ ) genau für  $x = \frac{1}{2}$  annimmt, gilt nämlich für alle  $\lambda \in ]0,1[$

$$Tr_\lambda(Q) \leq \exp \frac{C(Q) + \log 2}{1-\lambda}.$$

6. Der Hauptsatz über asymptotisch fehlerfreie Übertragung. Zunächst wird eine erste Ungleichung zwischen Sperrigkeit und Trennvermögen mit Abschätzungen der Fehlerwahrscheinlichkeit in Beziehung gesetzt. Es gilt der

6.1 Satz Bei gegebener Quellverteilung  $PEM(X)$  und gegebenem Kanal  $Q: X \rightarrow Y$  hat man für alle  $\lambda_1, \lambda_2 \in [0,1]$  die Gültigkeit der Implikationen

$$(a) \epsilon(P, Q) \leq \lambda_1 \lambda_2 \Rightarrow Sp_{\lambda_1}(P) \leq Tr_{\lambda_2}(Q) \quad \text{sowie}$$

$$(b) Sp_{\lambda_1}(P) \leq Tr_{\lambda_2}(Q) \Rightarrow \epsilon(P, Q) \leq \lambda_1 + \lambda_2.$$

Ziel der nachfolgenden Ausführungen ist die Beschreibung der Verringerung der "mittleren Fehler"  $\epsilon$  durch Übertragung längerer Worte. Dazu müssen variable Alphabete eingeführt werden.

Für jedes  $n \geq 1$  seien  $X^{(n)}, Y^{(n)}, Z^{(n)}, \dots$  nichtleere endliche Mengen. Zum Beispiel sei  $X^{(n)} := X^n$  das  $n$ -fache Produkt von  $X$  mit sich selbst,



welches die Menge der Worte der Länge  $n$  in Buchstaben aus  $X$  repräsentiert. Weiterhin seien für jedes  $n \geq 1$   $P^{(n)}$  eine Quellverteilung auf  $X^{(n)}$  (also  $\in M(X^{(n)})$ ) und  $Q^{(n)}$  ein Kanal  $Y^{(n)} \rightarrow Z^{(n)}$ . Es werden die Abkürzungen  $\mu := (P^{(n)})_{n \geq 1}$  und  $\nu := (Q^{(n)})_{n \geq 1}$  für die zugehörigen Folgen von Quellverteilungen bzw. Kanälen gewählt.  $\mu$  heißt dann *asymptotisch fehlerfrei durch  $\nu$  übertragbar* (in Zeichen:  $\mu \in T(\nu)$ ), falls

$$\lim_{n \rightarrow \infty} \varepsilon(P^{(n)}, Q^{(n)}) = 0.$$

Dabei ist die an früherer Stelle eingeführte Fehlergröße  $\varepsilon(P, Q)$  nunmehr für die Daten  $X := X^{(n)}$ ,  $Y := Y^{(n)}$ ,  $Z := Z^{(n)}$  und  $U := X^{(n)}$  auszuwerten.

**6.2 Satz** Mit den verabredeten Bezeichnungen  $\mu$  und  $\nu$  sind nachstehende Aussagen äquivalent:

(i)  $\mu \in T(\nu)$ .

(ii) Zu jedem  $\lambda \in ]0, 1[$  existiert ein  $n_0 := n_0(\lambda) \geq 1$  derart, daß für alle  $n \geq n_0$  die Ungleichung

$$Sp_\lambda(P^{(n)}) \leq Tr_\lambda(Q^{(n)})$$

gilt.

Im Spezialfall von Produktalphabeten  $X^{(n)} := X^n$ ,  $Y^{(n)} := Y^n$  und  $Z^{(n)} := Z^n$  für  $n \geq 1$  erhält man offenbar die Schranken  $Sp_\lambda(P^{(n)}) \leq |X|^n$  (bei beliebigem  $\lambda \in ]0, 1[$ ) und  $Tr_\lambda(Q^{(n)}) \leq |Z|^n$  (sofern  $\lambda \in ]0, 1[$ ). Um die "durchschnittliche Unbestimmtheit" verstehen zu können, mit der die Wahl des einzelnen Buchstaben behaftet ist, liegt eine Mittelung der Gestalt

$$\log Sp_\lambda(P^{(n)}) = O(n)$$

(für  $\lambda \in ]0, 1[$ ) nahe. Deshalb ist es ratsam, in dieser Situation die *obere* bzw. *untere Sperrigkeitsrate*  $\overline{sp}(\mu)$  bzw.  $\underline{sp}(\mu)$  einer Quellverteilungsfolge  $\mu$  in folgenden zwei Schritten einzuführen:

(1) Für jedes  $\lambda \in ]0, 1[$  sei

$$\overline{sp}_\lambda(\mu) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log Sp_\lambda(P^{(n)}) \text{ bzw.}$$

$$\underline{sp}_\lambda(\mu) := \liminf_{n \rightarrow \infty} \frac{1}{n} \log Sp_\lambda(P^{(n)})$$

Da die auf  $[0,1]$  definierten Funktionen  $\lambda \mapsto \overline{\text{sp}}_\lambda(\mu)$  und  $\lambda \mapsto \underline{\text{sp}}_\lambda(\mu)$  monoton fallen (mit  $\underline{\text{sp}}_1(\mu) = \overline{\text{sp}}_1(\mu) = -\infty$ ), existieren die Limiten

$$(2) \quad \overline{\text{sp}}(\mu) := \lim_{\lambda \rightarrow 0^+} \overline{\text{sp}}_\lambda(\mu) \text{ bzw.}$$

$$\underline{\text{sp}}(\mu) := \lim_{\lambda \rightarrow 0^+} \underline{\text{sp}}_\lambda(\mu)$$

und verdienen obige Bezeichnung.

In analoger Weise werden im Spezialfall von Produktalphabeten die obere bzw. untere Trennrate  $\overline{\text{tr}}(v)$  bzw.  $\underline{\text{tr}}(v)$  einer Kanalfolge  $v$  eingeführt. Man setzt für jedes  $\lambda \in [0,1]$

$$(3) \quad \overline{\text{tr}}_\lambda(v) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log \text{Tr}_\lambda(Q^{(n)}) \text{ bzw.}$$

$$\underline{\text{tr}}_\lambda(v) := \liminf_{n \rightarrow \infty} \frac{1}{n} \log \text{Tr}_\lambda(Q^{(n)}),$$

beobachtet, daß die auf  $[0,1]$  erklärten Funktionen  $\lambda \mapsto \overline{\text{tr}}_\lambda(v)$  und  $\lambda \mapsto \underline{\text{tr}}_\lambda(v)$  monoton steigen (wobei  $\underline{\text{tr}}_1(v) = \overline{\text{tr}}_1(v) = +\infty$  gilt) und schließt auf die Existenz der Limiten

$$(4) \quad \overline{\text{tr}}(v) := \lim_{\lambda \rightarrow 0^+} \overline{\text{tr}}_\lambda(v) \text{ und}$$

$$\underline{\text{tr}}(v) := \lim_{\lambda \rightarrow 0^+} \underline{\text{tr}}_\lambda(v).$$

Nachdem Spreizigkeits- und Trennrate als reelle Zahlen zur Verfügung stehen, deren Anordnung über die in Konkurrenz stehenden Qualitäten von Quelle und Kanal Auskunft gibt, ist der

### 6.3 Hauptsatz der Informationstheorie übersichtlich zu formulieren.

Es werden die Bezeichnungen  $\mu$  und  $v$  für die im Spezialfall von Produktalphabeten vorgegebenen Quellverteilungs- bzw. Kanalfolgen beibehalten. Dann gelten nachstehende Implikationen

$$(i) \quad \overline{\text{sp}}(\mu) < \underline{\text{tr}}(v) \Rightarrow \mu \in \text{T}(v).$$

$$(ii) \quad \overline{\text{sp}}(\mu) > \overline{\text{tr}}(v) \Rightarrow \mu \notin \text{T}(v).$$

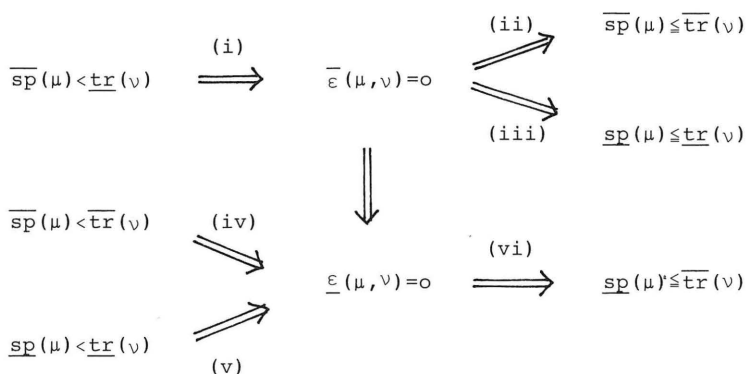
$$(iii) \quad \underline{\text{sp}}(\mu) > \underline{\text{tr}}(v) \Rightarrow \mu \notin \text{T}(v).$$

$$(iv) \quad \overline{\text{sp}}(\mu) < \overline{\text{tr}}(v) \Rightarrow \liminf_{n \rightarrow \infty} \varepsilon(P^{(n)}, Q^{(n)}) = 0.$$

$$(v) \quad \underline{\text{sp}}(\mu) < \underline{\text{tr}}(v) \Rightarrow \liminf_{n \rightarrow \infty} \varepsilon(P^{(n)}, Q^{(n)}) = 0.$$

(vi)  $\underline{\text{sp}}(\mu) > \overline{\text{tr}}(\nu) \Rightarrow \liminf_{n \rightarrow \infty} \varepsilon(P^{(n)}, Q^{(n)}) > 0$ .

Der Beweis des Hauptsatzes, der mit Hilfe der oben zitierten Resultate 6.1 und 6.2 geführt werden soll, verläuft günstigerweise nach folgendem Schema, in welches noch die Abkürzungen  $\bar{\varepsilon}(\mu, \nu)$  und  $\underline{\varepsilon}(\mu, \nu)$  für  $\limsup_{n \rightarrow \infty} \varepsilon(P^{(n)}, Q^{(n)})$  bzw.  $\liminf_{n \rightarrow \infty} \varepsilon(P^{(n)}, Q^{(n)})$  Eingang finden, womit im Falle  $\underline{\varepsilon}(\mu, \nu) = \bar{\varepsilon}(\mu, \nu)$  sogleich  $\varepsilon(\mu, \nu)$  erklärt ist.



Der Parallelität der Beweise der Aussagen (ii), (iii) bzw. (iv), (v) wegen zeigen wir nur (i), (ii), (iv) und (vi).

Ad (i). Sei  $\lambda \in ]0, 1]$ . Dann gilt nach Voraussetzung  $\overline{\text{sp}}_\lambda(\mu) < \underline{\text{tr}}_\lambda(\nu)$ , folglich gibt es eine natürliche Zahl  $n_0 := n_0(\lambda) \geq 1$ , so daß für alle  $n \geq n_0$

$$\frac{1}{n} \log \text{Sp}_\lambda(P^{(n)}) < \frac{1}{n} \log \text{Tr}_\lambda(Q^{(n)})$$

bzw.

$$\text{sp}_\lambda(P^{(n)}) < \text{tr}_\lambda(Q^{(n)})$$

vorliegt. Satz 7.2 liefert sodann die Behauptung  $\varepsilon(\mu, \nu) = 0$ , dh.  $\mu \in T(\nu)$ .

Ad (ii). Es sei eine positive reelle Zahl  $c$  mit  $\overline{\text{sp}}(\mu) > c > \overline{\text{tr}}(\nu)$  gewählt. Nach den Definitionen (2) und (4) existiert ein  $\lambda_0 \in ]0, 1]$ , so

daß die Ungleichungen  $\overline{sp}_{\lambda_0}(\mu) > c > \overline{tr}_{\lambda_0}(\nu)$  gelten. Nun gibt es gemäß (3) ein  $n_0 \geq 1$ , so daß für alle  $n \geq n_0$

$$\frac{1}{n} \log \text{Tr}_{\lambda_0}(Q^{(n)}) < c,$$

und gemäß (1) eine streng monoton steigende Folge  $(n_j)_{j \geq 0}$  natürlicher Zahlen, so daß

$$\frac{1}{n_j} \log \text{Sp}_{\lambda_0}(P^{(n_j)}) > c,$$

mithin

$$\text{Sp}_{\lambda_0}(P^{(n_j)}) > \text{Tr}_{\lambda_0}(Q^{(n_j)})$$

für alle  $j \geq 0$  erfüllt ist. Anwendung von Satz 6.2 führt auch in diesem Fall zur Behauptung  $\bar{\epsilon}(\mu, \nu) > 0$ , dh.  $\mu \notin T(\nu)$ .

Ad (iv). Wie zu (ii) zeigt man zunächst, daß eine streng monoton fallende Nullfolge  $(\lambda_k)_{k \geq 1}$  derart existiert, daß es zu jedem  $k \geq 1$  eine streng monoton steigende Folge  $(n_i)_{i \geq 1}$  von natürlichen Zahlen gibt, welche die Ungleichung

$$\text{Sp}_{\lambda_k}(P^{(n_i)}) < \text{Tr}_{\lambda_k}(Q^{(n_i)})$$

erfüllt. Nun greift Implikation (b) von Satz 6.1 und liefert  $\bar{\epsilon}(\mu, \nu) \leq 2\lambda_k$  für alle  $k \geq 1$ , also  $\underline{\epsilon}(\mu, \nu) = 0$ .

Ad (vi). Zunächst existiert eine positive reelle Zahl  $\lambda_0$  und zu jedem  $\lambda \in [0, \lambda_0]$  eine natürliche Zahl  $n_0 := n_0(\lambda) \geq 1$  derart, daß

$$\text{Sp}_{\lambda}(P^{(n)}) > \text{Tr}_{\lambda}(Q^{(n)})$$

für alle  $n \geq n_0$  erfüllt ist. Anwendung der Ungleichung (a) von Satz 6.1 ergibt die für alle  $\lambda \in ]0, \lambda_0[$  gültige Abschätzung  $\epsilon(P^{(n)}, Q^{(n)}) > \lambda^2$ , folglich  $\underline{\epsilon}(\mu, \nu) \geq \lambda^2 > 0$  und damit die Behauptung.  $\square$

Die im Werk von C.E. Shannon enthaltene, für die Informationstheorie richtungweisend gebliebene Aussage ergibt sich nun unmittelbar aus dem Hauptsatz als

6.4 Korollar. Unter der zusätzlichen Voraussetzung, daß Sperrigkeits- bzw. Trennrate  $sp(\mu) := \underline{sp}(\mu) = \overline{sp}(\nu)$  und  $tr(\nu) := \underline{tr}(\nu) = \overline{tr}(\nu)$  existieren und zudem  $sp(\mu) \neq tr(\nu)$  erfüllt ist, hat man die Äquivalenz

$$\mu \in T(\nu) \iff sp(\mu) < tr(\nu).$$

Damit ist die Entscheidung, ob eine Quellverteilungsfolge  $\mu$  durch einen Kanal  $\nu$  asymptotisch fehlerfrei übertragen werden kann oder nicht, allein mittels Inspektion der Anordnung zwischen  $sp(\mu)$  und  $tr(\nu)$  zu treffen. Dabei möge nicht übersehen werden, daß die bemerkenswerte Äquivalenz nur dann hilfreich sein kann, wenn die Existenz von  $sp(\mu)$  und  $tr(\nu)$  gesichert ist. An dieser Stelle eröffnet sich ein umfangreiches und tiefgründiges Forschungsgebiet innerhalb der Informationstheorie, in dem noch viele Fragen offen sind. Wie im nachfolgenden Paragraphen sichtbar werden wird, ist die Existenzfrage bisher nur für ausgewählte, aber zugleich anwendungsreiche Klassen von Quellverteilungsfolgen  $\mu$  und Kanalfolgen  $\nu$  gelöst.

## 7. Existenzaussagen für Sperrigkeits- und Trennrate

7.1 Für Quellverteilungen  $P^{(n)}$  in Produktgestalt  $P^{(1)} \otimes n$  auf  $X^n$ , bezüglich deren die einzelnen Buchstaben des Quellalphabets  $X$  unabhängig voneinander verteilt sind, existiert die  $\lambda$ -Sperrigkeitsrate in der Form

$$sp_\lambda(\mu) = \lim_{n \rightarrow \infty} \frac{1}{n} \log Sp_\lambda(P^{(n)}),$$

stellt sich als von  $\lambda$  unabhängig heraus, und stimmt mit der Entropie

$$H(P^{(1)}) = - \sum_{x \in X} P^{(1)}(\{x\}) \log P^{(1)}(\{x\})$$

überein.

7.2 Für Kanalfolgen  $\nu = (Q^{(n)})_{n \geq 1}$  in Produktgestalt (ohne Gedächtnis), also mit  $Q^{(n)} := Q^{(1)} \otimes n$ , wobei  $Q^{(1)}$  ein Kanal  $X \rightarrow Y$  ist, existiert die Trennrate  $tr(\nu) := \underline{tr}(\nu) = \overline{tr}(\nu)$  und stimmt mit der Kapazität von  $Q^{(1)}$  überein:

$$tr(\nu) = C(Q^{(1)}).$$

7.3 Unter der Synthese der in 7.1 und 7.2 genannten Voraussetzungen erhält man die in der klassischen Informationstheorie geläufige Implikation

$$H(P^{(1)}) < C(Q^{(1)}) \Rightarrow \mu \in T(\nu).$$

Um hinreichende Bedingungen an die Quellverteilungsfolge  $\mu$  bzw. die Kanalfolge  $\nu$  dafür stellen zu können, daß  $\text{sp}(\mu)$  bzw.  $\text{tr}(\nu)$  existieren, benötigt man den verallgemeinernden Übergang von endlichen Produktalphabeten  $X^n$  auf abzählbare Produktalphabeten  $X^{\mathbb{Z}}$ , also auf Mengen von abzählbaren Folgen  $(x_n)_{n \in \mathbb{Z}} = (\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots)$  von Elementen aus  $X$ .

7.4 Für Quellverteilungen  $P$  auf derartigen Alphabeten  $X^{\mathbb{Z}}$  führt man den Begriff der *Stationarität* oder allgemeiner der *Periodizität* ein und zeigt unter dieser Zusatzbedingung die Existenz der Sperrigkeitsrate  $\text{sp}(\mu)$ , wobei die Folgenglieder  $P^{(n)}$  von  $\mu$  als die Bilder  $\pi_n(P)$  von  $P$  unter den Projektionen  $\pi_n$  von  $X^{\mathbb{Z}}$  auf die endlichen Produktalphabeten  $X^n$  gewählt werden.

7.5 Für Kanäle  $Q: X^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$  lassen sich ebenfalls *Stationarität* und der umfassendere Begriff der *Periodizität* definieren. Nimmt man noch die Eigenschaft des *Abklingens* (Stetigkeit) des Kanals mit hinzu, so ergibt sich die gewünschte Existenz der Trennrate  $\text{tr}(\nu)$ , wobei die Folgenglieder  $Q^{(n)}$  von  $\nu$  gerade die Kompositionen  $Q[\pi_n]$  von  $Q$  mit den deterministischen Kanälen  $[\pi_n]$  (also Kanäle  $X^{\mathbb{N}} \rightarrow X^n$ ) sind.

## Referenzen

- [1] R. Ahlswede: Beiträge zur Shannonschen Informationstheorie  
im Falle nichtstationärer Kanäle  
Z. Wahrscheinlichkeitstheorie verw. Gebiete 10(1968),  
1-42
- [2] W. Böge: Informationstheorie  
Ausarbeitung einer Vorlesung  
Heidelberg 1967
- [3] E. Henze, H.H. Homuth: Einführung in die Informationstheorie  
3., überarbeitete und erweiterte Auflage  
Friedr. Vieweg und Sohn, Braunschweig 1970
- [4] H. Heyer: Wahrscheinlichkeitsrechnung und Grundzüge der Statistik  
Ausarbeitung einer Vorlesung  
Tübingen 1990
- [5] C.E. Shannon: A Mathematical Theory of Communication  
Bell Syst. Techn. J. 27(1948), 379-423, 623-656
- [6] K. Winkelbauer: Communication channel with finite past history  
In: Transactions of the Second Prague Conference on  
Information Theory, Statistical Decision Functions,  
Random Processes, pp. 685-831.  
Publishing House of the Czechoslovak Academy of Sciences  
Prague 1959
- [7] K. Winkelbauer: On discrete information sources  
In: Transactions of the Third Prague Conference on  
Information Theory, Statistical Decision Functions,  
Random Processes, pp. 765-830.  
Publishing House of the Czechoslovak Academy of Sciences  
Prague 1962
- [8] K. Winkelbauer: Axiomatic definition of channel capacity  
and entropy rate  
In: Transactions of the Fourth Prague Conference on  
Information Theory, Statistical Decision Functions,  
Random Processes, pp. 661-705. .  
Publishing House of the Czechoslovak Academy of Sciences  
Prague 1965

## Inhalt

<b>Udo Bayer/ Juliane Hansen/ Karl Gfesser</b>	5	Grußwort / Foreword
<b>Ottomar Hartwig</b>	7	Ein Bildzeichen für Elisabeth Walther-Bense zum 75. Geburtstag
<b>Gérard Deledalle</b>	8	Peirce, les Catégories et les Signes
<b>Rosemarie und Fried Alstaedter</b>	23	An Elisabeth
<b>Frieder Nake</b>	24	Der semiotische Charakter der informatischen Gegenstände
<b>Georg Nees</b>	36	Die Blindschleichen, das Eisenerz und die Zeichen. Semiotisch/kybernetische Erinnerungen und Vorahnungen
<b>Wil Frenken</b>	49	Für Elisabeth. PRO CAPTU LECTORIS HABENT SUA FATA LIBELLI
<b>Elisabeth Emter</b>	52	<i>Augenblick</i> . Eine Zeitschrift wider die metaphysische Behaglichkeit
<b>Armin Mehling</b>	60	Geburtstagsgruß
<b>Wojciech H. Kalaga</b>	61	Signification and Objects
<b>Betty Leirner</b>	71	espássaro
<b>Jan Peter Tripp</b>	73	<<Pauline>> (Noch 'ne Blume für E.)
<b>Dinda L. Gorrée</b>	74	Translation: Between Imaging, Modeling, and Manipulation
<b>Angelika Jakob</b>	84	Semiramis der Semiotik
<b>Hans Brög</b>	85	Am Rande der Semiotik
<b>Karel Trinkewitz</b>	91	Bernard Bolzanos Haus in Prag als angeblicher Tatort eines Mordes im Jahr 1848
<b>Dušan I. Bjelić</b>	94	The Levitational Physics of Icons and the Gravitational Theology of Newton
<b>Lee Lichterloh</b>	113	Komposition mit Schwarz
<b>Rudolf Haller</b>	114	Das Fortschreiten der Erkenntnis. Zur Verwendung semiotischer Zusammenhänge durch Benedictus de Spinoza
<b>Frue Cheng</b>	118	Neue Darstellung der Zeichenoperationen
<b>Angelika Karger</b>	128	Zeichenwirkung als philosophische Aufgabe
<b>Jens-Peter Mardersteig</b>	145	Faul im August
<b>Udo Bayer</b>	147	Zur Semiotik der Gartenkunst
<b>M. Drea</b>	165	Le monde en miniature



<b>Karl Herrmann</b>	167	Anwendung semiotischer Vorstellungen zur Erzeugung erkenntnistheoretischer Modelle
<b>Thomas Gil</b>	181	Der Zeichenbegriff in John Lockes empiristischer Erkenntnistheorie
<b>Solange Magalhães</b>	189	S/ Título
<b>Magdolna Orosz</b>	190	"Du kannst nur denken durch den Mittler Sprache." Vermittlung und zeichenhafte Welt in der deutschen Romantik
<b>Reinhard Döhl</b>	203	zuerst wurden die poetiken außer kurs gesetzt - dann kam der reim abhanden - schließlich fehlten sogar die worte. aprèslude
<b>Helmut Kreuzer</b>	209	Hiršals Jugendwelt. Oder eine "ungewohnte Form" der Autobiographie
<b>Almir Mavignier</b>	215	Konvex/Konkave Linie
<b>Ilse Walther-Dulk</b>	216	Auf der Suche nach der Philosophie Marcel Prousts
<b>Xu Hengchun</b>	232	Eine Skizze von Kulturuntersuchung
<b>Vera Molnar</b>	238	Variations Ste.-Victoire 1989-96
<b>Barbara Wichelhaus</b>	244	Der kreative Aufbau von Bedeutungen durch Malen und Zeichnen im Kindesalter
<b>Engelbert Kronthaler</b>	259	Du sollst Dir kein Bild machen ...
<b>Karl Gfesser</b>	274	Vorbemerkungen zu einer semiotischen Textanalyse
<b>Maria Heyer-Loos</b>	297	Montierte Landschaft
<b>Alfred Toth</b>	298	Auf dem Weg zur ersten semiotischen Grammatik
<b>Hariss Kidwaii</b>	311	Die Basistheorie der Semiotik und die Kleine Matrix
<b>Wolfgang Kiwus</b>	318	Computergrafiken
<b>Herbert Heyer</b>	320	Über asymptotisch fehlerfreie Übertragbarkeit von Information
<b>Josef Klein</b>	335	Über Intention und Intension in Ansehung des Aufbaus der deontischen Modalitäten - Zur normsemiotischen Kritik des Extensionalismus
<b>Gerald L. Eberlein/ Angelika Karger</b>	345	Semiotische Analyse eines sozio-kulturellen Phänomens am Beispiel von UFO-Gläubigkeit
<b>Anita Kernwein</b>	355	Bibliographie der Schriften Elisabeth Walthers