JACOBS, K.
Math. Zeitschr. 78, 151—170 (1962)

Über Kanäle vom Dichtetypus

Von
KONRAD JACOBS

Einleitung

In zwei bemerkenswerten Noten [1], [2] hat ROSENBLATT-ROTH Quellen und Kanäle vom Dichtetypus untersucht und ein *Coding Theorem* (Fundamentallemma von FEINSTEIN) bewiesen. Bekanntlich ist ein Coding Theorem allein nicht befriedigend; man muß auch seine *Umkehrung* zu beweisen suchen. Bei dem Versuch, die Umkehrung des Coding Theorem von ROSENBLATT-ROTH zu beweisen, fiel mir folgendes auf: Die Durchlaßkapazität wurde von ROSENBLATT-ROTH in einer Weise definiert, die gegen die *Abänderung* des betrachteten Kanals auf gewissen *Nullmengen invariant* ist; die von ihm gewählte Definition des *ε-Code* (der ε-unterscheidbaren Familie) ist jedoch gegen solche Abänderungen *nicht* invariant; vielmehr kann man durch solche Abänderungen u. U. beliebig lange Codes schaffen; die von ROSENBLATT-ROTH gegebene untere Abschätzung der Code-Längen mittels der Durchlaßkapazität ist zwar richtig, aber sicher nicht optimal, wie es die Umkehrung eines Coding Theorem besagen würde. Ich möchte daher die von ROSENBLATT-ROTH angegebene Definition des ε-Code in einer Weise einengen, daß a) zumindest die *Code-Längen* invariant gegen Abänderungen vom obgenannten Typus werden, b) die Abschätzung nach unten bestehen bleibt und c) die schwache Umkehrung[1]) des Coding Theorem für eine vernünftige Klasse von Kanälen (nämlich stationäre Kanäle mit endlichem Gedächtnis [letzteres nicht nur im Sinne einer Meßbarkeits-, sondern auch im Sinne einer Unabhängigkeitsforderung (*U*) (§ 3) zu verstehen], abzählbar-erzeugtem Borel-Körper des Ausgangs-Alphabets [diese Annahme kann im Fall von ROSENBLATT-ROTH (abzählbares Eingangsalphabet) übrigens stets erzwungen werden], und gewissen Eigenschaften (*J*) (§ 3), die zwar, wie ein Beispiel in § 4 zeigt, nicht notwendig sind, aber doch vernünftig erscheinen: Es ist zu erwarten, daß ein Kanal, der trotz kontinuierlicher Alphabete nur endliche ε-Codes besitzt, auch gewisse Kompaktheitseigenschaften hat) bewiesen werden kann. Die Tatsache, daß das Coding Theorem in [2] nur für Kanäle mit höchstens abzählbarem Eingangsalphabet aufgestellt wird, läßt vermuten, daß ROSENBLATT-ROTH das eben erörterte Problem gesehen hat; in diesem Falle gibt es nämlich eine größte Nullmenge im Eingangsalphabet, und man kann die Schwierigkeit mit einigen normierenden Vorschriften umgehen. — Um gewisse Konvergenzschwierigkeiten zu vermeiden, habe ich auch die Definition der Durchlaßkapazität gegenüber [2] ein wenig verändert: es wird nicht mehr verlangt, daß die Zeiträume am

[1]) Vgl. WOLFOWITZ [4].

Ausgang des Kanals stets ein zusammenhängendes Intervall bilden, vielmehr dürfen zeitweilig kurze Zeiten ohne Beobachtung verstreichen. Es entsteht eine Durchlaßkapazität, die a) nicht kleiner als die in [2] definierte ist, b) die bequeme Handhabung gewisser *periodischer* Eingangsquellen gestattet, und c), wie die Umkehrung des Coding Theorem zeigt, maximal ist, sowie d) bereits durch Betrachtung gewisser periodischer Eingangsquellen mit unabhängigen Zeichen-Blocks ermittelt werden kann*).

§ 1. Zur Struktur der Kanäle vom Dichtetypus

1. Grundbegriffe

Unter einem *Alphabet* verstehen wir ein Paar (A, C), wo A eine nichtleere Menge und C ein Borel-Körper in A ist. Aus einem Alphabet (A, C) bildet man den Produktraum

$$\Omega = \prod_{t\in\Gamma}^{\times} A_t = \{\omega = (\omega_t) = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots) \mid \omega_t \in A\}$$

mit $A_t = A$ $(t \in \Gamma)$. Hierbei bedeutet Γ die Menge der ganzen Zahlen. Natürlich kann man auch für jedes nichtleere $J \subseteqq \Gamma$ den Produktraum $\Omega(J) = \prod\limits_{t\in J}^{\times} A_t$ bilden. Man hat eine natürliche Abbildung φ_J von Ω auf $\Omega(J)$ vermöge $\varphi\omega = \eta$ mit $\eta_t = \omega_t$ $(t \in J)$. Wir bilden die Produkt-Borel-Körper $C(J) = \prod\limits_{t\in J}^{\times} C_t$ $(C_t = C)$ und setzen $B(J) = \varphi_J^{-1} C(J)$, $B(\Gamma) = B$. B wird von dem Mengenkörper $\bigcup\limits_{J \text{ endlich}} B(J)$ erzeugt. Für $B(\langle s, t\rangle)$ schreiben wir auch $B(s, t)$.

Nun erfolgt eine Festsetzung, die alle weiteren Betrachtungen entscheidend einengt: Wir wählen ein normiertes Maß n_0 auf C und setzen $m = \prod\limits_{t\in\Gamma}^{\times} n_t$ $(n_t = n_0)$. m ist eine Wahrscheinlichkeitsverteilung (WV) auf B. *Dies m wird ein für allemal festgehalten.*

Sei p eine beliebige WV auf B. Wir sagen, p sei vom $(m\text{-})$*Dichtetypus*, wenn p auf jedem $B(J)$ mit *endlichem* J totalstetig bezüglich m ist, in Zeichen: $p \ll m\ (B(J))$, d.h. aus $E \in B(J)$, $m(E) = 0$ folgt stets $p(E) = 0$. Hieraus folgt keineswegs $p \ll m\ (B)$, vielmehr sind z.B. verschiedene ergodische Maße stets trägerfremd; gewinnt man speziell p in der Form $p = \prod\limits_{t\in\Gamma}^{\times} q_t$ mit auf C vorgegebenem $q_t = q \neq n_0$, so tritt dieser Fall ein; ist $q \ll n_0$ (C), so ist p nichtsdestoweniger vom Dichtetypus.

Wir geben uns nun einen weiteren normierten Maßraum (A', C', n') vor und bilden in gleicher Weise wie oben Ω' mit den Borel-Körpern $B'(J)$, B', sowie die Produkt-WV m'. Wir haben damit auch den Begriff der WV vom $(m'\text{-})$ Dichtetypus auf B'.

Unter einem *Kanal* verstehen wir eine für $\omega \in \Omega$, $E' \in B'$ erklärte Funktion $P(\omega, E')$ mit folgenden Eigenschaften:

a) Für jedes $\omega \in \Omega$ ist $P(\omega, .)$ eine WV auf B'.

b) Für jedes $E' \in B'$ ist $P(., E')$ eine B-meßbare Funktion auf Ω.

*) *Zusatz während der Korrektur.* Diese Idee findet sich bereits in einer Abhandlung von Wolfowitz, vgl. [5].

Wir sagen, der Kanal P sei vom (m'-)*Dichtetypus*, wenn die WV $P(\omega, .)$ für jedes $\omega \in \Omega$ vom Dichtetypus ist. Zwei Kanäle P, Q heißen (m-) *äquivalent*, wenn für jedes $E' \in \bigcup_{J \text{ endlich}} B'(J)$ die Funktionen $P(., E')$ und $Q(., E')$ m-fast-überall übereinstimmen. Der Kanal P heißt *vorgriffsfrei mit dem endlichen Gedächtnis* $g \geqq 0$ (kurz: ein *g-Kanal*), wenn die Funktion $P(., E')$ für $E' \in B'(s, t)$ stets $B(s-g, t)$-meßbar ist. *Wir werden im weiteren nur mehr g-Kanäle vom Dichtetypus mit fixiertem $g \geqq 0$ betrachten.*

Wir betrachten jetzt den Produktraum $\widetilde{\Omega} = \Omega \times \Omega'$ mit den Produkt-Borel-Körpern $B(J) \times B'(J')$ (speziell $\widetilde{B} = B \times B'$) und dem Produktmaß $\overline{m} = m \times m'$.

Für ein beliebiges endliches nichtleeres $J \subseteqq \Gamma$, in welchem s das kleinste und t das größte Element ist, bezeichnen wir $\langle s-g, t\rangle$ mit $\bar{J}$ und setzen $B(\bar{J}) \times B'(J) = \widetilde{B}(J)$. Eine Funktion $f_J(\omega, \omega') \geqq 0$ auf $\widetilde{\Omega} = \Omega \times \Omega'$ heißt eine *Kanaldichte* für den Kanal P und den endlichen Zeitraum $J \subseteqq \Gamma$, wenn folgendes gilt:

a) $f_J(\omega, \omega')$ ist $B(\bar{J}) \times B'(J)$-meßbar, also insbesondere bei festem ω $B'(J)$-meßbar in ω', und bei festem ω' $B(\bar{J})$-meßbar in ω.

b) $\int m(d\omega)\, P(\omega, \widetilde{E}_\omega) = \int_{\widetilde{E}} f_J(\omega, \omega')\, \overline{m}(d\omega, d\omega') \qquad (\widetilde{E} \in \widetilde{B}(J))$

Hierbei bezeichnet $\widetilde{E}_\omega = \{\omega' \mid (\omega, \omega') \in \widetilde{E}\}$ den ω-Schnitt von $\widetilde{E}$.

c) $\int f_J(\omega, \omega')\, m'(d\omega') = 1 \qquad (\omega \in \Omega)$.

Für $f_{\langle s,t\rangle}$ schreiben wir auch ${}_s f_t$.

Da die linke Seite von b) offenbar eine bezüglich $\overline{m}$ totalstetige WV $\widetilde{m}$ auf $\widetilde{B}(J)$ definiert, ist die Kanaldichte f_J nach dem Satz von Radon-Nikodym stets vorhanden (die Bedingung c) läßt sich durch Abänderung auf einer Nullmenge aus $\widetilde{B}(J)$ erzwingen). Sie ist bis auf $\overline{m}$-Nullmengen eindeutig bestimmt. Äquivalente Kanäle liefern dieselben Kanaldichten.

Ist P ein Kanal und p eine WV auf B, so ist durch

$$\tilde{p}(\widetilde{E}) = \int p(d\omega)\, P(\omega, \widetilde{E}_\omega) \qquad (\widetilde{E} \in \widetilde{B})$$

eine WV $\tilde{p}$ auf $\widetilde{B}$ gegeben. Faßt man B und B' in üblicher Weise als Unter-Borel-Körper von $\widetilde{B}$ auf, so entsteht durch Einschränkung von $\tilde{p}$

a) auf B die WV p

b) auf B' eine WV, die wir mit p' bezeichnen.

Die WV p' ist stets vom Dichtetypus, wenn der Kanal es ist. Ist auch p vom Dichtetypus, so ist stets $\tilde{p} \ll \overline{m}$ $(\widetilde{B}(J))$.

Durch $(T\omega)_t = \omega_{t+1}$ ist eine eineindeutige Abbildung T von Ω auf sich gegeben, die $TB(J) = B(J-1)$ erfüllt und somit samt allen ihren Potenzen T^t $(t \in \Gamma)$ B-meßbar ist. Die entsprechenden Abbildungen in Ω' und $\widehat{\Omega}$ bezeichnen wir ebenfalls mit T. Aus einer WV p erhält man die WVen

$$p_t\colon \quad p_t(E) = p(T^{-t}E) \qquad (E \in B),$$

aus einem Kanal P erhält man die Kanäle

$$P_t\colon\quad P_t(\omega, E') = P(T^t\omega, T^t E') \qquad (\omega\in\Omega,\ E'\in B').$$

p bzw. P heißt periodisch mit der Periode $d\neq 0$, wenn $p_{kd}=p$ bzw. $P_{kd}=P$ $(k\in\Gamma)$ gilt; bei der Periode 1 spricht man von Stationarität[2]). Sind p und P periodisch bzw. stationär, so ist auch p, und damit auch p', periodisch bzw. stationär.

Wir erwähnen noch, wie man stationäre g-Kanäle vom Dichtetypus nach einem Verfahren von WOLFOWITZ [3], [4], [5] *konstruieren* kann. Sei $f(\xi_{-g},\ldots,\xi_0;\xi_0')\geqq 0$ eine auf $\Omega(-g,0)\times A'$ erklärte $C(-g,0)\times C'$-meßbare Funktion mit

$$\int f(\xi_{-g},\ldots,\xi_0;\xi_0')\,n'(d\xi_0') = 1 \qquad (\xi_{-g},\ldots,\xi_0\in A).$$

Wir setzen

$${}_sf_t(\omega,\omega') = \prod_{u=s}^{t} f(\omega_{u-g},\ldots,\omega_u;\omega_u')$$

$$P^{(t)}(\omega,E') = \int_{E'} {}_tf_t(\omega,\omega')\,m'(d\omega') \qquad \big(E'\in B'(t,t)\big)$$

$$P(\omega,.) = \prod_{t\in\Gamma}{}^{\times} P^{(t)}(\omega,.).$$

Dann ist P ein Kanal der verlangten Bauart mit den Kanaldichten ${}_sf_t$. Ändert man f auf einer Nullmenge ab, so entsteht ein äquivalenter Kanal.

2. *Die Entropie-Metrik in L^1_m*

Sei (Ω', B_0', m') ein beliebiger normierter Maßraum und $L^1_{m'}$ der Raum der m'-integrablen B_0'-meßbaren Funktionen f' mit der Norm $\|f'\| = \int |f'|\,dm$.

Die Funktion $z(x) = x\log x$ ist für $x\geqq 0$ erklärt und konvex (nach oben hohl), für $0\leqq x\leqq \frac{1}{e}$ fällt sie von 0 bis $-\frac{1}{e}$, um dann monoton ad infinitum, mit $z(1)=0$, zu steigen. Wir bilden jetzt

$$U' = \{g' \mid g'\in L^1_{m'},\ g'\geqq 0,\ z(g')\in L^1_{m'}\} \tag{1}$$

und führen in U' die neue Metrik *(Entropie-Metrik)*

$$|f',g'| = \|f'-g'\| + \|z(f')-z(g')\| \tag{2}$$

ein.

SATZ 1.1. *Ist $L^1_{m'}$ separabel, so ist auch U', versehen mit der Entropiemetrik, separabel.*

BEWEIS. Sei f_k' eine in $z(U')$ normdicht liegende Folge und $g_k'\in U'$ so gewählt, daß $z(g_k')=f_k'$ gilt. Wir setzen $V_{rk}' = \left\{g' \,\middle|\, \|z(g')-z(g_k')\| < \frac{1}{r}\right\}$ und bestimmen in V_{rk}' eine normdicht liegende Folge g_{rkj}'. Ist $g'\in U'$, $\varepsilon>0$ beliebig, so bestimme man r mit $\frac{3}{r}<\varepsilon$, sodann k mit $g'\in V_{rk}'$ und dann j mit $\|g'-g_{rkj}'\|<\frac{1}{r}$.

[2]) Die entsprechenden Begriffe erhält man für WVen p', $\tilde{p}$ über B', $\tilde{B}$.

Dann gilt

$$|g'_{rkj}, g'| \leqq \|z(g'_{rkj}) - z(g'_k)\| + \|z(g'_k) - z(g')\| + \|g'_{rkj} - g'\| < \frac{3}{r} < \varepsilon.$$

Nun beweisen wir einen allgemeinen Satz über separable pseudo-metrische Räume.

SATZ 1.2. *Sei U' ein separabler pseudo-metrischer Raum und D' der von der Pseudo-Metrik, d.h. von den Mengen der Bauart*

$$U'_\delta(g') = \{f' \mid f' \in U', \ |f', g'| < \delta\}$$

erzeugte Borel-Körper. Sei w' eine WV auf D'. Ein Punkt $g' \in U'$ heiße dicht, wenn $w'(U'_\delta(g')) > 0$ $(\delta > 0)$ ist. Dann gilt: w'-fastalle Punkte aus U' sind dicht.

BEWEIS. Sei $N' = \{g' \mid g' \in U', \ g' \text{ nicht-dicht}\}$. Zu jedem $g' \in N'$ gibt es ein maximales $\delta = \delta(g')$ mit $0 < \delta \leqq 1$ und $w'(U'_\delta(g')) = 0$. Sei $g'_k \in N'$ eine in N' pseudo-metrisch dichte Folge. Ist $|g'_k, g'| < \frac{1}{2}\delta(g')$, so folgt

$$U'_{\frac{1}{2}\delta(g')}(g'_k) \subseteqq U'_{\delta(g')}(g')$$

und somit $w'(U'_{\frac{1}{2}\delta(g')}(g'_k)) = 0$, also $\delta(g'_k) \geqq \frac{1}{2}\delta(g')$, also $g' \in U'_{\delta(g'_k)}(g'_k)$. Hieraus folgt

$$N' \subseteqq \bigcup_k U'_{\delta(g'_k)}(g'_k),$$

und damit $w'(N') = 0$, wie behauptet.

Im folgenden werden wir unter (Ω', B'_0, m') stets einen der Räume $(\Omega', B', (J), m')$ mit endlichem $J \subseteqq \Gamma$, wie sie in Nr. 1 erklärt wurden, verstehen, und U' gemäß (1), die Metrik gemäß (2) einführen.

3. Kanäle und meßbare Abbildungen

Bezeichnet man mit $R(B')$ die Gesamtheit aller endlichen σ-additiven reellen Funktionen auf B', so kann man einen Kanal $P(\omega, E')$ als eine Abbildung ψ: $\Omega \to R(B')$ mit $\psi\omega = P(\omega, .)$ auffassen. Ist $J \subseteqq \Gamma$, so liefert ein Kanal natürlich auch eine Abbildung ψ_J: $\Omega \to R(B'(J))$. Bezeichnet man mit E_J die Einschränkungsabbildung von $R(B')$ auf $R(B'(J))$, so gilt trivialerweise

$$\psi_J = E_J \psi.$$

Ist C' — und damit auch B', ebenso jedes $B'(J)$ — abzählbar erzeugt, so bewirkt der Übergang zu einem äquivalenten Kanal Q für jedes J die Abänderung von ψ_J auf einer m-Nullmenge. Ist nämlich $K' \subseteqq \bigcup_{J \text{ endlich}} B'(J)$ ein abzählbarer B' erzeugender Mengenkörper, und setzt man

$$N(E') = \{\omega \mid Q(\omega, E') \neq P(\omega, E')\},$$

so ist $N = \bigcup_{E' \in K'} N(E')$ eine m-Nullmenge und es gilt

$$Q(\omega, E') = P(\omega, E') \qquad (E' \in B'),$$

falls nur $\omega \notin N$ ist.

Nun wollen wir diesen simplen Sachverhalt in die Sprache der Dichten übertragen. Wieder beschränken wir uns auf g-Kanäle vom Dichtetypus. Mit $L'(J)$ bezeichnen wir die Gesamtheit aller $B'(J)$-meßbaren Funktionen aus $L^1_{m'}$, mit $U'(J)$ den entsprechend (1) erklärten Unterbereich von $L'(J)$. Zwischen den $g' \in R(B'(J))$ mit $g' \ll m'$ $(B'(J))$ und den $g' \in L'(J)$ besteht nach RADON-NIKODYM eine Isomorphie, so daß wir zugeordnete Elemente in der Bezeichnung gar nicht unterscheiden wollen: $L'(J)$ kann als Teilraum von $R(B'(J))$ aufgefaßt werden und ψ_J als Abbildung in $L'(J)$.

Für die weitere Untersuchung machen wir die

ANNAHME (A). *1. C' ist abzählbar erzeugt. 2. Die Kanäle, die wir betrachten, erfüllen $\psi_J \Omega \subseteqq U'(J)$, es existiert also stets*

$$H_J(\omega) = \int z(\psi_J \omega)\, m(d\omega').$$

Ferner wollen wir die Existenz von $\int H_J(\omega)\, m(d\omega)$ voraussetzen.

Nr. 1 hat zur Folge, daß auch die $B'(J)$ (J endlich) abzählbar erzeugt sind (ebenso B'). Damit sind nach Satz 1.1 auch die $U'(J)$ mit der Entropiemetrik separabel. Nr. 2 hat zur Folge, daß die ψ_J stets als Abbildungen in die $U'(J)$ aufzufassen sind. Wir wollen jetzt gewisse Meßbarkeitseigenschaften der Abbildungen ψ_J und E_J nachweisen. Den von der Entropiemetrik in $U'(J)$ erzeugten Borel-Körper wollen wir mit $D'(J)$ bezeichnen.

SATZ 1.3. *Sind $I, J \subseteqq \Gamma$ endlich mit $J \subseteqq I$, so gilt $E_J U'(I) \subseteqq U'(J)$. Als Abbildung von $U'(I)$ in $U'(J)$ aufgefaßt, ist E_J stetig bezüglich der Entropiemetrik, also gewiß $D'(I)$-$D'(J)$-meßbar.*

BEWEIS. Hier wird von der Annahme (A) noch kein Gebrauch gemacht. Jedenfalls ist $E_J L'(I) = L'(J)$, und diese Abbildung E_J ist nichts anderes als die bedingte Erwartung bezüglich $B'(J)$ und m'. Somit gilt nach der Jensenschen Ungleichung $z(E_J g') \leqq E_J(z(g'))$. Da z nach unten beschränkt ist, hat dies $E_J U'(I) \subseteqq U'(J)$ zur Folge. Sei nun $g'_k,\ g' \in U'(I)$ und $|g'_k, g'| \to 0$. Dann ist $z(g'_k)$ normkonvergent. Die Folge $E_J(z(g'_k))$ hat dieselbe Eigenschaft, ist also gleichgradig-integrabel, die von ihr majorisierte Folge $z(E_J(g'_k))$ also auch. Aus $g'_k \to g'$ (m'-stochastisch) folgt $z(E_J(g'_k)) \to z(E_J(g'))$ (m'-stochastisch), also $\|z(E_J(g'_k)) - z(E_J(g'))\| \to 0$, was mit $\|E_J g'_k - E_J g'\| \to 0$ zusammen die Konvergenz in der Entropiemetrik zur Folge hat. Also ist E_J stetig bezüglich der Entropiemetrik.

SATZ 1.4. *Zu jedem endlichen nichtleeren J gibt es eine $B(\bar{J})$-$D'(J)$-meßbare Abbildung $\bar{\psi}_J$ von Ω in $U'(J)$ mit*

$$\psi_J \omega = \bar{\psi}_J \omega \quad (m\text{-fastüberall}),$$

genauer: Ist $f_J(\omega, \omega')$ eine Kanaldichte, so gibt es eine Nullmenge $N \in B(\bar{J})$ mit

$$P(\omega, E') = \int_{E'} f_J(\omega, \omega')\, m'(d\omega') \qquad (\omega \notin N) \tag{3}$$

und man kann etwa

$$\bar{\psi}_J \omega = \begin{cases} f_J(\omega, .), & \textit{falls dies zu } U'(J) \textit{ gehört} \\ 1 & \textit{sonst} \end{cases} \tag{4}$$

setzen. Hieraus folgt noch: Sind P und Q äquivalent, so gibt es eine Nullmenge $N_0 \in B(\bar{J})$ mit

$$P(\omega, E') = Q(\omega, E') \qquad \big(\omega \notin N_0,\ E' \in B'(J)\big).$$

Beweis. Nun nützen wir die Annahme (A) voll aus. Aus der Bedingung b) in der Definition der Kanaldichte folgt: Zu jedem $E' \in B'(J)$ gibt es eine m-Nullmenge $N(E') \in B(\bar{J})$ mit (3) für $\omega \notin N(E')$. Sei K' ein $B'(J)$ erzeugender abzählbarer Mengenkörper und $N = \bigcup_{E' \in K'} N(E')$, also $N \in B(\bar{J})$, $m(N) = 0$. Die Gesamtheit aller $E' \in B'(J)$, die (3) für $\omega \notin N$ erfüllen, ist eine K' umfassende monotone Klasse, also $= B'(J)$. Um die Meßbarkeit der durch (4) gegebenen Abbildung $\bar{\psi}_J$ nachzuweisen, genügt es, für jedes $g' \in U'(J)$ und $\delta > 0$

$$\bar{\psi}_J^{-1}\, U'_\delta(g') \in B(\bar{J})$$

zu zeigen. Nun ist die Funktion

$$h(\omega, \omega') = |z\big(f_J(\omega, \omega')\big) - z\big(g'(\omega')\big)| + |f_J(\omega, \omega') - g'(\omega')|$$

$\widetilde{B}(J)$-meßbar, also die durch

$$h(\omega) = \begin{cases} \int h(\omega, \omega')\, m'(d\omega'), & \text{falls dies existiert} \\ \qquad |1, g'| & \text{sonst} \end{cases}$$

gegebene Funktion $B(\bar{J})$-meßbar. Es ist aber

$$\bar{\psi}_J^{-1}\big(U'_\delta(g')\big) = \{\omega \mid h(\omega) < \delta\}.$$

Wir machen uns jetzt folgenden allgemeinen Gedanken zunutze: Ist (Ω, B, m) irgendein normierter Maßraum, D' ein Borel-Körper in einer Menge U' und $\bar{\psi}$ eine B-D'-meßbare Abbildung von Ω in U', so ist durch $w' = \bar{\psi} m$ (nämlich $w'(E') = m(\bar{\psi}^{-1} E')$ $\big(E' \in D'\big)$) ein normiertes Maß w' auf D' gegeben. Ändert man $\bar{\psi}$ auf einer m'-Nullmenge ab, so bleibt w' ungeändert.

Nun wählen wir speziell $\big(\Omega, B(\bar{J}), m\big)$, $U' = U'(J)$, $D' = D'(J)$ wie bisher. Einem Kanal sind auf dem Wege über die Kanaldichten zahlreiche $B(\bar{J})$-$D'(J)$-meßbare Abbildungen von Ω in $U'(J)$ zugeordnet; je zwei dieser Abbildungen unterscheiden sich nur auf einer m-Nullmenge. Äquivalenten Kanälen sind dieselben Abbildungen zugeordnet. Wir erhalten daher den

Satz 1.5. *1. Zu jedem Kanal P gibt es genau eine WV w'_J auf $D'(J)$ derart, daß für jede gemäß Satz 1.4 $\big($z. B. vermöge (4)$\big)$ P zugeordnete Abbildung $\bar{\psi}_J$ die Beziehung*

$$w'_J = \bar{\psi}_J\, m$$

gilt. Äquivalente Kanäle liefern dieselben w'_J.

2. Sind $I, J \subseteqq \Gamma$ nichtleer und endlich, mit $J \subseteqq I$, so ist

$$w'_J = E_J\, w'_I.$$

(die rechte Seite ist wegen Satz 1.3 sinnvoll).

3. Für w'_J-fastalle $g' \in U'(J)$ gilt

(5) $$w'_J(U'_\delta(g')) > 0 \qquad (\delta > 0).$$

BEWEIS ZU NR. 2. Sind $\overline{\varphi}_I, \overline{\varphi}_J$ gemäß Satz 1.4 irgendwie bestimmt, so gilt m-fastüberall

$$\overline{\varphi}_J\,\omega = E_J\,\overline{\varphi}_I\,\omega.$$

Als handliches Ergebnis erhalten wir den

SATZ 1.6. *Ist $V' \in D'(J)$ derart, daß $w'_J(V') > 0$ gilt, so gibt es zu jeder Kanaldichte $f_J(\omega, \omega')$ mindestens ein ω mit $f_J(\omega, .) \in V'$.*

4. ε-Codes. Erklärung von $N(s, t, \varepsilon)$

Sei $M_J(P) = \{g' \mid g' \in U'(J),\ (5)\ \text{gilt}\}$.

DEFINITION 1. *1. Unter einem ε-Schema für den Zeitraum J versteht man eine Serie $\{E'(\nu) \mid \nu = 1, \ldots, N\} \subseteqq B'(J)$ von paarweise disjunkten Mengen mit folgender Eigenschaft: Zu jedem ν gibt es ein $g'_\nu \in M_J(P)$ mit*

(6) $$\int_{E'(\nu)} g'_\nu\, dm' > 1 - \varepsilon.$$

2. $N(J, \varepsilon)$ bezeichnet die maximale Länge N eines ε-Schemas für den Zeitraum J. Statt $N(\langle s, t\rangle, \varepsilon)$ schreibt man auch $N(s, t, \varepsilon)$.

3. Sei $f_J(\omega, \omega')$ eine Kanaldichte. Unter einem ε-Code zu $f_J(\omega, \omega')$ und dem Zeitraum J versteht man eine Serie von Paaren $\{(\omega(\nu), E'(\nu)) \mid \nu = 1, \ldots, N\}$ mit folgenden Eigenschaften:

a) Die $E'(\nu)$ bilden ein ε-Schema für den Zeitraum J.

b) Setzt man $g'_\nu(.) = f_J(\omega(\nu), .)$, so ist g'_ν eine Wahrscheinlichkeitsdichte, und es gilt (6).

Es ist nun a priori klar, daß ε-Schemata in ihrer Eigenschaft als solche unberührt bleiben, wenn man zu äquivalenten Kanälen übergeht; auch $N(J, \varepsilon)$ ist invariant gegen solche Änderungen. w'_J bleibt ja unverändert. Dagegen kann man einen ε-Code u.U. durch Abänderung der Kanaldichte auf einer Nullmenge zunichte machen. Wir zeigen jetzt, daß man dann aber sogleich einen ε-Code gleicher Länge wiedergewinnen kann, indem man die $\omega(\nu)$ durch in gewissem Sinne benachbarte Punkte ersetzt.

SATZ 1.7. *1. Jedes ε-Schema $\{E(\nu) \mid \nu = 1, \ldots, N\}$ kann bei vorgegebener Kanaldichte $f_J(\omega, \omega')$ durch passende Wahl der Punkte $\omega(\nu)$ zu einem ε-Code ergänzt werden.*

2. $N(J, \varepsilon)$ ist die maximale Länge eines ε-Codes für den Zeitraum J.

BEWEIS. Nr. 2 folgt aus Nr. 1. Seien $g'_\nu \in M_J(P)$ so gewählt, daß (6) gilt. Für hinreichend kleines $\delta > 0$ und $f'_\nu \in U'_\delta(g'_\nu)$ gilt immer noch

$$\int_{E'(\nu)} f'_\nu\, dm' > 1 - \varepsilon.$$

Wegen $w'_J(U'_\delta(g'_\nu)) > 0$ gibt es Punkte $\omega(\nu)$ mit $f_J(\omega(\nu), .) \in U'_\delta(g'_\nu)$. Nach Satz 1.4 (3) kann man erreichen, daß $f(\omega(\nu), .)$ eine Wahrscheinlichkeitsdichte ist.

SATZ 1.8. *Sind $I, J \subseteqq \Gamma$ endlich und nichtleer mit $J \subseteqq I$, so ist jedes ε-Schema für J auch ein ε-Schema für I. Insbesondere ist $N(I, \varepsilon) \geqq N(J, \varepsilon)$.*

BEWEIS. Sei $\{E'(\nu) \mid \nu = 1, \ldots, N\}$ ein ε-Schema für J. Wegen $B'(J) \subseteqq B'(I)$ ist es auch in $B'(I)$ enthalten. Man kann $g'_\nu \in M_J(P)$ und $\delta > 0$ so bestimmen, daß für $f'_\nu \in U'(J)$, $f'_\nu \in U'_\delta(g'_\nu) = U'_\nu$ stets

$$\int_{E'(\nu)} f'_\nu \, dm' > 1 - \varepsilon$$

gilt. Wegen $w'_J = E_J w'_I$ und $w'_J(U'_\nu) > 0$ ist $w'_I(E_J^{-1} U'_\nu) > 0$ und daher $E_J^{-1} U'_\nu \cap M_I(P) \neq 0$. Wählt man h'_ν aus dieser Menge, so ist

$$\int_{E'(\nu)} h'_\nu \, dm' = \int_{E'(\nu)} E_J h'_\nu \, dm' > 1 - \varepsilon .$$

§ 2. Das Coding Theorem

Bei der Untersuchung von $N(J, \varepsilon)$ für gewisse endliche Zeiträume J kann man sich für stationäre g-Kanäle stets auf den Fall beschränken, daß 0 das kleinste Element von J (und somit $-g$ das kleinste Element von $\bar{J}$) ist; denn ε-Schemata bleiben ε-Schemata, wenn man eine Verschiebung T^u auf sie ausübt. Wir befassen uns im weiteren nur mehr mit dem genannten Fall und untersuchen jetzt strikt monoton wachsende Folgen $\mathfrak{J} = \{J_k\}$ von Zeiträumen. Ansonsten wird innerhalb dieses Paragraphen kein Gebrauch von Stationaritätsannahmen gemacht. Bezeichnet allgemein $|J|$ die Mächtigkeit von J und setzt man $\hat{J}_k = \bar{J} - \{-g, \ldots, -1\}$ (d.i. das kleinste J_k umfassende Intervall), so folgt $|I_k| \to \infty$, $|\hat{J}_k| \to \infty$. Offenbar ist

$$d(\mathfrak{J}) = \liminf_k \frac{|J_k|}{|\hat{J}_k|}$$

ein vernünftiges Maß für die Dichte der Menge $\bigcup_k J_k$ auf der positiven Halbachse. Stets ist $0 \leqq d(\mathfrak{J}) \leqq 1$.

Wie schon in §1, Nr. 1 bemerkt wurde, ist für jede WV p auf B die zugehörige WV p' auf B' vom m-Dichtetypus. Ist $f'_J(\omega') \geqq 0$ eine m'-Dichte von p' bezüglich $B'(J)$, sowie $f_J(\omega, \omega')$ eine Kanaldichte, so bezeichnen wir die $B(J)$-meßbare Funktion

$$\bar{f}_J(\omega, \omega') = \begin{cases} \dfrac{f_J(\omega, \omega')}{f'_J(\omega')}, & \text{falls } f'_J(\omega') > 0, \\ 1 & \text{sonst} \end{cases}$$

als *Aposteriori-Kanaldichte* zu f_J und f'_J. Ist p vom m-Dichtetypus, so ist p' gegen den Übergang zu äquivalenten Kanälen invariant; $\bar{f}_J$ hat dann dieselbe Eigenschaft und ist $\bar{m}$-fasteindeutig durch p und P bestimmt, kann also auch sinnvoll als Aposteriori-Kanaldichte *zu p und P* bezeichnet werden.

DEFINITION 2. *Eine reelle Zahl $S \geqq -\infty$ heißt zulässig, wenn es zu jedem $\vartheta > 0$ eine WV p vom Dichtetypus auf B und eine Folge $\mathfrak{J} = \{J_k\}$ mit $d(\mathfrak{J}) > 1 - \vartheta$*

und beschränkten $|\hat{J}_{k+1}| - |\hat{J}_k|$, *sowie*

(1) $$\lim_k \tilde{p}\left(\left\{(\omega, \omega') \,\Big|\, \frac{1}{|J_k|} \log \bar{f}_{J_k}(\omega, \omega') > S\right\}\right) = 1$$

gibt. Offenbar ist $S = -\infty$ *immer zulässig. Die Zahl*

$$L = \sup S$$

über alle zulässigen S *wird als die Durchlaßkapazität des Kanals* P *bezeichnet.*

Da $\bar{f}_J$ $\bar{m}$-, also auch $\tilde{p}$-fasteindeutig durch die Äquivalenzklasse von P sowie durch p bestimmt ist [wegen $\tilde{p} \ll \bar{m}$ $(\tilde{B}(J))$], hängt die Zulässigkeit einer Zahl S nur von der Äquivalenzklasse von P ab. Äquivalente Kanäle haben somit die gleiche Durchlaßkapazität.

SATZ 2.1 (CODING THEOREM). *Sei* P *ein stationärer g-Kanal vom Dichtetypus sowie* L *seine Durchlaßkapazität. Dann gibt es zu beliebigen* $\varepsilon > 0$, $U < L$ *ein* $t_0 > 0$ *derart, daß*

$$N(0, t-1, \varepsilon) > e^{tU} \qquad (t \geqq t_0)$$

gilt.

BEWEIS. Offenbar ist der Satz nur für $L > U > 0$ von Interesse. Wir behandeln nur den Fall $L < \infty$. Der Fall $L = \infty$ geht völlig analog. Sei δ durch $U = L - \delta$ definiert. Man kann $N(0, t-1, \varepsilon) < \infty$ annehmen.

Wir wählen eine zulässige Zahl $S > L - \delta$ und $\vartheta > 0$ derart, daß auch noch

(2) $$(S - \vartheta)(1 - 2\vartheta) > L - \frac{\delta}{2}$$

gilt. Sodann wählen wir p und eine Folge $\mathfrak{J} = \{J_k\}$ derart, daß $d(\mathfrak{J}) > 1 - \vartheta$, $\sup_k (|\hat{J}_{k+1}| - |\hat{J}_k|) < \infty$ und (1) gilt. Die Kanaldichten $f_{J_k} = f_k$ und $\bar{f}_{J_k} = \bar{f}_k$ sowie die Dichten $f'_{J_k} = f'_k$ denken wir uns irgendwie fixiert. Die Menge

$$\tilde{M}(k) = \left\{(\omega, \omega') \,\Big|\, \frac{1}{|J_k|} \log \bar{f}_k(\omega, \omega') > S\right\}$$

gehört zu $\tilde{B}(J_k)$ und erfüllt

$$\tilde{p}(\tilde{M}(k)) \to 1.$$

Setzt man also

(3) $$M(k) = \left\{\omega \,\Big|\, \int_{\tilde{M}(k)_\omega} \bar{f}_k(\omega, \omega')\, m'(d\omega') > 1 - \frac{\varepsilon}{2}\right\},$$

so ist $M(k) \in B(\bar{J}_k)$ und

$$p(M(k)) \to 1.$$

Wir denken uns nun für einen Augenblick irgendein k festgehalten und wählen einen ε-Code $\{(\omega(\nu), E'(\nu)) \mid \nu = 1, \ldots, N\}$ für den Zeitraum J_k, mit

(4) $$\omega(\nu) \in M(k), \qquad E'(\nu) \subseteqq M(k)_{\omega(\nu)},$$

der unter allen möglichen ε-Codes mit dieser Eigenschaft (4) die maximale Länge N besitzt (wegen $N(J_k, \varepsilon) < \infty$ gibt es eine solche). Erinnern wir uns an die Menge $M_{J_k}(P) \subseteqq U'(J_k)$, die bei der Definition des ε-Codes eine Hauptrolle spielte: Es ist $f_k(\omega(\nu), .) \in M_{J_k}(P)$ und

$$m\left(\{\omega \mid f_k(\omega, .) \notin M_{J_k}(P)\}\right) = 0. \tag{5}$$

Ist $f_k(\omega, .) \in M_{J_k}(P)$, $\omega \in M(k)$, so folgt, daß die Menge

$$E' = \bigcup_{\nu=1}^{N} E'(\nu)$$

die Relation

$$\int_{E'} f_k(\omega, \omega')\, m'(d\omega') \geqq \frac{\varepsilon}{2}$$

erfüllt. Für $\omega = \omega(\nu)$ ist das wegen $\varepsilon < 1$ selbstverständlich; hätte man hier $< \frac{\varepsilon}{2}$ für ein $\omega \neq \omega(\nu)$, so wäre für $E'(N+1) = M(k)_\omega - E'$ wegen $\omega \in M(k)$ und (3) die Beziehung

$$\int_{E'(N+1)} f_k(\omega, \omega')\, m'(d\omega') > 1 - \varepsilon$$

erfüllt, und man könnte den ε-Code um das Glied $\left(\omega(N+1), E'(N+1)\right)$ mit $\omega(N+1) = \omega$ verlängern. Damit erhalten wir die Abschätzung

$$\begin{aligned} p'(E') &= \int \Big[\int_{E'} f_k(\omega, \omega')\, m'(d\omega')\Big]\, p(d\omega) \\ &\geqq \int_{M(k) \cap \{\omega \mid f_k(\omega, .) \in M_{J_k}(P)\}} \Big[\int_{E'} f_k(\omega, \omega')\, m'(d\omega')\Big]\, p(d\omega) \\ &\geqq \frac{\varepsilon}{2}\, p\left(M(k)\right), \end{aligned}$$

wenn wir noch (5) berücksichtigen. — Andererseits haben wir für $(\omega, \omega') \in \widetilde{M}(k)$

$$\log \bar{f}_k(\omega, \omega') > |J_k|\, S,$$

$$f_k(\omega, \omega') > e^{|J_k| S} f'_k(\omega')$$

also für $\nu = 1, \ldots, N$

$$\int_{E'(\nu)} f_k\left(\omega(\nu), \omega'\right) m'(d\omega') > e^{|J_k| S} \int_{E'(\nu)} f'_k(\omega')\, m'(d\omega') = e^{|J_k| S}\, p'\left(E'(\nu)\right).$$

Summation ergibt — man verschenkt offenbar nicht viel bei dieser Abschätzung —

$$\begin{aligned} N &> e^{|J_k| S}\, p'(E') \\ &> e^{|J_k| S}\, \frac{\varepsilon}{2}\, p\left(M(k)\right) \\ &= e^{|J_k| \left(S + \frac{1}{|J_k|} \log\left(\frac{\varepsilon}{2} p(M(k))\right)\right)}. \end{aligned}$$

Wegen $p\left(M(k)\right) \to 1$ bleibt $\log\left(\frac{\varepsilon}{2}\, p\left(M(k)\right)\right)$ $(<0!)$ für hinreichend große k beschränkt. Wir wählen nun k_0 so groß, daß für $k \geqq k_0$ folgende Relationen gelten:

$$\frac{|J_k|}{|\hat{J}_k|} > 1 - 2\vartheta \qquad \frac{1}{|J_k|} \log\left(\frac{\varepsilon}{2} p\left(M(k)\right)\right) > -\vartheta .$$

Wegen (2) hat man dann

$$N(J_k, \varepsilon) \geqq N > e^{|\hat{J}_k|\left(L - \frac{\delta}{2}\right)},$$

also erst recht, wegen Satz 1.8

$$N(\hat{J}_k, \varepsilon) \geqq e^{|\hat{J}_k|\left(L - \frac{\delta}{2}\right)}.$$

Um nun auf $N(0, t-1, \varepsilon)$ zu kommen, nützen wir die Endlichkeit der Größe $G = \sup_k (|\hat{J}_{k+1}| - |\hat{J}_k|)$ aus. Zu jedem t mit $\langle 0, t\rangle \supseteqq J_1$ gibt es genau ein k mit

$$\hat{J}_k \subseteqq \langle 0, t-1\rangle \subset \hat{J}_{k+1}$$

also $0 \leqq t - |\hat{J}_k| < G$, woraus sich $|\hat{J}_k| = t\frac{|\hat{J}_k|}{t} \geqq t\left(1 - \frac{G}{t}\right)$ ergibt. Nun wähle man t_0 derart, daß für $t \geqq t_0$ $\langle 0, t\rangle \supseteqq \hat{J}_{k_0}$ nebst der Relation

$$\left(L - \frac{\delta}{2}\right)\left(1 - \frac{G}{t}\right) > L - \delta$$

gilt. Wir haben dann $k \geqq k_0$ und somit

$$N(0, t-1, \varepsilon) \geqq N(\hat{J}_k, \varepsilon) \geqq e^{t(L-\delta)} \qquad (t \geqq t_0),$$

wie behauptet.

§ 3. Schwache Umkehrung des Coding Theorem

Wir fixieren zunächst einen endlichen Zeitraum J, eine Kanaldichte $f_J(\omega, \omega')$ und endlich viele Punkte $\omega(1), \ldots, \omega(N) \in \Omega$ mit $f_J(\omega(\nu), .) \in M_J(P)$. Auf $B(\bar{J})$ betrachten wir zwei WVen:

1. Gleichverteilung q_0 über die Punkte $\omega(\nu)$:

$$q_0(E) = \sum_{\omega(\nu) \in E} \frac{1}{N} \qquad (E \in B(\bar{J})).$$

2. Wir wählen $\vartheta > 0$ beliebig, setzen $U'(\nu) = U'_\vartheta\left(f_J(\omega(\nu), .)\right) = \{g' \mid g' \in U'(J), |g', f_J(\omega(\nu), .)| < \vartheta\}$, $U(\nu) = \{\omega \mid f_J(\omega, .) \in U'(\nu)\}$. Voraussetzungsgemäß ist dann — mit den Bezeichnungen von § 1, Nr. 3

$$m\left(U(\nu)\right) = m\left(\bar{\psi}_J^{-1} U'(\nu)\right) = w'_J\left(U'(\nu)\right) > 0.$$

Wir setzen

$$f_J(\omega) = \frac{1}{N} \sum_{\nu=1}^{N} \frac{1}{m\left(U(\nu)\right)} \chi_{U(\nu)}(\omega).$$

Dann ist f_J die m-Dichte einer WV p_0 auf $B(\bar{J})$.

Wir vergleichen nun zwei Größen, die schon in ihrem Aufbau eine gewisse Verwandtschaft zeigen, nämlich:

1. Sei $H_J(q_0) = -\sum_{\nu=1}^{N} z\big(q_0(\{\omega(\nu)\})\big) = \log N$.

Wir bilden

$$g_J'(\omega') = \int q_0(d\omega)\, f_J(\omega, \omega')$$
$$= \frac{1}{N} \sum_{\nu=1}^{N} f_J\big(\omega(\nu), \omega'\big);$$

dies ist die m'-Dichte einer WV q_0' auf $B'(J)$; nun bilden wir für jedes ω' mit $g_J'(\omega') > 0$ die (Aposteriori-)WV $q_0(.\,|\,\omega')$, die durch

$$q_0\big(\{\omega(\nu)\}\,|\,\omega'\big) = \frac{f_J(\omega(\nu), \omega')}{N\, g_J'(\omega')} \qquad (\nu = 1, \ldots, N)$$

gegeben ist, ferner den Ausdruck

$$H_J(q_0\,|\,\omega') = -\sum_{\nu} z\big(q_0(\{\omega(\nu)\}\,|\,\omega')\big).$$

Diese Funktion von ω' ist immerhin q_0'-fastüberall erklärt, so daß die Bildung

$$H_J(q_0\,|\,q_0') = \int H_J(q_0\,|\,\omega')\, q_0'(d\omega')$$

sinnvoll ist. Man berechnet

$$H_J(q_0\,|\,q_0') = -\frac{1}{N} \sum_{\nu} z\big(f_J(\omega(\nu), \omega')\big)\, m'(d\omega')$$
$$+ \int z\big(g_J'(\omega')\big)\, m'(d\omega') + \log N.$$

Unser Interesse gilt dem Ausdruck

$$H_J(q_0) - H_J(q_0\,|\,q_0') = \frac{1}{N} \sum_{\nu} \int z\big(f_J(\omega(\nu), \omega')\big)\, m'(d\omega') - \int z\big(g_J'(\omega')\big)\, m'(d\omega').$$

2. Sei $R_J(p_0) = -\int z\big(f_J(\omega)\big)\, m(d\omega)$. Die Funktion

$$f_J'(\omega') = \int f_J(\omega)\, f_J(\omega, \omega')\, m(d\omega)$$

ist die m'-Dichte einer WV p_0' auf $B'(J)$. Mit den Abkürzungen

$$S_\nu(h) = \frac{1}{m(U(\nu))} \int_{U(\nu)} h\, d\, m$$

$$h_\nu'(\omega') = S_\nu\big(z(f_J(.\,, \omega'))\big)$$

berechnen wir die Größe

$$R_J(p_0\,|\,p_0') = -\int \left[\int z\Big(\frac{f_J(\omega)\, f_J(\omega, \omega')}{f_J'(\omega')}\Big)\, m\,(d\omega)\right] p_0'(d\omega').$$

Das innere Integral ist zwar nur für $f'_J(\omega') > 0$ erklärt. Dies genügt aber zur Berechnung eines p'_0-Integrals. Zunächst ist für $f'_J(\omega') > 0$

$$
\begin{aligned}
-\int z\left(\frac{f_J(\omega)\, f_J(\omega,\omega')}{f'_J(\omega')}\right) m(d\omega) \\
&= -\int \frac{f_J(\omega)\, f_J(\omega,\omega')}{f'_J(\omega')} \log \frac{1}{f'_J(\omega')}\, m(d\omega) - \\
&\quad -\int \frac{f_J(\omega)\, f_J(\omega,\omega')}{f'_J(\omega')} \log\big(f_J(\omega)\, f_J(\omega,\omega')\big)\, m(d\omega) \\
&= \log f'_J(\omega') - \\
&\quad - \frac{1}{f'_J(\omega')} \frac{1}{N} \sum_\nu \frac{1}{m(U(\nu))} \int_{U(\nu)} - z\big(f_J(\omega,\omega')\big)\, m(d\omega) - \\
&\quad - \frac{1}{f'_J(\omega')} \int f_J(\omega,\omega')\, z\big(f_J(\omega)\big)\, m(d\omega) \\
&= \log f'_J(\omega') - \frac{1}{f'_J(\omega')} \frac{1}{N} \sum_\nu h'_\nu(\omega') - \\
&\quad - \frac{1}{f'_J(\omega')} \int f_J(\omega,\omega')\, z\big(f_J(\omega)\big)\, m(d\omega).
\end{aligned}
$$

Hieraus ergibt sich nach FUBINI

$$
R_J(p_0 | p'_0) = \int z\big(f'_J(\omega')\big)\, m'(d\omega) - \frac{1}{N} \sum_\nu \int h'_\nu\, dm' - \int z\big(f_J(\omega)\big)\, m(d\omega)
$$

und deshalb

$$
R_J(p_0) - R_J(p_0 | p') = \frac{1}{N} \sum_\nu \int h'_\nu\, dm' - \int z(f'_J)\, dm'.
$$

Wir wollen jetzt q_0 durch p_0 approximieren, indem wir ϑ gegen 0 streben lassen, und die untenstehende Relation (1) erreichen. Hierzu benötigen wir die

ANNAHME (J). *Man kann die Kanaldichte $f_J(\omega,\omega')$ stets so wählen, daß die Menge*

$$
\{z(f_J(\omega,.)) \,|\, \omega \in \Omega\} \subseteqq L^1_{m'},
$$

gleichgradig m'-integrabel wird.

Hinreichend hiefür ist z.B., daß man die Kanaldichte $f_J(\omega,\omega')$ als beschränkte Funktion wählen kann. In §4 wird gezeigt, daß schon diese Klasse von Kanälen Kanäle mit beliebig hohen Durchlaßkapazitäten enthält. Gleichbedeutend mit der Annahme (J) ist die gleichgradige m'-Integrabilität der Menge $z\big(M_J(P)\big)$.

SATZ 3.1. Gilt die Annahme (J), so ist

$$
\lim_{\vartheta \to 0+0} \big(R_J(p_0) - R_J(p_0 | p'_0)\big) = H_J(q_0) - H_J(q_0 | q'_0). \tag{1}
$$

BEWEIS. a) Beachtet man, daß für $\omega \in U(\nu)$

$$
\begin{aligned}
\|z\big(f_J(\omega,.)\big) - z\big(f_J(\omega(\nu),.)\big)\| &\leqq |f_J(\omega,.), f_J\big(\omega(\nu),.\big)| \\
&< \vartheta
\end{aligned}
$$

gilt, so folgt durch Mitteln

$$\|z(f_J(\omega(\nu),.)) - h'_\nu(.)\| < \vartheta \qquad (\nu = 1, \ldots, N)$$

und damit

$$\lim_{\vartheta\to 0+0} \frac{1}{N} \sum_\nu \int h'_\nu\, dm' = \frac{1}{N} \sum_\nu \int z(f_J(\omega(\nu),.))\, dm'.$$

b) Dieselbe Methode liefert

$$\lim_{\vartheta\to 0+0} \|f'_J - g'_J\| = 0,$$

also gewiß

$$\lim_{\vartheta\to 0+0} f'_J = g'_J \qquad (m'\text{-stochastisch}).$$

Hieraus folgt

$$\lim_{\vartheta\to 0+0} z(f'_J) = z(g'_J) \qquad (m'\text{-stochastisch}),$$

und dies hat wegen der Annahme (J)

$$\lim_{\vartheta\to 0+0} \|z(f'_J) - z(g'_J)\| = 0$$

und somit

$$\lim_{\vartheta\to 0+0} \int z(f'_J)\, dm' = \int z(g'_J)\, dm'$$

zur Folge. Damit ist alles gezeigt.

Der wesentliche Gedanke beim Beweis des nachstehenden Satzes 3.2 besteht darin, daß die Fixierung von q_0 dieselbe Situation schafft wie bei einem Kanal mit einem endlichen Eingangsalphabet der Länge N. Wir präzisieren dies zunächst folgendermaßen.

Hat man eine WV p_0 auf $B(\bar{J})$, und ist $|\bar{J}| = d$, so gibt es genau eine WV p der Periode d, die auf $B(\bar{J})$ mit p_0 übereinstimmt und folgende Eigenschaft hat: Sind $k_1, \ldots, k_r$ verschiedene ganze Zahlen, so sind die Borel-Körper $B(\bar{J} + k_\varrho d) = B_\varrho$ unter p unabhängig, d.h. es gilt

$$p\left(\bigcap_{\varrho=1}^{r} E\right) = \prod_{\varrho=1}^{r} p(E_\varrho) \qquad (E_\varrho \in B_\varrho).$$

Genau so gewinnt man q aus q_0. Für das weitere wird jetzt folgende Annahme von Bedeutung:

ANNAHME (U). *Es gibt eine Zahl $g_0 \geqq 0$ mit folgender Eigenschaft: Sind J_ϱ $(\varrho = 1, \ldots, r)$ ganzzahlige Intervalle, die paarweise mindestens den Abstand $g_0 + 1$ haben, so sind für m-fastalle $\omega \in \Omega$ die Borel-Körper $B'(J_\varrho)$ unabhängig unter $P(\omega, .)$.*

Man kann $g_0 = g$ annehmen, indem man eine der beiden Zahlen notfalls vergrößert. Wir wollen dies im folgenden stets tun. Die Bedingung (U) ist invariant gegen den Übergang zu äquivalenten Kanälen; sie ist gleichbedeutend mit der Auffindbarkeit von Kanaldichten $f_{\bigcup_\varrho J_\varrho}$, f_{J_ϱ} mit

$$f_{\bigcup_\varrho J_\varrho} = \prod_\varrho f_{J_\varrho}.$$

Die am Schluß von 1 Nr. 1 konstruierten Kanäle erfüllen stets (U) mit $g_0=0$. Sind die dort verwendeten Funktionen $f(\xi_{-g}, \ldots, \xi_0; .)$ gemeinsam beschränkt, oder sind auch nur die $z(f(\xi_{-g}, \ldots, \xi_0; .))$ gleichgradig n'-integrabel, so gilt auch die Annahme (J) für jedes J.

Wir denken uns jetzt J immer noch festgehalten, setzen $d=|\bar{J}|$ und $J_k=\bigcup_{j=0}^{k-1}(J+jd)$ $(k=1, 2, \ldots)$. Die Borel-Körper $B(\bar{J}_k)$ und $B'(\bar{J}_k)$ kann man sich im Falle, daß J ein Intervall mit dem linken Endpunkt 0 ist, bequem mittels des folgenden Schemas vorstellen. Wir bilden wie in § 2 die Größen $f'_k(\omega')$,

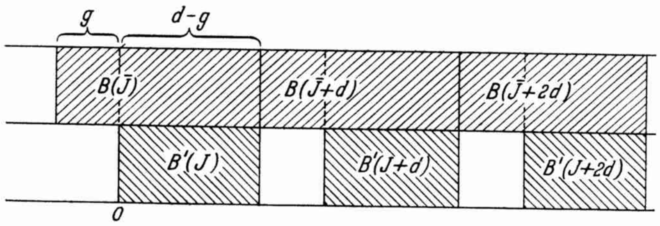

Fig. 1

$\bar{f}_k(\omega, \omega')$, ausgehend von der obengenannten WV p, die sich auf die in Nr. 2 konstruierte WV p_0 stützt und von einer Kanaldichte $f_k(\omega, \omega')$. Gestützt auf die Annahme (U) und die Konstruktion von p weist man sofort die Beziehungen

$$f_k(\omega, \omega') = \prod_{j=0}^{k-1} T^{jd} f_J(\omega, \omega') \qquad (\bar{m}\text{-fastüberall})$$

$$f'_k(\omega') = \prod_{j=0}^{k-1} T^{jd} f_J(\omega') \qquad (m'\text{-fastüberall})$$

$$\bar{f}_k(\omega, \omega') = \prod_{j=0}^{k-1} T^{jd} \bar{f}_J(\omega, \omega') \qquad (p\text{-fastüberall}),$$

sowie die Unabhängigkeit der $\tilde{B}(J+jd)$ unter $\tilde{p}$ nach. Hierbei wurde, wie üblich, $T^u g(\omega, \omega') = g(T^u\omega, T^u\omega')$ gesetzt. Ferner berechnet man

$$\begin{aligned}\int \log \bar{f}_J \, d\tilde{p} &= \int f_J(\omega) f_J(\omega, \omega') \log \frac{f_J(\omega, \omega')}{f'_J(\omega')} \, \bar{m}(d\omega, d\omega') \\ &= \int \left[\int z\left(\frac{f_J(\omega) f_J(\omega, \omega')}{f'_J(\omega')}\right) m(d\omega)\right] f'_J(\omega') \, m'(d\omega') - \\ &\quad - \int f_J(\omega) f_J(\omega, \omega') \log f_J(\omega) \, m(d\omega) \, m'(d\omega') \\ &= R_J(p_0) - R_J(p_0 | p'_0).\end{aligned}$$

Damit ergibt sich aus dem schwachen Gesetz der großen Zahl für unabhängige Zufallsvariable: Ist $S_0 < R_J(p_0) - R_J(p_0 | p'_0)$, so ist

$$\lim_k \tilde{p}\left(\left\{(\omega, \omega') \,\middle|\, \frac{1}{|J_k|} \log \bar{f}_k(\omega, \omega') > \frac{S_0}{|J|}\right\}\right) = 1. \tag{2}$$

Man hat hierbei nur die Beziehungen

$$|J_k| = k\,|J|$$

$$\log \bar{f}_k = \sum_{j=0}^{k-1} T^{jd}(\log \bar{f}_J)$$

zu beachten. Nun erinnern wir uns, daß in all diesen Bildungen noch folgende Parameter stecken: $\vartheta > 0$ und die Punkte $\omega(\nu)$ mit $f_J(\omega(\nu), .) \in M(P)$. Wir wollen nun annehmen, die $\omega(\nu)$ seien einem ε-Code $\{(\omega(\nu), E'(\nu)) \mid \nu = 1, \ldots, N\}$ für den Zeitraum J entnommen.

Satz 3.2. Es gibt ein $\vartheta > 0$ mit

$$(3) \qquad R_J(p_0) - R_J(p_0 | p_0') > \log N - \big(\varepsilon \log(N-1) - z(\varepsilon) - z(1-\varepsilon)\big).$$

Beweis. Wegen Satz 3.1 und $H_J(q_0) = \log N$ genügt es,

$$H_J(q_0 | q_0') < \varepsilon \log(N-1) - z(\varepsilon) - z(1-\varepsilon)$$

nachzuweisen. Dies aber folgt aus dem im Anhang bewiesenen Satz, wenn man dort

$$K' = B'(J), \quad p_\nu = \frac{1}{N}, \quad f(\nu, \omega') = f_J(\omega(\nu), \omega'),$$

$$Q(\nu, E') = \int_{E'} f(\nu, \omega')\, m'(d\omega')$$

setzt.

Satz 3.3 (Schwache Umkehrung des Coding Theorem). *Sei P ein stationärer g-Kanal vom Dichtetypus, der der Bedingung (U) und für jedes nichtleere endliche $J \subseteqq \Gamma$ auch der Bedingung (J) genügt. Sei L seine Durchlaßkapazität. Dann ist $L \geqq 0$, und es gibt zu jedem $\delta > 0$ ein $\varepsilon > 0$ und ein $t_0 > 0$ mit*

$$N(0, t-1, \varepsilon) < e^{t(L+\delta)} \qquad (t \geqq t_0).$$

Beweis. Wir halten $t > 0$ zunächst fest und setzen $J = \langle 0, t-1 \rangle$. Dann wählen wir zu einem zunächst beliebigen $\varepsilon > 0$ mit $\varepsilon < \frac{1}{e}$ und einer der Annahme (J) entsprechenden Kanaldichte $f_J(\omega, \omega')$ einen ε-Code $\{(\omega(\nu), E'(\nu)) \mid \nu = 1, \ldots, N\}$ für den Zeitraum J, mit $N = N(0, t-1, \varepsilon)$. Sodann bestimmen wir gemäß Satz 3.1 ein $\vartheta > 0$ derart, daß für das nunmehr festliegende p_0 die Relation (3) gilt. Aus (2) folgt dann

$$\begin{aligned}
&\lim_k p\Big(\Big\{(\omega, \omega') \,\Big|\, \frac{1}{|J_k|} \log \bar{f}_{J_k}(\omega, \omega') > \\
&> \frac{1}{t}\big[\log N - \big(\varepsilon \log(N-1) - z(\varepsilon) - z(1-\varepsilon)\big)\big]\Big\}\Big) \\
&= 1.
\end{aligned}$$

Offenbar gilt für die Folge $\mathfrak{J} = \{J_k\}$ die Beziehung

$$|\hat{J}_{k+1}| - |\hat{J}_k| = t + g$$

sowie

$$d(\mathfrak{J}) \geqq 1 - \frac{g}{t}.$$

Offenbar ist jedes

$$S < \limsup_t \frac{1}{t}\left[\log N(0, t-1, \varepsilon) - \left(\varepsilon \log (N(0, t-1, \varepsilon) - 1) - z(\varepsilon) - z(1-\varepsilon)\right)\right]$$

zulässig im Sinne der Definition 2. Für $\lim_t N(0, t-1, \varepsilon) < \infty$ ist hier die rechte Seite $=0$ und Satz 3.3 ist bewiesen (nach Satz 2.1 kann $L > 0$ nicht in Frage kommen). Für $\lim_t N(0, t-1, \varepsilon) = \infty$ ist offenbar

$$\begin{aligned}\limsup \frac{1}{t}&\left[\log N(0, t-1, \varepsilon) - \left(\varepsilon \log (N(0, t-1, \varepsilon) - 1) - z(\varepsilon) - z(1-\varepsilon)\right)\right] \\ &= \limsup_t \left[\frac{1}{t} \log N(0, t-1, \varepsilon)\right]\left[1 - \varepsilon - \frac{z(\varepsilon) - z(1-\varepsilon)}{\log N(0, t-1, \varepsilon)}\right] \\ &= (1-\varepsilon) \limsup_t \frac{1}{t} \log N(0, t-1, \varepsilon)\end{aligned}$$

und wir erhalten

$$L \geqq (1-\varepsilon) \limsup_t \frac{1}{t} \log N(0, t-1, \varepsilon).$$

Für $L = \infty$ ist nichts zu beweisen. Ist aber $L < \infty$, und $\delta > 0$ vorgegeben, so bestimme man $\varepsilon \in \left(0, \frac{1}{e}\right)$ derart, daß

$$\frac{L}{1-\varepsilon} + \frac{\delta}{2} < L + \delta$$

gilt und sodann $t_0 > 0$ derart, daß für $t \geqq t_0$

$$\frac{1}{t} \log N(0, t-1, \varepsilon) < \limsup_s \frac{1}{s} \log N(0, s-1, \varepsilon) + \frac{\delta}{2}$$

gilt. Man hat dann insgesamt für $t \geqq t_0$ die behauptete Abschätzung.

§ 4. Beispiele

Es bleibt nachzuweisen, daß die durch die Bedingungen (U) und (J) ($J \leqq \Gamma$) gekennzeichnete Klasse von Kanälen hinreichend umfangreich ist, also etwa nicht-pathologische Kanäle mit beliebig großer Durchlaßkapazität enthält. Zur Konstruktion solcher Kanäle verwenden wir das am Schluß von §1, Nr. 1 angegebene Verfahren.

Sei $N > 1$ beliebig und seien $\Omega(-g, 0) = \sum_{\nu=1}^{N} \Omega_\nu$, $A' = \sum_{\nu=1}^{N} A'_\nu$ disjunkte Zerlegungen mit

$$\text{(1)} \qquad \begin{cases} \Omega_\nu \in C(-g, 0), & n_0^{g+1}(\Omega_\nu) > 0 \\ A'_\nu \in C', & n_0'(A'_\nu) > 0. \end{cases}$$

Solche Zerlegungen gibt es sicher, wenn A und A' etwa Intervalle mit dem Lebesgue-Maß n_0 bzw. n' darauf sind. Seien $f'_\nu(\xi') \geqq 0$ auf A' erklärte Funktionen mit

$$\int f'_\nu \, dn' = 1, \qquad \int z(f'_\nu) \, dn' < \infty \quad \text{und} \quad f'_\nu(\xi') = 0 \qquad (\xi' \notin A'_\nu).$$

Dann ist durch

$$f(\xi_{-g}, \ldots, \xi_0; \xi'_0) = f'_\nu(\xi'_0) \qquad ((\xi_{-g}, \ldots, \xi_0) \in \Omega_\nu)$$

ein Kanal gegeben, der allen Forderungen genügt und offenbar $N(0, t-1, \varepsilon) \geqq N^t$ $(t > 0)$ erfüllt. Nach Satz 3.3 ist hier $L \geqq \log N$.

Daß die Bedingungen (J) $(J \subseteqq \Gamma)$ für die Gültigkeit der Aussage von Satz 3.3 nicht notwendig ist, kann man zeigen, indem man das Beispiel folgendermaßen modifiziert: Man wählt *abzählbare* Zerlegungen $\Omega(-g, 0) = \sum_{\nu=1}^{\infty} \Omega_\nu$ und $A' = \sum_{\nu=1}^{\infty} A'_\nu$ mit den Eigenschaften (1) und erhält offenbar zu jedem $\varepsilon > 0$ beliebig lange ε-Codes in jedem Zeitraum. Nun wähle man ein $N > 1$ und führe die Überlegungen des §3 unter Beschränkung auf $\sum_{\nu=1}^{N} \Omega_\nu$ bzw. $\sum_{\nu=1}^{N} A'_\nu$ durch. Man bemerkt, daß z. B. das, was man von der Annahme (J) braucht, um Satz 3.1 zu beweisen, auch hier erfüllt ist: für $\vartheta < 1$ ist sogar $f'_J = g'_J$. Man erhält $L = \infty$, d.h. die Aussage des Satz 3.3 ist wiederum erfüllt.

Anhang. Verallgemeinerung von FANO'S Lemma

Sei Ω' eine beliebige Menge und K' ein Borel-Körper in Ω'. $Q(\nu, .)$ $(\nu = 1, \ldots, N)$ seien WVen auf K'. Sei m' eine WV auf K', bezüglich welcher alle $Q(\nu, .)$ totalstetig sind $\left(\text{eine solche gibt es stets, z. B. } m' = \frac{1}{N} \sum_\nu Q(\nu, .)\right)$. Sei $f(\nu, \omega')$ eine m'-Dichte von $Q(\nu, .)$. Mit irgendwelchen Zahlen $p_\nu \geqq 0$, die $\sum_\nu p_\nu = 1$ erfüllen, bilden wir

$$p'(.) = \sum p_\nu Q(\nu, .)$$

mit der m'-Dichte

$$g'(\omega') = \sum p_\nu f(\nu, \omega'),$$

sowie

$$q(\nu, \omega') = \frac{p_\nu f(\nu, \omega')}{g'(\omega')},$$

$$H = -\int \left(\sum_\nu z\left(q(\nu, \omega')\right)\right) p'(d\omega').$$

Dann gilt folgende Verallgemeinerung eines bekannten Lemmas von FANO:

SATZ. *Sei* $0 < \varepsilon < \frac{1}{e}$, *es gebe* N *paarweise disjunkte Mengen* $E'_\nu \in K'$ *mit*

$$Q(\nu, E'_\nu) = \int_{E'_\nu} f(\nu, \omega')\, m'(d\omega') > 1 - \varepsilon \qquad (\nu = 1, \ldots, N).$$

Dann gilt

$$H < \varepsilon \log(N-1) - z(\varepsilon) - z(1-\varepsilon).$$

BEWEIS. Man kann $\sum_\nu E'_\nu = \Omega'$ annehmen. Setzt man

$$h'_\lambda(\omega') = \sum_{\nu \neq \lambda} q(\nu, \omega') \qquad (\nu = 1, \ldots, N),$$

so wird

$$\sum_\lambda \int_{E'_\lambda} h'_\lambda(\omega')\, p'(d\omega') = \sum_\lambda \sum_{\nu \neq \lambda} \int_{E'_\lambda} p_\nu f(\nu, \omega')\, m'(d\omega')$$
$$= \sum_\nu \sum_{\lambda \neq \nu} p_\nu Q(\nu, E'_\lambda) = \sum_\nu p_\nu Q(\nu, \Omega' - E'_\nu) < \varepsilon,$$

sowie

$$\sum_{\lambda} \int_{E'_\lambda} q(\lambda, \omega')\, p'(d\omega') = \sum_{\lambda} p_\lambda\, Q(\lambda, E'_\lambda) > 1 - \varepsilon.$$

Nimmt man bekannte Abschätzungen (die fast alle auf der Konvexität von $z(x)$ beruhen) zu Hilfe, so bekommt man

$$\begin{aligned}
H &= -\sum_{\lambda} \int_{E'_\lambda} \Big(z\big(q(\lambda,\omega')\big) + \sum_{\nu \neq \lambda} z\big(q(\nu,\omega')\big) \Big)\, p'(d\omega') \\
&= -\sum_{\lambda} p'(E'_\lambda) \frac{1}{p'(E'_\lambda)} \int_{E'_\lambda} z\big(q(\lambda,\omega')\big)\, p'(d\omega') - \\
&\quad - \sum_{\lambda} \int_{E'_\lambda} \Big(h'_\lambda(\omega') \sum_{\nu \neq \lambda} z\Big(\frac{q(\nu,\omega')}{h'_\lambda(\omega')}\Big)\Big)\, p'(d\omega') - \\
&\quad - \sum_{\lambda} p'(E'_\lambda) \frac{1}{p'(E'_\lambda)} \int_{E'_\lambda} z\big(h'_\lambda(\omega')\big)\, p'(d\omega') \leqq \\
&\leqq -\sum_{\lambda} p'(E'_\lambda)\, z\Big(\frac{1}{p'(E'_\lambda)} \int_{E'_\lambda} q(\lambda,\omega')\, p'(d\omega') \Big) + \\
&\quad + \sum_{\lambda} \int_{E'_\lambda} h'_\lambda(\omega')\, p'(d\omega') \log(N-1) - \\
&\quad - \sum_{\lambda} p'(E'_\lambda)\, z\Big(\frac{1}{p'(E'_\lambda)} \int_{E'_\lambda} h'_\lambda(\omega')\, p'(d\omega') \Big) \leqq \\
&\leqq -z\Big(\sum_{\lambda} \int_{E'_\lambda} q(\lambda,\omega')\, p'(d\omega') \Big) + \varepsilon \log(N-1) - \\
&\quad - z\Big(\sum_{\lambda} \int_{E'_\lambda} h'_\lambda(\omega')\, p'(d\omega') \Big) < \\
&< -z(1-\varepsilon) + \varepsilon \log(N-1) - z(\varepsilon).
\end{aligned}$$

Literatur

[1] ROSENBLATT-ROTH, M.: Die Entropie stochastischer Prozesse. Dokl. Akad. Nauk UdSSR **112**, 16–19 (1957) [Russ.].

[2] — Theorie der Informationsübertragung durch stochastische Kanäle. Dokl. Akad. Nauk UdSSR **112**, 202–205 (1957) [Russ.].

[3] WOLFOWITZ, J.: The coding of messages subject to chance errors. Ill. J. Math. **3**, 591–606 (1957).

[4] — Strong converse of the Coding Theorem for semicontinuous channels. Ill. J. Math. **3**, 477–489 (1959).

[5] — Coding Theorems of Information Theory. Ergebn. d. Math. und ihrer Grenzgeb. N. F., Heft 31. Berlin-Göttingen-Heidelberg 1961.

Institut für Mathematische Statistik an der Universität Göttingen

(Eingegangen am 20. Dezember 1960)