

Beiträge zur Shannonschen Informationstheorie im Falle nichtstationärer Kanäle *

RUDOLF AHLWEDE

Eingegangen am 15. April 1967

Abstract. Coding theorem and strong converse of the coding theorem are proved for nonstationary semicontinuous channels and for almost periodic channels with unknown phase. More general cases are considered.

Inhaltsverzeichnis

Einleitung	1
Kapitel I. Das Codingtheorem und seine starke Umkehrung für halbstetige nichtstationäre Kanäle mit unabhängigen Zeichen	3
§ 1. Grundbegriffe	4
§ 2. Eine allgemeine untere Abschätzung der maximalen Codelängen	5
§ 3. Beweis des Codingtheorems	6
§ 4. Beweis einer starken Umkehrung des Codingtheorems für nichtstationäre halbstetige Kanäle	10
Kapitel II. Translationsinvariante Codes fastperiodischer Kanäle	17
§ 1. Allgemeines über Simultankanäle	17
§ 2. Herleitung des Codingtheorems für einen fastperiodischen diskreten Kanal mit unabhängigen Zeichen bei zeittranslationsinvarianten Codes mit Hilfe des Maximalcodesatzes für nichtstationäre Simultankanäle	18
§ 3. Ein Beweis des Codingtheorems mit Hilfe einer Verallgemeinerung des π -Sequenzen-Verfahrens von WOLFOWITZ	22
§ 4. Die starke Umkehrung des Codingtheorems für den fastperiodischen Simultankanäle	33
Kapitel III. Nichtstationäre Simultankanäle	34
§ 1. Das Codingtheorem	34
§ 2. Die starke Umkehrung des Codingtheorems	36
Kapitel IV. Reichweite des Mittelungsverfahrens, Versagen der üblichen Kapazität-funktion, Beispiele	39
Literatur	41

Einleitung

Die Theorie stationärer Kanäle mit unabhängigen Zeichen und endlichem Eingangs- und Ausgangsalphabet ist von SHANNON [14], FEINSTEIN [5], WOLFO-

* Die Arbeit entstand teilweise im Rahmen eines von der Deutschen Forschungsgemeinschaft finanzierten Forschungsvorhabens.

WITZ [19], DOBRUSHIN [4] und STRASSEN [17] weit entwickelt worden. In vielen Fällen der Praxis sind die Störungseigenschaften der Nachrichtenübertragungsvorrichtungen (Kanäle) aber nicht zeitstationär. Man denke etwa an Funkübertragungen eines künstlichen Erdsatelliten nach festen Erdstationen. Derartige Nachrichtenübermittlungsanlagen bildeten für JACOBS den empirischen Anlaß, den mathematischen Begriff eines fastperiodischen Kanals einzuführen [8], [9]. JACOBS und der Verfasser haben unabhängig voneinander für fastperiodische Kanäle mit unabhängigen Zeichen und endlichen Alphabeten eine starke Umkehrung des Codingtheorems bewiesen (ohne Beweis angegeben in [10]). In [6], [21] wurden einige weitere Resultate für fastperiodische Kanäle hergeleitet. In Kapitel I der vorliegenden Arbeit wird für nichtstationäre Kanäle mit unabhängigen Zeichen und mit endlichem Eingangs- und beliebigem Ausgangsalphabet ein Codingtheorem (§ 3) und dessen starke Umkehrung (§ 4) bewiesen.

In der stationären Situation hat zuerst WOLFOWITZ [19] eine starke Umkehrung für den halbstetigen Kanal mit unabhängigen Zeichen bewiesen. KEMPERMAN [19] hat einen einfacheren Beweis gegeben. Unser Beweis der starken Umkehrung für den nichtstationären Fall orientiert sich am Beweis KEMPERMANS.

Das in Kap. I entwickelte nichtstationäre Modell hat gegenüber dem stationären den Nachteil, daß die Güte eines Codes nicht invariant gegenüber Zeittranslationen ist; ein λ -Code geht nicht notwendig in einen λ -Code über.

In Kap. II wird für translationsinvariante λ -Codes eines fastperiodischen Kanals Codingtheorem und starke Umkehrung des Codingtheorems bewiesen.

In Kap. III wird unter gewissen Kompaktheitsannahmen ein Codingtheorem und dessen starke Umkehrung für allgemeine nichtstationäre Simultankanäle bewiesen, in Kap. IV werden die Grenzen der in Kap. II, III benutzten Verfahren aufgezeigt.

Die Resultate aus Kap. II, III lassen sich auf eine weite Klasse von „zusammengesetzten“ Kanälen anwenden. Das habe ich bisher durchgeführt in den folgenden Fällen:

1. Nichtstationäre Simultankanäle, deren jeweilig steuernder Kanal
 - a) dem Empfänger,
 - b) dem Sender,
 - c) dem Empfänger und dem Sender

bekannt ist.

In der stationären Situation wurden diese Kanäle von WOLFOWITZ untersucht ([19]).

2. Nichtstationäre randomisierte Kanäle.

In der stationären Situation wurden diese Kanäle von SHANNON und WOLFOWITZ untersucht ([19]).

3. Kanäle mit Gedächtnis.

Insbesondere lassen sich die Ergebnisse von BLACKWELL, BREIMANN, THOMASIAN und WOLFOWITZ über Zustandskanäle ([2], [19]) auf den fastperiodischen Fall übertragen, falls die die Auswahl steuernde fastperiodische Markoffkette $(\omega^t(i|j))_{t=1,2,\dots}$ der Bedingung genügt:

$$\omega^t(i|j) \geq \alpha > 0 \quad \text{für alle } t, i, j.$$

4. KESTEN hat für stationäre halbstetige Simultankanäle Codingtheorem und schwache Umkehrung bewiesen ([12]). Die Ergebnisse aus Kap. II, III lassen sich auf nichtstationäre halbstetige Simultankanäle anwenden und liefern ebenfalls Codingtheorem und schwache Umkehrung.

Die unter 1., 2., 3., 4. genannten Ergebnisse werden in einer Fortsetzung dieser Arbeit bewiesen.

Kapitel I. Das Codingtheorem und seine Umkehrung für halbstetige nichtstationäre Kanäle mit unabhängigen Zeichen

Das nachstehende Kapitel befaßt sich mit halbstetigen (endliches Eingangsalphabet Ω und beliebiges Ausgangsalphabet Ω') Kanälen mit unabhängigen Zeichen, die nicht notwendig stationär sind (Definition in § 1), es wird ein Codingtheorem (Satz 1, § 3) und dessen starke Umkehrung (Satz 2, § 4) bewiesen. Das Codingtheorem wird hergeleitet mit Hilfe des Maximalcodesatzes (§ 2, Satz 0). Im stationären Fall hat sich die Informationsfunktion im Zeitraum $\langle 1, n \rangle$

$$\log \prod_{t=1}^n m(\cdot | x^t)$$

(Definition in § 3) zur asymptotischen Beschreibung der maximalen Codelängen als brauchbar erwiesen. Es liegt daher nahe, im nichtstationären Fall mit der Informationsfunktion

$$\log \prod_{t=1}^n m^t(\cdot | x^t)$$

zu arbeiten. Dabei werden Erwartungswert und Varianzabschätzungen der Informationsfunktion benötigt, die wesentlich von der Unabhängigkeit Gebrauch machen.

KEMPERMAN benutzt in seinem Beweis der starken Umkehrung für stationäre Kanäle die Gleichung

$$\begin{aligned} (*) \quad E \prod_{t=1}^n \log m(\cdot | x^t) &= n \sum_{i=1}^a \pi_i E \log m(\cdot | i) \\ &= n \tilde{E} \log m(\cdot | \cdot) \quad (\text{vgl. 8.4.6 in [19]}). \end{aligned}$$

Das Hauptproblem im nichtstationären Fall besteht darin, eine zu (*) „analoge“ Gleichung zu finden. Die Lösung bringt Hilfssatz 2 in § 4.

Die Schar der Übergangswahrscheinlichkeiten $F^t(\cdot | \cdot)$ ($t = 1, 2, \dots$) wird so in endlich viele Klassen zerlegt, daß die Informationsfunktionen der Elemente einer Klasse in der Supremumsnorm (L^∞) „benachbart“ sind; der nichtstationäre Kanal wird ausschnittsweise „stationarisiert“. Diese Zerlegung wird ermöglicht durch die Beobachtung, daß die Menge der Informationsfunktionen halbstetiger Kanäle in der Supremumsnorm totalbeschränkt ist.

Die Nichtstationarität hat zur Folge, daß man nicht mehr wie üblich mit einer konstanten Durchlaßkapazität auskommt. An ihre Stelle tritt eine Funktion

$$C(n) = \frac{1}{n} \sum_{t=1}^n C^t.$$

$C(n)$ wird durch 0 nach unten und durch $\log a$ nach oben beschränkt.

§ 1. Grundbegriffe

Seien Ω^t ($t = 1, 2, \dots$) Exemplare eines endlichen Alphabetes $\Omega = \{1, \dots, a\}$ und seien (Ω^t, B^t) ($t = 1, 2, \dots$) Exemplare eines beliebigen Meßbarkeitsbereiches (Ω', B') , d.h. Ω' ist eine beliebige Menge und B' ist ein Borelkörper in Ω' .

Seien $F^t(\cdot|1), \dots, F^t(\cdot|a)$ Wahrscheinlichkeitsverteilungen auf (Ω^t, B^t) ($t = 1, 2, \dots$).

Für jedes $u = (x^1, \dots, x^n) \in \Omega_n = \prod_{t=1}^n \Omega^t$ ist durch $F_n(\cdot|u) = \prod_{t=1}^n *F^t(\cdot|x^t)$ eine Wahrscheinlichkeitsverteilung auf

$$\left(\Omega'_n = \prod_{t=1}^n \Omega'^t, B'_n = \prod_{t=1}^n *B'^t \right)$$

definiert. Bei Betrachtung der so gebildeten Folge $F_n(\cdot|\cdot)$ von Übergangskernen von Ω_n nach Ω'_n ($n = 1, 2, \dots$) spricht man von einem halbstetigen (i. a. nicht-stationären) Kanal mit unabhängigen Zeichen. $F^t(\cdot|\cdot)$ heißt die Übergangsfunktion (des Kanals) für die Komponente t . Ist $\Omega' = \Omega'^t$ endlich, so spricht man von einem diskreten Kanal. Die Übergangsfunktion $F^t(\cdot|\cdot)$ ist dann durch die Übergangswahrscheinlichkeiten $F^t(k|i) = F^t(\{k\}|i)$ ($i \in \Omega, k \in \Omega'$) eindeutig festgelegt.

(0) Ein Code (n, N, λ) für den halbstetigen Kanal ist eine Menge

$$\{(u_1, A'_1), \dots, (u_N, A'_N)\},$$

in der $u_i \in \Omega_n$ und die A'_i disjunkte Mengen aus B'_n sind und für die $F_n\{A'_i|u_i\} \geq 1 - \lambda$ gilt. ($0 \leq \lambda \leq 1$) ($i = 1, 2, \dots, N$).

Man spricht von einem Code für den Zeitraum $\langle 1, n \rangle$, von der Länge N , zur Sicherheits- (oder Irrtums-) wahrscheinlichkeit λ , auch von einem λ -Code. Die maximale Länge eines Code (n, N, λ) wird mit $N(n, \lambda)$ bezeichnet.

(1) Eine bezüglich n asymptotische untere Abschätzung von $N(n, \lambda)$, die mittels einer Funktion $C(n)$ ($n = 1, 2, \dots$) in folgender Form gegeben ist:

Zu beliebigen λ, δ mit $0 < \lambda < 1, \delta > 0$ gibt es ein $n_0 = n_0(\lambda, \delta)$, derart, daß für $n \geq n_0$

$$N(n, \lambda) > \exp\{n(C(n) - \delta)\}$$

gilt

wird als ein Codingtheorem mit der Durchlaß-Kapazität(sfunktion) $C(n)$ bezeichnet.

(2) Hat man ein Codingtheorem mit einer Durchlaßkapazität $C(n)$ ($n = 1, 2, \dots$), so heißt eine bezüglich n asymptotische obere Abschätzung von $N(n, \lambda)$, die in folgender Form gegeben ist:

Zu beliebigen λ, δ mit $0 < \lambda < 1, \delta > 0$ gibt es ein $n_0 = n_0(\lambda, \delta)$, derart, daß für $n \geq n_0$

$$N(n, \lambda) < \exp\{n(C(n) + \delta)\}$$

gilt;

eine starke Umkehrung des betreffenden Codingtheorems.

Gelten mit einer Funktion $C(n)$ (1) und (2), so gelten für eine Funktion $C'(n)$

(1) und (2) genau dann, wenn gilt:

$$\lim_{n \rightarrow \infty} (C(n) - C'(n)) = 0.$$

§ 2. Eine allgemeine untere Abschätzung der maximalen Codelängen

Wir betrachten eine endliche Menge $\Omega = \{u \mid u = 1, \dots, a\}$, darin den Borelkörper B aller Teilmengen und eine Wahrscheinlichkeitsverteilung $\pi = (\pi_1, \dots, \pi_a)$, ferner eine beliebige Menge $\Omega' = \{u', \dots\}$ mit einem Borelkörper B' und darauf a Wahrscheinlichkeitsverteilungen $F(\cdot \mid 1), \dots, F(\cdot \mid a)$. π und die $F(\cdot \mid 1), \dots, F(\cdot \mid a)$ induzieren auf (Ω', B') die Wahrscheinlichkeitsverteilung

$$\nu(\cdot) = \sum_{u \in \Omega} \pi_u F(\cdot \mid u)$$

und in $\tilde{\Omega} = \Omega \times \Omega' = \{\tilde{u} = (u, u') \mid u \in \Omega, u' \in \Omega'\}$ (mit dem Produkt-Borelkörper $\tilde{B} = B \times B'$)

- a) die Produktverteilung $\tilde{F} = \pi \times \nu$,
 b) die Wahrscheinlichkeitsverteilung $\tilde{F} = \pi \times F$, gegeben durch die Formel

$$\tilde{F}(\tilde{A}) = \sum_{u \in \Omega} \pi_u F(\tilde{A}_u \mid u),$$

wobei $\tilde{A} \in \tilde{B}$ ist und

$$\tilde{A}_u = \{u' \mid u' \in \Omega', (u, u') \in \tilde{A}\}$$

den u -Schritt von \tilde{A} bezeichnet.

Man beachte nun, daß \tilde{F} bezüglich \tilde{F} totalstetig ist, so daß man die Radon-Nikodym-Dichte $\frac{d\tilde{F}}{d\tilde{F}}(\tilde{u})$ bilden kann.

Sei nun für beliebiges reelles S

$$\tilde{M} = \tilde{M}(\pi, F, S) = \left\{ \tilde{u} = (u, u') \mid \log \frac{d\tilde{F}}{d\tilde{F}}(\tilde{u}) > S \right\}.$$

Wir betrachten jetzt für ein festes λ mit $0 < \lambda < 1$ alle möglichen Systeme von Paaren (u_j, A'_j) ($1 \leq j \leq N$), für die $u_1, \dots, u_N \in \Omega$ und die A'_j paarweise disjunkte Mengen aus B' sind, sowie $F(A'_j \mid u_j) \geq 1 - \lambda$ gilt. Den Maximalwert, den N für solche Systeme annehmen kann, bezeichnen wir mit $N(\lambda)$. Dann gilt der

Satz. (Methode der maximalen Codes für halbstetige Kanäle). Für jedes positive $\lambda < 1$ gilt

$$N(\lambda) > e^S \{ \tilde{F}(\tilde{M}) - (1 - \lambda) \} = \exp[S + \log \{ \tilde{F}(\tilde{M}) - (1 - \lambda) \}]$$

(wobei der letzte Term nur für $\tilde{F}(\tilde{M}) - (1 - \lambda) > 0$ sinnvoll ist).

Beweis. Sei (u_i, A'_i) ($1 \leq i \leq N$) eines der oben beschriebenen Systeme mit

- (1) $A'_i \subset \tilde{M}_{u_i}$
 (2) $\pi_{u_i} > 0$

und mit unter diesen zusätzlichen Nebenbedingungen maximaler Länge.

Wir setzen $A' = \bigcup_{i=1}^N A'_i$. Dann gilt π -fast überall

(3) $F(\tilde{M}_{u_i} - A' \mid u) \leq 1 - \lambda,$

denn die Ungültigkeit von (3) für ein u mit $\pi_u > 0$ würde die Verlängerbarkeit des obigen Codes bedeuten.

Summation mit π liefert

$$\sum_{u \in \Omega} \pi_u F(\tilde{M}_u - A' | u) \leq 1 - \lambda.$$

Daraus folgt direkt

$$\sum_{u \in \Omega} \pi_u F(\tilde{M}_u - A' | u) = \sum_{u \in \Omega} \pi_u F(\tilde{M}_u | u) - \sum_{u \in \Omega} \pi_u F(A' | u) \leq 1 - \lambda$$

und nach Definition von \tilde{F} und ν

$$(4) \quad \nu(A') \geq \tilde{F}(\tilde{M}) - (1 - \lambda).$$

Wegen

$$\frac{d\tilde{F}}{d\tilde{F}}(u_i, u') = \frac{dF(\cdot | u_i)}{d\nu(\cdot)}(u') \quad (\pi_{u_i} > 0)$$

gilt

$$F(A'_i | u_i) = \int_{A_i} \frac{d\tilde{F}}{d\tilde{F}}(u_i, u') \nu(du') \quad (\pi_{u_i} > 0).$$

Nach Definition von \tilde{M} und A'_i und da $F(D' | u_i) \leq 1$ ist ($D' \in B'$, $u_i \in \Omega$), folgt

$$1 \geq F(A'_i | u_i) = \int_{A_i} \frac{d\tilde{F}}{d\tilde{F}}(u_i, u') \nu(du') \geq e^S \nu(A'_i) \quad (\pi_{u_i} > 0).$$

Durch Summation ergibt sich

$$N \geq e^S \nu(A') \geq e^S (\tilde{F}(\tilde{M}) - (1 - \lambda)).$$

Damit ist der Satz bewiesen.

§ 3. Beweis des Codingtheorems

Wir kehren nun zu der in § 1 erklärten Situation zurück. Für jedes $t = 1, 2, \dots$ sei $\pi^t = (\pi_1^t, \dots, \pi_a^t)$ eine Wahrscheinlichkeitsverteilung auf Ω^t , und seien $F^t(\cdot | 1) \dots F^t(\cdot | a)$ Wahrscheinlichkeitsverteilungen auf Ω'^t . $\pi^t, F^t(\cdot | 1) \dots F^t(\cdot | a)$ erzeugen auf (Ω^t, B^t) das Maß

$$\nu^t(\cdot) = \sum_{i=1}^a \pi_i^t F^t(\cdot | i).$$

Für $i \in H^t = \{i | \pi_i^t > 0\}$ sei $m^t(\cdot | i)$ die Radon-Nikodym-Dichte von $F^t(\cdot | i)$ bezüglich ν^t . $m^t(\cdot | i)$ sei auf Nullmengen so festgesetzt, daß $\sum_{i \in H^t} \pi_i^t m^t(\cdot | i) = 1$ gilt. Dann gilt

$$(1) \quad 0 \leq m^t(\cdot | i) \leq \frac{1}{\pi_i^t} \quad (t = 1, 2, \dots) \quad (i \in H^t).$$

Wir betrachten nun für jedes $i \in H^t$ die F^t -fastüberall auf Ω'^t definierte Funktion

$$V_i^t(\cdot) = \log m^t(\cdot | i).$$

π^t und F^t erzeugen auf dem Produktraum

$$\tilde{\Omega}^t = \Omega^t \times \Omega'^t = \{(i, x') | i \in \Omega^t, x' \in \Omega'^t\}$$

(mit dem Produktborelkörper $B^t \times B'^t$, wobei B^t das System aller Teilmengen der endlichen Menge Ω^t ist) eine Wahrscheinlichkeitsverteilung \tilde{F}^t , wobei für $\tilde{A} \in B^t$ — unter Verwendung der Schnitte $\tilde{A}_i = \{x' | x' \in \Omega^t, (i, x') \in \tilde{A}\} \in B'^t$ —

$$\tilde{F}^t(\tilde{A}) = \sum_{i=1}^a \pi_i^t F^t(\tilde{A}_i | i)$$

ist.

Wir schreiben auch $\tilde{F}^t(\cdot | \cdot)$, wobei der zweite Argumentpunkt jeweils mit einem $i \in \Omega^t$ zu besetzen ist. Die oben erklärten Funktionen V_i^t liefern uns nun vermöge

$$\tilde{V}^t(x' | i) = V_i^t(x') \quad (i \in H^t, x' \in \Omega^t)$$

eine \tilde{F}^t -fastüberall erklärte Funktion \tilde{V}^t auf $\tilde{\Omega}^t$. Wir interessieren uns nun für Erwartungswert und Varianz von $V_i^t(\cdot)$ bezüglich $F^t(\cdot | i)$ und Erwartungswert und Varianz von $\tilde{V}^t(\cdot | \cdot)$ bezüglich $\tilde{F}^t(\cdot | \cdot)$.

a) Erwartungswerte

1. Erwartungswert $E_i^t V_i^t$ von V_i^t bezüglich $F^t(\cdot | i)$:

$$\begin{aligned} E_i^t V_i^t &= \int_{\Omega^t} \log m^t(\cdot | i) dF^t(\cdot | i) \\ &= \int_{\Omega^t} (\log m^t(\cdot | i)) m^t(\cdot | i) d\nu^t(\cdot). \end{aligned}$$

Wegen (1) existiert $E V_i^t$ ($i \in H^t$).

Wegen $x \log x \geq -\frac{1}{e}$ für $x \geq 0$ erhält man sofort die Abschätzung

$$(2) \quad E_i^t V_i^t \geq -\frac{1}{e} \quad (t = 1, 2, \dots) \quad (i \in H^t).$$

Andererseits ist wegen (1)

$$(3) \quad E_i^t V_i^t = \int_{\Omega^t} \log m^t(\cdot | i) dF^t(\cdot | i) \leq -\log \pi_i^t \quad (i \in H^t).$$

2. Erwartungswert $\tilde{E}^t \tilde{V}^t$ von \tilde{V}^t bezüglich $\tilde{F}^t(\cdot | \cdot)$:

$$\begin{aligned} \tilde{E}^t \tilde{V}^t &= \sum_{i \in H^t} \pi_i^t \int_{\Omega^t} \log m^t(\cdot | i) dF^t(\cdot | i) \\ \tilde{E}^t \tilde{V}^t &= \sum_{i \in H^t} \pi_i^t E_i^t V_i^t \quad (\text{nach 1.}) \\ &\leq -\sum_{i=1}^a \pi_i^t \log \pi_i^t. \end{aligned}$$

Aus der Konvexität von $f(x) = x \log x$ ($x \geq 0$) folgt nun

$$\tilde{E}^t \tilde{V}^t \leq -a \left(\sum_{i=1}^a \frac{\pi_i^t}{a} \log \frac{\pi_i^t}{a} \right) = -a \cdot \frac{1}{a} \log \frac{1}{a} = \log a.$$

Wie üblich definiert man nun die Informationsfunktion in der Komponente t

$$I(\pi^t | F^t(\cdot | \cdot)) = \tilde{E}^t \tilde{V}^t$$

und die Durchlaßkapazität in der Komponente t

$$C^t = \sup_{\pi^t} I(\pi^t | F^t(\cdot | \cdot)).$$

In [19] wird auf S. 75f. bewiesen, daß $I(\pi^t | F^t(\cdot | \cdot))$ als Funktion von π_i^t stetig ist. Da die Menge aller $\pi^t = (\pi_1^t, \dots, \pi_a^t)$ ein Kompaktum im $(a - 1)$ -dimensionalen euklidischen Raum bildet, gilt sogar

$$C^t = \max_{\pi^t} I(\pi^t | F^t(\cdot | \cdot)).$$

b) *Varianzen*

1. Varianz $\sigma_i^2(V_i^t)$ von V_i^t bezüglich $F^t(\cdot | i)$: für $i \in H^t$ gilt

$$\begin{aligned} \sigma_i^2(V_i^t) &= \int_{\Omega^t} (V_i^t(x^t) - E V_i^t)^2 F^t(dx^t | i) \\ &\leq \int_{\Omega^t} (V_i^t(x^t) + \log \pi_i^t)^2 F^t(dx^t | i) \end{aligned}$$

(wegen der bekannten Minimaleigenschaft des Erwartungswertes).

Nach Definition von V_i^t geht es weiter mit

$$\begin{aligned} &= \int (\log \{\pi_i^t m^t(x^t | i)\})^2 F^t(dx^t | i) \\ &= \int m^t(x^t | i) (\log \{\pi_i^t m^t(x^t | i)\})^2 \nu^t(dx^t). \end{aligned}$$

Damit wird

$$\begin{aligned} \sum_{i \in H^t} \pi_i^t \sigma_i^2(V_i^t) &\leq \int \sum_{i \in H^t} \pi_i^t m^t(x^t | i) \cdot (\log \{\pi_i^t m^t(x^t | i)\})^2 \nu^t(dx^t) \\ &\leq \sup_{b_i \geq 0, \sum b_i = 1} \sum_{i=1}^a b_i (\log b_i)^2. \end{aligned}$$

Nun gilt aber das

Lemma. *Es gibt eine von a unabhängige Konstante L^* mit*

$$\sum_{i=1}^a b_i (\log b_i)^2 \leq \max(L^*, \log^2 a) \quad \text{für alle } b_i \geq 0 \quad \text{mit} \quad \sum_{i=1}^a b_i = 1.$$

(Einen Beweis findet man in [19] auf S. 106f.)

Es gilt also

$$(4) \quad \sum_{i \in H^t} \pi_i^t \sigma_i^2(V_i^t) \leq \max(L^*, \log^2 a).$$

2. Varianz $\tilde{\sigma}^2(\tilde{V}^t)$ von \tilde{V}^t bezüglich $\tilde{F}^t(\cdot | \cdot)$

$$\begin{aligned} \tilde{\sigma}^2(\tilde{V}^t) &= \sum_{i \in H^t} \pi_i^t \int_{\Omega^t} \{V_i^t - \tilde{E}^t \tilde{V}^t\}^2 F^t(dx^t | i) \\ &= \sum_{i \in H^t} \pi_i^t \int_{\Omega^t} \{(V_i^t - E_i^t V_i^t) + (E_i^t V_i^t - \tilde{E}^t \tilde{V}^t)\}^2 \cdot F^t(dx^t | i) \\ &\leq \sum_{i \in H^t} \pi_i^t \int_{\Omega^t} \{V_i^t - E_i^t V_i^t\}^2 F^t(dx^t | i) \\ &\quad + \sum_{i \in H^t} \pi_i^t \int_{\Omega^t} \{E_i^t V_i^t - \tilde{E}^t \tilde{V}^t\}^2 F^t(dx^t | i). \end{aligned}$$

Der gemischte Term verschwindet, da $\tilde{E}_i^t \tilde{V}_i^t - E_i^t V_i^t$ nicht von x^t abhängt und $E_i^t(V_i^t - E_i^t V_i^t) = 0$ ist.

Unter Ausnutzung der Minimaleigenschaft des Mittelwertes folgt

$$\tilde{\sigma}^2(\tilde{V}^t) \leq \sum_{i \in H^t} \pi_i^t \sigma_i^2(V_i^t) + \sum_{i \in H^t} \pi_i^t (E_i^t V_i^t)^2.$$

Wegen a) 1. (2), (3) und b) 1. (4) gilt

$$\begin{aligned}\tilde{\sigma}^2(\tilde{V}t) &\leq \max(L^*, \log^2 a) + \sum_{i=1}^a \pi_i^t \left(-\frac{1}{e} - \log \pi_i^t \right)^2 \\ &\leq \left(\frac{1}{e} \right)^2 + 2 \max(L^*, \log^2 a) + \frac{2}{e} \log a \\ &\leq k \log^2 a \text{ mit einer absoluten Konstanten } k.\end{aligned}$$

Damit sind die Beweismittel bereitgestellt für das

Theorem 1 (Codingtheorem). Sei $0 < \lambda \leq 1$. Für jeden halb stetigen nichtstationären Kanal mit unabhängigen Zeichen existiert eine Konstante $K > 0$, so daß für jedes n gilt:

$$N(n, \lambda) > \exp \left\{ n \left(C(n) - \frac{K}{\sqrt{n}} \right) \right\} \quad \left(C(n) = \frac{1}{n} \sum_{t=1}^n C^t \right).$$

Beweis. Auf $\Omega_n \times \Omega'_n$ betrachte man die Informationsfunktion

$$\log \prod_{t=1}^n m^t(\cdot | x^t) \quad (x^t \in H^t).$$

Für den Erwartungswert und die Varianz (zu bilden durch Integration mit

$\tilde{F}_n = \prod_{t=1}^n * \tilde{F}^t$) dieser Funktion gilt wegen der Unabhängigkeit

$$\begin{aligned}\tilde{E} \log \prod_{t=1}^n m^t(\cdot | \cdot) &= \sum_{t=1}^n I(\pi^t | F^t) \\ \tilde{\sigma}^2 \left(\log \prod_{t=1}^n m^t(\cdot | \cdot) \right) &\leq n c \quad \text{mit } c = k \log^2 a,\end{aligned}$$

so daß die Tschebyscheffsche Ungleichung für jedes α die Relation

$$\tilde{F}_n \left\{ \log \prod_{t=1}^n m^t(\cdot | \cdot) > \sum_{t=1}^n I(\pi^t | F^t) - \alpha \right\} \geq 1 - \frac{nc}{\alpha^2}.$$

liefert.

Die Anwendung des Satzes aus § 2 auf $(\Omega_n \times \Omega'_n, F_n)$ liefert

$$N(n, \lambda) > \exp \left\{ \sum_{t=1}^n I(\pi^t | F^t) - \alpha \right\} \left(1 - \frac{nc}{\alpha^2} - (1 - \lambda) \right).$$

Nun wähle man $\alpha = \sqrt{\frac{2nc}{\lambda}}$. Dann ist

$$\begin{aligned}N(n, \lambda) &> \exp \left\{ \sum_{t=1}^n I(\pi^t | F^t) - \sqrt{\frac{2c}{\lambda}} \sqrt{n} \right\} \cdot \frac{\lambda}{2} \\ &> \exp \left\{ \sum_{t=1}^n I(\pi^t | F^t) - \sqrt{\frac{2c}{\lambda}} \sqrt{n} + \log \frac{\lambda}{2} \right\}.\end{aligned}$$

Da dies für alle π^t ($t = 1, \dots, n$) gilt, kann man $I(\pi^t | F^t)$ durch C^t ersetzen.

Damit erhält man

$$N(n, \lambda) > \exp \left\{ n \left(\frac{1}{n} \sum_{t=1}^n C^t - \frac{K}{\sqrt{n}} \right) \right\} = \exp \left\{ n \left(C(n) - \frac{K}{\sqrt{n}} \right) \right\}$$

(man wähle etwa $K = \sqrt{\frac{2c}{\lambda}} + \left| \log \frac{\lambda}{2} \right|$), wie behauptet.

§ 4. Beweis einer starken Umkehrung des Codingtheorems
für nichtstationäre halbstetige Kanäle

a) Approximation der Informationsfunktion des halbstetigen Kanals durch die eines Kanals mit endlichem Ausgang.

Sei $\mu^t = F^t(\cdot|1) + \dots + F^t(\cdot|a)$ und seien $f^t(\cdot|1), \dots, f^t(\cdot|a)$ die Radon-Nikodym-Dichten von $F^t(\cdot|1), \dots, F^t(\cdot|a)$ bezüglich μ^t ($t = 1, 2, \dots$). Durch Abänderung auf μ^t -Nullmengen seien die $f^t(\cdot|i)$ so bestimmt, daß $f^t(\cdot|1) + \dots + f^t(\cdot|a) = 1$ gilt für alle $t = 1, 2, \dots$.

Nun sei $g^t(\cdot)$ die Dichte von ν^t bezüglich μ^t . Wir führen nun unser Problem auf den Fall eines endlichen Ausgangsalphabets folgendermaßen zurück:

Man wähle reelle Zahlen d_0, \dots, d_r mit $0 = d_0 < d_1 < \dots < d_r = 1$ und setze

$$J(j) := [d_{j-1}, d_j] \quad j = 1, \dots, r-1$$

$$J(r) := [d_{r-1}, d_r].$$

Sei

$$L^t(k_1 \dots k_a) = \{b | f^t(b|1) \in J(k_1), \dots, f^t(b|a) \in J(k_a)\} \quad (k_1, \dots, k_a = 1, \dots, r).$$

Die $L^t(k_1, \dots, k_a)$ bilden eine von t abhängige Zerlegung des Ausgangsalphabetes Ω^t in r^a disjunkte evtl. zum Teil leere Mengen.

Die Übergangswahrscheinlichkeit von i nach $L^t(k_1, \dots, k_a)$ ist dann gegeben durch

$$h^t(k_1, \dots, k_a | i) = \int_{L^t(k_1 \dots k_a)} dF^t(\cdot|i) = \int_{L^t(k_1 \dots k_a)} f^t(b|i) d\mu^t(b)$$

$$(i = 1, \dots, a; k_1, \dots, k_a = 1, \dots, r).$$

Die Mengen $L^t(k_1, \dots, k_a)$ repräsentieren, soweit sie nicht leer sind, eindeutig die a -Tupel (k_1, \dots, k_a) , aus denen wir uns nun das Ausgangsalphabet eines Kanals mit unabhängigen Zeichen und den Übergangswahrscheinlichkeiten $h^t(k_1, \dots, k_a | i)$ gebildet denken.

Es kommt jetzt darauf an, durch geschickte Wahl von d_0, \dots, d_r die Informationsfunktion des halbstetigen Kanals durch die Informationsfunktion dieses diskreten Kanals gut zu approximieren.

Seien $\Omega = \{1, \dots, a\}$, (Ω', B') beliebig, $F(\cdot|i)$ ($i = 1, \dots, a$) gegeben. Wir bestimmen μ und die $f(\cdot|i)$ analog wie oben μ^t und die $f^t(\cdot|i)$. Wir betrachten die Informationsfunktion

$$I(\pi_1, \dots, \pi_a | F(\cdot|\cdot)) = \sum_{i=1}^a \pi_i \int_{\Omega'} f(b|i) \log \frac{\pi_i f(b|i)}{\sum_{j=1}^a \pi_j f(b|j)} \mu(db).$$

Hilfssatz 1. Für jedes $\eta > 0$ lassen sich die d_1, \dots, d_r so wählen, daß

$$(1) \quad A = \left| \sum_{i=1}^a \pi_i \int_{\Omega'} f(b|i) \log \frac{f(b|i)}{\sum_{j=1}^a \pi_j f(b|j)} \mu(db) - \sum_{i=1}^a \pi_i \sum_{k_1 \dots k_a} h(k_1, \dots, k_a|i) \log \frac{h(k_1, \dots, k_a|i)}{\sum_{j=1}^a \pi_j h(k_1 \dots k_a|j)} \right| \leq \eta$$

gleichmäßig in $F(\cdot|\cdot)$ und π gilt.

Insbesondere läßt sich also die Informationsfunktion des halbstetigen Kanals in der Komponente t

$$\sum_{i=1}^a \pi_i^t \int_{\Omega^t} f^t(b|i) \log \frac{f^t(b|i)}{\sum_{j=1}^a \pi_j^t f^t(b|j)} \mu^t(db)$$

gleichmäßig in t bis auf η durch die Informationsfunktion des durch das obige Diskretisierungsverfahren gewonnenen diskreten Kanals approximieren.

Beweis. Addiert man zunächst zu beiden Termen auf der rechten Seite von Gl. (1) $\sum_{i=1}^a \pi_i \log \pi_i$, so erhält man nach Umformulierung wegen

$$\begin{aligned} \sum_{i=1}^a \pi_i \log \pi_i &= \int_{\Omega'} \sum_{i=1}^a \pi_i \log \pi_i f(b|i) d\mu(b) \\ &= \sum_{k_1, \dots, k_a=1}^r \sum_{i=1}^a \pi_i \log \pi_i h(k_1, \dots, k_a|i), \end{aligned}$$

$$A = \left| \sum_{i=1}^a \pi_i \int_{\Omega'} f(b|i) \log \frac{\pi_i f(b|i)}{\sum_{j=1}^a \pi_j f(b|j)} \mu(db) - \sum_{i=1}^a \pi_i \sum_{k_1 \dots k_a=1}^r h(k_1, \dots, k_a|i) \log \frac{\pi_i h(k_1 \dots k_a)}{\sum_{j=1}^a \pi_j h(k_1 \dots k_a|j)} \right|.$$

Wir berechnen weitere Größen, nämlich

$$B = \left| \sum_{\substack{k_1 \dots k_a \\ \mu(L(k_1 \dots k_a)) = 0}} \int_{L(k_1 \dots k_a)} \sum_{i=1}^a \pi_i f(b|i) \log \frac{\pi_i f(b|i)}{\sum_{j=1}^a \pi_j f(b|j)} \mu(db) - \sum_{\substack{k_1 \dots k_a \\ \mu(L(k_1 \dots k_a)) = 0}} \sum_{i=1}^a \pi_i h(k_1 \dots k_a|i) \log \frac{\pi_i h(k_1 \dots k_a|i)}{\sum_{j=1}^a \pi_j h(k_1 \dots k_a|j)} \right| = 0$$

$$C = \sum_{\substack{k_1 \dots k_a \\ \mu(L(k_1 \dots k_a)) > 0}} \int_{L(k_1 \dots k_a)} \sum_{i=1}^a \pi_i f(b|i) \log \sum_{j=1}^a \pi_j f(b|j) (\mu db).$$

Sei $d = \sup_{i=0, \dots, r-1} |d_{i+1} - d_i|$

$$(*) \quad \left(\sum_{i=1}^a \pi_i f(b|i) \right) \log \left(\sum_{j=1}^a \pi_j f(b|j) \right) \text{ und } \left(\frac{\sum_{i=1}^a \pi_i h(k_1 \dots k_a|i)}{\mu(L(k_1 \dots k_a))} \right) \log \left(\frac{\sum_{j=1}^a \pi_j h(k_1 \dots k_a|j)}{\mu(L(k_1 \dots k_a))} \right)$$

unterscheiden sich auf $L(k_1, \dots, k_a)$ um weniger als $-d \log d$ für hinreichend kleines d , denn $f(x) = x \log x$ ($0 \leq x \leq 1$) ist im Punkte $x = 0$ am steilsten und korrespondierende Klammern unterscheiden sich höchstens um d . Durch Aufsummieren erhält man

$$C \leq -d \log d \mu(\Omega') = -a \cdot d \cdot \log d \leq \frac{\eta}{2}$$

für hinreichend kleines d .

$$D = \left| \sum_{\substack{k_1 \dots k_a \\ \mu(L(k_1 \dots k_a)) > 0}} \int_{L(k_1 \dots k_a)} \sum_{i=1}^a \pi_i f(b|i) \log \pi_i f(b|i) \mu(db) \right. \\ \left. - \sum_{\substack{k_1 \dots k_a \\ \mu(L(k_1 \dots k_a)) > 0}} \sum_{i=1}^a \pi_i h(k_1 \dots k_a|i) \log \frac{\pi_i h(k_1 \dots k_a|i)}{\mu(L(k_1 \dots k_a))} \right|$$

ist ebenfalls $\leq -d \log d$ für hinreichend kleines d . Wegen $A \leq B + C + D$ ist Hilfssatz 1 bewiesen.

Hilfssatz 1 läßt sich noch etwas verschärfen.

Hilfssatz 1'. Zu jedem $\eta > 0$ lassen sich die d_1, \dots, d_r aus Hilfssatz 1 so wählen, daß

$$(1) \quad A = \left| \pi_i \int_{\Omega'} f(b|i) \frac{\log f(b|i)}{\sum_{j=1}^a \pi_j f(b|j)} \mu(db) \right. \\ \left. - \pi_i \sum_{k_1 \dots k_a} h(k_1 \dots k_a|i) \log \frac{h(k_1 \dots k_a|i)}{\sum_{j=1}^a \pi_j h(k_1 \dots k_a|j)} \right| \leq \eta$$

gleichmäßig in $F(\cdot|\cdot)$, π_i gilt ($i = 1, \dots, a$).

Zum Beweis benötigen wir das

Lemma. Seien

$$a_1, b_1 > 0, \quad a_2, b_2 \geq 0, \quad b_1 + b_2 \leq 1, \quad a_1 + a_2 \leq 1.$$

Ferner gelte

$$|a_1 - b_1| < \varepsilon, \quad |(a_1 + a_2) - (b_1 + b_2)| < \varepsilon,$$

dann ist

$$G = |a_1 \log(a_1 + a_2) - b_1 \log(b_1 + b_2)| \leq \varepsilon + 3 |\varepsilon \log \varepsilon|.$$

Beweis. O. B. d. A. sei $(a_1 + a_2) \leq (b_1 + b_2)$, dann ist wegen der Steigungseigenschaften von $f(x) = \log x$ ($x > 0$)

$$|\log(a_1 + a_2) - \log(b_1 + b_2)| \leq \left| \log \frac{a_1 + a_2 + \varepsilon}{a_1 + a_2} \right| \leq \log \left(1 + \frac{\varepsilon}{a_1 + a_2} \right) \leq \frac{\varepsilon}{a_1 + a_2}$$

$$\begin{aligned} & |a_1 \log(a_1 + a_2) - b_1 \log(b_1 + b_2)| \\ &= |a_1 \log(a_1 + a_2) - a_1 \log(b_1 + b_2) + (a_1 - b_1) \log(b_1 + b_2)| \\ &\leq a_1 |\log(a_1 + a_2) - \log(b_1 + b_2)| + |(a_1 - b_1) \log(b_1 + b_2)| \\ &\leq \frac{a_1}{a_1 + a_2} \varepsilon + \varepsilon |\log(b_1 + b_2)| \end{aligned}$$

$$(*) \quad \leq \varepsilon + \varepsilon |\log b_1|.$$

Fall 1. Ist $b_1 < \varepsilon$, so ist $a_1 \leq 2\varepsilon$ und deshalb

$$G \leq 2|\varepsilon \log \varepsilon| + |\varepsilon \log \varepsilon| = +3|\varepsilon \log \varepsilon|.$$

Fall 2. Ist $b_1 \geq \varepsilon$, dann ist

$$G \leq \varepsilon + |\varepsilon \log \varepsilon| \leq \varepsilon + 3|\varepsilon \log \varepsilon|.$$

Der Beweis von Hilfssatz I' verläuft jetzt wie der Beweis von Hilfssatz 1. Man hat lediglich statt (*)

$$(**) \quad \pi_i f(b|i) \log \left(\sum_{j=1}^a \pi_j h(k_1, \dots, k_a|j) \right)$$

und

$$\frac{\pi_i h(k_1, \dots, k_a|i)}{\mu(L(k_1, \dots, k_a))} \log \frac{\sum_{j=1}^a \pi_j h(k_1, \dots, k_a|j)}{\mu(L(k_1, \dots, k_a))}$$

gegeneinander abzuschätzen. Das wird aber durch das Lemma ermöglicht.

b) *Zerlegung der Schar der $F^t(\cdot|\cdot)$ ($t = 1, 2, \dots$) in endlich viele Klassen, so, daß die Informationsfunktionen der Elemente einer Klasse in der Supremumnorm (L^∞) benachbart sind.*

Sei

$$[0, \varepsilon), \quad [\varepsilon, 2\varepsilon), \dots, \quad [(L-1)\varepsilon, L\varepsilon = 1]$$

mit geeignetem ε und $L = L(\varepsilon)$ gegeben. (Die Äquidistanz der Intervalle ist nicht notwendig, aber man kann mit ihr das Gewünschte erreichen.)

Man definiere

$$M(l) = [\varepsilon(l-1), \varepsilon l) \quad (l = 1, \dots, L-1) \\ M(L) = [(L-1)\varepsilon, 1].$$

Die $M(l)$ ($l = 1, \dots, L$) bilden eine disjunkte Zerlegung des Einheitsintervalles. Seien die d_i wie oben gewählt. Die Zahlen $t = 1, 2, \dots$ seien in folgender Weise in disjunkte Klassen zerlegt: Zwei Zahlen t und s gehören derselben Klasse an, wenn für alle Indexkonstellationen (k_1, \dots, k_a) ($k_i = 1, \dots, r$), ($i = 1, \dots, a$) jeweils gleich indizierte Elemente $h^t(k_1, \dots, k_a|i)$, $h^s(k_1, \dots, k_a|i)$ im gleichen Teilintervall $M(l)$ liegen. Man erhält auf diese Weise $L(\varepsilon)^{a \cdot r^a} = R(\varepsilon, r) = R$ Klassen Γ_ϱ ($\varrho = 1, \dots, R(\varepsilon, r)$), von denen einige leer sein können.

Sei

$$\Gamma_\varrho^n = \Gamma_\varrho \cap \{1, \dots, n\}, \\ Z_\varrho^n = \{h^t | t \in \Gamma_\varrho^n\}.$$

Für ein beliebiges $u_0 = (x^1, \dots, x^n) \in \Omega_n$ betrachten wir nun die Anzahl $N_\varrho^e(i|u_0)$ der i lautenden Komponenten von u_0 im Ausschnitt Γ_ϱ^n , sei ferner $n_\varrho = |\Gamma_\varrho^n|$.

Man definiere

$$\pi_i^{(\varrho)} = \frac{N_\varrho^e(i|u_0)}{n_\varrho} \{i = 1, \dots, n\}, \quad (\varrho = 1, \dots, R)$$

und $\pi^t = \pi^{(\varrho)}$ für $t \in \Gamma_\varrho^n$ als Wahrscheinlichkeitsverteilung auf Ω^t für $t \in \Gamma_\varrho^n$. Mit diesem speziellen durch Auszählung erhaltenen π^t bilden wir nun $m^t(\cdot|\cdot)$ und betrachten

$$V_i^t(\cdot) = \log m^t(\cdot|i) \quad (i \in H^t = \{j | \pi^t > 0\})$$

V_i^t ist mit Ausnahme der $F^t(\cdot|i)$ -Nullmenge $\{x^t | m^t(x^t|i) = 0\}$ definiert.

Seien $V_i^{(\varrho)} \in \{V_i^t \mid t \in \Gamma_\varrho^n\}$ beliebig gewählte Repräsentanten; dann gilt der

Hilfssatz 2. *Man wähle $\pi_j^t = \pi_j^{(\varrho)}$ für $t \in \Gamma_\varrho^n$ und bilde den Erwartungswert von*

$$\sum_{t=1}^n \log m^t(x'^t \mid x^t) \quad \text{bezüglich} \quad F_n(\cdot \mid u_0) = \prod_{t=1}^n F^t(\cdot \mid x^t)$$

($m^t(\cdot \mid x^t)$ ist definiert, da x^t nach Definition von $\pi_i^{(\varrho)}$ und π_j^t in H^t liegt.)

Dann gilt

$$E \sum_{t=1}^n \log m^t(x'^t \mid x^t) = \sum_{\varrho=1}^R n_\varrho \sum_{i \in H^t} \pi_i^{(\varrho)} E V_i^{(\varrho)} + \sum_{\varrho=1}^R f(\varepsilon, \eta, n_\varrho),$$

wobei

$$|f(\varepsilon, \eta, n_\varrho)| \leq a(2r^a \varepsilon \log \varepsilon + \eta) n_\varrho$$

ist.

Beweis. Es ist

$$E_i^t V_i^t = \int_{\Omega^t} (\log m^t(\cdot \mid i)) m^t(\cdot \mid i) d\nu^t \quad (i \in H^t).$$

Nach Hilfssatz 1' ist $\pi_i^{(\varrho)} E_i^t V_i^t$ durch

$$(2) \quad \pi_i^{(\varrho)} \sum_{k_1, \dots, k_a} h^t(k_1, \dots, k_a \mid i) \log \frac{h^t(k_1, \dots, k_a \mid i)}{\sum_{j \in H^t} \pi_j^{(\varrho)} h^t(k_1, \dots, k_a \mid j)}$$

bis auf η approximierbar für jedes ϱ , $t \in \Gamma_\varrho^n$.

(3) Für h^t, h^s ($t, s \in \Gamma_\varrho^n$) gilt aber

$$\begin{aligned} & \pi_i^{(\varrho)} \sum_{k_1, \dots, k_a} h^t(k_1, \dots, k_a \mid i) \log \frac{h^t(k_1, \dots, k_a \mid i)}{\sum_{j \in H^t} \pi_j^{(\varrho)} h^t(k_1, \dots, k_a \mid j)} \\ & - \pi_i^{(\varrho)} \sum_{k_1, \dots, k_a} h^s(k_1, \dots, k_a \mid i) \log \frac{h^s(k_1, \dots, k_a \mid i)}{\sum_{j \in H^t} \pi_j^{(\varrho)} h^s(k_1, \dots, k_a \mid j)} \\ & = \sum_{k_1, \dots, k_a} \{ \pi_i^{(\varrho)} \{ h^t(k_1, \dots, k_a \mid i) \log h^t(k_1, \dots, k_a \mid i) \\ & \quad - h^s(k_1, \dots, k_a \mid i) \log h^s(k_1, \dots, k_a \mid i) \} \} \\ & + \left\{ \pi_i^{(\varrho)} h^s(k_1, \dots, k_a \mid i) \log \left(\sum_{j \in H^t} \pi_j^{(\varrho)} h^s(k_1, \dots, k_a \mid j) \right) \right. \\ & \quad \left. - (\pi_i^{(\varrho)} h^t(k_1, \dots, k_a \mid i)) \log \left(\sum_{j \in H^t} \pi_j^{(\varrho)} h^t(k_1, \dots, k_a \mid j) \right) \right\} \\ & \leq 2r^a(\varepsilon + 3|\varepsilon \log \varepsilon|) = r^a g(\varepsilon), \end{aligned}$$

wenn wir $g(\varepsilon) = 2(\varepsilon + 3|\varepsilon \log \varepsilon|)$ setzen.

Die letzte Abschätzung erhält man unter Benutzung des Lemmas.

Hintereinanderausführung von (2) und (3) liefert

$$|\pi_i^{(\varrho)} E_i^t V_i^t - \pi_i^{(\varrho)} E_i^s V_i^s| \leq r^a g(\varepsilon) + 2\eta(t, s \in \Gamma_\varrho^n).$$

Daraus folgt

$$\begin{aligned} & \left| E \sum_{t=1}^n \log m^t(x'^t \mid x^t) - \sum_{\varrho=1}^R n_\varrho \sum_{i \in H^t} \pi_i^{(\varrho)} E V_i^{(\varrho)} \right| \\ & \leq \sum_{\varrho=1}^R n_\varrho \sum_{i \in H^t} \pi_i^{(\varrho)} \max_{t, s \in \Gamma_\varrho^n} (E_i^t V_i^t - E_i^s V_i^s) \\ & \leq \sum_{\varrho=1}^R n_\varrho a(r^a g(\varepsilon) + 2\eta). \end{aligned}$$

Damit ist Hilfssatz 2 bewiesen.

$\sum_{i \in H^t} \pi_i^t E_i^t V_i^t = I(\pi^t | F^t(\cdot | i))$ ist Informationsfunktion für die Komponente t .

Hilfssatz 2 besagt, daß $E \sum_{t=1}^n \log m^t(x^t | x^t)$ um weniger als $na(r^a g(\varepsilon) + 2\eta)$ von der Informationsfunktion bezüglich $\pi_i^{(e)}$ im Zeitraum $\langle 1, n \rangle$ abweicht.

KEMPERMAN [19] erhält für die Varianz die Abschätzung

$$\sum_{i \in H^t} \pi_i^t \sigma_i^2(V_i^t) \leq \max(\log^2 3, \log^2 a)$$

(vgl. Gleichung (4) in § 3).

Wegen Hilfssatz 2 ist

$$E \sum_{t=1}^n \log m^t(x^t | x^t) \leq nC(n) + na(r^a g(\varepsilon) + 2\eta).$$

Nach der Tschebyscheffschen Ungleichung gilt nun

$$(4) \quad F_n \left\{ \sum_{t=1}^n \log m^t(x^t | x^t) \geq nC(n) + na(r^a g(\varepsilon) + 2\eta) + \sqrt{\frac{2n}{1-\lambda} (\max(\log^2 3, \log^2 a) + a(r^a g(\varepsilon) + 2\eta))} \right\} \leq \frac{1-\lambda}{2},$$

wenn man als grobe Abschätzung für die Varianz von

$$\sum_{e=1}^R f(\varepsilon, \eta, n_e) \quad na(r^a g(\varepsilon) + 2\eta)$$

wählt. Wir schreiben abkürzend $nC(n) + \dots$ statt

$$nC(n) + na(r^a g(\varepsilon) + 2\eta).$$

Sei

$$A' \in \Omega'_n \quad \text{mit} \quad F_n(A' | u_0) \geq 1 - \lambda.$$

Sei

$$B' = \left\{ u' = (x'^1, \dots, x'^n) \mid 0 < \prod_{t=1}^n m^t(x^t | x^t) \leq e^{nC(n) + \dots} \right\},$$

und sei

$$\bar{B}' = \left\{ u' = (x'^1, \dots, x'^n) \mid \prod_{t=1}^n m^t(x^t | x^t) \leq e^{nC(n) + \dots} \right\}.$$

$\bar{B}' - B'$ ist dann eine $F_n(\cdot | u_0)$ -Nullmenge. Es gilt also

$$(5) \quad F_n(\bar{B}' | u_0) = F_n(B' | u_0).$$

Auf $(A' \cap B')$ gilt nach Definition von B'

$$e^{-\{nC(n) + \dots\}} \leq \left(\prod_{t=1}^n m^t(x^t | x^t) \right)^{-1}.$$

Für A' gilt deshalb:

$$\begin{aligned} \int_{A'} \nu^1(dx'^1) \cdots \nu^n(dx'^n) &\geq \int_{A' \cap B'} \nu^1(dx'^1) \cdots \nu^n(dx'^n) \\ &= \int_{A' \cap B'} \frac{F_n(dx'^1, \dots, dx'^n | u_0 = (x^1, \dots, x^n))}{\prod_{t=1}^n m^t(x^t | x^t)} \geq F_n(A' \cap B' | u_0) \cdot e^{-\{nC(n) + \dots\}}. \end{aligned}$$

Dabei ist

$$\begin{aligned} F_n(A' \cap B' | u_0) &= F_n(A' \cap \bar{B}' | u_0) \quad (\text{nach Gleichung (5)}) \\ &\geq F_n(A' | u_0) - F_n((\bar{B}')^c | u_0) \geq 1 - \lambda - \left(\frac{1-\lambda}{2}\right) \\ & \quad (\text{nach Definition von } A', \bar{B}' \text{ und Gleichung (4)}) \\ &= \frac{1-\lambda}{2}. \end{aligned}$$

Wir erhalten also

$$(6) \quad \int_{A'} \nu_1(dx^1) \cdots \nu_n(dx^n) \geq \frac{1-\lambda}{2} e^{-\{nC(n)+\dots\}}.$$

Sei nun $(u_1^0, A_1^0), \dots, (u_M^0, A_M^0)$ ein Code (n, M, λ) , so daß

$$N^e(i | u_0) = N^e(i | u_j^0) \quad \{(i = 1, \dots, a), (j = 1, \dots, M)\}$$

gilt.

Die u_j^0 ($j = 1, \dots, M$) sind gerade so definiert worden, daß die gleichen Berechnungen, wie sie jetzt für u_0 durchgeführt wurden, auch für u_j^0 ($j = 1, \dots, M$) gemacht werden können.

Man erhält mit den gleichen $\nu_1, \nu_2, \dots, \nu_n$:

$$\int_{A_j'} \nu_1(dx^1) \cdots \nu_n(dx^n) \geq \frac{1-\lambda}{2} e^{-\{nC(n)+\dots\}} \quad (j = 1 \dots M).$$

Durch Aufsummieren ergibt sich:

$$1 \geq \sum_{j=1}^M \int_{A_j'} \nu_1(dx^1) \cdots \nu_n(dx^n) \geq M \frac{1-\lambda}{2} e^{-\{nC(n)+\dots\}}.$$

Dann ist $M \leq \frac{2}{1-\lambda} \exp\{nC(n) + \dots\}$.

Die Gesamtheit der $\pi^{(e)}$ mit der Eigenschaft $\pi_i^{(e)}$ ist ganzzahliges Vielfaches von $\frac{1}{n^e}$, ist kleiner als $(n+1)^a$. Die Gesamtheit der $(\pi^{(1)}, \dots, \pi^{(R)})$ ist also $\leq (n+1)^{aR^{(e)}}$. Daraus folgt

$$(7) \quad \begin{aligned} N &< (n+1)^{aR^{(e)}} \frac{2}{1-\lambda} \exp\{nC(n) + na(R^a g(\varepsilon) + 2\eta) \\ &+ \sqrt{\frac{2n}{1-\lambda} (\max(\log^2 3, \log^2 a) + a(r^a g(\varepsilon) + 2\eta))}\}. \end{aligned}$$

Theorem 2 (Starke Umkehrung des Codingtheorems). Für den nichtstationären halb stetigen Kanal mit unabhängigen Zeichen $F^t(\cdot | i)$ ($i = 1, \dots, a$; $t = 1, 2, \dots$) gilt die Abschätzung: zu λ ($0 < \lambda \leq 1$) und $\delta > 0$ existiert ein $n_0(\lambda, \delta)$, so daß für $n \geq n_0$

$$N(n, \lambda) < e^{n(C(n) + \delta)}$$

gilt.

Beweis. Man wähle $\eta < \frac{\delta}{12a}$ und $d_1 \dots d_r$ so, daß die Hilfssätze mit diesem η gelten, dann ε so, daß $ar^a g(\varepsilon) \leq \frac{\delta}{6}$ ist, schließlich n_0 so, daß $a \cdot R(\varepsilon) \log(n+1)$

$+\log \frac{2}{1-\lambda} < \frac{\delta}{3}$ und $\sqrt{\frac{2n}{1-\lambda} (\max(\log^2 3, \log^2 a) + a(r^a g(\varepsilon) + 2\eta))} \leq \frac{\delta}{3} n$ ist für $n \geq n_0$. Dann liefert (7) die Behauptung.

Kapitel II. Translationsinvariante Codes fastperiodischer Kanäle

Die in Kapitel I dargestellte Theorie hat gegenüber der Theorie stationärer Kanäle den Nachteil, daß die Güte eines Codes nicht invariant gegenüber Zeittranslationen ist. Die hiermit nahegelegte Invarianzforderung führt auf ein nichtstationäres Simultancodeproblem. Dieses Problem konnte ich bisher nicht in voller Allgemeinheit lösen (vgl. Kap. III, IV). In diesem Kapitel soll jedoch ein für die Praxis wichtiger Spezialfall vollständig erledigt werden: der Kanal möge fastperiodisch von der Zeit abhängen. Damit wird gerade die Situation, in der man sich befindet, wenn man Nachrichten, die von einem Satelliten nach festen Erdstationen gesendet werden, codiert, ohne die Phase des Satelliten zu kennen, modellmäßig erfaßt.

§ 1. Allgemeines über Simultankanäle

Seien Ω^t, Ω'^t ($t = 1, 2, \dots$) Exemplare einer endlichen Menge $\{1, 2, \dots, a\}$ mit $a \geq 2$ Elementen, die wir als Eingangs- bzw. Ausgangsalphabete für diskrete Kanäle benutzen werden. Wir betrachten hier nur den Fall, wo beide Alphabete gleich sind. Der allgemeine Fall geht genauso bzw. läßt sich auf diesen Spezialfall zurückführen.

Sei $I = \{s, \dots\}$ eine beliebige nichtleere (Index-)Menge und jedem $s \in I$ eine Folge $(\omega^t(\cdot | \cdot | s))_{t=1,2,\dots}$ von Matrizen zugeordnet, die die Eigenschaft hat, daß ihre Elemente $\omega^t(\cdot | \cdot | s)$ stochastisch von Ω^t auf Ω'^t sind für $t \in N$. Somit ist jedem $s \in I$ ein diskreter Kanal zugeordnet. Wenn wir Probleme, die das Simultanverhalten aller dieser Kanäle betreffen, behandeln, bezeichnen wir diese durch $s \in I$ indizierte Familie von Kanälen auch als einen *Simultankanal* (diskret mit unabhängigen Zeichen). Gemeinsame Eigenschaften der einzelnen beteiligten Kanäle werden auch dem Simultankanal beigelegt: Stationarität, Fastperiodizität etc.

Durch die Festsetzung

$$P_n(y_n | x_n | s) = \prod_{t=1}^n \omega^t(y^t | x^t | s)$$

$$s \in N$$

$$x_n = (x^1, \dots, x^n) \in \Omega_n = \prod_{t=1}^n \Omega^t$$

$$y_n = (y^1, \dots, y^n) \in \Omega'_n = \prod_{t=1}^n \Omega'^t$$

ist eine Menge von Kanälen $S_n = \{P_n(\cdot | \cdot | s) | s \in I\}$ im Zeitabschnitt $\langle 1, n \rangle$ definiert. Ein Code (n, N, λ) für den Simultankanal S_n ist eine Serie

$$\{(u_1, A_1), \dots, (u_N, A_N)\} \quad \text{mit} \quad u_i \in \Omega_n, A_i \subset \Omega'_n \quad \text{für} \quad i = 1, 2, \dots, N,$$

$A_i \cap A_j = \emptyset$ für $i \neq j$ und mit

$$P_n(A_i | u_i | s) = \sum_{v_i \in A_i} P_n(v_i | u_i | s) \geq 1 - \lambda \quad \text{für} \quad i = 1, 2, \dots, N, s \in I.$$

Die Theorie der stationären Simultankanäle ist durch [3], [19] weit entwickelt worden. Wir benötigen hier vor allem ein Ergebnis von BLACKWELL, BREIMANN, THOMASIAN über stationäre Simultankanäle mit endlicher Indexmenge I . Es läßt sich ohne weiteres (nach dem Schema von Kap. I) auf nichtstationäre Kanäle mit endlicher Indexmenge übertragen und lautet dann

Satz 0 (*Maximalcodesatz für nichtstationäre Simultankanäle mit endlicher Indexmenge*).

Zu jedem reellen $\delta > 0$ und zu jedem $n \in \mathbb{N}$ gibt es für den Simultankanal

$$S_n = \{P_n(\cdot | \cdot | s) | s \in I = \{1, 2, \dots, K\}\}$$

einen Code $(n, e^{(C(n)-\delta)^n}, \lambda_n)$, wobei

$$C(n) = \frac{1}{n} \max_{(\pi^1, \dots, \pi^n)} \inf_{s \in I} \left(\sum_{t=1}^n I^t(\pi^t | \omega^t(\cdot | \cdot | s)) \right)$$

(in Analogie zum stationären Fall definiert) die Simultankapazität im Zeitraum $\langle 1, n \rangle$ und

$$\lambda_n = 2 K^2 \cdot e^{\frac{(C(n)-\delta)^2 n}{16a^2}}$$

ist.

§ 2. Herleitung des Codingtheorems für einen fastperiodischen diskreten Kanal mit unabhängigen Zeichen bei zeittranslationsinvarianten Codes mit Hilfe des Maximalcodesatzes für nichtstationäre Simultankanäle

Sei $(\omega^t(\cdot | \cdot))_{t=1,2,\dots}$ eine fastperiodische Folge stochastischer $a \times a$ -Matrizen. Wir definieren allgemein $\omega^t(i | j | s) = \omega^{t+s}(i | j)$ für $1 \leq i, j \leq a, t \in \mathbb{N}, s \in \mathbb{N} \cup \{0\}$. Hierdurch erhalten wir einen nichtstationären Simultankanal

$$S = \{(\omega^t(\cdot | \cdot | s))_{t=1,2,\dots} | s \in \mathbb{N}\}$$

(im Sinne von § 1).

Für stationäre Simultankanäle mit beliebiger Indexmenge wurde der Beweis des Codingtheorems in [18], [3], erbracht. Der Beweis in [3] stützt sich auf den Maximalcodesatz für stationäre Simultankanäle mit endlicher Indexmenge. Bei der Übertragung dieses Verfahrens auf unsere Situation zeigt sich, daß der Beweis nur unter einer Zusatzvoraussetzung an den Kanal funktioniert.

Um den Maximalcodesatz anwenden zu können, muß man die unendlich vielen Elemente enthaltende Klasse von Kanälen $S_n = \{P_n(\cdot | \cdot | s) | s \in \mathbb{N}\}$ durch eine Klasse S_n^* mit endlich vielen Elementen approximieren. Dabei ist auf zweierlei zu achten:

1. Die Simultankapazitäten von S_n und S_n^* dürfen nur wenig voneinander abweichen.

2. Die Güte eines Codes für S_n^* darf sich beim Übergang zu S_n nicht nennenswert verringern.

Nach von NEUMANN ist eine beschränkte Funktion $f(x)$ auf einer Gruppe G fastperiodisch, wenn die Funktionenmenge $\{f(xy) | y \in G\}$ in der Topologie der gleichmäßigen Konvergenz relativ kompakt ist.

Die Definition läßt sich auch auf Halbgruppen übertragen. Da $\omega^t(\cdot|\cdot)$ fast-periodisch ist, existieren also zu jedem $\varepsilon > 0$ endlich viele Matrizenfolgen

$$(\omega^t(\cdot|\cdot|s^*))_{t=1,2,\dots} = (\omega^{t+s^*}(\cdot|\cdot))_{t=1,2,\dots}$$

($s^* = 0, 1, \dots, L(\varepsilon)$) so, daß zu jeder Matrizenfolge $(\omega^{t+s}(\cdot|\cdot))_{t=1,2,\dots}$ ein s^* mit $\|(\omega^t(\cdot|\cdot|s^*))_{t=1,2,\dots} - (\omega^{t+s}(\cdot|\cdot))_{t=1,2,\dots}\|_\infty = \sup_{\substack{1 \leq i, j \leq a \\ t \in N}} |\omega^t(i|j|s^*) - \omega^{t+s}(i|j)| < \varepsilon$

gefunden werden kann.

Wir können nun folgende Menge $S_n^*(\varepsilon)$ von endlich vielen Kanälen definieren:

$$\begin{aligned} S_n^*(\varepsilon) &= \{P_n(\cdot|\cdot|s^*) | P_n(y_n|x_n|s) \\ &= \prod_{t=1}^n \omega^t(y^t|x^t|s^*), s^* = 0, 1, \dots, L(\varepsilon)\}. \end{aligned}$$

Man schreibt ihr nach dem obigen Ansatz die Kapazität

$$C_n^*(\varepsilon) = \frac{1}{n} \max_{(\pi^1, \dots, \pi^n)} \inf_{s^*=0, 1, \dots, L(\varepsilon)} \left(\sum_{t=1}^n I^t(\pi^t | \omega^t(\cdot|\cdot|s^*)) \right).$$

zu. Da die Informationsfunktion $I(\pi|\omega)$ normstetig von ω abhängt, gilt für alle $n \in N$ $|C_n - C_n^*(\varepsilon)| \leq \eta$ mit $\eta \rightarrow 0$, falls $\varepsilon \rightarrow 0$, d.h. Bedingung 1. ist erfüllt.

Bedingung 2. wird aus dem nachstehenden Lemma folgen.

Lemma. Sei $b > 0$ reell. Es gibt eine Nullfolge reeller Zahlen $a_n > 0$ mit folgender Eigenschaft: Sind $n \in N$, $A \subset \Omega_n^*$, $x_n \in \Omega_n$, $s, s^* \in N$ derart, daß $P_n(A|x_n|s) > b$

$$|\omega^t(i|j|s^*) - \omega^t(i|j|s)| \leq \frac{a}{2\sqrt{n}} \quad (1 \leq i, j \leq a; t \in N)$$

gilt, so ist

$$\left| \frac{P_n(A|x_n|s^*)}{P_n(A|x_n|s)} - 1 \right| < a_n.$$

Beweis. Wir unterscheiden die Fälle:

$$\text{a) } \min_{\substack{1 \leq i, j \leq a \\ t=1, 2, \dots, n}} \omega^t(j|i|s) \geq n^{-2}.$$

Aus

$$\prod_{t=1}^n \frac{(\omega^t(y^t|x^t|s) - a \cdot 2^{-\sqrt{n}})}{\omega^t(y^t|x^t|s)} \leq \frac{P_n(y_n|x_n|s^*)}{P_n(y_n|x_n|s)} \leq \prod_{t=1}^n \frac{(\omega^t(y^t|x^t|s) + a \cdot 2^{-\sqrt{n}})}{\omega^t(y^t|x^t|s)}$$

folgt

$$\left(1 - \frac{a \cdot 2^{-\sqrt{n}}}{n^{-2}}\right)^n \leq \frac{P_n(y_n|x_n|s^*)}{P_n(y_n|x_n|s)} \leq \left(1 + \frac{a \cdot 2^{-\sqrt{n}}}{n^{-2}}\right)^n$$

für

$$\begin{aligned} n \geq n_0 &= \inf \left\{ k \mid \frac{a \cdot 2^{-\sqrt{k}}}{k^{-2}} < 1, k \in N \right\} \\ \exp \left\{ n \log \left(1 - \frac{a \cdot n^2}{2\sqrt{n}} \right) \right\} &\leq \frac{P_n(y_n|x_n|s^*)}{P_n(y_n|x_n|s)} \leq \exp \left\{ n \log \left(1 + \frac{a \cdot n^2}{2\sqrt{n}} \right) \right\} \quad (n \geq n_0). \end{aligned}$$

Nun ist

$$\begin{aligned} \left| n \log \left(1 \pm \frac{a \cdot n^2}{2\sqrt{n}} \right) \right| &= \frac{n^3 \cdot a}{2\sqrt{n}} \left| \sum_{v=1}^{\infty} \frac{(-1)^v}{v} \left(\pm \frac{a \cdot n^2}{2\sqrt{n}} \right)^{v-1} \right| \quad (n \geq n_0) \\ &\leq \frac{a \cdot n^3}{2\sqrt{n}} C \quad (n \geq n_0, C \text{ Konstante} > 0) \\ \lim_{n \rightarrow \infty} \frac{P_n(y_n | x_n | s^*)}{P_n(y_n | x_n | s)} &= 1. \end{aligned}$$

Es gibt deshalb eine Nullfolge $(a_n)_{n=1,2,\dots}$, so daß

$$\left| \frac{P_n(y_n | x_n | s^*)}{P_n(y_n | x_n | s)} - 1 \right| < a_n \quad (n \in N).$$

Wegen

$$|P_n(y_n | x_n | s^*) - P_n(y_n | x_n | s)| < a_n P_n(y_n | x_n | s)$$

ist

$$\begin{aligned} |P_n(A | x_n | s^*) - P_n(A | x_n | s)| \\ \leq \sum_{y_n \in A} |P_n(y_n | x_n | s^*) - P_n(y_n | x_n | s)| < a_n P_n(A | x_n | s) \end{aligned}$$

und deshalb

$$\left| \frac{P_n(A | x_n | s^*)}{P_n(A | x_n | s)} - 1 \right| < a_n,$$

wie behauptet.

$$\text{b) } \min_{\substack{1 \leq i, j \leq a \\ t = 1, 2, \dots, n}} \omega^t(j | i | s) < n^{-2}.$$

Die Menge

$$M = \{(i, j, t) | 1 \leq i, j \leq a, 1 \leq t \leq n, \omega^t(j | i | s) < n^{-2}\}$$

hat höchstens $a^2 n$ Elemente.

$$M_{x_n} = \{(i, j, t) | (i, j, t) \in M \text{ und } (i, j, t) = (x^t, j, t)\}$$

hat höchstens $a \cdot n$ Elemente. Jedem Element $(i, j, t) \in M_{x_n}$ ordnen wir eine Menge

$$B(i, j, t) = \{y_n = (y^1, \dots, y^n) | y^t = j\}$$

zu. Dann gilt nach Definition von $B(i, j, t)$:

$$P_n(B = \bigcup_{(i,j,t) \in M_{x_n}} B(i, j, t) | x_n | s^*) < a \cdot n \cdot n^{-2} = \frac{a}{n}$$

und

$$P_n(B = \bigcup_{(i,j,t) \in M_{x_n}} B(i, j, t) | x_n | s) < a n (n^{-2} + a \cdot 2^{-\sqrt{n}}) \leq \frac{2a}{n} \quad \text{für } n \geq n_0.$$

Das heißt aber, daß die Elemente aus B an A einen asymptotisch verschwindenden Anteil haben. Deshalb können wir schreiben

$$\frac{P_n(A | x_n | s^*)}{P_n(A | x_n | s)} = \frac{P_n(A - B | x_n | s^*) + b_n}{P_n(A - B | x_n | s) + c_n},$$

wobei $(b_n)_{n=1,2,\dots}$; $(c_n)_{n=1,2,\dots}$ Nullfolgen sind. Da $P_n(A | x_n | s) > b$ voraus-

gesetzt wurde, gibt es ein n_1 , derart, daß

$$P_n(A - B | x_n | s) > \frac{b}{2} \quad \text{für } n \geq n_1$$

gilt. Nach Fall a) ist

$$\lim_{n \rightarrow \infty} \frac{P_n(A - B | x_n | s^*)}{P_n(A - B | x_n | s)} = 1$$

und somit auch

$$\lim_{n \rightarrow \infty} \frac{P_n(A | x_n | s^*)}{P_n(A | x_n | s)} = 1.$$

Damit ist das Lemma bewiesen und deshalb Bedingung 2. erfüllt.

Damit sind die Beweismittel bereitgestellt für den folgenden

Satz 1 (Codingtheorem). *Sei S ein fastperiodischer Simultankanal, der der folgenden Bedingung genügt:*

(1) *zu jedem $\eta > 0$ gibt es ein $n_0(\eta)$, so daß für $n \geq n_0(\eta)$ $L(a \cdot 2^{-\sqrt{n}}) < e^{-\eta n}$ erfüllt ist.*

(Hinreichend dafür ist etwa $L(a \cdot 2^{-\sqrt{n}}) \sim e^{-\alpha n}$ ($0 < \alpha < 1$)). Dann gibt es zu λ ($0 < \lambda < 1$), $\delta > 0$ ein $n(\lambda, \delta)$, so daß für $n \geq n(\lambda, \delta)$ $N(n, \lambda) > e^{(C(n) - \delta)n}$ gilt.

Beweis. Man wende den Maximalcodesatz auf $S_n^*(a \cdot 2^{-\sqrt{n}})$ an. Aus der Forderung: $\lambda \geq \lambda_n$ für hinreichend große n , ergibt sich

$$L(a \cdot 2^{-\sqrt{n}}) = K \leq \exp \left\{ \frac{\delta^2}{32a^2} n + \log \frac{\lambda}{2} \right\}$$

(n hinreichend groß). Die Ungleichung ist wegen Bedingung (1) erfüllt. Berücksichtigt man 1., 2. so erhält man die Aussage des Satzes.

Die Klasse der (1) erfüllenden nichtperiodischen fastperiodischen Kanäle ist nicht leer, wie das folgende Beispiel zeigt.

Wir konstruieren zunächst eine nichtperiodische fastperiodische Funktion $f(n)$, $n \in N$, die Bedingung (1) erfüllt. Für $i \in N$ definieren wir

$$p_i(n) = \begin{cases} 1 & n \in \{i k \mid k \in N\} \\ 0 & \text{sonst.} \end{cases}$$

Falls $h(i) \geq 0$ für $i \in N$, $h(i) \neq 0$ für unendlich viele i und $\sum_{i=1}^{\infty} h(i) < \infty$, so ist $\sum_{i=1}^{\infty} h(i) p_i(n)$ fastperiodisch und nicht periodisch.

Bedingung (1) ist erfüllt, falls $L(a^{-\sqrt{n}}) < a^{\sqrt{n}}$ ($n \in N$). Dafür ist hinreichend $L(\varepsilon) \leq \frac{1}{\varepsilon}$ ($0 < \varepsilon < 1$). Man wähle

$$h(i) = \frac{1}{(i+1)^{i+1}}$$

und zu $\varepsilon > 0$ i_0 minimal mit der Eigenschaft:

$$\frac{1}{(i_0+1)^{i_0+1}} < \varepsilon.$$

Wegen

$$\sum_{i=i_0+1}^{\infty} \frac{1}{(i+1)^{i+1}} \leq \frac{1}{(i_0+1)^{i_0+1}}$$

gilt

$$\left| \sum_{i=i_0+1}^{\infty} \frac{1}{(i+1)^{i+1}} p_i(n) \right| < \varepsilon.$$

$$\sum_{i=1}^{i_0} \frac{1}{(i+1)^{i+1}} p_i(n).$$

hat eine Periodenlänge $\leq i_0!$

$$\left\{ \sum_{i=1}^{i_0} \frac{1}{(i+1)^{i+1}} p_i(n+l) \mid l \in N \right\}$$

liegt ε -dicht in

$$\left\{ \sum_{i=1}^{\infty} \frac{1}{(i+1)^{i+1}} p_i(n+l) \mid l \in N \right\}.$$

Deshalb ist $L(\varepsilon) \leq i_0! \leq i_0^{i_0} \leq \frac{1}{\varepsilon}$.

Seien nun f_1, \dots, f_{a-1} nicht negative fastperiodische Funktionen, die Bedingung (1) erfüllen. Durch Normierung kann man $\sum_{i=1}^{a-1} f_i \leq 1$ erreichen.

$f_a = 1 - \sum_{i=1}^a f_i$ erfüllt ebenfalls die Bedingung (1).

Seien nun $(f_{1j}, f_{2j}, \dots, f_{aj})$ in gleicher Weise konstruiert ($j = 1, \dots, a$). Dann ist $\omega^t(i|j) := f_{ij}(t)$ $i, j = 1, \dots, a$ $t = 1, 2, \dots$ ein fastperiodischer nichtperiodischer Kanal, der Bedingung (1) erfüllt.

§ 3. Ein Beweis des Codingtheorems mit Hilfe einer Verallgemeinerung des π -Sequenzen-Verfahrens von Wolfowitz

In § 2 wurde das Codingtheorem für fastperiodische Simultankanäle bewiesen, die Bedingung (1) genügen. Es stellt sich jetzt das Problem, diese Bedingung zu eliminieren. Zu diesem Zweck vergegenwärtigen wir uns einmal obigen Beweisgang. Wirklich benutzt von S_n wurde die folgende Eigenschaft:

(0) zu δ_1, δ_2 und zu der Schar S_n ($n = 1, 2, \dots$) gibt es eine Schar S_n^* ($n = 1, 2, \dots$) und ein n_0 , so daß für $n \geq n_0$ folgendes gilt:

- 1) S_n^* habe eine Mächtigkeit kleiner als $e^{\delta_2 n}$;
- 2) zu $P_n(\cdot | \cdot | s) \in S_n$ gibt es ein $P_n(\cdot | \cdot | s^*) \in S_n^*$, so daß $|P_n(A | u | s) - P_n(A | u | s^*)| \leq \delta_1$ ist für alle $u \in \Omega_n, A \subset \Omega_n'$.

Eigenschaft (0) wurde durch das Lemma und Bedingung (1) sichergestellt. Wenn (0) erfüllt ist, läßt sich für eine beliebige Menge nichtstationärer Kanäle Satz 0 anwenden und damit das Codingtheorem beweisen (Satz 2 in Kap. III, § 1).

Wegen der Fastperiodizität von $(\omega^t(\cdot | \cdot))_{t=1, 2, \dots}$ gibt es zu $\varepsilon > 0$ ein $L(\varepsilon)$ und zu jedem $l = 1, 2, \dots, L(\varepsilon)$ eine Menge $\nu S_n(\varepsilon) \subset S_n$, derart, daß

$$S_n = \{P_n(\cdot | \cdot | s) \mid s \in N\} = \bigcup_{l=1, 2, \dots, L(\varepsilon)} \nu S_n(\varepsilon)$$

und für 2 Elemente

$$P_n(\cdot | \cdot | s_1), P_n(\cdot | \cdot | s_2)$$

aus $\nu S_n(\varepsilon)$ ($l = 1, 2, \dots, L(\varepsilon)$) $|\omega^t(i, j|s_1) - \omega^t(i, j|s_2)| < \varepsilon$ ($1 \leq i, j \leq a$; $t = 1, 2, \dots, n$) gilt. Durch die Festsetzung

$$\|(\omega^t(\cdot|\cdot))_{t=1, 2, \dots, n}\|_n = \sup_{\substack{i, j=1, \dots, a \\ t=1, 2, \dots, n}} |\omega^t(i, j)|$$

wird eine Norm im Abschnitt $\langle 1, n \rangle$ definiert. Jedem $P_n(\cdot|\cdot|s)$ entspricht eine Folge $(\omega^{t+s}(\cdot|\cdot))_{t=1, 2, \dots, n}$. Der Menge von Kanälen $\nu S_n(\varepsilon)$ entspricht eine Menge von Folgen, die als Teilmenge einer ε - $\|\cdot\|_n$ -Umgebung einer Folge aufgefaßt werden kann. Es besteht die Hoffnung, die maximale Länge $N(n, \lambda)$ von λ -Simultancodes für $\nu S_n(\varepsilon)$ nach unten abschätzen zu können.

Die Norm $\|\cdot\|_n$ ist ein Begriff im Produktraum. Als Verfahren zum Beweis eines Codingtheorems, welches sich stark an die Produktraumsituation anpaßt, bietet sich das π -Sequenzen-Verfahren von WOLFOVITZ an ([19]). Es soll hier so verallgemeinert werden, daß es unser Simultancodeproblem löst.

Seien wie bisher $\Omega^t = \Omega'^t = \{1, 2, \dots, a\}$, $t = 1, 2, \dots$ endliche Alphabete, dann ist $\prod_{t=1}^n \Omega^t = \Omega_n$ die Menge aller gesendeten und $\prod_{t=1}^n \Omega'^t = \Omega'_n$ die Menge aller empfangenen Nachrichten im Zeitraum $\langle 1, n \rangle$.

Sei

$$S(\varepsilon) = \{(\omega^t(i, j) + \varepsilon_{ij}^t)_{t=1, 2, \dots, n} \mid \varepsilon_{ij}^t \leq \varepsilon, \omega^t(i, j) + \varepsilon_{ij}^t \text{ stochastisch}\},$$

$S(\varepsilon)$ ist Teilmenge einer ε -Normgebung von $(\omega^t(i, j))_{t=1, 2, \dots, n}$. Für $x_n \in \Omega_n$, $y_n \in \Omega'_n$, $s \in S$ sei die Übergangswahrscheinlichkeit (Kanal) im Zeitraum $\langle 1, n \rangle$ definiert als

$$P_n(y_n = (y^1 \dots y^n) \mid x_n = (x^1 \dots x^n) \mid s) = \prod_{t=1}^n \omega^t(y^t \mid x^t \mid s).$$

Die Elemente aus $S(\varepsilon)$ denke man sich durch eine Menge $I(\varepsilon) = \{s, \dots\}$ indiziert. $S(\varepsilon)$ induziert in $\langle 1, n \rangle$ die Kanalmenge

$$S_n(\varepsilon) = \{P_n(\cdot|\cdot|s) \mid s \in I(\varepsilon)\}.$$

Ein Code (n, N, λ) für den Simultankanal $S_n(\varepsilon)$ ist eine Serie $\{(u_1, A_1), \dots, (u_N, A_N)\}$ mit $u_i \in \Omega_n$, $A_i \subset \Omega'_n$ für $i = 1, 2, \dots, N$, $A_i \cap A_j = \emptyset$ für $i \neq j$ und mit

$$P_n(A_i \mid u_i \mid s) > 1 - \lambda \quad \text{für } i = 1, 2, \dots, N, s \in I(\varepsilon).$$

WOLFOVITZ hat im Falle eines stationären Kanals (d.h. $\omega^t(i, j) = \omega(i, j)$ für alle $t = 1, 2, \dots$; $\varepsilon = 0$) ein kombinatorisches Auszählverfahren entwickelt und mit dessen Hilfe das Codingtheorem und dessen starke Umkehrung bewiesen. Dieses Verfahren wird hier unseren Anforderungen entsprechend verallgemeinert.

Definitionen.

(a) $N(i|x_n) = |\{t \mid 1 \leq t \leq n, x^t = i\}|$, d.h. Anzahl der Komponenten von x_n mit Wert i .

(b) $N(i, j|x_n, y_n) = |\{t \mid 1 \leq t \leq n, x^t = i, y^t = j\}|$.

(c) Sei $Z^n = (Z_1^n, \dots, Z_R^n)$ eine disjunkte Zerlegung von $\langle 1, n \rangle$ in die Mengen Z_q^n

($\varrho = 1, \dots, R$) und sei $n_\varrho = |Z_\varrho^n|$.

$$N^\varrho(i, j | x_n, y_n) = |\{t | t \in Z_\varrho^n, x^t = i, y^t = j\}|.$$

$$N^\varrho(i | x_n) = |\{t | t \in Z_\varrho^n, x^t = i\}|,$$

(d) Ist $\pi^{(\varrho)} = (\pi_1^{(\varrho)}, \dots, \pi_a^{(\varrho)})$ ($\varrho = 1, \dots, R$) Wahrscheinlichkeitsverteilung, so ist für jedes $t \in (1, 2, \dots, n)$ durch $\pi^t = \pi^{(\varrho)}$ falls $t \in Z_\varrho^n$ eine WV auf Ω^t gegeben.

Sei $\pi_n = \prod_{t=1}^n \pi^t$. Bei gegebenem $\pi_n, Z^n, \delta_1 > 0$ sei

$$\begin{aligned} \Omega_n(\pi_n, Z^n, \delta_1) &= \{x_n | x_n \in \Omega_n, |N^\varrho(i | x_n) - n_\varrho \pi_i^{(\varrho)}| \\ &\leq \delta_1 \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})} \text{ für } i = 1, \dots, a; \varrho = 1, \dots, R\}. \end{aligned}$$

(Die Elemente aus $\Omega_n(\pi_n, Z^n, \delta_1)$ sind Verallgemeinerungen der π -Sequenzen von WOLFOWITZ.)

(e) Mit Hilfe von

$$\bar{\omega}^\varrho(i | j) = \sup_{t \in Z_\varrho^n} (\omega^t(i | j) + \varepsilon) \wedge 1$$

$$\underline{\omega}^\varrho(i | j) = \inf_{t \in Z_\varrho^n} (\omega^t(i | j) - \varepsilon) \vee 0$$

($\bar{\omega}^\varrho(\cdot | \cdot), \underline{\omega}^\varrho(\cdot | \cdot)$ sind natürlich nicht mehr stochastisch)

definieren wir

$$\begin{aligned} \Omega'_n((\omega^t(\cdot | \cdot))_{t=1, 2, \dots, n}; \varepsilon; Z^n; x_n; \delta_2) &= \{y_n | y_n \in \Omega'_n, N^\varrho(i | x_n) \omega^\varrho(i | j) - \delta_2 [N^\varrho(i | x_n) \bar{\omega}^\varrho(i | j)]^{1/2} \\ &< N^\varrho(i, j | x_n, y_n) < N^\varrho(i | x_n) \bar{\omega}^\varrho(i | j) \\ &+ \delta_2 [N^\varrho(i | x_n) \bar{\omega}^\varrho(i, j)]^{1/2} \text{ für alle } i, j = 1, \dots, a; \varrho = 1, \dots, R\}. \end{aligned}$$

Sei Z^n folgende spezielle Zerlegung: man teile das Intervall $[0, 1]$ in die Teilintervalle $[0, \varepsilon')$, $[\varepsilon', 2\varepsilon')$, \dots , $[(1/\varepsilon'), \varepsilon', 1]$.

Sind $s, t \in (1, 2, \dots, n)$, so heißen t und s ε' -äquivalent,

(2) falls $\omega^s(i | j), \omega^t(i | j)$ in gleichen Teilintervallen liegen für gleiche (i, j) . Diese Äquivalenzrelation definiert eine Zerlegung von $\langle 1, n \rangle$ in höchstens $R = (1/\varepsilon')^{a^2}$ Klassen. Wir benötigen noch folgende Definition:

zu $\pi_n, P_n(\cdot | \cdot | s)$ sei $\pi'_n(A | s) = \sum_{x_n \in \Omega_n} \pi_n(x_n)$. $P_n(A | x_n | s)$ ($A \subset \Omega'_n$) WV auf Ω'_n .

Damit sind die Hilfsmittel zur Formulierung folgender Hilfssätze bereitgestellt:

$$\text{Hilfssatz 0. } \pi_n(\Omega_n(\pi_n, Z^n, \delta_1)) \geq \left(1 - \frac{1}{\delta_1^2}\right)^R.$$

Beweis. Nach der Tschebyscheffschen Ungleichung gilt

$$\begin{aligned} \pi_n(\{x_n | N^\varrho(i | x_n) - n_\varrho \pi_i^{(\varrho)} > \delta_1 \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})}\}) \\ \leq \frac{1}{a \delta_1^2} \text{ für } i = 1, \dots, a; \varrho = 1, \dots, R. \end{aligned}$$

$$\pi_n(\{x_n | N^\varrho(i | x_n) - n_\varrho \pi_i^{(\varrho)} > \delta_1 \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})} \text{ für irgendein } i\}) \leq \frac{1}{\delta_1^2}$$

$$\pi_n(\{x_n | N^\varrho(i | x_n) - n_\varrho \pi_i^{(\varrho)} \leq \delta_1 \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})} \text{ für } i = 1, \dots, a\}) \geq 1 - \frac{1}{\delta_1^2}.$$

Aus der Unabhängigkeit der ZV $N^1(i|x_n), \dots, N^R(i|x_n)$ folgt die Behauptung.

$$\begin{aligned} \text{Hilfssatz 1. } P_n(\Omega'_n((\omega^t(\cdot|\cdot))_{t=1,2,\dots}; \varepsilon; Z^n; x_n \delta_2) | x_n | s) \\ \geq \left(1 - \frac{a^2}{\delta_2^2}\right)^R \quad (s \in I(\varepsilon)). \end{aligned}$$

Beweis. Ist

$$\begin{aligned} \Omega'_n((\omega^t(\cdot|\cdot))_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2) \\ = \{y_n | y_n \in \Omega'_n, N^\varrho(i|x_n) \underline{\omega}^\varrho(i|j) - \delta_2 [N^\varrho(i|x_n) \overline{\omega}^\varrho(i|j)]^{1/2} \\ < N^\varrho(i,j|x_n, y_n) < N^\varrho(i|x_n) \overline{\omega}^\varrho(i|j) + \delta_2 [N^\varrho(i|x_n) \overline{\omega}^\varrho(i,j)]^{1/2} \\ \text{für alle } i, j = 1, \dots, a\} \quad (\varrho = 1, \dots, R), \end{aligned}$$

so ist

$$\begin{aligned} (3) \quad P_n(\Omega'_n((\omega^t(\cdot|\cdot))_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2) | x_n | s) \\ \geq P_n(\{y_n | |N^\varrho(i,j|x_n, y_n) - \sum_{t \in Z^n_\varrho(x_n, i)} \omega^t(j|i|s)| \\ \leq \delta_2 [N^\varrho(i|x_n) \overline{\omega}^\varrho(i|j)]^{1/2} \text{ für alle } i, j = 1, \dots, a\} | x_n | s), \end{aligned}$$

wenn man mit $Z^n_\varrho(x_n, i)$ die Menge $\{t | t \in Z^n, x_t = i\}$ bezeichnet.

Nun ist aber nach der Tschebyscheffschen Ungleichung

$$\begin{aligned} P_n(\{y_n | N^\varrho(i,j|x_n, y_n) - \sum_{t \in Z^n_\varrho(x_n, i)} \omega^t(j|i|s)| \\ > \delta_2 [N^\varrho(i|x_n) \overline{\omega}^\varrho(i|j)]^{1/2} \text{ für irgendein } (i, j)\} | x_n | s) \leq \frac{a^2}{\delta_2^2}. \end{aligned}$$

Wegen (3) und der Unabhängigkeit der $N^\varrho(i, j|x_n, y_n)$ gilt Hilfssatz 1.

Hilfssatz 2. *Ist*

$$y_n \in \bigcup_{x_n \in \Omega_n(\pi_n, Z^n, \delta_1)} \Omega'_n((\omega^t(i|j))_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2),$$

so gibt es ein $n_0(\varepsilon, \varepsilon')$, derart, daß für $n \geq n_0(\varepsilon, \varepsilon')$ gilt:

$$\pi'_n(y_n | s) < \exp\{-H(\pi'_n) + K(\varepsilon, \varepsilon', \delta_1, \delta_2) n\} \quad (s \in I(\varepsilon))$$

mit $K(\varepsilon, \varepsilon', \delta_1, \delta_2) \rightarrow 0$ für $\varepsilon, \varepsilon' \rightarrow 0$. Dabei ist π'_n die durch $\pi_n, P_n(\cdot|\cdot)$ induzierte Ausgangswahrscheinlichkeit und H die Entropiefunktion.

Beweis.

$$\pi'_n(y_n | s) < \prod_{\varrho=1}^R \prod_{j=1}^a (\pi_j^{(\varrho)})^{N^\varrho(j|y_n)} \quad (s \in I(\varepsilon))$$

mit

$$\pi_j^{(\varrho)} = \sum_{i=1}^a \pi_i^{(\varrho)} \overline{\omega}^\varrho(i|j).$$

Aus den Definitionen (d), (e) folgt

$$\begin{aligned} N^\varrho(i, j|x_n, y_n) &\geq N^\varrho(i|x_n) \omega^\varrho(i|j) - \delta_2 [N^\varrho(i|x_n) \overline{\omega}^\varrho(i|j)]^{1/2} \\ &\geq [n_\varrho \pi_i^{(\varrho)} - \delta_1 \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})}] \overline{\omega}^\varrho(i|j) \\ &\quad - \delta_2 [N^\varrho(i|x_n) \overline{\omega}^\varrho(i|j)]^{1/2}. \end{aligned}$$

Wegen $|\omega^\varrho(i|j) - \bar{\omega}^\varrho(i|j)| \leq 2\varepsilon + \varepsilon'$ gilt

$$\begin{aligned} N^\varrho(i, j|x_n, y_n) &\geq n_\varrho \pi_i^{(\varrho)} \bar{\omega}^\varrho(i|j) \\ &\quad - [(2\varepsilon + \varepsilon') n_\varrho \pi_i^{(\varrho)} + \bar{\omega}^\varrho(i|j) \delta_1] \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})} \\ &\quad + \delta_2 [N^\varrho(i|x_n) \bar{\omega}^\varrho(i|j)]^{1/2}. \end{aligned}$$

$$\begin{aligned} N^\varrho(j|y_n) &= \sum_{i=1}^a N^\varrho(i, j|x_n, y_n) \geq \sum_{i=1}^a n_\varrho \pi_i^{(\varrho)} \bar{\omega}^\varrho(i|j) \\ &\quad - [(2\varepsilon + \varepsilon') n_\varrho \pi_i^{(\varrho)} + \bar{\omega}^\varrho(i|j) \delta_1] \sqrt{a n_\varrho \pi_i^{(\varrho)} (1 - \pi_i^{(\varrho)})} \\ &\quad + \delta_2 [N^\varrho(i|x_n) \bar{\omega}^\varrho(i|j)]^{1/2}. \end{aligned}$$

Da $0 \leq \pi_j^{(\varrho)} \leq 1$ gilt die Abschätzung

$$\pi'_n(y_n|s) < \exp \left\{ \sum_{\varrho=1}^R \sum_{j=1}^a n_\varrho \pi_j^{(\varrho)} \log \pi_j^{(\varrho)} + K'(\varepsilon, \varepsilon', \delta_1, \delta_2) n \right\}$$

mit $K(\varepsilon, \varepsilon', \delta_1, \delta_2) \rightarrow 0$ für $\varepsilon, \varepsilon' \rightarrow 0$ ($s \in I(\varepsilon)$).

Aus $|\pi_j^t - \pi_j^{(\varrho)}| \leq \varepsilon + \varepsilon'$ für $t \in Z_\varrho^n$ folgt

$$\pi'_n(y_n|s) < \exp \{ -H(\pi'_n) - n(\varepsilon + \varepsilon') \log(\varepsilon + \varepsilon') + K'(\varepsilon, \varepsilon', \delta_1, \delta_2) n \} \quad (s \in I(\varepsilon)).$$

Mit $K(\varepsilon, \varepsilon', \delta_1, \delta_2) = K - (\varepsilon + \varepsilon') \log(\varepsilon + \varepsilon')$ folgt die Behauptung.

Wir benötigen noch das folgende

Lemma. Seien $\omega^\varrho(i|j)$, $\varrho = 1, \dots, R$ Matrizen mit der Eigenschaft

$$\frac{1}{\sqrt{n_\varrho}} \leq \omega^\varrho(i|j) \leq 1 \quad \text{für } i, j = 1, \dots, a; \quad \varrho = 1, \dots, R.$$

Ist $x_n \in \Omega(\pi_n, Z^n, \delta_1)$, so gilt für die Mächtigkeit

$$\begin{aligned} M((\omega^\varrho(\cdot|\cdot))_{\varrho=1, \dots, R}, Z^n, x_n) \text{ der Menge } \Omega'_n((\omega(\cdot|\cdot))_{\varrho=1, \dots, R}, Z^n, x_n) \\ = \{y_n | y_n \in \Omega'_n, N^\varrho(i|x_n) \omega^\varrho(i|j) - \sqrt{n_\varrho} < N^\varrho(i, j|x_n, y_n) \\ < N^\varrho(i|x_n) \omega^\varrho(i|j) + \sqrt{n_\varrho} \text{ für alle } i, j = 1, \dots, a; \varrho = 1, \dots, R\} \end{aligned}$$

die Abschätzung:

$$\begin{aligned} M((\omega^\varrho(\cdot|\cdot))_{\varrho=1, \dots, R}, Z^n, x_n) \\ \leq \exp \left\{ \sum_{\varrho=1}^R \sum_{i=1}^a n_\varrho \pi_i^{(\varrho)} H(\omega^\varrho(\cdot|i)) + \left(\delta_1 a^{3/2} \cdot \frac{1}{c} + a \log \sqrt{n_\varrho} \right) \sqrt{n_\varrho} \right\}. \end{aligned}$$

Beweis.

$$\begin{aligned} \prod_{\varrho=1}^R \prod_{i, j=1}^a \omega^\varrho(j|i)^{N^\varrho(i, j|x_n, y_n)} \\ \geq \prod_{\varrho=1}^R \prod_{i, j=1}^a \omega^\varrho(j|i) \{ n_\varrho \pi_i^{(\varrho)} \omega^\varrho(j|i) + \delta_1 \sqrt{a n_\varrho \omega^\varrho(j|i)} + \sqrt{n_\varrho} \} \\ = \exp \left\{ \sum_{\varrho=1}^R \sum_{i, j=1}^a n_\varrho \pi_i^{(\varrho)} \omega^\varrho(j|i) \log \omega^\varrho(j|i) + \delta_1 \sqrt{a n_\varrho} \omega^\varrho(j|i) \right. \\ \left. \cdot \log \omega^\varrho(j|i) + \sqrt{n_\varrho} \log \omega^\varrho(j|i) \right\} \\ \geq \exp \left\{ - \sum_{\varrho=1}^R \sum_{j=1}^a n_\varrho \pi_j^{(\varrho)} H(\omega^\varrho(\cdot|j)) - \left(\delta_1 a^{3/2} \frac{1}{c} + a \log \sqrt{n_\varrho} \right) \cdot \sqrt{n_\varrho} \right\}. \end{aligned}$$

Durch Reziprokenbildung erhält man die gewünschte Abschätzung.

Hilfssatz 3. Ist $x_n \in \Omega(\pi_n, Z^n, \delta_1)$, so ist die Mächtigkeit

$$M_n((\omega^t)_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2) = |\Omega'_n((\omega^t)_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2)| \\ \leq \exp \left\{ \sum_{t=1}^n \sum_{i=1}^a \pi_i^t H(\omega^t(\cdot|i)) + c(\varepsilon, \varepsilon')n + d(\varepsilon', \delta_1) \sqrt{n} \cdot \log \sqrt{n} + f(\delta_2, n)n \right\}$$

mit

$$c(\varepsilon, \varepsilon') \rightarrow 0 \quad \text{für} \quad \varepsilon, \varepsilon' \rightarrow 0 \\ f(\delta_2, n) \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty.$$

Beweis. Ist

$$y_n \in \Omega'_n((\omega^t)_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2), \quad x_n \in \Omega(\pi_n, Z^n, \delta_1),$$

so liegt $N^e(i, j|x_n, y_n)$ in dem Intervall

$$[N^e(i|x_n)\omega^e(j|i) - \delta_2(N^e(i|x_n)\bar{\omega}^e(j|i))^{1/2}, \\ N^e(i|x_n)\bar{\omega}^e(j|i) + \delta_2(N^e(i|x_n)\bar{\omega}^e(j|i))^{1/2}] \quad (i, j = 1, \dots, a; \varrho = 1, \dots, R).$$

Nun zerlegen wir dieses Intervall in disjunkte Intervalle der Länge $\sqrt{n_\varrho}$. Da die Länge des Intervalles kleiner als $n_\varrho(2\varepsilon + \varepsilon') + 2\delta_2\sqrt{n_\varrho}$ ist, erhält man höchstens $\sqrt{n_\varrho}(2\varepsilon + \varepsilon') + 2\delta_2$ Teilintervalle ($i, j = 1, \dots, a; \varrho = 1, \dots, R$).

Wir erhalten eine Klasseneinteilung in $\Omega'_n((\omega^t)_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2)$, indem wir jeweils alle y_n zusammenfassen zu einer Klasse, für die $N^e(i, j|x_n, y_n)$ in gleichen Teilintervallen liegt für alle $i, j = 1, \dots, a; \varrho = 1, \dots, R$.

Es gibt höchstens $\prod_{\varrho=1}^R (\sqrt{n_\varrho}(2\varepsilon + \varepsilon') + 2\delta_2)^{a^2}$ Klassen. Eine Klasse wird charakterisiert durch ein Tupel

$$(t(i, j, \varrho)) \quad i, j = 1, \dots, a; \varrho = 1, \dots, R.$$

$1 \leq t(i, j, \varrho) \leq \sqrt{n_\varrho}(2\varepsilon + \varepsilon') + 2\delta_2$, $t(i, j, \varrho)$ ist ganzzahlig. Für die Elemente einer solchen Klasse gilt:

$$(4) \quad |N^e(i|x_n)\omega^e(j|i) - \delta_2[N^e(i|x_n)\bar{\omega}^e(j|i)]^{1/2} \\ + t(i, j, \varrho)\sqrt{n_\varrho} - N^e(i, j|x_n, y_n)| \leq \sqrt{n_\varrho} \quad (i, j = 1, \dots, a; \varrho = 1, \dots, R).$$

Wir definieren nun

$$\tilde{\omega}^e(j|i) = \frac{(N^e(i|x_n)\omega^e(j|i) - \delta_2[N^e(i|x_n)\bar{\omega}^e(j|i)]^{1/2} + t(i, j, \varrho)\sqrt{n_\varrho})}{N^e(i|x_n)} \text{ falls } N^e(i|x_n) > 0 \\ = \frac{1}{\sqrt{n_\varrho}} \text{ sonst; ist}$$

$$\omega^{*e}(j|i) = \left(\tilde{\omega}^e(j|i) \vee \frac{1}{\sqrt{n_\varrho}} \right) \wedge 1,$$

so ist

$$\frac{1}{\sqrt{n_\varrho}} \leq \omega^{*e}(j|i) \leq 1 \quad (i, j = 1, \dots, a; \varrho = 1, \dots, R).$$

Alle y_n , die (4) erfüllen, erfüllen auch

$$(5) \quad (N^e(i|x_n)\tilde{\omega}^e(j|i) - \sqrt{n_\varrho} \leq N^e(i, j|x_n, y_n) \leq N^e(i|x_n)\tilde{\omega}^e(j|i) + \sqrt{n_\varrho} \\ (i, j = 1, \dots, a; \varrho = 1, \dots, R)$$

und deshalb auch

$$N^e(i|x_n)\omega^{*e}(j|i) - \sqrt{n_\varrho} \leq N^e(i, j|x_n, y_n) \leq N^e(i|x_n)\omega^{*e}(j|i) + \sqrt{n_\varrho}.$$

Das ist für $N^e(i|x_n) > 0$ nach Definition von $\tilde{\omega}(j|i)$ unmittelbar klar. Für $N^e(i|x_n) = 0$ folgt es aus der Tatsache, daß

$$N^e(i, j|x_n, y_n) = 0 \quad \text{ist für} \quad y_n \in \Omega'_n((\omega^t)_{t=1, 2, \dots; \varepsilon; Z^n; x_n; \delta_2}).$$

Mit Hilfe des Lemmas folgt

$$M_n((\omega^t)_{t=1, 2, \dots; \varepsilon; Z^n; x_n; \delta_2}) \leq \prod_{\varrho=1}^R (\sqrt{n_\varrho} (2\varepsilon + \varepsilon') + 2\delta_2)^{a^2}.$$

$$\text{exp. } \sum_{\varrho=1}^n \sum_{i=1}^a \pi_i^t H(\omega^t(\cdot|i)) + R(\delta_1 a^{3/2}) \frac{1}{\varepsilon} + a \log \sqrt{n} \sqrt{n} + c(\varepsilon, \varepsilon') n + f(\delta_2, n) n$$

mit

$$c(\varepsilon, \varepsilon') \rightarrow 0, \quad \text{falls} \quad \varepsilon, \varepsilon' \rightarrow 0$$

$$f(\delta_2, n) \rightarrow 0, \quad \text{falls} \quad n \rightarrow \infty,$$

wie behauptet.

Mit den Hilfssätzen 1, 2, 3 sind die Beweismittel zum Beweis des Codingtheorems bereitgestellt.

Sei nun (n, N, λ) ein Simultancode für S_n , der den folgenden Bedingungen genügt:

- (I) $u_i, i = 1, 2, \dots, N$ sind Elemente aus $\Omega_n(\pi_n, Z^n, \delta_1)$.
- (II) $A_i = \left(\bigcup_{j=1}^{i-1} A_j \right)^c \cap \Omega_n((\omega^t)_{t=1, 2, \dots; Z^n; u_i; \delta_2})$.
- (III) Der Code sei maximal in dem Sinne, daß es unmöglich ist, ein Element (u_{N+1}, A_{N+1}) hinzuzufügen, so daß (u_{N+1}, A_{N+1}) (1), (2) erfüllt für $i = N + 1$ und $\{(u_1, A_1), \dots, (u_{N+1}, A_{N+1})\}$ Simultancode $(n, N + 1, \lambda)$ für S_n ist.

Nach Hilfssatz 1 existiert stets ein Code der Länge 1. Wähle δ_2 so groß, daß in Hilfssatz 1

$$\left(1 - \frac{a^2}{\delta_2^2} \right)^R \geq 1 - \frac{\lambda}{2}$$

ist. Dann ist

$$P_n \left(\left\{ \Omega'_n((\omega^t)_{t=1, 2, \dots; Z^n; n; \delta_2}) \cap \bigcup_{i=1}^N A_i \right\} | u | S \right) > \frac{\lambda}{2}$$

für alle $u \in \{u_1, \dots, u_N\}$. Für $u \notin \{u_1, \dots, u_N\}$, $u \in \Omega_n(\pi_n, Z^n, \delta_1)$ gibt es ein s , so daß

$$P_n \left(\left\{ \Omega'_n((\omega^t)_{t=1, 2, \dots; Z^n; u; \delta_2}) \cap \bigcup_{i=1}^N A_i \right\} | u | s \right) > \frac{\lambda}{2}$$

ist, da sich sonst ein Widerspruch zu (III) ergeben würde.

Definieren wir nun $P_n^*(A|u) = \max_{s \in I(\varepsilon)} P_n(A|u|s)$ ($A \in B'_n, u \in \Omega_n$), so können wir schreiben:

$$P_n^* \left(\left\{ \Omega'_n((\omega^t)_{t=1, 2, \dots; Z^n; u; \delta_2}) \cap \bigcup_{i=1}^N A_i \right\} | u \right) > \frac{\lambda}{2} \quad (u \in \Omega_u(\pi_n, Z^n, \delta_1)).$$

Ist ferner

$$\pi_n^*(A) := \sum_{u \in \Omega_n} \pi_n(u) P_n^*(A|u) \quad (A \in B'_n),$$

so folgt mit Hilfssatz 0

$$\pi_n^* \left(\bigcup_{i=1}^N A_i \right) > \frac{\lambda}{2} \left(1 - \frac{1}{\delta_1^2} \right)^R.$$

Unter Benutzung von

$$\pi'_n(y_n | s) \leq \pi_n^*(y_n) \leq \prod_{\varrho=1}^R \prod_{j=1}^a (\pi'_j(\varrho))^{N^{\varrho(j|y_n)}} \leq \exp \{ -H(\pi'_n) + K(\varepsilon, \varepsilon', \delta_1, \delta_2) n \}$$

(Hilfssatz 2) folgt weiter

$$\left| \bigcup_{i=1}^N A_i \right| > \frac{\lambda}{2} \left(1 - \frac{1}{\delta_1^2} \right)^R \exp \{ -H(\pi'_n) + K(\varepsilon, \varepsilon', \delta_1, \delta_2) n \}.$$

Andererseits liefert Hilfssatz 3

$$\left| \bigcup_{i=1}^N A_i \right| \leq N \exp \left\{ \sum_{t=1}^n \sum_{i=1}^a \pi_i^t H(\omega^t(\cdot | i)) + c(\varepsilon, \varepsilon') n + d(\varepsilon', \delta_1) \sqrt{n} \log \sqrt{n} + f(\delta_2, n) n \right\}.$$

Aus diesen beiden Ungleichungen folgt nach einer Supremumbildung über alle möglichen π_n der

Satz 2 (Codingtheorem für ε -Normgebungen eines nicht-stationären Kanals).
Die maximale Länge $N(n, \lambda)$ eines λ -Codes für den Simultankanal

$$S(\varepsilon) = \{ (\omega^t(i|j) + \varepsilon_{ij}^t)_{t=1,2,\dots} \mid |\varepsilon_{ij}^t| \leq \varepsilon,$$

$\omega^t(i|j) + \varepsilon_{ij}^t$ stochastisch} erfüllt die Abschätzung: zu $\lambda > 0$, $0 < \lambda \leq 1$, $\delta > 0 \exists n_0$:

$$N(n, \lambda) > \exp \{ (C(n) - \delta) n - g(\varepsilon) n \} \quad \text{für } n \geq n_0.$$

Dabei ist $\lim_{\varepsilon \rightarrow 0} g(\varepsilon) = 0$,

$$C_n = \frac{1}{n} \max_{\pi^1, \dots, \pi^n} \left\{ \sum_{t=1}^n I^t(\pi^t | \omega^t(\cdot | \cdot)) \right\}.$$

Bemerkung. Es gilt der

Hilfssatz 3'. $(\omega^t(\cdot | \cdot))_{t=1,2,\dots}$ erfülle die Bedingung (G): Für jedes Paar (i, j) gilt entweder $\omega^t(i|j) \geq g > 0$ für alle t , oder $\omega^t(i|j) = 0$ für alle t .

$$\begin{aligned} \text{Ist} \quad & x_n \in \Omega(\pi_n, Z^n, \delta_1), \\ & y_n \in \Omega_n^1((\omega^t)_{t=1,2,\dots}; \varepsilon; Z^n; x_n; \delta_2), \\ & \varepsilon < \frac{g}{2}, \end{aligned}$$

so ist

$$P_n(y_n | x_n | s) \geq \exp \left\{ - \sum_{t=1}^n \sum_{i=1}^a \pi_i^t H(\omega^t(\cdot | i)) + c(\varepsilon) n + d(\varepsilon) \sqrt{n} \right\}$$

mit

$$c(\varepsilon) \rightarrow 0 \quad \text{für } \varepsilon \rightarrow 0 \quad (s \in I(\varepsilon)).$$

Beweis.

$$P_n(y_n | x_n | s) \geq \prod_{\varrho=1}^R \prod_{i,j=1}^a (\omega^{\varrho}(j|i))^{N^{\varrho(i,j|x_n,y_n)}} \quad (s \in I(\varepsilon)).$$

Aus den Definitionen (d), (e) folgt:

$$\begin{aligned}
 P_n(y_n | x_n | s) &\geq \\
 \prod_{\varrho=1}^R \prod_{i,j=1}^a (\omega^\varrho(j|i)) &\{n_\varrho \pi_i^{(\varrho)} \bar{\omega}^\varrho(j|i) + 2\delta_2 \sqrt{\delta_1} \sqrt{an_\varrho} \bar{\omega}^\varrho(j|i) + \delta_1 \sqrt{an_\varrho} \bar{\omega}^\varrho(j|i)\} \\
 &\geq \exp \left\{ \sum_{\varrho=1}^R \sum_{i,j=1}^a n_\varrho \pi_i^{(\varrho)} \bar{\omega}^\varrho(j|i) \log \omega^\varrho(j|i) \right. \\
 &\quad \left. + (2\delta_2 \sqrt{\delta_1} \sqrt{an_\varrho} + \delta_1 \sqrt{an_\varrho}) \bar{\omega}^\varrho(j|i) \log \omega^\varrho(j|i) \right\}.
 \end{aligned}$$

An dieser Stelle ergibt sich eine Schwierigkeit durch den Ausdruck

$$\bar{\omega}^\varrho(j|i) \log \omega^\varrho(j|i),$$

der im allgemeinen nach unten unbeschränkt ist. Sie wird durch Bedingung (G) behoben; es folgt die Behauptung.

Für Kanäle, die Bedingung (G) erfüllen, liefern die Hilfssätze 1, 2, 3' ein Codingtheorem für $S(\varepsilon)$. Läßt sich Bedingung (G) abschwächen?

Da gerade diejenigen y_n , die von x_n mit kleiner Übergangswahrscheinlichkeit erreicht werden, Schwierigkeiten machen, könnte man erwarten, daß man zum Ziele kommt, wenn man auf diese y_n bei der Codierung ganz verzichtet. Das wäre damit gleichbedeutend, daß man Bedingung (G) durch folgende Abänderung erzwingt:

$$\begin{aligned}
 \omega'^t(i|j) &= \omega^t(i|j) \quad \text{falls} \quad \omega^t(i|j) \geq g \\
 &= 0 \quad \text{falls} \quad \omega^t(i|j) < g.
 \end{aligned}$$

Für den binären symmetrischen Kanal

$$(\omega(1|1) = \omega(2|2) = q, \quad \omega(1|2) = \omega(2|1) = p)$$

bedeutet das, falls etwa $p < g$ ist

$$\begin{aligned}
 \omega'(1|1) &= \omega'(2|2) = q. \\
 \omega'(1|2) &= \omega'(2|1) = 0.
 \end{aligned}$$

Die maximale Länge eines λ -Codes $N'(n, \lambda)$ erfüllt: $\lim_{n \rightarrow \infty} N'(n, \lambda) = 0$, während

$$N(n, \lambda) > \exp \{n(1 + p \log p + q \log q) - K(\lambda) \sqrt{n}\}$$

und

$$\lim_{q \rightarrow 0} (1 + p \log p + q \log q) = 1$$

ist. Auf Hilfssatz 3' können wir also beim Beweis von Satz 2 nicht verzichten.

Erinnern wir uns an unser Ausgangsproblem: das Codingtheorem für den phasenunabhängigen fastperiodischen Kanal. Es ist uns mit Satz 1 gelungen, das Codingtheorem für $\nu S_n(\varepsilon)$ ($l = 1, \dots, L(\varepsilon)$) zu beweisen. Ein Code, der gleichzeitig Code ist für alle $\nu S_n(\varepsilon)$ ($l = 1, \dots, L(\varepsilon)$), ist ein Code für S_n und umgekehrt. Davon werden wir im folgenden Gebrauch machen.

Seien $\bar{\pi}^t$ WV auf Ω^t ($t = 1, 2, \dots$) für die C^t angenommen wird. Die Menge $\{\bar{\pi}^t | t = 1, 2, \dots, n\}$ wollen wir geeignet zerlegen. Zunächst definiere man nach dem Muster (2) eine ε' -äquivalente Zerlegung νZ^n in $\langle 1, n \rangle$ für $\nu S_n(\varepsilon)$ ($l = 1, \dots, L(\varepsilon)$).

Dann teile man das Intervall $[0, 1]$ in die Teilintervalle

$$[0, \varepsilon''), \quad [\varepsilon'', 2\varepsilon''), \dots, \left[\left[\frac{1}{\varepsilon'} \right] \varepsilon', 1 \right].$$

$\bar{\pi}^t$ und $\bar{\pi}^s$ heißen ε'' -äquivalent, falls $\bar{\pi}_i^t$ und $\bar{\pi}_i^s$ in gleichen Teilintervallen liegen für alle $i = 1, \dots, a$. Diese Äquivalenzrelation definiert eine Zerlegung von $\langle 1, n \rangle$ in höchstens $\left(\frac{1}{\varepsilon''}\right)^a$ disjunkte Teilmengen. Schließlich bilde man die gemeinsame Verfeinerung dieser Zerlegung mit den Zerlegungen ${}_i Z^n$ ($i = 1, \dots, L(\varepsilon)$). Man erhält so höchstens $R = \left(\frac{1}{\varepsilon''}\right)^a \cdot R(\varepsilon') L(\varepsilon)$ disjunkte Teilmengen in $\langle 1, n \rangle$ und damit in natürlicher Weise eine Zerlegung der Menge $\{\bar{\pi}^t | t = 1, 2, \dots, n\}$ in höchstens $R = \left(\frac{1}{\varepsilon''}\right)^a \cdot R(\varepsilon') L(\varepsilon)$ disjunkte Teilmengen, die wir uns durch $\varkappa = 1, \dots, R$ durchnummeriert denken. Einen Repräsentanten der \varkappa -ten Teilmenge bezeichnen wir mit π^\varkappa , die Zerlegung mit Z^n . Sei nun (n, N, λ) ein Simultancode für S_n , der den folgenden Bedingungen genügt:

(I) $u_i, i = 1, \dots, N$ sind Elemente aus $\Omega_n(\pi^\varkappa, Z^n, \delta_1)$.

(II) $A_i = \left(\bigcup_{j=1}^{i-1} A_j \right)^c \cap \bigcup_{s^*=1, \dots, L(\varepsilon)} \Omega'_n((\omega^t(\cdot | \cdot | s^*))_{t=1, 2, \dots}; \varepsilon; Z^n; u; \delta_2)$.

(III) Der Code sei maximal in dem Sinn, daß es unmöglich ist, ein Element (u_{N+1}, A_{N+1}) hinzuzufügen, so daß (u_{N+1}, A_{N+1}) (1), (2) erfüllt für $i = N+1$ und $\{(u_1, A_1), \dots, (u_{N+1}, A_{N+1})\}$ Simultancode $(n, N+1, \lambda)$ für S_n ist.

Ist $u \in \Omega_n(\pi^\varkappa, Z^n, \delta_1)$, so gibt es wegen (III) ein s^* mit der Eigenschaft:

$$P_n \left(\bigcup_{s^*=1, \dots, L(\varepsilon)} \Omega'_n((\omega^t(\cdot | \cdot | s))_{t=1, 2, \dots}; \varepsilon; Z^n; u; \delta_2) \cap \left(\bigcup_{i=1}^N A_i \right) | u | s^* \right) > \frac{\lambda}{2}.$$

Die Wahl von s^* hängt von u ab. Da es weniger als $L(\varepsilon)$ verschiedene s^* gibt, existiert nach Hilfssatz 0 ein Index s_0^* und eine Menge $B_n \subset \Omega_n(\pi^\varkappa, Z^n, \delta_1)$, so, daß

$$\pi_n(B_n) \geq \left(1 - \frac{1}{\delta_1^2}\right)^R \cdot \frac{1}{L(\varepsilon)}$$

und so, daß für jedes $u \in B_n$

$$P_n \left(\Omega'_n((\omega^t(\cdot | \cdot | s_0^*))_{t=1, 2, \dots}; \varepsilon; Z^n; u; \delta_2) \cap \left(\bigcup_{i=1}^N A_i \right) | u | s_0^* \right) > \frac{\lambda}{2}.$$

Daher gilt:

$$(6) \quad \pi_n \left(\bigcup_{u \in \Omega_n(\pi^\varkappa, Z^n, \delta_1)} \Omega'_n(\omega^t(\cdot | \cdot | s_0^*))_{t=1, 2, \dots}; \varepsilon; Z^n; u; \delta_2 \right) \cap \left(\bigcup_{i=1}^N A_i \right) | s_0^* \right) > \frac{\lambda}{2} \left(1 - \frac{1}{\delta_1^2}\right)^R \cdot \frac{1}{L(\varepsilon)}.$$

Aus (6) und Hilfssatz 2 folgt:

$$(7) \quad \left(\bigcup_{i=1}^N A_i \right) \bigcup_{u \in \Omega_n(\pi^\varkappa, Z^n, \delta_1)} \Omega'_n((\omega^t(\cdot | \cdot | s_0^*))_{t=1, 2, \dots}; \varepsilon; Z^n; u; \delta_2)$$

enthält mindestens

$$\frac{\lambda}{2} \left(1 - \frac{1}{\delta_1^2}\right)^R \cdot \frac{1}{L(\varepsilon)} \exp\{H(\pi'_n(|s_0^*|) + K(\varepsilon, \varepsilon', \delta_1, \delta_2))\}$$

Elemente.

Wir leiten nun eine obere Schranke für die Anzahl der Elemente in

$$A_i \cap \bigcup_{u \in \Omega_n(\pi^\varkappa, Z_n, \delta_1)} \Omega'_n((\omega^t(\cdot | \cdot | s_0^*))_{t=1, 2, \dots; \varepsilon; Z^n; u; \delta_2}) \quad (i = 1, \dots, N)$$

her.

Hier entsteht eine Schwierigkeit, die im Beweis von Satz 1 nicht vorhanden ist. Ein Element aus A_i , welches in $\Omega'_n((\omega^t(\cdot | \cdot | s^*))_{t=1, 2, \dots; \varepsilon; Z^n; u; \delta_2})$ enthalten ist, kann auch in $\Omega'_n((\omega^t(\cdot | \cdot | s_0^*))_{t=1, 2, \dots; \varepsilon; Z^n; u; \delta_2})$ enthalten sein für ein $u \in \Omega_n(\pi^\varkappa, Z_n, \delta_1)$. Ein Index s^* für den dieses passieren kann, heiße mit s_0^* assoziiert. Die Menge der mit s_0^* assoziierten bezeichnen wir mit $A(s_0^*)$. Wir charakterisieren nun Indices s^* , die mit s_0^* assoziiert sind.

Dazu benötigen wir noch die Definitionen:

$$\begin{aligned} \bar{\omega}^\varkappa(i | j | s^*) &= \sup_{t \in Z^n} (\omega^t(i | j | s^*) + \varepsilon) \wedge 1, \\ \pi'^\varkappa(j | s) &= \sum_{i=1}^a \pi_i^\varkappa \bar{\omega}^\varkappa(i | j | s^*). \end{aligned}$$

Ist nun $u \in \Omega_n(\pi^\varkappa, Z_n, \delta_1)$

$$v \in \Omega'_n((\omega^t(\cdot | \cdot | s^*))_{t=1, 2, \dots; Z^n; u; \delta_2}),$$

so ist

$$\begin{aligned} n_\varkappa \cdot \pi^\varkappa(j | s^*) &+ 2 \delta_2 \sqrt{\delta_1} \sqrt{a n_\varkappa} \cdot a \sqrt[4]{\pi'^\varkappa(j | s^*)} + \delta_1 a \cdot \sqrt{a n_\varkappa} \sqrt[4]{\pi'^\varkappa(j | s^*)} \\ &\geq N^\varkappa(j | v) \geq n_\varkappa(\pi'^\varkappa(j | s^*) - (\varepsilon + \varepsilon')) - 2 \delta_2 \sqrt{\delta_1} \sqrt{a n_\varkappa} \cdot a \sqrt[4]{\pi'^\varkappa(j | s^*)} \\ &- \delta_1 a \cdot \sqrt{a n_\varkappa} \sqrt[4]{\pi'^\varkappa(j | s^*)} \quad (j = 1, \dots, a; \varkappa = 1, \dots, R; s^* = 1, \dots, L(\varepsilon)) \end{aligned}$$

und deshalb

$$\begin{aligned} &|\pi'^\varkappa(j | s_0^*) - \pi'^\varkappa(j | s^*)| \\ &\leq \frac{2 \delta_2 \sqrt{\delta_1} \sqrt{a} \cdot a + \delta_1 a \sqrt{a}}{\sqrt{n_\varkappa}} (\sqrt[4]{\pi'^\varkappa(j | s_0^*)} + \sqrt[4]{\pi'^\varkappa(j | s^*)}) + (\varepsilon + \varepsilon'), \end{aligned}$$

falls s^* mit s_0^* assoziiert ist.

Daraus folgt die Abschätzung:

$$(8) \quad |H(\pi'^\varkappa(\cdot | s_0^*)) - H(\pi'^\varkappa(\cdot | s^*))| < \frac{K(a, \delta_1, \delta_2)}{\sqrt{n_\varkappa}} - a(\varepsilon + \varepsilon') \log(\varepsilon + \varepsilon')$$

für eine geeignete nur von a, δ_1, δ_2 abhängige Funktion K ($\varkappa = 1, \dots, R; s^* \in A(s_0^*)$) (vgl. WOLFOVITZ [19]).

Sei s_{00}^* dasjenige Element aus $A(s_0^*)$, für welches

$$\sum_{\varkappa=1}^R n_\varkappa \sum_i \pi_i^\varkappa H(\omega^\varkappa(\cdot | i | s_{00}^*)) = \max_{s \in A(s_0^*)} \sum_{\varkappa=1}^R n_\varkappa \sum_{i=1}^a \pi_i^\varkappa H(\omega^\varkappa(\cdot | s)).$$

Nach Hilfssatz 3 und nach Definition von A_i ($i = 1, \dots, N$) folgt, daß

$$A_i \cap \left(\bigcup_{u \in \Omega_n(\pi^\varkappa, Z_n, \delta_1)} \Omega'_n((\omega^t(\cdot | \cdot | s_0^*))_{t=1, 2, \dots; \varepsilon; Z^n; u; \delta_2}) \right)$$

weniger als

$$L(\varepsilon) \exp \left\{ \sum_{\varkappa=1}^R n_\varkappa \sum_{i=1}^a \pi_i^\varkappa H(\omega^\varkappa(\cdot | i | s_{00}^*)) + c(\varepsilon, \varepsilon') n + d(\varepsilon', \delta_1) \sqrt{n} \log \sqrt{n} + f(\delta_2, n) \right\}$$

Elemente enthält.

$$\bigcup_{i=1}^N (A_i \cap (\bigcup_{t=1,2,\dots} \Omega'_n((\omega^t(\cdot | \cdot | s_0^*))_{t=1,2,\dots}; \varepsilon; Z^n; u; \delta_2)))$$

enthält weniger als das N -fache.

Mit (7) folgt daher

$$N > \frac{\lambda}{2} \left(1 - \frac{1}{\delta_1}\right)^R \frac{1}{L(\varepsilon)^2} \exp \left\{ H(\pi'_n(\cdot | s_0^*)) - \sum_{\kappa=1}^R n_{\kappa} \sum_{i=1}^a \pi_i^{\kappa} \cdot (H \bar{\omega}^{\kappa}(\cdot | i | s_{00}^*)) \right. \\ \left. + (K(\varepsilon, \varepsilon', \delta_1, \delta_2) - c(\varepsilon, \varepsilon')) n - d(\varepsilon, \delta_1) \sqrt{n} \log \sqrt{n} - f(\delta_2, n) n \right\}$$

und weiter mit (8)

$$(9) \quad N > \frac{\lambda}{2} \left(1 - \frac{1}{\delta_1}\right)^R \cdot \frac{1}{L(\varepsilon)^2} \exp \left\{ \sum_{\kappa=1}^R n_{\kappa} H(\pi^{\kappa}(\cdot | s_{00}^*)) \right. \\ \left. - \sum_{\kappa=1}^R n_{\kappa} \sum_{i=1}^a \pi_i^{\kappa} H(\bar{\omega}^{\kappa}(\cdot | i | s_{00}^*)) + K(\varepsilon, \varepsilon', \delta_1, \delta_2) n - c(\varepsilon, \varepsilon') n \right. \\ \left. - d(\varepsilon', \delta_1) \sqrt{n} \log \sqrt{n} - f(\delta_2, n) n + R K(a, \delta_1, \delta_2) \sqrt{n} \right. \\ \left. - a(\varepsilon + \varepsilon') \log(\varepsilon + \varepsilon') n \right\}.$$

Nach Definition von C_n , $\pi^{(\kappa)}$, $\bar{\omega}^{\kappa}(\cdot | \cdot)$ und π'^{κ} folgt:

$$(10) \quad n C_n \leq \sum_{\kappa=1}^R n_{\kappa} H(\pi'^{\kappa}(\cdot | s_{00}^*)) - \sum_{\kappa=1}^R n_{\kappa} \sum_{i=1}^a \pi_i^{\kappa} H(\bar{\omega}(\cdot | i | s_{00}^*)) + n K(\varepsilon, \varepsilon', \varepsilon''),$$

wobei $K(\varepsilon, \varepsilon', \varepsilon'') \rightarrow 0$ für $\varepsilon, \varepsilon', \varepsilon'' \rightarrow 0$.

Aus (10) und (11) folgt der

Satz 3 (Codingtheorem für den fastperiodischen Simultankanal).

Für die maximale Länge $N(n, \lambda)$ eines λ -Codes des fastperiodischen Simultankanals S_n gilt:

zu $\lambda > 0$, $\delta > 0$ gibt es ein $n_0(\lambda, \delta)$, so daß für $n \geq n_0$ $N(n, \lambda) > e^{(C_n - \delta)n}$ ist.

(Aus der Definition von C_n folgt, daß C_n monoton wachsend ist. Da andererseits $C_n \leq \log a$ ist, gilt $\lim_{n \rightarrow \infty} C_n = C$, d. h. unsere Kapazitätsfunktion kann durch eine Konstante C ersetzt werden.)

§ 4. Die starke Umkehrung des Codingtheorems für den fastperiodischen Simultankanal

Satz 4 (Starke Umkehrung des Codingtheorems).

Für die maximale Länge $N(n, \lambda)$ eines λ -Codes des fastperiodischen Simultankanals S_n gilt:

zu $\lambda > 0$, $0 \leq \lambda < 1$, $\delta > 0$ gibt es ein $n_0(\lambda, \delta)$, so daß für $n \geq n_0(\lambda, \delta)$

$N(n, \lambda) \leq e^{(C_n + \delta)n}$ ist.

Beweis. Wir benötigen die folgende — zur von Neumannschen Definition äquivalente — Definition für fastperiodisch. Die reellwertige Funktion $f(n)$ heißt fastperiodisch auf der Halbgruppe N , wenn es zu jedem $\varepsilon > 0$ ein System von

endlich vielen Teilmengen N_1, \dots, N_n von N gibt, für welches gilt:

$$(a) \quad N = \bigcup_{i=1}^n N_i,$$

$$(b) \quad |f(m + n_1) - f(m + n_2)| < \varepsilon,$$

falls n_1, n_2 aus einem beliebigen N_i und m beliebig in N gewählt sind.

Auf Grund dieser Definition gibt es eine Zerlegung des Abschnittes $\langle 1, 2, \dots, n \rangle$ in disjunkte Mengen Z_ϱ^n ($\varrho = 1, \dots, R(\varepsilon)$) derart, daß für t_1, t_2 aus einem Z_ϱ^n

$$\|\omega^{t_1+s}(\cdot|\cdot) - \omega^{t_2+s}(\cdot|\cdot)\| < \varepsilon \text{ ist für alle } \varepsilon > 0 \text{ und für alle } s \in N \cup \{0\}.$$

Wir betrachten die Menge von Eingangswahrscheinlichkeiten $V = \{(\pi^1, \dots, \pi^{R(\varepsilon)})$

$|\pi_i$ ist ganzzahliges Vielfaches von $\frac{1}{n_\varrho}$ für $i = 1, \dots, a; \varrho = 1, \dots, R(\varepsilon)\}$.

Jedem $(\pi^1, \dots, \pi^{R(\varepsilon)}) \in V$ entspricht eine Menge von Codeworten

$$\{u_i | i = 1, \dots, N\} \cap \left\{u \mid \frac{N_\varrho(i|u)}{n_\varrho} = \pi_\varrho^i, i = 1, \dots, a; \varrho = 1, \dots, R(\varepsilon)\right\}.$$

Wir denken uns diese Codeworte als u_{01}, \dots, u_{0M} gegeben und betrachten den Teilcode $\{(u_{01}, A_{01}), \dots, (u_{0M}, A_{0M})\}$. Sei nun $s \in N$ so gewählt, daß

$$H_n(\pi'_n(\cdot|s)) - \sum_{t=1}^n \sum_{i=1}^a \pi_i^t H(\omega^t(\cdot|i|s)) < C_n + \frac{1}{n}$$

ist. Den Ungleichungen der Hilfssätze 1, 2, 3 kann man Ungleichungen, die jeweils in entgegengesetzter Richtung verlaufen, an die Seite stellen.

Mit ihnen kann man — analog zum stationären Fall (vgl. WOLFWITZ [19] Kap. 4, § 4) — folgende Abschätzung herleiten:

$$M \leq \exp \{n C_n + K(\varepsilon) n + \sqrt{n}\}.$$

Deshalb gilt

$$N(n, \lambda) \leq (n+1)^{aR(\varepsilon)} \exp \{n C_n + K(\varepsilon) n + \sqrt{n}\}$$

mit $K(\varepsilon) \rightarrow 0$ für $\varepsilon \rightarrow 0$.

Hieraus folgt die Behauptung des Satzes. Auf eine ausführlichere Darstellung wurde verzichtet, da der Satz ein Spezialfall von Satz 1 aus Kap. III, § 2 ist.

Kapitel III. Nichtstationäre Simultankanäle

§ 1. Das Codingtheorem

In Kapitel I führte das Studium fastperiodischer Kanäle zu einer Idee, die sich für beliebige nichtstationäre Kanäle als brauchbar erweist. Im folgenden wird untersucht, wie weit sich die Ergebnisse aus Kapitel II verallgemeinern lassen. BLACKWELL — BREIMANN — THOMASIAN haben im Falle einer Klasse stationärer Kanäle Codingtheorem und schwache Umkehrung (Kap. I, § 1, Aussage (3)) bewiesen ([3]). WOLFWITZ bewies die starke Umkehrung ([18]). In der nichtstationären Situation kann man sich folgendes allgemeine Problem stellen:

Sei $S = \{(\omega^t(\cdot|s))_{t=1, 2, \dots} | s \in I\}$ eine Menge nichtstationärer Kanäle mit unabhängigen Zeichen. Unter welchen Voraussetzungen gilt das Codingtheorem, bzw. die schwache Umkehrung des Codingtheorems, bzw. die starke Umkehrung

des Codingtheorems? Beim Beweis des Codingtheorems in Kap. II wurde von der Fastperiodizität nicht voll Gebrauch gemacht. Bedient man sich der Definition für fastperiodisch aus Kap. II, § 4, so sieht man, daß für den fastperiodischen Kanal $(\omega^t(\cdot|\cdot))_{t=1, 2, \dots}$ die Translationsinvarianz einer ε -äquivalenten Zerlegung ungenutzt blieb. Zu allgemeineren Resultaten werden wir über die von Neumannsche Definition der Fastperiodizität (Kap. II, § 2) geführt.

Wir haben benutzt, daß $\{(\omega^{t+s}(\cdot|\cdot))_{t=1, 2, \dots} | s = 0, 1, 2, \dots\}$ relativ kompakt ist in der gleichmäßigen Norm. Diese Eigenschaft läßt sich zu einer Kompaktheit im Ausschnitt $\langle 1, n \rangle$ mit Wachstumseigenschaft abschwächen.

Mit $S_n = \{P_n(\cdot|\cdot|s) | s \in I\}$ sei der durch S im Ausschnitt $\langle 1, n \rangle$ induzierte Simultankanal bezeichnet. Jedem $P_n(\cdot|\cdot|s)$ entspricht eine Folge $(\omega^t(\cdot|\cdot|s))_{t=1, 2, \dots, n}$. Für die weitere metrische Behandlung der Menge $\{(\omega^t(\cdot|\cdot|s))_{t=1, 2, \dots} | s \in I\}$ legen wir den Supremumsabstand

$$\|(\omega^t(\cdot|\cdot|s))_{t=1, 2, \dots, n} - (\omega^t(\cdot|\cdot|s'))_{t=1, 2, \dots, n}\|_n = \sup_{\substack{i, j = 1, \dots, n \\ t = 1, 2, \dots, n}} |\omega^t(i|j|s) - \omega^t(i|j|s')|$$

zugrunde.

Dadurch ist auch S_n in natürlicher Weise metrisiert. $K(n, \varepsilon)$ gebe die minimale Anzahl von ε -Umgebungen an, die S_n überdecken. Ersetzt man in Kap. II, § 3 beim Beweis des Codingtheorems $L(\varepsilon)$ durch $K(n, \varepsilon)$, so erhält man die Abschätzung (9) aus Kap. II, § 3. Falls

$$R = \left(\frac{1}{\varepsilon'}\right)^a R(\varepsilon')^{K(n, \varepsilon)} = 0 \quad (\sqrt{n})$$

ist, folgt daraus das Codingtheorem.

Es gilt deshalb der

Satz 1 (Codingtheorem für nichtstationäre Simultankanäle).

Ist $S = \{(\omega^t(\cdot|\cdot|s))_{t=1, 2, \dots} | s \in I\}$ ein nichtstationärer Simultankanal, der

(W) $K(n, \varepsilon) = 0$ ($\log \sqrt{n}$) erfüllt, so gibt es zu λ ($0 < \lambda < 1$), $\delta > 0$ ein $n_0(\lambda, \delta)$, derart, daß für $n \geq n_0(\lambda, \delta)$

$$N(n, \lambda) > e^{(C_n - \delta)n}$$

gilt.

Definiert man in $S_n = \{P_n(\cdot|\cdot|s) | s \in I\}$ eine Metrik durch die Festsetzung

$$|P_n(\cdot|\cdot|s) - P_n(\cdot|\cdot|s')| = \sup_{\substack{A \subset \Omega_n \\ x_n \in \Omega_n}} |P_n(A|x_n|s) - P_n(A|x_n|s')|$$

und $K'(n, \varepsilon)$ als Anzahl einer minimalen ε -Überdeckung, so folgt aus dem Maximalcodesatz (Kap. II, § 2) der

Satz 2. Ist $S = \{(\omega^t(\cdot|\cdot|s))_{t=1, 2, \dots} | s \in I\}$ ein nichtstationärer Simultankanal, der

(W') $K'(n, \varepsilon) \leq e^{0(n)}$ erfüllt, so gibt es zu λ ($0 < \lambda < 1$), $\delta > 0$ ein $n_0(\lambda, \delta)$, derart, daß für $n \geq n_0(\lambda, \delta)$

$$N(n, \lambda) > e^{(C_n - \delta)n}$$

gilt.

Beispiele für Kanäle, die (W') erfüllen, sind:

- a) endlich viele nichtstationäre Kanäle;
- b) beliebig viele stationäre Kanäle;
- c) eine Menge nichtstationärer Kanäle, für die S_n aus höchstens a^{n^α} ($0 < \alpha < 1$) Kanälen besteht.

§ 2. Die starke Umkehrung des Codingtheorems

Der Beweis der starken Umkehrung für den fastperiodischen Simultankanal macht von der Translationsinvarianz der ε -äquivalenten Zerlegung Gebrauch, er ist nicht direkt auf die allgemeine Situation übertragbar. Folgende Überlegung macht hier die Gültigkeit einer starken Umkehrung plausibel. S bestehe aus den stationären Folgen

$$\begin{aligned} \omega(\cdot|\cdot|1), & \quad \omega(\cdot|\cdot|1) \dots \\ \omega(\cdot|\cdot|2), & \quad \omega(\cdot|\cdot|2) \dots \\ \omega(\cdot|\cdot|3), & \quad \omega(\cdot|\cdot|3) \dots \end{aligned}$$

und aus gewissen nichtstationären Folgen

$$T = \{(\omega^t(\cdot|\cdot|s))_{t=1,2,\dots} | s \in I\}$$

mit der Eigenschaft

$$\omega^t(\cdot|\cdot|s) \in \{\omega(\cdot|\cdot|i) | i = 1, 2, \dots\}.$$

In diesem Falle gilt

$$\begin{aligned} (1) \quad C_n &= \frac{1}{n} \max_{\pi^1, \dots, \pi^n} \inf_{s \in I \cup N} I(\pi_n | P_n(\cdot|\cdot|s)) \\ &= \max_{\pi} \inf_{s \in N} I(\pi | \omega(\cdot|\cdot|s)), \end{aligned}$$

denn

$$\begin{aligned} & \max_{\pi^1, \dots, \pi^n} \inf_{s \in I \cup N} I(\pi_n | P_n(\cdot|\cdot|s)) \\ &= \max_{\pi^1, \dots, \pi^n} \inf_{s \in I \cup N} \sum_{t=1}^n I(\pi^t | \omega^t(\cdot|\cdot|s)) \\ &= n \max_{\pi} \inf_{s \in N} I(\pi | \omega(\cdot|\cdot|s)). \end{aligned}$$

Da durch Hinzunahme der Kanäle aus T zu den stationären Kanälen die Güte von Codes nur schlechter werden kann, $N(n, \lambda)$ also kleiner wird, liefert hier der Wolfowitzsche Beweis der starken Umkehrung für stationäre Simultankanäle die obere Abschätzung:

$$(2) \quad \text{zu } \lambda (0 < \lambda < 1), \delta > 0 \text{ gibt es ein } n_0(\lambda, \delta), \text{ so daß für } n \geq n_0(\lambda, \delta) \\ N(n, \lambda) < e^{(C_n + \delta)n}$$

gilt.

Falls S die Voraussetzungen von Satz 1 oder Satz 2 aus § 1 erfüllt, gilt also die starke Umkehrung. Gleichung (1) zeigt, daß bei der Berechnung der Kapazität C_n gewisse Kanäle überhaupt keinen Beitrag leisten, C_n ist also vermutlich eher zu groß als zu klein auch für beliebige nichtstationäre S . (2) müßte also allgemein gelten. Das ist in der Tat der Fall.

In [1] findet man einen Beweis einer starken Umkehrung für *einen* nichtstationären Kanal — unter Bedingungen an die Alphabete, die allgemeiner sind als wir sie hier benötigen —, der ohne Stationarisierungsverfahren arbeitet. Die Untersuchungen beziehen sich auf „alphabetfreie“ Kanäle und sind in einer entsprechenden Terminologie abgefaßt. Wir übertragen das, was davon für uns brauchbar ist, auf unsere Situation.

Wir erinnern an die Definition eines halbstetigen nichtstationären Kanals und an die Definition eines Codes in Kap. I, § 1. Sei nun

$$\{(u_i = (x_i^1, \dots, x_i^n), A_i) \mid i = 1, \dots, N\}$$

ein Code (n, N, λ) maximaler Länge — d. h. $N = N(n, \lambda)$ — für den halbstetigen nichtstationären Kanal.

Man definiere:

$$q_n(\cdot) := q^1 x, \dots, x q^n = \left(\frac{1}{N} \sum_{i=1}^N F^1(\cdot \mid x_i^1) \right) \times \dots \times \left(\frac{1}{N} \sum_{i=1}^N F^n(\cdot \mid x_i^n) \right).$$

Ist $0 < b < 1$, so gilt die Abschätzung

$$(3) \quad \log(\lambda b N(n, \lambda)) \leq \frac{1}{N} \sum_{i=1}^N \int dF^n(\cdot \mid u_i) \log \frac{dF^n(\cdot \mid u_i)}{dq(\cdot)} \\ + \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{\lambda(1-b)} \int dF^n(\cdot \mid u_i) \left(\log \frac{dF^n(\cdot \mid u_i)}{dq_n(\cdot)} \right. \right. \\ \left. \left. - \int dF^n(\cdot \mid u_i) \log \frac{dF^n(\cdot \mid u_i)}{dq_n} \right)^2 \right]^{1/2}.$$

Der zweite Term der rechten Seite wird durch $\frac{1}{\lambda(1-b)} K \sqrt{n}$ nach oben abgeschätzt, der erste Term ist gleich

$$(4) \quad \sum_{t=1}^n \left(\frac{1}{N} \sum_{i=1}^N \int dF^t(\cdot \mid x_i^t) \log \frac{dF^t(\cdot \mid x_i^t)}{dq^t(\cdot)} \right) \leq \sum_{t=1}^n C^t.$$

Es gilt deshalb der

Satz 1 (*starke Umkehrung des Codingtheorems*).

Ist $0 < \lambda < 1$, $0 < b < 1$; dann ist

$$N(n, \lambda) \leq \exp \left\{ n C_n + \frac{1}{\lambda(1-b)} K \sqrt{n} - \log \lambda b \right\}.$$

Die Grundidee des Beweises besteht darin, statt der beim Fanoschen Lemma [19] benutzten Quellenwahrscheinlichkeit: $1/N$ Masse auf jedem Codewort, der bei einer starken Umkehrung erforderlichen Produktraumstruktur entsprechend folgende Quelle zu definieren:

$$\pi_j^t = \frac{|\{i \mid x^t = j, i \in (1, 2, \dots, N)\}|}{N}, \quad j = 1, \dots, a.$$

$(q^t(\cdot))$ errechnet sich zu $q^t(\cdot) = \sum_{j=1}^a \pi_j^t F^t(\cdot \mid j)$.

Mit $\pi_n = (\pi^1, \dots, \pi^n)$ gilt

$$\sum_{t=1}^n \left(\frac{1}{N} \sum_{i=1}^N \int dF^t(\cdot | x_i^t) \log \frac{dF^t(\cdot | x_i^t)}{dq^t(\cdot)} \right) = I(\pi_n | F^n(\cdot | \cdot))$$

und deshalb die für die Simultantheorie wichtige Abschätzung

$$N(n, \lambda) \leq \exp \left\{ I(\pi_n | F^n(\cdot | \cdot)) + \frac{1}{\lambda(1-b)} K \sqrt{n} - \log \lambda b \right\}.$$

Im Simultanfall gilt nämlich

$$N(n, \lambda) \leq \exp \left\{ I(\pi_n | F^n(\cdot | \cdot | s)) + \frac{1}{\lambda(1-b)} K \sqrt{n} - \log \lambda b \right\} \quad (s \in I),$$

folglich auch

$$\begin{aligned} N(n, \lambda) &\leq \exp \left\{ \inf_{s \in I} I(\pi_n | F^n(\cdot | \cdot | s)) + \frac{1}{\lambda(1-b)} K \sqrt{n} - \log \lambda b \right\} \\ &\leq \exp \left\{ n C_n + \frac{1}{\lambda(1-b)} K \sqrt{n} - \log \lambda b \right\}. \end{aligned}$$

Damit haben wir den

Satz 1' (*starke Umkehrung des Codingtheorems für einen nichtstationären Simultankanal*).

a) Für die maximale Länge $N(n, \lambda)$ eines λ -Codes eines beliebigen nichtstationären Simultankanals S_n gilt die Abschätzung:

zu $\lambda > 0$, $\delta > 0$ gibt es ein $n_0(\lambda, \delta)$, so daß für $n \geq n_0(\lambda, \delta)$

$$N(n, \lambda) \leq e^{(C_n - \delta)n}$$

ist.

b) Gilt für S_n ein Codingtheorem mit Kapazitätsfunktion C_n (etwa falls S_n (W) oder (W') erfüllt), so liefert a) die starke Umkehrung des Codingtheorems.

Der Vorteil des obigen Beweises gegenüber dem Wolfowitzschen Beweis im stationären Simultanfall [18] besteht darin, daß nicht erst Längen von Teilcodes, sondern direkt die Länge des Gesamtcodes abgeschätzt wird.

Die Ergebnisse von Kap. III sind echt allgemeiner als die in Kap. II, wie folgendes Beispiel zeigt.

$$\begin{array}{l} \text{Kanal 1: } \omega_1, \omega_2, \omega_1, \omega_2, \omega_1, \dots \\ \text{Kanal 2: } \omega_1, \omega_1, \omega_2, \omega_1, \omega_1, \omega_2, \dots \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ \text{Kanal } v: \underbrace{\omega_1, \omega_1, \dots, \omega_1, \omega_2, \omega_1, \dots}_{v\text{-mal}} \\ \quad \cdot \\ \quad \cdot \end{array}$$

Die Kanalmenge ist nicht relativ kompakt in der L^∞ -Norm, sie genügt aber der Wachstumsbedingung (W') mit $K'(n, \varepsilon) = n$.

Wir haben jetzt unter den Wachstumsbedingungen (W), (W') die Resultate aus Kap. II auf beliebige Klassen nichtstationärer Kanäle verallgemeinert. Dabei haben wir unsere ursprüngliche Problemstellung (Kap. II, Einleitung): für einen nichtstationären Kanal einen λ -Code translationsinvariant zu definieren und die maximale Länge solcher λ -Codes asymptotisch abzuschätzen, außer acht gelassen. Es wäre nun denkbar, daß man in dieser speziellen Situation

$$(S = \{(\omega^{t+s}(\cdot|\cdot))_{t=1,2,\dots} | s = 0, 1, 2, \dots\}$$

(wir sprechen kurz vom Phasenproblem) auf Bedingung (W) verzichten kann. Unsere Hoffnungen begräbt das folgende *Beispiel*. Seien ω_1, ω_2 verschiedene stochastische Matrizen. Sei S die Gesamtheit aller Matrizenfolgen, die aus ω_1, ω_2 gebildet werden können [„Arbitrarily varying channels“ nach KIEFER-WOLFOWITZ [13]].

Andererseits denke man sich einen Kanal gegeben, der folgendem Bildungsgesetz genügt: Man schreibe hintereinander (in irgendeiner Reihenfolge) alle Matrizenfolgen der Länge 1, dann der Länge 2 usw. Das Phasenproblem für diesen Kanal ist äquivalent mit dem Simultancodeproblem für S . Da aber S (W) nicht erfüllt, ist (W) auch beim Phasenproblem nicht erfüllt (gilt auch für W'). Es ist aber auch nicht jedes Simultancodeproblem als Phasenproblem deutbar.

Beispiele. 2 stationäre Kanäle

$$\begin{aligned} \omega_1, \omega_1 \dots \\ \omega_2, \omega_2 \dots \end{aligned}$$

(oder auch ein nichtstationärer Kanal) lassen sich nicht als Gesamtheit aller Kanäle auffassen, die aus einem Kanal durch Translation entstehen.

Kapitel IV. Reichweite des Mitteilungsverfahrens, Versagen der üblichen Kapazitätsfunktion, Beispiele

Der Beweis für den Maximalcodesatz Kap. II, § 2 beruht auf einem Mitteilungsverfahren. Dieses Verfahren ist ziemlich grob. Die Bedingung (W') ist so gewählt, daß das Verfahren gerade noch funktioniert. Zu allgemeineren Resultaten kann man nur über ein neues Verfahren gelangen. Dem Mitteilungsverfahren angepaßt ist die Definition der Kapazitätsfunktion

$$C_n = \frac{1}{n} \max_{\pi^1, \dots, \pi^n} \inf_{s \in S} I(\pi_n | P_n | s).$$

Die heuristische Betrachtung in Kap. III, § 2 veranschaulicht, wie grob C_n den Kanälen angepaßt ist; bei der dortigen Wahl von S geben die Kanäle aus T keinen Beitrag zur Kapazitätsfunktion. Ist die Bedingung (W) oder (W') erfüllt, so haben S und $S - T$ asymptotisch die gleichen Codelängen.

Besteht S aus der Menge aller Folgen, die aus ω_1, ω_2 gebildet werden können [arb. var. Channels], so ist (W) nicht erfüllt, unser Mitteilungsverfahren versagt, aber auch die Definition der Kapazitätsfunktion.

$$(1) \quad C_n = \frac{1}{n} \max_{\pi^1, \dots, \pi^n} \inf_{s \in S} I(\pi_n | P_n | s)$$

liefert nichts, das man für Codingtheorem und Umkehrung gebrauchen könnte, wie das folgende *Beispiel* zeigt.

$$\omega_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \omega_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Im stationären Simultanfall ist $N(n, \lambda) = \frac{2^n}{2} = \exp\{n \log 2 - \log 2\}$ und deshalb $C = \log 2$.

Für S gilt aber $N(n, \lambda) = 1$ und damit auch Codingtheorem und starke Umkehrung mit Kapazität $0 \neq C$. Die Brauchbarkeit der Formel (1) scheint also wesentlich an die Bedingung (W) gebunden zu sein. Andererseits ist (W) nicht notwendig für die Brauchbarkeit der C_n -Definition (1) als Simultankapazität, wie folgendes triviale Beispiel zeigt:

$$\omega_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad \omega_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sowohl im stationären Fall, als auch für S gilt $N(n, \lambda) = 1$, $C_n = 0$. Die Formel (1) ist also brauchbar, obwohl S Bedingung (W) nicht erfüllt. Mit Bedingung (W) allein ist also keine vollständige Charakterisierung der Leistungsfähigkeit obiger Kapazitätsformel zu erwarten.

Bemerkenswert an unseren Beispielen ist, daß es sich hier um Matrizen mit Elementen 0, 1 handelt. Die Schwierigkeiten liegen also bereits in der nichtstochastischen Theorie. (Zum Beispiel ist für

$$\omega_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \omega_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

im stationären Fall $N(2, \lambda) = 5$, im Falle S $N(2, \lambda) = 4$. Unsere Untersuchungen berühren sich hier mit der Theorie der Gruppencodes.)

Es sei noch vermerkt, daß für den Fall, wo S aus allen Folgen besteht, die aus t Matrizen $\omega_1 \dots \omega_t$ gebildet werden können, man zwar nicht Codingtheorem und Umkehrung, aber doch noch brauchbare Abschätzungen mit Hilfe der Mitteilungsmethode erhält, falls t klein gegen a ist:

$$N(n, \lambda) < \exp\{Cn - n \log t - \delta n\}.$$

C hat die Größenordnung $\log a$.

STRASSEN hat in [16] für eine spezielle Klasse von „arbitrarily varying channels“, Codingtheorem und schwache Umkehrung bewiesen. Die dortige Kapazitätsformel ermöglicht keinen Beweis einer starken Umkehrung. Die bisherigen Beweise für starke Umkehrungen im Simultanfall benutzen die obige Definition (1) der Kapazität. Wo diese Definition anwendbar ist, gilt automatisch eine starke Umkehrung. Ob es nichtstationäre Simultankanäle ohne starke Umkehrung gibt, ist noch unbekannt. Ein Beispiel für einen Kanal ohne starke Umkehrung gab JACOBS in [10], [20]. Wir geben noch ein Beispiel für „arbitrarily varying channels“ an, wo Codingtheorem und starke Umkehrung gelten.

a) Sei ω_1 die Einheitsmatrix, ω_2 eine Permutationsmatrix. Wegen $N(n, \lambda) = N^n(1, \lambda) = N^n(1, 0)$ genügt es, eine Zeitkomponente für $\lambda = 0$ zu betrachten.

Bezeichnet l_i ($i = 1, \dots, I$) die Zyklenlängen von ω_2 , so ist

$$N(1, 0) = \sum_{i=1}^I \left[\frac{l_i}{2} \right], \quad ([X] = \text{größte ganze Zahl } \leq x)$$

$$C = \log \sum_{i=1}^I \left[\frac{l_i}{2} \right].$$

b) Seien ω_k ($k = 1, \dots, K$) stochastische Matrizen, die nur 0, 1 Elemente haben. ω_k ist im allgemeinen nicht Permutationsmatrix.

$$\text{Ist } \omega(i|j) = \sum_{k=1}^K \omega_k(i|j),$$

so ist $N(1, 0) =$ Maximalzahl trägerfremder Zeilen von $\bar{\omega}$,
 $=$ Maximalzahl trägerfremder Zeilen der stochastischen Matrix
 $\bar{\omega} = \frac{1}{K} \omega$.

Es ist also $\log N(1, 0) =$ „zero error capacity“ von $\bar{\omega}$ ([15]).

Herrn Professor JACOBS gilt mein besonderer Dank für zahlreiche Anregungen.

Literatur

1. AUGUSTIN, U.: Gedächtnisfreie Kanäle für diskrete Zeit. Z. Wahrscheinlichkeitstheorie verw. Geb. **6**, 10–61 (1966).
2. BLACKWELL, D., L. BREIMANN, and J. THOMASIAN: Proof of Shannon's transmission theorem for finite-state indecomposable channels. Ann. math. Statistics **29**, 1209–1220 (1958).
3. — — — The capacity of a class of channels. Ann. math. Statistics **30**, 1229–1241 (1960).
4. DOBRUSHIN, R. L.: Asymptotic estimate of the probability of error for transmission of messages over a discrete memoryless communication channel with asymptotic transition probability matrix. Theor. Probab. Appl. **7**, 283–311 (1962).
5. FEINSTEIN, A.: Foundations of Information Theory. New York-Toronto-London: McGraw-Hill Book Co. 1958.
6. HU GUO-DING (HU KUO-TING), and SHEN SHI-YI (SHEN SHIH-I): Some coding theorems for almost-periodic channels. Chinese Math-Acta **6**, Amer. Math. Soc. Translat of acta mathematica sinica **15**, 136–152 (1965).
7. JACOBS, K.: Einführung in die Informationstheorie. Göttingen: Vorlesungsausarbeitung 1959 unveröffentlicht.
8. — Die Übertragung diskreter Informationen durch periodische und fast-periodische Kanäle. Math. Ann. **137**, 125–135 (1959).
9. — Über die Durchlaßkapazität periodischer und fastperiodischer Kanäle. Trans. 2nd Prague Conf. Information Theory, Statist. Decision Functions, Random Processes, 231–249 (1960).
10. — Almost periodic channels. Colloquium on Combinatorial Methods in Probability Theory, Aarhus, 118–126 (1962).
11. KEMPERMANN, J. H. B.: Studies in coding theory I. To appear in Illinois Math.
12. KESTEN, H.: Some remarks on the capacity of compound channels in the semicontinuous case. Inform. and Control **4**, 168–184 (1961).
13. KIEFER, J., and J. WOLFOWITZ: Channels with arbitrarily varying channel probability functions. Inform. and Control **5**, 44–54 (1962).
14. SHANNON, C. E.: A mathematical theory of communication. Bell System J. **27**, 379–423, 623–656 (1948).
15. — The zero capacity of a noisy channel IRE. IRE Trans. Inform. Theory, IT-2, 8–13 (1956).
16. STRASSEN, V.: Meßfehler und Information. Z. Wahrscheinlichkeitstheorie verw. Geb. **2**, 273–305 (1964).
17. — Asymptotische Abschätzungen in Shannons Informationstheorie. Trans. 3rd Prague Conf. Information Theory Statist. Decision Functions, Random Processes 683–723 (1963).

18. WOLFOWITZ, J.: Simultaneous channels. Arch. rat. Mech. Analysis, **4**, 371—386 (1960).
19. — Coding theorem of information theory. Erg. d. Math. u. ihrer Grenzgebiete. Berlin-Göttingen-Heidelberg: Springer 1961.
20. — On channels without a Capacity. Inform. and Control **6**, 49—54 (1963).
21. ZHANG ZHAO-ZHI (CHANG CHAO-CHIH): Some results obtained with almost-periodic channels. Chinese Math-Acta **6**, Amer. Math. Soc. Translat. of acta mathematica sinica **15**, 127—135 (1965).

RUDOLF AHLWEDE
The Ohio State University
Department of Mathematics
231 W. 18th Avenue,
Columbus, Ohio 43210, USA