Z. Wahrscheinlichkeitstheorie verw. Geb. 6, 10–61 (1966)

Gedächtnisfreie Kanäle für diskrete Zeit *

UDO AUGUSTIN

Eingegangen am 1. Oktober 1965

Einleitung

In dieser Arbeit werden für diskrete Zeit definierte nicht notwendig stationäre gedächtnisfreie Kanäle mit beliebigen Alphabeten behandelt. Das Hauptinteresse gilt den zeitasymptotischen Abschätzungen der maximalen Längen von ε-Codes. Hergeleitet werden diese Abschätzungen aus Gleichgradigkeits- bzw. Momentenbedingungen mit Hilfe der Tschebyscheff- und Jensenschen Ungleichung. Die Untersuchungen bewegen sich stets im Bereich der direkten Anwendbarkeit dieser Ungleichungen. Zunächst wird das Verhalten der maximalen Codelängen unter recht allgemeinen Voraussetzungen beschrieben. Danach werden die Forderungen schrittweise verschärft und dementsprechend genauere Codelängenabschätzungen durchgeführt. Schließlich ergeben sich auf diese Weise notwendige und hinreichende Bedingungen für die Gültigkeit der starken Umkehrung des Coding-Theorems und (unter einer Zusatzvoraussetzung) ebenfalls genaue Bedingungen für die Gültigkeit des Coding-Theorems. Die Bauart dieser letzteren Forderungen ist vergleichbar mit der Bauart der Bedingungen, die beim schwachen Gesetz der großen Zahlen für Doppelschemata unabhängiger Zufallsvariablen auftreten. Abgerundete Aussagen finden sich z. B. in Satz 2.1, Satz 7.2, Corr. 10.2.1, 10.3.1 und Beispiel 11.2.

Für den Kanal wird ein vom üblichen abweichendes einfacheres mathematisches Modell benutzt, das sich nach der Analyse des von JACOBS in [1] verwandten als zweckmäßig ergab: Für einen festen Zeitpunkt ist ein Kanal einfach durch eine Menge M von Wahrscheinlichkeitsverteilungen p auf einem Maßraum (X, B) gegeben. Diese Definition ist sicher dann sinnvoll, wenn die Buchstaben y aus einer Menge Y (Eingangsalphabet), die nach der Menge X (Ausgangsalphabet) übertragen werden sollen, exakt (δ-Verteilung) in das übermittelnde Nachrichtensystem eingegeben werden können. Denn in diesem Fall sind alle Nachrichten, die gleiche Übergangswahrscheinlichkeiten besitzen, vom Ausgang des Kanals her gesehen als ein und dieselbe Nachricht zu bewerten. Unter einem ε-Code ist dann eine Serie von Paaren (p^i, E_i) ($p^i \in M$, $E_i \in B$ paarweise disjunkt) mit $p^i(E_i) > \varepsilon$ zu verstehen.

Gewöhnlich denkt man sich außer dem Eingangsalphabet Y noch einen Borelkörper A in Y gegeben und definiert dann Kanal als stochastischen Kern $P(y, E)$ von (Y, A) nach (X, B). Durch das für $P(.,.)$ gegebene Eingangsquellensystem ist dann gerade festgelegt, mit welcher Genauigkeit die Nachrichten y in den Kanal eingegeben werden können. Vom Ausgang des Kanals her gesehen kann man aber in bezug auf Codierungsfragen die ungenaue Eingabe dem durch den

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

Kern beschriebenen statistischen Rauschen hinzurechnen. Die für Codierungen interessante Menge von Wahrscheinlichkeitsverteilungen ist dann gerade die Menge der von $P(.,.)$ und den Eingangsquellen induzierten Ausgangsverteilungen. Dieser entspricht die Menge M unseres Modells. Alphabete spielen danach also eine völlig untergeordnete Rolle. Gestützt werden diese heuristischen Überlegungen durch die Paragraphen 3 und 6. Der Grund dafür, daß wir nicht mit stochastischen Kernen arbeiten, liegt auch zum Teil darin, daß bei Kernen im Falle unendlicher Alphabete technische Schwierigkeiten (Meßbarkeits- und Separabilitätsfragen) auftreten, die mit dem eigentlichen Codierungsproblem nichts zu tun haben. Schwierigkeiten dieser Art ergeben sich in unserem Modell nicht, das ferner den Vorzug hat, daß die wesentlichen Abschätzungen mit endlichen Summen durchgeführt werden können.

Die Theorie der stationären gedächtnisfreien Kanäle mit endlichen Alphabeten findet man ausführlich bei WOLFOWITZ [2] abgehandelt. Die Abschätzungen der Codelängen werden dort ebenfalls mit Hilfe der Tschebyscheff-Ungleichung durchgeführt. Im allgemeinen werden dabei für die starken Umkehrungen des Coding-Theorems Stationarisierungsverfahren verwandt; wie der Vergleich mit den Beweisen des § 8 zeigt, sind die Abschätzungen aber grundsätzlich ebenso einfach ohne solche Verfahren.

Erstmalig wurden allgemeine nichtstationäre Kanäle (Kernmodell) mit endlichen Alphabeten von Herrn R. AHLSWEDE in seiner Diplomarbeit behandelt. Bei Einschränkung der hier benutzten Verfahren auf den Fall endlicher Alphabete ergeben sich schärfere Abschätzungen als dort.

Für die in dieser Arbeit vielfach benutzten einfachen Aussagen über konvexe Funktionen sei etwa auf KRASNOSELSKII-RUTICKII [3] verwiesen.

Besonders danke ich Herrn Professor Dr. K. JACOBS für die Förderung dieser Arbeit und einige anregende Bemerkungen.

Inhaltsverzeichnis

§ 1. Der Grundraum

Ein Tripel (X, B, M) bedeute im folgenden stets eine Menge X, versehen mit einem Borelkörper (BK) B, und eine nichtleere Menge M von Wahrscheinlichkeitsverteilungen (WVen) p, die auf (X, B) erklärt sind. Wir interpretieren X bzw. (X, B) als System von Ausgangssymbolen bei einer Nachrichtenübertragung und

M als System aller Übergangs-WVen bei der Übertragung von Eingangsnachrichten nach X. Die Eigenschaften der Nachrichtenübertragung sollen für uns die Eigenschaften von M sein. In dieser Auffassung ist es uninteressant, ob M eine Darstellung als $M = \{P(y,.) : y \in Y, P(y,.) \text{ ist WV auf } (X, B)\}$ hat mit einem speziellen System Y von Eingangssymbolen; allein auf die Struktur von M kommt es an. Unsere Aussagen beziehen sich i. a. auf den Fall, daß (X, B, M) „Produktstruktur" hat. Kanäle sind für uns Mengen „verträglicher" Tripel (X, B, M), wobei Verträglichkeit im Sinne einer Fortsetzung der Nachrichten auf größere Zeiträume zu verstehen ist. Dies sei jetzt genau formuliert.

Sei Z die Menge der ganzen Zahlen (die Zeit). Jedem $v \in Z$ sei ein Tripel (X_v, B_v, M_v) zugeordnet. Für jede nichtleere endliche Teilmenge $J \subseteqq Z$ sei

$$X_J := \prod_{v\in J}{}^{\times} X_v \qquad \text{(Produktraum)},$$

$$B_J := {}^{B}\Big(\prod_{v\in J}{}^{\times} B_v\Big) \qquad \text{(Produkt-BK)},$$

$$M_J := \prod_{v\in J}{}^{\times} M_v := \Big\{\prod_{v\in J}{}^{\times} p_v : p_v \in M_v\Big\} \quad \text{(Menge der mittels der } M_v (v \in J) \text{ auf } (X_J, B_J) \text{ erzeugten Produkt-WVen).}$$

Ausgehend von dieser Grundsituation definieren wir:

Definition 1.1 *(gedächtnisfreier und gedächtnisfreier stationärer Kanal).* Sei (X_v, B_v, M_v) für jedes $v \in Z$ vorgegeben. Dann heißt $\{(X_J, B_J, M_J) : \emptyset \neq J \subseteqq Z\}$ gedächtnisfreier Kanal. Der gedächtnisfreie Kanal heißt stationär, wenn alle (X_v, B_v, M_v) Exemplare von (X_0, B_0, M_0) sind.

Später beschäftigen wir uns im wesentlichen nur mit Mengen M_J der Form

$$M_{[1,t]} := M_1 \times \cdots \times M_t$$

für wachsendes t.

Der Begriff ε-Code wird für beliebige Tripel (X, B, M) erklärt.

Definition 1.2 *(ε-Code und $N(M, \varepsilon)$).* Sei (X, B, M) vorgegeben. Eine Serie von höchstens abzählbar vielen Paaren (p^i, E_i) $(p^i \in M, E_i \in B)$ mit paarweise disjunkten E_i heißt Code von M. Erfüllt ein Code $\{(p^i, E_i)\}$ von M zusätzlich die Bedingung

$$p^i(E_i) > \varepsilon$$

für ein ε mit $0 < \varepsilon < 1$, so heißt er ein Code für ε oder ε-Code von M. Das Supremum über die Mächtigkeiten aller für ein festes $\varepsilon (0 < \varepsilon < 1)$ möglichen ε-Codes von M wird mit $N(M, \varepsilon)$ bezeichnet. Statt von der Mächtigkeit eines Code sprechen wir auch von seiner Länge.

Einige einfache Eigenschaften von $N(.,.)$ erhält man sofort:

(a) Es ist $N(M, \varepsilon) \geqq 1$, da für jedes ε und beliebiges $p \in M$ stets $p(X) = 1$ ist.

(b) $N(M,.)$ ist eine monoton fallende Funktion von ε, denn jeder ε_0-Code ist ein ε-Code für alle $\varepsilon \leqq \varepsilon_0$.

(c) $N(M,.)$ ist von rechts her stetig an jeder Stelle ε_0 mit $N(M, \varepsilon_0) < \infty$, da ein Code maximaler Länge für ε_0 noch ein $(\varepsilon_0 + \delta)$-Code ist für ein hinreichend kleines $\delta > 0$.

(d) Bei Einschränkung von M auf einen Unter-BK von B wird $N(.,\varepsilon)$ höchstens kleiner, da dann für die Bildung der Codes weniger Ausgangsmengen zur Konkurrenz zugelassen werden.

(e) Es ist $N(M',\varepsilon) \leqq N(M,\varepsilon)$ für $M' \subseteqq M$.

Die Definition 1.2 schließt zunächst die Möglichkeit nicht aus, daß für ein ε $N(M,\varepsilon) = \infty$ ist, obwohl es nur ε-Codes endlicher Länge gibt. Daß dieses aber in Wahrheit nicht eintreten kann, zeigt das folgende

Lemma 1.1. *Wenn es zu festem ε beliebig lange ε-Codes gibt, so existiert auch ein unendlich langer ε-Code.*

Beweis. Es gebe beliebig lange ε-Codes, und es sei $\{(p^i, E_i)\}$ $(i = 1, \ldots, N)$ ein ε-Code. Zu zeigen ist, daß man diesen ε-Code verlängern kann. Dazu nehme man aus einem passenden anderen ε-Code ein (p'^k, E'_k) mit

$$(1) \qquad p^i(E'_k) < \frac{1}{3N} \left(\inf_{i=1,\ldots,N} p^i(E_i) - \varepsilon\right) = \varepsilon_N \qquad (i = 1, \ldots, N).$$

Dies ist stets möglich, denn anderenfalls gäbe es einen ε-Code

$$\{(\bar{p}^j, \bar{E}_j)\} \qquad \left(j = 1, \ldots, R \geqq \left[\frac{1}{\varepsilon_N} + 2\right] N\right)$$

mit der Eigenschaft, daß für mindestens ein zu $\{(p^i, E_i)\}$ gehöriges p^i

$$p^i\Big(\bigcup_{j=1}^{R} \bar{E}_j\Big) > 1$$

wäre im Widerspruch dazu, daß p^i WV ist. Den vorgegebenen Code $\{(p^i, E_i)\}$ verlängern wir nun um ein (p'^k, E'_k) mit (1), ersetzen aber dabei E_i durch

$$E_i^{(1)} = E_i - E_i \cap E'_k.$$

Durch Iteration bekommt man einen unendlich langen ε-Code $\{(p^i, E_i^0)\}$, wobei

$$E_i^0 = E_i \cap E_i^{(1)} \cap E_i^{(2)} \cap \cdots$$

ist und

$$E_i^{(n)} = E_i^{(n-1)} - E_i^{(n-1)} \cap E'_k,\ E'_k$$

jeweils geeignet im Iterationsprozeß gewählt sind. Denn wegen der rechten Seite der Ungleichung (1) gilt noch

$$p^i(E_i^0) > \varepsilon \quad (\text{nicht nur } \geqq \varepsilon)$$

für $i = 1, 2, \ldots$.

Ein wesentliches Hilfsmittel für eine einfache Beschreibung von Codelängenabschätzungen liefert ein Substitut für Simultanverteilungen. Ohne Benutzung dieses modifizierten Begriffes „Simultanverteilung bezüglich M", der im folgenden eingeführt wird, würden die späteren Abschätzungen erheblich unübersichtlicher.

Definition 1.3. *(Simultanverteilung bezüglich M).* Sei (X, B, M) vorgegeben. Sei $\{p^i\}$ $(i = 1, \ldots, n)$ eine endliche Teilmenge von M und jedem i $(i = 1, \ldots, n)$ eine reelle Zahl $a^i \geqq 0$ zugeordnet mit

$$\sum_{i=1}^{n} a^i = 1.$$

Die auf $(1, \ldots, n) \times X$ durch

$$\tilde{p}(i \times E) := a^i p^i(E) \qquad (E \in B)$$

definierte WV $\tilde{p}$ heißt die von $\{a^i\}$ und $\{p^i\}$ $(i = 1, \ldots, n)$ bestimmte Simultanverteilung bezüglich M. Ihre Randverteilung auf $(1, \ldots, n)$ wird mit $a(a(i) = a^i)$, ihre Randverteilung auf (X, B) mit $q(q = \sum a^i p^i)$ bezeichnet. Die Radon-Nikodymdichte

$$\tilde{f} := \frac{d\tilde{p}}{da \times dq}$$

heißt Simultandichte von $\tilde{p}$. Die Menge aller auf diese Weise mit Hilfe von M gebildeten Simultanverteilungen wird mit $\tilde{M}$ bezeichnet.

Wir merken noch an, daß

$$\tilde{f} \leqq \min_{a^i > 0} \frac{1}{a^i} \quad \tilde{p}\text{-fastüberall bzw. } a \times q\text{-fastüberall}$$

ist.

Von der Aussage des folgenden Satzes machen wir stets bei Abschätzungen von Codelängen nach unten Gebrauch.

Satz 1.2. (*Maximalcode-Satz* (siehe [*2*])). *Für jedes reelle S und jedes $\tilde{p} \in \tilde{M}$ ist*

$$S(\tilde{p}\{\tilde{f} > S\} - \varepsilon) < N(M, \varepsilon) \qquad (0 < \varepsilon < 1).$$

Beweis. Sei $\tilde{p} \in \tilde{M}$ die von $\{a^i\}$ und $\{p^i\}$ $(i = 1, \ldots, n)$ bestimmte Simultanverteilung bezüglich M und S beliebig reell. Ohne Einschränkung kann $S > 0$ und $\tilde{p}\{\tilde{f} > S\} > \varepsilon$ angenommen werden, da andernfalls die behauptete Ungleichung trivialerweise richtig ist. Für mindestens ein $p^j \in \{p^i\}$ ist dann

$$p^j \left\{\frac{dp^j}{dq} > S\right\} > \varepsilon .$$

Sei

$$\{(p^{i_k}, E_{i_k})\} \qquad (k = 1, \ldots, N')$$

ein ε-Code maximaler Länge von $\{p^1, \ldots, p^n\}$ unter der Nebenbedingung

$$E_{i_k} \subseteqq \left(\left\{\frac{dp^{i_k}}{dq} > S\right\} - (E_{i_1} \cup \cdots \cup E_{i_{k-1}})\right) \qquad (k = 2, \ldots, N').$$

Dann gilt

$$p^j\left(\left\{\frac{dp^j}{dq} > S\right\} - (E_{i_1} \cup \cdots \cup E_{i_{N'}})\right) \leqq \varepsilon \qquad (j = 1, \ldots, n),$$

da sich andernfalls der Code verlängern ließe. Also hat man

$$p^j \left\{\frac{dp^j}{dq} > S\right\} - \varepsilon \leqq p^j (E_{i_1} \cup \cdots \cup E_{i_{N'}}) \qquad (j = 1, \ldots, n)$$

und somit

$$\tilde{p}\{\tilde{f} > S\} - \varepsilon = \sum_{j=1}^{n} a^j p^j \left\{\frac{dp^j}{dq} > S\right\} - \varepsilon \leqq$$

$$\leqq \sum_{j=1}^{n} a^j p^j (E_{i_1} \cup \cdots \cup E_{i_{N'}}) = q(E_{i_1} \cup \cdots \cup E_{i_{N'}}).$$

Indem man noch ausnutzt, daß

$$\frac{dp^{i_k}}{dq} > S \quad \text{auf} \quad E_{i_k}$$

ist, gewinnt man hieraus die behauptete Ungleichung:

$$S(\tilde{p}\{\tilde{f} > S\} - \varepsilon) \leqq S\left(\sum_{k=1}^{N'} q(E_{i_k})\right) < \sum_{k=1}^{N'} p^{i_k}(E_{i_k}) \leqq$$
$$\leqq N' \leqq N(\{p^1, \ldots, p^n\}, \varepsilon) \leqq N(M, \varepsilon).$$

Uns interessiert insbesondere die Aussage von Satz 1.2 für Mengen $\tilde{M}$ der Bauart $M_{[1,t]} = M_1 \times \cdots \times M_t$. Sei für jedes v, $1 \leqq v \leqq t$, ein $\tilde{p}_v \in \tilde{M}_v$ vorgegeben; a_v sei die vordere Randverteilung von $\tilde{p}_v$ auf (einer endlichen Menge) A_v, $\tilde{f}_v$ die zu p_v gehörende Simultandichte. Dann ist die auf

$$\prod_{v=1}^{t}{}^{\times} (A_v \times X_v) \quad \text{bzw.} \quad \left(\prod_{v=1}^{t}{}^{\times} A_v\right) \times \left(\prod_{v=1}^{t}{}^{\times} X_v\right)$$

durch

$$\tilde{p}_{[1,t]} := \tilde{p}_1 \times \cdots \times \tilde{p}_t$$

definierte WV $\tilde{p}_{[1,t]}$ offenbar eine Simultanverteilung bezüglich $M_{[1,t]}$ mit zugehöriger Simultandichte

$$\tilde{f}_{[1,t]} := \prod_{v=1}^{t}{}^{\times} \tilde{f}_v .$$

Wir erhalten somit das

Corollar 1.2.1. *Es ist*

$$\tilde{M}_1 \times \cdots \times \tilde{M}_t \subseteqq \tilde{M}_{[1,t]},$$

und für jedes reelle S und jedes $\tilde{p}_1 \times \cdots \times \tilde{p}_t \in \tilde{M}_1 \times \cdots \times \tilde{M}_t$ gilt

$$N(M_{[1,t]}, \varepsilon) > S\left(\left(\prod_{v=1}^{t}{}^{\times} \tilde{p}_v\right)\left\{\prod_{v=1}^{t}{}^{\times} \tilde{f}_v > S\right\} - \varepsilon\right) \qquad (0 < \varepsilon < 1).$$

Wenn wir später die Aussage des Maximalcode-Satzes auf $M_{[1,t]}$ anwenden, so tun wir es stets in dieser letztgenannten Form.

§ 2. Unendlichlange ε-Codes

Sei für $M_v (v = 1, \ldots, t)$ bekannt, zu welchen ε es unendlichlange Codes gibt. Der folgende Satz 2.1 beantwortet die Frage, für welche ε dann $M_{[1,t]}$ unendlichlange Codes besitzt. Insbesondere liefert er, daß bei stationären gedächtnisfreien Kanälen, die unendlichlange ε-Codes haben, die zeitasymptotischen Abschätzungen der maximalen Codelängen für jedes feste ε trivial werden (sogar noch für wesentlich umfangreichere Klassen von Kanälen). Der im Beweis des Satzes beschriebene Aufschaukelungsprozeß, durch den man aus langen schlechten Codes ($\varepsilon \approx 0$) mit der Zeit lange gute Codes ($\varepsilon \approx 1$) erhält, ist für stationäre Kanäle unmittelbar plausibel: Hat man ein am Ausgang des Kanals nur mit geringer Sicherheit unterscheidbares Stichwortsystem, so läßt sich die Übertragungssicherheit dadurch verbessern, daß man jedes zu sendende Wort gleich mehrfach hintereinander eingibt, was natürlich auf Kosten der Zeit geht. Weiter geht aus dem Beweis hervor, daß selbst im Falle der Existenz nur endlicher ε-Codes die schlechten Codes stets zu beachten sind.

Setzen wir für jedes gegebene (X, B, M)

$$u(M) := \begin{cases} 0, & \text{wenn } N(M,\varepsilon) < \infty \quad \text{ist für jedes} \quad \varepsilon \in (0,1), \\ \sup\{\varepsilon : N(M,\varepsilon) = \infty\} & \text{sonst}, \end{cases}$$

so gilt der

Satz 2.1. *(Aufschaukelung unendlichlanger Codes). Es ist*

$$1 - u(M_{[1,t]}) = \prod_{v=1}^{t} (1 - u(M_v)). \tag{2}$$

Beweis. Wir weisen nach, daß die beiden folgenden Aussagen richtig sind, die zusammen offensichtlich zu (2) äquivalent sind:

(3) Für $\quad 0 < \varepsilon < 1 - \prod_{v=1}^{t} (1 - u(M_v)) \quad$ ist $\quad N(M_{[1,t]}, \varepsilon) = \infty.$

(4) Für $\quad 1 - \prod_{v=1}^{t} (1 - u(M_v)) < \varepsilon < 1 \quad$ ist $\quad N(M_{[1,t]}, \varepsilon) < \infty.$

1. Beweis von (3): Sei ohne Einschränkung

$$\{\varepsilon : 0 < \varepsilon < 1 - \prod_{v=1}^{t} (1 - u(M_v))\} \neq \emptyset,$$

dann ist

$$J := \{v : 1 \leqq v \leqq t,\ u(M_v) > 0\} \neq \emptyset.$$

Wegen

$$N(M_{[1,t]}, \varepsilon) \geqq N(M_J, \varepsilon)$$

können wir weiter annehmen, daß

$$u(M_v) > 0 \qquad (v = 1, \ldots, t)$$

ist. Sei dann für ε_1 $N(M_1, \varepsilon_1) = \infty$ und $\{(\bar{p}_1^i, E_i)\}$ ein ε_1-Code der Länge $N_1 < \infty$ für M_1. Setzen wir

$$p_1^i = (1-b)\,\bar{p}_1^i + b\,q_1 \ (0 < b < 1) \quad \text{mit} \quad q_1 := \frac{1}{N_1} \sum_{i=1}^{N_1} \bar{p}_1^i,$$

so ist $\{(p_1^i, E_i)\}$ ein $(1-b)\,\varepsilon_1$-Code der Länge N_1 (für die konvexe Hülle von M_1). Man betrachte $\{p_1^i\}$ und q_1 in bezug auf den von $\{E_i\}$ erzeugten BK $B_1(\{E_i\}) \subseteqq B_1$, desgleichen die mit $\{p_1^i\}$ und der vorderen Randverteilung $1/N_1$ auf jedem i gebildete Simultanverteilung $\tilde{p}_1$. Nach der Tschebyscheff-Ungleichung hat man

$$\frac{1}{N_1} \cdot \left| \left\{ i : q_1(E_i) > \frac{1}{d \cdot N_1} \sum_{k=1}^{N_1} q_1(E_k) \right\} \right| < d \quad (d > 0),$$

also

$$\frac{1}{N_1} \cdot \left| \left\{ i : \frac{1}{q_1(E_i)} > d \cdot N_1 \right\} \right| > 1 - d$$

und

$$\frac{1}{N_1}\cdot\left|\left\{i:\frac{p_1^i(E_1)}{q_1(E_i)}>d(1-b)\,\varepsilon_1 N_1\right\}\right|>1-d.$$

Folglich ist

$$\frac{1}{N_1}\sum_{i=1}^{N_1} p_1^i\left\{\frac{dp_1^i}{dq_1}>d(1-b)\,\varepsilon_1 N_1\right\}>(1-d)\,(1-b)\,\varepsilon_1\,,$$

das heißt

$$\tilde{p}_1\{\tilde{f}_1>d(1-b)\,\varepsilon_1 N_1\}>(1-d)\,(1-b)\,\varepsilon_1\,.$$

Man setze nun $d=b$ und behandele die anderen $M_v (v=2,\ldots,t)$ entsprechend wie M_1. Dann erhält man für $\tilde{p}_v$ (das in bezug auf M_v ebenso konstruiert ist wie $\tilde{p}_1$ in bezug auf M_1)

$$\tilde{p}_v\{\tilde{f}_v>b(1-b)\,\varepsilon_v N_v\}>(1-b)^2\,\varepsilon_v \qquad (v=1,\ldots,t)\,.$$

Zu vorgegebenem $K>0$ sei nun N_v so groß gewählt, daß

$$b\,\varepsilon_v N_v>K \qquad (v=1,\ldots,t)$$

ist. Beachtet man noch

$$\tilde{f}_v\geqq b \qquad (v=1,\ldots,t)\,,$$

so gilt ersichtlich

$$\tilde{p}_1\times\cdots\times\tilde{p}_t\left\{b^{-t}\prod_{v=1}^{t}\tilde{f}_v>(1-b)\,K\right\}\geqq\tilde{p}_1\times\cdots\times\tilde{p}_t\{\sup_v(b^{-1}\tilde{f}_v)>(1-b)\,K\}=$$
$$=\tilde{p}_1\times\cdots\times\tilde{p}_t\{\sup_v(\tilde{f}_v)>b(1-b)\,K\}>1-\prod_{v=1}^{t}(1-(1-b)^2\,\varepsilon_v)\,.$$

Da bei Einschränkung auf Unter-BK die Codelängen höchstens kleiner werden, können wir auf die linke Seite der letzten Ungleichung Corollar 1.2.1 anwenden: Sei M' eine Menge von WVen über $B_{[1,t]}$, in bezug auf die $\tilde{p}_1\times\cdots\times\tilde{p}_t$ Simultanverteilung ist. Dann gilt

$$N(M',\varepsilon')>b^t(1-b)\,K\,(1-\prod_{v=1}^{t}(1-(1-b)^2\,\varepsilon_v)-\varepsilon')\,.$$

Nun sind aber die WVen auf $B_{[1,t]}$, mit denen $\tilde{p}_1\times\cdots\times\tilde{p}_t$ erzeugt wurde, konvexe Kombinationen von WVen aus $M_{[1,t]}$. Daher ist

$$N(M_{[1,t]},\varepsilon')>b^t(1-b)\,K\left(1-\prod_{v=1}^{t}(1-(1-b)^2\,\varepsilon_v)-\varepsilon'\right),$$

denn für konvexe Kombination von WVen gilt:

(5) $$\text{Aus}\left(\sum_{j=1}^{r}c^j p^j\right)(E)>\varepsilon' \qquad \left(0\leqq c^j\leqq 1\,,\ \sum_{j=1}^{r}c^j=1\right)$$
$$\text{folgt}\quad p^j(E)>\varepsilon'\ \text{ für mindestens ein } j\,(1\leqq j\leqq r)\,.$$

Läßt man nun $K\to\infty$ gehen (Übergang zu neuen Codes) und $b(K)$ mit wachsendem K sehr langsam gegen Null, so ergibt sich aus der letzten Abschätzung, daß $M_{[1,t]}$ für ε' mit $0<\varepsilon'<1-\prod_{v=1}^{t}(1-\varepsilon_v)$ beliebig lange Codes hat. Hieraus ergibt sich die Behauptung (3), wenn man noch von den ε_v zu den $u(M_v)$ übergeht.

2. Beweis von (4): Wir beweisen (4) für $t = 2$. Allgemein ergibt sich dann (4) für beliebiges $t > 2$ durch Induktion. Außer dem Maximalcode-Satz benutzen wir einen typischen Schluß für obere Abschätzungen der Codelängen. Sei ohne Einschränkung

$$\left\{\varepsilon : 1 - \prod_{v=1}^{2}(1 - u(M_v)) < \varepsilon < 1\right\} \neq \emptyset,$$

also

$$u(M_v) < 1 \qquad (v = 1, 2).$$

Sei dann aus dieser nichtleeren Menge ein ε vorgegeben, und sei $\{(p^i, E_i)\}$ ein ε-Code der Länge $N < \infty$ für $M_1 \times M_2$. Wir zeigen, daß N nicht beliebig groß sein kann. Mit

$$p^i = : p_1^i \times p_2^i \qquad (i = 1, \ldots, N)$$

und

$$q := q_1 \times q_2 := \left(\frac{1}{N}\sum_{i=1}^{N} p_1^i\right) \times \left(\frac{1}{N}\sum_{i=1}^{N} p_2^i\right)$$

gilt für $S_1, S_2 > 0$

$$\begin{aligned} p^i\left\{\frac{dp^i}{dq} > S_1 S_2\right\} &= p^i\left\{\left(\frac{1}{S_1}\cdot\frac{dp_1^i}{dq_1}\right)\left(\frac{1}{S_2}\cdot\frac{dp_2^i}{dq_2}\right) > 1\right\} \leqq \\ &\leqq p^i\left\{\sup\left(\frac{1}{S_1}\cdot\frac{dp_1^i}{dq_1}, \frac{1}{S_2}\cdot\frac{dp_2^i}{dq_2}\right) > 1\right\} \\ &= p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1\right\} + p_2^i\left\{\frac{dp_2^i}{dq_2} > S_2\right\} - p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1\right\}\cdot p_2^i\left\{\frac{dp_2^i}{dq_2} > S_2\right\} = \\ &= 1 - \prod_{v=1}^{2}\left(1 - p_v^i\left\{\frac{dp_v^i}{dq_v} > S_v\right\}\right). \end{aligned}$$

Man setze

$$\varepsilon_v := u(M_v) \qquad (v = 1, 2).$$

Da nach Satz 1.2

$$\infty > N(M_1, \varepsilon_1 + b) > S_1\left(\frac{1}{N}\sum_{i=1}^{N} p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1\right\} - (\varepsilon_1 + b)\right)$$

gilt und somit für hinreichend großes S_1^0

$$\frac{1}{N}\sum_{i=1}^{N} p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1^0\right\} < \varepsilon_1 + 2b$$

ist, hat man nach der Tschebyscheff-Ungleichung

$$\frac{1}{N}\left|\left\{i : p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1^0\right\} > (\varepsilon_1 + 2b)(1 + b)\right\}\right| < \frac{1}{1+b},$$

also

$$\frac{1}{N}\left|F_1 := \left\{i : p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1^0\right\} < (\varepsilon_1 + 2b)(1 + b)\right\}\right| > 1 - \frac{1}{1+b} = \frac{b}{1+b}.$$

Setzen wir

$$\bar{q}_2 := \frac{1}{|F_1|}\sum_{i \in F_1} p_2^i,$$

so ist für $i \in F_1$

$$p_2^i\left\{\frac{dp_2^i}{dq_2} > S_2\right\} = p_2^i\left\{\frac{dp_2^i}{d\bar{q}_2} \cdot \frac{d\bar{q}_2}{dq_2} > S_2\right\} \leqq p_2^i\left\{\frac{dp_2^i}{d\bar{q}_2} > \frac{|F_1|}{N} S_2\right\}.$$

Wegen

$$\infty > N(M_2, \varepsilon_2 + b) > \frac{|F_1|}{N} S_2 \left(\frac{1}{|F_1|} \sum_{i \in F_1} p_2^i\left\{\frac{dp_2^i}{d\bar{q}_2} > \frac{|F_1|}{N} S_2\right\} - (\varepsilon_2 + b)\right)$$

gilt für hinreichend großes S_2^0

$$\frac{1}{|F_1|}\left|F_2 := \left\{i \in F_1 : p_2^i\left\{\frac{dp_2^i}{d\bar{q}_2} > \frac{|F_1|}{N} S_2^0\right\} < (\varepsilon_2 + 2b)(1 + b)\right\}\right| > \frac{b}{1+b}.$$

Folglich haben wir

$$\frac{1}{N}\left|\left\{i : p_1^i\left\{\frac{dp_1^i}{dq_1} > S_1^0\right\} < (\varepsilon_1 + 2b)(1 + b),\ p_2^i\left\{\frac{dp_2^i}{dq_2} > S_2^0\right\} < \right.\right.$$

$$\left.\left. < (\varepsilon_2 + 2b)(1 + b)\right\}\right| > \frac{b^2}{(1 + b)^2},$$

d. h. es ist

$$\frac{|F_2|}{N} > \frac{b^2}{(1 + b)^2}$$

und

$$p^i\left\{\frac{dp^i}{dq} > S_1^0 \cdot S_2^0\right\} \leqq 1 - (1 - (\varepsilon_1 + 2b)(1 + \mathrm{b}))(1 - (\varepsilon_2 + 2b)(1 + b)) \quad (i \in F_2).$$

Sei nun b von vornherein als so klein vorausgesetzt, daß die rechte Seite dieser Ungleichung $< \varepsilon - d$ wird mit passendem $d > 0$. Dann gilt

$$p^i\left\{\frac{dp^i}{dq} < S_1^0\, S_2^0\right\} > 1 - (\varepsilon - d) \qquad (i \in F_2),$$

also

$$p^i\left(E_i \cap \left\{\frac{dp^i}{dq} < S_1^0\, S_2^0\right\}\right) > d \qquad (i \in F_2).$$

Hieraus erhält man

$$q(E_i) \geqq q\left(E_i \cap \left\{\frac{dp^i}{dq} < S_1^0\, S_2^0\right\}\right) > \frac{1}{S_1^0 S_2^0}\, p^i\left(E_i \cap \left\{\frac{dp^i}{dq} < S_1^0\, S_2^0\right\}\right) > \frac{1}{S_1^0 S_2^0}\, d,$$

was die Ungleichungen

$$1 \geqq \sum_{i \in F_2} q(E_i) \geqq |F_2| \frac{d}{S_1^0 S_2^0} > \frac{b^2}{(1 + b)^2} \cdot N \cdot \frac{d}{S_1^0 S_2^0}$$

und

$$N < \frac{(1 + b)^2}{b \cdot d} \cdot S_1^0 \cdot S_2^0$$

nach sich zieht. Da $S_v^0 (v = 1, 2)$ nach Satz 1.2 nicht vom Code, sondern nur von $N(M_v, \varepsilon_v + 2b)$ und $\varepsilon_v + 2b$ abhängt, ist hiermit (4) für $t = 2$ bewiesen.

§ 3. Endliche Codes und Kompaktheit

Auf Grund der Tatsache, daß sich Kanäle mit unendlichlangen ε-Codes offenbar erheblich von den praktisch vorkommenden unterscheiden, und auf Grund des Resultats im vorigen Paragraphen ist es vernünftig, unsere Untersuchung von jetzt ab auf Kanäle mit nur endlichen ε-Codes einzuschränken. Wir werden zunächst die Mengen M mit $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$ topologisch charakterisieren.

Definition 3.1. ($L^{glgr} \ll \lambda$) Eine Menge L von Ladungsverteilungen auf (X, B) heißt gleichgradig totalstetig bezüglich der WV λ auf (X, B) (Bezeichnung: $L^{\mathrm{glgr}} \ll \lambda$), wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert derart, daß aus $\lambda(E) < \delta$ stets $|l(E)| < \varepsilon$ (gleichmäßig für $l \in L$) folgt.

Satz 3.1. *Es ist* $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$ *genau dann, wenn* $M^{glgr} \ll \lambda$ *ist für eine WV* λ *auf* (X, B).

Beweis. 1. Sei $N(M, \varepsilon)$ nicht endlich für ein $\varepsilon > 0$. Nach Lemma 1.1 gibt es dann einen unendlichlangen Code $\{(p^i, E_i)\}$ mit $p^i(E_i) > \varepsilon$. Hieraus folgt, da

$$\lim_i \lambda(E_i) = 0 \qquad (\lambda \text{ WV auf } B)$$

gilt, daß M bezüglich keiner WV auf B gleichgradig totalstetig ist.

2. Es existiere keine WV λ auf B mit $M^{\mathrm{glgr}} \ll \lambda$. Zu einem hinreichend kleinen $\varepsilon > 0$ gibt es dann eine Folge $\{p^i\}$ von WVen und eine monoton fallende Folge $\{E_i\}$ von Elementen aus B mit $\bigcap_{i=1}^{\infty} E_i = \emptyset$ derart, daß $p^i(E_i) > \varepsilon$ ist für jedes i. Für eine geeignete Teilfolge $\{p^{i_k}\}$ von $\{p^i\}$ kann man $p^{i_k}(\bar{E}^{i_k}) > \varepsilon/2$ erreichen, wobei $\bar{E}_{i_k} := E_{i_k} \cap \operatorname{compl}(E_{i_{k+1}})$ bedeutet. Es ist $\bar{E}_{i_k} \cap \bar{E}_{i_j} = \emptyset$ für $i_k \neq i_j$ und somit $\{(p^{i_k}, \bar{E}_{i_k})\}$ ein unendlichlanger $\varepsilon/2$-Code.

Für Ladungsverteilungen gilt die folgende Verallgemeinerung des Satzes von Arzelà (vgl. DUNFORD-SCHWARTZ [*4*]): Die in der schwachen Topologie des Banachverbands der Ladungsverteilungen auf B bedingt folgenkompakten Mengen sind gerade die normbeschränkten gleichgradig totalstetigen Mengen. Andererseits sind nach dem Satz von Eberlein für normbeschränkte Mengen in Banachräumen die Begriffe bedingt folgenkompakt und bedingt filterkompakt bezüglich der schwachen Topologie gleichbedeutend. Wir erhalten daher unter Verwendung von Satz 3.1 eine weitere Charakterisierung der Mengen M mit $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$:

Satz 3.2. *Es ist* $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$ *genau dann, wenn* M *bedingt kompakt ist in der schwachen Topologie des Banachverbandes der Ladungsverteilungen auf* (X, B).

Der Vollständigkeit halber merken wir hier noch an, daß sich die Eigenschaft bedingt schwachkompakt von M_1 und M_2 auf $M_1 \times M_2$ vererbt. Satz 3.2 und Satz 2.1 zusammen liefern z. B. einen Beweis dafür.

Zum Schluß dieses Paragraphen weisen wir noch einige Konvexitäts- und Stetigkeitseigenschaften von $N(.,.)$ nach. Für eine Menge M' von WVen bezeichne $\bar{M}'$ ihren schwachen Abschluß, $\operatorname{co}(M')$ ihre konvexe Hülle und im Falle, daß M' konvex schwachkompakt ist, $\operatorname{ex}(M')$ die Menge ihrer Extremalpunkte.

Lemma 3.3. *Es gilt für* $0 < \varepsilon < 1$

(a) $N(M, \varepsilon) = N(\bar{M}, \varepsilon)$,

(b) $N(M, \varepsilon) = N(\operatorname{co}(M), \varepsilon)$,

(c) $N(\overline{\operatorname{co}(M)}, \varepsilon) = N(\operatorname{ex}(\overline{\operatorname{co}(M)}), \varepsilon)$ *für bedingt schwachkompaktes* M.

Beweis. 1. Beweis von (a): Zu $\bar{p} \in \bar{M}$ mit $\bar{p}(E) > \varepsilon$ existiert ein schwachbenachbartes $p \in M$ mit $p(E) > \varepsilon$. Man ersetze im Code $(\bar{p}, E)$ durch (p, E). Dann ersieht man, daß $N(\bar{M}, \varepsilon) \leqq N(M, \varepsilon)$ ist. Andererseits ist $N(\bar{M}, \varepsilon) \geqq N(M, \varepsilon)$. Folglich gilt (a).

2. Die Behauptung (b) folgt mit (5).

3. Beweis von (c): Für bedingt schwachkompaktes M ist $\overline{\mathrm{co}(M)}$ konvex und kompakt. Nach dem Satz von Krein-Milman gilt dann

$$\overline{\mathrm{co}(M)} = \overline{\mathrm{co}(\mathrm{ex}(\overline{\mathrm{co}(M)}))},$$

woraus unter Benutzung von (a) und (b) die Aussage (c) folgt.

§ 4. Gleichgradigkeitsfunktionen

Satz 4.1 liefert weitere Charakterisierungen der Mengen M mit $M^{\mathrm{glgr}} \ll \lambda$ für eine WV λ. Wir beweisen den Satz ausführlich (obwohl es sich dabei um Standardverfahren handelt und obwohl wir nicht mehr direkt von den Aussagen des Satzes Gebrauch machen werden), da die Art der Schlüsse repräsentativ ist für die Schlußweisen bei einigen Beweisen in den folgenden Paragraphen. Später — vor allem in § 9 — begnügen wir uns dann an den entsprechenden Stellen mit einer Bemerkung. Die Bedingung (e) aus Satz 4.1 ist interessant für die Verwendung von Lemma 8.5, das zu verschärften Codelängenabschätzungen für $M_{[1,t]}$ führt. Die Äquivalenz der Bedingungen (a) bis (d) aus Satz 4.1 ist bekannt als Satz von de la Vallée Poussin.

Zur Vereinfachung der Formulierungen führen wir den Begriff der G-Funktion ein.

Definition 4.1. *(G-Funktion).* Unter einer G-Funktion verstehen wir eine Funktion g mit den folgenden Eigenschaften:

(a) g ist auf R^+ definiert und dort konvex, und es ist

$$g(0) = 0, \; g(y) > 0 \;\; (y > 0).$$

(b) Es gilt

$$\frac{g(y)}{y} \to \infty \quad (y \to \infty).$$

Die Funktionen mit (a) besitzen konkave Umkehrfunktionen g^{-1} auf R^+ und lassen sich gerade darstellen in der Form

$$g(y) = \int_0^y dz\, m(z),$$

wobei m eine auf R^+ definierte monoton steigende, für $z > 0$ strikt positive Funktion ist. Eine G-Funktion erfüllt wegen (b) zusätzlich

$$m(y) \to \infty \qquad (y \to \infty).$$

Satz 4.1. *Sei $\{f\}$ eine Familie von Funktionen auf (X, B), λ eine* WV *auf (X, B) und $\{\int_X d\lambda\, |f|\}$ gleichmäßig beschränkt für $f \in \{f\}$. Dann sind die folgenden Aussagen äquivalent:*

(a) *$\{f\}$ ist gleichgradig λ-integrabel, d. h., zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ derart, daß aus*

$$\lambda(E) < \delta \quad \textit{folgt:} \quad \int_E d\lambda |f| < \varepsilon \quad (\textit{glm. für } f \in \{f\}).$$

(b) *Zu jedem $\varepsilon > 0$ existiert eine Konstante $K > 0$ mit*

$$\int_{\{|f| > K\}} d\lambda |f| < \varepsilon \qquad (\textit{glm. für } f \in \{f\}).$$

(c) *Es gibt eine auf R^+ definierte monotone, für $y > 0$ strikt positive Funktion m mit $m(y) \to \infty$ $(y \to \infty)$ und eine Konstante $A > 0$ mit*

$$\int_X d\lambda |f| \, m(|f|) < A \qquad (f \in \{f\}).$$

(d) *Es existiert eine G-Funktion g und eine Konstante $D > 0$ mit*

$$\int_X d\lambda \, g(|f|) < D \qquad (f \in \{f\}).$$

(e) *Es existiert eine G-Funktion h, für die $h(\sqrt{.})$ konkav ist, und eine Konstante $D' > 0$ mit*

$$\int_X d\lambda \, h(|f|) < D' \qquad (f \in \{f\}).$$

Beweis. 1. Aus (a) folgt (b): Wäre (b) unter der Voraussetzung (a) nicht richtig, so gäbe es ein $\varepsilon > 0$ und für beliebig großes k noch ein $f_k \in \{f\}$ mit

$$\int_{\{|f_k| > k\}} d\lambda |f_k| > \varepsilon .$$

Wegen der Beschränktheit von $\{\int d\lambda |f|\}$ erhielte man daraus

$$\lambda\{|f_k| > k\} \xrightarrow{k \to \infty} 0$$

im Widerspruch zu (a).

2. Aus (b) folgt (c): Sei $\{\varepsilon_k\}$ eine monoton fallende Nullfolge mit

$$\sum_{k=1}^{\infty} k \, \varepsilon_k < \infty .$$

Wegen (b) kann man zu jedem ε_k ein $c_k > 0$ so wählen, daß

$$\int_{\{|f| > c_k\}} d\lambda |f| < \varepsilon_k \qquad (f \in \{f\})$$

ist und $\{c_k\}$ eine streng monotone Folge mit $c_k \to \infty$ $(k \to \infty)$ ist. Die durch

$$m(\mathrm{y}) := \begin{cases} 1 \;, & 0 \leqq y \leqq c_1 \\ k+1 \,, & c_k < y \leqq c_{k+1} \end{cases} \qquad (k \geqq 1)$$

auf R^+ definierte Funktion m ist monoton, strikt positiv für $y > 0$ und erfüllt $m(y) \to \infty$ $(y \to \infty)$. Es ist

$$\int_X d\lambda |f| \, m(|f|) \leqq \text{const} + \sum_{k=1}^{\infty} k \, \varepsilon_k \qquad (f \in \{f\}),$$

womit die Existenz der Konstanten A aus (c) gesichert ist.

3. Aus (c) folgt (a): Es gelte (c). Für $K > 0$ und gegebenes f sei

$$F := \{|f| < K\}, \quad \bar{F} := \{|f| \geqq K\}.$$

Dann gilt für $E \in B$

$$\int_E d\lambda |f| = \int_{E\cap F} d\lambda\, |f| + \int_{E\cap\bar{F}} d\lambda |f| \leqq K\,\lambda(E \cap F) + \frac{A}{m(K)}.$$

Man ersetze K durch $1/\sqrt{\lambda(E)}$ und wähle δ zu ε so klein, daß $\lambda(E) < \delta$ stets

$$\sqrt{\lambda(E)} + A \cdot \frac{1}{m(1/\sqrt{\lambda(E)})} < \varepsilon$$

zur Folge hat.

4. Aus (c) folgt (d): Die mit der Funktion m aus (c) durch

$$g(y) := \int_0^y m(z)\, dz$$

auf R^+ definierte Funktion g ist eine G-Funktion. Wegen $g(y) \leqq y \cdot m(y)$ wird (d) mit diesem g und $D = A$ erfüllt.

5. Aus (d) bzw. (e) folgt (c): Da $\frac{g'(y)}{y}$ für eine G-Funktion g' monoton wachsend und für $y > 0$ strikt positiv ist und $\frac{g'(y)}{y} \to \infty$ $(y \to \infty)$ erfüllt, folgt (c) aus (d) mit $m(y) := \frac{g(y)}{y}$, $A = D$ und aus (e) mit $m(y) := \frac{h(y)}{y}$, $A = D'$.

6. Aus (c) folgt (e): Wir zeigen, daß man beim Übergang von (b) nach (c) m stets als konkave Funktion konstruieren kann. Mit den in 2. eingeführten c_k und $c_0 = 0$ setzen wir für $k \geqq 0$

$$\bar{m}(y) := \inf_{0 \leqq i \leqq k} \left(\frac{1}{c_{i+1} - c_i}(y - c_i) + c_i\right) \qquad (c_k \leqq y \leqq c_{k+1}).$$

$\bar{m}$ ist (stückweise linear und) konkav auf R^+, liegt unterhalb der in 2. definierten Funktion m und erfüllt $\bar{m}(y) \to \infty$ $(y \to \infty)$. Für die durch

$$h(y) := \int_0^y dz\, \bar{m}(z)$$

definierte G-Funktion $h(.)$ ist dann $h(\sqrt{.})$ konkav, denn es ist

$$h(\sqrt{y}) = \int_0^{\sqrt{y}} dz\, \bar{m}(z) = \int_0^y d\sqrt{z}\, \bar{m}(\sqrt{z}) = \tfrac{1}{2}\int_0^y dz\, \frac{\bar{m}(\sqrt{z})}{\sqrt{z}};$$

und $\bar{m}(\sqrt{z})/\sqrt{z}$ bzw. $\bar{m}(z)/z$ ist dabei monoton fallend, da $\bar{m}$ konkav ist. (e) wird mit h und $D' = A$ erfüllt.

§ 5. Separable Darstellung des Codierungsproblems über dem Intervall [0,1]

Für die weiteren Betrachtungen spielen die folgenden Bemerkungen keine Rolle. Da alle Abschätzungen der Codelängen im Prinzip mit Ungleichungen arbeiten, für die Separabilitätsbedingungen keine Bedeutung haben, werden wir auch auf diesen Paragraphen nicht wieder zurückgreifen. Der Grund dafür, daß wir dennoch nicht auf die Einfügung dieses Paragraphen verzichtet haben, liegt darin, daß es bei der Konstruktion repräsentativer Beispiele von Nutzen sein kann zu wissen, daß die sämtlichen möglichen Fälle des Verhaltens der Codelängen durch ein einfaches, konkreteres Modell erfaßt werden können. Es besteht nämlich der folgende Sachverhalt:

Zu jedem von einer Serie $M_v (v \in Z)$ erzeugten gedächtnisfreien Kanal kann man einen in bezug auf sämtliche Nachrichtenübertragungseigenschaften äquivalenten Kanal angeben, der von einer Serie M'_v mit $M'_v \subseteqq L^1_{p_v}([0,1])$ erzeugt wird, wobei p_v WV auf ([0,1], natürlicher BK) ist $(v \in Z)$.

Ausgehend von einem durch $\{M_v : v \in Z\}$ gegebenen gedächtnisfreien Kanal zeigen wir zunächst, daß die M_v durch abzählbare Teilmengen $M'_v \subseteqq M_v$ so ersetzt werden können, daß der von $\{M'_v : v \in Z\}$ erzeugte Kanal dieselben maximalen Codelängen hat wie der von $\{M_v : v \in Z\}$ erzeugte. Die Konstruktion der M'_v kann folgendermaßen durchgeführt werden: Zu festem $J, \emptyset \neq J$ (endlich) $\subseteqq Z$, wählen wir eine Folge $\{\varepsilon_k\}$, die die rationalen Zahlen in $(0,1)$ und die Sprungstellen von $N(M_J, .)$ durchläuft. Jedem ε_k werde genau ein ε_k-Code von M_J maximaler Länge zugeordnet. Die Menge der dadurch erfaßten (auf B_J erklärten) WVen ist abzählbar. Für $v \in J$ bezeichne $P_{J,v}$ die Menge der (auf B_v erklärten) v-ten Komponenten dieser WVen. $P_{J,v}$ ist abzählbar, und $\prod\limits_{v \in J}{}^{\times} P_{J,v}$ hat dieselben maximalen Codelängen wie M_J. Wir lassen nun J die sämtlichen endlichen Teilmengen von Z durchlaufen. Wegen der Abzählbarkeit von $\{J : \emptyset \neq J \text{ (endl.)} \subseteqq Z\}$ und der von $P_{J,v}$ ist auch

$$M'_v := \bigcup_{\{J : v \in J \text{(endl.)} \subseteqq Z\}} P_{J,v}$$

abzählbar, und weiter hat $\prod\limits_{v \in J}{}^{\times} M'_v$ für jedes endliche $J \subseteqq Z$ dieselben maximalen Codelängen wie M_J. Bestimmt $\{M_v : v \in Z\}$ einen stationären Kanal, so läßt sich offenbar durch weiteres Hinzufügen geeigneter abzählbar vieler WVen aus M_v zu M'_v erreichen, daß $\{M'_v : v \in Z\}$ einen stationären Kanal definiert. Sei nun

$$M'_v =: \{p_v^1, p_v^2, \ldots\}, \quad p'_v := \sum_{i=1}^{\infty} \frac{1}{2^i} p_v^i .$$

Wegen $p_v^i \ll p'_v$ ist

$$M'_v \subseteqq L^1_{p'_v}(X_v, B_v) .$$

Da M'_v abzählbar ist, werden aber alle topologischen, linearen und Verbandsrelationen in M'_v schon beschrieben über einem abzählbar erzeugten Unter-BK $B'_v \subseteqq B_v$ (Treppenfunktionenkonstruktionen in $L^1_{p'_v}(X_v, B_v)$). Sei dann

$$p'_v = \hat{p}_v + \bar{p}_v$$

die Zerlegung von p'_v in atomfreien Anteil $\hat{p}_v$ und atomaren Anteil $\bar{p}_v$. Durch einen Maßalgebrenisomorphismus läßt sich $(X_v, B'_v, \hat{p}_v)$ abbilden auf ([0,1], natürlicher BK, $||\hat{p}_v|| \cdot \lambda$) ($\lambda$ das Lebesguemaß), und $\bar{p}_v$ kann abgebildet werden auf eine Massenbelegung der Punkte $1/n$ in [0,1]. Hiermit ist sogar etwas mehr als die Behauptung von oben bewiesen.

§ 6. Eingeschränktes Wachstum von $N(M, .)$

Wir untersuchen nun die Beziehungen zwischen dem Wachstum von $N(M, .)$ und dem Verhalten der zu M gehörigen Simultanverteilungen bzw. Simultandichten. Zum einen sind wir an dem Zusammenhang zwischen dem Wachstum von $N(M, .)$ und der Existenz der Suprema gewisser Integrale der Simultandichten interessiert und an direkten Abschätzungen von $N(M, .)$ durch diese Suprema, zum anderen daran, das Wachstum durch Ungleichungen, wie sie im Maximalcode-Satz auftreten, zu erfassen. Beides wird für die späteren asymptotischen Abschätzungen der maximalen Codelängen benötigt.

Den Anschluß an die Paragraphen 3 und 4 liefert das folgende Lemma:

Lemma 6.1. *Genau dann ist* $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$, *wenn eine Konstante* $K > 0$ *existiert und eine monotone Funktion* m *auf* $(0, \infty)$ *mit* $m(S) \to \infty$ $(S \to \infty)$, *so daß*

$$m(S)\, \tilde{p}\{\tilde{f} > S\} < K \qquad (\tilde{p} \in \tilde{M},\ S > 0) \tag{6}$$

gilt.

Beweis. 1. Sei $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$. Man setze

$$h(S') := \begin{cases} N\left(M, \frac{1}{S'}\right) & \text{für} \quad S' > 1 \\ 1 & \text{für} \quad 0 < S' \leqq 1 \end{cases},$$

nach dem Maximalcode-Satz hat man dann

$$h(S') > S'\, h(S') \left(\tilde{p}\{\tilde{f} > S'\, h(S')\} - \frac{1}{S'}\right) \qquad (S' > 1),$$

und somit

$$2 > S'\, \tilde{p}\{\tilde{f} > S'\, h(S')\},$$

was auch noch für $0 < S' \leqq 1$ richtig ist. Man definiere nun die Funktion m auf $(0, \infty)$ durch

$$m(S) := \sup\{S' : S'\, h(S') < S\} \qquad (S > 0).$$

m ist monoton, es gilt $m(S) \to \infty$ $(S \to \infty)$, und wegen

$$m(S)\, \tilde{p}\{\tilde{f} > S\} < 2 \qquad (\tilde{p} \in \tilde{M},\ S > 0)$$

wird (6) mit diesem m und $K = 2$ erfüllt.

2. Sei nun umgekehrt für eine monotone Funktion m auf $(0, \infty)$ mit $m(S) \to \infty$ $(S \to \infty)$ und für eine Konstante $K > 0$ (6) erfüllt. Zu gegebenem $\varepsilon (0 < \varepsilon < 1)$ wähle man $S > 0$ so groß, daß

$$\frac{K}{m(S)} \leqq \frac{\varepsilon}{2}$$

ist. Sei $\{(p^i, E_i)\}$ ein ε-Code der Länge $N < \infty$ und

$$q := \frac{1}{N} \sum_{i=1}^{N} p^i$$

gesetzt. Dann hat man

$$\frac{1}{N}\sum_{i=1}^{N} p^i\left\{\frac{dp^i}{dq} > S\right\} < \frac{K}{m(S)} \leqq \frac{\varepsilon}{2},$$

also nach der Tschebyscheffungleichung

$$\frac{1}{N}\left|F := \left\{i : p^i\left\{\frac{dp^i}{dq} > S\right\} < \frac{2}{3}\varepsilon\right\}\right| > \frac{1}{4}.$$

Es ist

$$p^i\left(E_i \cap \left\{\frac{dp^i}{dq} < S\right\}\right) > \frac{\varepsilon}{2} \qquad (i \in F),$$

mithin

$$q(E_i) > \frac{1}{S} p^i\left(E_i \cap \left\{\frac{dp^i}{dq} < S\right\}\right) > \frac{\varepsilon}{3S} \qquad (i \in F),$$

und

$$1 \geqq \sum_{i \in F} q(E_i) \geqq |F| \cdot \frac{\varepsilon}{3S} > N \cdot \frac{\varepsilon}{12S},$$

also

$$N(M, \varepsilon) < \frac{1}{\varepsilon} 12 S(\varepsilon) < \infty \qquad (0 < \varepsilon < 1).$$

Bemerkung: Setzt man für $\tilde{p} \in \tilde{M}$

$$H_{\tilde{p}}(S) := \tilde{p}\{\tilde{f} < S\} \qquad (S \text{ reell}),$$

so gilt noch: Es ist $N(M, \varepsilon) < \infty$ $(0 < \varepsilon < 1)$ genau dann, wenn $\{H_{\tilde{p}}(.) : \tilde{p} \in \tilde{M}\}$ totalbeschränkt ist in der Levymetrik der Verteilungsfunktionen. Von dieser Aussage werden wir keinen Gebrauch machen, vermerken aber ausdrücklich, daß im Grunde genommen unsere gesamten Wachstumsbeschreibungen von $N(M, \varepsilon)$ mit Hilfe von Charakterisierungen der Menge $\{H_{\tilde{p}}(.) : \tilde{p} \in \tilde{M}\}$ durchgeführt werden.

Lemma 6.2. *Sei m eine auf $(0, \infty)$ definierte strikt monotone, positive stetige Funktion mit $m(S) \to \infty$ $(S \to \infty)$, und sei h eine strikt positive Funktion auf $(0, \infty)$. Dann gelten die folgenden Aussagen:*

(a) *Falls*

$$m\left(S \cdot N\left(M, \frac{1}{S}\right)\right) < h(S) \qquad (S > 1) \tag{7}$$

gilt, so existiert eine Konstante $K > 0$ mit

$$S\,\tilde{p}\{m(\tilde{f}) > h(S)\} < K \qquad (\tilde{p} \in \tilde{M}, S > 0). \tag{8}$$

(b) *Falls* (8) *für eine Konstante $K > 0$ gilt, so ist*

$$h\left(\frac{2K}{\varepsilon}\right) > m\left(\frac{1}{12} \cdot \varepsilon \cdot N(M, \varepsilon)\right) \qquad (0 < \varepsilon < 1).$$

Beweis. 1. Aus den Eigenschaften von m folgt die Existenz eines $S' > 0$ mit

$$m^{-1}(m(S)) = m(m^{-1}(S)) \qquad (S \geqq S').$$

Hieraus und aus (7) erhält man

$$S\,N\left(M, \frac{1}{S}\right) < m^{-1}(h(S)) \qquad (S \geqq S').$$

Anwendung des Maximalcode-Satzes liefert

$$\frac{m^{-1}(h(S))}{S} > m^{-1}(h(S))\left(\tilde{p}\{\tilde{f} > m^{-1}(h(S))\} - \frac{1}{S}\right) \qquad (S \geqq S'),$$

woraus die Ungleichung

$$2 > S\,\tilde{p}\{\tilde{f} > m^{-1}(h(S))\} \qquad (S \geqq S')$$

folgt. Es existiert daher ein $K > 0$, so daß (8) auch für $S \leqq S'$ gilt.

2. Sei (8) für ein $K > 0$ richtig. Für einen ε-Code $\{(p^i, E_i)\}$ der Länge N setze man

$$q := \sum_{i=1}^{N} p^i \quad \text{und} \quad S := \frac{2K}{\varepsilon}.$$

Dann gilt

$$\frac{1}{N}\sum_{i=1}^{N} p^i\left\{m\left(\frac{dp^i}{dq}\right) > h(S)\right\} < \frac{K}{S}$$

und somit nach der Tschebyscheff-Ungleichung

$$\left| F := \left\{i : p^i\left\{m\left(\frac{dp^i}{dq}\right) > h(S)\right\} < \frac{3}{2}\cdot\frac{K}{S}\right\}\right| > \frac{1}{3}N.$$

Es ist

$$p^i\left(E_i \cap \left\{m\left(\frac{dp^i}{dq}\right) < h(S)\right\}\right) > \frac{2K}{S} - \frac{3}{2}\cdot\frac{K}{S} \qquad (i \in F).$$

Wegen

$$m\left(\frac{dp^i}{dq}(x)\right) < h(S) \qquad \left(x \in E_i \cap \left\{m\left(\frac{dp^i}{dq}\right) < h(S)\right\}\right)$$

ist folglich

$$m\left(\frac{p^i\left(E_i \cap \left\{m\left(\frac{dp^i}{dq}\right) < h(S)\right\}\right)}{q(E_i)}\right) < h(S),$$

und man erhält

$$m\left(\frac{K}{2Sq(E_i)}\right) < h(S) \quad \text{bzw.} \quad m^{-1}\left(m\left(\frac{K}{2Sq(E_i)}\right)\right) < m^{-1}(h(S)) \qquad (i \in F).$$

Daraus ergibt sich

$$1 \geqq \sum_{i=1}^{N} q(E_i) > |F|\cdot\frac{K}{2S}\,\frac{1}{m^{-1}(h(S))} > \frac{1}{3}N\,\frac{K}{2S}\,\frac{1}{m^{-1}(h(S))}$$

und schließlich die Behauptung

$$h\left(\frac{2K}{\varepsilon}\right) > m\left(\frac{1}{12}\,\varepsilon\,N(M,\varepsilon)\right) \qquad (0 < \varepsilon < 1).$$

Corollar 6.2.1. *Sei m wie in Lemma 6.2 gewählt. Erfüllt m zusätzlich die Bedingung*

(9) $$\frac{m(S)}{\log^n S} \to 0 \quad (S \to \infty) \quad \textit{für eine natürliche Zahl } n,$$

so gilt:

(a) *Es ist $m(N(M,\varepsilon)) = O(1/\varepsilon)$ $(\varepsilon \to 0)$ genau dann, wenn für ein $K > 0$ $S\,\tilde{p}\{m(\tilde{f}) > S\} < K$ ist (glm. für $\tilde{p} \in \tilde{M}$, $S > 0$).*

(b) *Es ist* $m(N(M,\varepsilon)) = o(1/\varepsilon)$ $(\varepsilon \to 0)$ *genau dann, wenn* $S\tilde{p}\{m(\tilde{f}) > S\} \to 0$ $(S \to \infty)$ *gilt (glm. für* $\tilde{p} \in \tilde{M}$*).*

Die zusätzliche Voraussetzung (9) an m bewirkt nämlich, daß $m\left(\frac{1}{\varepsilon} N(M,\varepsilon)\right) = O(1/\varepsilon)$ bzw. $= o(1/\varepsilon)$ genau dann ist, wenn $m(\varepsilon N(M,\varepsilon)) = O(1/\varepsilon)$ bzw. $= o(1/\varepsilon)$ ist. Die Aussagen folgen damit direkt aus Lemma 6.2, wenn man für $O(1/\varepsilon)$ bzw. $o(1/\varepsilon)$ ein geeignetes $h(1/\varepsilon)$ einsetzt.

Lemma 6.3. *m sei wie in Lemma 6.2 gewählt. Dann gilt*

(a) *Aus*

$$\int_0^1 d\varepsilon\, m\left(\frac{1}{\varepsilon} N(M,\varepsilon)\right) < \infty \tag{10}$$

folgt

$$\sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, m(\tilde{f}) < \infty. \tag{11}$$

(b) *Es ist*

$$\sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, m(\tilde{f}) > \frac{1-b}{1+b}\, \varepsilon\, m\left(\frac{b^2}{1+b}\, \varepsilon N(M,\varepsilon)\right) \qquad (0 < \varepsilon < 1,\ 0 < b < 1).$$

Beweis. 1. Unter der Voraussetzung (10) existiert auf (0,1) eine strikt monoton wachsende stetige Funktion $h(S)$ mit

$$m\left(\frac{1}{\varepsilon} N(M,\varepsilon)\right) < h\left(\frac{1}{\varepsilon}\right) \quad (0 < \varepsilon < 1) \quad \text{und} \quad \int_0^1 d\varepsilon\, h\left(\frac{1}{\varepsilon}\right) < \infty.$$

Nach Lemma 6.2 gilt für dieses h nun (8) für eine Konstante $K > 0$. Die Behauptung (11) folgt dann aus

$$\begin{aligned}\int d\tilde{p}\, m(\tilde{f}) &= -S\tilde{p}\{m(\tilde{f}) > S\}\Big|_0^\infty + \int_0^\infty dS\, \tilde{p}\{m(\tilde{f}) > S\} = \\ &= \int_0^\infty dS\, \tilde{p}\{m(\tilde{f}) > S\} \leqq K_1 + \int_{K_2}^\infty dS\, \frac{h(S)}{S^2}\, S\tilde{p}\,\{m(\tilde{f}) > h(S)\} \\ &< K_1 + K\int_{K_2}^\infty dS\, \frac{h(S)}{S^2} = K_1 + K\int_0^{1/K_2} d\varepsilon\, h\left(\frac{1}{\varepsilon}\right) < K_3 \text{ (glm. für } \tilde{p} \in \tilde{M})\end{aligned}$$

mit geeigneten Konstanten K_1, K_2, K_3.

2. Sei $\{(p^i, E_i)\}$ ein ε-Code der Länge N und $q := \sum_{i=1}^N p^i$. Es ist

$$p^i\left\{m\left(\frac{dp^i}{dq}\right) > \frac{1}{\varepsilon(1-b)} \int dp^i\, m\left(\frac{dp^i}{dq}\right)\right\} < \varepsilon(1-b) \qquad (0 < b < 1),$$

also

$$p^i\left(E_i \cap \left\{m\left(\frac{dp^i}{dq}\right) < \frac{1}{\varepsilon(1-b)} \int dp^i\, m\left(\frac{dp^i}{dq}\right)\right\}\right) > \varepsilon b$$

und somit

$$m\left(\frac{\varepsilon b}{q(E_i)}\right) < \frac{1}{\varepsilon(1-b)} \int dp^i\, m\left(\frac{dp^i}{dq}\right). \tag{12}$$

Summation über i liefert

$$\frac{1}{N} \sum_{i=1}^N m\left(\frac{\varepsilon b}{q(E_i)}\right) < \frac{1}{\varepsilon(1-b)}\, \frac{1}{N} \sum_{i=1}^N \int dp^i\, m\left(\frac{dp^i}{dq}\right) \leqq \frac{1}{\varepsilon(1-b)} \sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, m(\tilde{f}).$$

Folglich ist

$$\left| F := \left\{ i : m\left(\frac{\varepsilon b}{q(E_i)}\right) < \frac{1+b}{\varepsilon(1-b)} \sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, m(\tilde{f}) \right\} \right| > \frac{b}{1+b} N$$

und daher

$$q(E_i) > \varepsilon b \frac{1}{m^{-1}\left(\frac{1+b}{\varepsilon(1-b)} \sup\limits_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, m(\tilde{f})\right)} \qquad (i \in F)$$

und

$$1 \geqq \sum_{i \in F} q(E_i) > \frac{b}{1+b} N \varepsilon b \frac{1}{m^{-1}\left(\frac{1+b}{\varepsilon(1-b)} \sup\limits_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, m(\tilde{f})\right)}.$$

Dies liefert die in (b) behauptete Ungleichung.

Aus Lemma 6.3 und Corollar 6.2.1 ergibt sich das

Corollar 6.3.1. *Erfüllt m zusätzlich* (9), *so gilt:*

(a) *Die Bedingung* (11) *folgt aus*

$$\int_0^1 d\varepsilon\, m(N(M, \varepsilon)) < \infty. \tag{13}$$

(b) *Aus* (11) *folgt*

$$m(N(M, \varepsilon)) = O\left(\frac{1}{\varepsilon}\right) \qquad (\varepsilon \to 0).$$

Corollar 6.3.2. *m erfülle* (9).

(a) *Aus* (13) *folgt dann die Existenz einer G-Funktion g mit*

$$\sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, g(m(\tilde{f})) < \infty.$$

(b) *Aus* (11) *folgt*

$$g^{-1}(m(N(M, \varepsilon))) = o\left(\frac{1}{\varepsilon}\right) \quad (\varepsilon \to 0) \quad (g \text{ } G\text{-}Funktion).$$

Beweis. Die Aussage (b) ist klar. Zum Beweis von (a) nutzen wir aus, daß $m(N(M, \varepsilon))$ mit (13) als einzelne Funktion gleichgradig integrabel ist und daher für eine geeignete G-Funktion g_1 noch

$$\int_0^1 d\varepsilon\, g_1(m(N(M, \varepsilon))) < \infty$$

ist (Satz 4.1). $g_1(S)$ hat eine Darstellung

$$g_1(S) = \int_0^S dy\, m_1(y)$$

mit einer positiven monotonen Funktion $m_1(y)$, für die $m_1(y) \to \infty$ $(y \to \infty)$ gilt. Man setze

$$g(S) := \int_0^S dy \min(m_1(y), y^n).$$

g ist G-Funktion, und es ist

$$g(S) \leqq \frac{1}{n+1} y^{n+1}.$$

Dann genügt aber $g(m(S))$ der Voraussetzung (9). Die Behauptung (a) folgt daher aus Corollar 6.3.1 (a), indem man dort $g(m(.))$ anstelle von $m(.)$ einsetzt.

Von den obigen Resultaten werden wir im folgenden vor allem die Aussage (b) von Corollar 6.2.1, Lemma 6.3 und Corollar 6.3.1 mit $m = \log^+$ bzw. $m = g(\log^+(.))$ benutzen. Die Beispiele am Schluß der Arbeit zeigen, daß die hier vorliegenden Aussagen in gewisser Weise optimal sind.

§ 7. Coding-Theorem für stationäre Kanäle und schwache Umkehrung

Wir führen den Begriff der Kapazität ein und behandeln einige einfache Eigenschaften der Kapazität. Ihre Bedeutung ergibt sich für stationäre Kanäle anhand der Aussagen des Coding-Theorems und seiner schwachen Umkehrung. Im Anschluß an die Beweise dieser Theoreme werden noch kurz Codelängenabschätzungen diskutiert für den Fall, daß die Kapazität nicht endlich ist.

Abweichend vom allgemeinen Brauch erklären wir hier nicht den Begriff der Kapazität des Kanals, sondern den der Kapazität einer Menge M, da sich dies als günstiger erweist.

Definition 7.1 *(Kapazität von M).* Für (X, B, M) heißt

$$C(M) := \sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p} \log \tilde{f}$$

die Kapazität von M.

Lemma 7.1. (a) *Für zwei WVen p^1, p^2 mit $p^1 \ll p^2$ ist*

$$\int dp^1 \log \frac{dp^1}{dp^2} \geqq 0, \tag{14}$$

$$\int dp^1 \log \frac{dp^1}{dp^2} > 0 \text{ genau dann, wenn } p^1 \neq p^2 \text{ ist,} \tag{15}$$

$$\int dp^1 \log^+ \frac{dp^1}{dp^2} \leqq \int dp^1 \log \frac{dp^1}{dp^2} + \frac{1}{e}. \tag{16}$$

(b) *Für beliebiges (X, B, M) ist*

$$C(M) \geqq 0,$$

$$C(M) > 0 \quad \text{genau dann, wenn } |M| > 1 \text{ ist,} \tag{17}$$

$$\int d\tilde{p} \log^+ \tilde{f} \leqq C(M) + \frac{1}{e} \qquad (\tilde{p} \in \tilde{M}).$$

(c) *Für $M_{[1,t]} = M_1 \times \cdots \times M_t$ gilt*

$$C(M_{[1,t]}) = \sum_{v=1}^{t} C(M_v).$$

(d) *Gibt es zu M eine WV λ und eine Konstante $K \geqq 1$ derart, daß für jedes $p \in M$ die Dichte $dp/d\lambda$ existiert und λ-fastüberall durch K beschränkt ist, so gilt*

$$C(M) \leqq \log K.$$

Speziell ist

$$C(\{p^1, \ldots, p^r\}) \leqq \log r.$$

Beweis. 1. Beweis von (a): Die für $y > 0$ durch

$$z(y) := y \log y \qquad (y > 0)$$

definierte Funktion z läßt sich nach Null stetig fortsetzen mit $z(0) = 0$. Es ist $z(1) = 0$. $z(y)$ ist konvex auf R^+ und nimmt für $y = 1/e$ ihr Minimum $-1/e$ an. Unter Benutzung der Jensenschen Ungleichung erhält man

$$\int dp^1 \log \frac{dp^1}{dp^2} = \int dp^2 \frac{dp^1}{dp^2} \log \frac{dp^1}{dp^2} = \int dp^2 z\left(\frac{dp^1}{dp^2}\right) \geqq z\left(\int dp^2 \frac{dp^1}{dp^2}\right) = z(1) = 0,$$

also (14). Da $z(y)$ strikt konvex ist, gilt

$$\int dp^2 z\left(\frac{dp^1}{dp^2}\right) = z\left(\int dp^2 \frac{dp^1}{dp^2}\right)$$

höchstens dann, wenn dp^1/dp^2 p^2-fastüberall konstant ist. Daraus gewinnt man (15). Zum Nachweis von (16) beachte man

$$\int dp^2 \left| z^-\left(\frac{dp^1}{dp^2}\right)\right| \leqq \frac{1}{e}.$$

2. Beweis von (b): Man wende die Aussagen (a) an auf

$$\int d\tilde{p} \log \tilde{f} = \int d\tilde{p} \log \frac{d\tilde{p}}{da \times dq}.$$

Wenn $|M| > 1$ ist, so gibt es ein $\tilde{p} \in \tilde{M}$ mit $\tilde{p} \neq a \times q$. Damit ersieht man (17).

3. Beweis von (c): Nach Corollar 1.2.1 gilt

$$\tilde{M}_{[1,t]} \supseteqq \tilde{M}_1 \times \cdots \times \tilde{M}_t.$$

Mithin ist

$$\begin{aligned} C(M_{[1,t]}) &= \sup_{\tilde{p} \in \tilde{M}_{[1,t]}} \int d\tilde{p} \log \tilde{f} \geqq \sup_{\tilde{p}_1 \times \cdots \times \tilde{p}_t \in \tilde{M}_1 \times \cdots \times \tilde{M}_t} \int d\tilde{p}_1 \times \cdots \times d\tilde{p}_t \log \tilde{f}_1 \times \cdots \times \tilde{f}_t \\ &= \sum_{v=1}^{t} \sup_{\tilde{p}_v \in \tilde{M}_v} \int d\tilde{p}_v \log \tilde{f}_v = \sum_{v=1}^{t} C(M_v). \end{aligned}$$

Wir haben daher nur noch

$$C(M_{[1,t]}) \leqq \sum_{v=1}^{t} C(M_v) \tag{18}$$

zu zeigen. Sei $\tilde{p} \in \tilde{M}_{[1,t]}$ von $\{p^i \in M_{[1,t]}\}$ und $\{a^i\}$ $(i = 1, \ldots, n)$ erzeugt. p_v^i bezeichne die v-te Komponente von p^i. Dann ist

$$\int d\tilde{p} \log \tilde{f} = \sum_{i=1}^{n} a^i \int dp^i \log \frac{dp^i}{\sum_{j=1}^{n} a^j dp^j} = A - B$$

mit

$$A = \sum_{i=1}^{n} a^i \int dp^i \log \frac{dp^i}{\left(\sum_{j=1}^{n} a^j dp_1^j\right) \times \cdots \times \left(\sum_{j=1}^{n} a^j dp_t^j\right)},$$

$$B = \sum_{i=1}^{n} a^i \int dp^i \log \frac{\left(\sum\limits_{k=1}^{n} a\, dp^k\right)}{\left(\sum\limits_{j=1}^{n} a^j dp_1^j\right) \times \cdots \times \left(\sum\limits_{j=1}^{n} a^j dp_t^j\right)}.$$

Wegen (14) ist $B \geqq 0$ und somit

$$\int d\tilde{p} \log \tilde{f} \leqq A = \sum_{v=1}^{t} \left(\sum_{i=1}^{n} a^i \int dp_v^i \log \frac{dp_v^i}{\left(\sum\limits_{j=1}^{n} a^j dp_v^j\right)} \right) \leqq \sum_{v=1}^{t} C(M_v).$$

4. Beweis von (d): Entsprechend dem Schluß im Beweis von (18) erhält man

$$\int d\tilde{p} \log \tilde{f} = \int d\tilde{p} \log \frac{d\tilde{p}}{da \times dq} = \sum_{i=1}^{n} a^i \int dp^i \log = \frac{dp^i}{\sum\limits_{k=1}^{n} a^k dp^k} =$$

$$= \sum_{i=1}^{n} a^i \int dp^i \log \frac{dp^i}{d\lambda} - \sum_{i=1}^{n} a^i \int \mathrm{d}p^i \log \frac{\left(\sum\limits_{k=1}^{n} a^k dp^k\right)}{d\lambda} \leqq$$

$$\leqq \sum_{i=1}^{n} a^i \int dp^i \log \frac{dp^i}{d\lambda} \leqq \sum_{i=1}^{n} a^i \log K = \log K.$$

Für $M = \{p^1, \ldots, p^r\}$ wähle man speziell $\lambda := \frac{1}{r} \sum\limits_{j=1}^{r} p^j$. Dann ist $dp^i/d\lambda \leqq r$ λ-fastüberall.

Wir merken noch an, daß die Kapazität vor allem auf Grund von Lemma 7.1 (c) für untere und obere Abschätzungen der maximalen Codelängen geeignet ist.

Satz 7.2. *(Coding-Theorem für stationäre gedächtnisfreie Kanäle). Für einen stationären gedächtnisfreien Kanal gelten die Aussagen:*

(a) $$\liminf_{t\to\infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon) > S \qquad (S < C(M_1), 0 < \varepsilon < 1).$$

(b) *Ist $C(M_1) = \infty$, so gibt es eine G-Funktion g derart, daß für jedes ε $(0<\varepsilon<1)$ ein $t_0(\varepsilon)$ existiert mit*

$$\log N(M_{[1,t]}, \varepsilon) > g(t) \qquad (t > t_0(\varepsilon)).$$

D. h. also: Wenn M_1 bedingt schwachkompakt ist und $C(M_1) = \infty$, so wächst $N(M_{[1,t]}, \varepsilon)$ mit t stärker als exponentiell an, jedoch bleibt stets $N(M_{[1,t]}, \varepsilon) < \infty$.

Beweis. 1. Sei zu gegebenem $S < C(M_1)$ ein $\tilde{p}_1 \in \tilde{M}_1$ so gewählt, daß

$$\int d\tilde{p}_1 \log \tilde{f}_1 > S$$

ist. Für jedes $v > 1$ sei $\tilde{p}_v \in \tilde{M}_v$ ein Exemplar von $\tilde{p}_1$. Nach dem schwachen Gesetz der großen Zahlen gilt dann

$$\tilde{p}_1 \times \cdots \times \tilde{p}_t \{\tilde{f}_1 \times \cdots \times \tilde{f}_t > \exp(tS)\} = \tilde{p}_1 \times \cdots \times \tilde{p}_t \left\{ \frac{1}{t} \sum_{v=1}^{t} \log \tilde{f}_v > S \right\} \xrightarrow{t\to\infty} 1.$$

Der Maximalcode-Satz 1.2 bzw. Corollar 1.2.1 liefert für hinreichend großes t

$$N(M_{[1,t]}, \varepsilon) > \exp(tS) \frac{1-\varepsilon}{2},$$

woraus sich die Behauptung (a) ergibt.

2. Sei $C(M_1) = \infty$. Wegen (a) gibt es eine positive monoton wachsende Funktion $m(y)$ auf R^+ mit $m(y) \to \infty$ $(y \to \infty)$ derart, daß

$$\log N(M_{[1,t]}, \varepsilon) > t\, m(t) \qquad (t > t_0(\varepsilon))$$

ist.

$$g(t) := \int_0^t dy\, m(y)$$

ist G-Funktion mit

$$g(t) \leqq t\, m(t).$$

Das liefert zusammen die unter (b) behauptete Ungleichung.

Die schwache Umkehrung des Coding-Theorems beweisen wir hier für nicht notwendig stationäre Kanäle.

Satz 7.3 *(schwache Umkehrung des Coding-Theorems für gedächtnisfreie Kanäle). Für einen gedächtnisfreien Kanal gilt die Beziehung*

$$\text{(a)} \quad \log(\varepsilon b\, N(M_{[1,t]}, \varepsilon)) \leqq \frac{1}{\varepsilon(1-b)} \left(C(M_{[1,t]}) + \frac{1}{e} \right) \qquad (0 < \varepsilon < 1,\ 0 < b < 1)$$

bzw. schwächer

$$\text{(b)} \quad \log(N(M_{[1,t]}, \varepsilon)) \leqq \frac{1}{\varepsilon}\, C(M_{[1,t]})\,(1 + o_t(1)) + o_t(t) \qquad (0 < \varepsilon < 1).$$

Beweis. 1. Sei $\{(p^i, E_i)\}$ ein ε-Code der Länge N für $M_{[1,t]}$ und

$$q := \frac{1}{N} \sum_{i=1}^{N} p^i.$$

Entsprechend (12) ergibt sich

$$\log\left(\frac{\varepsilon b}{q(E_i)}\right) < \frac{1}{\varepsilon(1-b)} \int \mathrm{d}p^i \log^+ \frac{dp^i}{dq} \qquad (i = 1, \ldots, N),$$

also durch Summation

$$(19) \quad \frac{1}{N} \sum_{i=1}^{N} \log\left(\frac{\varepsilon b}{q(E_i)}\right) < \frac{1}{\varepsilon(1-b)} \frac{1}{N} \sum_{i=1}^{N} \int dp^i \log^+ \frac{dp^i}{dq} \leqq \frac{1}{\varepsilon(1-b)} \left(C(M_{[1,t]}) + \frac{1}{e} \right).$$

Hieraus erhält man die in (a) behauptete Ungleichung, indem man noch die Beziehung

$$(20) \qquad \frac{1}{N} \sum_{i=1}^{N} -\log q(E_i) \geqq -\log\left(\frac{1}{N} \sum_{i=1}^{N} q(E_i)\right) \geqq -\log\frac{1}{N} = \log N$$

benutzt, die sich aus der Jensenschen Ungleichung, angewandt auf die konvexe Funktion $-\log(.)$, ergibt.

2. Die schwächere Aussage (b) gewinnt man aus (a), indem man dort mit $t \to \infty$ b langsam gegen Null streben läßt.

Aus Satz 7.2, Satz 7.3 (b) und Lemma 7.1 (c) ergibt sich für stationäre gedächtnisfreie Kanäle:

$$\lim_{\varepsilon \to 1} \limsup_{t \to \infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon) = C(M_1).$$

Daraus ersieht man wegen der in Lemma 3.3 angegebenen Konvexitäts- bzw. Stetigkeitseigenschaften von $N(., \varepsilon)$ die Gültigkeit von

Lemma 7.4. *Für beliebiges M ist*

$$C(M) = C(\overline{\mathrm{co}(M)}),$$

für bedingt schwachkompaktes M ist

$$C(M) = C(\mathrm{ex}(\overline{\mathrm{co}(M)})).$$

Zum Schluß dieses Paragraphen geben wir noch eine Beziehung an, die bei nichtendlichen Kapazitäten sinnvolle obere Abschätzungen der maximalen Codelängen mit der Zeit liefert.

Lemma 7.5. *Für jede G-Funktion g ist*

$$\frac{1-b}{1+b}\,\varepsilon g^{-1}\left(\log^+\left(\varepsilon\frac{b^2}{1+b}N(M_{[1,t]},\varepsilon)\right)\right) \leqq$$
$$\leqq \sum_{v=1}^{t} [\sup_{\tilde{p}_v\in\tilde{M}_v} \int d\tilde{p}_v g^{-1}(\log^+ \tilde{f}_v)] \quad (0<\varepsilon<1, 0<b<1).$$

Beweis. Sei $\{(p^i, E_i)\}$ ein ε-Code der Länge N für $M_{[1,t]}$,

$$q := \left(\frac{1}{N}\sum_{i=1}^{N} p_1^i\right) \times \cdots \times \left(\frac{1}{N}\sum_{i=1}^{N} p_t^i\right).$$

Die Abschätzung von Lemma 6.3 (b) liefert

$$\frac{1-b}{1+b}\,\varepsilon g^{-1}\left(\log^+\left(\varepsilon\frac{b^2}{1+b}N\right)\right) \leqq \frac{1}{N}\sum_{i=1}^{N}\int dp^i g^{-1}\left(\log^+\frac{dp^i}{dq}\right).$$

Da

$$\log^+\frac{dp^i}{dq} \leqq \sum_{v=1}^{t}\log^+\frac{dp_v^i}{dq_v}$$

ist und weiter

$$g^{-1}(S_1+S_2) \leqq g^{-1}(S_1) + g^{-1}(S_2) \qquad (S_1, S_2 \geqq 0) \tag{21}$$

gilt (dies entnimmt man sofort der Darstellung

$$g^{-1}(S) = \int_0^S dy\, m(y)$$

von g^{-1} mit einer monoton fallenden Funktion $m(y)$:

$$\int_0^{S_1+S_2} dy\, m(y) = \int_0^{S_1} dy\, m(y) + \int_{S_1}^{S_1+S_2} dy\, m(y) \leqq \int_0^{S_1} dy\, m(y) + \int_0^{S_2} dy\, m(y)),$$

bekommt man

$$\frac{1}{N}\sum_{i=1}^{N}\int dp^i g^{-1}\left(\log^+\frac{dp^i}{dq}\right) \leqq \sum_{v=1}^{t}\frac{1}{N}\sum_{i=1}^{N}\int dp_v^i g^{-1}\left(\log^+\frac{dp_v^i}{dq_v}\right) \leqq$$
$$\leqq \sum_{v=1}^{t} [\sup_{\tilde{p}_v\in\tilde{M}_v}\int d\tilde{p}_v g^{-1}(\log^+\tilde{f}_v)].$$

Als Ergänzung zu Satz 7.1 (b) gewinnt man aus Lemma 7.5 das

Lemma 7.6. *Sei für einen stationären gedächtnisfreien Kanal mit einer G-Funktion g*

$$\log N(M_{[1,t]}, \varepsilon) > g(t) \quad \text{für} \quad t > t_0(\varepsilon).$$

Dann gilt mit jeder G-Funktion h, die der Bedingung

$$\frac{g(t)}{h(Kt)} \xrightarrow{t\to\infty} \infty \text{ für beliebig große } K$$

genügt,

$$\sup_{\tilde{p}_1 \in \tilde{M}_1} \int d\tilde{p}_1 \, h^{-1}(\log^+ \tilde{f}_1) = \infty.$$

§ 8. Verschärfte direkte Codelängenabschätzungen für $M_{[1,t]}$

Unter Verwendung der Tschebyscheff-Ungleichung mit Varianzen führen wir in diesem Paragraphen verschärfte direkte Codelängenabschätzungen für $M_{[1,t]}$ durch. Zur Vereinfachung der Formeln setzen wir

$$D^2(M) := \sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p} \, (\log \tilde{f} - \int d\tilde{p} \log \tilde{f})^2.$$

Für Kanäle, die etwa die Voraussetzung

$$C(M_v) < K, \quad D^2(M_v) < K \quad (v \in Z) \quad \text{für ein } K > 0 \tag{22}$$

erfüllen, liefert Satz 8.1 ein Coding-Theorem und Satz 8.2 eine starke Umkehrung des Coding-Theorems. Zusammen besagen die Sätze in diesem Falle, daß

$$|\log N(M_{[1,t]}, \varepsilon) - C(M_{[1,t]})| = O(\sqrt{t})$$

ist. Lemma 8.5 gestattet es, die Beweismethoden von Satz 8.1 und Satz 8.2 auch bei unendlichen $D^2(M_v)$ zu benutzen und Abschätzungen der Art

$$|\log N(M_{[1,t,]} \varepsilon) - C(M_{[1,t]})| = o(t)$$

zu beweisen.

Wir merken an, daß die in den folgenden Beweisen auftretenden Integrale alle endlich sind. Z. B. folgt

$$\int d\tilde{p} \log^2 \tilde{f} = \int da \times dq \, \tilde{f} \log^2 \tilde{f} < \infty \qquad (\tilde{p} \in \tilde{M})$$

daraus, daß jedes $\tilde{f}$ $a \times q$-fastüberall durch eine Konstante beschränkt ist und die Funktion $y \log^2 y$ in $(0,1]$ bei e^{-2} ihr einziges Maximum (von der Größe $4e^{-2}$) hat (ferner gilt $y \log^2 y \to 0$ $(y \to 0)$).

Satz 8.1. *Ist $C(M_v)$ endlich für jedes $v \in Z$, so gilt*

$$\log N(M_{[1,t]}, \varepsilon) \geqq C(M_{[1,t]}) - \left(\frac{1}{1-\varepsilon-b} \sum_{v=1}^{t} D^2(M_v)\right)^{1/2} + \log b \qquad (0 < \varepsilon < 1, \; 0 < b < 1 - \varepsilon).$$

Beweis. Zur Vereinfachung der Schreibweise setzen wir

$$\tilde{p}_1 \times \cdots \times \tilde{p}_t =: \tilde{p}_{[1,t]},$$

$$\log \tilde{f}_{[1,t]} - \int d\tilde{p}_{[1,t]} \log \tilde{f}_{[1,t]} =: h.$$

Nach der Tschebyscheff-Ungleichung ist

$$\tilde{p}_{[1,t]}\left\{|h| > \left(\frac{1}{1-\varepsilon-b}\int d\tilde{p}_{[1,t]}\, h^2\right)^{1/2}\right\} < 1-\varepsilon-b,$$

und daher gilt

$$\tilde{p}_{[1,t]}\left\{\log\tilde{f}_{[1,t]} > \int d\tilde{p}_{[1,t]}\log\tilde{f}_{[1,t]} - \left(\frac{1}{1-\varepsilon-b}\int d\tilde{p}_{[1,t]}\, h^2\right)^{1/2}\right\} > \varepsilon + b.$$

Mit dem Maximalcode-Satz 1.2 bzw. Corollar 1.2.1 erhält man dann
$\log N(M_{[1,t]},\varepsilon) \geqq$

$$\geqq \sup_{\tilde{p}_{[1,t]}\in\tilde{M}_1\times\cdots\times\tilde{M}_t}\left[\int d\tilde{p}_{[1,t]}\log\tilde{f}_{[1,t]} - \left(\frac{1}{1-\varepsilon-b}\int d\tilde{p}_{[1,t]}\, h^2\right)^{1/2}\right] + \log(\varepsilon+b-\varepsilon)$$

$$(23)\qquad \geqq C(M_{[1,t]}) - \sup_{\tilde{p}_{[1,t]}\in\tilde{M}_1\times\cdots\times\tilde{M}_t}\left(\frac{1}{1-\varepsilon-b}\int d\tilde{p}_{[1,t]}\, h^2\right)^{1/2} + \log(\varepsilon+b-\varepsilon).$$

Aus dieser Ungleichung ergibt sich wegen

$$\int d\tilde{p}_{[1,t]}\, h^2 = \sum_{v=1}^{t}\int d\tilde{p}_v(\log\tilde{f}_v - \int d\tilde{p}_v\log\tilde{f}_v)^2 \leqq \sum_{v=1}^{t} D^2(M_v)$$

die Behauptung.

Bemerkung: Aus dem obigen Beweis erhält man auch für den Fall $D^2(M_1)=\infty$ direkt das in § 7 angegebene Coding-Theorem für stationäre Kanäle: (23) liefert

$$\log N(M_{[1,t]},\varepsilon) > t\int d\tilde{p}_1\log\tilde{f}_1 - \mathrm{const}(\tilde{p}_1)\sqrt{t}.$$

Man lasse in dieser Ungleichung mit $t\to\infty$ das Integral in der Weise gegen $C(M_1)$ gehen, daß $\mathrm{const}(\tilde{p}_1) = o(\sqrt{t})$ bleibt.

Satz 8.2. *Es ist für* $0<\varepsilon<1,\ 0<b<1$

$$\log N(M_{[1,t]},\varepsilon) \leqq C(M_{[1,t]}) + \left(\frac{1}{\varepsilon(1-b)}\sum_{v=1}^{t} D^2(M_v)\right)^{1/2} - \log(\varepsilon\cdot b).$$

Beweis. Sei $\{(p^i, E_i)\}$ ein ε-Code der Länge N für $M_{[1,t]}$ und

$$q := q_1\times\cdots\times q_t := \left(\frac{1}{N}\sum_{i=1}^{N} p_1^i\right)\times\cdots\times\left(\frac{1}{N}\sum_{i=1}^{N} p_t^i\right).$$

Nach der Tschebyscheff-Ungleichung ist

$$p^i\left\{\log\frac{dp^i}{dq} - \int dp^i\log\frac{dp^i}{dq} > \left[\frac{1}{\varepsilon(1-b)}\int dp^i\left(\log\frac{dp^i}{dq} - \int dp^i\log\frac{dp^i}{dq}\right)^2\right]^{1/2}\right\} < \varepsilon(1-b).$$

Analog zum Beweis von Satz 7.3 ((19), (20)) erhält man

$$\log(\varepsilon b\, N(M_{[1,t]},\varepsilon)) \leqq \frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{\varepsilon b}{q(E_i)}\right) <$$

$$< \frac{1}{N}\sum_{i=1}^{N}\int dp^i\log\frac{dp^i}{dq} + \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{\varepsilon(1-b)}\int dp^i\left(\log\frac{dp^i}{dq} - \int dp^i\log\frac{dp^i}{dq}\right)^2\right]^{1/2}.$$

Hieraus gewinnt man die behauptete Ungleichung, da

$$\frac{1}{N}\sum_{i=1}^{N}\int dp^i\log\frac{dp^i}{dq} = \sum_{v=1}^{t}\left(\frac{1}{N}\sum_{i=1}^{N}\int dp_v^i\log\frac{dp_v^i}{dq_v}\right) \leqq \sum_{v=1}^{t} C(M_v) = C(M_{[1,t]})$$

und wegen der Konkavität von $\sqrt{y}$ nach der Jensenschen Ungleichung

$$\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{\varepsilon(1-b)}\sum_{v=1}^{t}\int dp_v^i\left(\log\frac{dp_v^i}{dq_v}-\int dp_v^i\log\frac{dp_v^i}{dq_v}\right)^2\right]^{1/2}\leqq$$

$$\leqq\left[\frac{1}{\varepsilon(1-b)}\sum_{v=1}^{t}\left\{\frac{1}{N}\sum_{i=1}^{N}\int dp_v^i\log^2\frac{dp_v^i}{dq_v}-\frac{1}{N}\sum_{i=1}^{N}\left(\int dp_v^i\log\frac{dp_v^i}{dq_v}\right)^2\right\}\right]^{1/2}$$

ist, der Wert dieses letzten Ausdrucks aber

$$\leqq\left[\frac{1}{\varepsilon(1-b)}\sum_{v=1}^{t}D^2(M_v)\right]^{1/2}$$

ist, was sich aus

$$-\frac{1}{N}\sum_{i=1}^{N}\left(\int dp_v^i\log\frac{dp_v^i}{dq_v}\right)^2\leqq-\left(\frac{1}{N}\sum_{i=1}^{N}\int dp_v^i\log\frac{dp_v^i}{dq_v}\right)^2$$

ergibt (Jensensche Ungleichung für y^2).

Daß für gedächtnisfreie Kanäle mit endlichen Alphabeten und allgemeiner für gedächtnisfreie Kanäle, deren M_v in bezug auf feste WVen λ_v durch gleichmäßig beschränkte Dichten darstellbar sind, tatsächlich vermittels Satz 8.1 und Satz 8.2 verschärfte Abschätzungen der Codelängen durch die Kapazität geliefert werden. die schon für relativ kurze Zeiträume brauchbar sind, zeigt das nächste Lemma,

Lemma 8.3. *Wenn es zu M eine* WV λ *und eine Konstante* $K\geqq 1$ *gibt, so daß für jedes $p\in M$ die Dichte $dp/d\lambda$ existiert und λ-fastüberall durch K beschränkt ist, so gilt*

$$D^2(M)\leqq\left(\frac{1}{e}+\log K\right)^2+7e^{-2}. \tag{24}$$

Speziell hat man

$$D^2(\{p^1,\dots,p_r\})\leqq\left(\frac{1}{e}+\log r\right)^2+7e^{-2}. \tag{25}$$

Bemerkung: Eine (24) entsprechende Ungleichung wird von Kemperman bei der Fehlerangabe für die starke Umkehrung des Coding-Theorems für stationäre gedächtnisfreie Kanäle mit endlichen Alphabeten benutzt (vgl. Wolfowitz [*2*]).

Beweis. Es ist für $\tilde{p}\in\tilde{M}$

$$\int d\tilde{p}(\log\tilde{f}-\int d\tilde{p}\log\tilde{f})^2\leqq\int d\tilde{p}\log^2\tilde{f}=$$

$$\int d\tilde{p}\log^2\frac{d\tilde{p}}{da\times dq}=\int d\tilde{p}\left(\log\frac{d\tilde{p}}{da\times d\lambda}-\log\frac{da\times dq}{da\times d\lambda}\right)^2.$$

Man setze

$$F_1:=\{\tilde{f}<1\},$$

$$F_2:=\left\{\tilde{f}\geqq 1,\quad \frac{da\times dq}{da\times d\lambda}\geqq 1\right\},$$

$$F_3:=\left\{\tilde{f}\geqq 1,\quad \frac{da\times dq}{da\times d\lambda}<1,\frac{d\tilde{p}}{da\times d\lambda}<1\right\},$$

$$F_4:=\left\{\tilde{f}\geqq 1,\quad \frac{da\times dq}{da\times d\lambda}<1,\frac{d\tilde{p}}{da\times d\lambda}\geqq 1\right\}.$$

Nach der Bemerkung vor Satz 8.1 hat man dann

$$\int_{F_1}d\tilde{p}\log^2\tilde{f}\leqq 4e^{-2};$$

wegen $\frac{d\tilde{p}}{da \times d\lambda} \leqq K$ $a \times \lambda$-fastüberall gilt ferner

$$\int_{F_2} d\tilde{p} \log^2 \tilde{f} \leqq \int_{F_2} d\tilde{p} \log^2 \frac{d\tilde{p}}{da \times d\lambda} \leqq \tilde{p}(F_2) \log^2 K,$$

$$\int_{F_3} d\tilde{p} \log^2 \tilde{f} \leqq \int_{F_3} d\tilde{p} \log^2 \frac{da \times dq}{da \times d\lambda}$$

und

$$\int_{F_4} d\tilde{p} \log^2 \tilde{f} \leqq \int_{F_4} d\tilde{p} \left(\log K - \log \frac{da \times dq}{da \times d\lambda}\right)^2 =$$

$$= \tilde{p}(F_4) \log^2 K - 2 \log K \int_{F_4} d\tilde{p} \log \frac{da \times dq}{da \times d\lambda} + \int_{F_4} d\tilde{p} \log^2 \frac{da \times dq}{da \times d\lambda}.$$

Zusammen ergibt das

$$\int d\tilde{p} \log^2 \tilde{f} \leqq 4 e^{-2} + \log^2 K - 2 \log K \int_{F_4} d\tilde{p} \log \frac{da \times dq}{da \times d\lambda} + \int_{F_3 \cup F_4} d\tilde{p} \log^2 \frac{da \times dq}{da \times d\lambda}. \tag{26}$$

Mit

$$0 \geqq \int_{F_4} d\tilde{p} \log \frac{da \times dq}{da \times d\lambda} \geqq \int_{\left\{\frac{da \times dq}{da \times d\lambda} < 1\right\}} d\tilde{p} \log \frac{da \times dq}{da \times d\lambda} = \int_{\left\{\frac{da \times dq}{da \times d\lambda} < 1\right\}} da \times dq \log \frac{da \times dq}{da \times d\lambda} \geqq -\frac{1}{e}$$

und

$$\int_{F_3 \cup F_4} d\tilde{p} \log^2 \frac{da \times dq}{da \times d\lambda} \leqq \int_{\left\{\frac{da \times dq}{da \times d\lambda} < 1\right\}} da \times dq \log^2 \frac{da \times dq}{da \times d\lambda} \leqq 4 e^{-2}$$

gewinnt man aus (26)

$$\int d\tilde{p} \log^2 \tilde{f} \leqq 8 e^{-2} + \log^2 K + \frac{2}{e} \log K = \left(\frac{1}{e} + \log K\right)^2 + 7 e^{-2}$$

und hieraus (24). Anwendung von (24) auf den Spezialfall $M = \{p^1, \ldots, p^r\}$ mit

$$\lambda := \frac{1}{r} \sum_{i=1}^{r} p^i$$

liefert gerade die Ungleichung (25).

Die Aussagen von Satz 8.1 und Satz 8.2 werden für unsere Untersuchungen unbrauchbar, wenn $D^2(M_v) = \infty$ ist für mindestens ein v, und sie liefern auch schon dann keine brauchbaren Abschätzungen der Codelängen mehr, wenn $D^2(M_v)$ für $v \to \infty$ zu stark anwächst. Daß wir jedoch diese Fälle nicht auszuschließen brauchen aus unseren Betrachtungen, wird durch die beiden folgenden Lemmata gewährleistet. Herrn V. STRASSEN verdanke ich den Hinweis, daß nach CHINTCHIN folgender Sachverhalt besteht:

Lemma 8.4. *Sind $f_1, \ldots, f_t$ bezüglich der WV p unabhängige Zufallsvariablen (ZVen) über (X, B) mit Erwartungswert*

$$\int_X dp\, f_v = 0 \qquad (v = 1, \ldots, t),$$

so gilt

$$\int_X dp \left| \sum_{v=1}^{t} f_v \right|^{1+b} \leqq 2^{1+b} \sum_{v=1}^{t} \int_X dp\, |f_v|^{1+b} \qquad (0 \leqq b \leqq 1).$$

Für unsere Zwecke zeigen wir etwas allgemeiner

Lemma 8.5. *Sind $f_1, \ldots, f_t$ bezüglich der WV p unabhängige ZVen über (X, B) mit Erwartungswert*

$$\int_X dp\, f_v = 0 \qquad (v = 1, \ldots, t),$$

so ist für jede G-Funktion g, für die $g(\sqrt{.})$ auf R^+ konkav ist,

$$\int_X dp\, g\left(\left|\sum_{v=1}^t f_v\right|\right) \leqq 4 \sum_{v=1}^t \int_X dp\, g(|f_v|).$$

Beweis. 1. Wir nehmen zunächst an, daß die f_v symmetrisch verteilt sind, daß also

$$p\{f_v < S\} = p\{-f_v < S\} \quad \text{für jedes} \quad S \in (-\infty, \infty)$$

ist ($v = 1, \ldots, t$). Seien $r_1, \ldots, r_t$ auf $[0,1]$ definierte integrierbare Funktionen, die nur die Werte $+1$ und -1 annehmen und für die

$$\int_{[0,1]} dy\, r_i(y)\, r_j(y) = 0 \qquad (i \neq j)$$

ist. Dann haben die ZVen

$$\sum_{v=1}^t f_v \quad \text{und} \quad \sum_{v=1}^t f_v r_v(y) \quad \text{für jedes feste} \quad y \in [0,1]$$

dieselben Verteilungsfunktionen. Folglich gilt

$$\begin{aligned}\int_X dp\, g\left(\left|\sum_{v=1}^t f_v\right|\right) &= \int_X dp\left[\int_{[0,1]} dy\, g\left(\left|\sum_{v=1}^t f_v(x)\, r_v(y)\right|\right)\right] = \\ &= \int_X dp\left[\int_{[0,1]} dy\, g\left(\sqrt{\left|\sum_{v=1}^t f_v(x)\, r_v(y)\right|^2}\right)\right] \leqq \\ &\leqq \int_X dp\, g\left(\sqrt{\left[\int_{[0,1]} dy\left(\sum_{v=1}^t f_v(x)\, r_v(y)\right)^2\right]}\right) = \int_X dp\, g\left(\sqrt{\sum_{v=1}^t |f_v(x)|^2}\right).\end{aligned}$$

Hierin folgt die Ungleichung aus der Anwendung der Jensenschen Ungleichung auf die konkave Funktion $g(\sqrt{.})$. Wegen (21) ist noch

$$\int_X dp\, g\left(\sqrt{\sum_{v=1}^t |f_v(x)|^2}\right) \leqq \sum_{v=1}^t \int_X dp\, g(|f_v|).$$

Für bezüglich p symmetrisch verteilte unabhängige ZVen $f_1, \ldots, f_t$ gilt also

$$\int_X dp\, g\left(\left|\sum_{v=1}^t f_v\right|\right) \leqq \sum_{v=1}^t \int_X dp\, g(|f_v|). \tag{27}$$

2. Zum Beweis des allgemeinen Falles nehmen wir uns ein weiteres Exemplar (X', B', p') von (X, B, p) her und bezüglich p' unabhängige ZVen $f_{t+1}, \ldots, f_{t+t}$ über (X', B'), für die

$$p\{f_v < S\} = p'\{-f_{t+v} < S\} \qquad (v = 1, \ldots, t)$$

gelte. Man setze

$$\begin{aligned}\bar{f}_v(x, x') &:= f_v(x) \qquad (\text{für jedes } x' \in X') \quad (v = 1, \ldots, t),\\ \bar{f}_{t+v}(x, x') &:= f_{t+v}(x') \quad (\text{für jedes } x \in X) \quad (v = 1, \ldots, t).\end{aligned}$$

Da die f_v den Erwartungswert 0 haben, liefert die Jensensche Ungleichung für bedingte Erwartung wegen der Konvexität von g

$$\int_X dp\, g\left(\left|\sum_{v=1}^{t} f_v\right|\right) \leqq \int_{X\times X'} dp \times dp' g\left(\left|\sum_{v=1}^{t} (\bar f_v + \bar f_{t+v})\right|\right).$$

Nun ist $(\bar f_v + \bar f_{t+v})$ bezüglich $p \times p'$ symmetrisch verteilt und somit nach (27) der Wert des letzten Ausdrucks

$$\leqq \sum_{v=1}^{t} \int_{X\times X'} dp \times dp'\, g(|\bar f_v + \bar f_{t+v}|).$$

Wegen

$$\int_{X\times X'} dp \times dp' g(|\bar f_v + \bar f_{t+v}|) \leqq \int_{X\times X'} dp \times dp' g\left(\frac{2|\bar f_v| + 2|\bar f_{t+v}|}{2}\right) \leqq$$
$$\leqq \int_{X\times X'} dp \times dp' \tfrac{1}{2}[g(2|\bar f_v|) + g(2|\bar f_{t+v}|)] = \int_X dp(2|f_v|)$$

und

$$g(2|f_v|) = g\left(\sqrt{4|f_v|^2}\right) \leqq 4g\left(\sqrt{|f_v|^2}\right) = 4g(|f_v|)$$

ergibt sich für unabhängige ZVen mit Erwartungswert 0 die Ungleichung

$$\int_X dp\, g\left(\left|\sum_{v=1}^{t} f_v\right|\right) \leqq 4 \cdot \sum_{v=1}^{t} \int_X dp\, g(|f_v|).$$

Unter Verwendung von Lemma 8.5 beweist man völlig analog zum Beweis von Satz 8.1 den

Satz 8.6. *Für den von $\{M_v : v \in Z\}$ erzeugten gedächtnisfreien Kanal sei $C(M_v) < \infty$ für jedes $v \in Z$. Dann ist*

$$\log N(M_{[1,t]}, \varepsilon) \geqq$$
$$\geqq C(M_{[1,t]}) - g^{-1}\left(\frac{1}{1-\varepsilon-b}\, 4 \sum_{v=1}^{t} [\sup_{\tilde p_v \in \tilde M_v} \int d\tilde p_v g(|\log \tilde f_v - \int d\tilde p_v \log \tilde f_v|)]\right) + \log b$$
$$(0 < \varepsilon < 1,\ 0 < b < 1 - \varepsilon)$$

für jede G-Funktion g, für die $g(\sqrt{.})$ konkav auf R^+ ist.

Ferner gilt der

Satz 8.7 *Für den von $\{M_v : v \in Z\}$ erzeugten gedächtnisfreien Kanal ist*

$$\log N(M_{[1,t]}, \varepsilon) \leqq$$
$$\leqq C(M_{[1,t]}) + g^{-1}\left(\frac{1}{\varepsilon(1-b)}\, 16 \sum_{v=1}^{t} [\sup_{\tilde p_v \in \tilde M_v} \int d\tilde p_v g(|\log \tilde f_v|)]\right) - \log(\varepsilon b)$$
$$(0 < \varepsilon < 1,\ 0 < b < 1)$$

für jede G-Funktion g, für die $g(\sqrt{.})$ konkav ist.

Der Beweis verläuft analog dem Beweis von Satz 8.2 mit Hilfe von Lemma 8.5. Man hat nur an der entsprechenden Stelle zu beachten, daß

$$\frac{1}{N}\sum_{i=1}^{N} \int dp_v^i g\left(\left|\log\frac{dp_v^i}{dq_v} - \int dp_v^i \log\frac{dp_v^i}{dq_v}\right|\right) \leqq$$
$$\leqq \frac{1}{N}\sum_{i=1}^{N} \int dp_v^i \frac{1}{2}\left[g\left(2\left|\log\frac{dp_v^i}{dq_v}\right|\right) + g\left(2\left|\int dp_v^i \log\frac{dp_v^i}{dq_v}\right|\right)\right] \leqq$$
$$\leqq \frac{1}{N}\sum_{i=1}^{N} \int dp_v^i g\left(2\left|\log\frac{dp_v^i}{dq_v}\right|\right) \leqq 4\frac{1}{N}\sum_{i=1}^{N} \int dp_v^i g\left(\left|\log\frac{dp_v^i}{dq_v}\right|\right)$$

ist.

Wir werden später zeigen (§ 11, Beispiel 11.4), daß bei nichtstationären gedächtnisfreien Kanälen die maximalen Codelängen mit der Zeit durchaus nicht exponentiell anzuwachsen brauchen, wie es etwa bei stationären gedächtnisfreien Kanälen (Satz 7.1) der Fall ist.

Das folgende Lemma beschreibt eine untere Abschätzung der maximalen Codelängen, die für nichtstationäre Kanäle noch interessant ist.

Lemma 8.8. *Sei h^{-1} konkav auf R, $h^{-1}(y) = y$ $(y \leqq 0)$, $h^{-1}(y) > 0$ $(y > 0)$. Sei g G-Funktion, für die $g(\sqrt{.})$ konkav ist. Für jedes $v \in Z$ sei*

$$\sup_{\tilde{p}_v \in \tilde{M}_v} \int d\tilde{p}_v h^{-1}(\log \tilde{f}_v) < \infty .$$

Dann ist

$$\log N(M_{[1,t]}, \varepsilon) \geqq t \cdot h\Bigg(\frac{1}{t} \sum_{v=1}^{t} [\sup_{\tilde{p}_v \in \tilde{M}_v} \int d\tilde{p}_v h^{-1}(\log \tilde{f}_v)] -$$
$$- \frac{1}{t} g^{-1}\Bigg(\frac{1}{1-\varepsilon-b} \sum_{v=1}^{t} [\sup_{\tilde{p}_v \in \tilde{M}_v} \int d\tilde{p}_v g(|h^{-1}(\log \tilde{f}_v) - \int d\tilde{p}_v h^{-1}(\log \tilde{f}_v)|)]\Bigg)\Bigg) + \log b$$

$$(0 < \varepsilon < 1,\ 0 < b < 1 - \varepsilon).$$

Man vergleiche die Aussage mit der von Satz 8.6 und beachte

$$\frac{1}{t} \sum_{v=1}^{t} h^{-1}(\log \tilde{f}_v) \leqq h^{-1}\left(\frac{1}{t} \sum_{v=1}^{t} \log \tilde{f}_v\right).$$

§ 9. An die schwache Topologie angepaßte Kapazität

Auf Grund der Rolle, die die stationären gedächtnisfreien Kanäle spielen, ergibt sich mit den Aussagen des § 7 unser Interesse an Abschätzungen der folgenden Art:

$$(28) \qquad |\log N(M_{[1,t]}, \varepsilon) - C(M_{[1,t]})| = o(t) \quad \textit{für jedes feste } \varepsilon, \qquad (0 < \varepsilon < 1).$$

Für einen gedächtnisfreien Kanal ist (28) nach § 8 erfüllt, wenn gilt:

Es gibt eine G-Funktion g, für die $g(\sqrt{.})$ konkav auf R^+ ist, derart, daß

$$(29) \qquad \sup_{v \in Z} \sup_{\tilde{p}_v \in \tilde{M}_v} \int d\tilde{p}_v g(|\log \tilde{f}_v|) < \infty$$

ist.

Die Forderung (29) ist, wie sich aus den Überlegungen der Paragraphen 4 und 6 ergibt, selbst vom gleichen Typus wie die Forderung

$$\sup_{v \in Z} C(M_v) < \infty .$$

Wegen der in § 8 beschriebenen Art der Abhängigkeit des Fehlergliedes $o(t)$ in (28) von den Funktionen g, mit denen (29) gilt, ist $C(M_{[1,t]})$ mit der Zeit ein guter Steuerungsparameter für das Verhalten der maximalen Codelängen, wenn der Kanal starke Gleichmäßigkeitsbedingungen von der Art wie (29) erfüllt. Die Bedingung (29) bzw. die weiter unten angegebenen Bedingungen (31) bis (35) lassen sich offenbar ansehen als „auf die schwache Topologie (des Raumes der LVen) bezogene Gleichgradigkeitsbedingungen an die Erzeugung der Kapazität“.

Oftmals lassen sich jedoch bei speziellen Modellen für Kanäle eher einige andere Gleichmäßigkeitseigenschaften ablesen, die wir hier ebenfalls aufführen werden und von denen wir zeigen, in welcher Weise sie sich der Bedingung (29) unterordnen.

Um einigermaßen deutlich zu machen, wie stark die Forderung (29) ist, sei noch folgendes angemerkt: Wenn

$$\sup_{v\in Z}\int_0^1 d\varepsilon\, h(\log N(M_v,\varepsilon)) < \infty \quad \textit{für eine G-Funktion } h \tag{30}$$

ist, so gilt (29). Wenn $\log N(M,\varepsilon) \geqq h(1/\varepsilon)$ ist für eine G-Funktion h und für beliebig kleine ε, dann ist $C(M)=\infty$. (Man erhält diese Aussagen mittels der Abschätzungen des § 6.)

Wir führen die Untersuchung nicht für Serien $\{M_v\}_{v\in Z}$, sondern nur für Mengen M durch.

Lemma 9.1. *Für (X, B, M) sind die folgenden Bedingungen äquivalent:*

(31) *Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ derart, daß aus $\tilde{p}(\tilde{E}) < \delta$ ($\tilde{E}$ aus dem Definitionsbereich von $\tilde{p}$) stets $\int_{\tilde{E}} d\tilde{p}\log\tilde{f} < \varepsilon$ (gleichmäßig für $\tilde{p}\in\tilde{M}$) folgt.*

(32) *M ist bedingt schwachkompakt, und zu jedem $\varepsilon > 0$ gibt es ein $\delta^0 > 0$ derart, daß für beliebige endlichviele WVen $p^1,\ldots,p^n\in M$ und Mengen $E_1,\ldots,E_n\in B$ mit $p^i(E_i) < \delta^0$ $(i=1,\ldots,n)$ stets*

$$\sum_{i=1}^n a^i\int_{E_i} dp^i\log\frac{dp^i}{\sum_{k=1}^n a^k\,dp^k} < \varepsilon$$

ist. (Hierbei sind die $a^i \geqq 0$ beliebig mit $\sum_{i=1}^n a^i = 1$).

(33) *Es gibt eine G-Funktion g mit*

$$\sup_{\tilde{p}\in\tilde{M}}\int d\tilde{p}\, g(|\log\tilde{f}|) < \infty.$$

(34) *Es gibt eine G-Funktion g mit*

$$\sup_{\tilde{p}\in\tilde{M}}\int d\tilde{p}\, g(\log^+\tilde{f}) < \infty.$$

(35) *Es gibt eine G-Funktion g, für die $g(\sqrt{.})$ konkav auf R^+ ist, mit*

$$\sup_{\tilde{p}\in\tilde{M}}\int d\tilde{p}\, g(|\log\tilde{f}|) < \infty.$$

Beweis. Da $y\log^2 y \leqq 4e^{-2}$ ist im Intervall [0,1], ergibt

$$\int d\tilde{p}\,|\log^-\tilde{f}|^2 = \int d\tilde{p}\frac{1}{\tilde{f}}\tilde{f}\,|\log^-\tilde{f}|^2 = \int da\times dq\,\tilde{f}\,|\log^-\tilde{f}|^2 \leqq 4e^{-2}.$$

(Für $\log^-\tilde{f}$ läßt sich also y^2 als G-Funktion benutzen.) Hiermit und mit der Übertragung des Beweises von Satz 4.1 erhält man die Äquivalenz von (33), (34) und (35). Daß (31) und (33) äquivalent sind, ist dann ebenfalls klar, wenn man beachtet, daß in (31) (anders als bei Satz 4.1) nicht mehr $C(M) < \infty$ gefordert werden muß, da man die WVen $\tilde{p}$ als atomfrei ansehen kann (durch weiteres Zerlegen der a^i, indem man p^i mehrfach auftreten läßt).

Es ist nun noch die Äquivalenz von (31) und (32) zu zeigen. Aus (31) folgt $C(M) < \infty$. Daher ist M bedingt schwachkompakt. Der Rest der Forderung (32) ist offenbar mindestens so schwach wie (31). Sei nun umgekehrt (32) erfüllt. Wir zeigen zunächst, daß dann $C(M) < \infty$ ist: Da M bedingt schwachkompakt ist, existieren zu jedem $\delta > 0$ höchstens endlichviele Atome $A_1, \ldots, A_r \in B$ mit

$$\sup_{p \in M} p(A_j) > \frac{\delta}{2}.$$

Sei λ WV auf B mit $\lambda(A_j) = 1/r$ $(j = 1, \ldots, r)$, $\{p^i\}$ eine endliche Menge $\subseteqq M$ und a WV auf $\{i\}$. Dann ist (mit $q := \sum a^i p^i$)

$$\sum a^i \int_{\cup A_j} dp^i \log^+ \frac{dp^i}{dq} \leqq \sum a^i \int_{\cup A_j} dp^i \log^+ \frac{dp^i}{d\lambda} + \frac{1}{e} \leqq \log r + \frac{1}{e}.$$

Ferner erhält man mit der Gleichgradigkeitsforderung in (32)

$$\sum a^i \int_{X - \cup A_j} dp^i \log^+ \frac{dp^i}{dq} < \varepsilon \cdot \frac{2}{\delta(\varepsilon)}.$$

Insgesamt ergibt sich damit

$$C(M) \leqq \varepsilon \cdot \frac{2}{\delta(\varepsilon)} + \log r_{(\delta(\varepsilon))} + \frac{1}{e}, \quad \text{also} \quad C(M) < \infty.$$

Sei (32) für M erfüllt und

$$\sum_{i=1}^{n} a^i p^i(E_i) < \frac{\delta^0(\varepsilon/2)}{S}$$

mit einem unten noch geeignet zu bestimmenden $S \geqq 1$, ferner sei

$$F_S := \left\{ i : (1 \leqq i \leqq n),\ p^i(E_i) > S \cdot \delta^0\left(\frac{\varepsilon}{2}\right) \right\};$$

dann ist

$$\sum_{i=1}^{n} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} = \sum_{i \in \operatorname{compl} F_S} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} + \sum_{i \in F_S} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} \leqq \frac{\varepsilon}{2} + \sum_{i \in F_S} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq}.$$

Wir schätzen jetzt den restlichen Summenterm ab und nutzen dabei

$$\sum_{i \in F_S} a^i < \frac{1}{S}$$

aus. Für $i \in F_S$ sei

$$b^i := \frac{a^i}{\sum_{k \in F_S} a^k}$$

gesetzt (ohne Einschränkung sei $a^i > 0$ $(i = 1, \ldots, n)$). Es ist dann

$$\begin{aligned}
\sum_{i \in F_S} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} &= \left(\sum_{i \in F_S} a^i\right) \left[\sum_{i \in F_S} b^i \int_{E_i} dp^i \log \frac{dp^i}{\sum_{k \in F_S} b^k dp^k}\right] + \\
&\quad + \left(\sum_{i \in F_S} a^i\right) \left(\sum_{i \in F_S} b^i \int_{E_i} dp^i \log \left[\frac{\sum_{k \in F_S} a^k dp^k}{\sum_{k=1}^{n} a^k dp^k} \cdot \frac{1}{\left(\sum_{i \in F_S} a^k\right)} \right] \right) \leqq \\
&\leqq \frac{1}{S}\left(C(M) + \frac{1}{e}\right) - z\left(\frac{1}{S}\right) \qquad \left(\text{für } \frac{1}{S} \leqq \frac{1}{e}\right).
\end{aligned}$$

Wegen $C(M) < \infty$ kann man S zu $\varepsilon/2$ in der Weise bestimmen, daß die rechte Seite der letzten Ungleichung $< \varepsilon/2$ ist. Daraus ergibt sich (31): Es ist

$$\int_{\tilde{E}} d\tilde{p} \log \tilde{f} < \varepsilon, \quad \text{falls nur} \quad \tilde{p}(\tilde{E}) < \frac{d^0(\varepsilon/2)}{S(\varepsilon/2)} \text{ ist}.$$

Eine weitere naheliegende Gleichmäßigkeitsforderung an M ist offenbar die folgende:

$$(36) \quad \begin{cases} \text{Zu jedem } \varepsilon' > 0 \text{ gibt es endlichviele Teilmengen } M^j (\subseteqq M) \ (j = 1, \ldots, K(\varepsilon')) \\ \text{mit} \\ \qquad C(M^j) < \varepsilon' \quad \text{und} \quad M = \bigcup_{j=1}^{K(\varepsilon')} M^j. \end{cases}$$

Für stationäre gedächtnisfreie Kanäle, deren Menge M_1 der Bedingung (36) genügt, erhält man in direkter Verallgemeinerung des Beweises von KEMPERMAN in [2] die starke Umkehrung des Coding-Theorems. Wir verzichten auf die Wiedergabe des Beweises dieser Umkehrung, da sie keine bessere Fehlerabschätzung als die Umkehrungen in § 8 liefert und da sich zudem die Bedingung (36) genau aus (32) bei Einschränkung auf bedingt normkompakte Mengen M ergibt, wie wir jetzt zeigen werden.

Lemma 9.2. *Genügt M der Bedingung* (36), *so ist M bedingt normkompakt.*

Beweis. Zum Beweis der Behauptung benutzen wir die folgende von PINSKER [5] angegebene Ungleichung, die man mittels der Tschebyscheff-Ungleichung herleiten kann:

$$(37) \qquad \int dp \left| \log \frac{dp}{dq} \right| \leqq \int dp \log \frac{dp}{dq} + \min\left(\frac{1}{e}, c \cdot \sqrt{\int dp \log \frac{dp}{dq}}\right),$$

(p, q WVen mit $p \ll q$, c universelle Konstante).

Sei M^j eine der nach (36) existierenden Mengen mit $C(M^j) < \varepsilon'$, und seien p^1, p^2 beliebig $\in M^j$. Die Behauptung folgt dann (mit $z(y) := y \log y$) aus

$$\begin{aligned} \|p^1 - p^2\| &= \|p^1 - \tfrac{1}{2}(p^1 + p^2)\| + \|p^2 - \tfrac{1}{2}(p^1 + p^2)\| = \\ &= 4 \sum_{i=1}^{2} \tfrac{1}{2} \int \tfrac{1}{2}(dp^1 + dp^2) \left[\frac{dp^i}{\frac{1}{2}(dp^1 + dp^2)} - 1\right]^+ \leqq \\ &\leqq 4 \sum_{i=1}^{2} \tfrac{1}{2} \int \tfrac{1}{2}(dp^1 + dp^2)\, z^+\left(\frac{dp^i}{\frac{1}{2}(dp^1 + dp^2)}\right) = \\ &= 4 \sum_{i=1}^{2} \tfrac{1}{2} \int dp^i \log^+ \frac{dp^i}{\frac{1}{2}(dp^1 + dp^2)} \leqq \\ &\leqq 4\left(C(M^j) + c \cdot \sqrt{C(M^j)}\right) \leqq 4\left(\varepsilon' + c \cdot \sqrt{\varepsilon'}\right). \end{aligned}$$

Lemma 9.3. *Erfüllt M die Bedingung* (36), *so auch die Bedingung* (32).

Beweis. Es gelte (36). Wir nehmen an, daß die M^j ($j = 1, \ldots, K(\varepsilon')$) paarweise disjunkt sind, was nach Lemma 7.4 zulässig ist. Es sei nun $p^i(E_i) \underset{\blacktriangle}{<} \delta^0$ und ohne Einschränkung $a^i > 0$ ($i = 1, \ldots, n$). Wir setzen

$$\begin{aligned} F_j &:= \{i : p^i \in M^j\} \qquad (j = 1, \ldots, K(\varepsilon')), \\ F' &:= \{j : 1 \leqq j \leqq K(\varepsilon'),\ F_j \neq \emptyset\}, \end{aligned}$$

und

$$c^j := \sum_{i \in F_j} a^i \quad \text{für} \quad j \in F' .$$

Es ist

$$\sum_{i=1}^{n} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} = A_1 + A_2$$

mit

$$\begin{aligned} A_1 &:= \sum_{F'} c^j \left[\sum_{F_j} \frac{a^i}{c^j} \int_{E_i} dp^i \log \frac{\sum_{F_j} \frac{a^i}{c^j} dp^i}{dq} \right] \\ &\leqq \sum_{F'} \left[\sum_{F_j} \frac{a^i}{c^j} \int_{E_i} dp^i c^j \log \frac{1}{c^j} \right] \\ &\leqq \sum_{F'} \left[\sum_{F_j} \frac{a^i}{c^j} p^i(E_i) \right] \cdot \left[-z \left(\min \left(\frac{1}{e}, c^j \right) \right) \right] \\ &\leqq K(\varepsilon') \cdot \delta_0 \cdot \frac{1}{e}, \\ A_2 &:= \sum_{F'} c^j \left[\sum_{F_j} \frac{a^i}{c^j} \int_{E_i} dp^i \log \frac{dp^i}{\sum_{F_j} \frac{a^i}{c^j} dp^i} \right] \\ &\leqq \sum_{F'} c^j \left(C(M^j) + c \cdot \sqrt{C(M^j)} \right) \leqq \varepsilon' + c \cdot \sqrt{\varepsilon'} \quad (\text{nach } (37)), \end{aligned}$$

also

$$\sum_{i=1}^{n} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} \leqq K(\varepsilon') \cdot \delta_0 \cdot \frac{1}{e} + \varepsilon' + c \cdot \sqrt{\varepsilon'} .$$

Man wähle nun ε' zu ε so, daß $\varepsilon' + c \cdot \sqrt{\varepsilon'} < \varepsilon/2$ ist, dann $\delta_0(\varepsilon')$ so, daß

$$K(\varepsilon') \cdot \delta_0(\varepsilon') \frac{1}{e} < \frac{\varepsilon}{2}$$

ist.

Lemma 9.4. *Ist M bedingt normkompakt und genügt M der Bedingung* (32), *so genügt M auch der Bedingung* (36).

Beweis. Wir benutzen die folgende Ungleichung (Pinsker [5]), die man sofort mittels der Tschebyscheff-Ungleichung erhält:

Für je zwei WVen p, q mit $p \ll q$ und für beliebiges $S > 0$ ist

$$p\left\{ \left| \log \frac{dp}{dq} \right| > S \right\} \leqq \frac{1+S}{S} \| p - q \| . \tag{38}$$

Seien $p^1, \ldots, p^n \in M^j \subseteq M$, $a^i \geqq 0$ $(i = 1, \ldots, n)$, $\sum_{i=1}^{n} a^i = 1$,

$$q := \sum_{i=1}^{n} a^i p^i \quad \text{und} \quad E_i := \left\{ \left| \log \frac{dp^i}{dq} \right| > \frac{\varepsilon}{2} \right\}.$$

Es ist

$$\begin{aligned} \sum_{i=1}^{n} a^i \int dp^i \log \frac{dp^i}{dq} &\leqq \sum_{i=1}^{n} a^i \int_{\mathrm{compl}(E_i)} dp^i \left| \log \frac{dp^i}{dq} \right| + \sum_{i=1}^{n} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} \leqq \\ &\leqq \frac{\varepsilon}{2} + \sum_{i=1}^{n} a^i \int_{E_i} dp^i \log \frac{dp^i}{dq} . \end{aligned}$$

Hat M^j einen Durchmesser $D(M^j)$ mit

$$\frac{1+\frac{\varepsilon}{2}}{\varepsilon/2}\cdot D(M^j) < \delta_0\left(\frac{\varepsilon}{2}\right),$$

(wobei $\delta_0(\varepsilon/2)$ nach (32) gewählt ist), dann folgt mit (38)

$$\sum_{i=1}^{n} a^i \int_{E_i} dp^i \log\frac{dp^i}{dq} < \frac{\varepsilon}{2}$$

und somit

$$\sum_{i=1}^{n} a^i \int_{E_i} dp^i \log\frac{dp^i}{dq} < \varepsilon\,.$$

Wir formulieren nun noch einige weitere Forderungen an M und stellen Relationen zu den schon angegebenen her.

(39) $\begin{cases}\textit{Es existiert eine WV } \lambda \textit{ derart, daß } p \ll \lambda \textit{ ist für } p\in M \textit{ und } \left\{z\left(\frac{dp}{d\lambda}\right) : p \in M\right\}\\ \textit{eine gleichgradig } \lambda\textit{-integrierbare Menge von Funktionen ist.}\end{cases}$

(40) $\{$ *M genügt* (39) *und ist bedingt normkompakt.*

(41) $\begin{cases}\textit{Es existiert eine WV } \lambda \textit{ so, daß jedes } p\in M \textit{ totalstetig bezüglich } \lambda \textit{ ist und } M\\ \textit{bezüglich der Metrik}\\ \qquad \operatorname{dist}|p^1,p^2|_\lambda^z := \int d\lambda \left| z\left(\frac{dp^1}{d\lambda}\right) - z\left(\frac{dp^2}{d\lambda}\right)\right| + \|p^1-p^2\| \;(p^1,p^2\in M)\\ \textit{totalbeschränkt ist.}\end{cases}$

Die Metrik $\operatorname{dist}|.,.|_\lambda^z$ wurde von JACOBS in [*1*] zum Beweis einer Umkehrung des Coding-Theorems verwandt.

Lemma 9.5. *Genügt M der Forderung* (39), *so auch der Forderung* (34).

Beweis. Mit dem gleichen Schluß wie zu Anfang des Nachweises, daß (31) aus (32) folgt (im Beweis von Lemma 9.1), ergibt sich, daß

$$\sup_{p\in M}\int d\lambda\, z\left(\frac{dp}{d\lambda}\right) < \infty$$

ist, wenn M der Bedingung (39) genügt. Nach Satz 4.1 gibt es dann wegen (39) eine G-Funktion g, für die $g(\sqrt{.})$ konkav auf R^+ ist, mit

$$\sup_{p\in M}\int dp\, g\left(\log^+\frac{dp}{d\lambda}\right) < K < \infty\,.$$

Es ist

$$\int d\tilde{p}\, g(\log^+\tilde{f}) = \sum a^i \int dp^i g\left(\log^+\frac{dp^i}{dq}\right)$$

und

$$g\left(\log^+\frac{dp^i}{dq}\right) = g\left(\left(\log\frac{dp^i}{d\lambda} - \log\frac{dq}{d\lambda}\right)^+\right) \leqq g\left(\log^+\frac{dp^i}{d\lambda} + \left|\log^-\frac{dq}{d\lambda}\right|\right) \leqq$$
$$\leqq \frac{1}{2} g\left(2\log^+\frac{dp^i}{d\lambda}\right) + \frac{1}{2} g\left(2\left|\log^-\frac{dq}{d\lambda}\right|\right) \leqq \frac{4}{2} g\left(\log^+\frac{dp^i}{d\lambda}\right) + \frac{4}{2} g\left(\left|\log^-\frac{dq}{d\lambda}\right|\right)$$

(wegen der Konkavität von $g(\sqrt{.})$). Mithin ist

$$\int d\tilde{p}\, g(\log^+ \tilde{f}) \leqq 2K + 2\sum a^i \int dp^i g\left(\left|\log^- \frac{dq}{d\lambda}\right|\right) = 2K + 2\int dq\, g\left(\left|\log^- \frac{dq}{d\lambda}\right|\right).$$

Nach dem anfangs im Beweis von Lemma 9.1 Gesagten bleibt

$$\int dq\, g\left(\left|\log^- \frac{dq}{d\lambda}\right|\right)$$

unterhalb einer Schranke, die nur von g, nicht aber von q abhängt.

Mittels der Lemmata 9.5, 9.3 und 9.1 folgt sofort das

Lemma 9.6. *Genügt M der Bedingung* (40), *so genügt M der Bedingung* (36).

Lemma 9.7. *Die Bedingungen* (40) *und* (41) *sind äquivalent.*

Beweis. 1. Aus (40) folgt (41): Unter der Voraussetzung (40) gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart, daß

$$\int d\lambda \left| z\left(\frac{dp^1}{d\lambda}\right) - z\left(\frac{dp^2}{d\lambda}\right)\right| < \varepsilon \quad \text{für} \quad \|p^1 - p^2\| < \delta \qquad (p^1, p^2 \in M)$$

ist. (Man schneide dazu die Spitzen der Funktionen $z(dp/d\lambda)$ in Abhängigkeit von ε in hinreichend großer Höhe gleichmäßig ab.)

2. Aus (41) folgt (40): Ist M bezüglich $\operatorname{dist}|.,.|_\lambda^z$ totalbeschränkt, so ist M auch bezüglich $\operatorname{dist}|.,.|_{\lambda'}^z$ mit

$$\lambda' := \tfrac{1}{2}(\lambda + p') \qquad (p' \in M \text{ fest})$$

totalbeschränkt, da

$$\int d\lambda' \left| z^+\left(\frac{dp^1}{d\lambda'}\right) - z^+\left(\frac{dp^2}{d\lambda'}\right)\right| \leqq \int d\lambda \left| z^+\left(2\frac{dp^1}{d\lambda}\right) - z^+\left(2\frac{dp^2}{d\lambda}\right)\right|$$

und

$$\int d\lambda' \left| z^-\left(\frac{dp^1}{d\lambda'}\right) - z^-\left(\frac{dp^2}{d\lambda'}\right)\right| \leqq \int d\lambda' \left|\left(\frac{dp^1}{d\lambda'} - 1\right)^- - \left(\frac{dp^2}{d\lambda'} - 1\right)^-\right| \leqq \|p^1 - p^2\|$$

ist. Für das zur Konstruktion von λ' benutzte p' ist jedenfalls

$$z\left(\frac{dp'}{\frac{1}{2}(d\lambda + dp')}\right) = z\left(\frac{dp'}{d\lambda'}\right) \lambda'\text{-integrierbar.}$$

Da nun $\{z(dp/d\lambda') : p \in M\}$ eine $\|.\|_{\lambda'}^1$-totalbeschränkte Menge ist, ist

$$\left\{z\left(\frac{dp}{d\lambda'}\right) : p \in M\right\}$$

eine gleichgradig λ'-integrable Funktionenmenge. Nach Definition der Metrik $\operatorname{dist}|.,.|_\lambda^z$ in (41) ist außerdem M bedingt normkompakt.

Anmerkung 1. Wir wollen noch durch den folgenden Hinweis deutlich machen, in welcher Weise die Metrik aus der Bedingung (41) Abstände zwischen WVen mißt. Sei

$$W(y) := (1 + y)\log(1 + y) \qquad (y \geqq 0)$$

und für meßbare Funktionen f

$$\|f\|_\lambda^W := \inf_{S > 0} \frac{1}{S}\left(1 + \int d\lambda\, W(S \cdot |f|)\right)$$

gesetzt. $\|f\|_\lambda^W$ ist eine Orlicz-Norm (vgl. etwa Krasnoselskii-Ruttckii [3]). (Die Normeigenschaften von $\|\,.\,\|_\lambda^W$ lassen sich sofort nachrechnen.) Die λ-integrablen Funktionen f mit $\|f\|_\lambda^W < \infty$ bilden bezüglich dieser Norm einen Banachraum. Durch ähnliche Schlüsse wie die bisherigen läßt sich zeigen, daß die Metrik

$$\operatorname{dist}|p^1, p^2|_\lambda^W := \left\| \frac{dp^1}{d\lambda} - \frac{dp^2}{d\lambda} \right\|_\lambda^W$$

für Mengen M, die der Bedingung (40) genügen, äquivalent ist zu der in (41) definierten Metrik.

Anmerkung 2. Die sämtlichen in diesem Paragraphen erwähnten Bedingungen vererben sich von M auf $\overline{co(M)}$ und von M_1 und M_2 auf $M_1 \times M_2$.

§ 10. Zeitasymptotische Abschätzungen der maximalen Codelängen

Will man entsprechende Sätze wie in § 8 direkt mit Hilfe der Bedingung (31) herleiten (also ohne Benutzung von G-Funktionen, speziell ohne ein solches Hilfsmittel wie Lemma 8.5), so ist es naheliegend, bei den Abschätzungen mit abgeschnittenen ZVen zu arbeiten. Es stellt sich dann heraus, daß jede der beiden folgenden Bedingungen (42) und (43), die nach Cor. 6.2.1 (b) äquivalent und nach Cor. 6.3.2 (b) mindestens so schwach wie (31) sind, schon ausreichend ist für die Durchführung der Abschätzungen.

$$\text{(42)} \qquad \{\textit{Es ist } C(M) < \infty \textit{ und } \log N(M, \varepsilon) = o\left(\frac{1}{\varepsilon}\right) \qquad (\varepsilon \to 0).$$

$$\text{(43)} \qquad \{\textit{Es ist } C(M) < \infty \textit{ und } \sup_{\tilde{p} \in \tilde{M}} S\,\tilde{p}\{\log \tilde{f} > S\} \xrightarrow{S \to 0} 0.$$

Das Beispiel 11.3 wird zeigen, daß (42) und (43) echt schwächere Einschränkungen als (31) bedeuten, Beispiel 11.2 erweist sie als echt schärfere Forderung als die, daß $C(M) < \infty$ ist. Im folgenden treten die Bedingungen (42) und (43) als über die Zeit gemittelte Bedingungen auf. Die in diesem Paragraphen benutzten Beweisverfahren beruhen auf einer Koppelung der Beweisideen zu Satz 2.1 mit den Abschätzungen aus § 7 bzw. § 8.

Im Laufe der Untersuchungen dieses Paragraphen werden wir eine weitere einschränkende Bedingung an $\{M_v\}_{v \in Z}$ stellen, die gewährleistet, daß in der Serie der $C(M_v)$ einzelne $C(M_v)$ nicht zu stark dominieren, und zwar ist es die folgende Forderung an $\{M_v\}_{v \in Z}$:

$$\text{(44)} \qquad \left\{ \begin{array}{l} \textit{Zu jedem } \varepsilon' > 0 \textit{ gibt es ein } \delta > 0 \textit{ derart, daß für } J \subseteqq [1, t] \textit{ mit } |J| < \delta t \\ \textit{stets} \\ \qquad C(M_J) < \varepsilon' [C(M_{[1,t]}) + t] \qquad (\textit{gleichmäßig bezüglich } \{t\}_{t > 1}) \\ \textit{ist.} \end{array} \right.$$

Als erstes beweisen wir jetzt analog zum ersten Teil des Beweises von Satz 2.1 zwei Aussagen über die Aufschaukelung der Codelängen.

Satz 10.1 *(Aufschaukelung der Codes). Die Serie $\{M_v\}_{v \in Z}$ habe die folgende Eigenschaft: Es gebe $b, c, d > 0$ und eine Folge von Zeitpunkten $1 \leqq t_r \to \infty$ derart, daß*

$$\left| J_r := \left\{ v : 1 \leqq v \leqq t_r,\ \log N\left(M_v, \frac{c}{t_r}\right) > b\, t_r \right\} \right| \geqq d t_r$$

ist. Dann gilt

$$\sup_{\varepsilon} \limsup_{t\to\infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon) = \infty .$$

Für einen stationären Kanal bedeutet die obige Voraussetzung gerade, daß nicht $\log N(M, \varepsilon) = o(1/\varepsilon)$ $(\varepsilon \to 0)$ ist.

Beweis. Zum Beweis der Behauptung können wir $c = 1$ annehmen (Übergang von t zu ct) und $d = 1$ (die Lage von $J_r \subseteqq [1, t_r]$ spielt keine Rolle, und es ist

$$\frac{1}{|J_r|} \log N(M_{J_r}, \varepsilon) \leqq \frac{1}{dt} \log N(M_{[1,t]}, \varepsilon)) .$$

Ferner können wegen $N(co(M_{[1,t]}), \varepsilon) = N(M_{[1,t]}, \varepsilon)$ die M_v konvex angenommen werden. Sei $t (= t_r)$ gegeben und $\{(p_1^i, E_i)\}$ ein $1/t$-Code für M_1 der Länge $N > e^{bt}$,

$$q_1 := \frac{1}{N} \sum_{i=1}^{N} \bar{p}_1^i , \quad p_1^i := \tfrac{1}{2} (\bar{p}_1^i + q_1) .$$

$\{(p_1^i, E_i)\}$ ist dann ein $1/2t$-Code für M_1 der Länge N. Es ist

$$\frac{1}{N} \left| \left\{ i : q_1(E_i) > t \cdot \frac{1}{N} \left(\sum_{i=1}^{N} q_1(E_i) \right) \right\} \right| < \frac{1}{t},$$

also

$$\frac{1}{N} \left| \left\{ i : \frac{p_1^i(E_i)}{q_1(E_i)} > \frac{1}{2t^2} e^{bt} \right\} \right| > \frac{1}{2t} \left(1 - \frac{1}{t} \right) .$$

Wir schränken jetzt M_1 von B_1 ein auf den durch die E_i erzeugten Unter-BK $B_1' := {}^{B}(\{E_i\}) \subseteqq B_1$. Die von $(1/N, \ldots, 1/N)$ und $\{p_1^i\}$ erzeugte Simultanverteilung $\tilde{p}_1$ ist über dem eingeschränkten Produkt-BK ${}^{B}((\{i\} \times E_l))$ zu verstehen. Wegen $2 \tilde{f}_1 \geqq 1$ ($\tilde{p}_1$-fastüberall) gilt dann für natürliches k $(1 \leqq k \leqq t)$, wenn die anderen M_v $(1 \leqq v \leqq t)$ genau in derselben Weise wie M_1 behandelt werden:

$$\begin{aligned}
&\left(\prod_{v=1}^{t}{}^{\times} \tilde{p}_v \right) \left\{ \log \left(2^t \prod_{v=1}^{t}{}^{\times} \tilde{f}_v \right) > k \cdot \log \left(\frac{1}{2t^2} e^{bt} \right) \right\} \geqq \\
&\qquad \geqq \left(\prod_{v=1}^{t}{}^{\times} \tilde{p}_v \right) \left(\bigcup \left\{ \text{für mindestens } k \text{ der Indizes } v \text{ ist } \tilde{f}_v > \frac{1}{2t^2} e^{bt} \right\} \right) > \\
&\qquad > \sum_{j=k}^{t} \binom{t}{j} \left[\frac{1}{2t} \left(1 - \frac{1}{t} \right) \right]^j \cdot \left[1 - \frac{1}{2t} \left(1 - \frac{1}{t} \right) \right]^{t-j} = \\
&\qquad = 1 - \sum_{j=0}^{k-1} \binom{t}{j} \left[\frac{1}{2t} \left(1 - \frac{1}{t} \right) \right]^j \cdot \left[1 - \frac{1}{2t} \left(1 - \frac{1}{t} \right) \right]^{t-j} .
\end{aligned}$$

Mithin wird bei festgehaltenem k für hinreichend großes $t (= t_r)$ schließlich

$$\begin{aligned}
&\left(\prod_{v=1}^{t}{}^{\times} \tilde{p}_v \right) \left\{ \log \left(\prod_{v=1}^{t}{}^{\times} \tilde{f}_v \right) > k \cdot bt + k \log \left(\frac{1}{2t^2} \right) - t \log 2 \right\} \geqq \\
&\qquad\qquad \geqq 1 - \left(\sum_{j=0}^{k-1} \frac{2^{-j}}{j!} \right) e^{-2} - \delta \qquad (> 0) .
\end{aligned}$$

Gehen wir von den B_v' zu den B_v zurück, so vergrößern wir höchstens die Codelängen. Anwendung des Maximalcodesatzes 1.2 auf die letzte Ungleichung liefert

$$N(M_{[1,t]}, \varepsilon) > \exp\left\{k \cdot bt + k \log\left(\frac{1}{2t^2}\right) - t \log 2\right\}\left(1 - \left(\sum_{j=0}^{k-1} \frac{2^{-j}}{j!}\right)e^{-2} - \delta - \varepsilon\right);$$

Daraus ergibt sich: Zu jedem $k' > 0$ existiert ein $\varepsilon(k')$ derart, daß

$$\limsup_{t\to\infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon(k')) > k'$$

ist.

Die folgende weitere Aussage über Aufschaukelung der Codes leitet man sofort mit Hilfe des letzten Satzes her.

Satz 10.2. *Die Serie* $\{M_v\}_{v\in Z}$ *genüge der Bedingung* (44), *und mit geeigneten* $c, b' > 0$ *gelte*

$$\frac{1}{t_r} \sum_{v=1}^{t} \log N\left(M_v, \frac{c}{t_r}\right) > b'[C(M_{[1,t]}) + t_r]$$

für unendlichviele $t_r > 0$. *Dann ist*

$$\sup_{\varepsilon} \limsup_{t\to\infty} \frac{\log N(M_{[1,t]}, \varepsilon)}{C(M_{[1,t]}) + t} = \infty .$$

Beweis. Die schwache Umkehrung des Coding-Theorems (Satz 7.3) liefert

$$\log\left(\frac{\varepsilon}{2} N(M_v, \varepsilon)\right) \leqq \frac{\varepsilon}{2}\left[C(M_{[1,t]}) + \frac{1}{e}\right],$$

also

$$\frac{1}{t} \log N\left(M_v, \frac{c}{t}\right) \leqq \frac{2}{c}\left[C(M_v) + \frac{1}{e}\right] + \log \frac{c}{2t}.$$

Mit

$$J := \left\{v \in [1, t_r] : \log N\left(M_v, \frac{c}{t_r}\right) > \frac{b'}{2}[C(M_{[1,t]}) + t_r]\right.$$

gilt

$$\frac{1}{t_r} \sum_{v\in J} \log N\left(M_v, \frac{c}{t_r}\right) \leqq \frac{2}{c}\left[C(M_J) + \frac{|J|}{e}\right] + \frac{|J|}{t_r} \log \frac{c}{2t_r}.$$

Wegen (44) folgt aus $|J| < \delta t$:

$$\frac{1}{t} \sum_{v\in J} \log N\left(M_v, \frac{c}{t}\right) \leqq \frac{2}{c} \cdot \left[\varepsilon'(C(M_{[1,t]}) + t) + \frac{\delta t}{e}\right] + \delta \log \frac{c}{2t}.$$

Also existiert ein $d > 0$, so daß

$$\left|\left\{v \in [1, t_r] : \log N\left(M_v, \frac{c}{t_r}\right) > \frac{b'}{2} \cdot [C(M_{[1,t]}) + t_r]\right\}\right| > dt_r$$

ist. Setzt man noch

$$\frac{b'}{2}[C(M_{[1,t]}) + t] = bt,$$

so folgt die Behauptung aus dem vorigen Satz.

Satz 10.3 *(starke Umkehrung des Coding-Theorems).* $\{M_v\}_{v\in Z}$ *genüge der Bedingung* (44) *und ferner der folgenden Bedingung:*

$$(45)\quad \left\{\begin{array}{l} \textit{Zu jedem } b,\, c > 0 \textit{ gibt es ein } t_0 > 0 \textit{ derart, daß für } t > t_0 \\ \qquad \frac{1}{t}\sum_{v=1}^{t} \log N\left(M_v, \frac{c}{t}\right) < b\,[C(M_{[1,t]}) + t] \\ \textit{ist.}\end{array}\right.$$

Dann gilt für jedes feste ε $(0 < \varepsilon < 1)$

$$\log N(M_{[1,t]}, \varepsilon) \leqq C(M_{[1,t]})\,(1 + o(1)) + t \cdot o(1)\,.$$

Beweis. 1. Wir zeigen zunächst, daß (44) und (45)

$$(46)\quad \left\{\begin{array}{l} \sum_{v=1}^{t} \sup_{\tilde p_v \in \tilde M_v} \tilde p_v\{\log \tilde f_v > o_t(1)\,[C(M_{[1,t]}) + t]\} \xrightarrow{t\to\infty} 0 \\ \textit{mit einer geeigneten Funktion } o_t(1)\end{array}\right.$$

zur Folge haben: Wegen (45) existiert eine Funktion $h(t) = o(1)$ mit

$$\frac{1}{t}\sum_{v=1}^{t} \log N\left(M_v, \frac{h(t)}{t}\right) \leqq h(t)\,[C(M_{[1,t]}) + t]\,.$$

Da dann für $K > 0$

$$\frac{1}{t}\left|\left\{v \in [1,t] : \log N\left(M_v, \frac{h(t)}{t}\right) < K \cdot h(t)\,[C(M_{[1,t]}) + t]\right\}\right| > 1 - \frac{1}{K}$$

ist, gibt es eine Funktion $\bar h(t) = o(1)$ mit

$$\frac{1}{t}\left| J_t := \left\{v \in [1,t] : \log N\left(M_v, \frac{\bar h(t)}{t}\right) < \bar h(t)\,[C(M_{[1,t]}) + t]\right\}\right| \xrightarrow{t\to\infty} 1\,.$$

Wir setzen

$$J_t' := [1,t] - J_t\,.$$

Sei $h'(t) = o(1)$ eine weitere Funktion mit

$$\frac{h'(t)}{\bar h(t)} \to \infty \quad (t \to \infty) \quad \text{und} \quad h'(t) > \frac{1}{t}\,.$$

Dann erhält man mit dem Maximalcode-Satz für $v \in J_t$:

$$\bar h(t)\,[C(M_{[1,t]}) + t] > \log N\left(M_v, \frac{\bar h(t)}{t}\right) > h'(t)\,[C(M_{[1,t]}) + t] +$$
$$+ \log\left[\tilde p_v\{\log \tilde f_v > h'(t)\,[C(M_{[1,t]}) + t]\} - \frac{\bar h(t)}{t}\right]^+,$$

mithin gilt gleichmäßig für $v \in J_t$:

$$\tilde p_v\{\log \tilde f_v > h'(t)\,[C(M_{[1,t]}) + t]\} \leqq \frac{1}{t} \cdot o_t(1) \quad (\text{gleichmäßig für } \tilde p_v \in \tilde M_v)\,,$$

und für $v \in J_t'$ ist

$$\tilde p_v\{\log \tilde f_v > h'(t)\,[C(M_{[1,t]}) + t]\} \leqq \frac{1}{h'(t)\,[C(M_{[1,t]}) + t]} \int d\tilde p_v \log^+ \tilde f_v\,.$$

Mit (44) folgt dann

$$\frac{1}{t}\sum_{v\in J'_t}\tilde{p}_v\{\log\tilde{f}_v > h'(t)\,[C(M_{[1,t]})+t]\}\xrightarrow{t\to\infty}0 \quad (\text{glm. bzgl. } \tilde{p}_v\in\tilde{M}_v).$$

Daraus ergibt sich (46).

2. Die Codelängenabschätzungen nach oben führen wir nun mit Hilfe von (46) durch. Zur Abkürzung setzen wir

$$r(t):=o_t(1)\,[C(M_{[1,t]})+t].$$

Sei $\{(p^i, E_i)\}$ ein ε-Code der Länge N für $M_{[1,t]}$,

$$q:=q_1\times\cdots\times q_t:=\left(\frac{1}{N}\sum_{i=1}^{N}p_1^i\right)\times\cdots\times\left(\frac{1}{N}\sum_{i=1}^{N}p_t^i\right)$$

und

$$E'_i:={}_1E'_i\times\cdots\times{}_tE'_i=\left\{\log\frac{dp_1^i}{dq_1}<r(t)\right\}\times\cdots\times\left\{\log\frac{dp_t^i}{dq_t}<r(t)\right\}.$$

Im folgenden schätzen wir bei laufendem t ab (— in § 8 hatten wir die Abschätzungen jeweils für festes t durchgeführt —), in dem Abschätzungsprozeß treten also in Abhängigkeit von t ständig andere Codierungs-WVen auf, die wir jedoch nicht durch neue Indizes kenntlich machen wollen. Für jedes $K>0$ ist

$$p^i\left(E'_i\cap\left\{\frac{dp^i}{dq}>\left(\frac{3}{\varepsilon}\right)^K\left(\int_{E'_i}dp^i\left(\frac{dp^i}{dq}\right)^{1/K}\right)\right\}\right)<\frac{\varepsilon}{3},$$

also

$$p^i\left\{\log\frac{dp^i}{dq}>K\log\frac{3}{\varepsilon}+K\sum_{v=1}^{t}\log\left(\int_{{}_vE'_i}dp_v^i\left(\frac{dp_v^i}{dq_v}\right)^{1/K}\right)\right\}<\frac{\varepsilon}{3}+p^i(\operatorname{compl}E'_i). \tag{47}$$

Da nach (46)

$$\sum_{v=1}^{t}\frac{1}{N}\sum_{i=1}^{N}p_v^i\left\{\log\frac{dp_v^i}{dq_v}>r(t)\right\}\xrightarrow{t\to\infty}0$$

ist, bekommt man für geeignetes $o_t(1)$:

$$\frac{1}{N}\left|F':=\left\{i:\sum_{v=1}^{t}p_v^i\left\{\log\frac{dp_v^i}{dq_v}>r(t)\right\}<o_t(1)\right\}\right|\xrightarrow{t\to\infty}1.$$

Zu jedem $\delta>0$ gibt es dann ein t_0 derart, daß für $t>t_0$

$$\prod_{v=1}^{t}\left(1-p_v^i\left\{\log\frac{dp_v^i}{dq_v}>r(t)\right\}\right)=p^i(E'_i)>1-\delta \qquad (i\in F'(t))$$

gilt. Mit (47) zusammen bedeutet das

$$\frac{1}{N}\left|F:=\left\{i:p^i\left\{\log\frac{dp^i}{dq}>K\log\frac{\varepsilon}{3}+K\sum_{v=1}^{t}\log\left(\int_{{}_vE'_i}dp_v^i\left(\frac{dp_v^i}{dq_v}\right)^{1/K}\right)\right\}<\frac{2}{3}\varepsilon\right\}\right|\xrightarrow{t\to\infty}1.$$

Für hinreichend großes t erhält man daher mit (20)

$$\begin{aligned}\frac{1}{|F|}&\geq\left(\prod_{i\in F}q(E_i)\right)^{1/|F|}>\\&>\frac{\varepsilon}{3}\exp\left\{-K\log\frac{3}{\varepsilon}-K\sum_{v=1}^{t}\frac{1}{|F|}\sum_{i\in F}\log\left(\int_{{}_vE'_i}dp_v^i\left(\frac{dp_v^i}{dq_v}\right)^{1/K}\right)\right\}.\end{aligned} \tag{48}$$

Wählt man nun K in Abhängigkeit von t derart, daß

$$K = K(t) = o_t(1)\,[C(M_{[1,t]}) + t] \quad \text{und}\ \frac{K(t)}{r(t)} \to \infty \qquad (t \to \infty)$$

ist, so gilt ersichtlich

$$\int\limits_{vE'_i} dp^i_v \left(\frac{dp^i_v}{dq_v}\right)^{1/K} \leqq \left(1 + \frac{1}{K}\int\limits_{vE'_i} dp^i_v K\left[\left(\frac{dp^i_v}{dq_v}\right)^{1/K} - 1\right]\right)$$
$$\leqq 1 + \frac{1}{K}\int\limits_{vE'_i} dp^i_v \left(\log \frac{dp^i_v}{dq_v}\right)(1 + o_t(1))$$

(glm. für alle p^i_v zu festem t) und ebenfalls gleichmäßig für alle i zu festem t

$$K \log\left(1 + \frac{1}{K}\int\limits_{vE'_i} dp^i_v \left(\log\frac{dp^i_v}{dq_v}\right)(1 + o_t(1))\right) \leqq \int dp^i_v \left(\log \frac{dp^i_v}{dq_v}\right)(1 + o_t(1)).$$

Setzt man diese Ungleichungen mit der Ungleichung (48) zusammen, so liefert das unter Beachtung von

$$\frac{N}{|F|} \to 1 \qquad (t \to \infty)$$

die Behauptung des Satzes.

Den beiden letzten Sätzen entnimmt man speziell das

Corollar 10.2.1, 10.3.1. *Für einen stationären gedächtnisfreien Kanal mit* $C(M_1) < \infty$ *gilt genau dann*

$$\limsup_{t \to \infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon) \leqq C(M_1) \quad \textit{für jedes} \quad \varepsilon\,(0 < \varepsilon < 1),$$

wenn

$$\log N(M_1, \varepsilon) = o\left(\frac{1}{\varepsilon}\right) \qquad (\varepsilon \to 0)$$

ist.

Zum Schluß dieses Paragraphen betrachten wir nun noch eine Bedingung für die Gültigkeit des Coding-Theorems (die natürlich nur für nichtstationäre Kanäle interessant ist).

Satz 10.4. $\{M_v\}_{v \in Z}$ *genüge der Forderung* (46). *Dann ist die Bedingung*

$$(49)\quad \left\{\begin{array}{l} \textit{Mit} \\ \qquad A(t,b) := \sum\limits_{v=1}^{t} \sup\limits_{\tilde p_v \in \tilde M_v} \int\limits_{\{\log \tilde f_v < b[C(M_{[1,t]})+t]\}} d\tilde p_v \log \tilde f_v \\ \textit{gilt für jedes } b > 0: \\ [C(M_{[1,t]}) + t]^{-1} \cdot [C(M_{[1,t]}) - A(t,b)] \xrightarrow{t \to \infty} 0 \end{array}\right.$$

notwendig und hinreichend dafür, daß es zu jedem $\delta > 0$ *ein* $t_0(\varepsilon)$ *gibt, so daß für* $t > t_0(\varepsilon)$

$$\log N(M_{[1,t]}, \varepsilon) > C(M_{[1,t]})(1 - \delta) - \delta t$$

ist.

Wir begnügen uns mit einer

Beweisskizze. 1. Wenn (46) erfüllt ist, so gilt nach dem Beweis von Satz 10.3

mit einer geeigneten Funktion $o_t(1)$:

$$\log N(M_{[1,t]}, \varepsilon) < A(t, b)(1 + o_t(1)) + t\, o_t(1).$$

Daraus ersieht man, daß (49) notwendig ist für die Gültigkeit der obigen unteren Codelängenabschätzung.

2. Daß (49) mit (46) auch hinreichend für diese Codelängenabschätzung ist, ergibt sich folgendermaßen: Mit $\tilde{p}_{[1,t]} = \tilde{p}_1 \times \cdots \times \tilde{p}_t$ ist für $\varepsilon + b' < 1$

$$\tilde{p}_{[1,t]} \left\{ \tilde{f}_{[1,t]}^{-1} < \left(\frac{1}{1 - \varepsilon - b'} \int_{\tilde{E}} d\tilde{p}_{[1,t]} \tilde{f}^{-1/K} \right)^K \right\} > \varepsilon + b' - \tilde{p}_{[1,t]}(\operatorname{compl} \tilde{E}),$$

also

$$\tilde{p}_{[1,t]} \Big\{ \log \tilde{f}_{[1,t]} > -K \log \frac{1}{1 - \varepsilon - b'} - \\ - K \log \Big(\int_{\tilde{E}} d\tilde{p}_{[1,t]} \tilde{f}^{-1/K}\Big) \Big\} > \varepsilon + b' - \tilde{p}_{[1,t]}(\operatorname{compl} \tilde{E}).$$

Wählt man nun $\tilde{E}$ speziell als

$$\tilde{E} := \{\log \tilde{f}_1 < b[C(M_{[1,t]}) + t], \ldots, \log \tilde{f}_t < b[C(M_{[1,t]}) + t]\},$$

wobei sich b in Abhängigkeit von t so ändere, daß

$$b = b(t) = o_t(1)$$

ist und (49) und (46) mit $b(t)$ erfüllt bleiben, so gilt

$$\tilde{p}_{[1,t]}(\operatorname{compl} \tilde{E}) \to 0 \ (t \to \infty) \quad (\text{glm. für } \tilde{p}_{[1,t]} \in \{\tilde{p}_{[1,t]}\});$$

ferner läßt sich wieder $K = K(t)$ wie im Beweis von Satz 10.3 so angeben, daß

$$K(t) = o_t(1)\,[C(M_{[1,t]}) + t]$$

und

$$-K \log \Big(\int_{\tilde{E}} d\tilde{p}_{[1,t]} \tilde{f}^{-1/K}\Big) > \int_{\tilde{E}} d\tilde{p}_{[1,t]} \log \tilde{f}(1 - o_t(1))$$

ist. Beachtet man nun noch, daß mit (49) für jedes $\delta > 0$

$$\sup_{\tilde{p}_{[1,t]}} \int_{\tilde{E}} d\tilde{p}_{[1,t]} \log \tilde{f}(1 - o_t(1)) > C(M_{[1,t]})(1 - \delta) - \delta t$$

ist, sobald t hinreichend groß ist, so erhält man nach Anwendung des Maximalcode-Satzes die behauptete Aussage.

Anmerkung: In den Beweisen dieses Paragraphen ist nirgends von der speziellen Lage der Zeitpunkte in Z Gebrauch gemacht worden (das gleiche gilt auch für die Beweise des § 8). Deshalb lassen sich die letzten Sätze auch für Zeiträume mit Lücken formulieren. Zieht man dann die Aussagen der Sätze 10.2 bis 10.4 zusammen, so ergibt sich die folgende Aussage:

Es sei $\mathfrak{S}$ ein System von endlichen Teilmengen $J \subseteqq Z$, das Mengen J mit beliebig großer Mächtigkeit $|J|$ enthalte. Ferner sei eine Serie $\{M_v\}_{v \in Z}$ vorgegeben, und es gelte: Zu jedem $\varepsilon' > 0$ existiert ein $\delta > 0$ derart, daß für $I \subseteqq J \in \mathfrak{S}$ mit $|I| < \delta |J|$ stets

$$C(M_I) < \varepsilon'[C(M_J) + |J|] \qquad (\text{glm. für } J \in \mathfrak{S})$$

ist. Notwendig und hinreichend dafür, daß für $J \in \mathfrak{S}$

$$|\log N(M_J, \varepsilon) - C(M_J)| \leqq o_{|J|}(1)\,[C(M_J) + |J|] \qquad (|J| \to \infty)$$

gilt, sind dann die beiden folgenden Bedingungen zusammen:

(a) Zu jedem $b, c > 0$ gibt es ein $t_0 > 0$, so daß für $J \in \mathfrak{S}$ mit $|J| > t_0$

$$\frac{1}{|J|} \sum_{v \in J} \log N\left(M_v, \frac{c}{|J|}\right) < b\,[C(M_J) + |J|]$$

gilt.

(b) Für jedes $b > 0$ gibt es ein $t_0 > 0$ derart, daß für $J \in \mathfrak{S}$ mit $|J| > t_0$

$$\sum_{v \in J} \sup_{\tilde{p}_v \in \tilde{M}_v} \int_{\{\log \tilde{f}_v < b[C(M_J) + |J|]\}} d\tilde{p} \log \tilde{f} > C(M_J)\,(1-b) - b\,|J|$$

gilt.

§ 11. Beispiele

Die Bedeutung und Stärke der in den letzten Paragraphen betrachteten Bedingungen wird zum Abschluß noch durch einige Beispiele erläutert. Insbesondere zeigt sich, daß es stationäre gedächtnisfreie Kanäle gibt mit endlicher Kapazität $C(M_1)$, für die die starke Umkehrung des Coding-Theorems nicht gilt, und daß es nichtstationäre gedächtnisfreie Kanäle gibt mit gleichmäßig beschränkten Kapazitäten $C(M_v)$, für die das Coding-Theorem nicht erfüllt ist.

Beispiel 11.1. *(Beispiel für eine nicht bedingt normkompakte Menge M, die der Bedingung (33) genügt.)*

Sei $X = [0,1]$, B der natürliche BK auf $[0,1]$, λ das Lebesguemaß und

$$M = \{p^j\}\ (j = 1, 2, \ldots),$$

wobei p^j definiert sei durch die folgende Dichte: für

$$j = 2i: \quad \frac{dp^{2i}}{d\lambda} := \begin{cases} 2 & \text{auf } [(2k-1)(\tfrac{1}{2})^i, 2k(\tfrac{1}{2})^i] \quad \text{für} \quad k = 1, \ldots, 2^{i-1} \\ 0 & \text{sonst,} \end{cases}$$

$$j = 2i - 1: \quad \frac{dp^{2i-1}}{d\lambda} := 2 - \frac{dp^{2i}}{d\lambda}.$$

Für $j_1 \neq j_2$ ist $\|p^{j_1} - p^{j_2}\| \geqq 1$, also ist M nicht normtotalbeschränkt. M genügt der Bedingung (39), folglich (33).

Anmerkung: Für den von $\{M_v : M_v = M\}_{v \in Z}$ erzeugten stationären gedächtnisfreien Kanal gilt eine besonders scharfe Form des Coding-Theorems und der starken Umkehrung:

1. Es ist

$$C(M) = \log 2,$$

denn es ist

$$\int d\tilde{p} \log \tilde{f} = \int d\tilde{p} \log \frac{d\tilde{p}}{da \times d\lambda} - \int d\tilde{p} \log \frac{da \times dq}{da \times d\lambda} \leqq \int d\tilde{p} \log \frac{d\tilde{p}}{da \times d\lambda} \leqq \log 2 \quad (\tilde{p} \in \tilde{M})$$

und

$$C(\{p^{2i}, p^{2i-1}\}) \geqq \tfrac{1}{2} \int dp^{2i} \log \frac{dp^{2i}}{d\lambda} + \tfrac{1}{2} \int dp^{2i-1} \log \frac{dp^{2i-1}}{d\lambda} = \log 2 \quad (p^{2i}, p^{2i-1} \in M).$$

2. Da p^{2i} und p^{2i-1} trägerfremde WVen sind, ist sicher

$$N(M_{[1,t]}, \varepsilon) \geqq 2^t, \quad \text{also} \quad \log N(M_{[1,t]}, \varepsilon) \geqq t\,C(M).$$

3. Sei $p \in M_{[1,t]}$, $E \in B_{[1,t]}$ und $p(E) > \varepsilon$. Dann ist

$$(\lambda \times \cdots \times \lambda)(E) > \frac{\varepsilon}{2^t}, \quad \text{folglich} \quad N(M_{[1,t]}, \varepsilon) < \frac{2^t}{\varepsilon},$$

d. h.

$$\log N(M_{[1,t]}, \varepsilon) < t\,C(M) - \log \varepsilon.$$

In den folgenden beiden Beispielen 11.2 und 11.3 bedeute X das halboffene Intervall [0,1), B den natürlichen BK über [0,1), λ das Lebes0uemaß; ferner sei T^y die durch

$$T^y : x \to x - y \pmod 1$$

definierte Abbildung von [0,1) auf sich. Für WVen p^0, $p^{0,s}$ sei

$$p^y := p^{0+y} := p^0\,T^y, \qquad M(p^0) := \{p^y : y \in [0,1)\}$$
$$p^{y,s} := p^{0,s}\,T^y, \qquad M(p^{0,s}) := \{p^{y,s} : y \in [0,1)\}$$

gesetzt. Für die weiteren Konstruktionen machen wir Gebrauch von dem folgenden

Lemma. *Sei p^0 WV auf* [0,1) *mit $p^0 \ll \lambda$. Dann gilt:*

1.

$$C(M(p^0)) \leqq \int dp^0 \log \frac{dp^0}{d\lambda}.$$

2. *Ist für ein $E \in B$*

$$p^0(E) < \delta \quad \textit{und} \quad \int_E dp^0 \log \frac{dp^0}{d\lambda} > \varepsilon > 0,$$

so existiert ein $\tilde{p} \in \tilde{M}(p^0)$ und dazu ein $\tilde{E}$ mit

$$\tilde{p}(\tilde{E}) < \delta \quad \text{und} \quad \int_{\tilde{E}} d\tilde{p} \log \tilde{f} > \varepsilon.$$

(Aus 1. und 2. ergibt sich noch $C(M(p^0)) = \int dp^0 \log dp^0/d\lambda$.)

Beweisskizze. Zu 1.: Es ist

$$C(M(p^0)) \leqq \sup_{\tilde{p} \in \tilde{M}(p^0)} \int d\tilde{p} \log \frac{d\tilde{p}}{da \times d\lambda} = \int dp^0 \log \frac{dp^0}{d\lambda}.$$

Zu 2.: Da $M(p^0)^{glgr} \ll \lambda$ ist und $\lambda(T^y E \cup E - T^y E \cap E) \to 0$ $(y \to 0)$ (glm. für $E \in B$) gilt, ergibt sich mittels eines Skalentricks sofort die Normkompaktheit von $M(p^0)$, genauer: Zu jedem $\varepsilon' > 0$ gibt es ein n so, daß

$$\|p^{y_1} - p^{y_2}\| < \varepsilon' \quad \text{ist, falls} \quad |y_1 - y_2| < \frac{1}{n} \text{ ist.}$$

Man teile nun [0,1) in Intervalle der Länge $1/n$ ein und nehme aus jedem der nTeilungsintervalle genau ein y, aus dem i-ten Teilungsintervall y_i. Mit Standardüberlegungen gilt dann

$$\left\| \frac{1}{n} \sum_{i=1}^{n} p^{y_i} - \lambda \right\| \xrightarrow{n} 0,$$

also

$$\left| \frac{dp^{y_k}}{\frac{1}{n} \sum_{i=1}^{n} dp^{y_i}} - \frac{dp^{y_k}}{d\lambda} \right| \xrightarrow{n} 0 \quad (\lambda\text{-stochastisch}).$$

Hieraus entnimmt man, daß für jedes $K > 0$

$$\frac{1}{n} \sum_{i=1}^{n} \int_{F(E, y_i, K)} dp^{y_i} \log \frac{dp^{y_i}}{\frac{1}{n} \sum_{k=1}^{n} dp^{y_k}} \xrightarrow{n} \int_{E \cap \left\{\frac{dp^0}{d\lambda} < K\right\}} dp^0 \log \frac{dp^0}{d\lambda}$$

gilt, wobei

$$F(E, y_i, K) := \left(T^{-y_i} E \cap \left\{ \frac{dp^{y_i}}{\frac{1}{n} \sum_{i=1}^{n} dp^{y_i}} < K \right\} \right)$$

gesetzt sei. Mithin ergibt sich die Behauptung.

Beispiel 11.2 (*Beispiel für eine Menge M mit $C(M) < \infty$ und* $\limsup_{\varepsilon \to 0} \varepsilon \log N(M, \varepsilon) > 0$).

Sei

$$M := \bigcup_{s \in [e, \infty)} M(p^{0,s})$$

mit

$$\frac{dp^{0,s}}{d\lambda}(x) := \begin{cases} s & \text{für} \quad x \in \left[0, \frac{1}{s \log s}\right] \\ 1 - \frac{s-1}{s \log s - 1} & \text{für} \quad x \in \left(\frac{1}{s \log s}, 1\right) \end{cases}.$$

Wegen

$$\| p^{0,s} - \lambda \| \xrightarrow{s \to \infty} 0 \quad \text{und} \quad \| p^{y,s} - p^{0,s} \| \xrightarrow{y \to 0} 0$$

(glm. für $s \in [e, \infty)$) ergibt sich, daß M bedingt normkompakt ist und durch Hinzunahme von λ kompakt wird. Man errechnet ferner unter Benutzung des Lemmas:

$$C(M) = \sup_{s \in [e, \infty)} \int dp^{0,s} \log \frac{dp^{0,s}}{d\lambda} = 1 .$$

Da

$$p^{0,s}\left(\left[0, \frac{1}{s \log s}\right]\right) = \frac{1}{\log s}$$

ist, erhält man (für $s \log s = n \in Z^+$)

$$N\left(M, \frac{1}{1 + \log s}\right) \geqq s \log s ,$$

also gibt es noch beliebig kleine $\varepsilon > 0$ mit

$$N(M, \varepsilon) \geqq \left(\frac{1}{\varepsilon} - 1\right) \exp\left\{\frac{1}{\varepsilon} - 1\right\}.$$

Anmerkung 1. Die Menge M aus Beispiel 11.2 erfüllt nicht die Bedingung (42). Mithin gilt für den von $\{M_v : v \in Z, M_v = M\}$ erzeugten stationären gedächtnisfreien Kanal nach Satz 10.1 nicht die starke Umkehrung des Coding-Theorems.

Anmerkung 2. Genügt eine Menge M' der Bedingung (31), so kann

$$\varepsilon \log N(M', \varepsilon)$$

beliebig langsam mit $\varepsilon \to 0$ gegen Null gehen, wie man nun leicht aus Beispiel 11.2 entnimmt:

Sei z. B. g eine G-Funktion mit $g(y)/y^2 \to 0$ $(y \to \infty)$ und

$$M' := \{\bar{p}^{y,s} : d\bar{p}^{y,s} = f^{y,s}\, d\lambda\}.$$

Hierbei sei

$$f^{y,s} := g^{-1}\left(\frac{dp^{y,s}}{d\lambda}\right) + \mathrm{const}(s)$$

mit $dp^{y,s}/d\lambda$ aus Beispiel 11.2 und $\mathrm{const}(s)$ derart, daß

$$\int d\lambda\, f^{y,s} = 1$$

ist. Eine einfache Umrechnung liefert

$$\int d\bar{p}^{y,s}\, g(|\log f^{y,s}|) \leqq c_1 \cdot \sup_s \int dp^{0,s} \log \frac{dp^{0,s}}{d\lambda} + c_2,$$

$$(c_1, c_2 \text{ geeignete Konstanten} > 0)$$

also gilt (31) für M'. Ferner erhält man für ein $c > 0$ $\log N(M', \varepsilon_j) > g^{-1}(c/\varepsilon_j)$ für eine Nullfolge $\{\varepsilon_j\}$.

Beispiel 11.3 *(Beispiel für eine Menge M mit $\log N(M, \varepsilon) = o(1/\varepsilon)$ $(\varepsilon \to 0)$, $C(M) < \infty$ und $\sup_{\tilde{p} \in \tilde{M}} \int d\tilde{p}\, g(|\log f|) = \infty$ für jede G-Funktion g).*

Die Konstruktion von M verläuft folgendermaßen: Wir geben zuerst eine WV p^0 auf $[0,1)$ an mit

$$\log N(M(p^0), \varepsilon) = o\left(\frac{1}{\varepsilon}\right) \quad \text{und} \quad C(M(p^0)) = \int dp^0 \log \frac{dp^0}{d\lambda} = \infty.$$

Mit Hilfe von p^0 werden dann WVen $p^{0,s}$ $(s \in (0, \frac{1}{2}])$ erzeugt, für die

$$\sup_{s \in (0, \frac{1}{2}]} \int dp^{0,s} \log \frac{dp^{0,s}}{d\lambda} < \infty,$$

$$p^{0,s}([0, s]) \xrightarrow{s \to 0} 0, \quad \text{aber} \quad \limsup_{s \to 0} \int_{[0,s]} dp^{0,s} \log \frac{dp^{0,s}}{d\lambda} \geqq \mathrm{const} > 0$$

ist, und für die außerdem das Wachstum der Dichten $dp^{0,s}/d\lambda$ majorisiert wird (glm. für $s \in (0, \frac{1}{2}]$) durch das Wachstum von $dp^0/d\lambda$.

$$M := \bigcup_{s \in (0, \frac{1}{2}]} M(p^{0,s})$$

hat die geforderten Eigenschaften.

1. Konstruktion von p^0: Sei p die WV auf $[0,1)$, die bestimmt ist durch

$$p\left(\left[0, \exp\left\{-\frac{1}{\varepsilon} + 1\right\}\right]\right) = \varepsilon \qquad (0 \leqq \varepsilon < 1).$$

Es ist

$$C(M(p)) = \int dp \log \frac{dp}{d\lambda} = \infty;$$

denn da es mindestens $\exp\{+1/\varepsilon - 1\} - 1$ paarweise disjunkte Teilintervalle $E \subseteq [0,1)$ der Länge $\exp\{-1/\varepsilon + 1\}$ mit $p\,T^y(E) > \varepsilon/2$ für wenigstens ein $y \in [0,1)$ gibt, ist

$$N\left(M(p), \frac{\varepsilon}{2}\right) + 1 > \exp\left\{\frac{1}{\varepsilon} - 1\right\},$$

deshalb ist $C(M(p)) = \int dp \log dp/d\lambda = \infty$, denn anderenfalls erfüllte $M(p)$ nach dem obigen Lemma die Bedingung (34), und es wäre somit $\log N(M(p), \varepsilon) = o(1/\varepsilon)$. Aus

$$p([0,x]) = \frac{1}{1 - \log x}$$

ersieht man, daß $(dp/d\lambda)(x)$ in einem Intervall $[0,c]$ monoton fällt und für $x > c$ gleichmäßig beschränkt ist. Sei $dp'/d\lambda$ monoton fallende WV-Dichte in $[0,1)$ mit

$$\frac{dp'}{d\lambda}(x) = \frac{dp}{d\lambda}(x) \quad \text{für kleine } x.$$

Dann ist

$$\int dp' \log \frac{dp'}{d\lambda} = \int d\lambda\, z\left(\frac{dp'}{d\lambda}\right) = \infty .$$

Nun gibt es sicherlich eine monoton wachsende Funktion $m(y)$ ($\geqq 0$) auf R^+ mit $1 \geqq m(y)/y \to 0$ $(y \to \infty)$, für die noch

$$\int d\lambda\, z\left(m\left(\frac{dp'}{d\lambda}\right)\right) = \infty$$

ist. Sei

$$f^0(x) := m\left(\frac{dp'}{d\lambda}(x)\right) + d, \quad d \geqq 0 \quad \text{derart, daß} \quad \int d\lambda\, f^0 = 1 \quad \text{ist.}$$

Dann ist

$$\int d\lambda\, z(f^0) = \infty,$$

und für die durch

$$dp^0 := d\lambda\, f^0$$

definierte WV p^0 ist

$$p^0\left(\left[0, \exp\left\{-\frac{1}{\varepsilon} + 1\right\}\right]\right) = o(\varepsilon) \qquad (\varepsilon \to 0)$$

(also $p^0([0,x]) = \bar{h}(x)/(1 - \log x)$ mit $\bar{h}(x) \to 0$ $(x \to 0)$). Da f^0 monoton fällt, ergibt sich

$$p^0(E) = \int_E d\lambda\, f^0 \leqq \int_{[0,\lambda(E)]} d\lambda\, f^0 = p^0([0, \lambda(E)]) \qquad (E \in B).$$

Aus den letzten Zeilen erhält man für jeden ε-Code $\{(p^i, E_i)\}_{i=1,\dots,N}$ von $M(p^0)$:

$$\varepsilon < p^i(E_i) \leqq p^0([0, \lambda(E_i)]) \leqq \frac{\bar{h}(\lambda(E_i))}{1 - \log \lambda(E_i)},$$

mithin

$$\log \frac{1}{\lambda(E_i(\varepsilon))} = o\left(\frac{1}{\varepsilon}\right), \quad \text{also } \log N \leqq o\left(\frac{1}{\varepsilon}\right).$$

Ferner ist

$$C(M(p^0)) = \int d\lambda\, z(f^0) = \infty .$$

2. Mittels p^0 bzw. f^0 konstruieren wir nun M wie folgt: Sei $C > 0$ vorgegeben und $f^{0,s}(0 < s \leqq \frac{1}{2})$ definiert durch

$$f^{0,s}(x) := \begin{cases} \min(f^0(x), K_s) & \text{für} \quad x \in [0, s] \\ \bar{K}_s & \text{für} \quad x \in (s, 1) \end{cases},$$

wobei die Konstante K_s so bestimmt wird, daß

$$\int_{[0,s]} d\lambda \, z(f^{0,s}) = C$$

ist (die Existenz einer derartigen Konstanten $K_s (> 1)$ für jedes $s \in (0, \frac{1}{2}]$ folgt aus $\int dp^0 \log f^0 = \infty$), die Konstante $\bar{K}_s$ nun so, daß

$$\int d\lambda \, f^{0,s} = 1$$

ist (es ist $0 \leqq \bar{K}_s < 2$). Sei

$$dp^{0,s} := d\lambda \, f^{0,s}.$$

Es gilt $p^0([0, s]) \to 0$ $(s \to 0)$ und daher nach Konstruktion der Dichten $f^{0,s}$

$$p^{0,s}([0, s]) \to 0 \quad (s \to 0).$$

Damit ist folgendes gezeigt: Zu jedem $\delta > 0$ gibt es ein s mit

$$p^{0,s}([0, s]) < \delta \quad \text{und} \quad \int_{[0,s]} dp^{0,s} \log f^{0,s} = C,$$

oder anders: zu jedem $\delta > 0$ gibt es ein s und dazu $\tilde{p} \in \tilde{M}(p^{0,s})$ und $\tilde{E}$ (aus dem Definitionsbereich von $\tilde{p}$) mit

$$\tilde{p}(\tilde{E}) < \delta \quad \text{und} \quad \int_{\tilde{E}} d\tilde{p} \log \tilde{f} \approx C$$

(vgl. Lemma). Die Menge

$$M := \bigcup_{s \in (0, \frac{1}{2}]} M(p^{0,s})$$

genügt also nicht der Bedingung (31). Es ist ferner

$$C(M) = \sup_{s \in (0, \frac{1}{2}]} \int dp^{0,s} \log f^{0,s} \leqq \sup_{s \in (0, \frac{1}{2}]} \int_{[0,s]} dp^{0,s} \log f^{0,s} + \log 2 = C + \log 2.$$

Zu zeigen ist noch

$$\log N(M, \varepsilon) = o\left(\frac{1}{\varepsilon}\right).$$

Dies folgt daraus, daß für jeden ε-Code $\{(p^i, E_i)\}_{i=1,\ldots,N}$ von M

$$\begin{aligned} \varepsilon < p^i(E_i) &\leqq p^0(E_i) + 2\lambda(E_i) \leqq p^0([0, \lambda(E_i)]) + 2\lambda(E_i) \\ &\leqq \frac{\bar{h}(\lambda(E_i))}{1 - \log \lambda(E_i)} + 2\lambda(E), \quad \text{also} \quad \log N \leqq o\left(\frac{1}{\varepsilon}\right) \end{aligned}$$

ist. Durch Vergleich mit dem letzten Beispiel sieht man noch, daß die hier konstruierte Menge M ebenfalls bedingt normkompakt ist und durch Hinzunahme von λ kompakt wird.

Beispiel 11.4 (*Beispiel für einen nichtstationären gedächtnisfreien Kanal mit* $C(M_v) = 1$ $(v \geqq 1)$ *und* $\limsup_{t\to\infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon) = 0$ $(0 < \varepsilon < 1)$).

Sei M die in Beispiel 11.2 betrachtete Menge. Es ist

$$C(\{p \in M : \| p - \lambda \| < \delta\}) = 1 \quad \text{für jedes } \delta > 0 .$$

Für $v \geqq 1$ setze man

$$M_v := \{p \in M : \| p - \lambda \| < 2^{-v}\} .$$

$\prod\limits_{v=1}^{\infty}{}^{\times} M_v$ ist bedingt normkompakt (jedes M_v ist normtotalbeschränkt, also auch jedes Produkt $M_{[1,t]}$, und es gilt

$$\text{Durchm}\,(\prod_{v=t}^{\infty}{}^{\times} M_v) \leqq \sum_{v=t}^{\infty} \text{Durchm}\,(M_v) < \sum_{v=t}^{\infty} 2^{-v} \xrightarrow{t\to\infty} 0) .$$

Folglich ist (nach § 3)

$$N(\prod_{v=1}^{\infty}{}^{\times} M_v, \varepsilon) < \infty \qquad (0 < \varepsilon < 1)$$

und wegen

$$\log N(M_{[1,t]}, \varepsilon) \leqq N(\prod_{v=1}^{\infty}{}^{\times} M_v, \varepsilon)$$

sicherlich

$$\limsup_{t\to\infty} \frac{1}{t} \log N(M_{[1,t]}, \varepsilon) = 0 \qquad (0 < \varepsilon < 1) .$$

Aber es ist

$$C(M_{[1,t]}) = t\, C(M_1) = t .$$

Das Coding-Theorem gilt also nicht.

Anmerkung: Setzt man mit M aus Beispiel 11.3

$$M_v := \{p \in M : \| p - \lambda \| < 2^{-v}\} \qquad (v \geqq 1) ,$$

so ist $C(M_v) \geqq C > 0$ $(v \geqq 1)$ (C die Konstante aus Beispiel 11.3). Wie in Beispiel 11.4 ist das Coding-Theorem ungültig. Da die Bedingung (45) erfüllt ist, ist nach Satz 10.4 die Bedingung (49) verletzt.

Literatur

[1] Jacobs, K.: Über Kanäle vom Dichtetypus. Math. Z. 78, 151–170 (1962).

[2] Wolfowitz, J.: Coding Theorems of Information Theory. Ergebn. d. Math. und ihrer Grenzgebiete, Bd. 31. Springer: Berlin-Göttingen-Heidelberg 1964.

[3] Krasnoselskii, M. A., and Ya. Rutuckii: Convex functions and Orlicz spaces. Groningen: P. Noordhoff Ltd. 1961.

[4] Dunford, N., and J. T. Schwartz: Linear operators Part I. Pure and applied Math. Vol. VII. New York: Interscience Publishers Inc. 1958.

[5] Pinsker, M. S.: Arbeiten zur Informationstheorie V. Math. Forschungsberichte XVIII. Berlin: VEB Deutscher Verlag d. Wiss. 1963.

Inst. f. Math. Statistik
34 Göttingen
Math. Inst. der Universität
852 Erlangen