Universität Bielefeld/IMW

Working Papers
Institute of Mathematical Economics

Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung

Nr. 134

STATISTICALLY VARYING GAMES WITH INCOMPLETE INFORMATION

Hans-Martin Wallmeier

May 1984

H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung
an der
Universität Bielefeld
Adresse/Address:
Universitätsstraße
4800 Bielefeld 1
Bundesrepublik Deutschland
Federal Republic of Germany

ABSTRACT

Within a game-theoretical model the influence of incomplete information on the payoff to two players is investigated. It is assumed that the states of nature - varying according to some probability-distribution - are only partially known when the controls are to be chosen. This partial information may be obtained by communicating the states of nature via two independent information-transmission systems owing some capacities, given in advance. Since the payoff to the first player is to be paid by the second (zero-sum assumption), a direct conflict of interests arises.

Admitting a finite, but arbitrarily large delay between the choice of states of nature and the selection of the contrals makes accessible information-theoretical methods. A coding theorem and its converse are proven and used to yield an upper bound to the "min max" of the average asymptotical payoff on one hand, and a lower bound to the "max min" of the average asymptotical payoff on the other. Both bounds are formulated as results on the distortion arising in a discrete, memoryless system and given in a computable manner. By use of the minmax-theorem of game-theory the upper- and lower bounds to the payoff are shown to coincide such that the existence of the value is established.

INTRODUCTION

In the game-theoretical literature "information" appears as a statical notion. A random-mechanism provides uncertainty concerning the value of some parameters of a "game" once and for all and information concerning this uncertainty is obtained by observing the values of some functions in the course of the game. These information-"transmitting" functions are assumed to be given in advance such that optimization concerning the transmission of information is performed by choosing appropriate values for the input-variables of these functions. Models like the one described above range under the term "Games with incomplete information" within game-theoretical literature; its investigation started with [1] ".

In contrast to these short-term apsects also a long-run consideration may be performed. Optimization concerning the choice of the information-transmitting function reflects the variation of system-parameters within some technically given bounds. Since optimization of storing and transmitting information is the field investigated by information-theory, the methods and the results of this discipline should be applied when problems of uncertainty and incomplete information arise within game-theoretical models.

As far as the application of information-theory within game-theoretical models is concerned the merely quantitative point of view of characterizing the information produced by sources or transmitted via noisy channels is not exhaustive. Rather, the qualitative aspects of information, the consequences of actions to be taken according to the available information, have to be considered. A tool for this proceeding sometimes can consist of using results from rate-distortion theory. It expresses quality, namely that of a re-production of a source by quantitative terms, i.e. the amount of "distortion" arising from that reproduction; a first attempt to handle a game-theoretical model following these lines was made in [4].

SECTION I : The Model

The description of a game consists of two parts

- the parameters of the game

and

- the rules according to which the game is played.

In the model to be described in this section, a two-person zero-sum game, the former consist of a finite set $\mathfrak{X}$, - which is interpreted as being the set of states of nature - , together with a probability-distribution $\mu(\cdot)$ on $\mathfrak{X}$. We assume further to be given two finite sets I and J , - sets of controls - and for each $x \in \mathfrak{X}$ an $|I| \times |J|$ -matrix (a_{ij}^x), - the matrix of payoffs.

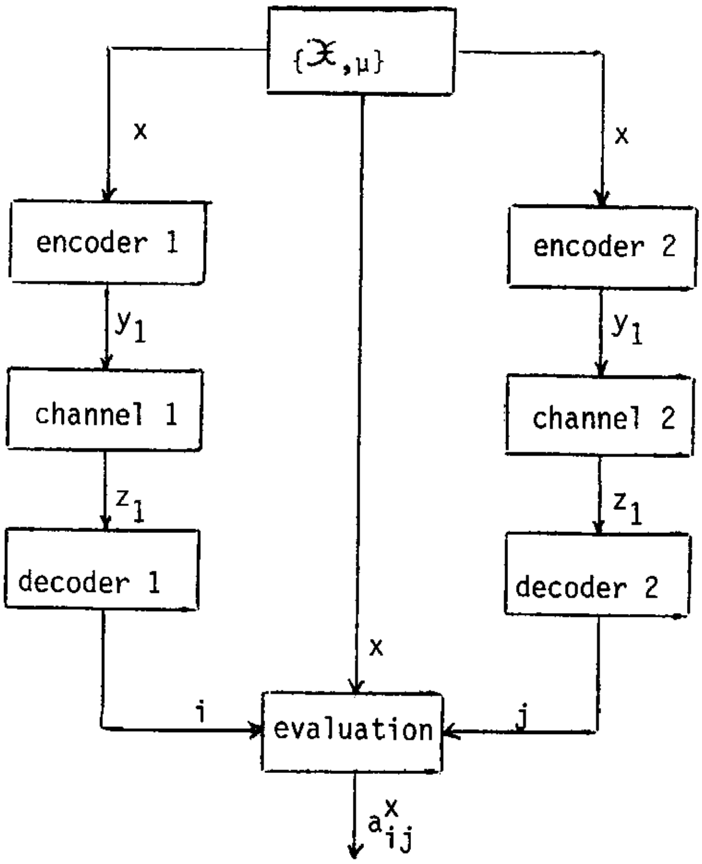
Information-transmission is performed by two (independent) discrete, memoryless channels W_κ with input-alphabets $\mathcal{Y}_\kappa$ and output-alphabets $\mathfrak{Z}_\kappa$, $\kappa = 1,2$.

For every time-unit some $x \in \mathfrak{X}$ is chosen according to μ independently of the previous states of nature. Further, for each time-unit some controls i and j have to be chosen, each without the other being known. We assume that some information concerning the state of nature can be made available by means of the two channels before the controls are to be chosen.

The controls i and j and the state x result in a payoff a_{ij}^x . For sequences i^n, j^n and x^n we define the payoff to be additive, that is

$$a_{i^n j^n}^{x^n} := \sum_t a_{i_t j_t}^{x_t} .$$

Graphically, the model is given as follows:



The power of the information-transmitting systems W_κ, $\kappa = 1,2$ being bounded, the problem of optimally organizing the processing of information arises. The source $\{\mathfrak{X}, \mu\}$, the channels W_κ and the payoff-matrices (a^x_{ij}) , $x \in \mathfrak{X}$ being given, the players I and II decide upon the "design" of the encoders and the decoders. Hallowed by tradition player I tries to maximize the payoff whereas player II tries to minimize it (zero-sum assumption!). The maxmin and the minmax of the payoff have to be investigated and, to obtain a complete solution, their coincidence has to be proven.

To solve the problem of determining the payoff which results from such a situation, the assumption of a finite but arbitrarily large admissible time-lag is made. This enables the players to build up encoders which transform the observed n-sequences x^n into channel-input sequences y^n_1 and y^n_2 , respectively. The latters are transmitted via the channels and, according to the channel-distribution, sequences z^n_1 and z^n_2 are obtained. On the basis of the channel output sequences z^n_1 and z^n_2 , respectively the decoders independently decide upon sequences i^n and j^n which, together with the sequence x^n of states, result in the payoff $a^{x^n}_{i^n j^n}$ for player I. This amount has to be paid by player II.

Our aim is to give a computable formula for the asymptotical average payoff per time-unit which can be ensured by both players to themselves. This payoff will be given as a function of the power (the "capacity") of the transmission system to be used.

Denoting the value of the asymptotical game by

$$\text{val}\,(\Gamma^\infty\,(\mu, W_1, W_2, A))$$

- it exists since max min and min max of the average payoffs per time-unit are shown to coincide,

we may formulate our main theorem.

Theorem

$$\text{val}\,(\Gamma^\infty\,(\mu, W_1, W_2, A)) = \max_{\substack{V_1 \\ I(\mu;V_1)\leq C_1}} \{\underline{D}_\mu^{d_{V_1}}\,(C_2)\}$$

where $\underline{D}^d(\cdot)$ is the distortion-rate function, - the inverse to the well-known rate-distortion function - with respect to the discrete memoryless source $\{\mathcal{X}, \mu\}$ and the single-letter fidelity criterion d, $d : \mathcal{X} \times J \to \mathbb{R}$. For a test-channel V_1 with input-alphabet $\mathcal{X}$ and output-alphabet I d_{V_1} is defined to be

$$d_{V_1}\,(x,j) := \sum_i V_1(i|x) \cdot a^x_{ij}\,.$$

SECTION II : Preliminaries

Let us be given finite sets $\mathcal{X}$ and $\hat{\mathcal{X}}$. Define

$$\mathcal{P}\,(\mathcal{X}) := \{\nu \mid \nu \text{ probability distribution on } \mathcal{X}\}\,.$$

A conditional probability distribution V on $\hat{\mathcal{X}}$ given $\mathcal{X}$ will be denoted as

$$V \mid \mathcal{X} \Rightarrow \hat{\mathcal{X}}$$

Define

$$\mathcal{W}(\hat{\mathcal{X}} \mid \mathcal{X}) := \{V \mid V \mid \mathcal{X} \rightarrow \hat{\mathcal{X}}\}$$

By $V \circ \nu$ we denote the marginal distribution on $\hat{\mathcal{X}}$ induced by $\nu \in \mathcal{P}(\mathcal{X})$, $V \in \mathcal{W}(\hat{\mathcal{X}} \mid \mathcal{X})$, i.e. $V \circ \nu\,(\hat{x}) := \sum_x V(\hat{x} \mid x)\ \nu(x)$.

Notation, definitions and results not being derived in this paper may be found in [2] and [3]. We shall repeat only:

$N(x \mid x^n) := |\{t \mid x_t = x\}|$, $x \in \mathcal{X}$, $x^n \in \mathcal{X}^n$, $\left(\frac{N(x \mid x^n)}{n}\right)_{x \in \mathcal{X}}$ is a

probability distribution on $\mathcal{X}^n$, being called type of x^n. The set of all types is denoted by $\mathcal{P}^n(\mathcal{X})$. Define

$$\mathcal{W}^n(\hat{\mathcal{X}} \mid \mathcal{X}) = \{V \in \mathcal{W}(\hat{\mathcal{X}} \mid \mathcal{X}) \mid \bigwedge_x V(\cdot \mid x) \in \mathcal{P}^n(\mathcal{X})\}.$$

For $\nu \in \mathcal{P}^n(\mathcal{X})$, $V \in \mathcal{W}^n(\hat{\mathcal{X}} \mid \mathcal{X})$

$$T_\nu = \{x^n \mid \bigwedge_x n^{-1} N(x \mid x^n) = \nu(x)\}$$

$$T_V(x^n) := \{\hat{x}^n \mid \bigwedge_{x,\hat{x}} n^{-1} N(x,\hat{x} \mid x^n,\hat{x}^n) = n^{-1} V(\hat{x} \mid x) \cdot N(x \mid x^n)\} .$$

For $\mu \in \mathcal{P}(\mathcal{X})$, $\delta > 0$ define

$$T_\mu^\delta := \{x^n \mid \bigwedge_x |\, n^{-1} N(x \mid x^n) - \mu(x)\,| < \delta\}$$

For $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}^n(\mathcal{X})$ define

$$\rho(\mu,\nu) := \max_x \{|\mu(x) - \nu(x)|\} ,$$

then $T_\mu^\delta = \bigcup_{\substack{\nu: \\ \rho(\mu,\nu)<\delta}} T_\nu$.

We shall use the inequalities:

(1) $(n+1)^{-|\mathcal{X}|} \exp\{n\, H(\nu)\} \leq |\, T_\nu \,| \leq \exp\{n\, H(\nu)\}$

(2) $\bigwedge_{x^n \in T_\nu} (n+1)^{-|\mathcal{X}| \cdot |\hat{\mathcal{X}}|} \exp\{n\, H(V|\nu)\} \leq |T_V(x^n)| \leq \exp\{n\, H(V|\nu)\}$

(3) $\bigwedge_{\varepsilon>0} \bigvee_{n_0} \bigwedge_{n \geq n_0} |\mathcal{P}^n(\mathcal{X})| \leq (n+1)^{|\mathcal{X}|} \leq \exp\{n \cdot \varepsilon\}$

(4) $\bigwedge_{\varepsilon,\delta>0} \bigvee_{n_0(|\mathcal{X}|,\varepsilon,\delta)} \bigwedge_{n \geq n_0} \mu(T_\mu^\delta) > 1 - \varepsilon.$

We shall derive some auxiliary results:

Let W be a compact, convex set.

Let $f,h : W \longrightarrow \mathbf{R}$

be convex and concave (and thereby continuous) and let

$$g,\ l : W \longrightarrow \mathbf{R}$$

be convex and continuous.

2.1. Lemma:

Define $F,G : \mathbf{R} \to \mathbf{R}$
by

$$F(r) := \max_{\substack{w:\\ g(w)\leq r}} \{f(w)\}$$

$$G(s) := \min_{\substack{w:\\ h(w)\geq s}} \{l(w)\} \tag{5}$$

Then $F(\cdot)$ is

- strictly monotonic increasing ("strictly isotonic") unless $\max\limits_{w\in W} \{f(w)\}$ is achieved
- concave

and

- G is the inverse function to F where F is invertible.

Define $K,L : \mathbf{R} \to \mathbf{R}$
by

$$K(r) := \min_{\substack{w:\\ l(w)\leq r}} \{k(w)\}$$

$$L(s) := \min_{\substack{w:\\ k(w)\leq s}} \{l(w)\} \tag{6}$$

Then $K(\cdot)$ is

- strictly monotonic decreasing ("strictly antitonic") unless $\min_{w \in W} \{k(w)\}$ is achieved
- convex

and

- L is the inverse function to K where K is invertible.

Proof:

We shall restrict ourselves on proving (5), the proof of (6) might be performed similarly. We begin with proving the concavity of $F(\cdot)$.

Let w_κ be such that

$$g(w_\kappa) \leq r_\kappa$$

$$f(w_\kappa) = F(r_\kappa), \quad \kappa = 1,2$$

Then $w_\lambda := \lambda \cdot w_1 + (1-\lambda)\, w_2 \in W$ and due to the convexity of g

$$g(w) \leq \lambda \cdot r_1 + (1-\lambda) \cdot r_2$$

and, using convexity and concavity of $f(\cdot)$:

$$\begin{aligned} F(w_\lambda) &\geq f(w_\lambda) \\ &= \lambda \cdot f(w_1) + (1-\lambda)\, f(w_2) \\ &= \lambda \cdot F(r_1) + (1-\lambda)\, F(r_2) \,. \end{aligned}$$

Per definition F is isotonic, combining this property with the concavity shows the strict isotonicity if the maxmin value has not been achieved yet. In the domain defined by this property $F(\cdot)$ is invertible.

It remains to show that $G = F^{-1}$ on the specified domain:

$$G(F(r)) = \min_{g(w) \geq \max_{w':\, g(w') \leq r} \{f(w')\}} \{g(w)\}$$

$$= \min_{\substack{w': \\ g(w')\geq r}} \{g(w)\}$$

$$= r$$

where, to derive the second equality, continuity and isotonicity of f and g were used.

2.2 Lemma

The functions

$$F, G, K, L : \mathbb{R} \to \mathbb{R}$$

are continuous.

Proof:

Obvious.

SECTION III: The Coding Theorem

Coding results of information-theory describe lower bounds to the power of information storing - or transmitting-systems. The coding theorem to be derived in this section is used to give explicitly a strategy of the "pre-playing" player, both, in the maxmin- and in the minmax-game. The coding- and decoding-rules given in the coding theorem give a way of selecting controls within the restrictions on the available information. To ensure optimal behaviour against any strategy of the "post-playing" (reacting) player random encoding of the states of nature is used. This random-behaviour against a malevolant being is in accordance with the coding result for arbitrarily varying channels. In contrast to the latter random encoding instead of using random-codes is shown to be sufficient. Random encoding excludes advantages of the post-playing player. In case of deterministic encoding he might use the structure of the code to obtain a higher payoff.

Let $\hat{\mathcal{X}}$ denote a finite set, for $\mu \in \mathcal{P}(\mathcal{X})$, $R > 0$ define

$$\mathcal{V}_{\hat{\mathcal{X}}}^{\mu}(R) := \{V \in W^n(\hat{\mathcal{X}} | \mathcal{X}) \mid I(\mu,V) < R\} .$$

3.1 Lemma

Let $\frac{\log e}{2} > \varepsilon > 0$, and $R > 0$ be given. There exists $n_0(|\mathcal{X}|,|\hat{\mathcal{X}}|,\varepsilon)$ such that for $n \geq n_0$, $\nu \in \mathcal{P}^n(\mathcal{X})$, $V \in \check{\mathcal{V}}_{\hat{\mathcal{X}}}(R-\varepsilon)$, $N \geq \exp\{n \cdot (R+\varepsilon)\}$ and independent, uniformly distributed random variables Y_k on $T_{V \circ \nu} \subset \hat{\mathcal{X}}^n$, $k = 1,\dots,N$:

$$\bigwedge_{x^n \in T_\nu} \sum_{\hat{x}^n \in \hat{\mathcal{X}}^n} \Pr\{\sum_k 1_{T_{V(x^n)}}(Y_k) < \exp\{n \cdot \varepsilon\}\}$$

$$< \exp\{n \cdot \log|\hat{\mathcal{X}}| - \frac{\varepsilon}{2}\exp\{n \cdot \varepsilon\}\}$$

Proof:

We shall make use of the well-known inequality

$$\Pr\{\sum_{k=1}^{N} M_k < \frac{1}{K} \cdot \alpha\}$$

$$< \exp\{\alpha \cdot (\frac{1}{K} - \frac{\log e}{2})\}$$

for independent, identically distributed random variables M_k with values in $\{0,1\}$, where α denotes the expectation of the sum of M_k, i.e.

$$\alpha = N \cdot E(M_k) .$$

Now

$$E(Y_k) = \frac{|T_V(x^n)|}{|T_{V \circ \nu}|}$$

yields

$$\alpha \geq N \cdot \exp\{n \cdot (H(V|\nu) - H(V \circ \nu) - \varepsilon)\}$$

$$\geq \exp\{n \cdot (R+\varepsilon - I(\nu,V) - \varepsilon)\}$$

$$= \exp\{n \cdot (R - I(\nu,V))\} .$$

Thus

$$\Pr \{ \sum_k 1_{T_{V(x^n)}}(Y_k) \leq \exp \{n \cdot \varepsilon\}\}$$

$$= \Pr \{ \sum_k 1_{T_{V(x^n)}}(Y_k) \leq \exp \{n \cdot (R-I(\nu,V) - R+I(\nu,V) +\varepsilon)\}\}$$

$$\leq \Pr \{ \sum_k 1_{T_{V(x^n)}}(Y_k) \leq \alpha \cdot \exp \{-n \cdot (R-I(\nu,V) -\varepsilon)\}\}$$

$$< \exp \{\alpha \cdot (\exp \{-n \cdot (R-I(\nu,V) -\varepsilon)\} - \frac{\log e}{2})\}$$

$$\leq \exp \{\alpha \cdot (- \frac{\varepsilon}{2})\} ,$$

where the last inequality holds, since $R > I(\nu,V) + \varepsilon$, for sufficiently large n.

Using $\alpha \geq \exp \{n \cdot (R-I(\nu,V))\}$
we infer

$$\Pr \{ \sum_k 1_{T_{V(x^n)}}(Y_k) \leq \exp \{n \cdot \varepsilon\}\}$$

$$< \exp \{- \exp \{n \cdot (R-I(\nu,V))\} \cdot \frac{\varepsilon}{2}\}$$

and

$$\sum_{\tilde{x}^n} \Pr \{ \sum_k 1_{T_{V(x^n)}}(Y_k) \leq \exp \{n \cdot \varepsilon\}\}$$

$$< |\hat{\mathcal{X}}^n| \cdot \exp \{- \exp \{n \cdot (R-I(\nu,V))\} \cdot \frac{\varepsilon}{2}\}$$

$$\leq \exp \{n \cdot \log |\hat{\mathcal{X}}| - \exp \{n \cdot (R-I(\nu,V))\} \cdot \frac{\varepsilon}{2}\}$$

$$\leq \exp \{n \cdot \log |\hat{\mathcal{X}}| - \exp \{n \cdot \varepsilon\} \cdot \frac{\varepsilon}{2}\}$$

For finite sets $\hat{\mathcal{X}}$, $\tilde{\mathcal{X}}$ and $x^n \in \mathcal{X}^n$ define

$$T^{\delta}(x^n,\tilde{x}^n) := \{\hat{x}^n \mid \bigwedge_{x,\hat{x},\tilde{x}} |n^{-1} N(x,\hat{x},\tilde{x}|x^n,\hat{x}^n,\tilde{x}^n)-n^{-1}V(\hat{x}|x)\cdot N(x,\tilde{x}|x^n,\tilde{x}^n)| < \delta\} ,$$

by $T^{\delta^c}(x^n,\tilde{x}^n)$ we denote the complementary set within $\hat{\mathcal{X}}^n$.

3.2 Lemma

For $\tau, \delta > 0$ there exists $n_0(|\mathfrak{X}|, |\hat{\mathfrak{X}}|, |\tilde{\mathfrak{X}}|, \tau, \delta)$ such that for $n \geq n_0$, $V \in \mathcal{W}^n(\hat{\mathfrak{X}}|\mathfrak{X})$, and $\hat{X}^n$ uniformly distributed on $T_V(x^n)$:

$$\bigwedge_{\tilde{x}^n \in \tilde{\mathfrak{X}}^n} E\,[1_{T^\delta(x^n,\tilde{x}^n)}(\hat{X}^n)] \geq 1 - \tau$$

Proof:

$$E\,[1_{T^\delta(x^n,\tilde{x}^n)}(\hat{X}^n)]$$

$$= 1 - Pr\,\{\hat{X}^n \notin \{\hat{x}^n | \bigwedge_{x,\hat{x},\tilde{x}} |n^{-1}N(x,\hat{x},\tilde{x}|x^n,\hat{x}^n,\tilde{x}^n) - n^{-1}V(\hat{x}|x)\cdot N(x,\tilde{x}|x^n,\tilde{x}^n)| < \delta\}\}$$

$$= 1 - Pr\,\{\hat{X}^n \in \{\hat{x}^n | \bigvee_{x,\hat{x},\tilde{x}} |\, N(x,\hat{x},\tilde{x}|x^n,\hat{x}^n,\tilde{x}^n) - V(\hat{x}|x)\, N(x,\tilde{x}|x^n,\tilde{x}^n)| \geq n\cdot\delta\}\}$$

$$= 1 - Pr\,\{\bigvee_{x,\hat{x},\tilde{x}} |\, N(x,\hat{x},\tilde{x}\,|\, x^n,\hat{X}^n,\tilde{x}^n) - V(\hat{x}|x)\cdot N(x,\tilde{x}|x^n,\tilde{x}^n)| \geq n\cdot\delta\}\}$$

For $(x^n,\tilde{x}^n) \in \mathfrak{X}^n \times \tilde{\mathfrak{X}}^n$ $N(x,\hat{x},\tilde{x}|x^n,\hat{X}^n,\tilde{x}^n)$ is a binomial-distributed random variable with expectation

$$N(x,\tilde{x}|x^n,\tilde{x}^n) \cdot V(\hat{x}|x)$$

and variance

$$N(x,\tilde{x}|x^n,\tilde{x}^n) \cdot V(\hat{x}|x) \cdot (1-V(\hat{x}|x)) \leq \frac{1}{4} \cdot N(x,\tilde{x}|x^n,\tilde{x}^n)$$

since $\hat{X}^n$ is uniformly distributed on $T_V(x^n)$.

Thus by Chebychev's inequality

$$Pr\,\{\bigvee_{x,\hat{x},x} |\, N(x,\hat{x},\tilde{x}|x^n,\hat{X}^n,\tilde{x}^n) - V(\hat{x}|x)\, N(x,\tilde{x}|x^n,\tilde{x}^n)| \geq n\cdot\delta\}$$

$$\leq |\mathfrak{X}|\cdot|\hat{\mathfrak{X}}|\cdot|\tilde{\mathfrak{X}}| \cdot \frac{N(x,\tilde{x}|x^n,\tilde{x}^n)}{4\cdot n^2\cdot \delta^2}$$

$$\leq \frac{|\mathfrak{X}|\cdot|\hat{\mathfrak{X}}|\cdot|\tilde{\mathfrak{X}}|}{4\cdot n \cdot \delta^2}$$

$< \tau$,

for sufficiently large n_0 .

The claim follows.

3.3 Lemma

Let $\varepsilon, \delta > 0$ be given. Let τ be such that $\log(1+\tau) < \frac{\varepsilon}{2}$.
There exists $n_0(|\mathcal{X}|, |\hat{\mathcal{X}}|, |\tilde{\mathcal{X}}|, \varepsilon, \delta, \tau)$ such that for all $n \geq n_0$, $V \in \mathcal{W}^n(\hat{\mathcal{X}} \mid \mathcal{X})$ and independent, on $T_V(x^n)$ uniformly distributed random variables $\hat{X}_l$, $l = 1, \ldots, L$, $L \geq \exp\{n \cdot \varepsilon\}$:

$$\bigwedge_{x^n \in \mathcal{X}^n} \sum_{\tilde{x}^n \in \tilde{\mathcal{X}}^n} \Pr\{\frac{1}{L} \sum_l 1_{T^\delta(x^n, x^n)}(\hat{X}_l) < 1 - \varepsilon\}$$

$$< \exp\{n \cdot \log|\tilde{\mathcal{X}}| - L \cdot \frac{\varepsilon}{2}\}$$

Proof:

$$\Pr\{\frac{1}{L} \sum_l 1_{T^\delta(x^n, x^n)}(\hat{X}_l) < 1 - \varepsilon\}$$

$$= \Pr\{\sum_l 1_{T^{\delta c}(x^n, x^n)}(\hat{X}_l) > L \cdot \varepsilon\}$$

$$< \exp\{-L \cdot \varepsilon\}\ E[\exp\{\sum_l 1_{T^{\delta c}(x^n, \tilde{x}^n)}(\hat{X}_l)\}],$$

the inequality being obtained using the Markov-inequality.

Due to the independency of the random variables $\hat{X}_l$ and using $\exp\{t\} \leq 1+t$ for $t \in [0,1]$ we obtain

$$E[\exp\{\sum_l 1_{T^{\delta c}(x^n, x^n)}(\hat{X}_l)\}]$$

$$\leq (1 + E[1_{T^{\delta c}(x^n, x^n)}(\hat{X}_l)])^L$$

whence

$$\sum_{\tilde{x}^n \in \tilde{\mathcal{X}}^n} \Pr\{\frac{1}{L} \sum_l 1_{T^{\delta c}(x^n, x^n)}(\hat{X}_l) < 1 - \varepsilon\}$$

$$\leq |\tilde{\mathcal{X}}^n| \cdot \exp\{-L \cdot (\varepsilon - \log(1 + E[1_{T^{\delta c}(x^n, x^n)}(\hat{X}_l)]))\}$$

The preceding lemma is applied to yield the upper bound

$$\exp\{n \cdot \log|\tilde{\mathfrak{X}}| - L(\varepsilon - \log(1+\tau))\}$$
$$< \exp\{n \cdot \log|\tilde{\mathfrak{X}}| - L \cdot \tfrac{\varepsilon}{2}\} .$$

We have now provided all material to prove the source-coding theorem. It will consist of two symmetric claims such that the proof of one of them is sufficient. The lemmata given above will be applied for $I = \hat{\mathfrak{X}}$, $J = \tilde{\mathfrak{X}}$ (and conversely for the second claim).

3.4 Theorem

For $\varepsilon, \delta > 0$, there exists $n_0(|\mathfrak{X}|, |I|, |J|, \varepsilon, \delta)$ such that for $n \geq n_0$, $R_1, R_2 > 0$, $V_1 \in \mathfrak{V}_I^{\mu}(R_1 - 2\varepsilon)$ and $V_2 \in \mathfrak{V}_J^{\mu}(R_2 - 2\varepsilon)$ there exist

$$\mathfrak{M}_I \subset I^n ,$$

$$|\mathfrak{M}_I| \leq \exp\{n \cdot (R_1 + 3\varepsilon)\}$$

(1) such that

$$\bigwedge_{x^n \in T_\mu^\delta} \bigwedge_{j^n \in J^n} \frac{|\mathfrak{M}_I \cap T_{V_1}(x^n) \cap T_I^\delta(x^n, j^n)|}{|\mathfrak{M}_I \cap T_{V_1}(x^n)|} \geq 1 - \varepsilon$$

and

$$\mathfrak{M}_J \subset J^n,$$

$$|\mathfrak{M}_J| \leq \exp\{n \cdot (R_2 + 3\varepsilon)\}$$

(2) such that

$$\bigwedge_{x^n \in T_\mu^\delta} \bigwedge_{i^n \in I^n} \frac{|\mathfrak{M}_J \cap T_{V_2}(x^n) \cap T_J^\delta(x^n, i^n)|}{|\mathfrak{M}_J \cap T_{V_2}(x^n)|} \geq 1 - \varepsilon .$$

Proof:

We may restrict ourselves on proving (1). The set $\mathfrak{M}$ will be defined to consist of elements of different types.

Using a random-mechanism we will show that for every ν such that $\rho(\mu,\nu)<\delta$ there exists a subset $\mathfrak{M}_\nu$ of $T_{V\circ\nu}$ with cardinality N_ν upperbounded by $\exp\{n \cdot (R_1 + 2\varepsilon)\}$ such that

$$(3) \qquad \bigwedge_{x^n \in T_\nu} \bigwedge_{j^n} \frac{|\mathfrak{M}_\nu \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|}{|\mathfrak{M}_\nu \cap T_V(x^n)|} \geq 1 - \varepsilon .$$

Having proven this, using

$$|\mathcal{P}^n(\mathfrak{X})| \leq \exp\{n \cdot \varepsilon\}$$

and the disjointness of $T_{V\circ\nu}$ and $T_{V\circ\nu'}$ for $\nu \neq \nu'$, leading to

$$\bigwedge_{x^n \in T_\nu} |\bigcup_{\nu'} \mathfrak{M}_{\nu'} \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)| = |\mathfrak{M}_\nu \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|$$

and

$$\bigwedge_{x^n \in T_\nu} |\bigcup_{\nu'} \mathfrak{M}_{\nu'} \cap T_V(x^n)| = |\mathfrak{M}_\nu \cap T_V(x^n)|$$

yield the desired property:

$$\bigwedge_{x^n \in T_\mu^\delta} \bigwedge_{j^n \in J^n} \frac{|\mathfrak{M} \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|}{|\mathfrak{M} \cap T_V(x^n)|} \geq 1 - \varepsilon, \quad \mathfrak{M} = \bigcup_{\nu'} \mathfrak{M}_{\nu'} .$$

Let us start proving (3). Since the mutual information $I(\cdot,\cdot)$ is a continuous function in the input-probability ν, $\nu \in \mathcal{P}(\mathfrak{X})$ and the transition probability V, $V \in W(\hat{\mathfrak{X}}|\mathfrak{X})$, compactness yields uniform continuity such that for any $\varepsilon > 0$ there exists ρ_ε such that

$$\rho(\mu,\nu) < \rho_\varepsilon \quad \text{implies} \quad |I(\nu,V) - I(\mu,V)| < \varepsilon$$

for all $V \in W(\hat{\mathfrak{X}}|\mathfrak{X})$.

Thus, for ν such that $\rho(\mu,\nu) < \delta_\varepsilon$ and $V \in \mathcal{V}^{\mu}(R_1-2\varepsilon)$: $V \in \mathcal{V}^{\nu}(R_1-\varepsilon)$.
For fixed ν obeying $\rho(\mu,\nu) < \delta_\varepsilon$ it is sufficient to show

$$1 > \Pr \{ \mathfrak{M}_\nu \,|\, |\mathfrak{M}_\nu| = N_\nu, \bigwedge_{x^n \in T_\nu} \bigwedge_{j^n} \frac{|\mathfrak{M}_\nu \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|}{|\mathfrak{M}_\nu \cap T_V(x^n)|} < 1 - \varepsilon\}$$

for an appropriate random-mechanism choosing N_ν-elementary sets.
We may upperbound the probability by

$$\sum_{x^n \in T_\nu} \sum_{j^n} \Pr \{ \mathfrak{M}_\nu \,|\, |\mathfrak{M}_\nu| = N_\nu \quad \frac{|\mathfrak{M}_\nu \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|}{|\mathfrak{M}_\nu \cap T_V(x^n)|} < 1 - \varepsilon\}$$

Define $N_\nu \in \mathbb{N}$ such that

$$\begin{aligned} &\exp\{n \cdot (R_1 + \varepsilon)\} \\ \leq\ &N_\nu \\ \leq\ &\exp\{n \cdot (R_1 + 2\varepsilon)\} \end{aligned}$$

and let us choose the elements of $\mathfrak{M}_\nu$ independently according to uniformly distributed random variables $(Y_k^\nu)_{k=1,\dots,N_\nu}$ with values in $T_{V\circ\nu}$.

Then

$$\Pr \{ \mathfrak{M}_\nu \,|\, |\mathfrak{M}_\nu| = N_\nu , \frac{|\mathfrak{M}_\nu \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|}{|\mathfrak{M}_\nu \cap T_V(x^n)|} < 1 - \varepsilon\}$$

$$= \Pr \{ \sum_{k=1}^{N_\nu} 1_{T_V(x^n) \cap T_I^\delta(x^n,j^n)}(Y_k^\nu) < (1-\varepsilon) \cdot \sum_{k=1}^{N_\nu} 1_{T_V(x^n)}(Y_k^\nu)\}$$

$$= \sum_{L=1}^{N_\nu} \sum_{\substack{A_L \subset \{1,\dots,N\} \\ |A_L| = L}} \Pr \{1_{T_V(x^n)}(Y_k^\nu) = 1 \Leftrightarrow k \in A_L, \frac{1}{L} \sum_{l \in A_L} 1_{T_I^\delta(x^n,j^n)}(Y_l^\nu) < 1 - \varepsilon\}$$

$$= \sum_{L=1}^{N_\nu} \sum_{A_L} \Pr \{1_{T_V(x^n)}(Y_k^\nu) = 1 \Leftrightarrow k \in A_L\}$$

$$\cdot \Pr \{\frac{1}{L} \sum_{l \in A_L} 1_{T_V(x^n) \cap T_I^\delta(x^n,j^n)}(Y_l) < 1-\varepsilon \mid Y_l^\nu = 1\}$$

$$\leq \sum_{L=1}^{\exp\{n\cdot\varepsilon\}} \sum_{A_L} \Pr\{1_{T_V(x^n)}(Y_k^\nu) = 1 \Leftrightarrow k \in A_L\}$$

$$+ \sum_{L\geq\exp\{n\cdot\varepsilon\}} \sum_{A_L} \Pr\{1_{T_V(x^n)}(Y_k^\nu) = 1 \Leftrightarrow k \in A_L\}$$

$$\cdot \Pr\{\frac{1}{L} \sum_{l\in A_L} 1_{T_I^\delta(x^n,j^n)}(Y_k^\nu) < 1-\varepsilon \mid Y_l^\nu = 1\}$$

$$\leq \Pr\{\sum_{k=1}^{N_\nu} 1_{T_V(x^n)}(Y_k^\nu) < \exp\{n\cdot\varepsilon\}$$

$$+ \max_{\substack{L: \\ L\geq\exp\{n\cdot\varepsilon\}}} \{\Pr\{\frac{1}{L}\sum_{l=1}^{L} 1_{T_I^\delta(x^n,j^n)}(Z_l) < 1-\varepsilon\}\},$$

where Z_l , $l=1,\dots,L$ denote independent, uniformly distributed random variables with values in $T_V(x^n)$.

In order to upperbound

$$(4) \quad \sum_{x^n\in T_\nu} \sum_{j^n} \Pr\{\mathfrak{M}_\nu \mid |\mathfrak{M}_\nu| = N_\nu \, , \frac{|\mathfrak{M}_\nu \cap T_V(x^n) \cap T_I^\delta(x^n,j^n)|}{|\mathfrak{M}_\nu \cap T_V(x^n)|} < 1-\varepsilon\}$$

we apply the lemmata 2.1 and 2.3 to the last expression of the preceding chain of inequalities. We obtain as a strict upperbound

$$|T_\nu| \cdot (\exp\{n\cdot\log|J| - \frac{\varepsilon}{2}\exp\{n\cdot\varepsilon\}\} + \exp\{n\cdot\log|J| - \exp\{n\cdot\varepsilon\}\frac{\varepsilon}{2})\}$$
$$\leq 2\exp\{n(\log|\mathcal{X}| + \log|J|) - \frac{\varepsilon}{2}\exp\{n\cdot\varepsilon\}\}$$

For sufficiently large n, depending on $|\mathcal{X}|$, $|J|$ and ε this expression may be bounded by 1.

The above theorem ensures the existence of a set of codewords $\mathfrak{M}_I$ ($\mathfrak{M}_J$) for each encoder such that for sequences $x^n \in T_\mu^\delta$ and arbitrary $j^n \in J^n$ ($i^n \in I^n$) chosen by the opponent, most of the potential codewords for x^n - i.e. the sequences $i^n \in \mathfrak{M}_I \cap T_{V_1}(x^n)$ ($j^n \in \mathfrak{M}_J \cap T_{V_2}(x^n)$) - have the product-distribution as the common distribution of (i^n, j^n) .

Since the controls are not chosen by the encoders but by the decoders, we have to transmit the information concerning the sequences of states to the latter. Thus we are enforced to prove an information transmission theorem, the channel coding theorem used herein will be formulated for the sake of completeness.

Let $\mathcal{Y}$ und $\mathcal{Z}$ be finite sets, $W \mid \mathcal{Y} \to \mathcal{Z}$ a discrete, memoryless channel with capacity denoted by C $(= \max_{P \in \mathcal{P}(\mathcal{Y})} \{I(P,W)\})$.

3.5 Theorem:

For all $\varepsilon > 0$, $0 < \lambda < 1$ there exists $n_0(|\mathcal{Y}|, |\mathcal{Z}|, \varepsilon, \lambda)$ such that for all $n \geq n_0$ there exists a code denoted by

$$\{(u_k, D_k) \mid k=1,\dots,N,\ u_k \in \mathcal{Y}^n,\ D_k \subset \mathcal{Z}^n,\ D_k \cap D_l = \emptyset \text{ for } k \neq l\}$$

such that

$$\lambda \geq \max_k \{W^n(D_k^c \mid u_k)\} ,$$

$$N \geq \exp\{n \cdot (C - \varepsilon)\} .$$

For $V_1 \in W^n(I \mid \mathcal{X})$, $V_2 \in W^n(J \mid \mathcal{X})$

we define

$$(5) \qquad \begin{aligned} d_{V_1}(x,j) &= \sum_i V_1(i \mid x) \quad a_{ij}^x , \\ d_{V_2}(x,i) &= \sum_j V_2(j \mid x) \quad a_{ij}^x \end{aligned}$$

and

$$d_{V_1}(x^n,j^n) = \sum_t d_{V_1}(x_t,j_t)$$

$$d_{V_2}(x^n,i^n) = \sum_t d_{V_2}(x_t,i_t).$$

In a well-known manner the source-coding theorem and the channel-coding theorem are clung together to yield the information transmission theorem.

3.6 Theorem

For $\varepsilon^* > 0$ and $\varepsilon = \varepsilon(\varepsilon^*)$, $\delta = \delta(\varepsilon)$ sufficiently small there exists $n_o\,(|\mathcal{X}|,|I|,|J|,|\mathcal{Y}_1|,|\mathcal{Y}_2|,|\mathcal{Z}_1|,|\mathcal{Z}_2|,\varepsilon^*,\varepsilon,\delta)$ such that for $n \geq n_o$, $V_1 \in \mathcal{V}^{\mu}_I(C_1-3\varepsilon)$ and $V_2 \in \mathcal{V}^{\mu}_J(C_2-3\varepsilon)$ there exist

$$P^1_E \mid \mathcal{X}^n \Longrightarrow \mathcal{Y}^n_1$$

$$P^1_D : \mathcal{Z}^n_1 \longrightarrow I^n$$

(5) such that

$$\bigwedge_{x^n \in T^\delta_\mu} \bigwedge_{j^n} n^{-1} \sum_{y^n_1} P^1_E(y^n_1|x^n) \sum_{z^n_1} W^n_1(z^n_1|y^n_1) \cdot a_{x^n P^1_D(z^n_1),j^n}$$

$$\geq n^{-1} d_{V_1}(x^n,j^n) - \varepsilon^*$$

and there exist

$$P^2_E \mid \mathcal{X}^n \Longrightarrow \mathcal{Y}^n_2$$

$$P^2_D : \mathcal{Z}^n_2 \longrightarrow J^n$$

(6) such that

$$\bigwedge_{x^n \in T^\delta_\mu} \bigwedge_{i^n} n^{-1} \sum_{y^n_2} P^2_E(y^n_2|x^n) \sum_{z^n_2} W^n_2(z^n_2|y^n_2) \cdot a_{x^n i^n,P^2_D(z^n_2)}$$

$$\leq n^{-1} d_{V_2}(x^n,i^n) + \varepsilon^* \quad ,$$

where C_1, C_2 denote the capacities of the channels W_1 and W_2, respectively.

Proof:

Define ε, λ such that

$$\lambda \cdot \max_{x,i,j} \{|a^x_{ij}|\} < \varepsilon,$$

$$\bigwedge_{x,j} (1-\lambda)\,(1-\varepsilon)\, d_{V_1}(x,j) - \varepsilon\,(\max_{x,i,j} \{|a^x_{ij}|\}+1) \geq d_{V_1}(x,j) - \varepsilon^*$$

(6) and

$$\bigwedge_{x,i} (1-\lambda)\,(1-\varepsilon)\, d_{V_2}(x,i) + \varepsilon\,(\max_{x,i,j} \{|a^x_{ij}|\} + 1) \leq d_{V_2}(x,i) + \varepsilon^*,$$

then, using symmetry, it is sufficient to prove

$$\bigwedge_{x^n \in T^\delta_\mu} \bigwedge_{j^n} n^{-1} \sum_{y^n_1} P^1_E(y^n_1|x^n) \sum_{z^n_1} W^n_1(z^n_1|y^n_1) \cdot a^{x^n}_{P^1_D(z^n_1),j^n}$$

$$\geq (1-\lambda)\,(1-\varepsilon)\, n^{-1}\, d_{V_1}(x^n,j^n) - \varepsilon \cdot (\max_{x,i,j} \{|a^x_{ij}|\} + 1).$$

(7) Let δ be such that $|\mathfrak{X}| \cdot |I| \cdot |J| \cdot \delta \cdot \max_{x,i,j} \{|a^x_{ij}|\} < \varepsilon$

Define $R_1 = C_1 - 3\varepsilon$ and let $\mathfrak{M} := \mathfrak{M}_I$ be chosen according to theorem 2.4.

Let

$$f : \mathfrak{M} \longrightarrow \{u_k \mid k = 1,\dots,N\}$$

be any injective function, the existence being ensured since

$$|\mathfrak{M}| \leq \exp\{n \cdot (R_1 + 2\varepsilon)\}$$
$$= \exp\{n \cdot (C_1 - \varepsilon)\}.$$

Further, define the encoding rule by

$$P_E^1(y_1^n|x^n) = \begin{cases} \dfrac{1}{|\mathfrak{M} \cap T_{V_1}(x^n)|} & \text{if} \quad f^{-1}(y_1^n) \in \mathfrak{M} \cap T_{V_1}(x^n) \\ 0 & \text{otherwise}, \qquad x^n \in T_\mu^\delta , \end{cases}$$

and the decoding rule by

$$P_D^1(z_1^n) = u_k \quad \text{if and only if} \quad z_1^n \in D_k .$$

Observe for $i^n \in T_I^\delta(x^n,j^n)$

(8) $$n^{-1}\, a^{x^n}_{i^n,j^n} > n^{-1}\, d_{V_1}(x^n,j^n) - \varepsilon ,$$

since

$$|n^{-1}\, a^{x^n}_{i^n j^n} - n^{-1}\, d_{V_1}(x^n,j^n)|$$

$$= |n^{-1} \sum_{x,i,j} N(x,i,j|x^n,i^n,j^n)\, a^x_{ij} - n^{-1} \sum_{x,j} N(x,j|x^n,j^n) \cdot d_{V_1}(x,j)|$$

$$= |\sum_{x,i,j} n^{-1} N(x,i,j|x^n,i^n,j^n) \cdot a^x_{ij} - n^{-1} V(i|x) \cdot N(x,j|x^n,j^n)\, a^x_{ij}|$$

$$\leq |\mathfrak{X}|\cdot|I|\cdot|J|\cdot \delta \cdot \max_{x,i,j} \{|a^x_{ij}|\}$$

$$< \varepsilon ,$$

the last inequality being due to (7).

SECTION IV : The Converse

The proof of a coding theorem only provides a lower bound to the power of all implementable coding procedures. To obtain a true description of the payoff induced by the given model a converse is to be proven. It gives those bounds which cannot be transgressed by any technical system.

For the model described in section I it is used to ensure that the reacting player - evaluating maxmin and minmax of the payoff - cannot avoid certain payoffs.

Since this impossibility result has to be applied for player I and player II as the reacting players, the "converse" to be given below will consist of two parts (just as the information-transmission theorem derived within the preceding section).

For $\mu \in \mathcal{P}(\mathcal{X})$,

$d : \mathcal{X} \times \hat{\mathcal{X}} \longrightarrow \mathbb{R}$

define

$\underline{R}^d_\mu(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}$

$\underline{D}^d_\mu(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$\underline{R}^d_\mu(D) := \min_{\substack{V: \\ E_{\mu,V}[d] \leq D}} \{I(\mu,V)\}$$

(1) and

$$\underline{D}^d_\mu(R) := \min_{\substack{V: \\ I(\mu,V) \leq R}} \{E_{\mu,V}[d]\},$$

where $I(\mu,V)$ is the mutual information function and

$E_{\mu,V}[d] = \sum_x \mu(x) \sum_{\hat{x}} V(\hat{x}|x)\, d_V(x,\hat{x})$ denotes the expected distortion.

$\underline{R}^d(\cdot)$ and $\underline{D}^d(\cdot)$ are the well-known rate-distortion- and distortion-rate-functions with respect to the discrete, memoryless source $\{\mathfrak{X}, \mu\}$ and the fidelity criterion d. According to section II $\underline{R}^d(\cdot)$ and $\underline{D}^d(\cdot)$ are inverse to each other on the domain where the inverse exists.

Additionally we define

$$\bar{R}^d_\mu(D) := \min_{\substack{V: \\ E_{\mu,V}[d] \geq D}} \{I(\mu,V)\}$$

$$\bar{D}^d_\mu(R) := \max_{\substack{V: \\ I(\mu,V) \leq R}} \{E_{\mu,V}[d]\}$$

These functions are inverse to each other too (see section II).

Let $\{\mathfrak{X}, \mu\}$ be a discrete, memoryless source, $\hat{\mathfrak{X}}$ a finite set and $d : \mathfrak{X} \times \hat{\mathfrak{X}} \longrightarrow \mathbb{R}$ a single-letter fidelity criterion. Let $W \mid \mathfrak{Y} \Longrightarrow \mathfrak{Z}$ denote a discrete, memoryless channel. For this setup the converse to the information-transmission theorem is given as follows:

<u>4.1 Theorem</u>

For every $n \in \mathbb{N}$, encoding- and decoding rules

$$P_E \mid \mathfrak{X}^n \longrightarrow \mathfrak{Y}^n$$
$$P_D \mid \mathfrak{Z}^n \Longrightarrow \hat{\mathfrak{X}}^n :$$

(1) $$n^{-1} \sum_{x^n} \mu(x^n) \sum_{y^n} P_E(y^n|x^n) \sum_{z^n} W^n(z^n|y^n) \sum_{\hat{x}^n} P_D(\hat{x}^n|z^n) \cdot d(x^n,\hat{x}^n) \leq \bar{D}^d_\mu(C)$$

and

(2) $$n^{-1} \sum_{x^n} \mu(x^n) \sum_{y^n} P_E(y^n|x^n) \sum_{z^n} W^n(z^n|y^n) \sum_{\hat{x}^n} P_D(\hat{x}^n|z^n)\, d(x^n,\hat{x}^n) \geq \underline{D}^d_\mu(C)$$

where $C = \max_P \{I(P,W)\}$ is the capacity of the channel W.

Proof:

For the first inequality a proof can be given which follows the lines of a well-known proof for the second (see [2],pp.70-71). Isotonicity of $\bar{R}^d_\mu(\cdot)$ instead of the antitonicity of the rate-distortion function $\underline{R}^d_\mu(\cdot)$ is to be used.

4.2 Corollary:

Under the assumptions of theorem 4.1:

$$(3) \qquad n^{-1} \sum_{x^n \in T^\delta_\mu} \mu(x^n) \sum_{y^n} P_E(y^n|x^n) \sum_{z^n} W^n(z^n|y^n) \cdot \sum_{\hat{x}^n} P_D(\hat{x}^n|z^n)\, d(x^n,\hat{x}^n) \leq \bar{D}^d_\mu (C) + \varepsilon$$

and

$$(4) \qquad n^{-1} \sum_{x^n \in T^\delta_\mu} \mu(x^n) \sum_{y^n} P_E(y^n|x^n) \sum_{z^n} W^n(z^n|y^n) \sum_{\hat{x}^n} P_D(\hat{x}^n|z^n)\, d(x^n,\hat{x}^n) \geq \underline{D}^d_\mu (C) - \varepsilon .$$

Proof:

The corollary is an immediate consequence of theorem 4.1 and of (4) , section II.

SECTION V: The value and its computability

In sections III and IV the information-theoretical basis for the determination of the value of the game defined in section I was given. In this section the information-transmission theorem and its converse will be combined to give a computable formula for an upperbound to the asymptotics of minmax of the payoff on one hand and a computable lower bound to the asymptotics of maxmin of the payoff on the other hand. Using the fact that maxmin never exceeds minmax, in order to give a computable formula for the value, we only have to prove the coincidence of the lower and upper bounds. This will be done by applying the minmax theorem after having shown that the functions over which maxmin and minmax have to be formed are in fact identical.

5.1 Theorem:

For any $\varepsilon^* > 0$ and $\varepsilon = (\varepsilon^*)$, $\delta = \delta(\varepsilon)$, sufficiently small there exists $n_o(|\mathcal{X}|,|I|,|J|,|\mathcal{Y}_1|,|\mathcal{Y}_2|,|\mathcal{Z}_1|,|\mathcal{Z}_2|, \varepsilon^*, \varepsilon, \delta)$ such that for $n \geq n_o$:

$$(1)\quad \sup_{(P_E^1,P_D^1)} \inf_{(P_E^2,P_D^2)} \{ \sum_{x^n} \mu(x^n)(\sum_{y_1^n} P_E^1(y_1^n|x^n) \sum_{z^n} W_1^n(z_1^n|y_1^n) \sum_{i^n} P_D^1(i^n|z_1^n) \cdot (\sum_{y_2^n} P_E^2(y_2^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n) \sum_{j^n} P_D^2(j^n|z_2^n)) \cdot a^{x^n}_{i^n j^n}\}$$

$$\geq \max_{V_1 \in \mathcal{V}_I^{\mu}(C_1-3\varepsilon)} \{ \underline{D}_{\mu}^{d_{V_1}}(C_2)\} - 3\varepsilon^*$$

and

$$(2)\quad \inf_{(P_E^2,P_D^2)} \sup_{(P_E^1,P_D^1)} \{ \sum_{x^n} \mu(x^n)(\sum_{y_1^n} P_E^1(y_1^n|x^n) \sum_{z_1^n} W_1^n(z_1^n|y_1^n) \sum_{i^n} P_D^1(i^n|z_1^n)) \cdot (\sum_{y_2^n} P_E^2(y_2^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n) \sum_{j^n} P_D^2(j^n|z_2^n))\, a^{x^n}_{i^n j^n}\}$$

$$\leq \min_{V_2 \in \mathcal{V}_J^{\mu}(C_2-3\varepsilon)} \{ \bar{D}_{\mu}^{d_{V_2}}(C_1)\} + 3\varepsilon^* .$$

<u>Proof:</u>

For symmetry it is sufficient to prove (1).

Assuming player I uses a strategy as being given in the information-transmission theorem we may start calculating

$$n^{-1} \sup_{(P_E^1,P_D^1)} \inf_{(P_E^2,P_D^2)} \{ \sum_{x^n} \mu(x^n) \, (\sum_{y_1^n} P_E^1(y_1^n|x^n) \sum_{z_1^n} W_1^n(z_1^n|y_1^n) \sum_{i^n} P_D^1(i^n|z_1^n)$$

$$(\sum_{y_2^n} P_E^2(y_2^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n) \sum_{j^n} P_D^2(j^n|z_2^n)) \, a^{x^n}_{i^n j^n}\}$$

$$\geq \sup_{\substack{P_E^1, \\ P_D^1 : \mathfrak{Z}_1^n \to I^n}} \inf_{(P_E^2,P_D^2)} \{\sum_{x^n \in T_\mu^\delta} \mu(x^n)(\sum_{y_2^n} P_E^2(y_2^n|x^n) \sum_{z_2^n} W^n(z_2^n|y_2^n) \sum_{j^n} P_D^2(j^n|z_2^n))$$

$$\cdot \, [n^{-1} \sum_{y_1^n} P_E^1(y_1^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n) \, a^{x^n}_{P_D^1(i^n) j^n}]$$

$$+ \sum_{x^n \notin T_\mu^\delta} \mu(x^n)(\sum_{y_2^n} P_E^2(y_2^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n) \sum_{j^n} P_D^2(j^n|z_2^n))$$

$$\cdot \, (\sum_{y_1^n} P_E^1(y_1^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n)) \, n^{-1} \min_{i^n} \{a^{x^n}_{i^n j^n}\}\}$$

$$\geq \sup_{\substack{V_1 \in \\ \mathcal{V}_I^\mu(C_1-3\varepsilon)}} \inf_{(P_E^2,P_D^2)} \{ \sum_{x^n \in T_\mu^\delta} \mu(x^n) \sum_{y_2^n} P_E^2(y_2^n|x^n) \sum_{z_2^n} W_2^n(z_2^n|y_2^n)$$

$$\bullet \sum_{j^n} P_D^2(j^n|z_2^n)) \, (d_{V_1}(x^n,j^n) - \varepsilon^*)$$

$$- \underset{x^n \notin T_\mu^\delta}{} \mu(x^n) \, n^{-1} \max_{i^n,j^n} \{|a^x_{i^n j^n}|\}\}$$

$$\geq \sup_{V_1 \in \mathcal{V}^{\mu}_{I}(C_1-3\varepsilon)} \{ \underline{D}_{\mu}^{d_{V_1}} (C_2)\} - 3\varepsilon^* .$$

The second inequality is due to inequality (5) of the information transmission theorem whereas the last is derived by the application of the corollary to the converse of the information transmission theorem.

According to this theorem the value of the asymptotic game - we shall denote it by val Γ^{∞} (μ, W_1, W_2, A) when it is proven to exist - is lower-bounded by

$$\sup_{\varepsilon^*, \varepsilon(\varepsilon^*)>0} \sup_{V_1 \in \mathcal{V}^{\mu}_{I}(C_1-\varepsilon)} \{ \underline{D}_{\mu}^{d_{V_1}} (C_2)\} - \varepsilon^*$$

and upper-bounded by

$$\inf_{\varepsilon^*, \varepsilon(\varepsilon^*)>0} \inf_{V_2 \in \mathcal{V}^{\mu}_{J}(C_2-\varepsilon)} \{ \bar{D}_{\mu}^{d_{V_2}} (C_1) + \varepsilon^* .$$

Recalling the definitions we observe that $\underline{D}^{\bullet}_{\mu}(C)$ and $\bar{D}^{\bullet}_{\mu}(C)$ are continuous functions, with respect to the fidelity criterion. Additionally remembering that $\mathcal{V}^{\mu}_{I}(\cdot)$ and $\mathcal{V}^{\mu}_{J}(\cdot)$ are defined by means of the mutual information function $I(\mu,V)$ which itself is continuous in the conditional probability distribution, we infer that the above bounds are equal to

$$\sup_{\substack{V_1 \in \\ \mathcal{V}_I^\mu(C_1)}} \{\underline{D}_\mu^{d_{V_1}}(C_2)\} = \max_{\substack{V_1: \\ I(\mu,V_1)\leq C_1}} \{\underline{D}_\mu^{d_{V_1}}(C_2)\}$$

and

$$\inf_{\substack{V_2 \in \\ \mathcal{V}_J^\mu(C_2)}} \{\bar{D}_\mu^{d_{V_2}}(C_1)\} = \min_{\substack{V_2: \\ I(\mu,V_2)\leq C_2}} \{\bar{D}_\mu^{d_{V_2}}(C_1)\} .$$

5.2 Corollary

$$\mathrm{val}\,(\Gamma^\infty(\mu,\, W_1,\, W_2,\, A)) = \max_{\substack{V_1: \\ I(\mu,V_1)\leq C_1}} \{\underline{D}_\mu^{d_{V_1}}(C_1)\}$$

$$= \min_{\substack{V_2: \\ I(\mu,V_2)\leq C_2}} \{\bar{D}_\mu^{d_{V_2}}(C_2)\}$$

Proof:

In view of theorem 5.1 and the preceding remarks there remains to show the second equality. Substituting the definitions of $\underline{D}_\mu^d$ and $\bar{D}_\mu^d$ it is to be shown:

$$\max_{\substack{V_1: \\ I(\mu,V_1)\leq C_1}} \min_{\substack{V_2: \\ I(\mu,V_2)\leq C_2}} \{E_{\mu,V_2}[d_{V_1}]\} = \min_{\substack{V_2: \\ I(\mu,V_2)\leq C_2}} \max_{\substack{V_1: \\ I(\mu,V_1)\leq C_1}} \{E_{\mu,V_1}[d_{V_2}]\}$$

Now

$$E_{\mu,V_2}[d_{V_1}] = \sum_x \mu(x) \sum_j V_2(j|x) \left(\sum_i V_1(i|x) \cdot a_{i\,j}^x\right) = E_{\mu,V_1}[d_{V_2}] .$$

Obviously, $E_{\mu,V_2}[d_{V_1}]$ is convex and concave in both, V_1 and V_2 .

Further, according to the convexity of $I(\mu,\cdot)$ in V , both $\vartheta^{\mu}_{I}(C)$ and $\vartheta^{\mu}_{J}(C)$ are convex sets for all $C > 0$. Thus the minmax theorem shows the identity of minmax and maxmin and thereby the existence of the value for the asymptotic game.

REFERENCES

[1] Aumann, R. / Maschler, M.: Repeated Games with Incomplete Information. A Survey of Recent Results, Report to the U.S. Arms Control and Disarmament Agency, Washington D.C. Final Report on Contract ACDA/St-116, chapter III, prepared by Mathematica, Princeton N.J., (1967), pp. 287-403.

[2] Berger, T.: Rate Distortion Theory, Prentice Hall, Englewood Cliffs, N.J., (1971).

[3] Csiszar, I. / Körner, J.: Information Theory, Akadêmiai Kiadō, Budapest, (1981).

[4] Wallmeier, H.-M.: Games with Informants, doctoral thesis, partly submitted to the Int.J. of Game Theory.

" WIRTSCHAFTSTHEORETISCHE ENTSCHEIDUNGSFORSCHUNG"

A series of books published by the Institute of Mathematical Economics, University of Bielefeld.

Wolfgang Rohde
Ein spieltheoretisches Modell eines Terminmarktes (A Game Theoretical Model of a Futures Market)
The model takes the form of a multistage game with imperfect information and strategic price formation by a specialist. The analysis throws light on theoretically difficult empirical phenomena.

Vol. 1 176 pages price: DM 24,80

Klaus Binder
Oligopolistische Preisbildung und Markteintritte (Oligopolistic Pricing and Market Entry)
The book investigates special subgame perfect equilibrium points of a three-stage game model of oligopoly with decisions on entry, on expenditures for market potential and on prices.

Vol. 2 132 pages price: DM 22,80

Karin Wagner
Ein Modell der Preisbildung in der Zementindustrie (A Model of Pricing in the Cement Industry)
A location theory model is applied in order to explain observed prices and quantities in the cement industry of the Federal Republic of Germany.

Vol. 3 170 pages price: DM 24,80

Rolf Stoecker
Experimentelle Untersuchung des Entscheidungsverhaltens im Bertrand-Oligopol (Experimental Investigation of Decision-Behavior in Bertrand-Oligopoly Games)
The book contains laboratory experiments on repeated supergames with two, three and five bargainers. Special emphasis is put on the end-effect behavior of experimental subjects and the influence of altruism on cooperation.

Vol. 4 197 pages price: DM 28,80

Angela Klopstech
Eingeschränkt rationale Marktprozesse (Market processes with Bounded Rationality)
The book investigates two stochastic market models with bounded rationality, one model describes an evolutionary competitive market and the other an adaptive oligopoly market with Markovian interaction.

Vol. 5 104 pages price: DM 29,80

Orders should be sent to:

Pfeffersche Buchhandlung, Alter Markt 7, 4800 Bielefeld 1, West Germany.