ISRAEL JOURNAL OF MATHEMATICS, Vol. 34, No. 3, 1979

ON INTRINSIC ERGODICITY OF PIECEWISE MONOTONIC TRANSFORMATIONS WITH POSITIVE ENTROPY

BY
FRANZ HOFBAUER

ABSTRACT

We consider a class of piecewise monotonically increasing functions f on the unit interval I. We want to determine the measures with maximal entropy for these transformations. In part I we construct a shift-space Σ_f^+ isomorphic to (I, f) generalizing the β-shift and another shift Σ_M over an infinite alphabet, which is of finite type given by an infinite transition matrix M. Σ_M has the same set of maximal measures as (I, f) and we are able to compute the maximal measures of Σ_M. In part II we try to bring these results back to (I, f). There are only finitely many ergodic maximal measures for (I, f). The supports of two of them have at most finitely many points in common. If (I, f) is topologically transitive it has unique maximal measure.

PART I

0. Introduction

We consider the dynamical system (I, f), where $I = [0, 1]$, the transformation f on I is piecewise monotonically increasing, i.e. there are disjoint intervals $J_1, J_2, \cdots, J_n$ satisfying $\bigcup J_i = I$ and $f|J_i$ is strictly increasing and continuous. We need two further conditions:

(a) $\bigcup_{m=0}^{\infty} f^{-m}\{j_0, j_1, \cdots, j_n\}$ is dense in I, where $0 = j_0 < j_1 < \cdots < j_n = 1$ are the end points of the J_i's, i.e. $(J_1, \cdots, J_n)$ is a generator for the dynamical system (I, f),

(b) $h_{\text{top}}(f) > 0$.

Our goal is to determine the set of maximal measures, i.e. the set of those invariant measures on (I, f), whose entropy h_μ is equal to the topological entropy $h_{\text{top}}(f)$ of (I, f).

Received June 30, 1978

For a dynamical system (X, T) we shall call $N \subset X$ a small set, if for every T-invariant measure μ with $\mu(N) = 1$ the entropy h_μ is zero. An isomorphism modulo small sets (meaning that the two dynamical systems are isomorphic after having taken away a small set from each one) preserves the set of maximal measures. (Suppose μ is maximal; $\mu = r\nu_1 + (1-r)\nu_2$, where ν_2 is concentrated on a small set N and ν_1 on $X - N$. Because $h_{\nu_2} = 0$ we have $h_{\nu_1} = (1/r)h_\mu$ and because μ is maximal we have $r = 1$ and $\mu = \nu_1$. Hence every small set is a null set for a maximal measure.)

We construct two isomorphisms modulo small sets

$$\varphi : (I, f) \to (\Sigma_f^+, \sigma)$$

where Σ_f^+ is a subshift of $\Sigma_n^+ = (\{1, \cdots, n\}^N, \sigma)$ and σ the shift transformation, and

$$\psi : (\Sigma_f, \sigma) \to (\Sigma_M, \sigma)$$

where Σ_f is the natural extension of Σ_f^+ (there is a 1-1 correspondence between the set of invariant measures of a shift space and that of its natural extension leaving the entropy invariant) and Σ_M is a finite type subshift of C^Z, where C is a countable compact set ($C \cong \{1, \frac{1}{2}, \frac{1}{3}, \cdots, 0\} \subset R$). M denotes the corresponding infinite transition matrix.

Hence we have that (Σ_M, σ) has the same set of maximal measures as (I, f) via $\mu \to \mu \circ \varphi^{-1} =: \nu$ and $\nu \to \nu \circ \psi^{-1}$.

$$\Sigma_f^+ = \{x \in \Sigma_n^+ : a^{x_m} \leqq x_m x_{m+1} x_{m+2} \cdots \leqq b^{x_m}\}$$

for certain $a^i, b^i \in \Sigma_n^+$ $(1 \leqq i \leqq n)$. $\leqq$ is with respect to the lexicographic ordering in Σ_n^+. This is a generalization of the β-shift (cf. [5]). Σ_f^+ and Σ_M have already been constructed in [2]. The β-shift case of ψ and Σ_M can be found in [3] and [6]. The transformation $u \to uM$ is an operator on the space of all continuous summable functions u on C with 1-norm. Let $r(M)$ be the spectral radius of this operator. We prove that

$$h_{\text{top}}(\Sigma_f^+) = \log r(M) = \lim \frac{1}{k} \log \| M^k \|_1$$

and therefore $h_{\text{top}}(f) = h_{\text{top}}(\Sigma_f^+) = h_{\text{top}}(\Sigma_M) = \log r(M)$.

If M is not irreducible, we may divide M into irreducible submatrices $M^1, M^2, M^3, \cdots$. For every submatrix M^i there is a corresponding subshift Σ_{M^i} of Σ_M.

The maximal measures of a dynamical system form a compact convex subset of the set of all invariant measures together with the weak* topology. Its extremal points are the ergodic maximal measures and such measures are always concentrated on irreducible subsystems. Hence it suffices to consider the Σ_{M^i}'s.

We prove that every ergodic maximal measure is concentrated on a Σ_{M^i} with $r(M^i) = r(M)$ and that on such an M^i there is at most one maximal measure which is Markov given by

$$P_{jk} = \frac{M^i_{jk} v_k}{r(M) v_j} \qquad \text{and} \qquad \pi_j = u_j v_j,$$

where u and v are vectors satisfying $uM^i = r(M^i)u$, $M^i v = r(M^i)v$, $u_1 = 1$ and the inner product $\langle u, v \rangle = 1$.

Hence to determine the set of maximal measures of (I, f) one has to look for the irreducible submatrices M^i of M satisfying $r(M^i) = r(M)$. To do this we have to investigate the structure of M. This problem is left for the second part of this paper.

1. The isomorphism φ

Let (I, f) be the system described above satisfying (a). Define $i : I \to \{1, 2, \cdots, n\}$ by $i(x) = k$ if $x \in J_k$. Define $\varphi : I \to \Sigma_n^+ = \{1, 2, \cdots, n\}^N$ by

$$\varphi(x) = i(x) i(f(x)) i(f^2(x)) \cdots .$$

Because of (a) φ is injective. Define

$$\Sigma_f^+ = \overline{\varphi(I)} \subset \Sigma_n^+. \tag{1.1}$$

It is easy to see that $\sigma \circ \varphi = \varphi \circ f$ and that φ and $\varphi^{-1} | \varphi(I)$ are measurable.

We introduce the lexicographical ordering on Σ_n^+ and Σ_f^+.

LEMMA 1. *$x < y$ in I $\Leftrightarrow$ $\varphi(x) < \varphi(y)$ in Σ_f^+ for $x, y \in I$.*

PROOF. Set $\mathbf{x} = \varphi(x)$, $\mathbf{y} = \varphi(y)$. Choose k so that $x_i = y_i$ $(0 \leq i \leq k - 1)$ and $x_k \neq y_k$, i.e. $f^i(x)$ and $f^i(y)$ are in the same J_m for $0 \leq i \leq k - 1$. Because $f | J_m$ is strictly increasing, $x < y \Leftrightarrow f^k(x) < f^k(y)$ and this is equivalent to $x_k < y_k$, i.e. $\mathbf{x} < \mathbf{y}$.

LEMMA 2. *$\Sigma_f^+ - \varphi(I)$ is countable and hence small.*

PROOF. Let $\mathbf{x} \in \Sigma_f^+ - \varphi(I)$, i.e. there is a sequence (r_k) in I with $\varphi(r_k) \to \mathbf{x}$ in

Σ_f^+. $\varphi(r_k)$ is a Cauchy sequence, hence also r_k, since the topology of Σ_f^+ is generated by cylinder sets ${}_0[i_0 i_1 \cdots i_{k-1}]$ and these sets correspond to intervals $\subset I$, whose diameter decreases to 0 as $k \to \infty$ (cf. (1.2) below). Let $y = \lim r_k$ in I and $\boldsymbol{y} = \varphi(y) \neq \boldsymbol{x}$, since $\boldsymbol{x} \not\in \varphi(I)$. Let m be the smallest integer with $y_m \neq x_m$ $(m \geqq 0)$, i.e. $f^m(y) \in J_p$ and there is an N with $f^m(r_k) \in J_q$ for $k \geqq N$ $(p \neq q)$. But $f^m(r_k) \to f^m(y)$ in I. Hence $f^m(y)$ has to be an endpoint of J_q, which belongs already to J_p. Hence

$$\Sigma_f^+ - \varphi(I) = \bigcup_{m=0}^{\infty} \sigma^{-m}\{j'_1, j'_2, \cdots, j'_{n-1}\},$$

where

$$j'_i = \lim_{t \downarrow j_i} \varphi(t) \quad \text{if } j_i \in J_{i-1} \quad \text{and} \quad j'_i = \lim_{t \uparrow j_i} \varphi(t) \quad \text{if } j_i \in J_i.$$

This set is countable.

Hence we have

THEOREM 1. *$\varphi : (I, f) \to (\Sigma_f^+, \sigma)$ is an isomorphism modulo small sets.*

To give a description generalizing that of the β-shift we have to introduce some more notation.

$${}_0[x_0 x_1 \cdots x_{m-1}] = \{\boldsymbol{y} \in \Sigma_f^+ \text{ or } \Sigma_n^+ : y_i = x_i \text{ for } 0 \leqq i \leqq m-1\}$$

denotes a cylinder set.

$$J_{x_0 x_1 \cdots x_{m-1}} = J_{x_0} \cap f^{-1} J_{x_1} \cap \cdots \cap f^{-m+1} J_{x_{m-1}} \subset I.$$

We have

$$\text{(1.2)} \qquad \varphi^{-1}({}_0[x_0 \cdots x_{m-1}]) = J_{x_0 \cdots x_{m-1}}, \quad \overline{\varphi(J_{x_0 \cdots x_{m-1}})} = {}_0[x_0 \cdots x_{m-1}].$$

Define for $k = 1, 2, \cdots, n$

$$\text{(1.3)} \qquad \begin{aligned} \boldsymbol{a}^k &= \lim_{t \in J_k, t \downarrow j_{k-1}} \varphi(t) \qquad (= \varphi(j_{k-1}) \text{ if } j_{k-1} \in J_k), \\ \boldsymbol{b}^k &= \lim_{t \in J_k, t \uparrow j_k} \varphi(t) \qquad (= \varphi(j_k) \text{ if } j_k \in J_k). \end{aligned}$$

Set $A = \{\boldsymbol{a}^1, \cdots, \boldsymbol{a}^n\}$, $B = \{\boldsymbol{b}^1, \cdots, \boldsymbol{b}^n\}$. Remark that $\boldsymbol{a}^k$ and $\boldsymbol{b}^k$ begin with k.

Define $G_{x_0 \cdots x_{m-1}} \subset \Sigma_f^+$ by ([,] denotes a closed interval in Σ_f^+ or Σ_n^+ with respect to the lexicographic ordering)

$$(1.4)\qquad \begin{aligned} G_{x_0} &= [\sigma a^{x_0}, \sigma b^{x_0}] = \sigma([a^{x_0}, b^{x_0}]), \\ G_{x_0\cdots x_{m-1}} &= \sigma([a^{x_{m-1}}, b^{x_{m-1}}] \cap G_{x_0\cdots x_{m-2}}), \end{aligned}$$

and $G'_{x_0\cdots x_{m-1}}$ by the same formulas, but as subsets of Σ_n^+. Remark that $G_{x_0\cdots x_{m-1}}$ and $G'_{x_0\cdots x_{m-1}}$ are empty or intervals in Σ_f^+ or Σ_n^+ respectively.

LEMMA 3. $G_{x_0\cdots x_{m-1}} = \sigma^m ({}_0[x_0 \cdots x_{m-1}])$ $({}_0[x_0 \cdots x_{m-1}] \subset \Sigma_f^+)$.

PROOF BY INDUCTION. The case $m = 1$ is trivial.

$$\begin{aligned} G_{x_0\cdots x_{m-1}} &= \sigma({}_0[x_{m-1}] \cap G_{x_0\cdots x_{m-2}}) \\ &= \sigma({}_0[x_{m-1}] \cap \sigma^{m-1}({}_0[x_0 \cdots x_{m-2}])) \\ &= \sigma^m ({}_0[x_0 \cdots x_{m-1}]). \end{aligned}$$

THEOREM 2. *The following are equivalent*:

(i) $x \in \Sigma_f^+$,

(ii) $a^{x_m} \leqq \sigma^m x \leqq b^{x_m}$ *for each* $m \geqq 0$,

(iii) $\sigma^m x \in G'_{x_0\cdots x_{m-1}}$ *for each* $m \geqq 1$.

PROOF. We show (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i).

Let $x \in \Sigma_f^+$. Then $\sigma^m x \in \Sigma_f^+$, i.e. $\sigma^m x = \lim \varphi(y_k)$ for suitable $y_k \in J_{x_m}$, i.e. $j_{x_m-1} \leqq y_k \leqq j_{x_m}$. Because of Lemma 1 we have $a^{x_m} \leqq \varphi(y_k) \leqq b^{x_m}$, i.e. (ii).

We show the next step by induction. Because both a^{x_0} and b^{x_0} begin with x_0 we have that $x \in [a^{x_0}, b^{x_0}]$ implies $\sigma x \in G'_{x_0} = \sigma([a^{x_0}, b^{x_0}])$. Suppose $\sigma^m x \in G'_{x_0\cdots x_{m-1}}$. Because of (ii) we have $\sigma^m x \in [a^{x_m}, b^{x_m}] \cap G'_{x_0\cdots x_{m-1}}$. This set is a subinterval of $[a^{x_m}, b^{x_m}]$ and hence its endpoints both begin with x_m, because a^{x_m} and b^{x_m} do. Therefore $\sigma^{m+1} x \in \sigma([a^{x_m}, b^{x_m}] \cap G'_{x_0\cdots x_{m-1}}) = G'_{x_0\cdots x_m}$. This is (iii).

For the third implication we have to show $x \in \Sigma_f^+$, i.e. there are $y_k \in I$ satisfying $\varphi(y_k) \to x$ in Σ_n^+. Choose $y_k \in J_{x_0\cdots x_{k-1}}$. Then $\varphi(y_k) \in {}_0[x_0 \cdots x_{k-1}]$, hence $\varphi(y_k) \to x$. It suffices to show $J_{x_0\cdots x_{k-1}} \neq \emptyset$. If this set is empty we have also $G_{x_0\cdots x_{k-1}} = \sigma^k (\overline{\varphi(J_{x_0\cdots x_{k-1}})}) = \emptyset$ (cf. (1.2) and Lemma 3) and hence $G'_{x_0\cdots x_{k-1}} = \emptyset$ (cf. (1.4)). This contradicts (iii). Hence (i) follows.

We conclude this section with a lemma we shall need later.

LEMMA 4. *Let* $x \in \Sigma_f^+$ *and* $a, b \in A$ *(or* B*). If*

$$(1.5)\qquad x_j x_{j+1} \cdots x_{j+r} = a_0 a_1 \cdots a_r$$

and

(1.6) $$x_{j+r}x_{j+r+1}\cdots x_{j+r+s}=b_0b_1\cdots b_s$$

then

(1.7) $$x_{j+r}\cdots x_{j+r+s}=b_0\cdots b_s=a_ra_{r+1}\cdots a_{r+s}.$$

PROOF. Suppose (1.7) is not satisfied, i.e. there is an i ($r\leqq i\leqq r+s-1$) such that $x_{j+r}\cdots x_{j+i}=b_0\cdots b_i=a_r\cdots a_{r+i}$ and $x_{j+i+1}=b_{i+1}>a_{r+i+1}$ (cf. (ii) of Theorem 2 and (1.5)), i.e. $\sigma^r\boldsymbol{a}>\boldsymbol{b}$. But this again contradicts Theorem 2, since $\sigma^r\boldsymbol{a}\in\Sigma_f^+$. The case $\boldsymbol{a},\boldsymbol{b}\in B$ is similar.

This lemma implies

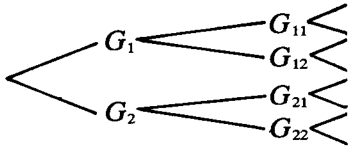
(1.8) *two initial segments of points in A (or B) contained in $\boldsymbol{x}\in\Sigma_f^+$ are disjoint or the one contains the other.*

2. The isomorphism ψ

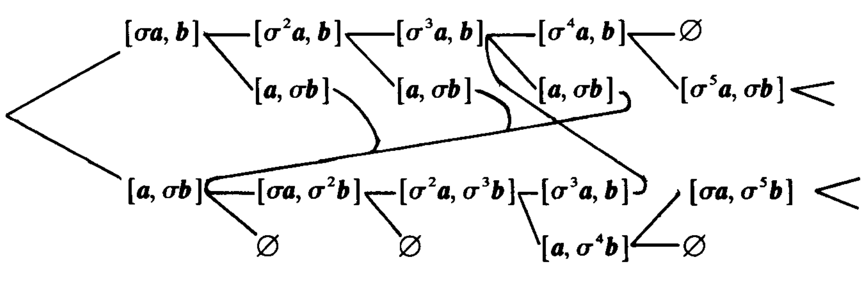
Before we go into it we consider an example. Define $f:I\to I$ by

$$f(x)=\begin{cases}\frac{16}{15}x+\frac{1}{5} & \text{for } 0\leqq x\leqq\frac{3}{4},\\ \sqrt{x-\frac{3}{4}} & \text{for } \frac{3}{4}<x\leqq 1.\end{cases}$$

One easily sees that $\boldsymbol{a}^1=111121121\cdots$, $\boldsymbol{a}^2=2\boldsymbol{a}^1$, $\boldsymbol{b}^1=1\boldsymbol{b}^2$ and $\boldsymbol{b}^2=2112112112\cdots$. We draw a diagram for Σ_f^+ using the $G_{x_0\cdots x_{m-1}}$:



For our example we get using (1.4) for the construction (we have written $\boldsymbol{a}$ for $\boldsymbol{a}^1$ and $\boldsymbol{b}$ for $\boldsymbol{b}^2$)



Every path in this diagram ending not with an empty set corresponds to an element of Σ_f^+ writing 1 if the way goes up and 2 if the way goes down.

We define ψ first on Σ_f^+ by

$$\psi(\boldsymbol{x}) = (x_0, G_{x_0})(x_1, G_{x_0 x_1})(x_2, G_{x_0 x_1 x_2}) \cdots . \tag{2.1}$$

It means we represent $\boldsymbol{x}$ not by the edges lying on a way in the above diagram, but by the vertices lying on this way.

By (1.4) one easily sees that each $G_{x_0 \cdots x_{m-1}}$ is of the form $[\sigma^k \boldsymbol{a}, \sigma^l \boldsymbol{b}]$, where $\boldsymbol{a} \in A$, $\boldsymbol{b} \in B$, $a_{k-1} = b_{l-1} = x_{m-1}$, $a_{k-j} = b_{l-j}$ $(1 \leqq j \leqq \min(k, l))$ and that $G_{x_0 \cdots x_m}$ is then

$$\begin{array}{ll} \varnothing & \text{if } x_m < a_{k+1} \text{ or } x_m > b_{l+1}, \\ [\sigma^{k+1}\boldsymbol{a}, \sigma \boldsymbol{b}^{x_m}] & \text{if } x_m = a_{k+1} \text{ and } x_m < b_{l+1}, \\ [\sigma \boldsymbol{a}^{x_m}, \sigma^{l+1}\boldsymbol{b}] & \text{if } x_m = b_{l+1} \text{ and } x_m > a_{k+1}, \\ [\sigma^{k+1}\boldsymbol{a}, \sigma^{l+1}\boldsymbol{b}] & \text{if } x_m = a_{k+1} = b_{k+1}, \\ [\sigma \boldsymbol{a}^{x_m}, \sigma \boldsymbol{b}^{x_m}] & \text{if } a_{k+1} < x_m < b_{l+1}. \end{array} \tag{2.2}$$

We set

$$\begin{aligned} D = \{(i, [\sigma^k \boldsymbol{a}, \sigma^l \boldsymbol{b}]) : \boldsymbol{a} \in A, \boldsymbol{b} \in B, i = a_{k-1} = b_{l-1}, k, l \geqq 1 \\ \text{and } a_{k-j} = b_{l-j} \text{ for } 1 \leqq j \leqq \min(k, l)\}. \end{aligned} \tag{2.3}$$

Then ψ is a map from Σ_f^+ to D^N, but the image $\psi(\Sigma_f^+)$ is not σ-invariant. So we have to change over to the natural extension

$$\Sigma_f = \{\boldsymbol{x} \in \Sigma_n = \{1, \cdots, n\}^Z : x_k x_{k+1} \cdots \in \Sigma_f^+ \text{ for each } k \in Z\}.$$

Now let $\boldsymbol{x} \in \Sigma_f$ and set for each $k \leqq 0$

$$\boldsymbol{y}^k = \psi(x_k x_{k+1} \cdots) = (x_k, G_{x_k})(x_{k+1}, G_{x_k x_{k+1}}) \cdots \in \prod_k^\infty D,$$

i.e., we have the diagram

$$\begin{array}{cccccccc} & & y_0^0 & y_1^0 & y_2^0 & y_3^0 & \cdots & = \boldsymbol{y}^0 \\ & y_{-1}^{-1} & y_0^{-1} & y_1^{-1} & y_2^{-1} & y_3^{-1} & \cdots & = \boldsymbol{y}^{-1} \\ y_{-2}^{-2} & y_{-1}^{-2} & y_0^{-2} & y_1^{-2} & y_2^{-2} & y_3^{-2} & \cdots & = \boldsymbol{y}^{-2} \\ \cdots & & \cdots & & \cdots & & \cdots & \end{array}$$

Let N be the following invariant subset of Σ_f:

$$N = \{x \in \Sigma_f : \exists m \in Z, \text{so that } \forall j < m, \exists k \leqq j \text{ with } x_k \cdots x_m = a_0 \cdots a_{m-k},$$
$$a \in A \cup B \text{ and } \cdots x_{m-1}x_m \text{ is not periodic for each } m\}$$

LEMMA 5. *For $x \in \Sigma_f - N$ there is a $y \in D^Z$ with $y^k \to y$ $(k \to -\infty)$, i.e. the columns in the diagram above are ultimately constant.*

PROOF. We show that for every $m \in Z$ there is a K with $y_i^k = y_i$ for all $i \geqq m$ and all $k \leqq K$.

As $x \notin N$ there is a K such that for each $k \leqq K$ we have $x_k \cdots x_m \neq a_0 \cdots a_{m-k}$ for every $a \in A \cup B$ or $\cdots x_{m-1}x_m$ is periodic for an m and there is a periodic point in $A \cup B$ having the same period (if not, we have the first case). We consider the first case. For each $k \leqq K$ divide $x_k x_{k+1} \cdots$ into initial segments of points in A. This can be done in a unique way starting at the beginning of $x_k x_{k+1} \cdots$ and choosing every such initial segment maximal (cf. the remark after (1.3)). Let l be a $k \leqq K$ such that the first initial segment $x_l x_{l+1} \cdots x_r$ of a^{x_l} of the above division has the largest possible r $(r < m)$. Then the first initial segment $x_k \cdots x_q$ of a^{x_k} for all other $k \leqq r$ satisfies $q \leqq r$. For $k < l$ this is because we have chosen l in such a way. For $l < k \leqq r$ $(r \geqq K)$ this is because of (1.8). Hence we have for all $k \leqq K$ (in fact for all $k \leqq r$)

$$y^k = (x_k, [\sigma a^{x_k}, \quad]) \cdots \quad (x_r, [\sigma^t a^{x_l}, \quad]) \quad (x_{r+1}, [\sigma a^{x_{r+1}}, \quad]) \quad \cdots .$$

The same arguments hold if we divide $x_k x_{k+1} \cdots$ into initial segments of points in B. We find an $s < m$ such that for all $k \leqq K$

$$y^k = (x_k, [\quad , \sigma b^{x_k}]) \quad \cdots (x_s, [\quad , \sigma^v b^{x_j}]) \quad (x_{s+1}, [\quad , \sigma b^{x_{s+1}}]) \quad \cdots .$$

Hence for all $k \leqq K$ the beginning points of the above intervals do not depend on k from the $(r+1)$st coordinate onward and the endpoints of these intervals from the $(s+1)$st coordinate onward (cf. (2.1) and (2.2)). As $r < m$ and $s < m$ we have proved the first case.

The second case, where $\cdots x_{m-1}x_m$ is periodic for an m and there is a periodic point in $A \cup B$ having the same period, is easy and omitted. It may happen that for the beginning points of the above intervals the first case occurs and for the endpoints the second one, or vice versa.

Now we define (y as in Lemma 5)

$$\text{(2.4)} \qquad \psi : \Sigma_f - N \to D^Z \quad \text{by} \quad \psi(x) = y.$$

By (2.2) one sees easily that a given element $(i, [\sigma^k a, \sigma^l b])$ of D may be followed in a point $y \in \psi(\Sigma_f - N)$ by

$$(a_k, [\sigma^{k+1}\boldsymbol{a}, \sigma\boldsymbol{b}^{a_k}]), \quad (i, [\sigma\boldsymbol{a}^i, \sigma\boldsymbol{b}^i]) \qquad \text{for } a_k < i < b_l;$$

$$(2.5) \qquad (b_l, [\sigma\boldsymbol{a}^{b_l}, \sigma^{l+1}\boldsymbol{b}]) \qquad \text{if } a_k < b_l,$$

$$(a_k, [\sigma^{k+1}\boldsymbol{a}, \sigma^{l+1}\boldsymbol{b}]) \qquad \text{if } a_k = b_l.$$

Denote the corresponding transition matrix with index set D by M'. ψ^{-1} is given by the projection to the first component of the elements of D.

LEMMA 6. *Let $\boldsymbol{y} \in D^Z$ satisfy (2.5). Then $\boldsymbol{x} = \psi^{-1}(\boldsymbol{y}) \in \Sigma_f$.*

PROOF. We have to show $\boldsymbol{a}^{x_k} \leqq x_k x_{k+1} \cdots \leqq \boldsymbol{b}^{x_k}$ for all $k \in Z$ ((ii) of Theorem 2). Suppose $y_k = (x_k, [\sigma^r\boldsymbol{a}, \sigma^s\boldsymbol{b}])$, $\boldsymbol{a} \in A$, $\boldsymbol{b} \in B$. Then $x_k = a_{r-1} = b_{s-1}$ (cf. (2.3)) and hence $\boldsymbol{a}^{x_k} \leqq \sigma^{r-1}\boldsymbol{a}$ and $\boldsymbol{b}^{x_k} \geqq \sigma^{s-1}\boldsymbol{b}$ ((ii) of Theorem 2). Therefore it suffices to show

$$(2.6) \qquad \sigma^r\boldsymbol{a} \leqq x_{k+1}x_{k+2}\cdots \leqq \sigma^s\boldsymbol{b}.$$

Because of (2.3) and (2.5) we have $a_r \leqq x_{k+1} \leqq b_s$. If equality does not hold our lemma is proved. Suppose for $t \geqq 1$

$$(2.7) \qquad a_r a_{r+1} \cdots a_{r+t-1} = x_{k+1}x_{k+2}\cdots x_{k+t}.$$

We have to show $a_{r+t} \leqq x_{k+t+1}$. Because of (2.3) and (2.5) it follows from (2.7) that $y_{k+t} = (x_{k+t}, [\sigma^{r+t}\boldsymbol{a}, \quad])$. As above we have from (2.3) that $a_{r+t} \leqq x_{k+t+1}$.

The same arguments show that $x_{k+1}x_{k+2}\cdots \leqq \sigma^s\boldsymbol{b}$.

It is easy to see that ψ and ψ^{-1} are measurable and that ψ commutes with the shift transformation. Hence ψ is an isomorphism between $\Sigma_f - N$ and the finite type subshift of D^Z given by (2.5).

We make D^Z compact. D is a countable set. Let C be its one-point compactification by adding say 0. We take the closure of $\psi(\Sigma_f - N)$ in C^Z. This is again a finite type subshift of C^Z with transition matrix M (index set C) satisfying $M \mid D = M'$. We denote it by Σ_M. It is easy to see that one gets M by adding to (2.5): 0 follows only after 0 and 0 may be followed by every element of D which may itself follow after infinitely many elements of D and by 0.

Hence $\Sigma_M - \psi(\Sigma_f - N)$ consists of all points, for which there is an $m \in Z$ with $y_k = 0$ for all $k \leqq m$. Hence

LEMMA 7. *$\Sigma_M - \psi(\Sigma_f - N)$ is small.*

To prove that N is small we remark that $N = \bigcup N_{\boldsymbol{a}}$, the union taken over all

$a \in A \cup B$, which are not periodic, where

$$N_a = \bigcup_{j \in Z} \sigma^j \bigcap_{l=1}^{\infty} \bigcup_{i=l}^{\infty} \sigma^{-i}({}_0[a_0 a_1 \cdots a_{i-1}]).$$

It suffices to prove that N_a is small (N_a is invariant). This is done in [3] for the β-shift and exactly the same proof works for N_a using Lemma 4 or (1.8) instead of the lemma used there. Hence

LEMMA 8. *N is small.*

All this together gives

THEOREM 3. *$\psi : (\Sigma_f, \sigma) \to (\Sigma_M, \sigma)$ is an isomorphism modulo small sets.*

3. A formula for h_{top}

To compute the topological entropy $h_{\text{top}}(\Sigma_f^+) = \lim \log \sqrt[k]{n_k}$, where n_k is the number of admissible blocks of length k in Σ_f^+, we introduce the following sets of blocks, where $k \geqq 1$ and $i \in D$:

$$\mathcal{N}_k^i = \{x_0 x_1 \cdots x_{k-1} : {}_0[x_0 \cdots x_{k-1}] \neq \emptyset \text{ in } \Sigma_f^+,\ \psi(x_0 \cdots x_{k-1}) \text{ ends with } i\}.$$

Set $N_k^i = \operatorname{card} \mathcal{N}_k^i$ and $N_k = (N_k^i)_{i \in D}$, a vector in l^1.

We obtain the set $\mathcal{N}_{k+1}^j$ by adding an x_k to $x_0 \cdots x_{k-1} \in \bigcup_i \mathcal{N}_k^i$, such that ${}_0[x_0 \cdots x_k] \subset \Sigma_f^+$ is not empty and that $\psi(x_0 \cdots x_k)$ ends with j. This is possible for all blocks $x_0 \cdots x_{k-1}$ in $\mathcal{N}_k^i$, if i may be followed by j in Σ_M, i.e. $M_{ij} = 1$. Hence we have, as x_k has to be equal to the first component of j (cf. (2.1)),

$$N_{k+1}^j = \sum_{i, M_{ij}=1} N_k^i = \sum_{i \in D} M_{ij} N_k^i, \qquad \text{i.e.} \quad N_{k+1} = N_k M'.$$

From this one gets $N_k = N_1 M'^k$ ($N_1 = (1, \cdots, 1, 0, 0, \cdots)$ n ones). As $n_k = \| N_k \|_1 = \Sigma_{i \in D} N_k^i$ (N_k has only finitely many entries $\neq 0$) we have

$$h_{\text{top}}(\Sigma_f^+) = \lim \frac{1}{k} \log \| N_1 M'^k \|_1 = \lim \frac{1}{k} \log \| M'^k \|_1$$

because with increasing k, $N_1 M'^k = N_k$ becomes positive on every coordinate and hence on every irreducible subset of the index set D of M'. We have

THEOREM 4. $h_{top}(f) = h_{top}(\Sigma_f^+) = h_{top}(\Sigma_M) = \log r(M')$.

REMARK. If we consider M as an operator on the space of all continuous summable functions on $C \cong \{1, \frac{1}{2}, \frac{1}{3}, \cdots, 0\} \subset R$ with 1-norm, then $\|M^k\|_1 = \|M'^k\|_1$, hence $r(M) = r(M')$ and $h_{top}(\Sigma_f^+) = r(M)$. In what follows we can restrict ourselves again to D and M', because $\Sigma_M - \Sigma_{M'}$ is a nullset for every maximal measure. We have introduced C and M only to get a compact dynamical system.

4. Eigenvectors of M'

Now divide M' into irreducible submatrices $M^1, M^2, M^3, \cdots$ (if we divide M into irreducible submatrices we get the same M^i's plus the 1×1 matrix indexed by $\{0\} = C - D$). We are interested in ergodic maximal measures, which are concentrated on subshifts Σ_{M^i} of Σ_M corresponding to the submatrices M^i. Fix an M^i and denote it by L and its index set by $E \subset D$. What follows is a slight generalization of a result in [6]. Suppose there is a maximal measure μ concentrated on Σ_L.

LEMMA 9. *μ is a Markov measure given by*

$$\mu({}_0[y_0 \cdots y_{k-1}]) = \pi_{y_0} P_{y_0 y_1} P_{y_1 y_2} \cdots P_{y_{k-2} y_{k-1}}$$

for suitable $\pi = (\pi_i)_{i \in E}$ and $P = (P_{ij})_{(i,j) \in E \times E}$ satisfying

$$(4.1) \qquad \begin{aligned} &\pi_i \geqq 0, \quad P_{ij} \geqq 0 \quad \textit{and} \quad P_{ij} = 0 \qquad \textit{if } L_{ij} = 0, \\ &\sum_{i \in E} \pi_i P_{ij} = \pi_j, \quad \sum_{i \in E} \pi_i = 1, \quad \sum_{j \in E} P_{ij} = 1 \qquad \textit{for each } i \in E. \end{aligned}$$

PROOF. Set $\pi_i = \mu({}_0[i])$ and $P_{ij} = \mu({}_0[ij])/\mu({}_0[i])$. Then (4.1) is satisfied. Let m be the corresponding Markov measure.

Set $\alpha = \{{}_0[i]\}_{i \in E}$. Then

$$h_\mu = H_\mu(\alpha \mid \bigvee_{k \geqq 1} \sigma^{-k}\alpha) \leqq H_\mu(\alpha \mid \sigma^{-1}\alpha) = H_m(\alpha \mid \sigma^{-1}\alpha) = h_m$$

and equality holds only if $\mu = m$ by a theorem of Parry as in [6]. Because μ is maximal we have $\mu = m$ is Markov.

Now assign to every pair $(i, j) \in E \times E$ with $L_{ij} = 1$ a variable X_{ij} satisfying

$$(4.2) \qquad X_{ij} \geqq 0, \quad \sum_{j, L_{ij} = 1} X_{ij} = \sum_{j, L_{ij} = 1} X_{ji} \quad (i \in E), \quad \sum_{i, j, L_{ij} = 1} X_{ij} = 1.$$

Set $X = (X_{ij})$. There is a 1-1 correspondence between (π, P) satisfying (4.1) and X satisfying (4.2) via $X_{ij} = \pi_i P_{ij}$ and $\pi_i = \Sigma_k X_{ik}$, $P_{ij} = X_{ij}/\Sigma_k X_{ik}$. The entropy of μ is $H(\pi, P) = -\Sigma_{i,j}\pi_i P_{ij}\log P_{ij}$. Define

$$H(X) = -\sum_{i,j} X_{ij}\log X_{ij} + \sum_i\Big(\sum_j X_{ij}\Big)\log\Big(\sum_j X_{ij}\Big)$$

the entropy of X. Then corresponding (π, P) and X have the same entropy. Now every maximal measure concentrated on Σ_L corresponds to an X satisfying (4.2), for which $H(X)$ attains its supremum $\log r(M')$.

LEMMA 10. *Suppose H attains its supremum in X. Then $X_{ij} > 0$ for every (i, j) with $L_{ij} = 1$.*

PROOF. Suppose $X_{i_0i_1} = 0$ and $L_{i_0i_1} = 1$. Because of the irreducibility of L we can find $i_2, i_3, \cdots, i_r = i_0$ such that $L_{i_li_{l+1}} = 1$ $(0 \leqq l \leqq r-1)$ and not all $X_{i_li_{l+1}} = 0$. Without loss of generality assume $X_{i_1i_2} \neq 0$. Set $K = \{(i_l, i_{l+1}) : 0 \leqq l \leqq r-1\}$. Then by (4.2) $\Sigma_j X_{ji_1} = \Sigma_j X_{i_1j} \geqq X_{i_1i_2} > 0$. Define $X'(\varepsilon)$ by

$$X'_{ij} = \begin{cases} X_{ij}/(1+r\varepsilon) & (i,j)\notin K, \\ (X_{ij}+\varepsilon)/(1+r\varepsilon) & (i,j)\in K; \end{cases}$$

$X'(\varepsilon)$ satisfies (4.2). Set $f(\varepsilon) = H(X'(\varepsilon)) - H(X)$. One has, because of $X_{i_0i_1} = 0$ and $\Sigma_j X_{ji_1} > 0$, $f'(\varepsilon) \to \infty$ as $\varepsilon \to 0$. Hence there is an $\varepsilon > 0$ with $H(X'(\varepsilon)) > H(X)$, a contradiction to X maximal.

Now we know that H attains its supremum only in the interior of its domain determined by (4.2). By Lagrange's method the function

$$f(X, \lambda_i, \kappa) = H(X) + \sum_{i\in E}\lambda_i\Big(\sum_j X_{ij} - \sum_j X_{ji}\Big) + \kappa\Big(\sum_{i,j} X_{ij} - 1\Big)$$

has to satisfy for a maximal X:

$$\frac{\partial f}{\partial X_{ij}}(X) = 0, \qquad \frac{\partial f}{\partial \lambda_i}(X) = 0, \qquad \frac{\partial f}{\partial \kappa}(X) = 0.$$

This gives

$$\text{(4.3)} \qquad \sum_{j, L_{rj}=1} X_{rj} = \exp(\lambda_s - \lambda_r - \kappa)X_{rs}, \qquad \forall r, s \text{ with } L_{rs} = 1;$$

$$\text{(4.4)} \qquad \sum_{j, L_{ij}=1} X_{ij} = \sum_{j, L_{ji}=1} X_{ji}, \qquad \forall i \in E;$$

$$\sum_{i,j,L_{ij}=1} X_{ij} = 1. \tag{4.5}$$

By (4.3) we have

$$\exp(\lambda_s - \lambda_r - \kappa)X_{rs} = \exp(\lambda_t - \lambda_r - \kappa)X_{rt}, \qquad \text{if } L_{rs} = L_{rt} = 1,$$

i.e. $\exp(\lambda_s - \lambda_t)X_{rs} = X_{rt}$, and hence summing over t with $L_{rt} = 1$,

$$\exp(\lambda_s)X_{rs} \sum_{t,L_{rt}=1} \exp(-\lambda_t) = \sum_{t,L_{rt}=1} X_{rt}$$

and again by (4.3)

$$\exp(\lambda_s)X_{rs} \sum_{t,L_{rt}=1} \exp(-\lambda_t) = \exp(\lambda_s - \lambda_r - \kappa)X_{rs} \qquad \text{if } L_{rs} = 1.$$

Because $X_{rs} > 0$ we get

$$\sum_{t\in E} L_{rt} \exp(-\lambda_t) = \exp(-\kappa)\exp(-\lambda_r), \qquad \forall r \in E. \tag{4.6}$$

Again (4.3) gives

$$\exp(\lambda_r)\sum_j X_{rj} = \exp(\lambda_s - \kappa)X_{rs} \qquad \text{if } L_{rs} = 1,$$

or

$$\sum_{r,L_{rs}=1} \exp(\lambda_r) \sum_{j,L_{rj}=1} X_{rj} = \exp(\lambda_s - \kappa) \sum_{r,L_{rs}=1} X_{rs}, \qquad \forall s \in E,$$

i.e.

$$\sum_{r\in E} \Big(\exp(\lambda_r) \sum_{j,L_{rj}=1} X_{rj}\Big) L_{rs} = \exp(-\kappa)\Big(\exp(\lambda_s) \sum_{r,L_{rs}=1} X_{rs}\Big). \tag{4.7}$$

Hence for the eigenvalue $\exp(-\kappa)$ the matrix L has a left eigenvector u ($u_r = \exp(\lambda_r)\Sigma_j X_{rj}$) by (4.7) and a right eigenvector v ($v_r = \exp(-\lambda_r)$) by (4.6) satisfying $\langle u, v\rangle = \Sigma_{r,j} X_{rj} = 1$. Furthermore $\exp(-\kappa)$ has to be the spectral radius λ of L. We have got that, if there is a maximal measure (π, P) concentrated on Σ_L, there exists such a pair (u, v) of eigenvectors of L as above (for the eigenvalue $\lambda = r(L)$) satisfying $\langle u, v\rangle = 1$. One gets

$$\pi_i = u_i v_i, \qquad P_{ij} = L_{ij} v_j / \lambda v_i. \tag{4.8}$$

This measure has entropy $\log \lambda = \log r(L)$. Hence by its maximality $r(L) =$

$r(M')$. On the other hand one sees easily that (π, P) given by (4.8) gives rise to a measure μ on Σ_L with entropy $\log \lambda$. Hence we have

THEOREM 5. *Every ergodic maximal measure of Σ_M is concentrated on a Σ_{M^i} satisfying $r(M^i) = r(M)$. For such an M^i there is a* 1-1 *correspondence between the set of maximal measures on Σ_{M^i} and $\{(u, v) : uM^i = r(M^i)u,\ M^i v = r(M^i)v,\ \langle u, v\rangle = 1,\ u_1 = 1\}$.*

Now we can use a result of Krieger [4] for the irreducible submatrices M^k of M' satisfying $r(M^k) = r(M')$.

Suppose μ_1 and μ_2 are two ergodic maximal measures on Σ_{M^k}. Then $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ is also maximal. Let ν be the Markov measure given by $P_{ij} = \mu({}_0[ij])/\mu({}_0[i])$ and $\pi_i = \mu({}_0[i])$. ν is ergodic by theorem 7.16 of [1], because P_{ij} is irreducible ($\mu({}_0[ij]) > 0$ for all (i, j) with $M^k_{ij} = 1$ by Lemma 10). Hence $\nu \neq \mu$, because μ is not ergodic. By the proof of Lemma 9 we have $h_\nu > h_\mu$, a contradiction to the assumption that μ_1, μ_2 and hence also μ have maximal entropy. Hence we have

THEOREM 6. *If M^i satisfies $r(M^i) = r(M')$, then there is at most one maximal measure on Σ_{M^i}, which is Markov given by* (4.8).

REMARK. From this and Theorem 5 it follows that for such an M^i there is at most one pair of vectors (u, v) satisfying the conditions in Theorem 5. I have also a direct proof for this result using the special structure of M', but this is much more complicated than Krieger's method.

Summarizing the results about $\Sigma_{M'}$ we get

THEOREM 7. (i) $h_{\mathrm{top}}(\Sigma_{M'}) = \log r(M')$.

(ii) *Every ergodic maximal measure is concentrated on a Σ_{M^i} satisfying $r(M^i) = r(M')$.*

(iii) *For every such M^i there is at most one maximal measure concentrated on Σ_{M^i}.*

PART II

The second part of this paper is devoted to the investigation of the matrix M' found in the first part.

First we introduce a more convenient notation for the index set D of M'. Then we look for the irreducible submatrices M^i and determine the corresponding invariant subsets of (I, f) getting a nice picture of the structure of this transformation. We deduce that (I, f) has only finitely many ergodic maximal measures and determine their supports. If (I, f) is topologically transitive, it has unique maximal measure, which is positive on every open subset of I.

5. We introduce a new notation for the index set D of M'. Consider $A = \{\boldsymbol{a}^1, \cdots, \boldsymbol{a}^n\}$ and $B = \{\boldsymbol{b}^1, \cdots, \boldsymbol{b}^n\}$ (cf. (1.3)). Take an $\boldsymbol{a}^i$ and divide it into initial segments of $\boldsymbol{b}^j$'s in the following way. $\boldsymbol{a}^i$ begins with an initial segment of $\boldsymbol{b}^i$ of length at least one. Denote the length of this segment by $r(i, 1)$. i.e. $a^i_k = b^i_k$ for $k = 0, \cdots, r(i, 1) - 1$ and $a^i_{r(i,1)} \neq b^i_{r(i,1)}$. At this $r(i, 1)$-th coordinate there begins an initial segment of $\boldsymbol{b}^j$ for $j = a^i_{r(i,1)}$ of length $r(i, 2) \geqq 1$, and so on. So we get a division of $\boldsymbol{a}^i$ into initial segments of $\boldsymbol{b}^j$'s in a unique way. For $m \geqq 0$ we have

$$(5.1)\qquad \begin{aligned} &a^i_{r(i,1)+\cdots+r(i,m)+k} = b^j_k, \quad 0 \leqq k \leqq r(i, m+1) - 1, \quad j = a^i_{r(i,1)+\cdots+r(i,m)}, \\ &a^i_{r(i,1)+\cdots+r(i,m+1)} \neq b^j_{r(i,m+1)}, \end{aligned}$$

Similarly we divide $\boldsymbol{b}^j \in B$ into initial segments of $\boldsymbol{a}^i$'s and denote their lengths by $s(j, k)$ for $1 \leqq j \leqq n$ and $k \geqq 1$.

We have $r(i, 1) = s(i, 1)$, $1 \leqq i \leqq n$ and $r(i, k) \geqq 1$, $s(j, k) \geqq 1$. We can assume $r(i, k) < \infty$, $s(j, k) < \infty$. If $r(i, k) = \infty$ for some i, k we have for $m = r(i, 1) + \cdots + r(i, k-1)$, $\sigma^m \boldsymbol{a}^i = \boldsymbol{b}^j$ for $j = a^i_m$, i.e. ${}_0[a^i_0 \cdots a^i_m]$ consists of one point. Substituting $\boldsymbol{a}^i$ by

$$\tilde{\boldsymbol{a}}^i = a^i_0 \cdots (a^i_m + 1) a^i_0 \cdots (a^i_m + 1) a^i_0 \cdots$$

we get a new shift space forbidding the block $a^i_0 \cdots a^i_m$, i.e. we take away countably many points. Note that $a^i_m + 1 \leqq b^j_{r(i,k-1)} \leqq n$, where $j = a^i_{r(i,1)+\cdots+r(i,k-2)}$.

Lemma 11. *There are maps $R, S : N \to N$ such that*

$$r(i, m) = s(j, 1) + \cdots + s(j, R(m)), \qquad j = a^i_{r(i,1)+\cdots+r(i,m-1)}.$$

$$s(j, m) = r(i, 1) + \cdots + r(i, S(m)), \qquad i = b^j_{s(j,1)+\cdots+s(j,m-1)}.$$

Proof. We prove the first equation. We have by (5.1)

$$(5.2)\qquad \begin{aligned} &a^i_{r(i,1)+\cdots+r(i,m-1)+k} = b^j_k, \qquad 0 \leqq k < r(i, m), \\ &a^i_{r(i,1)+\cdots+r(i,m)} < b^j_{r(i,m)}. \end{aligned}$$

Choose l so that

$$(5.3)\qquad s(j,1)+\cdots+s(j,l)\leqq r(i,m)<s(j,1)+\cdots+s(j,l+1)$$

and set $R(m)=l$. This is uniquely determined. We have by definition that $s(j,1)+\cdots+s(j,R(m))\leqq r(i,m)$. Suppose

$$(5.4)\qquad s(j,1)+\cdots+s(j,R(m))<r(i,m).$$

By (5.3) we have $r(i,m)<s(j,1)+\cdots+s(j,R(m)+1)$ and by (5.1)

$$b^j_{s(j,1)+\cdots+s(j,R(m))+k}=a^h_k \quad \text{for some } h \text{ and } 0\leqq k<s(j,R(m)+1).$$

Hence

$$(5.5)\qquad b^j_k=a^h_{k-s(j,1)-\cdots-s(j,R(m))},\quad s(j,1)+\cdots+s(j,R(m))\leqq k\leqq r(i,m).$$

We get by (5.2) and (5.5)

$$a^i_{r(i,1)+\cdots+r(i,m-1)+k}=a^h_{k-s(j,1)-\cdots-s(j,R(m))}$$

for $s(j,1)+\cdots+s(j,R(m))\leqq k<r(i,m)$ (this set of k's is not empty because of (5.4)) and

$$a^i_{r(i,1)+\cdots+r(i,m)}<a^h_{r(i,m)-s(j,1)-\cdots-s(j,R(m))}.$$

This is a contradiction to $\boldsymbol{a}^i\in\Sigma^+_f$ (cf. Theorem 2). This proves the converse inequality and the lemma is proved.

LEMMA 12. *Let $y=(a^i_{k-1},[\sigma^k\boldsymbol{a}^i,\sigma^l\boldsymbol{b}^j])\in D$ such that it occurs in $\Sigma_{M'}$ and suppose $k\geqq l$. There is a unique m such that*

$$r(i,1)+\cdots+r(i,m-1)<k\leqq r(i,1)+\cdots+r(i,m-1)+r(i,m).$$

Then

(i) *$l=k-r(i,1)-\cdots-r(i,m-1)$ and $j=a^i_{k-l}$,*

(ii) *if $k=r(i,1)+\cdots+r(i,m)$ then y has successors $(j',[\sigma^{k+1}\boldsymbol{a}^i,\sigma\boldsymbol{b}^{j'}])$, $j'=a^i_k$, $(i',[\sigma\boldsymbol{a}^{i'},\sigma^{r(i,m)+1}\boldsymbol{b}^j])$, $i'=b^j_{r(i,m)}$ and $(t,[\sigma\boldsymbol{a}^t,\sigma\boldsymbol{b}^t])$ for $j'<t<i'$ and if $k<r(i,1)+\cdots+r(i,m)$ then y has only $(a^i_k,[\sigma^{k+1}\boldsymbol{a}^i,\sigma^{l+1}\boldsymbol{b}^j])$ as successor.*

Similar statements hold for $k\leqq l$ using $s(j,k)$ instead of $r(i,k)$. For $k=l$ one gets the same using either $r(i,k)$ or $s(j,k)$, because $r(i,1)=s(i,1)$.

PROOF. Suppose y occurs in $\boldsymbol{x}\in\Sigma_{M'}$ and $y=x_0$. If $l>1$ then by (2.5)

$x_{-1} = (a^i_{k-2}, [\sigma^{k-1}a^i, \sigma^{l-1}b^j])$. If $l-1>1$ we can do the same and so on and get $x_{-l+1} = (j, [\sigma^{k-l+1}a^i, \sigma b^j])$. Furthermore $a^i_{t+k-l} = b^j_t$ for $0 \leqq t \leqq l-1$, because $y \in D$ (cf. (2.3)). If $k = l$ then $a^i_t = b^j_t$ for $0 \leqq t \leqq l-1$, hence $i = j$. $r(i,1) = s(i,1) \geqq l$ and $m = 1$. (i) becomes $k = l$ and $j = i$. If $k > l$ set $z = \psi^{-1}(x) \in \Sigma_f$ (cf. §2). We have

$$z_t = a^i_{t+k-1} \quad \text{for } -k+1 \leqq t \leqq 0 \quad \text{and} \quad z_t = b^j_{t+l-1} \quad \text{for } -l+1 \leqq t \leqq 0. \tag{5.6}$$

Hence

$$b^j_t = a^i_{t+k-l} \qquad \text{for } 0 \leqq t \leqq l-1, \tag{5.7}$$

in particular $j = b^j_0 = a^i_{k-l}$, the second part of (i).

We show that for some $l' \geqq l$, $b^j_0 \cdots b^j_{l'-1}$ occurs in the division of a^i into initial segments of b^j's constructed above. Otherwise there would be a j' and an $h > l'$ (which we choose maximal) such that

$$b^{j'}_t = a^i_{t+k-h} \qquad \text{for } 0 \leqq t \leqq h-1$$

(t must run until $h-1$ because of (5.7) and Lemma 4). But then we have by (5.6)

$$z_t = b^{j'}_{t+h-1} \qquad \text{for } -h+1 \leqq t \leqq 0$$

and by definition of ψ, x_0 would be $(a^i_k, [\sigma^k a^i, \sigma^h b^{j'}])$, a contradiction to $x_0 = y$, because $h > l$. Hence at the $(k-l)$-th coordinates of a^i there begins an initial segment of b^j of the subdivision of a^i constructed above, i.e. for some m, $k - l = r(i,1) + \cdots + r(i, m-1)$ and $r(i,m) \geqq l$ by (5.7). This is (i). Part (ii) is only a translation of (2.5) and easily verified using the definitions of $r(i,k)$ and $s(j,k)$.

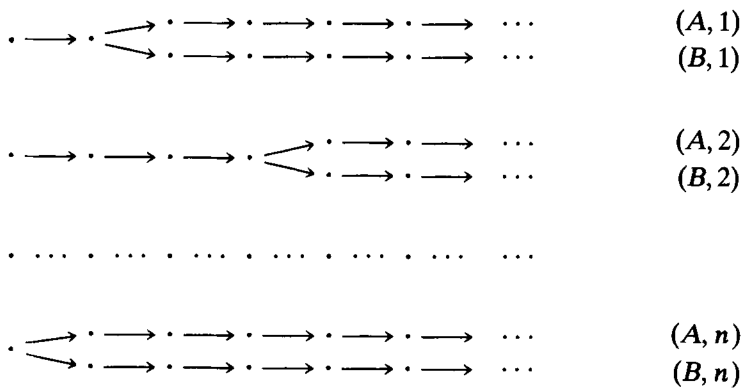
If $k \geqq l$ one sees that l and j in y are determined by k and i. Denote y by (A, i, k). If $k \leqq l$, k and i are determined by l and j. Denote y by (B, j, l). We have to identify $(A, i, k) = (B, i, k)$ for $1 \leqq k \leqq r(i,1) = s(i,1)$. Furthermore, if $\sigma^p a^i = \sigma^q a^j$ ($\sigma^p b^i = \sigma^q b^j$) then $(A, i, p+k) = (A, j, q+k)$ ($(B, i, p+k) = (B, j, q+k)$) for $k \geqq 1$. The matrix M' can be rewritten as ((ii) of Lemma 12)

(5.8)
(A, i, k) may be followed by $(A, i, k+1)$ and additional if $k = r(i,1) + \cdots + r(i,m)$ by (setting $j = a^i_{r(i,1)+\cdots+r(i,m-1)}$) $(B, j, r(i,m)+1)$ and $(A, t, 1) = (B, t, 1)$ for $a^i_k < t < b^j_{r(i,m)}$;
(B, j, l) may be followed by $(B, j, l+1)$ and additional if $l = s(j,1) + \cdots + s(j,m)$ by (setting $i = b^j_{s(j,1)+\cdots+s(j,m-1)}$) $A, i, s(j,m)+1)$ and $(A, t, 1) = (B, t, 1)$ for $a^i_{s(j,m)} < t < b^j_l$.

Remark that we need Lemma 11 to show writing i' for $b^j_{r(i,m)}$ that

$$(i', [\sigma a^{i'}, \sigma^{r(i,m)+1} b^j]) = (B, j, r(i,m)+1).$$

The diagram representing M' (cf. §2) looks like this:



Only the arrows from (A, i, k) to $(A, i, k+1)$ and from (B, j, l) to $(B, j, l+1)$ are indicated. If $\sigma^p a^i = \sigma^q b^j$ ($\sigma^p b^i = \sigma^q b^j$) for some i, j we can do the identifications mentioned above decreasing the number of rows of length ∞ in the diagram (this is important for applications). $(A, i) = \{(A, i, k) : k \geqq 1\}$, $(B, j) = \{(B, j, l) : l \geqq 1\}$.

Let us now return to maximal measures. Trivial examples of (I, f) with more than one maximal measure are given by the graphs in Fig. 1. The first one is the disjoint union of two copies of the same dynamical system. The second one contains also a wandering set. We shall see that every (I, f) with more than one maximal measure is essentially of this form.

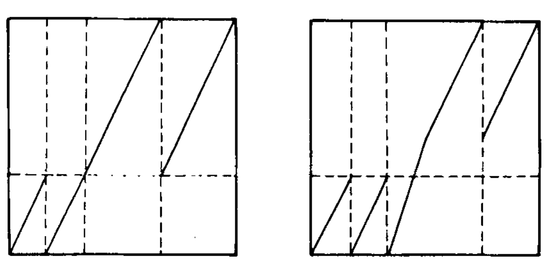
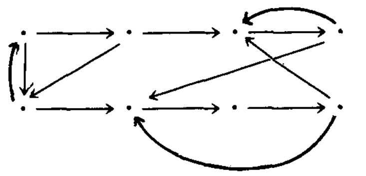


Fig. 1.

A more complicated example is the following one. For $n = 2$ let $\mathbf{a}^1 = \mathbf{a} = 1112121212\cdots$, $\mathbf{b}^1 = 1\mathbf{b}$, $\mathbf{a}^2 = 2\mathbf{a}$ and $\mathbf{b}^2 = \mathbf{b} = 2211211211211\cdots$. The diagram is (identifications for $\sigma\mathbf{a}^2 = \mathbf{a}^1$, $\sigma\mathbf{b}^1 = \mathbf{b}^2$, $\sigma^2\mathbf{a}^1 = \sigma^4\mathbf{a}^1$ and $\sigma\mathbf{b}^2 = \sigma^4\mathbf{b}^2$)



(A, 1)

(B, 2)

This has two irreducible submatrices:

$$M^1 = M'/D^1, \qquad \text{where } D^1 = \{(A,1,1),(A,1,2),(B,2,1)\},$$

$$M^2 = M'/D^2, \qquad \text{where } D^2 = D\backslash D^1.$$

$r(M^1) = r(M^2) = r(M')$, hence we have exactly two ergodic maximal measures. It is not difficult to construct a corresponding (I, f). We have in Σ_f^+

$$\mathbf{a} < \sigma^2\mathbf{b} < \sigma\mathbf{a} < \sigma^3\mathbf{b} < \sigma^2\mathbf{a} = \sigma^4\mathbf{a} < 1\mathbf{b}, \qquad 2\mathbf{a} < \sigma\mathbf{b} = \sigma^4\mathbf{b} < \sigma^3\mathbf{a} < \mathbf{b}.$$

Set $u_i = \varphi^{-1}(\sigma^i(1\mathbf{a}))$ and $v_i = \varphi^{-1}(\sigma^i(1\mathbf{b}))$ for $0 \leqq i \leqq 4$. Then the u_i and v_i satisfy the same order relations, because φ preserves the ordering (cf. Lemma 1), and $f(u_i) = u_{i+1}$ $(u_4 := u_2)$, $f(v_i) = v_{i+1}$ $(v_4 := v_1)$. We join the points (u_i, u_{i+1}), $(v_i, v_{i+1}) \in I \times I$ with straight lines to get the graph of an f (Fig. 2).

One can choose u_i, v_i so that the slope of f is everywhere greater than 1. It is easy to see that $I^2 = [u_1, v_3] \cup [u_2, v_4] \cup [u_3, v_2] \cup [u_4, v_1]$ is an invariant set for f. It is the support of one of the two ergodic maximal measures. Let I^1 be the closure of $I\backslash I^2$. Then the support of the other ergodic maximal measure is a Cantor-like subset of I^1, which remains after having taken away the wandering points of I^1.

Exchanging the eight subintervals of I in Fig. 2 we get, putting the intervals of I^2 before those of I^1, the graph in Fig. 3.

Exchanging these intervals is an isomorphism modulo small sets (cf. §0). We get a disjoint union of two dynamical systems (I^2 and a Cantor-like subset of I^1) together with a wandering set (rest of I^1).

Now we return to the investigation of $\Sigma_{M'}$ determining the irreducible submatrices of M' and the corresponding invariant subsets of (I, f). We work with the diagram above. This is easier than to work with the matrix M'.

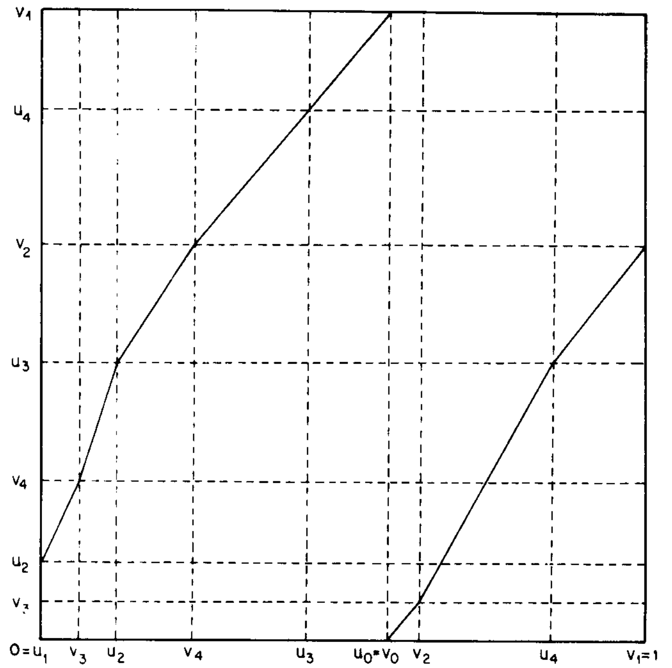


Fig. 2.

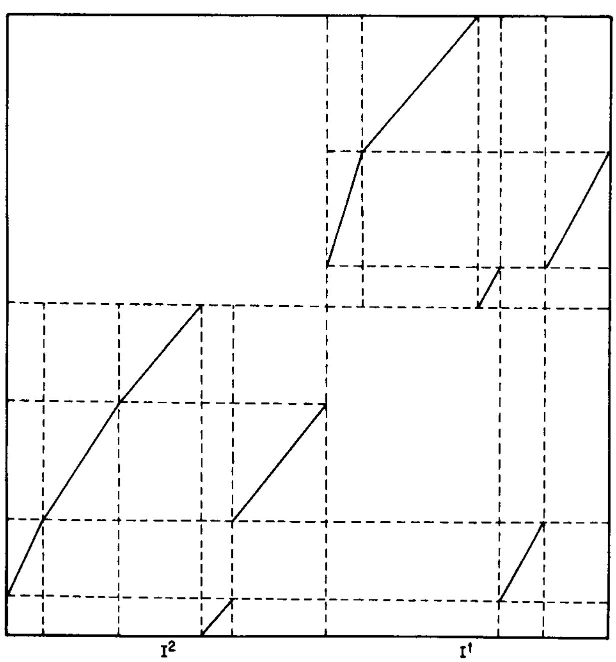


Fig. 3.

If M' is reducible, we can divide D into two disjoint subsets D^1 and D^2 such that M'/D^1 is irreducible and such that there may be transition from D^1 to D^2, but not from D^2 to D^1. Hence, if $(A,i,k)\in D^2$ then $(A,i,m)\in D^2$ for all $m \geqq k$. Similarly for (B,j,k). Therefore

$$D^2 = \bigcup_i \{(A,i,k): k \geqq k_i\} \cup \bigcup_j \{(B,j,k): k \geqq m_j\}$$

for some k_i, m_j, $1 \leqq i,j \leqq n$ with $1 \leqq k_j, m_j \leqq \infty$ ($k_i = \infty$ (or $m_j = \infty$) means $\{(A,i,k): k \geqq k_i\} = \emptyset$). Set $M^1 = M'/D^1$ and $M^2 = M'/D^2$. Remember the maps $\varphi:(I,f)\to\Sigma_f^+$ (§1) and $\psi:\Sigma_f\to\Sigma_{M'}$ (§2). Let $\pi:\Sigma_f\to\Sigma_f^+$ be the projection to the positive coordinates. Consider

$$\Sigma_{M'} \xrightarrow{\psi^{-1}} \Sigma_f \xrightarrow{\pi} \Sigma_f^+ \xrightarrow{\varphi^{-1}} (I,f)$$

and denote it by χ. We want to determine $\chi(\Sigma_{M^2})$. To apply ψ^{-1} means to represent a way in the diagram, which corresponds to a point in $\Sigma_{M'}$ by the edges and not by the vertices (cf. §2). The edges in (A,i) are numbered by a_1^i, a_2^i, $a_3^i,\cdots$ and in (B,j) by b_1^j, b_2^j, $b_3^j,\cdots$. Furthermore we have

$$k_i = r(i,1)+\cdots+r(i,K_i) \quad \text{and} \quad m_j = s(j,1)+\cdots+s(j,M_j)$$

for some K_i and M_j. $r(i,k)\geqq m_j$ for $k > K_i$, where $j = a^i_{r(i,1)+\cdots+r(i,k-1)}$ and $s(j,m)\geqq k_i$ for $m > M_j$, where $i = b^j_{s(j,1)+\cdots+s(j,m-1)}$, because there is no transition from D^2 to D^1. Set

$$B^j = \bigcup [\sigma^{r(i,1)+\cdots+r(i,l)}\boldsymbol{a}^i, \boldsymbol{b}^j],$$

where $j = a^i_{r(i,1)+\cdots+r(i,l)}$, the union taken over all i, l such that $l \geqq K_i$;

$$A^i = \bigcup [\boldsymbol{a}^i, \sigma^{s(j,1)+\cdots+s(j,l)}\boldsymbol{b}^j],$$

where $i = b^j_{s(j,1)+\cdots+s(j,l)}$, the union taken over all j, l such that $l \geqq M_j$. Then

$$\sigma^{m_j}B^j \subset B^p \cup A^q \cup \bigcup_{p<t<q} [\boldsymbol{a}^t, \boldsymbol{b}^t] = B^p \cup A^q \cup \bigcup_{p<t<q} A^t = B^p \cup A^q \cup \bigcup_{p<t<q} B^t$$

for some p,q with $1 \leqq p \leqq q \leqq n$. Set

$$\Sigma^2 = \bigcup_{m_j<\infty} \bigcup_{l=0}^{m_j-1} \sigma^l B^j \cup \bigcup_{k_i<\infty} \bigcup_{l=0}^{k_i-1} \sigma^l A^i.$$

Σ^2 is σ-invariant and furthermore $\pi\circ\psi^{-1}(\Sigma_{M^2}) = \Sigma^2$, because $\pi\circ\psi^{-1}(\Sigma_{M^2})\subset\Sigma_f^+$ consist of points $\boldsymbol{x}$ one gets by starting at any point in D^2 and going any way (which must be in D^2) and writing down the numbers of the edges on this way. $F^j := \varphi^{-1}(B^j)$ and $E^i := \varphi^{-1}(A^i)$ are subintervals of I. The same is true for $f^l(F^j)$, $l \leqq m_j - 1$ and $f^l(E^i)$, $l \leqq k_i - 1$. Set $\tilde{I}^2 = \chi(\Sigma_{M^2})$. Then

$$\tilde{I}^2 = \bigcup_{m_j < \infty} \bigcup_{l=0}^{m_j - 1} f^l(F^j) \cup \bigcup_{k_i < \infty} \bigcup_{l=0}^{k_i - 1} f^l(E^i).$$

This is a finite union of intervals.

Let I^1 be the closure of $I \backslash \tilde{I}^2$. This is again a finite union of intervals. But there may be points $x \in I^1$ with $f^l(x) \in \tilde{I}^2$ for some l. Set $\Omega_1 = \chi(\Sigma_{M^1})$. Then

$$\Omega_1 = I^1 \backslash \{x \in I^1 : f^l(x) \in I^2 \text{ for some } l\} = \bigcap_{l=0}^{\infty} f^{-l}(I^1).$$

Ω_1 is invariant under f and $I^1 \backslash \Omega_1$ is the set of wandering points contained in I^1.

It is not difficult to see that χ is continuous and surjective. M^1 is irreducible, hence there is a $y \in \Sigma_{M^1}$ with $\{\sigma^l y : l \geqq 0\}$ is dense in Σ_{M^1}. By continuity of χ, $\{f^l(\chi(y)) : l \geqq 0\}$ is dense in Ω_1. Therefore Ω_1 is topologically transitive.

Set $U_1 = I^1 \backslash \Omega_1$ and call it the unstable set for Ω_1, because the points in U_1 wander from Ω_1 to the invariant set $\tilde{I}^2$. If M^2 is irreducible then $\tilde{I}^2$ is topologically transitive. Call it Ω_2 and set $S_2 = U_1$, the stable set of Ω_2 (cf. the above example).

If M^2 is reducible, one can divide D^2 again into an irreducible subset D^3 and into D^4 such that there may be transition from D^3 to D^4, but not from D^4 to D^3. $\tilde{I}^3 = \chi(\Sigma_{M^4})$ is again a finite union of intervals, $f(\tilde{I}^3) \subset \tilde{I}^3$. Let I^2 be the closure of $\tilde{I}^2 \backslash \tilde{I}^3$ and $\Omega_2 = \bigcap f^{-l}(I^2)$. $U_2 = I^2 \backslash \Omega_2$ and $S_2 = \{x \in U_1 : f^l(x) \in \Omega_2$ for some $l \geqq 1\}$. Now repeat this procedure again for M^4 and so on. We get Ω_1, Ω_2, $\Omega_3, \cdots \subset I$, which are f-invariant and topologically transitive. $\bigcup S_i = \bigcup U_i = I \backslash \bigcup \Omega_i$ is the set of all wandering points. Remark that in general the Ω_i's are not disjoint. For $i \neq j$, $\Omega_i \cap \Omega_j$ is finite or empty.

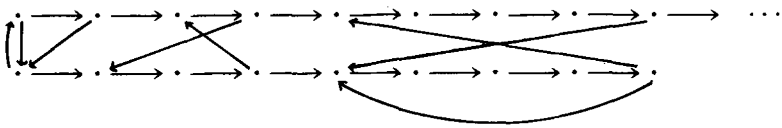
Let us consider another example, again for $n = 2$. Let $\boldsymbol{a}^1 = \boldsymbol{a}$, $\boldsymbol{b}^1 = 1\boldsymbol{b}$, $\boldsymbol{a}^2 = 2\boldsymbol{a}$, $\boldsymbol{b}^2 = \boldsymbol{b}$, where

$$\boldsymbol{a} = 1112122111221121111212211 \cdots$$

and

$$\boldsymbol{b} = 2211211122111221112 \cdots.$$

The diagram is



We have three irreducible submatrices:

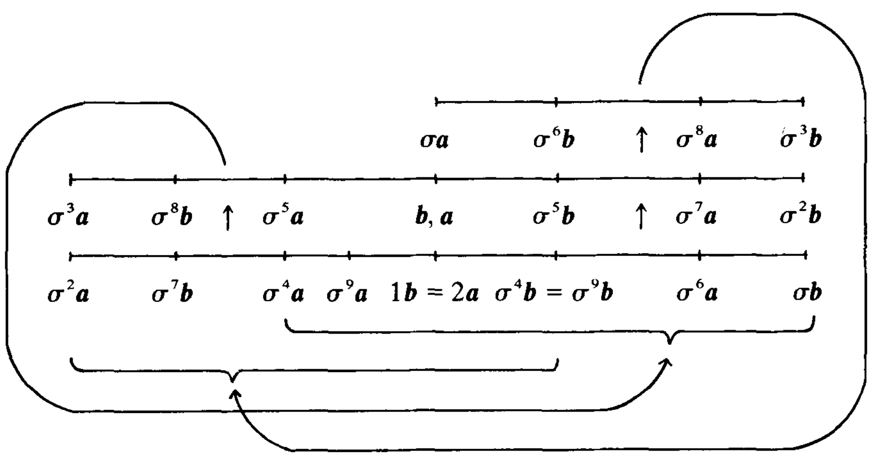
$$M^1 = M'/D^1,\ D^1 = \{(A,1,1),(A,1,2),(B,2,1)\},$$

$$M^2 = M'/D^2,\ D^2 = \{(A,1,k),(B,2,l): 3 \leqq k \leqq 4, 2 \leqq l \leqq 4\},$$

$$M^3 = M'/D^3,\ D^3 = \{(A,1,k),(B,2,l): k \geqq 5, 5 \leqq l \leqq 9\}.$$

One can get (I,f) as above.

$\chi(\Sigma_{M^2 \cup M^3})$ is as I^2 in the first example. We can draw the following picture of this set:



Ω_1 is a Cantor-like set as that described in the first example. Also $U_1 = I^1 \backslash \Omega_1$, where $I^1 = I \backslash \chi(\Sigma_{M^2 \cup M^3})$; $\Omega_2 = \varphi^{-1}\{\sigma^4 b, \sigma^5 b, \sigma^6 b, \sigma^7 b, \sigma^8 b\}$, a periodic orbit;

$$U_2 = \varphi^{-1}(]\sigma^4 b, \sigma^6 a[\cup]\sigma^5 b, \sigma^7 a[\cup]\sigma^6 b, \sigma^8 a[\cup]\sigma^7 b, \sigma^4 a[\cup]\sigma^8 b, \sigma^5 a[),$$

which wanders to Ω_3; $\Omega_3 = \chi(\Sigma_{M^2 \cup M^3}) \backslash U_2 =$ above picture $\backslash U_2$. We have $\Omega_1 \cap \Omega_2 = \emptyset$ and $\Omega_2 \cap \Omega_3 = \Omega_2$.

Let us return to maximal measures. By the above results we know that, if (I,f) is topologically transitive, M' must be irreducible. Hence

THEOREM 8. *If (I,f) is topologically transitive, then it has unique maximal measure m, which is positive on every open subset of I.*

The second part follows, because $m \circ \chi$ is positive on every cylinder set in $\Sigma_{M'}$ and χ is continuous and surjective. Now we prove that (I, f) has always only finitely many ergodic maximal measures. We need a lemma.

LEMMA 13. *Let* $D_m = \{(A, i, k), (B, j, k): k \geqq m\}$ *and* $M_m = M'/D_m$. *Then* $r(M_m) \to 1$, *as* $m \to \infty$.

PROOF. $r(M_m) = \lim \sqrt[k]{\|M_m^k\|_1}$. Denote the number of admissible blocks $xy_1 \cdots y_k$ of length $k+1$ beginning with $x \in D_m$ such that every $y_l \in D_m$ by N_k^x. We have

$$\|M_m^k\|_1 = \sup_{x \in D_m} N_k^x. \tag{5.9}$$

Define T_k by $T_1 = \|M_m\|_1$, $T_k = 0$ for $k \leqq 0$ and $T_{k+1} = T_k + T_{k-m}$. We show by induction $\|M_m^k\|_1 \leqq T_k$.

Suppose $\|M_m^l\|_1 \leqq T_l$ for every $l \leqq k$. Let $x = (A, i, l)$. If x may be followed only by $(A, i, l+1)$ then

$$N_{k+1}^x \leqq \sup_{y \in D_m} N_k^y \leqq T_k \leqq T_{k+1}.$$

If x may also be followed by (B, j, t) say, then an arrow beginning in (A, i) after $x = (A, i, l)$ which lands at $(B, k, r) \in D_m$ begins at $(A, i, l+m)$ or later (cf. (5.8)). It may be that we can add any block of length $k-1$ of elements in D_m after (B, j, t), but after (A, i, l) there must follow $(A, i, l+1), \cdots, (A, i, l+m)$ and then it may be that we can add any block as above. Hence

$$N_{k+1}^x \leqq \sup_{y \in D_m} N_k^y + \sup_{y \in D_m} N_{k-m}^y \leqq T_k + T_{k-m} = T_{k+1}.$$

Therefore $N_{k+1}^x \leqq T_{k+1}$ for all $x \in D_m$. We get the desired inequality using (5.9).

Hence we have $r(M_m) \leqq \lim \sqrt[k]{T_k} \leqq \max\{|\alpha_1|, \cdots, |\alpha_{m+1}|\}$, where $\alpha_1, \cdots, \alpha_{m+1}$ are the roots of $x^{m+1} - x^m - 1 = 0$. From this we get $|\alpha_i| \leqq 1 + |\alpha_i|^{-m}$, i.e.

$$\max\{|\alpha_1|, \cdots, |\alpha_{m+1}|\} \to 1, \qquad \text{if } m \to \infty.$$

By Lemma 13 we can choose an m, such that $r(M_m) < r(M')$ and all but finitely many D^i (index set of the irreducible submatrix M^i) are subsets of D_m. For such a $D^i \subset D_m$ the corresponding subspace Σ_{M^i} cannot be the support of an ergodic maximal measure. Together with Theorem 6 we get

THEOREM 9. *There are only finitely many ergodic maximal measures on* (I, f). *Their supports are* Ω_i*'s constructed above, which are finite unions of intervals or Cantor-like sets. Two of them have at most finitely many points in common.*

Exchanging intervals as in the first example above, we get an isomorphic transformation $\tilde{f}$, which is again piecewise monotonic, such that the supports of the ergodic maximal measures are contained in intervals $I^1, I^2, I^3, \cdots$, which satisfy $\tilde{f}(I^j) \subset \bigcup_{k \geqq j} I^k$.

REFERENCES

1. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass., 1968.
2. F. Hofbauer, *Schifträume und Gleichgewichtszustände*, Dissertation, Wien, 1976.
3. F. Hofbauer, *β-shifts have unique maximal measure*, to appear in Monatsh. Math.
4. W. Kreiger, *On the uniqueness of the equilibrium state*, Math. Systems Theory **8** (1974), 97–104.
5. W. Parry, *On β-extensions of real numbers.* Acta Math. **11** (1960), 401–416.
6. Y. Takahashi, *Isomorphism of β-automorphisms to Markov automorhpisms*, Osaka J. Math. **10** (1973), 175–184.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT WIEN
STRUDLHOFGASSE 4
A–1090 WIEN, AUSTRIA