SOME EXAMPLES OF SEQUENCE ENTROPY AS AN ISOMORPHISM INVARIANT

BY

F. M. DEKKING

ABSTRACT. With certain geometrically diverging sequences $A$ and the shift $T$ on dynamical systems arising from substitutions we associate a Markov shift $S$ such that the $A$-entropy of $T$ equals the usual entropy of $S$. We present examples to demonstrate the following results. Sequence entropy can distinguish between an invertible ergodic transformation and its inverse. $A$-entropy does not depend monotonically on $A$. The variational principle for topological sequence entropy need not hold.

1. Introduction. Let $(X, \mu)$ be a Lebesgue space with invertible $\mu$-preserving map $T$. Let $A = \{a_n\}_{n=1}^{\infty}$ be a sequence of integers. Then the sequence entropy $h_A(T)$ of $T$ with respect to $A$ is defined as follows [14]:

$$h_A(T, \xi) = \lim_{n \to \infty} \sup \frac{1}{n} H(T^{-n} \xi \lor \cdots \lor T^{-1} \xi), \quad \xi \in Z,$$

$$h_A(T) = \sup_{\xi \in Z} h_A(T, \xi)$$

where $Z$ is the collection of measurable partitions of $X$ with finite entropy. We also call $h_A(T)$ the $A$-entropy of $T$.

Clearly sequence entropy is an isomorphism invariant of dynamical systems and one recovers the usual entropy $h(T)$ of $T$ by taking $A = \{0, 1, 2, \ldots \}$.

Kušnirenko [14] proved that $T$ has discrete spectrum if and only if $h_A(T) = 0$ for every sequence $A$. On the other hand we have the following theorem of Newton and Krug [13], extending earlier work in [16]. If $h(T) > 0$, then $h_A(T) = K(A)h(T)$ where $K(A)$ does not depend on $T$. (A simplification of their proof can be found in [12].) This result implies that sequence entropy is uninteresting as a new invariant in case $T$ has positive entropy. However, little is known in case $h(T) = 0$ and $T$ has a continuous part in its spectrum, which in some sense is the case where new invariants are most needed. We briefly survey the known results for such $T$.

Kušnirenko showed that the $\{2^n\}$-entropy of the horocycle flow on a two-dimensional manifold of constant negative curvature lies between log 2 and log 64, and that the $\{2^n\}$-entropy of the transformation $(x, y) \to (x + \alpha, x + y + \beta)$ equals log 2. Recently Hulse [8] extended this result and determined a class of sequences such that $h_A(T) = \infty$ if $T$ has quasi-discrete spectrum. The existence of such sequences in case $T$ has quasi-discrete spectrum follows from a theorem of Pickel' [18] stating that $\sup_A h_A(T) = \log k$ for some positive integer $k$, or equals $\infty$. 

Received by the editors February 8, 1979.


Key words and phrases. Sequence entropy, topological sequence entropy, substitution, Markov shift.

© 1980 American Mathematical Society

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
In the present paper we calculate the sequence entropies for a special class of zero-entropy transformations and a number of sequences. These results reveal some new properties of sequence entropy. Our main example shows that sequence entropy can distinguish between \( T \) and \( T^{-1} \) (Example 8).

We do not know the answer to the following "intermediate value" question. If \( h(T) = 0 \) and \( h_A(T) = \infty \) for some \( A \), does there exist \( A' \) such that \( 0 < h_{A'}(T) < \infty \)? If this holds then the conjecture that \( T \times T \) is not isomorphic to \( T \) for ergodic \( T \) with finite entropy could easily (cf. the proof of Theorem 1 of [14]) be shown to be true. A related question is the following. Is \( h_A(T) > h_B(T) \), if \( A \) is a subsequence of \( B \)? It follows from Newton's work that the answer is positive when \( h(T) > 0 \), \( B \) is strictly increasing and \( A = \{t_n\} \) has bounded gaps or satisfies \( t_{n+1} - t_n \to \infty \). Example 7 however shows that the answer can be negative for such \( A \) and \( B \) if \( h(T) = 0 \).

We also consider the topological analogue of sequence entropy. Let \( X \) be a compact metric space, \( T \) a continuous map on \( X \). For \( \alpha \) an open cover of \( X \), let \( N(\alpha) \) denote the minimal cardinality of any subcover of \( \alpha \). Let \( A = \{t_n\}_{n=1}^\infty \) be a sequence of integers. Then \( h^\text{top}_A(T) \), the topological sequence entropy of \( T \) with respect to \( A \), is defined by

\[
h^\text{top}_A(T, \alpha) = \limsup_{n \to \infty} \frac{1}{n} \log N(T^{-1} \alpha \lor \cdots \lor T^{-1} \alpha),
\]

\[
h^\text{top}_A(T) = \sup_{\alpha \in W} h^\text{top}_A(T, \alpha)
\]

where \( W \) is the collection of open covers of \( X \) [6].

Let \( M \) denote the collection of \( T \)-invariant Borel probability measures, and \( h^\alpha_A(T) \) the \( A \)-entropy of \( T \) w.r.t. \( \mu \) in \( M \). Then Goodman [6] showed (with a restriction that can be removed [4]) that for any \( A \)

\[
h^\text{top}_A(T) > \sup_{\mu \in M} h^\alpha_A(T)
\]

with equality in case \( h^\text{top}(T) > 0 \), where \( h^\text{top}(T) \) denotes the usual topological entropy of \( T \). If \( h^\text{top}(T) = 0 \), then the variational principle for topological sequence entropy need not hold. Goodman gives an example with \( h^\text{top}_A(T) = \log 2 \) but \( \sup_{\mu \in M} h^\alpha_A(T) = 0 \), where \( T \) has discrete spectrum. We provide several other examples. If, for instance, \( T \) is the shift on the closed orbit of the Morse sequence (see for example [1, p. 92]) then

\[
h^\text{top}_{\{2\}}(T) = \log \left( \frac{1}{2} + \frac{1}{2} \sqrt{5} \right) > \sup_{\mu \in M} h^\alpha_{\{2\}}(T) = \frac{3}{2} \log 2.
\]

2. Preliminaries.

2.1. Substitution shifts. Let \( X = \{0, 1\}^\mathbb{Z} \) with shift \( T \). Let \( \bar{0} = 1, \bar{1} = 0 \). We extend this map—called mirroring—to blocks, i.e. finite sequences of zeroes and ones, and to elements of \( X \) by component wise application.

A map \( \theta : \{0, 1\} \to \{0, 1\}^l \) is called a substitution (on two symbols) of length \( l \). In this paper we only consider continuous substitutions [1]. A substitution is continuous iff \( \theta \bar{1} \) is the mirror image of \( \theta \bar{0} \) and, moreover, \( \theta \) is not finite i.e. \( \theta \bar{0} \neq \bar{00} \ldots 0, \theta \bar{0} \neq 0101 \ldots 010 \) or mirror images of these blocks. As we did in the case of
mirroring we extend the domain of substitutions to blocks and infinite sequences, and define $\theta^n$ by $\theta^n i = \theta^{n-1} \theta^i$, $i = 0, 1$, for each positive integer $n$.

If $\theta$ is a continuous substitution of length $l$ then $\theta^2$ starts and ends in 0 and we can define an infinite sequence $w_{\theta}$ by

$$w_{\theta}[-l^{2m}, l^{2m}-1] = \theta^{2m}(00), \quad m = 1, 2, \ldots \quad (1)$$

Here we used the following notation. If $x$ is an element of $X$ then $x[k, n] := x_k x_{k+1} \cdots x_n$.

Let $X_{\theta}$ be the closure (in the product topology on $X$) of the orbit $\{T^k w_{\theta} : k \in \mathbb{Z}\}$ of $w_{\theta}$. It is well known that $X_{\theta}$ is minimal under $T$ and that there is a unique $T$-invariant Borel probability measure $\mu$ on $X_{\theta}$ (see [1]). We call the dynamical system $(X_{\theta}, T, \mu)$—often abbreviated to $T_{\theta}$—a substitution shift. Note that $w_{\theta^m} = w_{\theta}$ and hence $T_{\theta^m} = T_{\theta}$ for $n = 2, 3, \ldots$.

The strict ergodicity of $T_{\theta}$ enables one to read many properties of $T_{\theta}$ from the sequence $w_{\theta}$ or from the blocks $\theta^m 0$ or $\theta^m 1$, $m = 1, 2, \ldots$. For example, consider simple cylinders $[b_0 b_1 \ldots b_n] := \{x \in X : x_0 = b_0, x_1 = b_1, \ldots, x_n = b_n\}$. Then such a cylinder has a nonempty intersection with $X_{\theta}$ iff the block $B = b_0 b_1 \ldots b_n$ occurs in $w_{\theta}$ iff it occurs in some $\theta^m 0$ iff it occurs in some $\theta^m 1$. We call such a block admissible. Furthermore, the $\mu$-measure of $[B] = [b_0 b_1 \ldots b_n]$ equals the relative frequency of $B$ in $w_{\theta}$ and

$$\mu([B]) = \lim_{m \to \infty} \frac{1}{m} N_{\theta}(\theta^m 0) = \lim_{m \to \infty} \frac{1}{m} N_{\theta}(\theta^m 1) \quad (2)$$

where $N_{\theta}(C)$ denotes the number of occurrences of $B$ in a block $C$. Since $\theta^m 1$ is the mirror image of $\theta^m 0$ the following lemma is obvious.

**Lemma 1.** Let $\theta$ be a continuous substitution. A block $B$ is admissible iff its mirror image is and $\mu([B]) = \mu([\overline{B}])$.

The next lemma is essential to the structure of $T_{\theta}$ (cf. [1, L.53]). In the following form the lemma is proved in [2, L.27]).

**Lemma 2.** Let $\theta$ be a nonfinite substitution of length $l$. For each positive integer $m$ the collection of open sets $\{T^j \theta^m X_{\theta} : 0 < j < l^m-1\}$ is a partition of $X_{\theta}$.

**Lemma 3.** Let $\theta$ be a nonfinite substitution of length $l$ and let $N > 0$. If $x \in T^i \theta^m X_{\theta}$ for some $m > 0$ and $0 < j < l^m$ then there exists an integer $p$ such that $x[0, N] = w_{\theta}[p, p + N]$ and $p = j$ modulo $l^m$.

**Proof.** Since $X_{\theta^m} = X_{\theta}$ we may assume $m = 1$. Let $x \in T^i \theta X_{\theta}$. By minimality of $T$ on $X_{\theta}$ we have $x = \lim_{i \to \infty} T^k w_{\theta}$ for some sequence $(k_i)$. Therefore $x[0, N] = w_{\theta}[k_i, k_i + N]$ for large $i$. On the other hand $T^i w_{\theta} \in T^i \theta X_{\theta}$ for large $i$, since $T^i \theta X_{\theta}$ is open. By Lemma 2, $k_i = j$ modulo $l^m$ for such $i$. \qed

Throughout this paper $\simeq$ denotes topological isomorphism, $\approx$ measure-theoretic isomorphism. (See [1] for definitions.) Unique ergodicity of substitution shifts implies that if $T_{\theta} \simeq T_{\theta'}$ then $T_{\theta} \approx T_{\theta'}$.

For a block $B = b_0 b_1 \ldots b_n$ let $B^- = b_n \ldots b_1 b_0$ be its reversal. If $\theta$ is a continuous substitution we define $\theta^{-} \sim \theta \sim 0 = (\theta 0)^{-}$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Lemma 4. Let \( \theta \) be a continuous substitution. Then \( T_\theta^{-1} \cong T_\theta \).

Proof. Let \( \phi \) be the reversal function on \( X \) defined by \( \phi(x) = x^\sim \), where \( x^\sim = x_{-k-1} \) for \( k \in \mathbb{Z} \). Then \( \phi \) is obviously continuous, invertible and \( \phi T^{-1} = T \phi \). It is easily checked that \( (\theta T^{-1})^{2m}(00) = (\theta^{2m}00)^\sim \). From (1) we see that \( \phi(w_\theta) = w_\theta \). The minimality of \( T^{-1} \) on \( X_\theta \) and \( T \) on \( X_\theta \) then implies that \( \phi \) is an isomorphism from \( (X_\theta, T^{-1}) \) to \( (X_\theta, T) \).

2.2. Sequence entropy of substitution shifts. Let \( A = \{ t_i \} \) be an infinite sequence of integers. From now on \( \xi \) will have a fixed meaning: \( \xi = \{ [0], [1] \} \), the time-0 partition of \( X \). Note that \( \xi \) is an open cover of \( X \). Let \( \xi_k := \xi \cup T^{-1}\xi \cup \ldots \cup T^{-k+1}\xi \). According to [6, P.2.1] and [14, L.2.1] we then have

Lemma 5. Let \( A \) be an infinite sequence of integers, \( T \) a subshift on \( X \), \( \xi = \{ [0], [1] \} \). Then

\[
\begin{align*}
h_A^{\top}(T) &= \lim_{k \to \infty} h_A^{\top}(T, \xi_k) \\
h_A(T) &= \lim_{k \to \infty} h_A(T, \xi_k).
\end{align*}
\]

Given \( A = \{ t_i \}, \eta \) a partition of \( X \), \( \eta^n \) will always denote the partition \( \eta^n = T^{-t_0}\eta \cup T^{-t_1}\eta \cup \ldots \cup T^{-t_n}\eta \). The first conclusion of Lemma 5 then can be rewritten as

\[
h_A^{\top}(T) = \lim_{k \to \infty} \limsup_{n \to \infty} \frac{1}{n} \log N(\xi^\eta_k).
\]

Since \( \xi^\eta_k \) is a partition, \( N(\xi^\eta_k) \) equals the number of nonempty atoms of \( \xi^\eta_k \). An atom of \( \xi^\eta_k \) is a general cylinder i.e. has the form

\[
\{ x \in X : x[t_j, t_j+k-1] = B^j, j = 1, \ldots, n \}
\]

where \( B^1, \ldots, B^n \) are blocks of length \( k \). We denote this atom by \( [B^1; B^2; \ldots; B^n] \).

Let \( \theta \) be a substitution, \( w = w_\theta \) as defined in (1). An atom \([B^1; B^2; \ldots; B^n]\) of \( \xi^\eta_k \) intersected with \( X_\theta \) is nonempty iff there is an integer \( p \) such that \( w[p + t_j, p + t_j+k-1] = B^j \), \( j = 1, \ldots, n \). If this is the case we say that the \( \xi^\eta_k \)-block \( (B^1; \ldots; B^n) \) occurs in \( w \), or is admissible. Let \( N_{(B^1; \ldots; B^n)}(C) \) denote the number of occurrences of the \( \xi^\eta_k \)-block \( (B^1; \ldots; B^n) \) in a block \( C \). Since a general cylinder is a finite disjoint union of simple cylinders, equation (2) generalizes to

\[
\mu([B^1; \ldots; B^n]) = \lim_{m \to \infty} \frac{1}{m} N_{(B^1; \ldots; B^n)}(\theta^m i) \quad \text{for } i = 0, 1,
\]

and we obtain analogously

Lemma 6. Let \( \theta \) be a continuous substitution, \( A \) an infinite sequence of integers. A \( \xi^\eta_k \)-block is admissible iff its mirror image is, and the corresponding atoms have equal measure.

3. \( \{ l^n \} \)-entropy for substitution shifts. In the sequel, \( \theta \) is a continuous substitution of length \( l \), \( \theta_0 = a_0a_1 \ldots a_{l-1} \) with \( a_0 = 0 \) (this is no restriction since \( T_\theta = T_\tilde{\theta} \), where \( \tilde{\theta} \) is defined by \( \tilde{\theta}0 = \theta(0) \)) and \( w = w_\theta \) is the sequence generated by \( \theta \).
Recall that $\xi = \{[0], [1]\}$. We consider $A = \{1, l, l^2, \ldots \}$, thus $\xi^n = T^{-1}\xi \vee T^{-2}\xi \vee \ldots \vee T^{-n}\xi$. Atoms are called nonempty if their intersection with $X_n$ is nonempty.

If $[b_0; b_1; \ldots ; b_n]$ is a nonempty atom of $\xi^{n+1}$ then trivially $[b_0; b_1; \ldots ; b_{n-1}]$ is a nonempty atom of $\xi^n$.

**Basic Property.** Let $[b_0; b_1; \ldots ; b_n]$ be a nonempty atom of $\xi^{n+1}$. Then $[b_1; b_2; \ldots ; b_n]$ is a nonempty atom of $\xi^n$.

**Proof.** If $[b_0; b_1; \ldots ; b_n]$ is a nonempty atom of $\xi^{n+1}$ then there exist integers $p$ and $0 < q < l$ such that $w_{p+q+1} = b_j, j = 0, 1, \ldots , n$. Since $\theta w = w$, we have for any integer $s$

$$w[sl, sl + l - 1] = \begin{cases} 0, & a_0 a_1 \ldots a_{l-1} \\
1, & \bar{a}_0 \bar{a}_1 \ldots \bar{a}_{l-1} \end{cases} \quad (4)$$

and (recall that $a_0 = 0$)

$$w_{sl} = w_s. \quad (5)$$

Hence, if $a_q = 0$, then

$$w_{p+q+1} = w_{p+q+1} = b_j, \quad j = 1, 2, \ldots , n,$$

and if $a_q = 1$, then

$$w_{p+q+1} = w_{p+q+1} = \bar{b}_j, \quad j = 1, 2, \ldots , n.$$  

Therefore, if $a_q = 0$ then $[b_1; b_2; \ldots ; b_n]$; if $a_q = 1$ then $[\bar{b}_1; \bar{b}_2; \ldots ; \bar{b}_n]$ is a nonempty atom of $\xi^n$. The claimed property follows in the last case from Lemma 6.

**Theorem 1.** Let $\theta$ be a continuous substitution of constant length $l$ and let $A = \{1^n\}$. Then

$$h^{\text{top}}(T_{\theta}) = h^{\text{top}}(T_{\theta}, \xi).$$

**Proof.** Let $\xi_n = \xi \vee T^{-1}\xi \vee \ldots \vee T^{-n+1}\xi$. According to Lemma 4

$$h^{\text{top}}(T_{\theta}) = \lim_{k \to \infty} h^{\text{top}}(T_{\theta}, \xi_k)$$

for any $A$. Since $\xi_{k-1}$ refines $\xi_k$, it suffices to prove (with $A = \{1^n\}$)

$$h^{\text{top}}(T_{\theta}, \xi_k) < h^{\text{top}}(T_{\theta}, \xi) \quad \text{for } k = 2, 3, \ldots ,$$

We first consider $k = 2$. Let $(b_0 c_0; b_1 c_1; \ldots ; b_n c_n)$ be the $\xi^2$-block corresponding to a nonempty atom of $\xi_{2}$, i.e. there exist integers $p$ and $0 < q < l$ such that

$$w_{p+q+1} = w_{p+q+1} = b_j, \quad j = 0, 1, \ldots , n.$$

**Case 1.** $0 < q < l - 1$. Then $c_0, \ldots , c_n$ are completely determined by the symbol $c_1$ plus the $\xi^n$-block $(b_1; b_2; \ldots ; b_n)$. For the $\xi^n$-block $(c_1; c_2; \ldots ; c_n)$ equals $(b_1; b_2; \ldots ; b_n)$ or its mirror image, depending on whether $a_q = a_{q+1}$ or not.

**Case 2.** $q = l - 1$. Then $(b_0 c_0; b_1 c_1; \ldots ; b_n c_n)$ is completely determined by an admissible $\xi^2$-block. This is the $\xi^2$-block $(b_1 c_1; \ldots ; b_n c_n)$ if $a_{l-1} = 0$, the $\xi^2$-block $(\bar{b}_1 c_1, \ldots , \bar{b}_n c_n)$ if $a_{l-1} = 1$ (cf. the proof of the basic property). The symbols $b_0$ and $c_0$ are determined by $b_1$ since all three symbols occur in the same $\theta 0$ or $\theta 1$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Combining these two cases we obtain
\[ N(\xi_n^2) < 2 \cdot 2^N(\xi_n^n) + N(\xi_n^2) \]
\[ < 4N(\xi_n^n) + 4N(\xi_n^{n-1}) + \cdots + 4N(\xi_n^2) + N(\xi_n^2) \]
\[ < 4nN(\xi_n^n). \]
Therefore
\[ h^{\text{top}}_A(T_\theta, \xi_n) = \limsup_{n \to \infty} \frac{1}{n} \log N(\xi_n^n) \]
\[ < \limsup_{n \to \infty} \frac{1}{n} \log(4(n-1)N(\xi_n^{n-1})) \]
\[ = h^{\text{top}}_A(T_\theta, \xi). \]

Now let \( k > 2 \). Choose \( m \) such that \( l^m > k \). Let \( B^j = b_1^j b_2^j \ldots b_k^j \) be a block of length \( k \) for \( j = 1, \ldots, n \) and let \((B^1; B^2; \ldots; B^n)\) be the \( \xi_n^n \)-block corresponding to a nonempty atom of \( \xi_n^n \). Then there exist integers \( p \) and \( 0 < q < l^m \) such that
\[ w[p^m + q + l^j, p^{m-1} + q + l^j + k - 1] = B^j, \quad j = 1, \ldots, n. \]
The \( \xi_n^n \)-block is completely determined by the first \( m-1 \) blocks \( B^1, \ldots, B^{m-1} \) and the \( \xi_n^{n-m+1} \)-block
\[ \begin{cases} (b_{m-q}^m b_{m-1}^m \ldots b_{m-q+1}^m) & \text{if } q > l^m - k, \\ (b_1^m b_2^m \ldots b_q^m b_{q+1}^m) & \text{if } 0 < q < l^m - k. \end{cases} \]
(Actually the \( \xi_n^{n-m+1} \)-block \((b_1^m; \ldots; b_q^m)\) would suffice in the second case.) This leads to the following bound.
\[ N(\xi_n^n) < (2^k)^{m-1} N(\xi_n^{n-m+1}). \]
Therefore
\[ h^{\text{top}}_A(T_\theta, \xi_n) < h^{\text{top}}_A(T_\theta, \xi). \]

**Theorem 2.** Let \( \theta \) be a continuous substitution of length \( l \) and let \( A = \{l^n \} \). Then
\[ h_A(T_\theta, \xi_n) = h_A(T_\theta, \xi). \]

**Proof.** As in the proof of Theorem 1 it suffices to show (with \( A = \{l^n \} \)) that
\[ h_A(T_\theta, \xi_n) < h_A(T_\theta, \xi) \]
for \( k = 2, 3, \ldots \). Let \( k > 2 \) and \( m \) be arbitrary such that \( l^m > k \). Let
\[ F = \bigcup_{j=l^m-k+1}^{l^m-1} T^j_\theta X_\theta \]
and \( \eta = \{F, X_\theta \setminus F\} \). We shall estimate \( H(\xi_n^n \setminus \eta, \xi_n^n \setminus \eta) \) for \( n > m \).

An atom of \( \xi_n^n \setminus \eta \) which lies in \( X_\theta \setminus F \) splits into at most \( 2^{(k-1)m} \) atoms of \( \xi_n^n \).
For if \([B^1; B^2; \ldots; B^n] \) is a nonempty atom of \( \xi_n^n \) \((B^j = b_1^j b_2^j \ldots b_k^j \) for \( j = 1, \ldots, n \)), and \( x \) belongs to the nonempty atom \([b_1^j; b_2^j; \ldots; b_k^j] \) of \( \xi_n^n \) and to \( T^j_\theta X_\theta \) then it follows from Lemma 3 that the blocks \( B^m, \ldots, B^n \) all occur at places \( j \) modulo \( l^m \) in \( w \). Since \( 0 < j < l^m - k \), these blocks are completely
determined by the $\xi^n$-block $(b_1^1; b_1^2; \ldots; b_1^n)$ and the block $b_2^m b_3^m \ldots b_k^m$. Furthermore there are at most $(2^{k-1})^{m-1}$ possibilities for $b_2^1 b_3^1 \cdots b_k^1$, \ldots, $b_2^m b_3^m \cdots b_k^m$. 

An atom of $\xi^n \cup \eta$ which lies in $F$ splits into at most $2^{kn}$ atoms of $\xi_k^n$, as there can be no more atoms in $\xi_k^n$. Estimating the entropy of a partition with the logarithm of the number of its nonempty atoms we obtain

$$H(\xi_k^n | \xi^n \cup \eta) \leq \mu(X_\theta \setminus F) \log 2^{(k-1)m} + \mu(F) \log 2^{nk}$$

$$\leq (k-1)m \log 2 + nk \mu(F) \log 2.$$  

According to standard properties of entropy

$$H(\xi_k^n) < H(\xi_k^n | \xi^n \cup \eta) + H(\xi^n) + H(\eta).$$

So

$$H(\xi_k^n) < H(\xi^n) + (k-1)m \log 2 + nk \mu(F) \log 2 + \log 2,$$

and

$$h_A(T_\theta, \xi_k) < h_A(T_\theta, \xi) + k \mu(F) \log 2.$$  

Since $T$ preserves $\mu$ and the $T^i \theta^m X_\theta$ are disjoint according to Lemma 2, we have $\mu(T^i \theta^m X_\theta) = l_m^{-m}$ for $j = 0, \ldots, l_m^{-m} - 1$. Therefore

$$h_A(T_\theta, \xi_k) < h_A(T_\theta, \xi) + (k-1)kl_m^{-m} \log 2.$$  

Since we can choose $m$ arbitrarily large we obtain the desired

$$h_A(T_\theta, \xi_k) < h_A(T_\theta, \xi), \text{ where } A = \{l_m\}. \quad \square$$

We now turn to the task of calculating $h_{(l_m^r)}^{(l_m^r)}(T_\theta, \xi)$ and $h_{(l_m^r)}(T_\theta, \xi)$. If $[b_0; b_1; \ldots; b_n]$ is a nonempty atom of $\xi^{n+1}$ then $[b_1; b_2; \ldots; b_n]$ is a nonempty atom of $\xi^n$, as noted in the Basic Property. Conversely, if $[b_1; b_2; \ldots; b_n]$ is a nonempty atom of $\xi^n$, then either $[0; b_1; \ldots; b_n]$ or $[1; b_1; \ldots; b_n]$ or both are nonempty atoms of $\xi^{n+1}$. As we shall show below it is possible to decide which one of these possibilities occurs, according to an algorithm which does not depend on $n$ but only on the first element of the atom of $\xi^n$ under consideration, i.e. on $b_1$.

First we split some atoms into two: instead of $\xi^n = T^{-1} \xi \vee \ldots \vee T^{-i} \xi$ we consider $\xi \vee \xi^n$, denoting a nonempty atom of this partition by

$$[b_0(b_1; b_2; \ldots; b_n)] := \{x \in X_\theta: x_0 = b_0, x_1 = b_1, x_2 = b_2, \ldots, x_{n-1} = b_n\},$$

and the corresponding $\xi \vee \xi^n$-block by $b_0(b_1; b_2; \ldots; b_n)$. We then fuse pairs of atoms of $\xi \vee \xi^n$. Let $\eta_n$ be the partition whose atoms are

$$[b_0(b_1; b_2; \ldots; b_n)] \cup [\bar{b}_0(b_1; b_2; \ldots; b_n)]. \quad (6)$$

Note that

$$|\log N(\xi^n) - \log N(\eta_n)| < \log 2, \quad |H(\xi^n) - H(\eta_n)| < \log 2.$$  

Therefore

$$h_{(l_m^r)}^{(l_m^r)}(T_\theta, \xi) = \limsup_{n \to \infty} \frac{1}{n} \log N(\eta_n).$$
and

\[ h_{(T^p)}(T_\theta, \xi) = \limsup_{n \to \infty} \frac{1}{n} H(\eta_n). \]

The following lemma describes the announced splitting algorithm. In the sequel, + between symbols denotes addition mod 2.

**Lemma 7.** Let \( 0(b_1; b_2; \ldots; b_n) \) be an admissible \( \xi \vee \xi^n \)-block. Then

\[
0(a_q + a_{q+1}; b_1; b_2; \ldots; b_n), \quad q = 0, 1, \ldots, l - 2,
\]

\[
0(a_{l-1} + b_1; b_1; b_2; \ldots; b_n) \tag{7}
\]

are \( l \) admissible \( \xi \vee \xi^{n+1} \)-blocks. Moreover any admissible \( \xi \vee \xi^{n+1} \)-block \( 0(c; b_1; b_2; \ldots; b_n) \) with \( c = 0 \) or 1 is among (7).

**Proof.** Let \( 0(b_1; b_2; \ldots; b_n) \) be an admissible \( \xi \vee \xi^n \)-block, i.e. there exists an integer \( p \) such that \( w_p = 0, w_{p+q-1} = b_j, j = 1, 2, \ldots, n \). By definition

\[
w_{p+q}(w_{p+q+1}; \ldots; w_{p+q+n}), \quad q = 0, \ldots, l - 1, \tag{8}
\]

are admissible \( \xi \vee \xi^{n+1} \)-blocks. But it follows with (4) and (5) that

\[
w[p, lp + l - 1] = a_0a_1 \ldots a_{l-1} \quad \text{and}
\]

\[
w_{p+q+\nu} = \begin{cases} b_j \text{ if } a_q = 0, \\ b_j \text{ if } a_q = 1, \end{cases} \quad j = 1, \ldots, n \text{ and } q = 0, \ldots, l - 1.
\]

If + between symbols denotes addition modulo 2 we can concisely write \( w_{p+q+\nu} = b_j + a_q \) and we see that the blocks in (8) can be identified as

\[
a_q(a_{q+1}; b_1 + a_q; b_2 + a_q; \ldots; b_n + a_q), \quad q = 0, \ldots, l - 2,
\]

\[
a_{l-1}(b_1; b_1 + a_{l-1}; b_2 + a_{l-1}; \ldots; b_n + a_{l-1}). \tag{9}
\]

Considering the cases \( a_q = 0 \) and \( a_q = 1 \) separately, mirroring the \( \xi \vee \xi^{n+1} \)-block in the latter case and using Lemma 6, we obtain that the blocks in (7) are admissible. The last assertion of the lemma follows from the fact that any admissible \( \xi \vee \xi^{n+1} \)-block \( 0(c; b_1; b_2; \ldots; b_n) \) is among (8) for some \( p \) and \( 0 < q < l \), and the Basic Property. □

**Theorem 3.** Let \( \theta \) be a continuous substitution of length \( l \). Then

\[
h_{(T^p)}(T_\theta) = \begin{cases} \log \left( \frac{1}{2} + \frac{1}{2} \sqrt{5} \right) & \text{if } \theta 0 = 0101 \ldots 01, \\ \log 2 & \text{in all other cases.} \end{cases}
\]

**Proof.** Let \( \alpha \) be the number of zeros in \( \{a_0 + a_1, a_1 + a_2, \ldots, a_{l-2} + a_{l-1}, a_{l-1}\} \), \( \beta \) the number of zeroes in \( \{a_0 + a_1 + a_2, \ldots, a_{l-2} + a_{l-1}, a_{l-1} + 1\} \). Then \( 0 < \alpha, \beta < l \). In case \( 0 < \alpha, \beta < l \) it follows from Lemma 7 by induction that any \( \xi \vee \xi^n \)-block is admissible. (The \( \xi \vee \xi^1 \)-blocks \( 00 \) and \( 01 \) are admissible for any nonfinite \( \theta \).) Therefore \( N(\eta_n) = 2^\alpha \) and \( h_{(T^p)}(T_\theta) = \log 2 \). Here we used Theorem 1.

It is quickly verified that the cases \( \alpha = l, \beta = 0 \) lead to finite substitutions and that \( \beta = l \) is impossible. The only case left is \( \alpha = 0 \) with \( \theta 0 = 0101 \ldots 01 \). Then
all blocks in (7) are equal to 0(1; 0; \ldots ; b_n) if b_1 = 0, while both
0(0; 1; b_2; \ldots ; b_n) and 0(1; 1; b_2; \ldots ; b_n) occur if b_1 = 1. It follows by induction
that 0(b_0; b_1; \ldots ; b_{n-1}) is an admissible \xi \vee \xi^n\omega-block if and only if no two
consecutive b's are 0. This easily implies that \eta_n has exactly F_{n+1} nonempty atoms,
where F_n is the nth Fibonacci number. Therefore, if \theta_0 = 0101 \ldots 01, then
\[ h_{\Omega}^{\text{top}}(T_\theta) = \limsup_{n \to \infty} \frac{1}{n} \log(F_{n+1}) = \log\left(\frac{1}{2} + \frac{1}{2} \sqrt{5}\right). \]

**Corollary.** Let \theta be a continuous substitution of length l. Then
\[ h_{\Omega}^{\text{top}}(T_\theta) = \log 2. \]

**Proof.** Immediate from T_\theta = T_{\theta^0}.

**Remark.** The reader acquainted with topological Markov chains (introduced as
intrinsic Markov chains in [17]) will have noticed that in the case \theta_0 = 0101 \ldots 01
the nonempty atoms of \eta_n are in one-to-one correspondence with the nonempty
atoms of (\xi \vee T^{-1}\xi \vee \ldots \vee T^{-n+1}\xi) \cap X_\sigma, where \xi is the time-0 partition of
a shift space on two symbols with shift T and X_\sigma the topological Markov-chain
determined by the matrix \sigma = (\xi_i^j).

We now turn to measure-theoretic (l^n)-entropy. Given a continuous substitution
of length l, we define a substitution T_{\theta_d} of length l by
\[ \theta_d(0) = (a_0 + a_1)(a_1 + a_2) \cdots (a_{l-2} + a_{l-1})(a_{l-1} + a_0), \]
\[ \theta_d(1) = (a_0 + a_1)(a_1 + a_2) \cdots (a_{l-2} + a_{l-1})(a_{l-1} + a_0). \]
In [1] \theta_d is called the discrete substitution associated with \theta. If \theta is any substitution
then we call the matrix L = L(\theta) = (l_{ij}) defined by l_{ij} = N(l_{ij}), i, j = 0, 1, the
matrix of \theta. Note (remember a_0 = 0) that L(\theta_d) = (\alpha_{ij}^d) = (\alpha_{ij}^{l-\alpha_{ij}^d}),
where \alpha and \beta are as in the proof of Theorem 3.

On \{0, 1\} we define a probability vector \pi by
\[ \pi_0 = \mu([00] \cup [11]), \quad \pi_1 = \mu([01] \cup [10]) \quad (10) \]
and a Markov matrix \Pi = \Pi(\theta) = (\pi_{ij}) by
\[ \Pi := L^{-1}L(\theta_d) = L^{-1}\begin{pmatrix} \alpha & l - \alpha \\ \beta & l - \beta \end{pmatrix}. \quad (11) \]

**Lemma 8.** Let \pi and \Pi be as in (10) and (11). Then \Pi(\theta^m) = \Pi^m(\pi), \pi\Pi = \pi and
\[ p_0 = \beta(l - \alpha + \beta)^{-1}. \]

**Proof.** It is not hard to verify that (\theta^m)_d = (\theta_d)^m for m = 1, 2, \ldots .
Consequently \Pi(\theta^m) = L^{-m}L((\theta^m)_d) = \Pi^m(\theta_d) = \Pi^m(\theta). It follows from (2) that
\[ \Pi(\theta^m) \to \begin{pmatrix} p_0 & p_1 \\ p_0 & p_1 \end{pmatrix} \quad \text{if} \ m \to \infty; \]
which implies \pi\Pi = \pi, and \[ p_0 = \beta(l - \alpha + \beta)^{-1} \] follows by direct computation. \[ \square \]

Let \nu be the Markov measure and T_{\nu_\Pi} be the Markov shift defined by \pi and \Pi [3]
(\nu is uniquely determined by \Pi since \Pi is irreducible). Let \xi = \{[0], [1]\} and
\[ \xi_n = \xi \vee T^{-1}\xi \vee \ldots \vee T^{-n+1}\xi. \]

The distinction between \xi and \xi is made to
indicate that atoms of $\xi$ will be measured with $\mu$ and those of $\zeta$ and $\zeta_n$ by $\nu$. We define a map $\phi$ from $\eta_n$ (with atoms (6)) to $\zeta_n$ by

$$\phi: \left[0(b_1; b_2; \ldots ; b_n)\right] \cup \left[1(b_1; b_2; \ldots ; b_n)\right] \rightarrow [b_n b_{n-1} \ldots b_2 b_1].$$

In particular $\phi([00] \cup [11]) = [0]$ and $\phi([01] \cup [10]) = [1]$. We claim that $\phi$ is one-to-one and measure preserving between $\eta_n$ and $\zeta_n$, i.e. if $B$ is an atom of $\zeta_n$ then $\phi^{-1}B$ is an atom of $\eta_n$ such that $\mu(\phi^{-1}B) = \nu(B)$. The proof is by induction on $n$. For $n = 1$ our claim is true by (10). Let $n > 2$. Suppose

$$\mu\left([0(b_1; b_2; \ldots ; b_n) \cup [1(b_1; b_2; \ldots ; b_n)\right] = \nu([b_n \ldots b_2 b_1])$$

for all atoms $[0(b_1; b_2; \ldots ; b_n)] \cup [1(b_1; b_2; \ldots ; b_n)]$ of $\eta_n$. We have to show

$$\mu\left([0(b_0; b_1; \ldots ; b_n) \cup [1(b_0; b_1; \ldots ; b_n)\right] = \nu([b_n \ldots b_1 b_0])$$

for $b_0 = 0, 1$. Now

$$\nu([b_n \ldots b_1 b_0]) = \pi_{b_0} \nu([b_n \ldots b_2 b_1])$$

$$= \pi_{b_0} \mu([0(b_1; b_2; \ldots ; b_n)] \cup [1(b_1; b_2; \ldots ; b_n))],$$

the last equality by the induction hypothesis. So we prove our claim, if we show

$$\mu([0(b_0; b_1; \ldots ; b_n)] \cup [1(b_0; b_1; \ldots ; b_n)])$$

$$= \mu([0(b_1; b_2; \ldots ; b_n)] \cup [1(b_1; b_2; \ldots ; b_n)]).$$

Let $m$ be arbitrary but large. Any $\xi \vee \xi^{n+1}$-block occurring in $\theta^{m+1}(01)$ is generated by a $\xi \vee \xi^n$-block occurring in $\theta^m(01)$ as in the proof of the Basic Property. Conversely, any $\xi \vee \xi^n$-block occurring in $\theta^m(01)$ generates $l \xi \vee \xi^{n+1}$-blocks in $\theta^{m+1}(01)$ as in (9). Considering this equation and its mirror image and taking into account that $\xi \vee \xi^n$-blocks and their mirror images occur in pairs in $\theta^m(01) = \theta^m 0 0^m 0$ and in $\theta^{m+1}(01)$, we obtain the following refinement of Lemma 7. Suppose there occur $k \xi \vee \xi^n$-blocks $0(b_1; \ldots ; b_n)$ in $\theta^m(01)$. Then there occur exactly $k \xi \vee \xi^n$-blocks $0(a_q + a_{q+1}; b_1; \ldots ; b_n)$ for each $q = 0, 1, \ldots , l - 2$ separately, and $k$ blocks $0(a_{l-1} + b_1; b_1; \ldots ; b_n)$ in $\theta^{m+1}(01)$. Therefore

$$N_0(b_0; b_1; \ldots ; b_n)(\theta^{m+1}(01))$$

$$= (\text{number of } b_0 \text{ in } \{a_0 + a_1, \ldots , a_{j-2} + a_{j-1} + b_1\})N_0(b_1; \ldots ; b_0)(\theta^m(01))$$

$$= l_\eta \eta_0 N_0(b_1; \ldots ; b_0)(\theta^m(01)),$$

by the definition of $\Pi$ in (11). Dividing both sides by $l^{m+1}$ and letting $m \rightarrow \infty$ yields (12) by (3) and Lemma 6.

The fact that $\phi: \eta_n \rightarrow \zeta_n$ is one-to-one and measure preserving for $n = 1, 2, \ldots$ implies immediately

**Theorem 4.** Let $\theta$ be a continuous substitution of length $l$ with associated discrete substitution $\theta_\eta$. Then

$$h_{1^{\ell^*}}(T_\theta) = h(T_\Pi)$$

where $T_\Pi$ is a Markov shift with transition matrix $\Pi = l^{-1}L(\theta_\eta).$
sequence entropy as an isomorphism invariant

Example 1. High \{l^n\}-entropy is obtained when \( \alpha = l - \beta \) iff

\[
\Pi = \frac{1}{2} \left( \frac{l - 1}{l + 1} \right) + \frac{1}{2} \left( \frac{l + 1}{l - 1} \right).
\]

For any odd \( l \) a continuous \( \theta \) with \( L(\theta) = l \Pi \) as above is easily found. In this case \( p = \left( \frac{1}{2}, \frac{1}{2} \right) \) and

\[
h_{(l^n)}(T_{\theta}) = \log 2 + \frac{1}{2} \log \frac{l^2}{l^2 - 1} - \frac{1}{2l} \log \frac{l + 1}{l - 1},
\]

which can be arbitrarily close to \( \log 2 \).

Example 2. Low \{l^n\}-entropy is obtained when \( 00 = 0101 \ldots 01 \). Then

\[
h_{(l^n)}(T_{\theta}) = \frac{l}{l + 1} \log l - \frac{l - 1}{l + 1} \log(l - 1),
\]

which can be arbitrarily close to 0.

Example 3. Theorems 3 and 4 provide us with many examples of \( h^{\text{top}}_{\alpha}(T) \neq h_{\alpha}(T) = \sup_{\mu \in \mathcal{M}} h_{\mu}(T) \). For example if \( \theta \) is defined by \( \theta 0 = 01 \), then \( h^{\text{top}}_{(l^n)}(T_{\theta}) = \log(\frac{l}{2} + \frac{1}{2} \sqrt{5}) \) and \( h_{(l^n)}(T_{\theta}) = \frac{3}{5} \log 2 \) (set \( l = 2 \) in Example 2).

Example 4. \{l^n\}-entropy is not a complete invariant for \( T_{\theta'} \)'s with \( \theta \) of length \( l \). It is particularly obvious that \( \Pi(\theta) = \Pi(\theta'^{-}) \), so \( T_{\theta} \) cannot be distinguished from \( T_{\theta'^{-}} \) (Lemma 4).

4. Other sequences. In this section we consider sequences

\[ A = \{ k(1) \cdot l^n, k(2) \cdot l^n, \ldots, k(g) \cdot l^n \} \]

with \( 1 = k(1) < k(2) < \ldots < k(g) < l \). The restriction \( k(1) = 1 \) is mainly for technical reasons. Recall that \( \xi = \{ [0], [1] \} \). The basic property now is: Let

\[
[ b_1; b_2; \ldots; b_{g(n+1)} ] = \{ x \in X: x_{k(i)^{l^n}} = b_{g(i)+j}, 1 < i < g, 0 < j < n \}
\]

be a nonempty atom of \( \xi^{g(n+1)} \). Then \( [ b_{g+1}; b_{g+2}; \ldots; b_{g(n+1)} ] \) is a nonempty atom of \( \xi^{g^n} \). The proof is a slight generalization of the proof of the basic property for \( A = \{ l^n \} \).

Theorem 5. Let \( \theta \) be a continuous substitution of length \( l \) and let \( A \) be as in (\#). Then

\[
h^{\text{top}}_{\alpha}(T_{\theta}) = h^{\text{top}}_{\alpha}(T_{\theta}, \xi) \quad \text{and} \quad h_{\alpha}(T_{\theta}) = h_{\alpha}(T_{\theta}, \xi).
\]

Proof. Note that

\[
\lim_{n \to \infty} \frac{1}{n} \log N(\xi^n) = \lim_{n \to \infty} \frac{1}{gn} \log N(\xi^{gn})
\]

and

\[
\lim_{n \to \infty} \frac{1}{n} H(\xi^n) = \lim_{n \to \infty} \frac{1}{gn} H(\xi^{gn}).
\]

We can follow the proofs of Theorems 1 and 2 with \( n \) replaced by \( gn \) everywhere and \( A \) as in (\#). Details are left to the reader. □

In the calculation of \( h^{\text{top}}_{\alpha}(T_{\theta}, \xi) \) and \( h_{\alpha}(T_{\theta}, \xi) \) the following notions play an important role. Let \( a_0a_1 \ldots a_{l-1} \) be a block of zeroes and ones. The auto correlation
matrix $M = (m_{ij})_{i=1, q=0}^{l, q=0}$ of $a_0 a_1 \ldots a_{l-1}$ is defined by

$$m_{ij} = a_i + a_{(q+j) \mod l}, \quad 1 \leq i < l, \quad 0 < q < l,$$

where + between symbols stands for addition mod 2. The cross correlation matrix $N = (n_{ij})_{i=1, q=0}^{l, q=0}$ of $a_0 a_1 \ldots a_{l-1}$ is defined by

$$n_{ij} = \begin{cases} m_{ij} & \text{if } i + q < l, \\ \overline{m}_{ij} & \text{if } i + q > l. \end{cases}$$

Given $1 = k(1) < k(2) \ldots < k(g) < l$ and a matrix $M = (m_{ij})_{i=1, q=0}^{l, q=0}$ we call the matrix $\hat{M} = (m_{k(q)+j})_{i=1, q=0}^{l, q=0}$ the reduced matrix.

As in the last section we split the atoms of $\xi^{gn}$ by $\xi$ and consider $\eta_n$ with atoms

$$\left[ b_0(b_1; \ldots ; b_{gn}) \right] \cup \left[ \overline{b}_0(\overline{b}_1; \ldots ; \overline{b}_{gn}) \right].$$

Then we have

$$h_A^{\top}(T_\theta, \xi) = \frac{1}{g} \limsup_{n \to \infty} \frac{1}{n} \log N(\eta_n),$$

$$h_A(T_\theta, \xi) = \frac{1}{g} \limsup_{n \to \infty} \frac{1}{n} H(\eta_n).$$

Let $[0(b_1; \ldots ; b_{gn})]$ be a nonempty atom of $\xi \vee \xi^{gn}$. This is equivalent to the existence of an integer $p$ such that

$$w_p = 0, \quad w_{p+k(i)} = b_{i+1}, \quad 1 \leq i < g, \quad 0 < j < n.$$  

For each $q$ with $0 < q < l$

$$w_{p+q} w_{p+q+1} \ldots ; w_{p+q+k(i)} \ldots ; w_{p+q+k(g)i} (1 < i < g, 0 < j < n)$$

is an admissible $\xi \vee \xi^{gn}$-block. Now equations (4) and (5) imply that

$$w_p = 0,$$

$$w_{p+q+k(i)} = \begin{cases} a_{q+k(i)} & \text{if } q + k(i) < l, \\ a_{q+k(i)-l} + b_1 & \text{if } q + k(i) > l, \end{cases}$$

$$w_{p+q+k(i)} = b_{i+1} + a_q, \quad 1 \leq i < g, 1 < j < n.$$  

Considering the cases $a_q = 0$ and $a_q = 1$ separately for $q = 0, \ldots , l-1$, we obtain as in the last section

**Lemma 9.** Let $\theta$ be a continuous substitution of length $l$ and $A$ as in ($\#$). Let $\hat{M}$ and $\hat{N}$ be the reduced auto, respectively cross correlation matrices of $\theta 0$. Let $0(b_1; \ldots ; b_{gn})$ be an admissible $\xi \vee \xi^{gn}$-block. Then, if $b_1 = 0$

$$0(\hat{m}_1; \hat{m}_2; \ldots ; \hat{m}_g; b_1; \ldots ; b_{gn}), \quad q = 0, \ldots , l-1,$$

and if $b_1 = 1$

$$0(\hat{n}_1; \hat{n}_2; \ldots ; \hat{n}_g; b_1; \ldots ; b_{gn}), \quad q = 0, \ldots , l-1,$$

are admissible $\xi \vee \xi^{gn}$-blocks. Moreover, if $0(c_1; \ldots ; c_g; b_1; \ldots ; b_{gn})$ with $c_q = 0$ or $1$ for $1 < i < g$ is an admissible $\xi \vee \xi^{gn}$-block, then this block is among (17) or (18).
With the columns occurring in the reduced auto or cross correlation matrices \( \hat{M} \) or \( \hat{N} \) of \( \theta \) we associate nonnegative integers \( s \) with \( 0 < s < 2^g \) called states simply by considering a column as the binary notation of \( s \). Thus

\[
(1, 2, \ldots, 2^{g-1}) \hat{M} =: (s_1, s_2, \ldots, s_g) =: v_0
\]
is the vector of states associated with \( \hat{M} \). Analogously \( (1, 2, \ldots, 2^{g-1}) \hat{N} =: v_1 \) is the vector of states associated with \( \hat{N} \). Let \( S \) be the set of states which occur in \( v_0 \) or \( v_1 \). We partition \( S \) into two sets \( S_0 \) and \( S_1 \) by \( s \in S_r \) iff \( s = r \) (mod 2). States in \( S_0 \) are called even, those in \( S_1 \) odd. We illustrate this terminology with an example.

**Example 5.** Let \( \theta_0 = 0011 \), \( A = (4^n, 3 \cdot 4^n)_{n \geq 0} \). Then the auto and cross correlation matrices of \( \theta_0 \) are

\[
M = \begin{pmatrix}
0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0
\end{pmatrix}, \quad N = \begin{pmatrix}
0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1
\end{pmatrix}.
\]
The reduced matrices are

\[
\hat{M} = \begin{pmatrix}
0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0
\end{pmatrix} \quad \text{and} \quad \hat{N} = \begin{pmatrix}
0 & 1 & 0 & 0 \\
1 & 1 & 0 & 1
\end{pmatrix}.
\]

\( v_0 = (2, 1, 2, 1), \ v_1 = (2, 3, 0, 2) \) and \( S_0 \cup S_1 = S = \{0, 2, 1, 3\} \).

We consider \( S^\mathbb{Z} \) with shift \( T \) and time-0 partition \( \zeta = \{[s]; s \in S\} \). We set \( \xi_n := \zeta \sqrt{T^{-1}\zeta \cdot \ldots \cdot T^{-n+1}\zeta} \).

Any admissible \( \xi \sqrt{T\xi} \)-block or its mirror image occurs in

\[
a_0a_1 \ldots a_{\ell-1}a_0a_1 \ldots a_{k(g)-1} \quad \text{or} \quad a_0a_1 \ldots a_{\ell-1}a_0a_1 \ldots a_{k(g)-1}.
\]
The columns of \( \hat{M} \) have just the same components as the \( \xi^g \)-blocks preceded by 0, or their mirror images, in case they are preceded by 1, which occur in the former block. Analogously columns of \( \hat{N} \) correspond to \( \xi^g \)-blocks occurring in the latter block. We therefore have a one-to-one correspondence between nonempty atoms of \( \eta_1 \) (see (13)) and states in \( S_r \), or equivalently, with nonempty atoms of \( \xi \). Moreover, atoms \([0(b_1; \ldots; b_g)] \cup [1(b_1; \ldots; b^g)] \) with \( b_1 = 0 \) correspond to even states and those with \( b_1 = 1 \) to odd states. Let \([0(b_1; \ldots; b^g_g)] \) be a nonempty atom of \( \xi \sqrt{T\xi} \). As in the Basic Property it follows that \( 0(b_1; \ldots; b^g_g), 0(b^g_{g+1}; \ldots; b^g_{2g}), \ldots, 0(b^g_{g(n-1)+1}; \ldots; b^g_{g(n)}) \) are admissible \( \xi \sqrt{T\xi} \)-blocks. If \( s_n, s_{n-1}, \ldots, s_1 \) are the corresponding states (note the order inversion), then we let the atom \([0(b_1; \ldots; b^g_g)] \cup [1(b_1; \ldots; b^g_g)] \) of \( \eta_n \) correspond to the atom \([s_1] \cap T^{-1}[s_2] \cap \ldots \cap T^{-n+1}[s_n] \) of \( \xi_n \).

Let \( a \) be a matrix on \( S \times S \) defined by

\[
a_{st} = \begin{cases} 
1 & \text{if } s \in S_r \text{ and } t \text{ occurs in } v_r \text{ for } r = 0 \text{ or } 1, \\
0 & \text{otherwise}.
\end{cases}
\]

Let \( T_a \) be the topological Markov shift defined by \( a \) ([3], [15]). With Lemma 9 it is easily shown by induction that the above defined correspondence is one-to-one between nonempty atoms of \( \eta_n \) and of \( \xi_n \) w.r.t. \( T_a \). With the aid of Theorem 5 we obtain

**Theorem 6.** Let \( \theta \) be a continuous substitution of length \( l \) and let \( A \) be as in (\#). Let \( T_a \) be the topological Markov shift determined by (19). Then \( h^\text{top}(T_a) = g^{-1}h^\text{top}(T_\theta) \).
It is well known ([3], [17]) that \( h_{\text{top}}(T_\theta) \) equals \( \log \lambda \), where \( \lambda \) is the largest eigenvalue of \( \sigma \). The matrix \( \sigma \) in (16) has a special form: transitions only depend on whether a state is even or odd. Let \( \sigma^o \) be the \( 2 \times 2 \) matrix defined by \( \sigma^o = \sum_{r \in S} a^o_r \) where \( s \in S_r \). Then the largest eigenvalue of \( \sigma^o \) equals \( \lambda \). To prove this, note that \( \lambda \) is an eigenvalue of \( \sigma^o \) with a nonnegative eigenvector. Since \( \sigma^o \) is irreducible (a component of \( \sigma^o \) equals zero iff the corresponding component in \( \langle \beta_r \rangle \) equals zero, where \( \alpha \) and \( \beta \) are as in the proof of Theorem 3), the uniqueness of nonnegative eigenvectors [5, p. 63] implies that \( \lambda \) is the largest eigenvalue of \( \sigma^o \). This observation facilitates the calculation of \( h_{\text{top}}(T_\theta) \).

**Example 5 (continued).** Here

\[
\begin{pmatrix}
0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 \\
1 & 1 & 0 & 1
\end{pmatrix}, \quad \sigma^o = \begin{pmatrix}
1 & 1 \\
2 & 1
\end{pmatrix}.
\]

Hence \( h_{\text{top}}(T_\theta) = \frac{1}{2} \log(1 + \sqrt{2}) \).

**Example 6** [9, §7, Example 3]. Let \( \theta = 001, A = \{3^n, 2 \cdot 3^n\} \). Then \( M = \hat{M} = \begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}, N = \hat{N} = \begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}, v_0 = (2 1 3), v_1 = (2 3 0), S = \{0, 2, 3\} \) and \( \sigma^o = \begin{pmatrix}
0 & 1 \\
1 & 1
\end{pmatrix} \). Hence \( h_{\text{top}}(T_\theta) = \frac{1}{2} \log 3 \).

We turn to the calculation of measure-theoretic sequence entropy. Let a probability vector \( \mu = (p_s)_{s \in S} \) be defined by

\[
p_s = \mu(B_s), \quad s \in S,
\]

where \( B_s \) is the atom of \( \eta_1 \) corresponding to \( s \). Let a Markov matrix \( \Pi = \Pi(\theta, A) = (\pi_{st})_{s,t \in S} \) be defined by

\[
\pi_{st} = \ell^{-1} \text{ (number of times } t \text{ occurs in } v_r) \quad \text{if } s \in S_r, \text{ for } r = 0 \text{ or } 1.
\]

**Lemma 10.** Let \( p \) and \( \Pi \) be defined by (20) and (21). Then \( p \Pi = p \).

**Proof.** Let \( \Pi^o \) be the Markov matrix obtained by lumping [11] all states in \( S_r \) to a single state \( r \) for \( r = 0, 1 \). Let \( p^o \) be the probability vector \( (\Sigma_{s \in S_0} p_s, \Sigma_{s \in S_1} p_s) \). Then \( \Pi^o = \Pi(\theta, \{^1T^\theta\}_{s \in S}) \) and \( p^o = (\mu([00] \cup [11]), \mu([01] \cup [10])) \). According to Lemma 8 \( p^o \Pi^o = p^o \). This suffices to have \( p \Pi = p \).

Let \( T_\Pi \) be the Markov shift on \( S^Z \) defined by \( p \) and \( \Pi \). As in the last section one shows that our one-to-one correspondence between \( \eta_1 \) and \( \xi_1 \) is measure preserving. One then has

**Theorem 7.** Let \( \theta \) be a substitution of length \( l \), \( A \) as in (4). Let \( T_\Pi \) be the Markov shift defined by (20) and (21). Then \( h_\Pi(T_\theta) = g^{-1} h(T_\Pi) \).

In our situation the formula for the entropy of \( T_\Pi \) [3, Proposition 12.3] simplifies to

\[
h(T_\Pi) = -2\mu([00]) \sum_{t \in S} \pi_{st} \log \pi_{st} - 2\mu([01]) \sum_{t \in S} \pi_{st} \log \pi_{st},
\]

where \( s_0 \) is any even and \( s_1 \) any odd state.
SEQUENCE ENTROPY AS AN ISOMORPHISM INVARIANT

Example 6 (continued). Here $\Pi = \frac{1}{3} \sigma$. Therefore $h_\sigma(T_\theta) = \frac{1}{3} \log 3$. Note that in this example $h_\sigma^{\text{top}}(T_\theta) = h_\sigma(T_\theta)$, i.e. the variational principle holds for this substitution shift with $A = \{3^n, 2 \cdot 3^n\}$. It does not hold with $A = \{3^n\}$, cf. Theorem 3 and Example 1.

Example 7. Let $\theta_0 = 00001$. Then

\[
M = \begin{bmatrix}
0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 \\
\end{bmatrix}, \quad N = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 0 \\
\end{bmatrix},
\]

$\alpha$ = number of zeroes in first row of $M = 3$, $\beta$ = number of zeroes in first row of $N = 4$. So by Lemma 8, $2\mu((00)) = \beta(l - \alpha + \beta)^{-1} = \frac{2}{3}$. First let $A = \{5^n\}$. Then $\hat{M} = (0011)$, $\hat{N} = (00011)$, $S = \{0, 1\}$ and $\Pi = \frac{1}{2}(\alpha \beta)$. Hence

\[
h_{\{5^n\}}(T_\theta) = \log 5 + \frac{2}{3} \left[-\frac{2}{3} \log 3 - \frac{2}{3} \log 2\right] + \frac{1}{3} \left[-\frac{4}{3} \log 4\right] = 0.61547 \ldots .
\]

Now let $A = \{5^n, 3 \cdot 5^n\}$. Then

\[
\hat{M} = \begin{bmatrix}
0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 \\
\end{bmatrix}, \quad \hat{N} = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 \\
\end{bmatrix},
\]

$\nu_0 = (02013)$, $\nu_1 = (02230)$, $S = \{0, 2, 1, 3\}$. The transition probabilities from 0 and 2 are therefore $\frac{1}{2}(2, 1, 1, 1)$ and those from 1 and 3 are $\frac{1}{2}(2, 2, 0, 1)$. Hence

\[
h_{\{5^n, 3 \cdot 5^n\}}(T_\theta) = \frac{1}{2} \left\{ \log 5 + \frac{3}{2} \left[-\frac{2}{3} \log 2\right] + \frac{1}{2} \left[-\frac{4}{3} \log 2\right] \right\} = 0.61987 \ldots .
\]

In this example we have $h_{\{5^n\}}(T_\theta) < h_{\{5^n, 3 \cdot 5^n\}}(T_\theta)$, although $\{5^n\}$ is a subsequence of $\{5^n, 3 \cdot 5^n\}$.

Example 8. Let $\theta_0 = 0111001$, $\theta_2 = 0110001$ and $A = \{7^n, 4 \cdot 7^n, 6 \cdot 7^n\}$. Then $\alpha(\theta_0) = l - \beta(\theta_0) = 2\mu((00)) = \frac{1}{2}$ for $i = 1, 2$, where $\mu_i$ is the $T$-invariant measure on $X_{\theta_i}$. Furthermore,

\[
\hat{M}(\theta_0) = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 \\
\end{bmatrix}, \quad \hat{N}(\theta_0) = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 0 & 1 \\
\end{bmatrix},
\]

$\nu_0(\theta_0) = (5 \ 6 \ 0 \ 3 \ 6 \ 3 \ 5)$, $\nu_1(\theta_0) = (5 \ 2 \ 4 \ 5 \ 0 \ 5 \ 2)$, $S(\theta_0) = \{0, 2, 4, 6, 3, 5\}$. The transition probabilities from an even state are uniformly distributed over the positive states and those from an odd state are $\nu_1(\theta_0)$. Now for $\theta_2$,

\[
\hat{M}(\theta_2) = \begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 \\
\end{bmatrix}, \quad \hat{N}(\theta_2) = \begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 \\
\end{bmatrix},
\]

$\nu_0(\theta_2) = (5 \ 6 \ 1 \ 4 \ 2 \ 3 \ 7)$, $\nu_1(\theta_2) = (5 \ 2 \ 5 \ 2 \ 4 \ 5 \ 0)$, $S(\theta_2) = \{0, 2, 4, 6, 1, 3, 5, 7\}$. The transition probabilities from an even state are uniformly distributed over the positive states and those from an odd state are uniformly distributed over the positive states and those from an odd state are $\nu_1(\theta_2)$. Therefore

\[
h_\sigma(T_\theta) - h_\sigma(T_\theta) = \frac{1}{3} \left\{ \frac{1}{2} \left[-\frac{6}{7} \log \frac{6}{7} - \frac{1}{3} \log \frac{1}{3}\right] \right\} - \frac{1}{3} \left\{ \frac{1}{2} \log 7\right\} = \frac{1}{3} \log 2.
\]

Hence $T_{\theta_1}$ and $T_{\theta_2}$ are not isomorphic. However, $\theta_2^2 = (\theta_1^{-1})^2$, so it follows by Lemma 4 that $T_{\theta_1} \simeq T_{\theta_2}^{-1}$.
Example 8 strengthens a result of Gottschalk [7] who proved that \( T_\theta \cong T_\theta^{-1} \) if \( \theta \) is the continuous substitution defined by \( \theta^0 = 0010 \). Actually one can show that \( T_\theta \cong T_\theta^{-1} \), and consequently \( T_\theta \cong T_\theta^{-1} \), for this \( \theta \) by calculating the \( \{16^n, 11 \cdot 16^n, 13 \cdot 16^n\} \)-entropy for \( \theta^2 \) and \( (\theta^{-2})^2 \).

5. The isomorphism problem for substitution shifts. We once more only consider systems arising from continuous substitutions on two symbols. In [1] the problem of topological classification of such systems is solved for substitutions having the same length. It is stated without proof that two systems arising from continuous substitutions are measure-theoretically isomorphic iff the substitutions have a common power. This would imply the corresponding topological result which has been proved in [15]. Concerning the measure-theoretical classification there are some results in [10] where an invariant is obtained from calculations on the spectral measure of \( T_\theta \). If \( M \) and \( N \) are the auto and cross correlation matrices of \( \theta^0 \), let us call \( M' := \frac{1}{2}(M + N) \) the aperiodic auto correlation matrix of \( \theta^0 \). It is shown in [9] that the vector of row sums of \( M' \) is an invariant for \( T_\theta \)'s arising from continuous substitutions of the same length \( l \). Obviously \( M'(\theta) = M'(\theta^{-1}) \), so this invariant does not distinguish between \( T_\theta \) and \( T_\theta^{-1} \) or \( T_\theta^{-1} \).

Consider the collection of sequences

\[
\{ A = \{ k(1)l^m, \ldots, k(g)l^m \} : m > 1, 1 = k(1) < \cdots < k(g) < l^m \}.
\]

We conjecture that the corresponding collection of sequences' entropies is complete among \( T_\theta \)'s arising from substitutions of the same length \( l \), i.e. if \( T_\theta \cong T_\theta \) then there is an \( A \) in this collection such that \( h_\lambda(T_\theta) = h_\lambda(T_\theta) \). The essential problem is to compare two \( T_\theta \)'s arising from substitutions whose lengths are different but possess the same prime factors (the latter is the case iff their pure point spectrum is equal, see [1, p. 96]). We do not know an \( A \) such that the \( A \)-entropy can be computed for both systems.

ADDED IN PROOF. From [19] follows that \( T_\theta \cong T_\theta \) for two continuous substitutions of the same length iff \( \theta^0 \cong \theta^0 \).

REFERENCES


Mathematisch Instituut, Katholieke Universiteit, Toernooiveld, 6525 ED Nijmegen, The Netherlands