APPROXIMATING THE MAXIMUM ERGODIC AVERAGE VIA PERIODIC ORBITS

D. COLLIER AND I.D. MORRIS

ABSTRACT. Let $\sigma: \Sigma_A \to \Sigma_A$ be a subshift of finite type, let $\mathcal{M}_\sigma$ be the set of all $\sigma$-invariant Borel probability measures on $\Sigma_A$, and let $f: \Sigma_A \to \mathbb{R}$ be a Hölder continuous observable. There exists at least one $\sigma$-invariant measure $\mu$ which maximises $\int f \, d\mu$. The following question was asked by B.R. Hunt, E. Ott and G. Yuan: how quickly can the maximum of the integrals $\int f \, d\mu$ be approximated by averages along periodic orbits of period less than $p$? We give an example of a Hölder observable $f$ for which this rate of approximation is slower than stretched-exponential in $p$.

AMS subject classification: 37B10, 37D20, 40J30.

1. INTRODUCTION

Let $\sigma: \Sigma_A \to \Sigma_A$ be a subshift of finite type, let $\mathcal{M}_\sigma$ be the set of all $\sigma$-invariant Borel probability measures on $\Sigma_A$, and let $f: \Sigma_A \to \mathbb{R}$ be Hölder continuous. There is at least one invariant measure $\mu \in \mathcal{M}_\sigma$, which we term a maximising measure for $f$, such that

$$\int f \, d\mu = \beta(f) = \sup_{\nu \in \mathcal{M}_\sigma} \int f \, d\nu.$$ 

In this paper we investigate the problem of finding invariant measures supported on periodic orbits which approximately realise this maximum. More specifically, if we let $\mathcal{M}_{\sigma,p}$ be the set of all ergodic $\sigma$-invariant measures supported on points fixed by $\sigma^p$, we investigate the quantity

$$E_f(p) = \sup_{\nu \in \mathcal{M}_\sigma} \int f \, d\nu - \sup_{\nu \in \mathcal{M}_{\sigma,p}} \int f \, d\nu,$$

introduced by G. Yuan and B.R. Hunt [11], being the difference between the maximum integral $\beta(f)$ of $f$ and the ‘best approximating’ periodic orbit whose period divides $p$.

Our motivation is twofold. Firstly, the above problem is of interest in control theory: if we are able to control the dynamical system $\sigma: \Sigma_A \to \Sigma_A$, we will wish to direct the trajectories of points in $\Sigma_A$ toward particular unstable orbits which optimise some given quantity. For reasons of computational simplicity, it is desirable that such orbits should be periodic and have small length. This leads us to ask how close to optimal orbits of a specified period may be. These considerations motivate the work of G. Yuan, B. R. Hunt and E. Ott [7, 11].

Secondly, the quantity $\beta(f)$ has shown itself to be of interest in a number of situations in ergodic theory, both intrinsically [1, 2, 6, 8] and in application to existing problems [3, 4, 9]. It is therefore of interest to be able to compute this quantity accurately in numerical experiments. One obvious approach to this task would be
to exhaustively compute ergodic averages of \( f \) along periodic orbits of length up to \( n \), and take the supremum of these averages as an approximation to \( \beta(f) \). The error incurred in this approximation would therefore equal \( \inf_{1 \leq p \leq n} E_p(f) \).

It is classical [10] that \( \cup_{p=1}^{\infty} \mathcal{M}_{\sigma,p} \) is dense in \( \mathcal{M}_{\sigma} \) in the weak-* topology, and so \( E_f(p) \to 0 \) as \( p \to \infty \) for all continuous \( f \). We wish to investigate the rate at which this convergence occurs when \( f \) is Hölder.

The behaviour of \( E_f(p) \) as \( p \to \infty \) is at present poorly understood. On the strength of numerical experiments [7] combined with rigorous analysis, Yuan and Hunt [11] observed that the sequence \( E_f(p) \) often decays to zero at an exponential rate, but were unable to prove this in generality. They then asked whether it could be shown that \( E_f(p) \) always decays exponentially.

The purpose of this paper is to answer this question in the negative. We have

**Theorem 1.** Let \( \sigma : \Sigma_A \to \Sigma_A \) be a subshift of finite type. There exists Hölder continuous \( f : \Sigma_A \to \mathbb{R} \) such that \( E_f(p) \) tends to 0 at a slower than stretched-exponential rate: \( \log E_f(p) = o(p^\epsilon) \) for every \( \epsilon > 0 \).

We note that a result similar to Theorem 1 has recently been proved by X. Bressaud and A. Quas in [5], in which the quantity \( \inf_{1 \leq p \leq n} E_p(f) \) is considered. Bressaud and Quas are able to obtain upper and lower bounds which are superior to those in the present article, but with the weakness that their lower bound applies only along subsequences of integers \( n \).

### 2. Notation and definitions

Let \( A \) be an irreducible aperiodic \( N \times N \) matrix of zeroes and ones. We define the shift space associated to \( A \) to be the set

\[
\Sigma_A := \{ x = (x_i)_{i=1}^n : x_i \in \{1, \ldots, N\} \text{ and } A(x_i, x_{i+1}) = 1 \text{ for all } i \geq 1 \},
\]

and define the shift map \( \sigma : \Sigma_A \to \Sigma_A \) by

\[
(\sigma x)_i = x_{i+1}
\]

for all \( i \geq 1 \). Given \( \theta \in (0,1) \) we define the \( \theta \)-metric \( d_\theta \) on \( \Sigma_A \) by

\[
d_\theta ((x_i)_{i \geq 1}, (y_i)_{i \geq 1}) = \theta^{\inf \{ n \geq 1 : x_n \neq y_n \}}.
\]

We say that a function \( f : \Sigma_A \to \mathbb{R} \) is \( \theta \)-Hölder continuous if it is Lipschitz continuous with respect to the metric \( d_\theta \). We fix \( \theta \in (0,1) \) for the remainder of this paper.

We define a finite word to be a finite sequence \( \omega = (\omega_i)_{i=1}^n \) taking values in the set \( \{1, 2, \ldots, N\} \). We say that \( \omega = (\omega_i)_{i=1}^n \) is compatible with the matrix \( A \) if \( A(\omega_i, \omega_{i+1}) = 1 \) for all \( i < n \). We define the length of the word \( \omega \) to be \( n \). We will on occasion describe elements of \( \Sigma_A \) as infinite words. We say that finite words \( \omega^1, \omega^2 \) of equal length \( n \) are rotation equivalent if there exists a nonnegative integer \( r < n \) such that \( \omega^1_i = \omega^2_{i+r} \) whenever \( 1 \leq i \leq n - r \) and \( \omega^1_i = \omega^2_{i+n-r} \) whenever \( n-r < i \leq n \). When \( \omega^1 \) and \( \omega^2 \) are rotation equivalent we write \( \omega^1 \simeq \omega^2 \). Words of distinct lengths are never rotation equivalent.

Given two words \( \omega^1, \omega^2 \) with lengths \( n_1 \) and \( n_2 \), we define their concatenation \( \omega^1 \cdot \omega^2 \) to be the word of length \( n_1 + n_2 \) given by \( [\omega^1 \cdot \omega^2]_i = \omega^1_i \) when \( 1 \leq i \leq n_1 \), and \( [\omega^1 \cdot \omega^2]_i = \omega^2_i \) when \( n_1 + 1 \leq i \leq n_1 + n_2 \). Given a finite list of words \( \omega^1, \ldots, \omega^m \), we denote the compound concatenation \( \omega^1 \cdot \omega^2 \cdot \ldots \cdot \omega^m \) by \( \prod_{k=1}^m \omega^k \). Concatenation is
associative. If $\omega^1, \omega^2$ are finite words with lengths $n_1, n_2$ compatible with $A$, their concatenation is compatible with $A$ if and only if $A(\omega^1_{n_1}, \omega^2_{n_2}) = 1$.

For each $p > 0$ we let $\Omega_p$ be the set of all words of length $p$ which are compatible with $A$. We let $\Omega_p$ be the set of all words $\omega \in \Omega_p$ such that $A(\omega_p, \omega_1) = 1$. Note that $\omega \in \Omega_p$ if and only if both $\omega \in \Omega_p$ and $\omega \cdot \omega \in \Omega_{2p}$.

If $x \in \Sigma_A$ it is clear that that $\sigma^p x = x$ if and only if $x_i = x_{i+p}$ for all $i \geq 1$, if and only if there exists $\omega \in \Omega_p$ such that $x_{i+kp} = \omega_i$ for all $1 \leq i \leq p$ and $k \geq 0$. In this case we write $x = \pi(\omega)$. This defines a relationship between the sets $\text{Fix}_p = \{x \in \Sigma_A : \sigma^p x = x\}$ and $\Omega_p$, which may readily be seen to be bijective. Moreover, we have $\pi(\omega^1) = \sigma^j \pi(\omega^2)$ for some $j \geq 0$ if and only if $\omega^1 \simeq \omega^2$.

If $a = (a_i)_{i=1}^n$ is a finite word and $b$ is either a finite word or an element of $\Sigma_A$, we write $a \prec b$ if there is $k \geq 0$ such that $a_i = b_{i+k}$ for all $1 \leq i \leq n$. In this case we say that $a$ is a subword of $b$.

### 3. Proof of Theorem 1

We begin with the following:

**Proposition 3.1.** Let $K = \sigma K$ be a closed nonempty subset of $\Sigma_A$ and define $f_K(x) = -d_0(K, x)$ for all $x \in \Sigma_A$. Clearly $f_K$ is $\theta$-Hölder continuous. For each $\omega \in \Omega_p$ we define

$$\xi(\omega, K) := \inf_{\omega' \subset \omega \subset \omega_{\text{lex}}} \{ \ell > 0 : \exists a \in \Omega_{\ell} \text{ such that } a \prec \omega' \text{ and } \forall x \in K, \ a \not\prec x \}. $$

Let $p > 0$, and suppose that $\sup_{\omega \in \Omega_p} \xi(\omega, K) \leq p$. Then

$$E_{f_K}(p) \geq \frac{1}{p} \sup_{\omega \in \Omega_p} \xi(\omega, K) = \frac{1}{p} \inf_{\omega \in \Omega_p} \xi(\omega, K).$$

**Proof.** Let $p > 0$ and $\omega \in \Omega_p$ where $\xi(\omega, K) \leq p$. We have

\[
\log_{\theta} \sup_{\omega \subset \omega_{\text{lex}}} d(\pi(\omega'), K) = \log_{\theta} \sup_{\omega \subset \omega_{\text{lex}}} \inf_{x \in K} d_0(\pi(\omega'), x) \\
\quad = \inf_{\omega \subset \omega_{\text{lex}}} \sup_{x \in K} \inf_{\ell > 0} \{ \ell > 0 : x_\ell \not= \pi(\omega')_\ell \} \\
\quad = \inf_{\omega \subset \omega_{\text{lex}}} \sup_{x \in K} \{ \ell > 0 : \exists x \in K \text{ s.t. } x_i = \pi(\omega')_i \forall 1 \leq i < \ell \} \\
\quad = \inf_{\omega \subset \omega_{\text{lex}}} \{ \ell > 0 \ : \ \exists a \in \Omega_{\ell} \text{ s.t. } a = \pi(\omega') \forall 1 \leq i \leq \ell \text{ and } \forall x \in K, a \not\prec x \} \\
\quad = \inf_{\omega \subset \omega_{\text{lex}}} \{ \ell > 0 \ : \ \exists a \in \Omega_{\ell} \text{ s.t. } a \prec \pi(\omega') \text{ and } a \not\prec x \ \forall \ x \in K \} \\
\quad = \inf_{\omega \subset \omega_{\text{lex}}} \{ \ell > 0 \ : \ \exists a \in \Omega_{\ell} \text{ s.t. } a \prec \omega' \text{ and } a \not\prec x \ \forall \ x \in K \} \\
\quad = \xi(\omega, K),
\]

where we have used $\xi(\omega, K) \leq p$ in the second-from-last equality. Since $K$ is closed, $\sigma$-invariant and nonempty, the Krylov-Bogolioubov Theorem shows that there exists
\[ \mu \in \mathcal{M}_\sigma \text{ such that } \mu(K) = 1. \text{ It follows that } \beta(f) = 0, \text{ and so} \]

\[
E_{f_\epsilon}(p) = \inf_{\sigma^j x = a} \frac{1}{p} \sum_{j=0}^{p-1} d(\sigma^j x, K) \\
\geq \frac{1}{p} \inf_{\sigma^j x = a} \sup_{0 \leq j < p} d(\sigma^j x, K) \\
= \frac{1}{p} \inf_{\omega \in \Omega_p} \sup_{\omega', \omega \in \omega} d(\pi(\omega'), K) \\
\geq \frac{1}{p} \inf_{\omega \in \Omega_p} \beta(\omega, K),
\]
as required. \qed

It follows that to prove Theorem 1, it suffices to construct a nonempty compact set \( K = \sigma K \subseteq \Sigma_A \) such that \( \sup_{\omega \in \Omega_p} \xi(\omega, K) = o(p^d) \) as \( p \to \infty \). The remainder of this section is dedicated to this task. We will construct the invariant set \( K \) recursively, using a sequence of sets of words \( M_n \) of increasing length. The set \( K \) will then arise as a limit of these sets.

Let \( M_1 \) be a subset of \( \Omega_{\ell_1} \) for some positive integer \( \ell_1 \) such that \( m_1 = \text{Card} \ M_1 \) is divisible by \( 2^{16} \). We require that \( M_1 \) have the following properties:

**Definition 3.2.** There exists an integer \( P \geq 1 \) and word \( z = (z_i)_{i=1}^P \in \Omega_p \) such that:

(i) Every \( \omega \in M_1 \) satisfies \( z_i = \omega_i \) for all \( 1 \leq i \leq P \).

(ii) If \( \omega = \omega^1 \cdot \omega^2 \) where \( \omega^1, \omega^2 \in M_1 \), and \( z_i = \omega_{i+r} \) for all \( 1 \leq i \leq P \), then either \( r = \ell_1 \) or \( r = 0 \).

(iii) \( A(\omega^1, \omega^2) = 1 \) for all \( \omega^1, \omega^2 \in M_1 \); that is, \( \omega^1 \cdot \omega^2 \in \Omega_{2\ell_1} \) for every pair \( \omega^1, \omega^2 \in M_1 \).

The reader may verify that such a set \( M_1 \) may be constructed for any prescribed \( m_1 \) and matrix \( A \).

We will define a sequence of sets \( M_n \) with cardinalities \( m_n \) consisting of words of length \( \ell_n \) in an inductive fashion starting with the set \( M_1 \). We begin by defining some integer sequences which will be crucial to our construction.

**Definition 3.3.** Given an integer \( m_1 \) divisible by \( 2^{16} \), let \( q_1 = t_1 = m_1/4 \). Define sequences \( (m_n)_{n \geq 1}, (q_n)_{n \geq 1} \) and \( (t_n)_{n \geq 1} \) as follows. Given the integer \( t_n \), we let \( m_{n+1} = 4^{q_n} \) and \( q_{n+1} = \left\lfloor \frac{m_{n+1}}{4} \right\rfloor - 1 \), and let \( \tau_{n+1} \) be the unique positive real number such that

\[
\tau_{n+1}^2 = 4^{q_n} - 1 = q_{n+1}.
\]

We then define

\[
t_{n+1} = 4^{\left\lfloor \frac{t_n}{\tau_{n+1}} \right\rfloor - 1}.
\]

The key properties of these sequences are stated in the lemma below.

**Lemma 3.4.** For each \( n \geq 1 \), we have \( t_n \mid q_n, m_n/q_n = 4 \) and \( t_n \geq 2^{14} \). Moreover, the sequence \( t_n \) satisfies \( t_n/\sqrt{t_{n+1}} \leq 16 t_n^{1/4} \) for all \( n \geq 1 \) and

\[
\lim_{n \to \infty} \frac{t_n}{t_{n+1}} = 0.
\]
Proof. The first two statements are clear. We consider the sequence \((t_n)_{n \geq 1}\); the definition implies that
\[
\frac{1}{16} t_{n+1} = 4 \sqrt{t_{n+1}} - 2 \leq 4 \left| \frac{t_n}{\sqrt{t_{n+1}}} \right|^{-1} = t_{n+1} \leq 4 \sqrt{t_{n+1}} - 1 \leq t_{n+1},
\]
for every \(n \geq 1\). We proceed inductively. Given that \(t_n \geq 2^{14}\), we notice that
\[
\sqrt{t_{n+1}} \log t_{n+1} = (t_n - 1) \log 4.
\]
We thus have
\[
t_n = \sqrt{t_{n+1}} \log_4 t_{n+1} + 1 > \sqrt{t_{n+1}},
\]
and therefore
\[
\log_4 t_{n+1} < 2 \log_4 t_n.
\]
We deduce
\[
\sqrt{t_{n+1}} \geq \frac{1}{4} \sqrt{t_{n+1}} = \frac{t_n - 1}{4 \log_4 t_{n+1}} \geq \frac{t_n - 1}{8 \log_4 t_n} \geq \frac{1}{8} \frac{3^n}{4} - \frac{1}{8},
\]
where we have used the elementary inequality \(\log t \leq t^{1/4}\) for all \(t \geq 16\). Thus
\[
t_{n+1} \geq \frac{1}{64} \left( \frac{3^n}{4} - 1 \right)^2 \geq 2^{-6} \left( 2^{2n} - 2^{11} \sqrt{2} + 1 \right) > 2^{14}
\]
for every \(n \geq 1\), making it clear that \(t_n/\sqrt{t_{n+1}} \leq 16 t_n^{1/4}\). One may easily use the above to show that \(t_{n+1} \geq t_n + \frac{1}{4^n}\) for each \(n \geq 2\), implying that \(t_n \to \infty\); since for each \(n\) we have
\[
0 \leq \frac{t_n}{t_{n+1}} \leq \frac{4t_n}{t_n^{3/2} - 2t_n^{3/4} + 1},
\]
it follows that \(\lim_{n \to \infty} t_n/t_{n+1} = 0\) as required.

For an integer \(n \geq 1\), a finite word \(a = (a_i)_{i=1}^m\) and a finite or infinite word \(b\), we shall write \(a \prec_n b\) if there is \(k \geq 0\) such that \(a_i = b_{i+k t_n}\) for all \(1 \leq i \leq m\).

The distinction between \(\prec\) and \(\prec_n\) will be important since we will construct words \(\omega \in M_{n+1}\) as concatenations of words \(a, b, c, \ldots \in M_n\). For example, if \(a, b, c \in M_n\), then it is true that \(b \cdot c \prec_n a \cdot b \cdot c \cdot a\) and \(c \cdot b \not\prec_n a \cdot b \cdot c \cdot a\); however, the statement \(c \cdot b \prec a \cdot b \cdot c \cdot a\) could be either true or false, depending on the subword structure of the words \(a, b\) and \(c\).

Given \(n \geq 1\) and the set \(M_n\), we construct the set \(M_{n+1}\) as follows. Recalling that \(m_n = 4 q_n\), partition \(M_n\) into \(q_n\) disjoint sets \(C^k\) of cardinality 4, where \(1 \leq k \leq q_n\). For each \(k\), we will write \(C^k = \{c^j_k : 1 \leq j \leq 4\}\). Define
\[
I_n = \{(i_1, i_2, \ldots, i_{q_n}) \in \{1, 2, 3, 4\}^{q_n} : i_m = i_{m+t_n} \text{ for all } 1 \leq m \leq q_n - t_n\}.
\]
For each \(q_n\)-tuple \((i_1, \ldots, i_{q_n}) \in I_n\), we construct the word
\[
\omega(i_1, \ldots, i_{q_n}) := \prod_{k=1}^{q_n} c^k_{i_k}.
\]
We let \(M_{n+1} = \{\omega(i_1, \ldots, i_{q_n}) : (i_1, \ldots, i_{q_n}) \in I_n\}\). Clearly,
\[
\text{Card } M_{n+1} = \text{Card } I_n = 4^{q_n} = m_{n+1}
\]
in accordance with Definition 3.4. We remark that Definition 3.2(iii) implies that \(M_n \subseteq \Omega_{t_n}\) for every \(n \geq 1\). The key features of the above construction are summarised in the following lemma.
Lemma 3.5. The following are direct consequences of the definition of the sets \( M_n \):

(i) If \( \prod_{j=1}^{k} c_{j}^{k} \in M_{n} \), where \( n \geq 2 \) and each \( c_{j}^{k} \in C_{n-1}^{k} \), then \( c_{j}^{k} \in C_{n-1}^{1} \) if and only if \( k = 1 \).

(ii) If \( \omega_{1} = c_{j}^{k_{1}} \cdot c_{j}^{k_{2}} \) with \( c_{j}^{k_{1}} \in C_{n}^{k_{1}}, c_{j}^{k_{2}} \in C_{n}^{k_{2}} \), then there exists \( N > n \) and \( \omega_{2} \in M_{N} \) such that \( \omega_{1} \prec \omega_{2} \) if and only if \( k_{2} = 1 + k_{1} \) mod \( q_{n} \).

(iii) If \( \omega_{1} = \prod_{k=0}^{t_{n}} c_{j}^{k} \) where each \( c_{j}^{k} \in C_{n}^{k} \), and \( k_{0} \leq q_{n} - t_{n} \), then \( \omega_{1} \neq \omega_{2} \) for all \( \omega_{2} \in M_{N} \), for all \( N \geq n + 1 \).

(iv) If \( \omega_{1} \prec \omega_{2} \in M_{n+1} \), \( \omega_{1} \prec \omega_{3} \in M_{n+1} \) and \( \omega_{1} \) has length at least \( t_{n} \), then \( \omega_{2} = \omega_{3} \).

(v) For each \( n \geq 1 \) we have

\[ \ell_{n+1} = q_{n} \ell_{n} = \frac{1}{4} m_{n} \ell_{n} \]

and hence

\[ \ell_{n+1} = \ell_{1} \prod_{k=1}^{n} q_{k} \cdot \]

The proof is clear.

The following lemma allows us to pass from the relation \( \prec \) to the relation \( \prec_{n} \), and thus make use of Lemma 3.5.

Lemma 3.6. Let \( a = (a_{i})_{i=1}^{\ell_{n}} \in M_{n} \) and \( \omega = (\omega_{i})_{i=1}^{\ell_{N}} \in M_{N} \) where \( n < N \). Suppose that there is \( r \geq 0 \) such that \( a_{i} = \omega_{r+i} \) for all \( 1 \leq i \leq \ell_{n} \). Then \( \ell_{n} \mid r \).

Proof. We first prove the case \( n = 1 \). Let \( \omega = \prod_{k=1}^{m} \omega^{k} \) where each \( \omega^{k} \in M_{1} \), and suppose \( a_{i} = \omega_{r+i} \) for all \( 1 \leq i \leq \ell_{1} \). There exists \( k^{*} \) such that \( k^{*} \ell_{1} \leq r + i < (k^{*} + 2) \ell_{1} \) for all \( 1 \leq i \leq \ell_{1} \), so that if we let \( \tilde{\omega} = \omega^{k^{*}} \cdot \omega^{k^{*}+1} \) then \( z_{i} = a_{i} = \tilde{\omega}_{i+r-k} \) for all \( 1 \leq i \leq P \leq \ell_{1} \) by Definition 3.2 (i). By Definition 3.2 (ii) we have either \( r - k^{*} \ell_{1} = \ell_{1} \) or \( r - k^{*} \ell_{1} = 0 \) and so \( \ell_{1} \mid r \) as required.

We proceed by induction on \( n \). Let \( a = \prod_{k=1}^{m} a^{k} \) and \( \omega = \prod_{k=1}^{m} \tilde{\omega}^{k} \) with each \( \tilde{\omega}^{k} \in C_{n-1}^{k} \). Let \( \tilde{a}^{1} = (\tilde{a}^{1})_{i=1}^{\ell_{n-1}} \). Since \( \tilde{a}^{1}_{i} = \omega_{r+i} \) for all \( 1 \leq i \leq \ell_{n-1} \), we have \( \ell_{n-1} \mid r \) by the induction hypothesis. Since each \( \tilde{\omega}^{k} \) has length \( \ell_{n-1} \), it follows that there is \( s = r/\ell_{n-1} \) such that \( \prod_{k=1}^{m} a^{k} = \prod_{k=s}^{m} \tilde{\omega}^{k} \). By Lemma 3.5 (i) we have \( \tilde{a}^{k} \in C_{n-1}^{k} \) if and only if \( k = 1 \), and similarly \( \tilde{\omega}^{k} \in C_{n-1}^{k} \) if and only if \( k \equiv 1 \) mod \( q_{n} \). Since \( \tilde{\omega}^{r+1} \in C_{n-1}^{r+1} \) it follows that \( s \equiv 0 \) mod \( q_{n} \). Since \( r = s \ell_{n-1} \) and \( \ell_{n} = q_{n} - \ell_{n-1} \), we deduce that \( \ell_{n} \mid r \).

Henceforth, we shall say that a finite word \( \omega \) is admissible if and only if it occurs as a subword of some \( \omega_{n} \in M_{n} \) for some \( n \) (hence for all sufficiently large \( n \)). We deduce:

Corollary 3.7. Let \( \omega = (\omega_{i})_{i=1}^{p} \) be admissible. Let \( a, b \in M_{n} \), and write \( a = (a_{i})_{i=1}^{\ell_{n}}, b = (b_{i})_{i=1}^{\ell_{n}} \). If there are \( r, s > 0 \) such that \( a_{i} = \omega_{i+r}, b_{i} = \omega_{i+s} \) for all \( 1 \leq i \leq \ell_{n} \), then \( \ell_{n} \mid r - s \).

We may now prove:

Lemma 3.8. The following constraints on admissibility hold:

(i) If \( \omega = c_{j}^{k} \cdot c_{j}^{k} \) with \( c_{j}^{k} \in C_{n}^{k}, c_{j}^{k} \in C_{n}^{k} \), then \( \omega \) is admissible if and only if \( k_{2} = 1 + k_{1} \) mod \( q_{n} \).
(ii) If \( \omega = \prod_{k=k_0}^{k_0+t_n} c_j^k \) with each \( c_j^k \in C_n \), and \( j_{k_0} \neq j_{k_0+t_n} \), and \( k_0 \leq q_n-t_n \), then \( \omega \) is not admissible.

**Proof.** Apply Lemma 3.5 and Lemma 3.6.

**Lemma 3.9.** Let \( a, b \in M_n \) where \( n \geq 2 \), and suppose that the word \( a \cdot b \) is inadmissible. Then \( a \cdot b \) has an inadmissible subword of length less than or equal to \( 2t_{n-1} \ell_{n-1} \).

**Proof.** Let \( a = \prod_{k=q_n-1}^{q_n-2} a_k \), \( b = \prod_{k=q_n-1}^{q_n-2} b_k \) where each \( a_k, b_k \in M_{n-1} \). Define \( u = \prod_{k=q_n-1-t_{n-1}+1}^{q_n-2} a_k \) and \( v = \prod_{k=q_n-1}^{q_n-2} b_k \), each being an admissible word of length \( t_{n-1} \ell_{n-1} \). Clearly \( u \cdot v < a \cdot b \). Suppose, for a contradiction, that \( u \cdot v < \omega_1 \cdot \omega_2 = \omega \), say, where \( \omega_1, \omega_2 \in M_n \) and \( \omega \) is admissible. Let \( (u \cdot v)_i = \omega_{i+r} \) for all \( 1 \leq i \leq 2\ell_{n-1} \). By Lemma 3.6, \( r = \ell_{n-1} \) for some \( s \). Suppose that \( s < q_n \); then \( v < \omega_{n-1} \omega_2 \) and so \( \omega_2 = b \) by Lemma 3.5(iv). Thus \( s = q_n \) and so \( u < \omega_{n-1} \omega_1 \) and \( \omega_1 = a \). Thus \( \omega = a \cdot b \) and so is not admissible. It follows that \( u \cdot v \) is inadmissible; this word has length \( 2t_{n-1} \ell_{n-1} \) as required.

**Lemma 3.10.** Let \( \omega \) be a word of length \( \ell \geq \ell_n \), where \( n \geq 2 \). Then either \( \omega \approx \prod_{k=1}^{m} \omega^k \) with each \( \omega^k \in C_{n-1} \), or there exists \( \omega' \approx \omega \) which has an inadmissible subword of length less than or equal to \( 3\ell_{n-1} \).

**Proof.** Suppose that for all \( \omega' \approx \omega \), every subword of \( \omega' \) of length \( 3\ell_{n-1} \) is admissible. Let \( \ell = s\ell_{n-1} + r \) with \( 0 \leq r < \ell_{n-1} \). We claim that there exist \( \omega^k \in M_{n-1} \) and a word \( \omega^s \) of length \( r \) such that \( \omega \approx \prod_{k=1}^{m} \omega^k \cdot \omega^s \). We will show that \( (\omega_i)_{i=1}^{m\ell_{n-1}} \approx \prod_{k=1}^{m} \omega^k \) for all \( m \leq s \) by induction on \( m \).

Clearly, any admissible word of length \( 3\ell_{n-1} \) must include some \( a \in M_{n-1} \) as a subword, and so \( \omega \) must include such a subword. Taking a rotation equivalent of \( \omega \) if necessary, we deduce that there exists \( \omega^1 \in M_{n-1} \) such that \( \omega_i = \omega_i^1 \) for all \( 1 \leq i \leq \ell_{n-1} \). This proves the case \( m = 1 \).

Given that \( (\omega_i)_{i=1}^{m\ell_{n-1}} = \prod_{k=1}^{m} \omega^k \) with each \( \omega^k \in M_{n-1} \), consider the word \( b = (b_i)_{i=1}^{2\ell_{n-1}} \) defined by \( b_i = \omega_{(m-1)\ell_{n-1}+i} \) for all \( 1 \leq i \leq 2\ell_{n-1} \), which is well-defined as long as \( m+1 \leq s \). Since \( b \) has length \( 2\ell_{n-1} < 3\ell_{n-1} \) it is admissible, and so there exist \( N, \ell \), \( t > 0 \) and \( a = \prod_{k=1}^{M} a_k \in M_N \) with each \( a_k \in M_{n-1} \) such that \( b_i = a_{i+t} \) for all \( 1 \leq i \leq 2\ell_{n-1} \). By Lemma 3.6 we have \( t = \ell_{n-1} \ell \) for some \( \ell \), so that \( b = a^{m+1} \cdot a^{r+2} \). Thus \( (\omega_i)_{i=m\ell_{n-1}+1}^{(m+1)\ell_{n-1}} = (b_i)_{i=m\ell_{n-1}+1}^{(m+1)\ell_{n-1}} = a^{r} \in M_{n-1} \) as required to prove the case \( m+1 \). This completes the induction step and proves our claim.

Since for every \( \omega' \approx \omega \), every subword of \( \omega' \) with length less than or equal to \( 3\ell_{n-1} \) is admissible, the word \( z = \omega^s \cdot \omega^s \cdot \omega^1 \) must be admissible. By Corollary 3.7 this implies \( \ell_{n-1} \mid r \) and hence \( r = 0 \). This completes the proof.

**Lemma 3.11.** Let \( \omega = \prod_{k=1}^{m} \omega^k \) for some \( \omega^1, \ldots, \omega^m \in M_n \) with \( m \geq 1 \). Then at least one of the following holds:

(a) There exists \( \omega' \approx \omega \) such that \( \omega' \) has an inadmissible subword with length less than or equal to \( 2t_{n-1} \ell_{n-1} \), or
(b) there exists \( \omega' \approx \omega \) such that \( \omega' \) has an inadmissible subword with length less than or equal to \( (1 + t_n)\ell_n \), and \( m \geq q_n \), or
(c) there exist \( \tilde{m} > 0 \) and \( \tilde{\omega}^1, \ldots, \tilde{\omega}^{\tilde{m}} \in M_{n+1} \) such that \( m = \ell_{n+1} \tilde{m} \) and \( \omega \approx \prod_{k=1}^{\tilde{m}} \tilde{\omega}^k \).
Proof. Suppose that the case (a) does not hold. Writing \( \omega = \prod_{i=1}^{m} c_{j_i}^{k_i} \) with every \( c_{j_i}^{k_i} \in C_{n_i}^{k_i} \), this assumption implies via Lemma 3.8(i) and Lemma 3.9 that for every \( i < m \) we have \( k_i = 1 + k_i \mod q_n \). Clearly, there exists a rotation equivalent of \( \omega \) which includes the word \( c_{j_m}^{k_m} \cdot c_{j_i}^{k_i} \). Again by Lemma 3.8(i) and Lemma 3.9, our assumption that case (a) does not hold implies that this word is admissible, and hence \( k_1 = k_m + 1 \mod q_n \). It follows that the sequence \( k_i \) must take every value in the range \( 1, \ldots, q_n \) an equal number of times. Taking a rotation equivalent if necessary, we have

\[
\omega = \prod_{s=1}^{S} \prod_{k=1}^{q_n} c_{j_{k_s,s}}^{k_s},
\]

for some integer \( S \geq 1 \) and some sequence \( j_{k,s} \). Note in particular that \( m \geq q_n \) (as required for (b) to hold).

We now suppose also that (b) does not hold. To show that \( \omega \) is a concatenation of elements of \( M_{n+1} \), and hence that (c) holds, it remains to show that \( j_{k,s} = j_{k+t_s,s} \) for all \( 1 \leq s \leq S \) and \( 1 \leq k \leq q_n - t_n \). If this is not the case, then \( \omega \) must include a subword of the form

\[
\prod_{k=k_0+t_n}^{k_0+t_n} c_{j_{k+s,s}}^{k_s}
\]

for some fixed \( s \) such that \( j_{k_0,s} \neq j_{k_0+t_n,s} \), which is inadmissible by Lemma 3.8(ii) and has length \((1 + t_n)\ell_n\), implying the case (b). □

Combining the above lemmata, we obtain:

**Proposition 3.12.** Define a set \( K \subseteq \Sigma_A \) by letting \( x \in K \) if and only if every subword of \( x \) is admissible. Then \( K \) is closed, satisfies \( \sigma K = K \), and is nonempty. Moreover,

\[
\sup_{\omega \in \Omega_{\ell_n}} \xi(\omega, K) = o(p^\varepsilon),
\]

for every \( \varepsilon > 0 \).

Proof. That \( K \) is a nonempty subset of \( \Sigma_A \) follows from the fact that \( M_n \subseteq \Omega_{\ell_n} \) for every \( n \geq 1 \); the proof that \( K \) is closed and \( \sigma K = K \) is straightforward. Given \( \omega \in \Omega_{\ell_n} \) and \( \ell_n \leq p < \ell_{n+1} \), we will attempt to bound the quantity

\[
\xi(\omega, K) = \inf_{\omega' \geq \omega} \{ \ell > 0 : \exists a \in \Omega_\ell \text{ such that } a \prec \omega' \text{ and } a \text{ is not admissible} \}.
\]

Suppose that \( \xi(\omega, K) > 3\ell_{n-1} \). Then by Lemma 3.10 we have \( \omega \approx \prod_{i=1}^{m} c_{j_i}^{k_i} \) where each \( c_{j_i}^{k_i} \in C_{n_i-1}^{k_i} \) and \( m = p/\ell_{n-1} \). Lemma 3.11 then implies that either \( \xi(\omega, K) \leq 2\ell_{n-2}t_{n-2} \), or \( \xi(\omega, K) \leq \ell_{n-1}(1 + t_{n-1}) \), or \( \omega \approx \prod_{i=1}^{m} c_{j_i}^{k_i} \) where each \( c_{j_i}^{k_i} \in C_{n_i}^{k_i} \) and \( \hat{m} = p/\ell_n \). In the last of these three cases, we apply Lemma 3.11 again to see that, since \( p < \ell_{n+1} \), either \( \xi(\omega, K) \leq 2\ell_{n-1}t_{n-1} \), or \( p \geq \ell_n q_n \) and \( \omega \approx \prod_{i=1}^{m} c_{j_i}^{k_i} \). In either case we conclude that in any case

\[
\xi(\omega, K) \leq \max \left\{ 3\ell_{n-1}, 2\ell_{n-2}t_{n-2}, \ell_{n-1}(t_{n-1} + 1), 2\ell_{n-1}t_{n-1}, \frac{2pt_n}{q_n} \right\}
\]

\[
= \max \left\{ 2\ell_{n-1}t_{n-1}, \frac{2pt_n}{q_n} \right\}.
\]
Hence for every $n \geq 3$ and $\varepsilon > 0$,

$$\sup_{\ell_n \leq k < \ell_{n+1}} \frac{1}{p^q} \frac{\mathcal{L}(\omega, K)}{p^q} \leq \max \left\{ 2\ell_{n-1} - \varepsilon, 2 \frac{\ell_{n-1}^2}{q_n} \right\} \leq 2 \max \{ \ell_{n-1}^2 q_n^{-\varepsilon}, \ell_{n-1} q_n^{-\varepsilon} \} = 2 \ell_{n-1} q_n^{-\varepsilon}.$$

To complete the proof, therefore, it suffices to show that $t_n q_n^{-\varepsilon} \ell_n \to 0$ for every $\varepsilon > 0$. Using Lemma 3.5(vii), Lemma 3.4 and Definition 3.3, we have

$$\log_4 \left( t_n q_n^{-\varepsilon} \ell_n \right) \leq \log_4 \left( \frac{t_{n-1}^{1/4} q_n^{-\varepsilon} \ell_1 \prod_{k=2}^{n-1} q_k}{4 \varepsilon t_{n-1} + \sum_{k=1}^{n-2} t_k} \right)$$

$$= -\varepsilon t_{n-1} + \sum_{k=1}^{n-2} t_k + O \left( t_{n-1}^{1/4} \right).$$

Elementary analysis then shows that since $t_{n-1} / t_n$ tends to zero, $\sum_{k=2}^{n-2} t_k = o(t_{n-1})$ and so

$$\log_4 \left( t_n q_n^{-\varepsilon} \ell_n \right) \to -\infty$$

as required. \hfill \Box

REFERENCES


D. COLLIER, SCHOOL OF MATHEMATICS, UNIVERSITY OF MANCHESTER, MANCHESTER M13 9PL, U.K.

I. D. MORRIS, MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, U.K.