Proceedings of the Edinburgh Mathematical Society (2002) **45**, 653–671 ©
DOI:10.1017/S001309150000119X Printed in the United Kingdom

CONTINUED FRACTIONS WITH BOUNDED PARTIAL QUOTIENTS

J. L. DAVISON

Department of Mathematics and Computer Science, Laurentian University, Sudbury, Ontario, Canada P3E 2C6 (les@cs.laurentian.ca)

(Received 28 November 2000)

Abstract Precise bounds are given for the quantity

$$L(\alpha) = \frac{\limsup_{m\to\infty}(1/m)\ln q_m}{\liminf_{m\to\infty}(1/m)\ln q_m},$$

where (q_m) is the classical sequence of denominators of convergents to the continued fraction $\alpha = [0, u_1, u_2, \dots]$ and (u_m) is assumed bounded, with a distribution.

If the infinite word $\boldsymbol{u} = u_1 u_2 \dots$ has arbitrarily large instances of segment repetition at or near the beginning of the word, then we quantify this property by means of a number γ, called the segment-repetition factor.

If α is not a quadratic irrational, then we produce a specific sequence of quadratic irrational approximations to α, the rate of convergence given in terms of L and γ. As an application, we demonstrate the transcendence of some continued fractions, a typical one being of the form $[0, u_1, u_2, \dots]$ with $u_m = 1 + \lfloor m\theta \rfloor \bmod n$, $n \geqslant 2$, and θ an irrational number which satisfies any of a given set of conditions.

Keywords: transcendence; approximation; distribution

AMS 2000 *Mathematics subject classification:* Primary 11A55
Secondary 11B37

1. Introduction

Suppose Σ is a finite set of positive integers. If $(u_m)_{m\geqslant 1}$ is an infinite sequence with $u_m \in \Sigma$ for each $m \geqslant 1$, then we let $\boldsymbol{u}$ be the infinite word $u_1 u_2 \dots$ and say that $\boldsymbol{u}$ takes its values from Σ. Suppose the continued fraction $\alpha = [0, u_1, u_2, \dots]$ has convergents

$$\left(\frac{p_m}{q_m}\right)_{m\geqslant 0}.$$

Then we define

$$L(\alpha) = \frac{\limsup_{m\to\infty}(1/m)\ln q_m}{\liminf_{m\to\infty}(1/m)\ln q_m}$$

and set $L(\boldsymbol{u}) = L(\alpha)$.

It is evident that $L(\alpha) \geqslant 1$ and that $L(\alpha) = 1$ if and only if $\lim_{m\to\infty}(1/m)\ln q_m$ exists. It is of interest to be able to provide precise upper estimates for $L(\alpha)$. Of course, if we only know that $a \leqslant u_m \leqslant b$ for all $m \geqslant 1$, then it is easy to see that

$$L(\alpha) \leqslant \ln(\tfrac{1}{2}(b + \sqrt{b^2 + 4}))/\ln(\tfrac{1}{2}(a + \sqrt{a^2 + 4}))$$

and that this bound can be attained.

If we assume that $|\Sigma| = 2$ and that $\boldsymbol{u}$ takes each of the two values with a frequency (see Definition 3.1), then the authors of [**1**] have shown that $L(\boldsymbol{u}) < 1.13$, irrespective of the particular elements of Σ. In §§ 2 and 3, we extend this result to any finite set, on the assumption that $\boldsymbol{u}$ takes each of the values in Σ with a frequency. In § 4 we give a more precise estimate, provided that $\boldsymbol{u}$ is uniformly distributed (that is, each value is taken with the same frequency).

Our main goal in this paper is to prove transcendence of a certain family of continued fractions. For this, we also need to consider the property that the infinite word $\boldsymbol{u}$ has arbitrarily large instances of segment repetition near the beginning of $\boldsymbol{u}$. A special case of this concept was discussed in [**1**]. More formally we make the following definition.

Definition 1.1. Suppose $\gamma \in \mathbb{R}$ with $\gamma \geqslant 1$. The infinite word $\boldsymbol{w} = w_1 w_2 w_3 \dots$ is said to have a *segment expansion factor* greater than or equal to γ if there exist three infinite sequences of finite words $\{U_k\}_{k\geqslant1}$, $\{V_k\}_{k\geqslant1}$, $\{W_k\}_{k\geqslant1}$ which satisfy all the following conditions.

(1) $U_k V_k W_k$ is a prefix of $\boldsymbol{w}$.

(2) $\lim_{k\to\infty} |V_k| = \infty$, where $|V_k|$ is the length of V_k.

(3) W_k is a prefix of V_k^s for some positive integer s.

(4) $\displaystyle\liminf_{k\to\infty} \frac{|U_k V_k W_k|}{|U_k| + |U_k V_k|} = \gamma.$

Finally, we will say $\boldsymbol{w}$ has a *prefix expansion factor* greater than or equal to γ if we can take $U_k = \lambda$ for all $k \geqslant 1$.

In § 5 we provide explicit computations of the segment expansion factor for the infinite word $\boldsymbol{w} = w_1 w_2 \dots$, where $w_m = \lfloor m\theta \rfloor \bmod n$ and θ is an irrational with $0 < \theta < 1$.

In the final section of the paper we first prove that if α is not a quadratic irrational (that is, $\boldsymbol{u}$ is not ultimately periodic) and if $\boldsymbol{u}$ has a segment expansion factor greater than or equal to γ, then there is a sequence of quadratic irrationals (α_k) which satisfy

$$|\alpha - \alpha_k| < \frac{1}{H(\alpha_k)^{2\gamma/L(\alpha)}},$$

where $H(\alpha_k)$ denotes the height of α_k.

With the aid of Schmidt's Theorem [**9**], we then obtain a transcendence result for a special class of continued fractions derived from the words studied in § 5.

2. Basic terminology and the trace inequality

Let $\Sigma = \{a_1, a_2, \dots, a_n\}$ be a finite set of $n \geqslant 2$ positive integers, ordered so that $1 \leqslant a_1 < a_2 < \cdots < a_n$. Let $\boldsymbol{u} = (u_m)_{m\geqslant 1} \in \Sigma^{\mathbb{N}}$ be any infinite sequence with values in Σ.

Consider the sequence $(q_m)_{m\geqslant -1}$ defined by

$$q_{-1} = 0, \qquad q_0 = 1, \qquad q_m = u_m q_{m-1} + q_{m-2} \quad \text{for } m \geqslant 1. \tag{2.1}$$

The sequence $(q_m)_{m\geqslant -1}$ so defined is the sequence of denominators of the convergents to the continued fraction $[0, u_1, u_2, \dots]$. Readers can consult [**6**] or [**8**] for information on standard continued fraction theory. We will say $\boldsymbol{u}$ *generates the sequence* (q_m). The statement (2.1) can be expressed in matrix form by

$$\begin{bmatrix} q_0 \\ q_{-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} q_m \\ q_{m-1} \end{bmatrix} = \begin{bmatrix} u_m & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_{m-1} \\ q_{m-2} \end{bmatrix} \quad \text{for } m \geqslant 1.$$

If we write

$$A_i = \begin{bmatrix} a_i & 1 \\ 1 & 0 \end{bmatrix} \quad \text{for } 1 \leqslant i \leqslant n,$$

then it can be shown that the semigroup $S_n = S(A_1, A_2, \dots, A_n)$ generated by the matrices $A_1, A_2, \dots, A_n$ is free, so we can identify the matrices in S_n with the corresponding words (strings) in the symbols $A_1, A_2, \dots, A_n$. The length of such a word W, denoted by $|W|$, is the number of symbols (counting repetitions) that occur in W. If $\mathcal{S}_n^-$ denotes the set of those matrices in S_n with determinant equal to -1, then $W \in \mathcal{S}_n^-$ if and only if $|W|$ is odd. The *trace of* W, denoted by $\operatorname{tr}(W)$, is the trace of the matrix W.

We can write the preceding matrix recurrence in the form

$$\begin{bmatrix} q_m \\ q_{m-1} \end{bmatrix} = W_m \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{2.2}$$

where $W_m(A_1, A_2, \dots, A_n)$ is a word of length m in the matrices $A_1, A_2, \dots, A_n$. We will say that W_m is associated with q_m.

If $\rho(M)$ denotes the spectral radius of the real square matrix M, then the L^2-norm of M, written as $\|M\|$, equals $\sqrt{\rho(M^{\mathrm{t}}M)}$. In particular, $\|A_i\| = \rho(A_i) = \frac{1}{2}(a_i + \sqrt{a_i^2 + 4})$.

The first result, proved in [**1**], shows the connection between q_m and W_m.

Proposition 2.1. *The following inequalities hold:*

(a) $q_m \leqslant \|W_m\|$; *and*

(b) $q_m \geqslant \frac{1}{2} \operatorname{tr} W_m$.

In order to proceed further, it is thus essential to consider the trace of words in $\mathcal{S}_n$. As much of the first part is easily derivable from the $n = 2$ case described in some detail in [**1**], we will be brief in our exposition.

If

$$X = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in \mathcal{S}_n,$$

let Φ_X be the map $x \to ((\delta x + \gamma)/(\beta x + \alpha))$. Then $\Phi_{MN} = \Phi_M \circ \Phi_N$ and Φ_X has two fixed points x_X, y_X with $x_X < y_X$. For $1 \leqslant i, j \leqslant n$ put

$$x_{ij} = x_{A_iA_j} = \tfrac{1}{2}\Big(-a_j - \sqrt{a_j^2 + 4a_j/a_i}\Big) \quad \text{and} \quad y_{ij} = y_{A_iA_j} = \tfrac{1}{2}\Big(-a_j + \sqrt{a_j^2 + 4a_j/a_i}\Big).$$

Note that $x_{ii} = x_{A_i}$ and $y_{ii} = y_{A_i}$ for $1 \leqslant i \leqslant n$.

Lemma 2.2. *For $1 \leqslant i, j \leqslant n$ the following hold:*

(a) $(1/(x_{ij} + a_j)) = x_{ji}$;

(b) $y_{ij}x_{ji} = -1$; *and*

(c) $x_{ij} > -(a_j + (1/a_i))$.

Lemma 2.3. *The fixed points (x_{ij}), (y_{ij}) are totally ordered as follows:*

(a) $-a_j - 1 < x_{1j} < x_{2j} < \cdots < x_{nj} < -a_j$, $1 \leqslant j \leqslant n$; *and*

(b) $(1/(a_i + 1)) < y_{i1} < y_{i2} < \cdots < y_{in} < (1/a_i)$, $1 \leqslant i \leqslant n$.

The proofs of these three lemmas are omitted.

Proposition 2.4. *We have*

$$\operatorname{tr}(A_1A_nX) \geqslant \rho(A_1A_n)\operatorname{tr}(X) \quad \textit{for any } X \in \mathcal{S}_n^-(A_1, A_2, \dots, A_n).$$

Proof. As in [**1**], it suffices to show that (i) $\beta x_{1n} + \alpha < 0$, and (ii) $x_{n1} \leqslant \Phi_X(x_{1n}) \leqslant y_{1n}$, where

$$X = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}.$$

These two statements are proved by induction on the (odd) length of X. Lemmas 2.2 and 2.3 give the basis case ($|X| = 1$) and, if we let $\mathcal{U}_n = \{A_iA_j;\ 1 \leqslant i,\ j \leqslant n\}$, the inductive step hinges on the fact that

$$\min\left\{\frac{m_1}{m_2} : M = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \in \mathcal{U}_n\right\} = a_1 + \frac{1}{a_n}.$$

The details are left to the reader. □

So far the extension to more than two arguments is fairly direct. We now come to the proposition that allows us to take out an A_1A_n term when the word does not have an adjacent pair A_1, A_n.

Proposition 2.5. *Suppose $n \geqslant 3$, and let W^{R} denote the transposed (or reverse) string of W. Then $\operatorname{tr}(A_1 W A_n X) \geqslant \operatorname{tr}(A_1 A_n W^{\mathrm{R}} X)$ for all $W \in \mathcal{S}(A_2, \dots, A_{n-1})$ and for all $X \in \mathcal{S}(A_1, A_2, \dots, A_n)$ that do not start with A_1.*

Proof. Let

$$W = \begin{bmatrix} u & v \\ w & x \end{bmatrix}.$$

Then (noting that $a_2 \geqslant 2$) it is easy to establish by induction that

$$w < u < a_n w + x. \tag{2.3}$$

Also

$$W A_n - A_n W^{\mathrm{R}} = \begin{bmatrix} 0 & u - a_n w - x \\ a_n w + x - u & 0 \end{bmatrix},$$

which by (2.3) is of the form

$$\begin{bmatrix} 0 & -y_n \\ y_n & 0 \end{bmatrix}$$

with $y_n > 0$.

Since $\operatorname{tr}(A_1(W A_n - A_n W^{\mathrm{R}})X) = y_n(\alpha - a_1\gamma - \delta)$, it suffices to show that $\alpha \geqslant a_1\gamma + \delta$ for all X not starting with A_1. If $X = A_i$ for some i, $2 \leqslant i \leqslant n$, then $a_i \geqslant a_1 + 1$, so the result holds in this case. If $|X| \geqslant 2$, then we can write $X = A_i U A_j$, where $2 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$, and $U \in \mathcal{S}_n \cup \{I\}$. It is easy to see that $\alpha \geqslant a_1\gamma + \delta$ in this situation. □

3. Infinite words with frequency

As mentioned in § 1, we must impose some condition on $\boldsymbol{u} = (u_m)_{m \geqslant 1}$ in order to expect a better estimate for $L(\boldsymbol{u})$. It turns out that a natural condition to impose is that each $a_i \in \Sigma$ occurs in $\boldsymbol{u}$ with a frequency α_i. Specifically, we make the following definition.

Definition 3.1. For $1 \leqslant i \leqslant n$ put $\alpha_i(m) = |\{1 \leqslant k \leqslant m : u_k = a_i\}|$. If for each i, $\lim_{m\to\infty}(\alpha_i(m)/m)$ exists, say equal to α_i, then we say that $\boldsymbol{u}$ is a word with frequency

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n).$$

Let us write $\boldsymbol{A} = (A_1, A_2, \dots, A_n)$, where, as usual, A_i denotes the matrix

$$\begin{bmatrix} a_i & 1 \\ 1 & 0 \end{bmatrix}.$$

Furthermore, put $M(\boldsymbol{A}, \boldsymbol{\alpha}) = \sum_{i=1}^{n} \alpha_i \ln \rho(A_i)$. Then we have the following proposition (cf. [**1**, **3**]).

Proposition 3.2. *Suppose $\boldsymbol{u}$ is an infinite word with values taken from Σ with frequency $\boldsymbol{\alpha}$ and suppose $\boldsymbol{u}$ generates (q_m). Then $\limsup_{m\to\infty}(1/m)\ln q_m \leqslant M(\boldsymbol{A}, \boldsymbol{\alpha})$.*

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

Proof. Using our previously introduced notation we have

$$\|W_m\| \leqslant \rho(A_1)^{\alpha_1(m)}\rho(A_2)^{\alpha_2(m)}\dots\rho(A_n)^{\alpha_n(m)}.$$

By Proposition 2.1 it follows that

$$\limsup_{m\to\infty}\left(\frac{1}{m}\right)\ln q_m \leqslant \limsup_{m\to\infty}\frac{\ln\|W_m\|}{m} \leqslant \sum_{i=1}^{n}\alpha_i \ln\rho(A_i) = M(\boldsymbol{A},\boldsymbol{\alpha}).$$

□

The task ahead will be to construct a piecewise linear (in $\boldsymbol{\alpha}$) function $H(\boldsymbol{A},\boldsymbol{\alpha})$ that satisfies the inequality $\liminf_{m\to\infty}(1/m)\ln q_m \geqslant H(\boldsymbol{A},\boldsymbol{\alpha})$.

First we note the following lemma from Proposition 2.1.

Lemma 3.3. $\liminf_{m\to\infty}\left(\frac{1}{m}\right)\ln q_m \geqslant \liminf_{\substack{m\to\infty\\ m\ \text{odd}}}\left(\frac{1}{m}\right)\ln \operatorname{tr} W_m.$

Proof. (q_m) is an increasing sequence so that

$$\liminf_{m\to\infty}\left(\frac{1}{m}\right)\ln q_m = \liminf_{\substack{m\to\infty\\ m\ \text{odd}}}\left(\frac{1}{m}\right)\ln q_m \geqslant \liminf_{\substack{m\to\infty\\ m\ \text{odd}}}\left(\frac{1}{m}\right)\ln \operatorname{tr} W_m.$$

□

We now use Proposition 2.4 and Proposition 2.5 to get the following theorem.

Theorem 3.4. *Let $W_m(A_1, A_2, \dots, A_n)$ be a word of odd length m.*

(a) *If $\alpha_1(m) \leqslant \alpha_n(m)$, then*

$$\operatorname{tr} W_m(A_1, A_2, \dots, A_n) \geqslant \rho(A_1A_n)^{\alpha_1(m)} \operatorname{tr} W_{m-2\alpha_1(m)}(A_2, A_3, \dots, A_n),$$

where the number of occurrences of A_i in $W_{m-2\alpha_1(m)}$ is $\alpha_i(m)$ for $2 \leqslant i \leqslant n-1$, and $\alpha_n(m) - \alpha_1(m)$ for $i = n$.

(b) *If $\alpha_n(m) \leqslant \alpha_1(m)$, then*

$$\operatorname{tr} W_m(A_1, A_2, \dots, A_n) \geqslant \rho(A_1A_n)^{\alpha_n(m)} \operatorname{tr} W_{m-2\alpha_n(m)}(A_1, A_2, \dots, A_{n-1}),$$

where the number of occurrences of A_i in $W_{m-2\alpha_n(m)}$ is $\alpha_1(m) - \alpha_n(m)$ for $i = 1$ and $\alpha_i(m)$ for $2 \leqslant i \leqslant n-1$.

Proof. If W_m has an adjacent A_1A_n or A_nA_1, we use Proposition 2.4, since tr is invariant under cyclic permutation and under transpose (or reverse), to obtain

$$\operatorname{tr} W_m \geqslant \rho(A_1A_n) \operatorname{tr} W_{m-2}(A_1, \dots, A_n).$$

We can continue to remove adjacent A_1A_n in such a manner until no adjacency remains. At that point we can use Proposition 2.5 instead, which will produce an adjacency between A_1, A_n for which Proposition 2.4 will again be used. Going back and forth in this manner, we will either exhaust the A_1s first ($\alpha_1(m) < \alpha_n(m)$) or the A_ns ($\alpha_n(m) < \alpha_1(m)$) or possibly exhaust them together ($\alpha_1(m) = \alpha_n(m)$). The desired inequalities are now clear. □

The effect of Theorem 3.4 is to reduce the number of variables in the argument by one, or possibly two. Consider for example a word $\boldsymbol{u}$ with frequency $\boldsymbol{\alpha}$, where $\alpha_1 < \alpha_n$. Then $\alpha_1(m) < \alpha_n(m)$ for all $m \geqslant m_0$, say, and also $m^* = m - 2\alpha_1(m) \to \infty$, since $\alpha_1 < \frac{1}{2}$. Thus

$$\liminf_{\substack{m\to\infty \\ m \text{ odd}}} \left(\frac{1}{m}\right) \ln \operatorname{tr} W_m \geqslant \alpha_1 \ln \rho(A_1 A_n) + (1-2\alpha_1) \liminf_{\substack{m^*\to\infty \\ m^* \text{ odd}}} \left(\frac{1}{m^*}\right) \ln \operatorname{tr} W_{m^*}, \tag{3.1}$$

and the frequencies of A_i in W_{m^*} are $\alpha_i/(1-2\alpha_1)$ for $2 \leqslant i \leqslant n-1$ and $(\alpha_n - \alpha_1)/(1-2\alpha_1)$ for $i = n$.

The right-hand side of equation (3.1) has the makings of part of the recursive definition of $H(\boldsymbol{A}, \boldsymbol{\alpha})$, but we must be careful about the relationship between the different $\boldsymbol{\alpha}$s. We therefore consider the various domains of definition of the proposed $H(\boldsymbol{A}, \boldsymbol{\alpha})$.

For $n \geqslant 1$ let us put

$$D_n = \left\{ \boldsymbol{\alpha} \in R^n : \alpha_i \geqslant 0 \text{ for } 1 \leqslant i \leqslant n \text{ and } \sum_{i=1}^{n} \alpha_i = 1 \right\}.$$

Note, in particular, that $D_1 = \{1\}$. For $n \geqslant 2$ we also need the following special points in D_n:

(a) for $1 \leqslant i \leqslant n$, put $\boldsymbol{e}_i = (\alpha_k)$, where $\alpha_k = 1$ if $k = i$ and 0 otherwise; and

(b) for $1 \leqslant i < j \leqslant n$, put $\boldsymbol{f}_{ij} = \frac{1}{2}(\boldsymbol{e}_i + \boldsymbol{e}_j)$.

We are going to break up the set D_n into 2^{n-1} subdomains, which we can usefully parametrize using binary strings of length $(n-1)$.

Definition 3.5.

(a) If λ denotes the empty string, put $\Delta_\lambda = \{1\}$.

(b) Let $n \geqslant 2$ and suppose $\Delta_{\boldsymbol{b}}$ has been defined for all binary strings $\boldsymbol{b}$ of length $(n-2)$. Then we set

$$\Delta_{1\boldsymbol{b}} = \{\boldsymbol{\alpha} \in D_n : \alpha_1 \leqslant \alpha_n \text{ and } (\alpha_2, \alpha_3, \dots, \alpha_{n-1}, \alpha_n - \alpha_1) \in (1-2\alpha_1)\Delta_{\boldsymbol{b}}\}$$

and

$$\Delta_{0\boldsymbol{b}} = \{\boldsymbol{\alpha} \in D_n : \alpha_n \leqslant \alpha_1 \text{ and } (\alpha_1 - \alpha_n, \alpha_2, \dots, \alpha_{n-1}) \in (1-2\alpha_n)\Delta_{\boldsymbol{b}}\}.$$

Proposition 3.6. *Let $n \geqslant 1$. Then*

(a) $D_n = \bigcup \Delta_{\boldsymbol{b}}$, *taken over all binary strings $\boldsymbol{b}$ of length $(n-1)$; and*

(b) *if $\boldsymbol{b}$ is a binary string of length $(n-1)$, then $\Delta_{\boldsymbol{b}}$ is a simplex with vertices in the set $\{\boldsymbol{f}_{ij} : 1 \leqslant i < j \leqslant n\} \bigcup \{\boldsymbol{e}_i : 1 \leqslant i \leqslant n\}$.*

Proof. Both parts are proved by induction. □

Note that if $\alpha \in \Delta_{1\boldsymbol{b}}$, then we can write

$$\boldsymbol{\alpha} = 2\alpha_1 \boldsymbol{f}_{1n} + (1 - 2\alpha_1)(0, \boldsymbol{\beta}) \quad \text{for some } \boldsymbol{\beta} \in \Delta_{\boldsymbol{b}}, \tag{3.2}$$

and if $\alpha \in \Delta_{0\boldsymbol{b}}$, then we have

$$\boldsymbol{\alpha} = 2\alpha_n \boldsymbol{f}_{1n} + (1 - 2\alpha_n)(\boldsymbol{\beta}, 0) \quad \text{for some } \boldsymbol{\beta} \in \Delta_{\boldsymbol{b}}. \tag{3.3}$$

These decompositions are unique if $\alpha_1 < \frac{1}{2}$ in (3.2) and $\alpha_n < \frac{1}{2}$ in (3.3). We are now ready to define H recursively on D_n.

Definition 3.7.

(a) If $n = 1$, then set $H(A, 1) = \ln \rho(A)$, for A a matrix of the form

$$\begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix}$$

with a a positive integer.

(b) Suppose $n \geqslant 2$ and assume $H(\boldsymbol{A}, \boldsymbol{\alpha})$ has been defined for $\boldsymbol{\alpha} \in D_{n-1}$. If $\boldsymbol{\alpha} \in D_n$, then by Proposition 3.6 either $\boldsymbol{\alpha} \in \Delta_{1\boldsymbol{b}}$ or $\boldsymbol{\alpha} \in \Delta_{0\boldsymbol{b}}$ for some $\boldsymbol{b}$ of length $n - 2$.

If $\boldsymbol{\alpha} \in \Delta_{1\boldsymbol{b}}$, then in light of (3.2) we set

$$H(\boldsymbol{A}, \boldsymbol{\alpha}) = \alpha_1 \ln \rho(A_1 A_n) + (1 - 2\alpha_1) H(\boldsymbol{B}, \boldsymbol{\beta}),$$

where $\boldsymbol{B} = (A_2, A_3, \dots, A_n)$.

If $\boldsymbol{\alpha} \in \Delta_{0\boldsymbol{b}}$, then using (3.3) we set

$$H(\boldsymbol{A}, \boldsymbol{\alpha}) = \alpha_n \ln \rho(A_1 A_n) + (1 - 2\alpha_n) H(\boldsymbol{C}, \boldsymbol{\beta}) \quad \text{where } \boldsymbol{C} = (A_1, A_2, \dots, A_{n-1}).$$

Proposition 3.8. *Let $n \geqslant 1$. Then*

(a) *H is a well-defined function on D_n; and*

(b) *H is linear in $\boldsymbol{\alpha}$ on $\Delta_{\boldsymbol{b}}$ for any $\boldsymbol{b}$. In other words, H is piecewise linear in $\boldsymbol{\alpha}$ on D_n.*

Proof.

(a) Easily checked (by induction).

(b) If, for example, $\boldsymbol{\alpha} \in \Delta_{1\boldsymbol{b}}$ and $\alpha_1 < \frac{1}{2}$, then (3.2) gives

$$\boldsymbol{\beta} = \frac{1}{1 - 2\alpha_1}(\alpha_2, \alpha_3, \dots, \alpha_{n-1}, \alpha_n - \alpha_1).$$

If $H(\boldsymbol{B}, \boldsymbol{\beta})$ is linear in $\boldsymbol{\beta}$ on $\Delta_{\boldsymbol{b}}$, then $H(\boldsymbol{A}, \boldsymbol{\alpha})$ is linear in $\boldsymbol{\alpha}$ on $\Delta_{1\boldsymbol{b}}$.

□

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

If we set $n = 2$ in Definition 3.7, we obtain

$$H(\boldsymbol{A}, \boldsymbol{\alpha}) = \begin{cases} \alpha_1 \ln \rho(A_1 A_2) + (\alpha_2 - \alpha_1) \ln \rho(A_2) & \text{if } \alpha_1 \leqslant \alpha_2, \\ \alpha_2 \ln \rho(A_1 A_2) + (\alpha_1 - \alpha_2) \ln \rho(A_1) & \text{if } \alpha_2 \leqslant \alpha_1. \end{cases}$$

We are now ready to prove the following theorem.

Theorem 3.9. *Let $\boldsymbol{u}$ be an infinite word with values taken from Σ with frequency $\boldsymbol{\alpha}$ and suppose $\boldsymbol{u}$ generates (q_m).*

$$\text{Then } \lim_{m\to\infty} \inf \left(\frac{1}{m}\right) \ln q_m \geqslant H(\boldsymbol{A}, \boldsymbol{\alpha}).$$

Proof. By Lemma 3.3 it suffices to show that

$$\liminf_{\substack{m\to\infty \\ m \text{ odd}}} \ln \operatorname{tr} W_m(A_1, A_2, \dots, A_n) \geqslant H(\boldsymbol{A}, \boldsymbol{\alpha}).$$

We will prove this by induction on n.

If $n = 1$, then in fact

$$\lim_{m\to\infty} \left(\frac{1}{m}\right) \ln \operatorname{tr}(A_1^m) = \ln \rho(A_1) = H(A_1, 1),$$

as can easily be checked by the reader (A_1 is diagonalizable).

Assume now that $n \geqslant 2$ and that the result has been established for the case of $(n-1)$ arguments (in both $\boldsymbol{A}$ and $\boldsymbol{\alpha}$).

We will consider four possibilities for $\boldsymbol{\alpha} \in D_n$.

Case 1 ($\boldsymbol{\alpha_1 < \alpha_n}$ (and so $\boldsymbol{\alpha_1 < \frac{1}{2}}$)). Equation (3.1) is then applicable, and it follows by induction that

$$\liminf_{\substack{m\to\infty \\ m \text{ odd}}} \left(\frac{1}{m}\right) \ln \operatorname{tr} W_m \geqslant \alpha_1 \ln \rho(A_1, A_n) + (1 - 2\alpha_1) H(\boldsymbol{B}, \boldsymbol{\beta}), \tag{3.4}$$

where, as before, $\boldsymbol{B} = (A_2, A_3, \dots, A_n)$ and

$$\boldsymbol{\beta} = \left(\frac{\alpha_2}{1 - 2\alpha_1}, \dots, \frac{\alpha_{n-1}}{1 - 2\alpha_1}, \frac{\alpha_n - \alpha_1}{1 - 2\alpha_1}\right).$$

But the right-hand side of (3.4) is $H(\boldsymbol{A}, \boldsymbol{\alpha})$, by Definition 3.7, and so we have established the required inequality.

Case 2 ($\boldsymbol{\alpha_n < \alpha_1}$ (and so $\boldsymbol{\alpha_n < \frac{1}{2}}$)). This case is similar to Case (1). We proceed from Theorem 3.4 (b) and use $\boldsymbol{C}$ instead of $\boldsymbol{B}$.

Case 3 ($\boldsymbol{\alpha_1 = \alpha_n < \frac{1}{2}}$). Let $I = \{m \in \mathbb{N} : \ m \text{ odd and } \alpha_1(m) \leqslant \alpha_n(m)\}$ and $J = \{m \in \mathbb{N} : \ m \text{ odd and } \alpha_n(m) < \alpha_1(m)\}$. At least one of I or J is infinite. If I is infinite, then equation (3.1) gives

$$\begin{aligned} \liminf_{m\in I} \left(\frac{1}{m}\right) \ln \operatorname{tr} W_m &\geqslant \alpha_1 \ln \rho(A_1 A_n) + (1 - 2\alpha_1) \liminf_{\substack{m\to\infty \\ m \text{ odd}}} \left(\frac{1}{m^*}\right) \ln \operatorname{tr} W_{m^*} \\ &\geqslant H(\boldsymbol{A}, \boldsymbol{\alpha}) \end{aligned}$$

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

by the inductive hypothesis. If also J is infinite, we use α_n instead and get

$$\liminf_{m\in J}\left(\frac{1}{m}\right)\ln\operatorname{tr}W_m \geqslant H(\boldsymbol{A},\boldsymbol{\alpha}),$$

and hence

$$\liminf_{\substack{m\to\infty\\ m\text{ odd}}}\left(\frac{1}{m}\right)\ln\operatorname{tr}W_m \geqslant H(\boldsymbol{A},\boldsymbol{\alpha}).$$

If one of I, J is finite, then we just have one $\liminf$ to consider.

Case 4 ($\boldsymbol{\alpha_1 = \alpha_n = \frac{1}{2}}$). Let $\epsilon > 0$. Then for $m \geqslant m_0(\epsilon)$ we have that

$$\frac{\alpha_1(m)}{m} > \tfrac{1}{2} - \epsilon \quad\text{and}\quad \frac{\alpha_n(m)}{m} > \tfrac{1}{2} - \epsilon.$$

By Theorem 3.4, where we replace the final trace by 1, we easily have $\operatorname{tr}W_m \geqslant \rho(A_1A_n)^{(m/2)-\epsilon m}$ and hence

$$\liminf_{\substack{m\to\infty\\ m\text{ odd}}}\left(\frac{1}{m}\right)\ln\operatorname{tr}W_m \geqslant \tfrac{1}{2}\ln\rho(A_1A_n) = H(\boldsymbol{A},\boldsymbol{f}_{1n}),$$

which is the required bound.

□

Let us now set $F(\boldsymbol{A},\boldsymbol{\alpha}) = (M(\boldsymbol{A},\boldsymbol{\alpha})/H(\boldsymbol{A},\boldsymbol{\alpha}))$. Then we have the following corollary.

Corollary 3.10. $L(\boldsymbol{u}) \leqslant F(\boldsymbol{A},\boldsymbol{\alpha})$.

Proof. Clear. □

Theorem 3.11.

$$L(\boldsymbol{u}) \leqslant \max_{1\leqslant i<j\leqslant n}\left\{\frac{\ln\rho(A_i)\rho(A_j)}{\ln\rho(A_iA_j)}\right\}.$$

Proof. $F(\boldsymbol{A},\boldsymbol{\alpha})$ is the piecewise quotient of two linear functions (in $\boldsymbol{\alpha}$) and hence attains its maximum at a vertex of one of the defining $\Delta_{\boldsymbol{b}}$ simplexes. By Proposition 3.6 the vertices of $\Delta_{\boldsymbol{b}}$ are elements of $S_n = \{\boldsymbol{f}_{ij},\ 1\leqslant i<j\leqslant n\}\bigcup\{\boldsymbol{e}_i : 1\leqslant i\leqslant n\}$. We then compute

$$F(\boldsymbol{A},\boldsymbol{e}_i) = \frac{\alpha_i\ln\rho(A_i)}{H(\boldsymbol{\alpha},\boldsymbol{e}_i)} = \frac{\alpha_i\ln\rho(A_i)}{\alpha_i\ln\rho(A_i)} = 1.$$

$$F(\boldsymbol{A},\boldsymbol{f}_{ij}) = \frac{\frac{1}{2}\ln\rho(A_i)+\frac{1}{2}\ln\rho(A_j)}{H(\boldsymbol{A},\boldsymbol{f}_{ij})} = \frac{\ln\rho(A_i)\rho(A_j)}{\ln\rho(A_iA_j)},$$

and so deduce that

$$\max_{\boldsymbol{\alpha}\in S_n} F(\boldsymbol{A},\boldsymbol{\alpha}) = \max_{1\leqslant i<j\leqslant n}\frac{\ln\rho(A_i)\rho(A_j)}{\ln\rho(A_iA_j)}.$$

Hence, by Corollary 3.10, the result follows. □

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

In [**1**] it was shown that

$$\max_{1\leqslant a_1<a_2}\left(\frac{\ln\rho(A_1)\rho(A_2)}{\ln\rho(A_1A_2)}\right)<1.129$$

and that the maximum is attained when $a_1 = 1$, $a_2 = 13$. Thus it follows from Theorem 3.11 that $L(\boldsymbol{u}) < 1.129$ for all infinite words with frequency and whose values come from a finite set of positive integers. For $n \geqslant 3$, we can make a slight improvement on the estimate for $L(\boldsymbol{u})$ if we can also assume that all symbols occur with the same frequency. We discuss this in our next section.

4. Infinite words with uniform distribution

If $|\Sigma| = n$ and the infinite word $\boldsymbol{u}$ has values from Σ with frequency $\alpha_i = (1/n)$ for $1 \leqslant i \leqslant n$, then we say that $\boldsymbol{u}$ is a word with uniform distribution.

If we set

$$\boldsymbol{g}_n=\left(\frac{1}{n},\frac{1}{n},\dots,\frac{1}{n}\right)\in D_n,$$

then from their respective definitions, we have

$$M(\boldsymbol{A},\boldsymbol{g}_n)=\frac{1}{n}\sum_{i=1}^{n}\ln\rho(A_i)$$

and

$$H(\boldsymbol{A},\boldsymbol{g}_n)=\begin{cases}\dfrac{1}{n}\displaystyle\sum_{i=1}^{\lfloor n/2\rfloor}\ln\rho(A_iA_{n+1-i}) & \text{if } n \text{ is even,}\\[2ex] \dfrac{1}{n}\left\{\displaystyle\sum_{i=1}^{\lfloor n/2\rfloor}\ln\rho(A_iA_{n+1-i})+\ln\rho(A_{\lceil n/2\rceil})\right\} & \text{if } n \text{ is odd.}\end{cases}$$

For $1 \leqslant x \leqslant y$, let

$$X=\begin{bmatrix}x & 1\\ 1 & 0\end{bmatrix},\qquad Y=\begin{bmatrix}y & 1\\ 1 & 0\end{bmatrix}$$

and consider the function $l(x, y) = 1.1 \ln\rho(XY) - \ln\rho(X)\rho(Y)$. Then

$$l(x,y)=1.1\ln(\tfrac{1}{2}(xy+2+\sqrt{x^2y^2+4xy}))-\ln(\tfrac{1}{4}(x+\sqrt{x^2+4})(y+\sqrt{y^2+4})).$$

The function l has various properties, summarized in the following lemma.

Lemma 4.1. *The function $l(x, y)$ satisfies the following conditions.*

(a) *If $1 \leqslant x_1 \leqslant x_2 \leqslant y$, then $l(x_1, y) \leqslant l(x_2, y)$.*

(b) *If $2 \leqslant x \leqslant y_1 \leqslant y_2$, then $l(x, y_1) \leqslant l(x, y_2)$.*

(c) $\min_{y\geqslant 2} l(1,y) > -0.083$.

(d) $\min_{y\geqslant 3} l(2,y) = l(2,3)$ (> 0.19).

Proof. Omitted. □

Proposition 4.2. *If $n \geqslant 3$, then $F(\boldsymbol{A}, \boldsymbol{g}_n) < 1.1$.*

Proof.

(1) $n = 3$. Then

$$F(\boldsymbol{A}, \boldsymbol{g}_3) = \frac{\ln \rho(A_1)\rho(A_3) + \ln \rho(A_2)}{\ln \rho(A_1 A_3) + \ln \rho(A_2)}.$$

So $F(\boldsymbol{A}, \boldsymbol{g}_3) < 1.1$ is equivalent to $l(a_1, a_3) + 0.1 \ln \rho(A_2) > 0$. From parts (b) and (d) of Lemma 4.1 we see that $l(a_1, a_3) > 0$ if $a_1 \geqslant 2$. If $a_1 = 1$, then $a_2 \geqslant 2$ so that $\rho(A_2) \geqslant 1 + \sqrt{2}$ and thus

$$\begin{aligned} l(a_1, a_3) + 0.1 \ln \rho(A_2) &\geqslant \min_{y\geqslant 2} l(1,y) + 0.1 \ln(1+\sqrt{2}) \\ &> -0.083 + 0.088 = 0.05 > 0. \end{aligned}$$

Thus in either case we have established that $F(\boldsymbol{A}, \boldsymbol{g}_3) < 1.1$.

(2) n odd, $n \geqslant 5$. Then $F(\boldsymbol{A}, \boldsymbol{g}_n) < 1.1$ if and only if

$$\sum_{l=1}^{\lfloor n/2 \rfloor} l(a_i, a_{n+1-i}) + 0.1 \ln \rho(A_{\lceil n/2 \rceil}) > 0. \tag{4.1}$$

The left-hand side of (4.1) is at least $l(a_1, a_n) + 0.1 \ln \rho(A_2)$ and hence is greater than 0 by part (1).

(3) n even, $n \geqslant 4$. Then

$$F(\boldsymbol{A}, \boldsymbol{g}_n) < 1.1 \;\Leftrightarrow\; \sum_{i=1}^{\lfloor n/2 \rfloor} l(a_i, a_{n+1-i}) > 0. \tag{4.2}$$

We will show that $l(a_1, a_n) + l(a_2, a_{n-1}) > 0$.

By Lemma 4.1,

$$\begin{aligned} l(a_1, a_n) + l(a_2, a_{n-1}) &\geqslant \min_{y\geqslant 3} l(1,y) + \min_{y\geqslant 3} l(2,y) \\ &> -0.083 + 0.19 > 0. \end{aligned}$$

□

Corollary 4.3. *If $\boldsymbol{u}$ is an infinite word taking values in the finite set Σ with uniform distribution and $|\Sigma| \geqslant 3$, then $L(\boldsymbol{u}) < 1.1$.*

It is possible to show that $L(\boldsymbol{u}) < 1.09$ if $|\Sigma| \geqslant 4$, and presumably we would get smaller bounds as $|\Sigma| \to \infty$.

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

5. Expansion factors for $\{\lfloor m\theta\rfloor \bmod n\}$

Suppose θ is irrational with $0 < \theta < 1$ and $n \geqslant 2$. We set $w_m = \lfloor m\theta\rfloor \bmod n$ and $\boldsymbol{w} = w_1w_2\dots$. If $\Phi : \{0, 1, \dots, n-1\} \to \Sigma$ is a bijection, then the infinite word $\boldsymbol{u}$ is defined by setting $u_m = \Phi(w_m)$ for $m \geqslant 1$. $\boldsymbol{u}$ is said to be *derived from* $\boldsymbol{w}$ and we also write $\boldsymbol{u} = \Phi \circ \boldsymbol{w}$. From Definition 1.1 it is evident that $\boldsymbol{u}$ has a segment expansion factor greater than or equal to γ if and only if $\boldsymbol{w}$ has a segment expansion factor greater than or equal to γ.

It will turn out that, in order to obtain the transcendence of continued fractions associated with such $\boldsymbol{u}$, we need to obtain conditions on θ to guarantee that $\boldsymbol{w}$ will have a segment expansion factor greater that $\frac{3}{2}$.

Let the continued fraction of θ be $[0, b_1, b_2, \dots]$ with convergents $(P_k/Q_k)_{k\geqslant 0}$. If t is a non-negative integer, then we set

$$P_{k,t} = tP_{k+1} + P_k, \qquad Q_{k,t} = tQ_{k+1} + Q_k.$$

Thus, in particular,

$$P_{k,0} = P_k, \qquad P_{k,b_{k+2}} = P_{k+2}$$

and

$$Q_{k,0} = Q_k, \qquad Q_{k,b_{k+2}} = Q_{k+2}.$$

If $b_{k+2} \geqslant 2$ and $1 \leqslant t \leqslant b_{k+2} - 1$, then $P_{k,t}/Q_{k,t}$ is called a *median convergent to* θ (see [**8**] for further information).

The following proposition generalizes what was proved in [**3**].

Proposition 5.1. *Suppose $k \geqslant 0$ and $0 \leqslant t \leqslant b_{k+2} - 1$. Then for all integers m satisfying $1 \leqslant m < Q_{k,t+1}$ we have*

$$\lfloor m\theta\rfloor = \begin{cases} \left\lfloor m\dfrac{P_{k,t}}{Q_{k,t}}\right\rfloor & \text{if } k \text{ is even,} \\ \left\lceil m\dfrac{P_{k,t}}{Q_{k,t}}\right\rceil - 1 & \text{if } k \text{ is odd.} \end{cases}$$

Proof.

(1) First consider when k is even. From the classical inequalities

$$0 < Q_{k+2}\theta - P_{k+2} < Q_{k+2}^{-1}$$

we derive

$$\lfloor m\theta\rfloor = \left\lfloor m\frac{P_{k+2}}{Q_{k+2}}\right\rfloor \quad \text{for } 1 \leqslant m \leqslant Q_{k+2}. \tag{5.1}$$

We now show that if t satisfies $0 \leqslant t \leqslant b_{k+2} - 1$, then

$$\left\lfloor m\frac{P_{k,t}}{Q_{k,t}}\right\rfloor = \left\lfloor m\frac{P_{k,t+1}}{Q_{k,t+1}}\right\rfloor \quad \text{for } 1 \leqslant m < Q_{k,t+1}. \tag{5.2}$$

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

From the basic theory we have

$$\frac{P_{k,t+1}}{Q_{k,t+1}} - \frac{P_{k,t}}{Q_{k,t}} = \frac{1}{Q_{k,t+1}Q_{k,t}},$$

and therefore if $1 \leqslant m < Q_{k,t+1}$, we obtain

$$m\frac{P_{k,t}}{Q_{k,t}} < m\frac{P_{k,t+1}}{Q_{k,t+1}} < m\frac{P_{k,t}}{Q_{k,t}} + \frac{1}{Q_{k,t}},$$

from which (5.2) follows immediately.

If we put $t = b_{k+2} - 1$ in (5.2) and use (5.1), we find that $\lfloor m\theta \rfloor = \lfloor m(P_{k,t}/Q_{k,t}) \rfloor$ for $1 \leqslant m < Q_{k,t+1}$. We can then put $t = b_{k+2} - 2$ in (5.2) and see that the required result holds. Continuing in this manner we establish the result for all $t : 0 \leqslant t \leqslant b_{k+2} - 1$.

(2) Now consider the case when k is odd. We can similarly show that if $1 \leqslant m \leqslant Q_{k+2}$, then

$$\lfloor m\theta \rfloor = \left\lceil m\frac{P_{k+2}}{Q_{k+2}} \right\rceil - 1, \tag{5.3}$$

and if $1 \leqslant m < Q_{k,t+1}$, $0 \leqslant t \leqslant b_{k+2} - 1$, then

$$\left\lceil m\frac{P_{k,t}}{Q_{k,t}} \right\rceil = \left\lceil m\frac{P_{k,t+1}}{Q_{k,t+1}} \right\rceil \tag{5.4}$$

and the result follows as before.

□

Corollary 5.2. *Suppose $k \geqslant 0$ and $0 \leqslant t \leqslant b_{k+2} - 1$. Then $\lfloor (Q_{k,t} + r)\theta \rfloor = P_{k,t} + \lfloor r\theta \rfloor$ for $1 \leqslant r \leqslant Q_{k+1} - 1$.*

Proof. Omitted. □

For each $k \geqslant 0$ and for $0 \leqslant t \leqslant b_{k+2} - 1$ we define the following prefixes of $\boldsymbol{w}$:

$$X_{k,t} = \{\lfloor m\theta \rfloor \bmod n\}_{1 \leqslant m \leqslant Q_{k,t}},$$
$$Z_{k,t} = \{\lfloor m\theta \rfloor \bmod n\}_{1 \leqslant m < Q_{k,t+1}}.$$

For convenience we write $X_k = X_{k,0}$ and $Z_k = Z_{k,0}$. The prefix partial order will be denoted by '$\leqslant$', so it is evident that $X_{k,t} \leqslant Z_{k,t}$.

We are now able to prove the following proposition.

Proposition 5.3. *Suppose $\theta = [0, b_1, b_2, \dots]$ with convergents $\{P_k/Q_k\}_{k \geqslant 0}$ and $\boldsymbol{w} = \{\lfloor m\theta \rfloor \bmod n\}_{m \geqslant 1}$.*

(a) *If there is an infinite number of k with $P_k \equiv 0 \bmod n$, then $\boldsymbol{w}$ has a prefix expansion factor greater than or equal to 2.*

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

(b) *Suppose* $M = \limsup_{r\to\infty} b_r < \infty$. *If there is an infinite number of* k *with* $P_k + P_{k+1} \equiv 0 \bmod n$ *and* $b_{k+2} \geqslant 2$, *then* $\boldsymbol{w}$ *has a prefix expansion factor greater than or equal to* $\frac{3}{2} + 1/(8M)$.

(c) *If there is an infinite number of* k *with* $b_{k+1} \geqslant \frac{3}{2}n$, *then* $\boldsymbol{w}$ *has a prefix expansion factor greater than or equal to* $\frac{3}{2} + 1/n$.

Proof.

(a) Suppose $P_k \equiv 0 \bmod n$. By Corollary 5.2 (with $t = 0$), we can write $Z_k = X_k Y_k$ with $Y_k \leqslant X_k^s$ for some integer s. Let $U_k = \lambda$, $V_k = X_k$, $W_k = Y_k$. Then

$$\frac{|V_k W_k|}{|V_k|} = \frac{Q_{k+1} + Q_k - 1}{Q_k} > 2.$$

Since there are an infinite number of such k, we can conclude that $\boldsymbol{w}$ has a prefix expansion factor greater than or equal to 2.

(b) Suppose $P_k + P_{k+1} \equiv 0 \bmod n$ and $b_{k+2} \geqslant 2$. Applying Corollary 5.2 with $t = 1$ gives $\lfloor (Q_{k,1} + r)\theta \rfloor \equiv \lfloor r\theta \rfloor \bmod n$ for $1 \leqslant r \leqslant Q_{k+1} - 1$. Hence $Z_{k,1} = X_{k,1} Y_{k,1}$ with $Y_{k,1} \leqslant X_{k,1}^s$ for some positive integer s. Set $U_k = \lambda$, $V_k = X_{k,1}$, $W_k = Y_{k,1}$ and we find that

$$\frac{|V_k W_k|}{|V_k|} = \frac{2Q_{k+1} + Q_k - 1}{Q_{k+1} + Q_k}.$$

The assumption that $\limsup_{r\to\infty} b_r = M \geqslant 2$ gives, for all sufficiently large k,

$$\frac{Q_{k+1}}{Q_k} \geqslant 1 + \frac{1}{M+1}$$

and hence that

$$\frac{|V_k W_k|}{|V_k|} \geqslant \frac{3}{2} + \frac{1}{8M}.$$

(c) Suppose $b_{k+1} \geqslant \frac{3}{2}n$. Put $\tilde{X}_k = \{\lfloor m\theta \rfloor \bmod n\}_{1 \leqslant m \leqslant nQ_k}$. Then $|Z_k| = Q_{k+1} + Q_k - 1 > nQ_k = |\tilde{X}_k|$ so $\tilde{X}_k \leqslant Z_k$. Thus $Z_k = \tilde{X}_k \tilde{Y}_k$ and $\tilde{Y}_k \leqslant \tilde{X}_k^s$ for some positive integer s. Setting $U_k = \lambda$, $V_k = \tilde{X}_k$, $W_k = \tilde{Y}_k$ we obtain

$$\frac{|V_k W_k|}{|V_k|} = \frac{Q_{k+1} + Q_k - 1}{nQ_k} \geqslant \frac{\frac{3}{2}nQ_k + Q_{k-1} + Q_k - 1}{nQ_k} \geqslant \frac{3}{2} + \frac{1}{n}.$$

$\square$

Remark 5.4.

(1) Proposition 5.3 only refers to results involving $t = 0$ and $t = 1$. In fact if we consider $t \geqslant 2$ (so that necessarily $b_{k+2} \geqslant 3$) we find that

$$\frac{|Z_{k,t}|}{|X_{k,t}|} = \frac{Q_{k,t+1} - 1}{Q_{k,t}} < \frac{3}{2}$$

and we do not have an instance to demonstrate that $\boldsymbol{w}$ has a prefix expansion factor greater than or equal to $\frac{3}{2}$.

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

(2) It is possible to obtain analogous results to Proposition 5.3 when $\theta > 1$, but we have suppressed the details.

We now conclude this section by finding a situation where $\boldsymbol{w}$ has a segment expansion factor greater than $\frac{3}{2}$. From Corollary 5.2 (with $t = 0$, $k \geqslant 0$) we see that

$$\lfloor (Q_{k+1} + r)\theta \rfloor = \lfloor (Q_k + r)\theta \rfloor + P_{k+1} - P_k \quad \text{for } 1 \leqslant r \leqslant Q_{k+1} - 1,$$

and we can now prove the following proposition.

Proposition 5.5. *Suppose* $M = \limsup_{r\to\infty} b_r < \infty$. *If there are an infinite number of* k *satisfying* $P_k \equiv P_{k+1} \bmod n$ *and* $b_{k+1} \geqslant 3$, *then* $\boldsymbol{w}$ *has a segment expansion factor greater than or equal to* $\frac{3}{2} + 1/(12M)$.

Proof. Assume $P_k \equiv P_{k+1} \bmod n$ and let $U_k = \{\lfloor m\theta \rfloor \bmod n\}_{1 \leqslant m \leqslant Q_k}$,

$$V_k = \{\lfloor m\theta \rfloor \bmod n\}_{Q_k < m \leqslant Q_{k+1}}, \qquad W_k = \{\lfloor m\theta \rfloor \bmod n\}_{Q_{k+1} < m < 2Q_{k+1}}.$$

Then $U_k V_k W_k \leqslant \boldsymbol{w}$. By the observation just preceding Proposition 5.5 it is evident that $W_k \leqslant V_k^s$ for some positive integers s. Also

$$\frac{|U_k V_k W_k|}{|U_k| + |U_k V_k|} = \frac{2Q_{k+1} - 1}{Q_k + Q_{k+1}} \geqslant \frac{3}{2} + \frac{1}{12M},$$

since under our present assumption

$$\frac{Q_{k+1}}{Q_k} \geqslant 3 + \frac{1}{M+1}$$

for large k instances. □

6. Approximation by quadratic irrationals

We now proceed to make the connection between segment/prefix expansion factors and quadratic approximation. If η is a quadratic irrational satisfying the minimal equation $a\eta^2 + b\eta + c = 0$, where $a, b, c \in \mathbb{Z}$ with $\gcd(a, b, c) = 1$, then we set $H(\eta) = \max\{|a|, |b|, |c|\}$. As before, $\boldsymbol{u} = u_1 u_2 \dots$, $\alpha = [0, u_1, u_2, \dots]$ and $\{p_m/q_m\}_{m \geqslant 0}$ is the sequence of convergents of α.

The following result is a refinement of the estimate given by Baker in [**2**].

Lemma 6.1. *Let*

$$\eta = [0, u_1, u_2, \dots, u_{h-1}, \overline{u_h, \dots, u_{h+k-1}}].$$

Then $H(\eta) < 2q_{h-1}q_{h+k-1}$.

Proof. Let $\eta_h = [\overline{u_h, u_{h+1}, \dots, u_{h+k-1}}]$. Then

$$\eta = \frac{p_{h-1}\eta_h + p_{h-2}}{q_{h-1}\eta_h + q_{h-2}} = \frac{p_{h+k-1}\eta_h + p_{h+k-2}}{q_{h+k-1}\eta_h + q_{h+k-2}}.$$

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

Eliminating η_h, we obtain $P\eta^2 + Q\eta + R = 0$, where

$$P = q_{h-2}q_{h+k-1} - q_{h-1}q_{h+k-2},$$
$$Q = q_{h-1}p_{h+k-2} + p_{h-1}q_{h+k-2} - p_{h-2}q_{h+k-1} - q_{h-2}p_{h+k-1},$$
$$R = p_{h-2}p_{h+k-1} - p_{h-1}p_{h+k-2}.$$

Now $0 < \eta < 1$ so $p_r \leqslant q_r$ for $r \geqslant 0$. Hence

$$|P| \leqslant q_{h-1}q_{h+k-1}, \qquad |R| \leqslant q_{h-1}q_{h+k-1}$$

and

$$|Q| \leqslant \max\{2q_{h-1}q_{h+k-2}, 2q_{h-2}q_{h+k-1}\} \leqslant 2q_{h-1}q_{h+k-1}.$$

□

Lemma 6.2. *Suppose $\{U_k\}_{k\geqslant 1}$ and $\{V_k\}_{k\geqslant 1}$ are two families of words (in Σ) satisfying the two conditions:*

(i) $U_kV_k < \boldsymbol{u}$*; and*

(ii) $\lim_{k\to\infty} |V_k| = \infty$.

Then

$$\limsup_{k\to\infty}\left(\frac{\ln q_{|U_k|}q_{|U_kV_k|}}{|U_k| + |U_kV_k|}\right) \leqslant \limsup_{m\to\infty}\left(\frac{1}{m}\right)\ln q_m.$$

Proof. Let $M = \limsup_{m\to\infty}(1/m)\ln q_m$ and let $\epsilon > 0$. Then there exists $m_0 \geqslant 1$ such that if $m \geqslant m_0$, then $(1/m)\ln q_m < M + \frac{1}{2}\epsilon$. Let $A = \max_{0\leqslant m\leqslant m_0-1}\ln q_m$. Since $\lim_{k\to\infty}|V_k| = \infty$, then there exists k_0 such that if $k \geqslant k_0$, then $|V_k| \geqslant m_0$ and $A < \frac{1}{2}\epsilon|V_k|$. Now put $I = \{k : |U_k| < m_0\}$ and $J = \{k : |U_k| \geqslant m_0\}$. If $k \in I$, $k \geqslant k_0$ we have

$$\frac{\ln q_{|U_k|}q_{|U_kV_k|}}{|U_k| + |U_kV_k|} \leqslant \frac{A + \ln q_{|U_k|V_k|}}{|U_kV_k|} < \tfrac{1}{2}\epsilon + M + \tfrac{1}{2}\epsilon = M + \epsilon,$$

whereas, if $k \in J$, $k \geqslant k_0$ we have

$$\begin{aligned}\frac{\ln q_{|U_k|}q_{|U_kV_k|}}{|U_k| + |U_kV_k|} &= \frac{|U_k|}{|U_k| + |U_kV_k|}\frac{\ln q_{|U_k|}}{|U_k|} + \frac{|U_kV_k|}{|U_k| + |U_kV_k|}\frac{\ln q_{|U_kV_k|}}{|U_kV_k|}\\ &< M + \tfrac{1}{2}\epsilon,\end{aligned}$$

since it is a convex combination of two numbers, both less than $M + \frac{1}{2}\epsilon$. □

The following theorem generalizes Theorem 4 of [**1**].

Theorem 6.3. *If α is not a quadratic irrational (that is, $\boldsymbol{u}$ is not ultimately periodic) and if $\boldsymbol{u}$ has a segment expansion factor greater than or equal to γ, then there is a sequence of quadratic irrationals (α_k) which satisfies*

$$|\alpha - \alpha_k| < \frac{1}{H(\alpha_k)^{2\gamma/L(\alpha)}}, \tag{6.1}$$

where $H(\alpha_k)$ denotes the height of α_k.

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

Proof. By assumption there are three families of words $\{U_k\}_{k\geqslant 1}$, $\{V_k\}_{k\geqslant 1}$, $\{W_k\}_{k\geqslant 1}$ satisfying the requirements of Definition 1.1. Set

$$\alpha_k = [0, u_1, u_2, \dots, u_{|U_k|}, \overline{u_{|U_k|+1}, \dots, u_{|U_k V_k|}}].$$

By standard theory, α_k is a quadratic irrational. Furthermore, since $\lim_{k\to\infty} |V_k| = \infty$, it is evident that there an infinite number of distinct α_ks. By Lemma 6.1, $H(\alpha_k) < 2q_{|U_k|}q_{|U_k V_k|}$. In addition, from Definition 1.1 it is clear that α and α_k have the same first $|U_k V_k W_k|$ partial quotients. Thus

$$|\alpha - \alpha_k| \leqslant \frac{1}{q^2_{|U_k V_k W_k|}}. \tag{6.2}$$

The required result will then follow if we can show that

$$q^2_{|U_k V_k W_k|} \geqslant (2q_{|U_k|}q_{|U_k V_k|})^{2\gamma/L(\alpha)}$$

for all k that are large enough. Using Lemma 6.2,

$$\begin{aligned}
\liminf_{k\to\infty} \frac{2\ln q_{|U_k V_k W_k|}}{\ln(2q_{|U_k|}q_{|U_k V_k|})} &= 2\liminf_{k\to\infty}\left\{\frac{\ln q_{|U_k V_k W_k|}}{\ln(q_{|U_k|}q_{|U_k V_k|})}\right\} \\
&\geqslant 2\liminf_{k\to\infty}\left\{\left(\frac{\ln q_{|U_k V_k W_k|}}{|U_k V_k W_k|}\bigg/\frac{\ln(q_{|U_k|}q_{|U_k V_k|})}{|U_k| + |U_k V_k|}\right)\frac{|U_k V_k W_k|}{|U_k| + |U_k V_k|}\right\} \\
&\geqslant 2\frac{\liminf_{m\to\infty}(1/m)\ln q_m}{\limsup_{m\to\infty}(1/m)\ln q_m}\cdot\gamma \\
&= \frac{2\gamma}{L(\boldsymbol{\alpha})},
\end{aligned}$$

as required. □

Theorem 6.4. *Let $0 < \theta < 1$ be irrational with continued fraction expansion $\theta = [0, b_1, b_2, \dots]$ and principal convergents $(P_k/Q_k)_{k\geqslant 0}$. Let n be an integer greater than or equal to 2 and $\Sigma \subset \mathbb{Z}^+$ with $|\Sigma| = n$. Suppose $\Phi : \{0, 1, \dots, n-1\} \to \Sigma$ is a bijection, $\boldsymbol{w} = \{\lfloor m\theta\rfloor \bmod n\}_{m\geqslant 1}$ and $\boldsymbol{u} = \Phi \circ \boldsymbol{w}$ with associated continued fraction α.*

Then α is transcendental if any of the following conditions hold for an infinite number of positive integers k:

(a) $P_k \equiv 0 \bmod n$;

(b) $b_{k+1} \geqslant \frac{3}{2}n$;

(c) $P_k + P_{k+1} \equiv 0 \bmod n$;

(d) $P_k \equiv P_{k+1} \bmod n$.

Proof. Since θ is irrational it is clear that α is neither a rational nor a quadratic irrational. In view of Theorem 6.3 and Schmidt's Theorem [**9**], the transcendence of α will be proved provided $(2\gamma/L(\alpha)) > 3 + \delta$, for some $\delta > 0$. For any irrational θ, it can

https://doi.org/10.1017/S001309150000119X Published online by Cambridge University Press

be shown by classical ergodic theory that $L(\boldsymbol{u}) = 1$ (cf. [**5**, **7**]). (I thank the referee for bringing this to my attention.) A purely elementary proof of this fact can be found in [**4**]. Thus we need only demonstrate that $\boldsymbol{w}$ has a segment expansion factor $\gamma \geqslant \frac{3}{2} + \frac{1}{2}\delta$ for some $\delta > 0$.

If (a) or (b) holds for an infinite number of k, then Proposition 5.3 (a), (c) gives the required result. Suppose that (c) holds, but neither (a) nor (b) holds for an infinite number of k. Then $\limsup_{r\to\infty} b_r \leqslant \frac{3}{2}n$ and $P_r \not\equiv 0 \bmod n$ for all sufficiently large r. Now if $b_{k+2} = 1$, then $P_{k+2} = P_{k+1} + P_k \equiv 0 \bmod n$. So we must have $b_{k+2} \geqslant 2$ for all sufficiently large instances of (c). Proposition 5.3 (b) then yields the result.

Finally, suppose that (d) holds but none of (a)–(c) hold for an infinite number of k. It is then easy to check that we must have $b_{k+1} \geqslant 3$ for all sufficiently large instances of (d). We can then use Proposition 5.5 to complete the proof. □

Theorem 6.4 immediately gives the result that α is always transcendental when n equals 2 or 3. If $n = 5$ and $\theta = [0, 1, 3, 1, 1, 1, \dots]$, then for all $k \geqslant 1$, we find that none of the conditions (a)–(d) hold; so it is an open question whether the corresponding α can be shown to be transcendental by the methods of this paper.

Acknowledgements. I thank the referee for the many constructive comments and suggestions on how to improve this paper. Some of the work was done while visiting the University of Edinburgh and I express my gratitude for the hospitality shown to me.

References

1. J.-P. Allouche, J. L. Davison, M. Queffélec and L. Q. Zamboni, Transcendence of Sturmian or morphic continued fractions, *J. Number Theory* **91** (2001), 39–66.
2. A. Baker, Continued fractions of transcendental numbers, *Mathematika* **9** (1962), 1–8.
3. J. L. Davison, A class of transcendental numbers with bounded quotients, in *Number Theory and Applications, Banff, AB, 1988*, NATO Advanced Science Institutes Series C, Mathematical and Physical Sciences, no. 265 pp. 365–371 (Kluwer, Dordrecht, 1989).
4. J. L. Davison, Levy constants for classes of continued fractions (unpublished).
5. C. Faivre, Distribution of Lévy constants for quadratic numbers, *Acta Arithm.* **61** (1992), 113–134.
6. G. Hardy and E. Wright, *An introduction to the theory of numbers* (Oxford University Press, 1975).
7. H. Jager and P. Liardet, Distributions arithmétiques des dénominateurs de convergents de fractions continues, *Nederl. Akad. Wetensch. Indag. Math.* **50** (1988), 181–197.
8. S. Lang, *An introduction to Diophantine approximation* (Addison-Wesley, 1966).
9. W. Schmidt, On simultaneous approximation of two algebraic numbers by rationals, *Acta Math.* **119** (1967), 27–50.