Journal of Statistical Physics, Vol. 66, Nos. 5/6, 1992

A Multidimensional Continued Fraction and Some of Its Statistical Properties

P. R. Baldwin[1]

Received December 11, 1990; final November 7, 1991

The problem of simultaneously approximating a vector of irrational numbers with rationals is analyzed in a geometrical setting using notions of dynamical systems theory. We discuss here a (vectorial) multidimensional continued-fraction algorithm (MCFA) of additive type, the generalized mediant algorithm (GMA), and give a geometrical interpretation to it. We calculate the invariant measure of the GMA shift as well as its Kolmogorov–Sinai (KS) entropy for arbitrary number of irrationals. The KS entropy is related to the growth rate of denominators of the Euclidean algorithm. This is the first analytical calculation of the growth rate of denominators for any MCFA.

KEY WORDS: Continued fractions; entropy; algorithm.

1. INTRODUCTION

The problem of simultaneously approximating a single irrational or a set of irrationals has an illustrious history and given rise to much beautiful mathematics.[1–3] For the case of approximating a single irrational, our understanding is quite complete. There is a unique algorithm which gives "best" approximations to any given irrational. This algorithm is the ordinary continued-fraction (OCF) expansion and was discussed by, among others, Gauss. The ergodic theory (invariant measure and Kolmogorov–Sinai entropy) as well as the growth of denominators is completely understood. Its dynamical systems interpretation is due to Khintchine.[6]

Although there has been a tremendous amount of work on the problem, the subject of approximating *sets* of irrationals is much less developed. An approximation procedure was given by Jacobi and put on a

[1] Department of Physics, University of Akron, Akron, Ohio 44325-4001.

0022-4715/92/0300-1463$06.50/0

more rigorous basis by Perron, resulting in the famous Jacobi–Perron[2] algorithm of which there are many variants. However, all proposed algorithms suffer various deficiencies. Moreover, a rigorous result states that none of these algorithms may possess *all* of the approximation properties which the OCF does for the approximation of a single irrational (for a nice summary see Szekeres[9] and Lagarias[10]). The field of multidimensional continued-fraction algorithms (MCFA) has given rise to a myriad of generalized algorithms (F-expansions)[11] with different merits.[14] In summary, the bulk of the most interesting questions remain unanswered. The foremost question is, does there exist a convenient algorithm that lists the entire set of best convergents, and only this set? Or at least an algorithm which is able to list a sequence, which contains a subsequence which shares a subsequence with the sequence of best convergents?

There are several goals to this paper. First, we propose and study a simple MCFA, the generalized mediant algorithm (GMA). The algorithm has a simple geometrical interpretation and has the very pleasant feature that its invariant measure and Kolmogorov–Sinai entropy may be explicitly calculated (for any number of irrationals). As is known, finding an analytical form for the invariant measure is never guaranteed.[7] Knowing the KS entropy yields the growth rate of denominators via a simple relation (see ref. 8, especially the discussion relating to Corollary 7.10)

$$\lambda_1 = e^{h/(I+1)}$$

where I is the number of irrationals to be approximated, λ_1 is the eigenvalue governing the growth rate of denominators, and h is the KS entropy for the unique absolutely continuous invariant (ACI) measure. These are the first analytical results on the growth rate of denominators for an MCFA.

This work gives explicit closed-form analytical results on convergence properties of an MCFA. In future work, we will show how the dynamical systems perspective enables us to evaluate a convergence exponent. The philosophy of our "program" is to state and approach all the Diophantine metric properties completely within a dynamical systems context.

Section 2 reviews the one-irrational (two-dimensional) case. We discuss the geometric aspects of the Farey shift and OCF, and their relation to the study of eigenvalues of shear matrices. Expositing the material this way helps us develop the results of the GMA algorithm, since the OCF case is easier to visualize. The general approach is the same in both cases.

In Section 3, we proceed to explain the GMA and its geometrical interpretation. We state Proposition 2: any triple of integers with greatest common factor unity may be written as a sum of the rows of a suitable product of elementary matrices. The GMA provides one way of carrying this procedure out. It is analogous to the Farey shift geometrically, since the geometrical rule for both algorithms is "move vertices to mediants" (hence the name, generalized mediant algorithm). For the GMA, we give the relation between products of elementary shear matrices and the GMA shift map. The associated shift map is characterized in the following fashion: (i) the support of its invariant measure $\mathscr{S}$ is given by the ordered unit hypercube, with the additional restriction that the sum of the two smallest irrationals is larger than unity, i.e.,

$$z > \cdots > b > a > 1 - b \tag{1.1}$$

where a, b are the smallest and next to smallest irrationals, respectively, and z represents the greatest irrational. (ii) The GMA shift is a two-to-one map on $\mathscr{S}$ (the inverse image of almost all points of $\mathscr{S}$ contains two points).

In Section 4, we discuss the relation between the eigenvalues of the shift map and the eigenvalues of the corresponding string of products of elementary matrices (which we term the E-string). Let λ_1 be the ("average") eigenvalue of the E-string lying outside the unit circle and let $\{\lambda_i\}^{i=2,\dots,d}$ be the set of $d-1=I$ other eigenvalues (all lying within the unit circle). Then the eigenvalues of the shift are given by $\{\lambda_1/\lambda_{d-j+2}\}^{j=2,\dots,d}$. In order to formulate properly the notion of eigenvalues for dynamical systems, we must also introduce in this section the positive, real, Oseledec eigenvalues.

In Section 5, we determine an equation for the invariant density of the shift. With the invariant measure in hand and knowing the relation of Section 4, we are able to study the KS entropy, and hence the growth rate of denominators. We give an analytical expression for and plot the KS entropy as a general function of dimension. Also, we analyze analytically the behavior as I (the number of irrationals) approaches 1 or grows large. The KS entropy is easily related to the largest eigenvalue of the E-string as stated above. Section 6 is a discussion and conclusion. Proofs for statements in Section 3 are given in Appendices A and B, whereas proofs for Section 4 are given in Appendix C. For convenience, a glossary is appended.

The question of approximating sets of numbers is in itself of fundamental importance. Let us point out very briefly, though, why the question may be important for someone interested in the field of Hamiltonian systems, for example. For over 9 years,[16] researchers have tried to understand questions regarding the breaking of KAM tori in 3-degree-of-freedom systems, without success. In 2-degree-of-freedom systems one sees a magnificent relation between stability (robustness) and the Diophantine properties of the winding numbers (in dissipative systems: Arnold tongues).[12] These Diophantine properties are understood via continued-fraction expansions and the Farey shift.[13] It seems likely that understanding the metric properties of sets of irrationals may help in understanding analyticity breaking in higher-degree-of-freedom systems.

This paper is the first of two papers discussing MCFAs. In this first paper, we give an algorithm with positive entropy implying an exponential growth rate of the denominators of the rational approximants given by the algorithm. The second paper determines an important convergence exponent for MCFAs for the case of two irrationals for several different algorithms. Employing a dynamical systems viewpoint, we express this exponent in terms of the Lyapunov exponents of the shift. Algorithms with best approximation properties must have a certain value for this convergence exponent. In particular, this straightforward analysis shows numerically that neither JP nor the GMA algorithm find best Diophantine approximations for almost all reals.

2. APPROXIMATION OF TWO-DIMENSIONAL VECTORS (OCF)

In this section, we review some geometrical and statistical properties of the Farey shift (FS) and the ordinary continued fraction (OCF). We present these well-known results since they lead naturally to the GMA. Moreover the results are presented in such a manner as to reflect our dynamical systems philosophy to MCFAs.

We begin our exposition with a definition followed by a well-known result.

Definition 1. *Elementary matrices.* Define the matrices $\mathsf{E}_{ij}^{(d)}$ by

$$(\mathsf{E}_{ij}^{(d)})_{kl} = \delta_{kl} + \delta_{ik}\delta_{jl} \tag{2.1}$$

for

$$1 \leqslant i, j, k, l \leqslant d \quad \text{and} \quad i \neq j \tag{2.2}$$

We call the E_{ij} elementary matrices.

We will suppress the superscript d when the dimension of the space is evident. The E_{ij} have unit entries on the main diagonal with also exactly one off-diagonal unit entry (in the ijth place). In the case $n=2$, $\mathsf{E}_{12} = [\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}]$, $\mathsf{E}_{21} = [\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}]$.

Proposition 1 (ref. 5, Chapter 3). Any relatively prime pair of integers (M, N) may be uniquely represented as a sum of the columns of the following matrix:

$$\mathsf{K} = \begin{bmatrix} M_l & N_l \\ M_r & N_r \end{bmatrix}$$

where the matrix K is written as the product of a string of (E_{ij}), beginning with E_{12} (i.e., the rightmost member of the matrix product). We call this decomposition into elementary matrices the E-string.

The proposition has an elegant geometrical interpretation discussed by Minkowski (see Fig. 1 and ref. 5, Chapter 3). We wish to approximate a certain irrational number which is equal to the slope of a ray (not drawn in Fig. 1) from the origin. Consider a parallelogram formed by the origin, two vertices (M_l, N_l), (M_r, N_r), as well as that vertex which we shall call *the focus* (M, N) $(\equiv M_l + M_r,\ N_l + N_r)$. We can thus identify the parallelogram with a matrix by writing

$$\mathsf{K} = \begin{bmatrix} M_l & N_l \\ M_r & N_r \end{bmatrix}$$

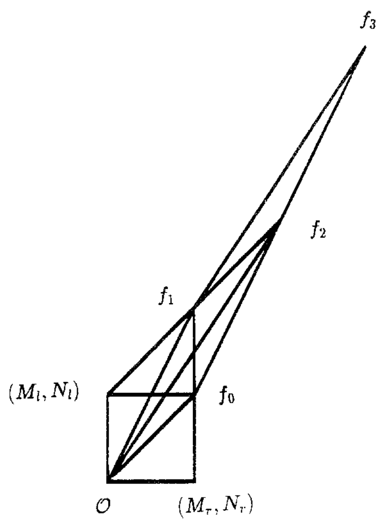

Fig. 1. A converging simplex in two dimensions. $\mathcal{O}$ labels the origin, and (M_l, N_l) and (M_r, N_r) the left and right vertexes of the initial parallelogram. The point $f_0 = (M_l + M_r, N_l + N_r)$ is the initial focus. Three shears have been applied to this original parallelogram. Each shear leaves the origin fixed, but moves (M_l, N_l), (M_r, N_r) to new values. We have labeled the new position of the focus after each shear by f_i.

as done in Proposition 1. The parallelogram is updated by applying one of the matrices (E_{ij}) to K, and reading off the new coordinates from the newly formed matrix. The update is done such that the ray always pierces one of the sides of the parallelogram. There is always a unique way to do this, if the slope of the line is irrational. In this way the coordinates of the focus (when written as a fraction) form rational approximants to the irrational slope of the ray.

For example, consider the unit square which is one of the regions depicted in Fig. 1. Three elementary shears have been performed, and the original parallelogram (the square) has been sheared into the most elongated parallelogram in the figure (whose final focus is given by f_3). Subsequent shears continued to stretch (shear) the parallelogram. The entire history of shears is recorded in the E-string whose product is the matrix K of Proposition 1. Since the matrix K is built up from a product of elementary matrices, the absolute value of the determinant of this matrix is unity, implying that the area inside the parallelogram is unity. The eigenvalues of the matrix K only change when there is a switch in the application of E_{12} to E_{21} (or vice versa).

There is a relation between this sequence of application of shear matrices and a mapping of the unit interval into itself. We will make this correspondence clear by way of examples.

2.1. The Farey Shift

If we watch a progression of the foci as the label n of Proposition 1 runs from larger values to smaller values, we note the following behavior (for $A < B$):

$$(A, B) \to (A', B') = (A, B - A)$$

that is, we subtract the smaller from the larger. We call this procedure *the Euclidean algorithm.*

Suppose we wish to analyze the map induced on $x = \min(A, B)/\max(A, B)$ by the above procedure. Defining $x' = \min(A', B')/\max(A', B')$, we find that x' may be expressed in terms of x:

$$\begin{aligned} x' &= F(x) \\ &= \begin{cases} T_0(x), & 0 < x \leqslant 1/2 \\ T_1(x), & 1/2 < x < 1 \end{cases} \end{aligned} \tag{2.3}$$

where we have defined two elementary shift operations:

$$T_0(x) = \frac{x}{1-x}; \qquad T_1(x) = \frac{1-x}{x} \tag{2.4}$$

Each shift function is continuous on the open unit interval. The function $T_0(x)$ maps the closed interval [0, 1/2] to the closed unit interval, whereas $T_1(x)$ maps [1/2, 1] to the closed unit interval. The function $F(x)$ is thus a tentlike two-to-one function on the unit interval with maximum at 1/2. We call $F(x)$ *the associated shift of the Farey Euclidean algorithm.* The shift we define in (2.3) is the *Farey shift.*

We give a specific example of the relation between the E-string and the Farey shift in Fig. 2. Reading downward from the top, we see that if we subtract one entry from the other such that the formerly larger entry becomes smaller, then this corresponds to an application of $T_1(x)$ and the matrix type of the E-string is changed. Otherwise $F(x)$ corresponds to $T_0(x)$ and the elementary matrix of the E-string remains unchanged. To demonstrate Proposition 1, we write

$$\mathsf{E}_{12}^2 \mathsf{E}_{21}^{12} \mathsf{E}_{12}^6 = \begin{bmatrix} 25 & 152 \\ 12 & 73 \end{bmatrix}$$

The sum of the columns is given by (37, 225), the fraction to be approximated. Thus, the expansion by the Farey shift T_0, T_1 of any irrational corresponds to a Euclidean algorithm with a nice geometrical interpretation. Observe that by keeping track of the sequence of the T_i which we employed in reducing the irrational with the larger denominator, we can invert the procedure and recover the irrational.

There are of course many advantages to studying a shift map on the unit interval rather than a Euclidean algorithm or the ergodic theory of a matrix group. For example, one may calculate the ergodic properties such as invariant measure, and the Lyapunov spectrum.

Example 1. $\frac{37}{225} = \frac{1}{6} + \frac{1}{12} + \frac{1}{3}$

M	N	$F(x)$	E-string
37	225	T_0	E_{12}
37	188	T_0	E_{12}
37	151	T_0	E_{12}
37	114	T_0	E_{12}
37	77	T_0	E_{12}
37	40	T_1	E_{12}
37	3	T_0	E_{21}
34	3	T_0	E_{21}
31	3	T_0^9	E_{21}
4	3	T_1	E_{21}^9
1	3	T_0	E_{12}
1	2	T_0	E_{12}
1	1		E_{12}

Fig. 2. An example of OCF: reduction of matrix pair (37, 225).

2.2. Farey Shift—Statistical Properties

We define the invariant measure[19] of the shift map as

$$d\mu(x) = d\mu(F^{-1}(x)) \tag{2.5}$$

and the invariant density as

$$\rho(x) = \frac{d\mu(x)}{dx} \tag{2.6}$$

The invariant density for an absolutely continuous invariant measure should satisfy

$$\rho(x) = \sum_{T(z)=x} \rho(z) \frac{dz}{dx} \tag{2.7}$$

which reduces to

$$\rho(x) = \left(\frac{1}{1+x}\right)^2 \left[\rho\left(\frac{1}{1+x}\right) + \rho\left(\frac{x}{1+x}\right)\right] \tag{2.8}$$

for the Farey map. There is a continuous measure on the open unit interval satisfying Eq. (2.8) given by

$$\rho(x) = 1/x \tag{2.9}$$

Although this density cannot be normalized, it can be used to find relative densities on the open interval.

The KS entropy for a one-dimensional map is given by

$$h = \int d\mu \ln |J| \tag{2.10}$$

where J is the Jacobian of the map given by $J(x) = dT(x)/dx$. For the Farey shift, it is easy to show that the entropy of the Farey shift is zero, using a regularization procedure. Thus the growth of the denominators for the Farey shift is, on the average, subexponential. In fact,

$$\lim_{n\to\infty} \frac{\ln n \cdot \ln(\text{denominators of Farey shift})}{n} = \frac{\pi^2}{12} \tag{2.11}$$

2.3. Ordinary Continued-Fraction Shift

The ordinary-continued fraction Euclidean algorithm is given by (for $A < B$)

$$(A, B) \to (A', B') = (A, B - [B/A] A)$$

where [•] is the Gauss integer symbol meaning the largest integer smaller than •. Thus we subtract the smaller integer from the larger enough times until it is no longer the larger of the two integers. The two integers never become equal because by supposition they are relatively prime.

The map induced on $x = \min(A, B)/\max(A, B)$ by the above procedure is

$$x' = T_{\mathrm{OCF}}(x) = \frac{1}{x} - \left[\frac{1}{x}\right] \tag{2.12}$$

The ordinary continued-fraction map demonstrates all the best approximation properties. We refer the reader to Lagarias[10] and Brentjes.[14] Note that the OCF can be expressed in terms of the same primitive operators T_0, T_1 as the Farey shift:

$$T_{\mathrm{OCF}}(x) = T_1 T_0^{[1/x]-1}(x) \tag{2.13}$$

Equation (2.13) may be checked by direct substitution using (2.4) and (2.12). The appearance of T_1 in the string signals the switch in the matrix type and thus the growth of the eigenvalue of the E-string. From these simple considerations it is clear that the continued fraction map has positive entropy. We proceed to show this by direct calculation.

2.4. OCF—Statistical Properties

For the OCF, the equation for the invariant density reads

$$\rho(y) = \sum_{k=1}^{\infty} \rho\left(\frac{1}{k+y}\right)\left(\frac{1}{k+y}\right)^2 \tag{2.14}$$

with a properly normalized solution given by

$$\rho(y) = \frac{1}{\ln 2} \frac{1}{1+y} \tag{2.15}$$

The result for the entropy is given by

$$h_{\mathrm{OCF}} = \int d\mu \ln J \tag{2.16}$$

$$= \frac{1}{\ln 2} \int_0^1 \frac{dx}{1+x} \ln \frac{1}{x^2} \tag{2.17}$$

$$= \frac{\pi^2}{6 \ln 2} = 2.37 \tag{2.18}$$

and the denominators grow like

$$\lim_{n \to \infty} \frac{\ln(\text{denominators of OCF})}{n} = \frac{\pi^2}{12 \ln 2} \tag{2.19}$$

For a thorough discussion see Khintchine.[6]

The dynamical systems approach we will employ throughout our discussion of continued fractions is not the usual number-theoretic approach to Diophantine properties. Moreover, this approach naturally suggests the GMA, which we construct in the next section from geometric considerations.

3. APPROXIMATION OF HIGHER-DIMENSIONAL VECTORS (GMA)

The following proposition paves the way for the GMA for three dimensions. We treat the $d=3$ and $d>3$ cases separately in this section, since the general dimensional algorithm is most easily made clear geometrically by its $d=3$ version, which fortunately we may readily visualize. A reader not interested in the full set of technical details contained in this section should glean the following, before moving to the next section. The GMA algorithm for a set of d integers is the process of subtracting the smallest element from the largest, with additional rules in case of ties. The associated $(d-1)$-dimensional shift is given by Eqs. (3.5)–(3.9) or (3.10)–(3.14).

We proceed now to a full exposition of GMA. The following is a proposition similar to Proposition 1.

Proposition 2. Any integer triplet (P_1, P_2, P_3) with greatest common factor unity and

$$P_1 + P_2 \geqslant P_3; \qquad P_1 \leqslant P_2 \leqslant P_3 \tag{3.1}$$

may be represented as a sum of the columns of a matrix K, which may be written as the product of a string of E_{ij}: $\mathsf{K} = \prod_n \mathsf{E}_{i_n j_n}$. The product is ordered in the following way:

$$\mathsf{K} = \cdots \mathsf{E}_{i_2 j_2} \mathsf{E}_{i_1 j_1} \mathsf{E}_{i_0 j_0}$$

Remark 1. The proposition may be straightforwardly stated in any dimension. The d integers with greatest common factor unity are ordered $1 \leqslant P_1 \leqslant P_2 \leqslant \cdots < P_d$. The sum of the smallest two integers must be greater than or equal to the greatest integer, $P_1 + P_2 \geqslant P_d$.

The GMA, described below, provides precisely one route to the matrix of the type described in Proposition 2 (see below).

Consider the points given by the origin, and the three lattice points each given by an integer triplet (L_i, M_i, N_i); $i = 1, 2, 3$. We call these three points *the vertices*. Consider also the three lattice points given by the mediants of the vertices [i.e., $(L_i + L_j,\ M_i + M_j,\ N_i + N_j)$; $i \neq j$]. We call these points *the mediants* (for short, we write M_{ij} for mediants). These $1 + 3 + 3 = 7$ lattice points, along with the *focus* $(L_1 + L_2 + L_3,\ M_1 + M_2 + M_3,\ N_1 + N_2 + N_3)$, from the 8 corners of a parallelepiped. By designating the three vertices (L_i, M_i, N_i), then, one may construct this parallelepiped in a unique way. According to Proposition 2, there is at least one prescription for reconstructing any given focus under the restrictions (3.1). Moreover, the process of reconstructing any given focus may be viewed as a dynamical system, where the dynamics is that of the motion of one focus to another.

Consider a direction given by a unit vector $(a, b, 1)/(1 + a^2 + b^2)^{1/2}$ with a, b irrational. The GMA is a particular prescription for moving one of two vertices to a particular mediant in such a manner that the ray from the origin stays within the new parallelepiped constructed from the new vertices. (There are other "simplex-splitting" algorithms also.)

Definition 2. *Generalized mediant algorithm* (GMA), $d = 3$. Consider an integer triplet (P_1, P_2, P_3) satisfying the assumptions of Proposition 2: $P_1 \leqslant P_2 \leqslant P_3$; $P_1 + P_2 \geqslant P_3$. The algorithm may be described as follows:

1. We use a superscript to number iterates of the GMA procedure. Define

$$(P_1^{(0)}, P_2^{(0)}, P_3^{(0)}) \equiv (P_1, P_2, P_3)$$

Also define

$$i_{-1} = 1, \qquad j_{-1} = 2$$

This is convenient so that rule 2 which follows holds for all $n \geqslant 0$.

2. Define j_n to be the label which is not a member of the set of the two elements $\{i_{n-1}, j_{n-1}\}$. Define i_n to be the label such that $P_{i_n}^{(n)} < P_p^{(n)}$ for all $p \neq i_n$. That is, i_n labels the smallest element of the set of the three integers for the nth iterate of GMA. If there is no uniquely smallest integer, define $i_n = i_{n-1}$. Now let

$$(P_1^{(n+1)}, P_2^{(n+1)}, P_3^{(n+1)}) = (P_1^{(n)}, P_2^{(n)}, P_3^{(n)}) \cdot \mathsf{E}_{i_n j_n}^{-1}$$

The procedure subtracts the smallest (i_n) from the largest (j_n) entry and records the position of the two entries (as $\mathsf{E}_{i_n j_n}^{-1}$). In the case of ties, additional rules are provided.

3. Repeat this procedure until the triplet is reduced to $(1, 1, 1) = (P_1^{(N)}, P_2^{(N)}, P_3^{(N)})$.

Remark 2. The GMA provides the existence of a matrix K of Proposition 2:

$$\mathsf{K} = \mathsf{E}_{i_{N-1} j_{N-1}} \cdot \mathsf{E}_{i_{N-2} j_{N-2}} \cdots \mathsf{E}_{i_1 j_1} \cdot \mathsf{E}_{i_0 j_0}$$

An example is provided in Fig. 3a. The triplet (7, 22, 23) is reduced via the GMA. One always subtracts a smallest from a largest entry. The matrix K of Proposition 2 is constructed as specified:

$$\mathsf{K}_{(7,22,23)} = \mathsf{E}_{13}\,\mathsf{E}_{12}(\mathsf{E}_{31}\,\mathsf{E}_{32})^3\,\mathsf{E}_{13}(\mathsf{E}_{12}\,\mathsf{E}_{13})^2 \tag{3.2}$$

$$= \begin{bmatrix} 4 & 12 & 13 \\ 0 & 1 & 0 \\ 3 & 9 & 10 \end{bmatrix} \tag{3.3}$$

The sum of the columns of K now gives the desired result (7, 22, 23). The rows of the matrix of Eq. (3.3) are the vertices of the parallelogram which has (7, 22, 23) as its focus. A proof of the statements of Proposition 2 and Definition 2 is given in Appendix A. No subtle points appear in the rather tedious proofs.

Example 2.

A	B	C	$F(x)$	E-string
7	22	23	T_0	E_{13}
7	22	16	T_0	E_{12}
7	15	16	T_0	E_{13}
7	15	9	T_0	E_{12}
7	8	9	T_1	E_{13}
7	8	2	T_0	E_{32}
7	6	2	T_0	E_{31}
5	6	2	T_0	E_{32}
5	4	2	T_0	E_{31}
3	4	2	T_0	E_{32}
3	2	2	T_1	E_{31}
1	2	2	T_0	E_{12}
1	1	2	T_0	E_{13}
1	1	1		

Fig. 3. An example of GMA: reduction of triplet (7, 22, 23).

We give a geometric picture of the algorithm. Consider once again the initial parallelepiped as described below Proposition 2 and a unit vector given by $(a, b, 1)/(1+a^2+b^2)^{1/2}$. At each application of the GMA, only one of the faces is pierced by the ray generated by the unit vector. Every face that is pierced contains a single vertex at a corner. Suppose this corner is V_j. Then (a) E_{ij} means move V_i to M_{ij} and (b) E_{kj} means move V_k to M_{kj}. The correct choice is given by the criterion that either in the (a) case, V_k is the new corner vertex, or in the (b) case, V_i is the new corner vertex. Cases (a) and (b) cannot both be satisfied such that the sum of the two smallest integers of the focus is larger than the third.

In the 2d case of the Farey shift, we saw that associated to the E-string and a Euclidean algorithm there is a shift map, e.g., the Farey shift. Exactly the same method may be used to construct a shift map for the GMA. The reduction of the focus (as in the above example) is given by $(L<M<N)$

$$(L, M, N) \to (L', M', N') = (L, M, N-L)$$

that is, we subtract the smallest from the largest. This Euclidean algorithm is in some sense the generalization of the Farey Euclidean algorithm (see Section 2). Suppose we reorder the matrix after each subraction. Consider the ordered version of the (L, M, N) above: (A, B, C). The associated GMA shift maps $(A/C, B/C)$ to $(A'/C', B'/C')$.

Definition 3. *GMA shift* $(d=3)$. Consider the set $\mathscr{S}$ defined by

$$\mathscr{S} = \{(x, y) \mid 0 < 1-y \leqslant x \leqslant y < 1\} \tag{3.4}$$

Define for $(x, y) \in \mathscr{S}$

$$T_0(x, y) = \left(\frac{x}{y}, \frac{1-x}{y}\right) \tag{3.5}$$

$$T_1(x, y) = \left(\frac{1-x}{y}, \frac{x}{y}\right) \tag{3.6}$$

The GMA shift is defined as

$$T_{\mathrm{GMA}}(x, y) = T_0(x, y) \qquad \text{for} \quad 0 < x \leqslant 1/2 \tag{3.7}$$

$$= T_1(x, y) \qquad \text{for} \quad 1/2 < x < 1 \tag{3.8}$$

In Fig. 4, we have drawn the invariant set $\mathscr{S}$. On the left (Fig. 4a), we have drawn the set $0 \leqslant 1-y \leqslant x \leqslant y$ and divided it into two regions, $x \leqslant 1/2$ and $x \geqslant 1/2$, with equal area. We have also labeled five line segments in the diagram which delimit the boundaries of each piece. On the right, we have

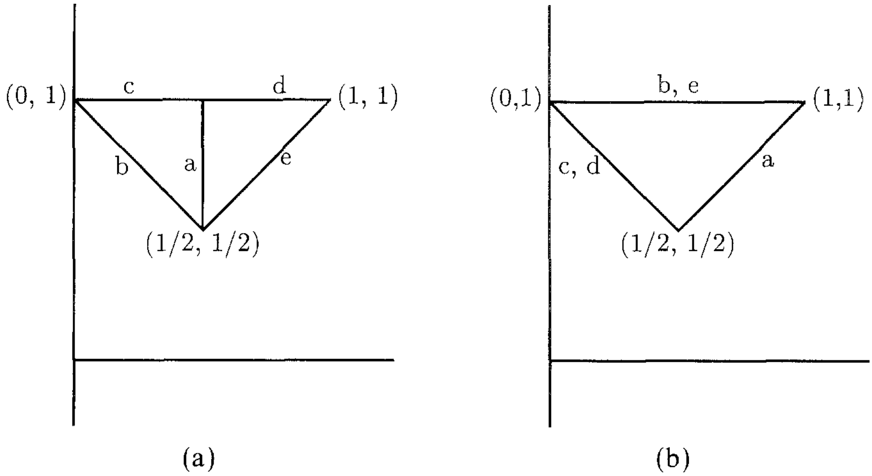

Fig. 4. The invariant set of the GMA shift for two irrationals. (a) The set $\mathscr{S}$ is mapped two to one to itself. (b) See the discussion below Definition 3.

iterated the set $\mathscr{S}$ and shown where each of the line segments from Fig. 4a are mapped to. For example, the line segment c connects the two points $(0, 1)$ and $(1/2, 1)$. It is mapped to the line segment joining $(0, 1)$ and $(1/2, 1/2)$.

The extension of these procedures for dimensions higher than 3 is straightforward.

Definition 4. *GMA Procedure*, $d > 3$. Given a d-integer multiplet of positive integers $(P_1, P_2, \ldots, P_d)$ satisfying the assumptions of Remark 1: $P_1 \leqslant \cdots \leqslant P_d$; $P_1 + P_2 \geqslant P_d$.

1. Again we use superscripts to number iterates of the GMA procedure. Define

$$P_j^{(0)} \equiv P_j \qquad \text{for} \quad 1 \leqslant j \leqslant d$$

Also define

$$i_{-l} = 1, \qquad j_{-l} = l + 1 \qquad \text{for} \quad 1 \leqslant l \leqslant d - 2$$

2. Define j_n to be the label which is not a member of the set of elements $\{i_{n-l}, j_{n-l}\}^{1 \leqslant l \leqslant d-2}$. Define i_n to be the label such that $P_{i_n}^{(n)} < P_l^{(n)}$ for all $l \neq i_n$. That is, i_n labels the smallest element of the set of integers for the nth iterate of GMA. If there is no uniquely smallest integer, define $i_n = i_{n-1}$. Now let

$$(P_1^{(n+1)}, \ldots, P_d^{(n+1)}) = (P_1^{(n)}, \ldots, P_d^{(n)}) \cdot \mathsf{E}_{i_n j_n}^{-1}$$

The procedure subtracts a smallest (labeled by i_n) from a largest (labeled by j_n) entry and records the position of the two entries in the E-string.

3. Repeat this procedure until the triplet is reduced to $(1,..., 1) = (P_1^{(N)},..., P_d^{(N)})$.

The straightforward proofs for the above statements may be found in Appendix B.

As for the $d=3$ case, there exists an associated shift.

Definition 5. *GMA shift* ($d>3$). Define the set

$$\mathscr{S} = \{(a, b, c,..., z) \mid 0 < 1-b \leqslant a \leqslant b \leqslant c \leqslant \cdots z < 1\} \tag{3.9}$$

for I irrationals. Define

$$T_0(a, b, c,..., z) = \left(\frac{a}{z}, \frac{1-a}{z}, \frac{b}{z},..., \frac{y}{z}\right) \tag{3.10}$$

$$T_1(a, b, c,..., z) = \left(\frac{1-a}{z}, \frac{a}{z}\frac{b}{z},..., \frac{y}{z}\right) \tag{3.11}$$

The GMA shift is defined as

$$T_{\text{GMA}}(a,...) = T_0(a,...) \qquad \text{for} \quad 0 < a \leqslant 1/2, \tag{3.12}$$

$$= T_1(a,...) \qquad \text{for} \quad 1/2 < a \leqslant 1 \tag{3.13}$$

What is the geometrical picture for dimension greater than 3? A generalized parallelepiped in d dimensions must have 2^d vertices (e.g., for $d=2$, we have $2^2=4$). For $d=4$ with 4 vertices we may define $4 \cdot 3/1 \cdot 2 = 6$ mediants, precisely as before. We must, however, define also $4 \cdot 3 \cdot 2/1 \cdot 2 \cdot 3 = 4$ mediants$_2$ as the sum of any three vertices. The focus, as before, will be the sum of all 4 vertices. Including the origin, vertices, mediants, mediants$_2$, and focus, we have a total of $1+4+6+4+1=16=2^4$ points, enough to define a generalized parallelepiped. In d dimensions, there will be one focus, d vertices, $[{d \atop 2}]$ mediants,..., $[{d \atop d-1}]$ mediants$_{d-1}$, yielding $\sum_{j=0}^{d} [{d \atop j}] = 2^d$ corners of a parallelepiped. The GMA moves a vertex to a mediant and reconstructs the parallelepiped by the above procedure. The rule for the vertex move reads precisely the same as the three-dimensional case (see below Remark 2).

4. RELATION BETWEEN EIGENVALUES OF E-STRING AND GMA SHIFT

In this section we relate the eigenvalues of the E-string to the eigenvalues of the GMA shift. What we would like to claim is that if we start with an initial multiplet of d arbitrarily large integers, then the eigenvalues of the GMA Euclidean algorithm satisfy a simple relation to the eigenvalues of the GMA shift [see (4.23)–(4.25)]. This is a property common to different simplex splitting algorithms of which the GMA is an example.

Suppose we have a triplet $d=3$ of integers $(N_1^{(0)}, N_2^{(0)}, N_3^{(0)})$ with $N_1 < N_2 < N_3$; $N_1 + N_2 \geqslant N_3$, which is reduced to another triplet of integers $(R_1^{(L)}, R_2^{(L)}, R_3^{(L)})$ after L steps of the GMA:

$$(R_1^{(L)}, R_2^{(L)}, R_3^{(L)}) = (N_1^{(0)}, N_2^{(0)}, N_3^{(0)})\, \mathsf{E}_{i_0 j_0}^{-1} \cdots \mathsf{E}_{i_{L-1} j_{L-1}}^{-1} \tag{4.1}$$

$$= (N_1^{(0)}, N_2^{(0)}, N_3^{(0)})\, \mathsf{K}_L^{-1} \tag{4.2}$$

Let us call $(N_1^{(L)}, N_2^{(L)}, N_3^{(L)})$ the ordered version of $(R_1^{(L)}, R_2^{(L)}, R_3^{(L)})$, so that $N_1^{(L)} \leqslant N_2^{(L)} \leqslant N_3^{(L)}$. Then we can rewrite (4.2) in the following way:

$$\begin{bmatrix} N_1^{(L)} \\ N_2^{(L)} \\ N_3^{(L)} \end{bmatrix} = \mathsf{S}_L \begin{bmatrix} N_1^{(0)} \\ N_2^{(0)} \\ N_3^{(0)} \end{bmatrix} \tag{4.3}$$

Here $\mathsf{S}_L = \mathbf{Perm}(\mathsf{K}_L^{-1})^T$, **Perm** is some permutation matrix, the superscript T denotes the matrix transpose, and K_L is the E-string $\mathsf{K}_L = \mathsf{E}_{i_{L-1} j_{L-1}} \cdots \mathsf{E}_{i_0 j_0}$.

Let us examine what behavior is induced on the shift. We define

$$x_0 = \frac{N_1^{(0)}}{N_3^{(0)}}, \qquad y_0 = \frac{N_2^{(0)}}{N_3^{(0)}}; \qquad x_L = \frac{N_1^{(L)}}{N_3^{(L)}}, \qquad y_L = \frac{N_2^{(L)}}{N_3^{(L)}} \tag{4.4}$$

Now x_L, y_L are simply the Lth iterates of the GMA *shift* on x_0, y_0. So we can define

$$(x_L, y_L) = T_{\mathrm{GMA}}^L(x_0, y_0) \equiv T_{\mathrm{GMA}} \circ \cdots \circ T_{\mathrm{GMA}}(x_0, y_0) \tag{4.5}$$

Thus we may also define the following matrix:

$$\mathsf{T}_L = \begin{bmatrix} \partial x_L/\partial x_0 & \partial x_L/\partial y_0 \\ \partial y_L/\partial x_0 & \partial y_L/\partial y_0 \end{bmatrix} \tag{4.6}$$

Clearly

$$\mathsf{T}_L = \mathsf{T}_{L,L-1} \mathsf{T}_{L-1,L-2} \cdots \mathsf{T}_{1,0} \tag{4.7}$$

where $\mathsf{T}_{j,j-1}$ is the Jacobian of the mapping from (x_{j-1}, y_{j-1}) to (x_j, y_j):

$$\mathsf{T}_{j,j-1} = \begin{bmatrix} \partial x_j/\partial x_{j-1} & \partial x_j/\partial y_{j-1} \\ \partial y_j/\partial x_{j-1} & \partial y_L/\partial y_{j-1} \end{bmatrix} \tag{4.8}$$

We would like to define "average eigenvalues" for the Euclidean algorithm and shift. Unfortunately, eigenvalues of $\mathsf{S}_L, \mathsf{T}_L$ may be complex or negative, so that taking the Lth roots of $\mathsf{S}_L, \mathsf{T}_L$ may be an ill-defined procedure. The idea carried through by Oseledec[20] is to multiply $\mathsf{S}_L, \mathsf{T}_L$ by their transposes to get positive, real, symmetric matrices. From those matrices one may define square roots, etc., in a unique way to arrive at new positive, real, symmetric matrices. Thus one defines

$$\mathsf{S} \equiv \lim_{L \to \infty} (\mathsf{S}_L^T \mathsf{S}_L)^{1/2L} \tag{4.9}$$

$$\mathsf{T} \equiv \lim_{L \to \infty} (\mathsf{T}_L^T \mathsf{T}_L)^{1/2L} \tag{4.10}$$

One also defines

$$\mathsf{E} \equiv \mathsf{S}^{-1} \tag{4.11}$$

$$= \lim_{L \to \infty} (\mathsf{K}_L^T \mathsf{K}_L)^{1/2L} \tag{4.12}$$

$$= \lim_{L \to \infty} (\mathsf{E}_{i_0 j_0}^T \cdots \mathsf{E}_{i_{L-1} j_{L-1}}^T \mathsf{E}_{i_{L-1} j_{L-1}} \cdots \mathsf{E}_{i_0 j_0})^{1/2L} \tag{4.13}$$

In deriving (4.12), we have used the definition of the matrix $\mathsf{S}_L = \mathbf{Perm}(\mathsf{K}_L^{-1})^T$, and the fact that permutation matrices are orthogonal $\mathbf{Perm}^T\, \mathbf{Perm} = \mathsf{I}$.

Because the shift map is ergodic and satisfies the other criterion of the Oseledec theorem,[20] we should expect that the eigenvalues of E and T (with probability one) do not depend on which set of irrationals are being approximated. That is, the set of real multiplets which are exceptions have measure zero relative to the absolutely continuous invariant measure (which we calculate in Section 5). Thus, it is meaningful to discuss these matrices without reference to any particular starting vectors.

The eigenvalues of E describe the dynamics of the parallelogram as it is sheared along some "irrational" vector as we described earlier. By the way GMA is defined we expect that there should be one eigenvalue greater than unity describing the stretching of the parallelogram, with all the other eigenvalues less than unity describing the contraction. The parallelogram becomes long and thin (as in Fig. 1 for the 2d case). The eigenvalues of T, on the other hand, describe the expansive nature of the shift map and so we expect that all of these eigenvalues should be larger than unity (just as the OCF shift is expansive). The question we address is: how are the eigenvalues of E, the E-string, related to the eigenvalues of T, the shift?

What we find is that if $\lambda_1 > 1 > \lambda_2 \geqslant \lambda_3$ are the eigenvalues of E, the eigenvalues of T are given by

$$\sigma_1 = \lambda_1/\lambda_3, \qquad \sigma_2 = \lambda_1/\lambda_2, \qquad \sigma_1 \geqslant \sigma_2 > 1 \tag{4.14}$$

Equations (4.14) are very important if one wishes to calculate the entropy. The result tells us that the eigenvalues of the shift are all indeed greater than one and that all the Lyapunov exponents are positive. The entropy of the shift is then simply the sum of the positive exponents, which is in turn the sum of *all* the exponents:

$$h = \ln \sigma_1 + \ln \sigma_2 \tag{4.15}$$

(note that σ_1 is an eigenvalue, $\ln \sigma_1$ is an exponent). Thus we calculate, using (4.14) and (4.15),

$$h = \ln \frac{\lambda_1}{\lambda_3} + \ln \frac{\lambda_1}{\lambda_2} = \ln \frac{\lambda_1^2}{\lambda_2 \lambda_3} = \ln \frac{\lambda_1^3}{\lambda_1 \lambda_2 \lambda_3} = \ln \lambda_1^3 = 3 \ln \lambda_1 \tag{4.16}$$

In the above equation we have used

$$\lambda_1 \lambda_2 \lambda_3 = 1 \tag{4.17}$$

This holds because

$$(\det \mathsf{E})^L = \prod_{k=0}^{L} (\det \mathsf{E}_{i_k j_k}) = \prod_{k=0}^{L} 1 = 1 \tag{4.18}$$

$$\det \mathsf{E} = 1 \tag{4.19}$$

Thus, since each $\mathsf{E}_{i_k j_k}$ has determinant one, then E also has unit determinant, and the product of the eigenvalues is also unity.

Also from (4.15)

$$h = \ln(\sigma_1 \sigma_2) = \ln \det \mathsf{T} = \lim_{L \to \infty} \frac{1}{L} \ln \det \mathsf{T}_L = \lim_{L \to \infty} \frac{1}{L} \sum_{j=1}^{L} \ln \det \mathsf{T}_{j,j-1} \tag{4.20}$$

Using the Birkhoff theorem, we see that if we can calculate the absolutely continuous invariant measure, then we can calculate the entropy analytically: Eq. (4.20) converges to an integral, which we can perform over the phase space,

$$h = \int d\mu(x) \ln \det \mathsf{J}(x) \tag{4.21}$$

where J is the Jacobian. Combining (4.16) and (4.21) gives an expression for the largest eigenvalue of the Euclidean algorithm in terms of an integral:

$$\ln \lambda_1 = \frac{1}{3} \int d\mu(x) \ln \det \mathsf{J}(x) \tag{4.22}$$

If the shift map had some negative exponents and some positive exponents, then one could never proceed through an argument like (4.20) to arrive at an expression for the entropy in terms of an integral over phase space. So to give an expression for the eigenvalues of the shift as in (4.14), showing that they are all greater than one, is very important.

General Dimensions. The above discussion is similar for any dimension.

Theorem 1. If

$$\lambda_1 > 1 > \lambda_2 \geqslant \cdots \geqslant \lambda_d \tag{4.23}$$

$$\lambda_1 \cdots \lambda_d = 1 \tag{4.24}$$

are the eigenvalues of E, then the eigenvalues of T, the shift, are given by

$$\sigma_1 = \lambda_1/\lambda_d, \qquad \sigma_2 = \lambda_1/\lambda_{d-1},..., \qquad \sigma_{d-1} = \lambda_1/\lambda_2 \tag{4.25}$$

This yields for the entropy

$$h = \sum_{j=1}^{d-1} \ln \frac{\lambda_1}{\lambda_{d-j+2}} = \ln \frac{\lambda_1^{d-1}}{\lambda_d \cdots \lambda_2} = \ln \frac{\lambda_1^d}{\lambda_d \cdots \lambda_1} = \ln \lambda_1^d = d \ln \lambda_1 \tag{4.26}$$

or

$$h = \ln \det \mathsf{T} = \lim_{L \to \infty} \frac{1}{L} \ln \det \mathsf{T}_L = \frac{1}{L} \sum_{j=1}^{L} \lim_{L \to \infty} \ln \det \mathsf{T}_{j,j-1} \tag{4.27}$$

Again we have arrived at a relation for λ_1, the growth rate of the denominators, in terms of the entropy h,[8] via Eq. (4.26):

$$\lambda_1 = e^{h/d} \tag{4.28}$$

where h can be evaluated from converting Eq. (4.27) to an integral using the Birkhoff theorem,

$$h = \int d\mu(x) \ln \det \mathsf{J}(x) \tag{4.29}$$

It is Eq. (4.29) that we will evaluate in the next section. Combining (4.26) and (4.29) yields

$$\ln \lambda_1 = \frac{1}{d} \int d\mu(x) \ln \det \mathsf{J}(x) \tag{4.30}$$

More detailed arguments for the above statements are given in Appendix C, where the relation (4.25) is derived from (4.23). In order to implement the proofs of Appendix C, we need to assume (4.23), that there is a single eigenvalue λ_1 strictly greater than all the others. The way that GMA is constructed, it is clear that this should be so, as we have already discussed in this section. Moreover, numerical calculations[23] convincingly support (4.23). Strictly speaking, however, we have provided no analytical proof that (4.23) holds. The author believes that no such proof exists for (4.23) for any MCFA and that numerical proofs may be needed.

The proofs in Appendix C pertain to a class of continued-fraction algorithms, and so one hopes are of general interest. They should be interesting to a wide audience in the field of MCFAs. For this paper we shall focus on particular characteristics of the GMA.

5. STATISTICAL PROPERTIES

In Section 3, we introduced the GMA algorithm and the GMA shift. The results of Section 4 related the eigenvalues of the GMA shift to the eigenvalues of the E-string. Due to the way the E-string of the GMA shift is constructed, it was clear that all the eigenvalues of the E-string but one lie inside the unit circle. Due to the relation (4.25), then, we know that all the eigenvalues of the shift lie outside the unit circle. This allows us simply to find the KS entropy by taking the determinant of the Jacobian of the map as in (4.29). Remarkably, we are able to express the entropy of the shift as a single integral for the approximation of any number of irrationals I.

5.1. Invariant Measure for GMA Shift in Dimension $\geqslant 3$

Three-Dimensional Case. From Eqs. (3.5)–(3.9), we note that after one application of T_{GMA} on a point $(x, y) \in \mathcal{S}$ we have

$$x' + y' = \frac{1}{y} \geqslant 1 \qquad \text{and} \qquad x' \leqslant y' < 1 \tag{5.1}$$

Thus, T_{GMA} is a mapping of $\mathcal{S}$ to itself. For $d = 3$, the mapping restricted to the invariant set $\mathcal{S}$ reads

$$\begin{aligned} &\left(x < \frac{1}{2}\right) \qquad (x', y') = \left(\frac{x}{y}, \frac{1-x}{y}\right) \\ &\left(x \geqslant \frac{1}{2}\right) \qquad (x', y') = \left(\frac{1-x}{y}, \frac{x}{y}\right) \end{aligned} \tag{5.2}$$

The invariant density $\rho(x) = d\mu(x)/dx$ satisfies the Perron–Frobenius equation:

$$\rho(x, y) = \sum_{T(r,s)=(x,y)} \rho(r, s)\, |\det J(r, s \mid x, y)| \tag{5.3}$$

where

$$\det J(r, s \mid x, y) = \frac{\partial r}{\partial x}\frac{\partial s}{\partial y} - \frac{\partial r}{\partial y}\frac{\partial s}{\partial x} \tag{5.4}$$

which yields the following equation for the invariant measure on $\mathscr{S}$:

$$\rho(x, y) = \frac{1}{(x+y)^3}\left(\rho\left(\frac{y}{x+y}, \frac{1}{x+y}\right) + \rho\left(\frac{x}{x+y}, \frac{1}{x+y}\right)\right) \tag{5.5}$$

This has as a solution

$$\rho(x, y) = \frac{1}{\text{norm}}\frac{1}{xy} \tag{5.6}$$

General Dimensions. The solution for higher numbers of irrationals is very much the same. The invariant set $\mathscr{S}$ is defined by

$$1 \geqslant z \geqslant y \geqslant \cdots \geqslant b \geqslant a \geqslant 1 - b \tag{5.7}$$

for the $I = d - 1$ irrationals $(a, b,..., z)$. On the invariant set the map reads

$$\begin{aligned}(a', b', c', d',..., y', z') &= \left(\frac{a}{z}, \frac{1-a}{z}, \frac{b}{z}, \frac{c}{z},..., \frac{x}{z}, \frac{y}{z}\right)\\ &\text{or } \left(\frac{1-a}{z}, \frac{a}{z}, \frac{b}{z}, \frac{c}{z},..., \frac{x}{z}, \frac{y}{z}\right)\end{aligned} \tag{5.8}$$

depending on the relative magnitude of a and $1-a$. The analogue of Eq. (5.4) reads

$$\begin{aligned}\rho(a,..., y, z) = \frac{1}{(a+b)^{I+1}}\bigg(&\rho\left(\frac{a}{a+b}, \frac{c}{a+b},..., \frac{y}{a+b}, \frac{1}{a+b}\right)\\ &+ \rho\left(\frac{b}{a+b}, \frac{c}{a+b},..., \frac{y}{a+b}, \frac{1}{a+b}\right)\bigg)\end{aligned} \tag{5.9}$$

This is solved by

$$\rho(a, b,..., y, z) = \frac{1}{ab \cdots yz} \tag{5.10}$$

as can be easily verified. The invariant measure is then given by

$$d\mu(a, b, \ldots, y, z) = \frac{1}{\text{norm}(I)} \frac{da}{a} \cdots \frac{dz}{z} \tag{5.11}$$

where "norm" is the normalization and the integration region is given by (5.7),

$$\text{norm}(I) = \int^{\mathscr{S}} \frac{da}{a} \cdots \frac{dz}{z} \tag{5.12}$$

We will consider the normalization of the invariant density below.

Normalization. We consider the case of I irrationals. The last $I-2$ of the integrals over the integration region $\mathscr{S}$ [Eq. (5.7)] are trivial to perform, yielding for the normalization

$$\text{norm}(I) = \int \frac{da}{a} \frac{db}{b} \frac{[\ln(1/b)]^{I-2}}{(I-2)!}; \qquad 1 - b < a < b < 1 \tag{5.13}$$

Next the a integral is performed and the substitution $b = 1/(1 + b')$ is made, yielding

$$\text{norm}(I) = \int_0^1 \frac{db'}{1+b'} \frac{[\ln(1+b')]^{I-2}}{(I-2)!} \ln \frac{1}{b'} \tag{5.14}$$

An integration by parts yields

$$\text{norm}(I) = \int_0^1 \frac{db'}{b'} \frac{[\ln(1+b')]^{I-1}}{(I-1)!} \tag{5.15}$$

The formula for the measure given by (5.11) is now complete.

5.2. KOLMOGOROV–SINAI ENTROPY OF THE GMA SHIFT

The Kolmogorov–Sinai entropy may be heuristically interpreted as the loss in information per iteration of the map. The following formula is exact under certain technical restrictions of the map[18]:

$$h = \int d\mu \ln \det J_+ \tag{5.16}$$

where J_+ is the Jacobian of the expanding subspace of the map. In Section 4, we concluded that all the eigenvalues of the shift map lay outside the unit circle. Thus we may equate $J_+ \equiv J$. Remarkably, we may explicitly evaluate the entropy analytically. This is one pleasant feature of this map.

Jacobian. Here we find the Jacobian for the mapping given by Eq. (5.8). Note in (5.8) that only one primed variable is dependent on $b, c,..., y$, etc. This observation allows us to write

$$\begin{aligned} J &= \left| \frac{\partial(a', b', c',..., y', z')}{\partial(a, b, c,..., y, z)} \right| \\ &= \left| \frac{\partial c'}{\partial b} \cdots \frac{\partial z'}{\partial y} \right| \left| \frac{\partial(a', b')}{\partial(a, z)} \right| = \frac{1}{z^{I-2}} \left| \frac{\partial(a', b')}{\partial(a, z)} \right| \end{aligned} \tag{5.17}$$

From this we can directly calculate

$$J = \frac{1}{z^{I+1}} \tag{5.18}$$

where I is the number of irrationals.

Entropy. With the formula for the Jacobian (5.17) and the formula for the measure (5.11) and (5.12) we can calculate the entropy:

$$\begin{aligned} h &= \int \ln |J| \, d\mu \\ &= (I+1) \int^{\mathscr{S}} \ln \frac{1}{z} \, d\mu \end{aligned} \tag{5.19}$$

with the integration region given by (5.7). The integrals are performed in exactly the same order as in as the last subsection. Considering Eqs. (5.11), (5.12), and (5.19), it follows that

$$\begin{aligned} h(I) &= (I+1) \cdot \frac{\operatorname{norm}(I+1)}{\operatorname{norm}(I)} \\ &= \frac{(I+1)\, G(I)}{(I)\, G(I-1)} \end{aligned} \tag{5.20}$$

where

$$G(I) = \int_0^1 \frac{ds}{s} [\ln(1+s)]^I \tag{5.21}$$

The $I \to 1$ Limit. It is interesting to consider noninteger values of the RHS of (5.20). In that vein, we show

$$\lim_{\varepsilon \to 0} \varepsilon G(\varepsilon) = 1 \tag{5.22}$$

By definition,

$$G(\varepsilon)=\int_0^1 \frac{ds}{s}\,[\ln(1+s)]^{\varepsilon}$$

$$=\int_0^1 \frac{ds}{s^{1-\varepsilon}}\,[K(s)]^{\varepsilon} \tag{5.23}$$

Here

$$K(s)=\frac{\ln(1+s)}{s} \tag{5.24}$$

is a bounded function of s on the unit interval. Now we can expand $[K(s)]^{\varepsilon}=1+\varepsilon \ln K(s)+\mathcal{O}(\varepsilon^2)$ to find

$$G(\varepsilon)=\int_0^1 \frac{ds}{s^{1-\varepsilon}}+\varepsilon\int_0^1 \frac{ds}{s^{1-\varepsilon}} \ln K(s)+\mathcal{O}(\varepsilon^2) \tag{5.25}$$

Now we may integrate the first term exactly. Moreover, the second integrand is well behaved, since, as $s\to 0$, $[\ln K(s)]/s\to -1/2$. So

$$G(\varepsilon)=\frac{1}{\varepsilon}+\varepsilon\int_0^1 ds\,\frac{\ln K(s)}{s}+\mathcal{O}(\varepsilon^2) \tag{5.26}$$

Equation (5.22) follows directly.

Thus, if we consider I to be a real variable, then

$$\lim_{I\to 1}\frac{h(I)}{(I-1)\ln 2}=2\,\frac{G(1)}{\ln 2}=\frac{\pi^2}{6\ln 2}=h_{\mathrm{OCF}} \tag{5.27}$$

We mention the above result for the following reason. The GMA gives a way for studying certain properties of the Farey shift, and the author speculates that results such as (2.11) may be recovered from studying the above limit.

Large-I Limit. In this section, we establish the result for large I. We have

$$h=\ln 2-\frac{(\ln 2)^2}{I^2}+\mathcal{O}\left(\frac{1}{I^3}\right) \tag{5.28}$$

Our analytical answer has been verified numerically and been found to be in good agreement with the analytical form for $I<150$. We begin by evaluating the integral in (5.21), $G(I)$, by parts to yield

$$G(I)\doteq\frac{1}{I+1}\left\{2(\ln 2)^{I+1}+\int_0^1 \frac{ds}{s^2}\,[\ln(1+s)]^{I+1}\right\} \tag{5.29}$$

Thus, from (5.21) and (5.29)

$$h(I) = \ln 2 \, \frac{1 + H(I+1)}{1 + H(I)} \tag{5.30}$$

$$H(p) = \frac{1}{2} \int_0^1 \frac{ds}{s^2} \left(\frac{\ln(1+s)}{\ln 2} \right)^p \tag{5.31}$$

Now the quantity $H(p)$ approaches 0 uniformly as p approaches ∞. Thus, we may write, to lowest order,

$$h(I) = \ln 2 + \ln 2 \, [H(I+1) - H(I)] \tag{5.32}$$

Now the integral occurring in H is dominated by s near 1. To lowest order, then, we replace the quantity $1/(2s^2)$ by $1/(1+s)$. Then the integral may be performed exactly, yielding

$$H(p) = \frac{\ln 2}{p} + \mathcal{O}\left(\frac{1}{p^2} \right) \tag{5.33}$$

Equation (5.28) follows immediately from the last two equations.

The largest eigenvalue of the E-string is related to the entropy by[3]

$$\lambda_1 = \exp\left(\frac{h}{d} \right) \tag{5.34}$$

$$= \exp\left(\frac{G(I)}{(I) \, G(I-1)} \right) \tag{5.35}$$

The largest eigenvalue gives valuable information regarding the Diophantine properties of the Euclidean algorithm. A full discussion of all the eigenvalues will be deferred to a later time.

6. DISCUSSION, CONCLUSION

In this paper we have investigated a multidimensional continued-fraction algorithm completely from a dynamical systems perspective. The attractive feature of this algorithm is the fact that we may explicitly calculate its statistical properties. In Section 5, we calculated the observable invariant measure and the KS entropy for the algorithm. Remarkably, we find that we can reduce the expression for the entropy (for the approximation of any numbers of irrationals) to a single integral. This yields in turn an expression for the growth rate of denominators as discussed in Section 4.

In Figs. 5a and 5b we have plotted the entropy for arbitrary dimension. Although we may discuss the entropy inherent in the approximation of a set of I irrationals only when I is an integer, it is still interesting to consider the functional dependence of the expression for the entropy [as given by (5.20) and (5.21)] as we *continuously* vary I. In Fig. 5a, as the number of irrationals increases, the entropy rises to an asymptotic value. In Fig. 5b, we have divided the entropy of the GMA shift by the number of integers which are unchanged at each application of the GMA procedure of Definition 2 (there are $d-2=I-1$ of these integers). For large I, this curve shows a $1/I$ dependence. As one lowers I, one sees that the entropy of the ordinary continued fraction lies on the analytic continuation of this curve to $I=1$. This intriguing feature will be discussed at length at a later time.

The entropy is directly related to the growth rate of denominators by

$$\lambda_1 = e^{h/(I+1)}$$

where h is the KS entropy, I is the number of irrationals to be approximated, and λ_1 is the eigenvalue governing the growth rate of denominators. This quantity has never been calculated for any MCFA. Our results for the entropy for large I yield that

$$\lambda_1 = 2^{1/(I+1)} \tag{6.1}$$

Since we should expect that the GMA shift is Bernoulli (the sequence of E_{ij} in the E-string is asymptotically as random as a coin toss), it may also be interesting to compare our results on the GMA shift to the recent results on the Lyapunov spectrum of products of random matrices.[21] Since we can explicitly write down the invariant measure for the GMA case, it might be interesting to compare our rigorous results for the eigenvalues of the GMA shift with those found for the random matrix case (where the distribution is sometimes constructed somewhat artificially).

In this paper we have simply examined the behavior of the determinant of the Jacobian. That is, the product of these eigenvalues. In a following work we provide a calculation and dynamical systems interpretation for an important number-theoretic convergence exponent. An MCFA with best approximation properties *must* have a certain value for this exponent. Thus, this quantity enables us to quantify the quality of the approximation properties of any given MCFA, and specifically enables us to compare the approximation properties of the GMA and JP algorithms.

Near the completion of this work, the work of Brun and Selmer (22) on essentially the same algorithm were made known to the author. No calculation for the eigenvalues has ever been given.

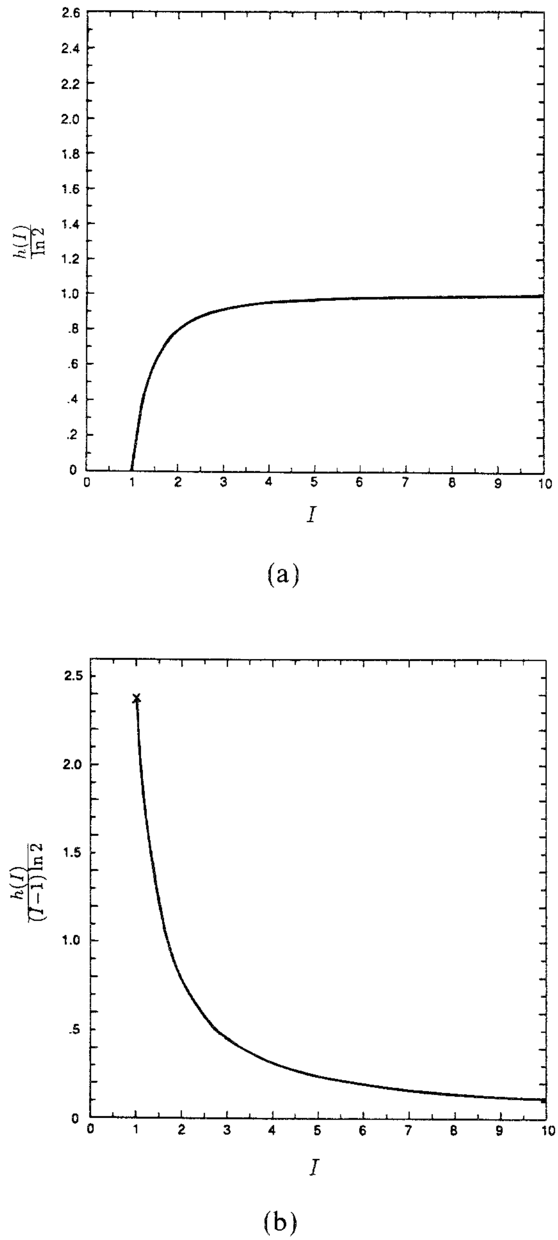

Fig. 5. Two entropy plots. (a) Plot of $h(I)/\ln 2$ vs. I, where I is the number of irrationals and $h(I)$ is the analytic continuation of the entropy expression for GMA as a function of I. The function $h(I)/\ln 2$ approaches unity for I large. (b) Plot of $h(I)/[(I-1)\ln 2]$ vs. I. As I approaches unity, the ordinate approaches h_{OCF}. As I grows large, the ordinate goes as $1/I$.

APPENDIX A. PROOF OF PROPOSITION 2, STATEMENTS IN DEFINITION 2

In this Appendix we prove some of the assertions in Definition 2 and show how the GMA provides a proof for Proposition 2. None of these proofs involve sophisticated ideas. One must only do some careful bookkeeping.

We first show that the conditions imposed on the initial element of the GMA are passed on along to each iterate. That is, for every iteration of the GMA, the following hold [(A.1)–(A.6)]:

$$i_{n-1} \neq j_{n-1} \tag{A.1}$$

That makes the choice for j_n, which is defined to be the number between 1 and 3 inclusive not in the set $\{i_{n-1}, j_{n-1}\}$ unique. Also,

$$i_n \in \{i_{n-1}, j_{n-1}\} \tag{A.2}$$

Now define k_n as

$$k_n \in \{i_{n-1}, j_{n-1}\} \qquad \text{and} \qquad k_n \neq i_n$$

This is meaningful due to (A.2) [see also (A.4) below]. We also define $k_{-1} = 3$ for convenience. We also want to show

$$j_n = k_{n-1} \tag{A.3}$$

$$i_n, j_n, k_n \text{ are 3 distinct integers} \tag{A.4}$$

$$P_{i_n}^{(n)} \leqslant P_{k_n}^{(n)} \leqslant P_{j_n}^{(n)} \tag{A.5}$$

$$P_{i_n}^{(n)} + P_{k_n}^{(n)} \geqslant P_{j_n}^{(n)} \tag{A.6}$$

In the above, i_n is defined by the GMA as the label for the uniquely smallest entry of the nth iterate of the GMA, or i_{n-1} when there is no uniquely smallest element. Some of these statements are rather redundant [for example, (A.3) implies (A.1)]; however, it is convenient to set things up this way so that the proof for $d \geqslant 4$ is as similar as possible. The most important observation is (A.5) and (A.6), the sum of the two smallest elements is always larger than or equal to the greatest, for any iteration of the GMA.

Proof by Induction $n = 0$. Since $i_{-1} = 1$, $j_{-1} = 2$, (A.1) is verified, and we have $j_0 = 3$, the complement to the set $\{i_{-1}, j_{-1}\}$. Since $j_0 = 3 = k_{-1}$, (A.3) is verified.

We are given that $P_1^{(0)} \leqslant P_2^{(0)} \leqslant P_3^{(0)}$. If $P_1^{(0)} < P_2^{(0)}$, then $P_1^{(0)}$ is the smallest element and $i_0 = 1 = i_{-1}$: the smallest entry is in the first position.

If $P_1^{(0)} = P_2^{(0)}$, then $i_0 = i_{-1} = 1$ according to GMA. Thus $i_0 = 1 = i_{-1}$, verifying (A.2), and we calculate $k_0 = 2$. Thus, $i_0 = 1$, $j_0 = 3$, $k_0 = 2$, and (A.4) is verified. The fact that $P_{i_0}^{(0)} \leqslant P_{j_0}^{(0)} \leqslant P_{k_0}^{(0)}$ verifies Eq. (A.5). Also, due to the given $P_{i_0}^{(0)} + P_{j_0}^{(0)} \geqslant P_{k_0}^{(0)}$, verifying Eq. (A.6).

$n \geqslant 1$. We assume (A.1)–(A.6) true for all $0 \leqslant r \leqslant n$, we wish to show (A.1)–(A.6) hold for $n+1$. Now (A.4) tells us that i_n, j_n, and k_n are three distinct integers from 1 to 3 [so (A.1) is automatically implied for $n+1$]. By the GMA prescription

$$(P_1^{n+1}, P_2^{n+1}, P_3^{n+1}) = (P_1^n, P_2^n, P_3^n)\, \mathsf{E}_{i_n j_n}^{-1} \tag{A.7}$$

yielding

$$\begin{aligned} P_{i_n}^{(n+1)} &= P_{i_n}^{(n)} \\ P_{k_n}^{(n+1)} &= P_{k_n}^{(n)} \\ P_{j_n}^{(n+1)} &= P_{j_n}^{(n)} - P_{i_n}^{(n)} \end{aligned} \tag{A.8}$$

We have subtracted one of the smallest entries [see (A.5)] from one of the largest entries. Since i_n and j_n are different, then we can uniquely define $j_{n+1} = k_n$ [verifies (A.3)].

Case 1. Suppose

$$2P_{i_n}^{(n)} \leqslant P_{j_n}^{(n)} \tag{A.9}$$

Then

$$P_{i_n}^{(n+1)} = P_{i_n}^{(n)} \leqslant P_{j_n}^{(n)} - P_{i_n}^{(n)} = P_{j_n}^{(n+1)} \leqslant P_{k_n}^{(n)} = P_{k_n}^{(n+1)} \tag{A.10}$$

We have used in order Eqs. (A.8), (A.9), (A.8), (A.5), and (A.8) to achieve respectively each relation in (A.10). Since $P_{i_n}^{(n+1)}$ is either the uniquely smallest element or is one of the smallest elements of the $(n+1)$th iterate, the GMA ensures that $i_{n+1} = i_n$. So we have $i_{n+1} = i_n$; $j_{n+1} = k_n$. Since i_n and k_n are different, then i_{n+1} and j_{n+1} are different. Thus we can uniquely define a $k_{n+1} = j_n$. Thus (A.2) and (A.4) are satisfied for Case 1.

Case 2. Suppose

$$2P_{i_n}^{(n)} > P_{j_n}^{n} \tag{A.11}$$

Then $i_{n+1} = j_n$, $j_{n+1} = k_n$, and $k_{n+1} = i_n$, since

$$P_{j_n}^{(n+1)} = P_{j_n}^{(n)} - P_{i_n}^{(n)} < P_{i_n}^{(n)} = P_{i_n}^{(n+1)} \leqslant P_{k_n}^{(n)} = P_{k_n}^{(n+1)} \tag{A.12}$$

We have used (A.8), (A.11), (A.8), (A.5), and (A.8). This time $i_{n+1}=j_n$, since $P^{(n+1)}_{j_n}$ is the smallest element. Thus, for Case 2, (A.2) and (A.4) are also satisfied for $n+1$.

For either Case 1 or Case 2, we can use the new labels and show

$$P^{(n+1)}_{i_{n+1}} \leqslant P^{(n+1)}_{k_{n+1}} \leqslant P^{(n+1)}_{j_{n+1}}$$

This verifies Eq. (A.5). Also, for either case

$$P^{(n+1)}_{i_{n+1}} + P^{(n+1)}_{k_{n+1}} = P^{(n+1)}_{i_n} + P^{(n+1)}_{j_n} = P^{(n)}_{j_n} \geqslant P^{(n)}_{k_n} = P^{(n+1)}_{k_n} = P^{(n+1)}_{j_{n+1}}$$

This verifies Eq. (A.6) for $n+1$.

We always have that the sum of the two smallest entries is larger than or equal to the greatest entry.

The next thing we need to show is that the GMA terminates on (1, 1, 1). The initial (P_1, P_2, P_3) are relatively prime by assumption. Since GMA subtracts a smallest element from a largest element, then the subsequent (P_1, P_2, P_3) can never all have the same common factor. Thus (a, a, a) can never occur (if $a>1$) and so 0 never appears as an entry. Therefore in less than $P_1+P_2+P_3$ steps a 1 will appear as an entry. Thereafter 1 will be subtracted from other entries until (1, 1, 1) is reached.

Now we can show the existence of the matrix K of Proposition 2.

Proof of Proposition 2. Define the function *Sum*, which adds the columns of a matrix and returns a vector:

$$(\text{Sum } A)_l \equiv \sum_{k=1}^{3} A_{kl} \tag{A.13}$$

Notice that

$$(\text{Sum } AB)_l = \sum_{k=1}^{3} (AB)_{kl} = \sum_{k,r=1}^{3} A_{kr} B_{rl} = ((\text{Sum } A)B)_l \tag{A.14}$$

We have shown that after a certain number of steps, the GMA terminates onto (1, 1, 1):

$$(1, 1, 1) = (P_1, P_2, P_3) \cdot \mathsf{E}^{-1}_{i_0 j_0} \cdots \mathsf{E}^{-1}_{i_{N-1} j_{N-1}} \tag{A.15}$$

for some $N \geqslant 1$. Or

$$\begin{aligned}(P_1, P_2, P_3) &= (1, 1, 1) \cdot \mathsf{E}_{i_{N-1} j_{N-1}} \cdots \mathsf{E}_{i_0 j_0} \\ &= (\text{Sum } I)\mathsf{K} \qquad (\text{A.16}) \\ &= (\text{Sum } \mathsf{K}) \qquad (\text{A.17})\end{aligned}$$

In the last equality, we have used (A.14). The matrix K of Proposition 2 is probably not unique. The GMA is attractive since there is a nice geometrical interpretation.

Similar proofs for the higher-dimensional GMA are given in Appendix B.

APPENDIX B. PROOF OF STATEMENTS IN DEFINITION 4 AND REMARK 2

In this Appendix we prove some of the assertions made in Definition 4. None of these proofs involve sophisticated ideas.

We first show that the conditions imposed on the initial element are passed along to each iterate of the GMA. Thus we wish to show that for all $n \geqslant 0$ the following hold [(B.1)–(B.6)]:

$$\text{the set} \quad \{i_{n-1},..., i_{n-d+2}, j_{n-1},..., j_{n-d+2}\} \text{ has } d-1 \text{ distinct elements} \tag{B.1}$$

That makes the choice for j_n, which is defined to be the number between 1 and d not in the above set, unique. Recall that i_n, j_n are the labels of E_{ij}. The label i_n is defined to be the label for the uniquely smallest entry of the nth iterate of GMA, or i_{n-1} when there is no uniquely smallest element. We also wish to show

$$i_n \in \{i_{n-1}, j_{n-1}\} \tag{B.2}$$

Now define k_n as

$$k_n \in \{i_{n-1}, j_{n-1}\} \quad \text{and} \quad k_n \neq i_n$$

This is meaningful due to (B.2) [see also (B.4) below]. We initialize $k_{-l} = l+2$ for $d-3 \leqslant l < 0$ to be consistent with the definitions for i_p, j_p for p negative. We also define $k_{2-d} = d$ for convenience. We want to show

$$j_n = k_{n+2-d} \tag{B.3}$$

$$i_n, j_n, k_n, k_{n-1},..., k_{n-d+3} \text{ are } d \text{ distinct integers} \tag{B.4}$$

$$P_{i_n}^{(n)} \leqslant P_{k_n}^{(n)} \leqslant P_{k_{n-1}}^{(n)} \leqslant \cdots \leqslant P_{k_{n-d+3}}^{(n)} \leqslant P_{j_n}^{(n)} \tag{B.5}$$

$$P_{i_n}^{(n)} + P_{k_n}^{(n)} \geqslant P_{j_n}^{(n)} \tag{B.6}$$

Proof by Induction. $n = 0$. There has been defined for negative labels

$$(i_{-l}, j_{-l}, k_{-l}) = (1, l+1, l+2) \qquad \text{for} \quad l = 1,..., d-2$$

This verifies (B.1) immediately for $n=0$. We find $j_0=d=k_{2-d}$. This verifies (B.3). Now $P_1^{(0)}$ is at least as small as any $P_j^{(0)}$. If it is strictly smaller, then $i_0= =1=i_{-1}$. If it is not strictly smaller, we conclude $i_0=i_{-1}=1$ as well. Thus (B.2) holds. We have found

$$i_0, k_0,..., k_{3-d}, j_0 \equiv 1,..., d$$

are d distinct integers, verifying (B.4). Also,

$$P_{i_0}^{(0)} \leqslant P_{k_0}^{(0)} \leqslant P_{k_{-1}}^{(0)} \leqslant \cdots \leqslant P_{k_{3-d}}^{(0)} \leqslant P_{j_0}^{(0)}$$
$$P_{i_0}^{(0)} + P_{k_0}^{(0)} \geqslant P_{j_0}^{(0)}$$

due to the given (see Definition 4), verifying (B.5) and (B.6) for $n=0$.

$n \geqslant 1$. We assume (B.1)–(B.6) true for all r, with $0 \leqslant r \leqslant n$. We wish then to show (B.1)–(B.6) hold for $n+1$. Now, according to (B.1), we have

the set $\{j_{n-1},..., j_{n-d+2}, i_{n-1},..., i_{n-d+2}\}$ has $d-1$ distinct elements

Using (B.2) repeatedly yields

$$\{j_{n-1},..., j_{n-d+2}, i_{n-d+2}\} \quad \text{has } d-1 \text{ distinct elements}$$

And since j_n must not belong to the above set,

$$\{j_n, j_{n-1},..., j_{n-d+2}, i_{n-d+2}\} \quad \text{has } d \text{ distinct elements}$$

(i) Suppose $i_{n-d+3}=i_{n-d+2}$. Then $k_{n-d+3}=j_{n-d+2}$, due to the way the index k is defined, and

$$\{j_n, j_{n-1},..., j_{n-d+2}, i_{n-d+3}\} \quad \text{has } d \text{ distinct elements}$$

So

$$\{j_n, j_{n-1},..., j_{n-d+3}, i_{n-d+3}\} \quad \text{has } d-1 \text{ distinct elements}$$

and is missing the element $j_{n-d+2}=k_{n-d+3}$. From this last statement one uses repeated applications of (B.2), yielding

$$\{j_n,..., j_{n-d+3}, i_n,..., i_{n-d+3}\} \quad \text{has } d-1 \text{ distinct elements}$$

Therefore (B.1) is verified for $n+1$, and we see for case (i) that $j_{n+1}= j_{n-d+2}=k_{n-d+3}$, verifying (B.3).

(ii) In the other case we have $i_{n-d+3}=j_{n-d+2}$, $k_{n-d+3}=i_{n-d+2}$, and

$$\{j_n,..., j_{n-d+3}, i_{n-d+3}, i_{n-d+2}\} \quad \text{has } d \text{ distinct elements}$$
$$\{j_n,..., j_{n-d+3}, i_{n-d+3}\} \quad \text{has } d-1 \text{ distinct elements}$$

and is missing $i_{n-d+2}=k_{n-d+3}$,

$$\{j_n,..., j_{n-d+3}, i_n,..., i_{n-d+3}\} \quad \text{has } d-1 \text{ distinct elements}$$

Again (B.1) is verified, and we have $j_{n+1}=i_{n-d+2}=k_{n-d+3}$, verifying (B.3).

Now by (B.4)

$$i_n, j_n, k_n, k_{n-1},..., k_{n-d+3} \quad \text{are } d \text{ distinct integers} \tag{B.7}$$

Since

$$(P_1^{n+1},..., P_d^{n+1})=(P_1^n,..., P_d^n)\, \mathsf{E}_{i_n j_n}^{-1} \tag{B.8}$$

then

$$\begin{aligned} &P_{i_n}^{(n+1)}=P_{i_n}^{(n)} \\ &P_{j_n}^{(n+1)}=P_{j_n}^{(n)}-P_{i_n}^{(n)} \\ &P_{k_{n-l}}^{(n+1)}=P_{k_{n-l}}^{(n)} \qquad \text{for} \quad 0\leqslant l\leqslant d-3 \end{aligned} \tag{B.9}$$

Case 1. Suppose

$$2P_{i_n}^{(n)}\leqslant P_{j_n}^{(n)} \tag{B.10}$$

Then

$$P_{i_n}^{(n+1)}=P_{i_n}^{(n)}\leqslant P_{j_n}^{(n)}-P_{i_n}^{(n)}=P_{j_n}^{(n+1)}\leqslant P_{k_n}^{(n)}\leqslant P_{k_{n-1}}^{(n)}\leqslant \cdots \leqslant P_{k_{n-d+3}}^{(n)} \tag{B.11}$$

In deriving each relation in (B.11), we have used in order (B.9), (B.10), (B.9), and (B.6), and repeated applications of (B.4). By the way GMA is defined, if i_n labels one of the smallest entries of $P^{(n+1)}$, then $i_{n+1}=i_n$. Thus, by definition, $k_{n+1}=j_n$. Now we already showed $j_{n+1}=k_{n-d+3}$. Thus

$$i_n, j_n, k_n, k_{n-1},..., k_{n-d+3}\equiv i_{j+1}, k_{n+1}, k_n,..., k_{n-d+4}, j_{n+1} \tag{B.12}$$

Using the new set of indices on the RHS of (B.12), one easily shows (B.5) for $n+1$ using (B.11) and the last relation in (B.9).

Case 2. Suppose

$$2P_{i_n}^{(n)}>P_{j_n}^{(n)} \tag{B.13}$$

Then

$$P_{j_n}^{(n+1)}=P_{j_n}^{(n)}-P_{i_n}^{(n)}<P_{i_n}^{(n)}=P_{i_n}^{(n+1)}\leqslant P_{k_n}^{(n)}\leqslant P_{k_{n-1}}^{(n)}\leqslant \cdots \leqslant P_{k_{n-d+3}}^{(n)} \tag{B.14}$$

In deriving each relation in (B.14), we have used in order (B.9), (B.13), (B.9), and (B.4). Since j_n labels the smallest entry of $P^{(n+1)}$, then $i_{n+1} = j_n$. Thus, by definition, $k_{n+1} = i_n$. We showed $j_{n+1} = k_{n-d+3}$. Thus,

$$j_n, i_n, k_n, k_{n-1}, \ldots, k_{n-d+3} \equiv i_{n+1}, k_{n+1}, k_n, \ldots, k_{n-d+4}, j_{n+1} \tag{B.15}$$

For Case 2, with the indices on the RHS of (B.15), one shows (B.5) for $n+1$ using (B.14) and the last relation in (B.9).

For either Case 1 or Case 2,

$$P^{(n+1)}_{i_{n+1}} + P^{(n+1)}_{k_{n+1}} = P^{(n+1)}_{i_n} + P^{(n+1)}_{j_n} = P^{(n)}_{j_n} \geqslant P^{(n)}_{k_{n-d+3}} = P^{(n+1)}_{k_{n-d+3}} = P^{(n+1)}_{j_{n+1}}$$

verifying Eq. (B.6) for $n+1$.

The proofs that (a) 1, 1,..., 1 is reached after a finite number of applications of GMA and (b) Remark 2 following Proposition 2 holds run exactly the same as in the $d=3$ case. See Appendix A.

APPENDIX C. PROOF OF THEOREM 1 (END OF SECTION 4)

The results of Oseledec[20] are indispensable to this section. Define

$$\mathsf{S}_L = (\mathsf{K}_L^{-1})^T \tag{C.1}$$

$$\mathsf{K}_L = \mathbf{Perm}\ \mathsf{E}_{i_{L-1} j_{L-1}} \cdots \mathsf{E}_{i_0 j_0} \tag{C.2}$$

where the E's are elementary matrices and **Perm** is a permutation matrix (see Section 4). Now from Section 4

$$X_j^{(L)} = (\mathsf{S}_L)_{jk} X_k^{(0)} \tag{C.3}$$

where $X_j^{(L)} \leqslant X_{j+1}^{(L)}$, etc., is the Lth GMA iterate after rearranging the GMA iterates from smallest to largest [see Section 4, (4.1)–(4.3)]. We define also the following matrix:

$$(\mathsf{T}_L)_{ij} = \frac{\partial(X_i^{(L)}/X_d^{(L)})}{\partial(X_j^{(0)}/X_d^{(0)})} \tag{C.4}$$

where the partial derivative is taken holding all $X_k^{(0)}$ with $k \neq j$ fixed. One sees that this is precisely the same as

$$(\mathsf{T}_L)_{ij} = \frac{\partial(x_i^{(L)})}{\partial(x_j^{(0)})} \tag{C.5}$$

where

$$(x_1^{(L)}, \ldots, x_{d-1}^{(L)}) = \underbrace{T_{\mathrm{GMA}} \circ \cdots \circ T_{\mathrm{GMA}}}_{L \text{ times}}(x_1, \ldots, x_{d-1}) \tag{C.6}$$

and

$$x_i = X_i / X_d \tag{C.7}$$

That is, the set $X_1^{(L)}/X_d^{(L)},..., X_{d-1}^{(L)}/X_d^{(L)}$ is the Lth iterate of the GMA shift beginning from $X_1^{(0)}/X_d^{(0)},..., X_{d-1}^{(0)}/X_d^{(0)}$. Studying T_L will yield the Oseledec eigenvalues of the GMA shift, whereas studying S_L will yield the eigenvalues of the GMA Euclidean algorithm. We want to understand how these eigenvalues are related. By straightforward calculation one finds

$$(\mathsf{T}_L)_{ij} = \frac{X_d}{X_d^{(L)}} \left(\frac{\partial X_i^{(L)}}{\partial X_j} - \frac{X_i^{(L)}}{X_d^{(L)}} \frac{\partial X_d^{(L)}}{\partial X_j} \right) \tag{C.8}$$

for $1 \leqslant i, j \leqslant d-1$. We can rewrite (C.8) using (C.3),

$$(\mathsf{T}_L)_{ij} = \frac{X_d}{X_d^{(L)}} \left((S_L)_{ij} - \frac{X_i^{(L)}}{X_d^{(L)}} (S_L)_{dj} \right) \tag{C.9}$$

A similar calculation yields for T_L^{-1}

$$(\mathsf{T}_L^{-1})_{rs} = \frac{X_d^{(L)}}{X_d} \left((S_L^{-1})_{rs} - \frac{X_r^{(0)}}{X_d^{(0)}} (S_L^{-1})_{ds} \right) \tag{C.10}$$

Define now a matrix V_L by

$$\mathsf{T}_L^{-1} = \frac{X_d^{(L)}}{X_d^{(0)}} \mathsf{V}_L^T \tag{C.11}$$

Using (C.1), (C.2), (C.10), and (C.11), we find

$$(\mathsf{V}_L)_{rs} = (\mathsf{V}_L^T)_{sr} = (S_L^{-1})_{sr} - \frac{X_s}{X_d} (S_L^{-1})_{dr} \tag{C.12}$$

$$= (\mathsf{K}_L^T)_{sr} - \frac{X_s}{X_d} (\mathsf{K}_L^T)_{dr} \tag{C.13}$$

$$= (\mathsf{K}_L)_{rs} - \frac{X_s}{X_d} - \frac{X_s}{X_d} (\mathsf{K}_L)_{rd} \tag{C.14}$$

Define also

$$Y_j^{(m)} = X_j^{(L-m)} \tag{C.15}$$

So we have

$$Y^{(L)} = (\mathsf{K}_L^T)\, Y^{(0)} \tag{C.16}$$

$$\mathsf{T}_L^{-1} = \frac{Y_d^{(0)}}{Y_d^{(L)}} \mathsf{V}_L^T \tag{C.17}$$

$$\mathsf{V}_{L_{rs}} = (\mathsf{K}_L)_{rs} - \frac{Y_s^{(L)}}{Y_d^{(L)}} (\mathsf{K}_L)_{rd}, \qquad 1 \leqslant r, s \leqslant d-1 \tag{C.18}$$

Lemma 1. If $v_1,..., v_{d-1}$ are the Oseledec eigenvalues of V_L, then the Oseledec eigenvalues of T_L are given by $(\lambda_1/v_{d-1},..., \lambda_1/v_1)$, where λ_1 is the growth rate of the denominators of the GMA Euclidean algorithm.

Proof.

$$\lim_{L\to\infty} (\mathsf{T}_L^T \mathsf{T}_L)^{1/2L} = (\mathsf{T}_L^{-1} \mathsf{T}_L^{-1^T})^{-1/2L} \tag{C.19}$$

$$= \lim_{L\to\infty} \left(\frac{Y_d^{(0)}}{Y_d^{(L)}}\right)^{-1/L} \lim_{L\to\infty} (\mathsf{V}_L^T \mathsf{V}_L)^{-1/2L} \tag{C.20}$$

$$= \lim_{L\to\infty} \left(\frac{Y_d^{(L)}}{Y_d^{(0)}}\right)^{1/L} \lim_{L\to\infty} [(\mathsf{V}_L^T \mathsf{V}_L)^{1/2L}]^{-1} \tag{C.21}$$

The eigenvalues of the matrix inside parentheses on the RHS are by assumption $v_1,..., v_{d-1}$. So the eigenvalues of the inverse of this matrix are $1/v_j$. The first parentheses is simply λ_1, the growth rate of denominators. Remember that the X's are decreasing (being stripped) and the Y's are increasing. This establishes the result of Lemma 1. ▌

Let us next define a matrix W_L which has a similar definition to V_L, but is a $d \times d$ matrix,

$$(\mathsf{W}_L)_{rs} = (\mathsf{K}_L)_{rs} - \frac{Y_s^{(L)}}{Y_d^{(L)}} (\mathsf{K}_L)_{rd} \tag{C.22}$$

for $1 \leqslant r,\ s \leqslant d$.

Lemma 2. The Oseledec eigenvalues of W_L are given by $v_1,..., v_{d-1}$, 0, where the v_i are the Oseledec eigenvalues of V_L.

Proof. Clearly 0 is an eigenvalue of W_L, since W_L has a column of zeros, $(\mathsf{W}_L)_{rd} = 0$. The matrix W_L may be written

$$\mathsf{W}_L = \begin{bmatrix} I_{(d-1)\times(d-1)} & 0 \\ -y_j^{(0)} & 0 \end{bmatrix} \begin{bmatrix} \mathsf{V}_L & 0 \\ 0 & 1 \end{bmatrix} \tag{C.23}$$

where for convenience we have defined

$$y_j^{(0)} = Y_j^{(0)} / Y_d^{(0)} \tag{C.24}$$

We have introduced block matrix notation in (C.23). For example, the djth element of the first matrix on the RHS of (C.23) is $y_j^{(0)}$ for $j \leqslant d-1$. The upper $(d-1)\times(d-1)$ block of that matrix has a one on the diagonal and zeros everywhere else.

Now the Oseledec eigenvalues of a matrix are the same as the Oseledec eigenvalues of its transpose. Thus, let us investigate

$$\mathsf{W}_L \mathsf{W}_L^T = \begin{bmatrix} I_{(d-1)\times(d-1)} & 0 \\ -y_j^{(0)} & 0 \end{bmatrix} \begin{bmatrix} \mathsf{V}_L \mathsf{V}_L^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_{(d-1)\times(d-1)} & -y_j^{(0)} \\ 0 & 0 \end{bmatrix} \tag{C.25}$$

Now we can go to a basis where $\mathsf{V}_L \mathsf{V}_L^T$ is diagonal via a rotation. Thus,

$$\begin{bmatrix} \mathsf{V}_L \mathsf{V}_L^T & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathsf{R}_L^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1^{(2L)} & \cdots & 0 & 0 \\ \vdots & \ddots & 0 & 0 \\ 0 & 0 & v_{d-1}^{(2L)} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathsf{R}_L & 0 \\ 0 & 1 \end{bmatrix} \tag{C.26}$$

where the R_L are $(d-1)\times(d-1)$ rotation matrices. That V_L has Oseledec eigenvalues is a given of Lemma 2. Thus,

$$\lim_{L\to\infty} (v_1^{(2L)})^{1/2L} \to v_1, \quad \text{etc.} \tag{C.27}$$

And using (C.25) and (C.26),

$$\begin{aligned} &\begin{bmatrix} \mathsf{R}_L & 0 \\ 0 & 1 \end{bmatrix} \mathsf{W}_L \mathsf{W}_L^T \begin{bmatrix} \mathsf{R}_L^T & 0 \\ 0 & 1 \end{bmatrix} \\ &\quad = \begin{bmatrix} I & 0 \\ \alpha_j & 0 \end{bmatrix} \begin{bmatrix} v_1^{(2L)} & \cdots & 0 & 0 \\ \vdots & \ddots & 0 & 0 \\ 0 & 0 & v_{d-1}^{(2L)} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I & \alpha_j \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{C.28}$$

Note that each $v_i^{(2L)}$ and α_j is real. One next calculates the eigenvalues of the RHS. Whereas this is very tedious, it is completely straightforward. We demonstrate this for a particular case; the proof for the general case is exactly the same. Suppose we wish to calculate the eigenvalues of

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & 0 \end{bmatrix} \begin{bmatrix} m_1^{2L} & 0 & 0 \\ 0 & m_2^{2L} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{bmatrix} \tag{C.29}$$

where a, b are real. A straightforward calculation for the eigenvalues e_1, e_2, e_3 yields

$$\begin{aligned} e_1 &= 0 \\ e_{2,3} &= \frac{m_1^{2L}(1+a^2)+m_2^{2L}(1+b^2)}{2} \\ &\quad \pm \left[\left(\frac{m_1^{2L}(1+a^2)-m_2^{2L}(1+b^2)}{2}\right)^2 + a^2b^2m_1^{2L}m_2^{2L}\right]^{1/2} \end{aligned} \tag{C.30}$$

If $m_1 > m_2$, then

$$e_2^{(2L)} \simeq m_1^{2L}(1+a^2), \qquad e_3^{(2L)} = \frac{1+a^2+b^2}{1+b^2}\, m_2^{2L}$$

yielding $e_i = \lim_{L\to\infty}(e_i^{(2L)})^{1/2L} = (0, m_1, m_2)$ as the eigenvalues. If $m_1 = m_2 = m$, then

$$e_{2,3}^{(2L)} = m^{2L}\left\{1+\frac{a^2+b^2}{2} \pm \left[\left(\frac{a^2-b^2}{2}\right)^2 + a^2b^2\right]^{1/2}\right\} \tag{C.31}$$

Obviously e_2, e_3 are real. Now

$$\begin{aligned} 1 &+ \frac{a^2+b^2}{2} - \left[\left(\frac{a^2-b^2}{2}\right)^2 + a^2b^2\right]^{1/2} \\ &= 1+\frac{a^2+b^2}{2} - \left[\left(\frac{a^2+b^2}{2}\right)^2 - a^2b^2\right]^{1/2} \\ &\geqslant 1+\frac{a^2+b^2}{2} - \left[\left(\frac{a^2+b^2}{2}\right)^2\right]^{1/2} = 1 \end{aligned} \tag{C.32}$$

Thus, both $e_{2,3}^{(2L)}$ are proportional to m^{2L}. [The above calculation for (C.32) was to show that the proportionality constant cannot be zero.] So even in the special case, we have $e_2 = m_2$, $e_3 = m_3$, as expected. ∎

What we next show is that the eigenvalues of W_L are $\lambda_2, ..., \lambda_d$, 0, where the λ_i are the eigenvalues of the E-string. That would mean that the Oseledec eigenvalues of V_L are $\lambda_2, ..., \lambda_d$ via Lemma 2 and that the Oseledec eigenvalues of T_L are $\lambda_1/\lambda_d, ..., \lambda_1/\lambda_2$, via Lemma 1. That is the result that establishes Theorem 1 of Section 4.

From (C.22) and (C.16), one can write

$$(\mathsf{W}_L)_{rs} = \frac{1}{Y_d^{(L)}} \sum_{a=1}^{d} (K_{rs}K_{ad} - K_{rd}K_{as})\, Y_a^{(0)} \tag{C.33}$$

Suppose K_L is similar to a diagonal matrix:

$$\mathsf{K}_L = \mathsf{P}\begin{bmatrix} \lambda_1^{(L)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_d^{(L)} \end{bmatrix}\mathsf{P}^{-1} \tag{C.34}$$

for some matrix P, where

$$\lim_{L\to\infty} (|\lambda_i^{(L)}|)^{1/L} = \lambda_i; \qquad i = 1,..., d \tag{C.35}$$

are the Oseledec eigenvalues of the E-string. Let us put W_L also in this basis. Then we find

$$(\mathsf{P}^{-1}\mathsf{W}_L\mathsf{P})_{qt} = \frac{F_{11}\lambda_1^{(L)}}{Y_d^{(L)}}\left[\sum_{u=1}^{d} \frac{F_{uu}}{F_{11}} \frac{\lambda_u^{(L)}\lambda_t^{(L)}}{\lambda_1^{(L)}}\delta_{qt} - \frac{F_{tq}\lambda_q^{(L)}\lambda_t^{(L)}}{F_{11}\lambda_1^{(L)}}\right] \tag{C.36}$$

$$\equiv \frac{F_{11}\lambda_1^{(L)}}{Y_d^{(L)}} C_{qt} \tag{C.37}$$

where we have defined

$$F_{tq} = \sum_{a=1}^{d} Y_a^{(0)} P_{at}(P^{-1})_{qd} \tag{C.38}$$

and we have defined the matrix C implicitly in going from (C.36) to (C.37). Note that $F_{tq}F_{qt} = F_{tt}F_{qq}$.

Thus, we must calculate the eigenvalues of C_{qt}. Now,

$$\begin{aligned} C_{q=1,t=1} &= \sum_{u=2}^{d} \frac{F_{uu}}{F_{11}}\lambda_u^{(L)} \\ C_{q=1,t\neq 1} &= -\lambda_t^{(L)}\frac{F_{t1}}{F_{11}} \\ C_{q\neq 1,t=1} &= -\lambda_q^{(L)}\frac{F_{1q}}{F_{11}} \\ C_{q\neq 1,t\neq 1} &= \delta_{qt}\lambda_t^{(L)} + \delta_{qt}\lambda_t^{(L)} \sum_{u=2}^{d} \frac{\lambda_u^{(L)}}{\lambda_1^{(L)}}\frac{F_{uu}}{F_{11}} - \lambda_t^{(L)}\frac{\lambda_q^{(L)}}{\lambda_1^{(L)}}\frac{F_{tq}}{F_{11}} \end{aligned} \tag{C.39}$$

Let us break C up into two pieces

$$\mathsf{C} = \mathsf{C}^0 + \mathsf{C}^1 \tag{C.40}$$

We put into $\mathbf{C}^1$ those terms which are down by an order of $\lambda_j^{(L)}/\lambda_1^{(L)}$. Thus, let

$$
\begin{aligned}
C^0_{q=1,t=1} &= \sum_{u=2}^{d} \frac{F_{uu}}{F_{11}} \lambda_u^{(L)} \\
C^0_{q=1,t\neq 1} &= -\lambda_t^{(L)} \frac{F_{t1}}{F_{11}} \\
C^0_{q\neq 1,t=1} &= -\lambda_q^{(L)} \frac{F_{1q}}{F_{11}} \\
C^0_{q\neq 1,t\neq 1} &= \delta_{qt} \lambda_t^{(L)} \\
C^1_{q=1,t=1} &= 0 \\
C^1_{q=1,t\neq 1} &= 0 \\
C^1_{q\neq 1,t=1} &= 0 \\
C^1_{q\neq 1,t\neq 1} &= \lambda_t^{(L)} \left(\delta_{qt} \sum_{u=2}^{d} \frac{\lambda_u^{(L)}}{\lambda_1^{(L)}} \frac{F_{uu}}{F_{11}} - \frac{\lambda_q^{(L)}}{\lambda_1^{(L)}} \frac{F_{tq}}{F_{11}} \right)
\end{aligned} \tag{C.41}
$$

We will assume that $\lambda_1 > 1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$. That is, the parallelograms of Section 3 are getting long (in the eigendirection of the eigenvalue λ_1) and thin (in every other eigendirection). Our assumption is obvious the way GMA is defined and is consistent with all our numerical results, although strictly speaking we have not supplied a proof of this assertion. Consideration of (C.41) makes it clear that $\mathbf{C}^1$ has a vanishing effect on the spectrum as $\mathrm{L} \to \infty$.

We will need to show that eigenvalues of $\mathbf{C}^0$ (and hence $\mathbf{C}$) are $\lambda_2, \ldots, \lambda_d$. A straightforward calculation yields

$$
\frac{-1}{\alpha} \det(\mathbf{C}^0 - \alpha I) = \left(\prod_{i=2}^{d} |\lambda_i^n - \alpha| \right) \cdot \left\{ 1 - \sum_{u=2}^{d} \frac{F_{uu}/F_{11}}{\alpha/\lambda_u^{(L)} - 1} \right\} \tag{C.42}
$$

If the λ_i of (C.35) are all different, then it is easy to solve for roots for α on the RHS of (C.42),

$$
\alpha = \alpha_b = \lambda_{b+1}^{(L)} \sum_{q=1}^{b+1} F_{qq} \Big/ \sum_{q=1}^{b} F_{qq}; \qquad 1 \leqslant b \leqslant d-1 \tag{C.43}
$$

Thus, the Oseledec eigenvalues of $\mathbf{C}^0$ and $\mathbf{C}$ are $\lambda_2, \ldots, \lambda_d$, 0. And the Oseledec eigenvalues of $\mathbf{W}$ are $\lambda_2, \ldots, \lambda_d$, 0, yielding the Oseledec eigenvalues for $\mathbf{V}$ as $\lambda_2, \ldots, \lambda_d$. Thus, the Oseledec eigenvalues for $\mathbf{T}$ are $\lambda_1/\lambda_d, \ldots, \lambda_1/\lambda_2$.

In the case where some of the λ_i are equal, it is not easy to solve explicitly for the roots as in (C.43). The conclusion, however, that the Oseledec eigenvalues of C are $\lambda_2,..., \lambda_d$, 0 remains the same. (That is, if two roots, say λ_3 and λ_4, are equal, then there will be two roots for α with $|\alpha/\lambda_3^{(L)}|$ a number which always stays of order unity as $L \to \infty$.) The rest of the argument runs parallel to the above. Theorem 1 of Section 4 is proved.

GLOSSARY

$\mathscr{Z}^+$	set of positive integers
$[\cdot]$	Gauss integer symbol (Section 2)
h	entropy
I	# of irrationals to be simultaneously approximated
d	dimension of the vector of convergents (equal to $I+1$)
$\mathscr{I}^p$	unit hypercube in p dimensions
$\mathscr{S}$	support of the invariant measure (see Section 5)
E_{ij}	elementary matrix, with klth component $\delta_{kl}+\delta_{ik}\delta_{jl}$
E-string	product of elementary matrices given by the algorithm
vertices V_i	corners of the elementary simplex adjoined to the origin (Section 3)
mediants M_{ik}	a direct sum of any two of the vertices (Section 3)
focus	sum of all the vertices (Section 3)
Euclidean algorithm	reverse of the E-string procedure (see Section 2)
OCF	ordinary continued-fraction algorithm
GMA	generalized mediant algorithm: the subject of this paper
JP	Jacobi–Perron: the most well-studied MCFA
MCFA	Multidimensional continued-fraction algorithm
KS entropy	Kolmogorov–Sinai entropy
T_{OCF}	ordinary continued-fraction shift map
FS	Farey shift map
$(a,..., z)$	irrational vector with I components; each element is an irrational
$d\mu(x)$	invariant measure
$\rho(x)$	invariant density $[= d\mu(x)/dx]$
$\lambda_1, \lambda_2,..., \lambda_d$	the d Oseledec eigenvalues of the E-string (see Section 4) ordered $\lambda_1 > 1 > \lambda_2 \geqslant \lambda_3 \geqslant \cdots$
$\sigma_1,..., \sigma_{d-1}$	Oseledec eigenvalues of the shift map (Section 4) ordered greatest to smallest; all the $\sigma_i > 1$, and $\sigma_i = \lambda_1/\lambda_{d-i+1}$
$\ln \sigma_1,..., \ln \sigma_{d-1}$	Oseledec *exponents* of the shift map (Section 4)
Perm	a permutation matrix (Section 4).

ACKNOWLEDGMENTS

This work was supported in part by DOE grant DE-AC03-84 ER 40182 and NSF grant DMR-84-05355 (Polymers Program—University of Illinois). The author wishes to thank J. C. Lagarias, Bambi Hu, and Jian Min Mao for discussions, and Y. Oono for his encouragement at early stages of the work. The author would also like to thank S. Hariharan and the Department of Mathematical Sciences at the University of Akron for the use of their computing services, as well as the referees for their many useful comments.

REFERENCES

1. H. Minkowski, *Diophantische Approxiomationen* (Leipzig, Teubner, 1982).
2. L. Bernstein, *The Jacobi–Perron Algorithm and Its Application* (Springer-Verlag, Berlin, 1980).
3. W. Schmidt, *Diophantine Approximation* (Springer-Verlag, Berlin, 1980).
4. L. K. Hua and W. Yuan, *Applications of Number Theory to Numerical Analysis* (Springer-Verlag, Berlin, 1981).
5. G. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Clarendon Press, Oxford, 1979).
6. A. Y. Khintchine, *Continued Fractions* (P. Nordhoff, Groningen, 1963); *Compositio Mathematica* **1**:361–382 (1935); **3**:275–285 (1936); P. Lévy, *Théorie de l'Addition des Variables Aléatives* (Paris, 1937).
7. M. S. Waterman, *Rocky Mtn. J. Math.* **6**:181 (1976).
8. F. Schweiger, *Dynam. Syst. Ergodic Theory* **23**:283 (1989); F. Schweiger, *The Metrical Theory of Jacobi–Perron Algorithm* (Springer-Verlag, Berlin, 1973).
9. G. Szekeres, *Ann. Univ. Sci. Budap. Rolando Eötvös Sect. Math.* **13**:113 (1970).
10. J. C. Lagarias, Geodesic multidimensional continued fractions, *Acta Math.*
11. M. S. Waterman, *Z. Wahrsch. Verw. Geb.* **16**:77 (1970).
12. Hao Bai-Lin, *Chaos* (World Scientific, Singapore, 1984).
13. M. Feigenbaum, *J. Stat. Phys.* **52**:527 (1988); S. Ostlund and S. Kim, *Physica Scripta* **T9**:193–198 (1985); P. Cvitanovic, B. Shraiman, and B. Soderberg, *Physica Scripta* **32**:263 (1985).
14. A. J. Brentjes, *Multi-Dimensional Continued Fraction Algorithms* (Mathematisch Centrum, Amsterdam, 1981).
15. L. K. Hua, *Introduction to Number Theory* (Springer-Verlag, Berlin, 1982).
16. J. Guckenheimer, B. Hu, and J. Rudnick, Quasiperiodic transitions to chaos with three incommensurate frequencies, Preprint (1982).
17. S. Kim and S. Ostlund, *Phys. Rev. Lett.* **55**:1165 (1985).
18. Y. Oono, *Prog. Theor. Phys.* **60**:1944 (1978).
19. J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57**:617 (1985).
20. V. I. Oseledec, *Moscow Math. Soc.* **19**:197 (1968).
21. A. Crisanti, G. Paladin, and A. Vulpiani, *Phys. Rev. A* **39**:6491 (1989); A. Crisanti, G. Paladin, and A. Vulpiani, *J. Stat. Phys.* **53**:583 (1988); G. Paladin and A. Vulpiani, *J. Phys. A* **19**:2033 (1986); G. Paladin and A. Vulpiani, *J. Phys. A* **21**:363 (1988); G. Parisi and A. Vulpiani, *J. Phys. A* **21**:363 (1988); J. Deutsch and G. Paladin, *Phys. Rev. Lett.*

62:695 (1989); A. Crisanti, G. Paladin, and A. Vulpiani, *Phys. Rev. B* **35**:7164 (1987); K. Ravishankar, *J. Stat. Phys.* **54**:531 (1989); J. P. Eckmann and C. E. Wayne, *J. Stat. Phys.* **50**:853 (1988); B. Derrida, K. Mecheri, and J. L. Pichard, *J. Phys.* (Paris) **48**:733 (1987).

22. E. S. Selmer, *Nord. Mat. Tidskr.* **9**:37 (1961); F. Schweiger, *Ber. Öster. Akad. Wiss. Math. Naturw. Kl. Abt. II, Math. Phys. Tech. Wiss.* **191**:325 (1982).
23. P. R. Baldwin, A convergence exponent for multidimensional continued-fraction algorithms, *J. Stat. Phys.*, this issue.