

A Generalization of the Simple Continued Fraction Algorithm

By Theresa P. Vaughan

Dedicated to Professor Leonard Carlitz on the occasion of his seventieth birthday

Abstract. In this paper we present a generalization of the continued fraction algorithm, based on a geometric and matrix-theoretic approach.

We first give a geometric representation in the plane R^2 , of the simple continued fraction algorithm, described in terms of geometric and arithmetic properties of 2×2 matrices with nonnegative integer entries and determinant 1. The algorithm of this paper is then derived as a natural generalization of the situation in R^2 . We describe a computational procedure for our algorithm, and give several examples.

In this paper we present a generalization of the continued fraction algorithm, based on a geometric and matrix-theoretic approach.

There are several ways in which the simple continued fraction algorithm (CFA) may be represented geometrically in R^2 . We summarize briefly, in Section 3 of this paper, one such interpretation, which is described entirely in terms of geometric and arithmetic properties of 2×2 matrices with nonnegative entries and determinant 1. Our approach displays, in an easy and natural manner, all the main properties of continued fractions: the coordinates of the continued fraction for a positive real number x , its convergents, and the convergence of the algorithm. The question of periodicity is not so simple, and we do not discuss it here.

In Section 1 we give a simple computational procedure for our algorithm, and some examples of its use.

In Section 4 we show how this algorithm is a natural generalization to R^3 of the geometric and matrix-theoretic situation in R^2 given in Section 3.

In Section 5 we show that our algorithm is always convergent. Geometrically, the idea of the proof is very simple. The actual details are elementary but rather tedious.

Basically, our algorithm bears a strong resemblance to the Jacobi-Perron algorithm in its matrix-theoretic aspects, and we hope in another paper to investigate the Jacobi-Perron algorithm itself from the point of view of this paper. We have not yet worked out a theoretical basis for comparing our algorithm to the Jacobi-Perron algorithm, but the examples of Section 2 show some interesting differences.

For instance, for $(\sqrt[3]{2}, \sqrt[3]{4})$, our algorithm became periodic after 13 steps, while the Jacobi-Perron algorithm takes only 3 steps. On the other hand, our rational approximations at comparable stages are superior in accuracy.

Received June 1, 1977; revised September 27, 1977.

AMS (MOS) subject classifications (1970). Primary 10F20, 10A30.

Copyright © 1978, American Mathematical Society

In [1, p. 69] it is stated that it is not known whether or not the Jacobi-Perron algorithm for $(\sqrt[3]{4}, \sqrt[3]{16})$ is periodic, although it has been carried out to 150 steps. Our algorithm for this vector becomes periodic after 91 steps.

1. The Algorithm. Define the matrices E_{ij} ($i, j \in \{1, 2, 3\}, i \neq j$) by: E_{ij} is 3×3 and

$$(E_{ij})_{rs} = \begin{cases} 1, & r = s, \\ 1, & r = j, s = i, \\ 0, & \text{otherwise.} \end{cases}$$

The algorithm defines by recursion, two sequences $\{x_n\}$, $\{y_n\}$ of nonnegative real numbers, a sequence $\{k_n\}$ of positive integers, and a sequence $\{A_n\}$ of 3×3 matrices, as follows. Initially, let x_0, y_0 be nonnegative real numbers, and A_0 the 3×3 identity matrix. Then for $i = 0, 1, 2, \dots$

I. If $0 < x_i < y_i < 1$, put

$$k_i = \left[\frac{1}{y_i} \right], \quad y_{i+1} = \left(\frac{1}{y_i} - k_i \right)^{-1}, \quad x_{i+1} = x_i(1 + k_i y_{i+1}), \quad A_{i+1} = A_i E_{31}^{k_i}.$$

II. If $0 < y_i < x_i < 1$, put

$$k_i = \left[\frac{1}{x_i} \right], \quad x_{i+1} = \left(\frac{1}{x_i} - k_i \right)^{-1}, \quad y_{i+1} = y_i(1 + k_i x_{i+1}), \quad A_{i+1} = A_i E_{21}^{k_i}.$$

III. If $0 < x_i < 1 < y_i$, put

$$k_i = [y_i], \quad x_{i+1} = x_i, \quad y_{i+1} = y_i - k_i, \quad A_{i+1} = A_i E_{13}^{k_i}.$$

IV. If $0 < y_i < 1 < x_i$, put

$$k_i = [x_i], \quad x_{i+1} = x_i - k_i, \quad y_{i+1} = y_i, \quad A_{i+1} = A_i E_{12}^{k_i}.$$

V. If $0 < 1 < x_i < y_i$, put

$$k_i = [y_i/x_i], \quad x_{i+1} = x_i, \quad y_{i+1} = y_i - k_i x_i, \quad A_{i+1} = A_i E_{23}^{k_i}.$$

VI. If $0 < 1 < y_i < x_i$, put

$$k_i = [x_i/y_i], \quad x_{i+1} = x_i - k_i y_i, \quad y_{i+1} = y_i, \quad A_{i+1} = A_i E_{32}^{k_i}.$$

Evidently, given $x_i, y_i > 0$, then in order to find x_{i+1}, y_{i+1} it is not necessary to know the matrix A_i . We shall see later, however, that the matrices A_i carry the information equivalent to the knowledge of the i th convergent of a continued fraction.

Note also that in the six steps described above, the inequalities are strict. One can proceed perfectly well, however, if some x_i, y_i are 0 or 1 or equal to each other. In these cases one finds that either the algorithm must stop (rational x_i, y_i) or involves only two steps (x_i or $y_i = 0$ and the algorithm reduces to continued fractions) or the i th step is not uniquely determined. Procedure in these cases, then, is left to the user.

We also note that the user may find it more convenient to compute the matrices A_i^{-1} rather than A_i ; A_i^{-1} may have negative entries, but their magnitude is smaller. The matrices A_i themselves carry other information, which we describe briefly below.

Suppose A_i is the matrix produced at the i th step of the algorithm, whose 0th step was $x_0, y_0, A_0 = I_3$. Then A_i is a matrix with determinant 1, with nonnegative integer entries, which we assume for the moment are all positive.

Suppose we have

$$A_i = \begin{bmatrix} a_{1i} & b_{1i} & c_{1i} \\ a_{2i} & b_{2i} & c_{2i} \\ a_{3i} & b_{3i} & c_{3i} \end{bmatrix}.$$

Then the ratios

$$\frac{a_{2i}}{a_{1i}}, \frac{b_{2i}}{b_{1i}}, \frac{c_{2i}}{c_{1i}}$$

are approximations to x_0 , and the ratios

$$\frac{a_{3i}}{a_{1i}}, \frac{b_{3i}}{b_{1i}}, \frac{c_{3i}}{b_{1i}}$$

are approximations to y_0 .

The excellence of these approximations is proven later to be at least of the order

$$\left(\frac{2}{3}\right)^{\lfloor i/2 \rfloor},$$

and we conjecture that in fact they are at least of order

$$\left(\frac{2}{3}\right)^i.$$

Remark. The algorithm we have described has the property that the periodicity of the algorithm is independent of the order of the entries of v_0 . That is, the algorithms for the vectors $(1, x_0, y_0), (1, y_0, x_0), (1, x_0^{-1}, y_0 x_0^{-1}), (1, y_0 x_0^{-1}, x_0^{-1}), (1, y_0^{-1}, x_0 y_0^{-1}), (1, x_0 y_0^{-1}, y_0^{-1})$ are either all periodic, or none of them is. This fact enables us to predict periodicity before a repetition of steps actually takes place. E.g. for the vector $v_0 = (1, \sqrt[3]{4}, 2\sqrt[3]{2})$, we find $v_1 = (1, \sqrt[3]{4}, 2\sqrt[3]{2} - \sqrt[3]{4})$, and $v_{31} = (1, \sqrt[3]{4} - 1, 1/\sqrt[3]{4}) = (1/\sqrt[3]{4}) \cdot (\sqrt[3]{4}, 2\sqrt[3]{2} - \sqrt[3]{4}, 1)$. That is, except for the multiplier, the entries of v_{31} are those of v_1 permuted by $\sigma = (132)$. The order of the permutation σ is 3, so we predict periodicity, with $v_{91} = v_1$. In fact, putting

$$T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

we find for this algorithm that

$$E_{23}(E_{23}^{-1}A_{31}T^2)^3 = A_{91}$$

(since the rows of T^2X are the rows of X permuted by σ).

2. **Examples.** (i) $x_0 = 1/3, y_0 = 5/7$. Then $0 < x_0 < y_0 < 1, k_1 = \lceil 7/5 \rceil = 1, A_1 = E_{31}^1, y_1 = 5/2, x_1 = 7/6$. Now $1 < x_1 < y_1$, so $k_2 = \lfloor y_1/x_1 \rfloor = 2, x_2 = x_1 = 7/6, y_2 = y_1 - 2x_1 = 1/6, A_2 = E_{31}^1 E_{23}^2$. Now $0 < y_2 < 1 < x_2, k_3 = \lfloor x_2 \rfloor = 1, x_3 = x_2 - 1 = 1/6, y_2 = 1/6, A_3 = E_{31}^1 E_{23}^2 E_{12}^1$. Now $0 < x_2 \leq y_2 < 1, k_4 = \lfloor 1/y_2 \rfloor = 6, A_4 = E_{31}^1 E_{23}^2 E_{12}^1 E_{31}^6$. Now $1/y_2 - 6$ is 0, so we can proceed no further.

We may rewrite the procedure above as follows:

$$\begin{aligned} \left(1, \frac{1}{3}, \frac{5}{7}\right) &= \frac{2}{7} \cdot E_{31} \left(1, \frac{7}{6}, \frac{5}{2}\right) \\ &= \frac{2}{7} \cdot E_{31} E_{23}^2 \left(1, \frac{7}{6}, \frac{1}{6}\right) = \frac{2}{7} E_{31} E_{23}^2 E_{12} \left(1, \frac{1}{6}, \frac{1}{6}\right) \\ &= \frac{2}{7} \cdot \frac{1}{6} E_{31} E_{23}^2 E_{12} E_{31}^6 (0, 1, 1). \end{aligned}$$

As a check, we compute

$$E_{31} E_{23}^2 E_{12} E_{31}^6 = \begin{bmatrix} 3 & 2 & 19 \\ 1 & 1 & 6 \\ 2 & 2 & 13 \end{bmatrix} = A_4;$$

and we indeed have $A_4(0, 1, 1) = (21, 7, 15)$ so that

$$\frac{2}{7} \cdot \frac{1}{6} A_4(0, 1, 1) = \left(1, \frac{1}{3}, \frac{5}{7}\right).$$

We could, from this point of view, go one step further, and write

$$(0, 1, 1) = E_{23}(0, 1, 0);$$

then

$$A_4 E_{23} = \begin{bmatrix} 3 & 21 & 19 \\ 1 & 7 & 6 \\ 2 & 15 & 13 \end{bmatrix}$$

and now the center column is $21 \times (1, 1/3, 5/7)$.

In fact, this is always possible, although the last step or two may be accomplished in more than one way (e.g. above, one can write $(0, 1, 1) = E_{32}(0, 0, 1)$ or even one step back, $(6, 1, 1) = E_{21}^6(0, 1, 1)$). For this reason, we have not attempted to describe this procedure in the algorithm proper.

Thus, given rational x_0, y_0 , the algorithm may be used to produce a matrix with nonnegative integer entries and determinant 1 having one column equal to an integer multiple of $(1, x_0, y_0)$.

(ii) $x_0 = \sqrt{2}, y_0 = 0$. $0 \leq y_0 < 1 < x_0, k_1 = \lfloor \sqrt{2} \rfloor = 1, x_1 = \sqrt{2} - 1, y_1 = y_0 = 0, A_1 = E_{12}$. Now $0 \leq y_1 < x_1 < 1, k_2 = \lfloor 1/x_1 \rfloor = 2, x_2 = \sqrt{2} + 1, y_2 = y_1 \cdot (1 + 2x_2) = 0, A_2 = E_{12} E_{21}^2; 0 \leq y_2 < 1 < x_2, \lfloor x_2 \rfloor = 2, x_3 = x_2 - 2 = \sqrt{2} - 1 = x_1, y_3 = y_2 = 0 = y_1, A_3 = E_{12}^2$.

Note that we use only Step II and Step IV of the algorithm. The matrix A_i will have the form

$$A_i = \begin{bmatrix} a_i & b_i & 0 \\ c_i & d_i & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and the quotients $c_i/a_i, d_i/b_i$ are actually the i th and the $(i-1)$ st convergents of the continued fraction for $\sqrt{2}$ (in an order described in Section 3).

(iii) $x_0 = \sqrt[3]{2}, y_0 = \sqrt[3]{4}$. We do not repeat the calculations here, but the algorithm repeats after thirteen steps.

We find the matrix

$$A = E_{23}^1 E_{12}^1 E_{31}^3 E_{23}^1 E_{32}^3 E_{23}^1 E_{12}^3 E_{31}^1 E_{13}^1 E_{21}^1 E_{32}^1 E_{13}^1 E_{21}^1.$$

The product after the initial E_{23} repeats.

We compute: $A = E_{23}B$ where

$$B = \begin{bmatrix} 281 & 400 & 177 \\ 354 & 504 & 223 \\ 92 & 131 & 58 \end{bmatrix}.$$

Now $(1, \sqrt[3]{2}, \sqrt[3]{4}) = E_{23}(1, \sqrt[3]{2}, \sqrt[3]{4} - \sqrt[3]{2})$ and $v = (1, \sqrt[3]{2}, \sqrt[3]{4} - \sqrt[3]{2})$ is an eigenvector for B . We compute the vector

$$B(v) = \begin{bmatrix} 281 + 223 & \sqrt[3]{2} + 177 & \sqrt[3]{4} \\ 354 + 281 & \sqrt[3]{2} + 223 & \sqrt[3]{4} \\ 92 + 73 & \sqrt[3]{2} + 58 & \sqrt[3]{4} \end{bmatrix}.$$

Now the quantity $281 + 223\sqrt[3]{2} + 177\sqrt[3]{4} = \alpha$ is an integer unit in $\mathbf{Q}(\sqrt[3]{2})$ and is an eigenvalue of B . Indeed, if we put

$$M = \begin{bmatrix} 281 & 223 & 177 \\ 2 \cdot 177 & 281 & 223 \\ 2 \cdot 223 & 2 \cdot 177 & 281 \end{bmatrix},$$

then M is a matrix representing multiplication by α (in the ordered basis $(\sqrt[3]{4}, \sqrt[3]{2}, 1)$ for $\mathbf{Q}(\sqrt[3]{2})$ over \mathbf{Q}), M has determinant one, and one has

$$M^{-1} = \begin{bmatrix} 19 & -5 & -8 \\ -16 & 19 & -5 \\ -10 & -16 & 19 \end{bmatrix}.$$

We note that M is similar to B :

$$E_{23}BE_{23}^{-1} = M;$$

the characteristic polynomial of B is

$$x^3 - 843x^2 + 57x - 1,$$

and $\alpha \approx 842.932$ is very close to the trace of B .

Using an HP-25 hand calculator, we find:

$$\left(\frac{2}{3}\right)^{18} = .000676, \quad \sqrt[3]{2} \approx 1.2599210, \quad \sqrt[3]{4} - \sqrt[3]{2} \approx 0.3274800,$$

$$\frac{\sqrt[3]{4} - \sqrt[3]{2}}{\sqrt[3]{2}} = \sqrt[3]{2} - 1 = .2599210,$$

$\frac{354}{281} = 1.2597864$	$\frac{92}{281} = .3274021$	$\frac{92}{354} = 0.2598870$
$\frac{504}{400} = 1.26$	$\frac{131}{400} = 0.3275$	$\frac{131}{504} = 0.2599206$
$\frac{223}{177} = 1.2598870$	$\frac{58}{177} = 0.2376836$	$\frac{58}{223} = .26008969$

TABLE 1

that is, as expected, the quotients $y/x, z/x$ of the column vectors of B are within .0006 of the same quotients for v .

It is interesting to compare these computations with the similar ones of the Jacobi-Perron algorithm. One finds:

$$a_1^{(0)} = \sqrt[3]{2}, \quad a_2^{(0)} = \sqrt[3]{4}; \quad b_1^{(0)} = 1, \quad b_2^{(0)} = 1,$$

$$a_1^{(1)} = \frac{\sqrt[3]{4} - 1}{\sqrt[3]{2} - 1}, \quad a_2^{(1)} = \frac{1}{\sqrt[3]{2} - 1}, \quad b_1^{(1)} = 2, \quad b_2^{(1)} = 3,$$

$$a_1^{(2)} = \sqrt[3]{2} + 2, \quad a_2^{(2)} = \sqrt[3]{4} + \sqrt[3]{2} + 1, \quad b_1^{(2)} = 3, \quad b_2^{(2)} = 3.$$

The corresponding matrices are

$$B_0 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 3 \end{bmatrix}$$

(actually the JPA finds the inverses of the B_i). One finds the $a_i^{(j)}$ repeat after the third step above: $B_3 = B_4 = B_5 \dots$.

B_2 has characteristic polynomial

$$x^3 - 3x^2 - 3x - 1,$$

with root $\lambda = \sqrt[3]{4} + \sqrt[3]{2} + 1$; and the vector $(1, \sqrt[3]{2} + 2, \sqrt[3]{4} + \sqrt[3]{2} + 1)$ is an eigenvector for B_2 . For large k , the column vectors of B_2^k will satisfy

$$\frac{y}{x} \approx \sqrt[3]{2} + 2, \quad \frac{z}{x} \approx \sqrt[3]{4} + \sqrt[3]{2} + 1.$$

Our interest here is: how good is the approximation after a number of steps comparable to the algorithm of this paper? The total length of A (the sum of the exponents on the E_{ij}) is 18. On the other hand, we may write

$$B_2 = E_{12}^3 \cdot E_{13}^3 \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

which has, roughly speaking, length 6. Consider then B_2^3 . We find

$$B_2^3 = \begin{bmatrix} 1 & 3 & 12 \\ 3 & 10 & 39 \\ 3 & 12 & 40 \end{bmatrix}.$$

By calculator, $\sqrt[3]{2} + 2 \approx 3.259921$, $\sqrt[3]{4} + \sqrt[3]{2} + 1 \approx 3.8473221$, and the quotients y/x , z/x from B_2^3 are

$$\frac{3}{1} = 3, \quad \frac{10}{3} = 3.33 \dots, \quad \frac{39}{12} = 3.25 \dots,$$

$$\frac{3}{1} = 3, \quad \frac{12}{3} = 4, \quad \frac{40}{12} = 3.33 \dots$$

The best is $39/12 = 3.25$ which is within .01 of $\sqrt[3]{2} + 2$.

The JPA in this case is inferior as a tool of approximation, and superior in demonstrating periodicity. But we also have the following example.

(iv) In [1, p. 69] it is stated that the pair $(\sqrt[3]{4}, \sqrt[3]{16})$ has not yet been shown to have a periodic JPA, although the calculations have been carried out to 150 steps. We have found that our procedure produces periodicity in 91 steps (with $x_{91} = x_1$, $y_{91} = y_1$).

(v) For our last example, we give a table of both the JPA and our algorithm, for vectors $(1, \lambda, \lambda^2)$ where λ is the largest root of the listed polynomial. All these polynomials have their roots in the same field (all have three real roots).

Notation. JPA, of course, stands for the Jacobi-Perron algorithm, and TVA for the algorithm described in this paper. T is the following matrix:

$$T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The steps of the JPA are marked off with parentheses (); the periodic parts of both algorithms are enclosed in square brackets.

The program used to calculate these is straightforward, and can only handle integers up to sixteen decimal digits long. Nevertheless, the overflows for $P_7(x)$ and $P_8(x)$ are curious, in view of the extreme shortness and simplicity of the algorithms for the other $P_i(x)$.

i	$P_i(x)$	JPA	TVA
1	$x^3 - 3x^2 - 4x - 1$	$(E_{12}^4 E_{13}^6 T)(E_{12}^8 E_{13}^2 T)[E_{12}^9 E_{13}^2 T]$	$E_{23}^4 [E_{12}^4 E_{31}^5 E_{23}^4 E_{12}^5 E_{31}^5 E_{23}^4]$ $= E_{23}^4 [(E_{12}^4 E_{31}^5 T)^3]$
2	$x^3 + 4x^2 + 3x - 1$	$(T)(E_{13}^4 T)[(E_{13}^4 T)(E_{13}^5 T)]$	$E_{21}^4 [E_{32}^4 E_{13}^5 E_{21}^4 E_{32}^5 E_{13}^4 E_{21}^5]$ $= E_{21}^4 [(E_{32}^4 E_{13}^5 T)^3]$
3	$x^3 - 6x^2 + 5x - 1$	$(E_{12}^5 E_{13}^2 T)(E_{12}^1 E_{13}^2 T)[E_{12}^9 E_{13}^2 T]$	$E_{23}^5 [E_{12}^5 E_{31}^4 E_{23}^5 E_{12}^4 E_{31}^4 E_{23}^5]$ $= E_{23}^5 [(E_{12}^5 E_{31}^4 T)^3]$
4	$x^3 - 5x^2 + 6x - 1$	$(E_{12}^3 E_{13}^1 T)(E_{12}^2 E_{13}^4 T)[(E_{13}^5 T)(E_{13}^4 T)]$	$E_{23}^3 [E_{12}^3 E_{31}^3 E_{23}^3 E_{12}^3 E_{31}^3 E_{23}^3]$ $= E_{23}^3 [(E_{12}^3 E_{31}^3 T)^3]$
5	$x^3 + 2x^2 - x - 1$	$(T)(E_{13} T)[(E_{13} T)(E_{13}^3 T)]$	$E_{21} [E_{32} E_{13}^3 E_{21} E_{32}^3 E_{13}^3 E_{21}^3]$ $= E_{21} [(E_{32} E_{13}^3 T)^3]$
6	$x^3 + x^2 - 2x - 1$	$(E_{12} E_{13} T)(E_{12}^2 E_{13}^4 T)[(E_{13}^4 T)(E_{13}^5 T)]$	$E_{23} [E_{12} E_{31}^3 E_{23} E_{12}^3 E_{31}^3 E_{23}^3]$ $= E_{23} [(E_{12} E_{31}^3 T)^3]$
7	$x^3 - 20x^2 - 9x - 1$	machine overflow at step 3	machine overflow at step 14
8	$x^3 + 9x^2 + 20x - 1$	$(T)(E_{13}^2 T)[E_{12}^9 E_{13}^2 T]$	machine overflow at step 12

3. A Geometric Approach to Simple Continued Fractions. Our algorithm was motivated by the actual construction of the continued fraction algorithm when it is viewed as an aspect of the geometry associated with 2×2 matrices with positive integer entries. To be more precise, we now give a brief summary of our geometric interpretation of continued fractions. (This can be done in several ways ([2], [3])); but the particular approach given here, putting emphasis on linear algebra, seems to be new.)

Put

$$E_{12} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad E_{21} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Put

$$Q = \{(x, y) | x \geq 0, y \geq 0\}, \quad S_1 = \{(1, y) | 0 \leq y \leq 1\},$$

$$S_2 = \{(x, 1) | 0 \leq x \leq 1\}, \quad S = S_1 \cup S_2.$$

For each nonsingular 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with nonnegative integer entries, define

a map $A^*: S \rightarrow S$ as follows. Let $v = (x, y) \in S$. Then

$$Av = (ax + by, cx + dy) = (z, w) \in Q.$$

Both z and w are nonnegative, and one of them must be positive. Let $M = \max\{z, w\}$. Then

$$A^*v = \left(\frac{z}{M}, \frac{w}{M}\right) \in S.$$

Thus we have: E_{ij} is the result of performing the elementary operation “add row i to row j ” on the identity matrix I_2 , Q is the first quadrant of R^2 , S_1 and S_2 are the sides of the unit square in Q , $A(Q) \subset Q$, and $A^*(v)$, for v in S , is the intersection of the line through $(0, 0)$ and $A(v)$, with S .

We list the following statements without proof. All these statements have very direct and elementary proofs.

3.1. For A, B 2×2 matrices with nonnegative integer entries, and $v \in S$, we have $(AB)^*(v) = A^*(B^*v)$.

3.2. $E_{12}^*(S) \cap E_{21}^*(S) = \{(1, 1)\}$.

3.3. $E_{12}^*(S) = S_2$ and $E_{21}^*(S) = S_1$,

3.4. $E_{12}^{n*}(S_1) \cap E_{12}^{(n+1)*}(S_1) = \{(1/(n+1), 1)\}$, $E_{21}^{n*}(S_2) \cap E_{21}^{(n+1)*}(S_2) = \{(1, 1/(n+1))\}$,

3.5. $S_2 = \bigcup_{n=1}^{\infty} \{E_{12}^{n*}(S_1)\}$, $S_1 = \bigcup_{n=1}^{\infty} \{E_{21}^{n*}(S_2)\}$,

and the sets $E_{ij}^{k*}(S_i)$ are intervals in S_j with disjoint interiors; if $|r - s| > 1$, then $E_{ij}^{r*}(S_i) \cap E_{ij}^{s*}(S_i) = \emptyset$.

3.6. Let $\langle 0, n_1, n_2, \dots, n_k \rangle$ denote the continued fraction

$$\frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \dots + \frac{1}{n_k}}}}$$

of length k (where the n_i are positive integers); and let P_k be the set of all continued fractions of length k in $(0, 1]$. Then the family $\{E_{ij}^{n*}(S_i) | n = 1, 2, \dots\}$ induces the partition P_1 on S_j .

3.7. Let T_1 be the partition of S given by:

$$T_1 = \{E_{12}^{n*}(S_1) | n = 1, 2, \dots\} \cup \{E_{21}^{n*}(S_2) | n = 1, 2, \dots\}$$

and for $k = 2, 3, \dots$ define T_k recursively by:

$$T_k = \left\{ E_{12}^{n*}(X) \left| \begin{array}{l} n = 1, 2, \dots \\ X \in T_{k-1}, X \subset S_1 \end{array} \right. \right\} \cup \left\{ E_{21}^{n*}(X) \left| \begin{array}{l} n = 1, 2, \dots \\ X \in T_{k-1}, X \subset S_2 \end{array} \right. \right\}.$$

Then T_k induces the partition P_k on the sides S_1, S_2 (respectively) of S , and the members of T_k are intervals with disjoint interior.

3.8. Suppose $x > 0$ is a real number with continued fraction expansion to k places:

$$x = n_1 + \frac{1}{n_2 + \frac{1}{\ddots + \frac{1}{n_k + x_{k+1}}}}$$

Put x^* equal to the intersection of the line L through $(0, 0)$ and $(1, x)$ with S . Then x^* is contained in the following member of T_k :

$$X = (E_{i_1 j_1}^{n_1} E_{i_2 j_2}^{n_2} E_{i_3 j_3}^{n_3} \cdots E_{i_k j_k}^{n_k})^*(S_{i_k}),$$

where $(i_r, j_r) \neq (i_{r+1}, j_{r+1})$; $(i_1, j_1) = (1, 2)$ when $x \geq 1$, and $(i_1, j_1) = (2, 1)$ when $x < 1$. If A is the matrix $= E_{i_1 j_1}^{n_1} \cdots E_{i_k j_k}^{n_k}$, then the entries of A are the numerators and denominators of the k th and $(k - 1)$ st convergents to x , arranged in the order described in 3.9.

3.9.

$$E_{12}^{n_1} \cdots E_{ij}^{n_k} = \begin{cases} \begin{bmatrix} q_k & q_{k-1} \\ p_k & p_{k-1} \end{bmatrix}, & k \text{ odd,} \\ \begin{bmatrix} q_{k-1} & q_k \\ p_{k-1} & p_k \end{bmatrix}, & k \text{ even,} \end{cases}$$

$$E_{21}^{n_1} \cdots E_{ij}^{n_k} = \begin{cases} \begin{bmatrix} p_k & p_{k-1} \\ q_k & q_{k-1} \end{bmatrix}, & k \text{ even,} \\ \begin{bmatrix} p_{k-1} & p_k \\ q_{k-1} & q_k \end{bmatrix}, & k \text{ odd,} \end{cases}$$

where

$$\frac{p_k}{q_k} = n_1 + \frac{1}{n_2 + \frac{1}{\ddots + \frac{1}{n_k}}},$$

$p_0 = 1, q_0 = 0$.

3.10. The statement “the continued fraction algorithm converges” may be proven by using the fact that $\text{mesh } T_k \leq 1/k(k + 1) \rightarrow 0$ as $k \rightarrow \infty$.

4. The Three-Dimensional Case. In this section we generalize the situation summarized in Section 3 to R^3 . It will be seen that several generalizations are possible; our choice was guided by a wish to mimic as closely as possible the statements 3.1–3.10 of Section 3.

Corresponding to E_{12} and E_{21} in the two-dimensional case, we consider here the six elementary matrices E_{ij} defined by: E_{ij} is a 3×3 matrix with

$$(E_{ij})_{rs} = \begin{cases} 1, & r = s, \\ 1, & r = j, s = i, \\ 0, & \text{otherwise,} \end{cases}$$

that is, E_{ij} is the result of performing the elementary row operation “add row i to row j ” on the identity matrix I_3 .

Now in the two-dimensional case, E_{12} and E_{21} are noncommuting matrices which neatly correspond one to each side of the unit square in the first quadrant (Property 3.3). In the three-dimensional case, we have six matrices E_{ij} , some of which commute with others, and we have only three “faces” of the unit cube in the first octant. Evidently we cannot generalize the two-dimensional case *mutatis mutandis*.

To facilitate further discussion, we define the following sets in the first octant of R^3 :

$$\begin{aligned} \mathcal{O} &= \{(x, y, z) \mid x \geq 0, y \geq 0, z \geq 0\}, \\ F_1 &= \{(1, y, z) \mid 0 \leq y, z \leq 1\}, \\ F_2 &= \{(x, 1, z) \mid 0 \leq x, z \leq 1\}, \\ F_3 &= \{(x, y, 1) \mid 0 \leq x, y \leq 1\}, \end{aligned}$$

so that \mathcal{O} is the first octant, and F_1, F_2, F_3 are the three faces of the unit cube in \mathcal{O} . We may occasionally refer to, e.g. F_1 as the “ x -face” (it is the face with x -values all equal to one), and so on. We put

$$F = F_1 \cup F_2 \cup F_3.$$

In analogy with Section 3, for each 3×3 nonsingular matrix A with nonnegative integer entries, define a mapping $A^*: F \rightarrow F$ as follows. If $v = (x, y, z) \in F$, then $Av = w = (r, s, t)$ is a vector in \mathcal{O} with nonnegative entries, at least one of which is positive. Put $M = \max\{r, s, t\}$ and define A^*v by:

$$A^*v = \left(\frac{r}{M}, \frac{s}{M}, \frac{t}{M} \right),$$

that is, A^*v is the intersection of the line along v with F . Again it is clear that

$$(AB)^*v = A^*(B^*v).$$

Consider now, for example, the sets $E_{12}^{k*}(F_1), E_{12}^{k*}(F_2), E_{12}^{k*}(F_3)$. We have

$$\begin{aligned} X_k &= E_{12}^{k*}(F_1) = \left\{ \left(\frac{1}{k+y}, 1, \frac{z}{k+y} \right) \mid (1, y, z) \in F_1 \right\}, \\ Y_k &= E_{12}^{k*}(F_2) = \left\{ \left(\frac{x}{kx+1}, 1, \frac{z}{kx+1} \right) \mid (x, 1, z) \in F_2 \right\}, \\ Z_k &= E_{12}^{k*}(F_3) = \{(x, kx+y, 1) \mid (x, y, 1) \in F_3; kx+y \leq 1\} \\ &\cup \left\{ \left(\frac{x}{kx+y}, 1, \frac{1}{kx+y} \right) \mid (x, y, 1) \in F_3; kx+y \geq 1 \right\}. \end{aligned}$$

Below, we have sketched the sets X_2, Y_2, Z_2 .

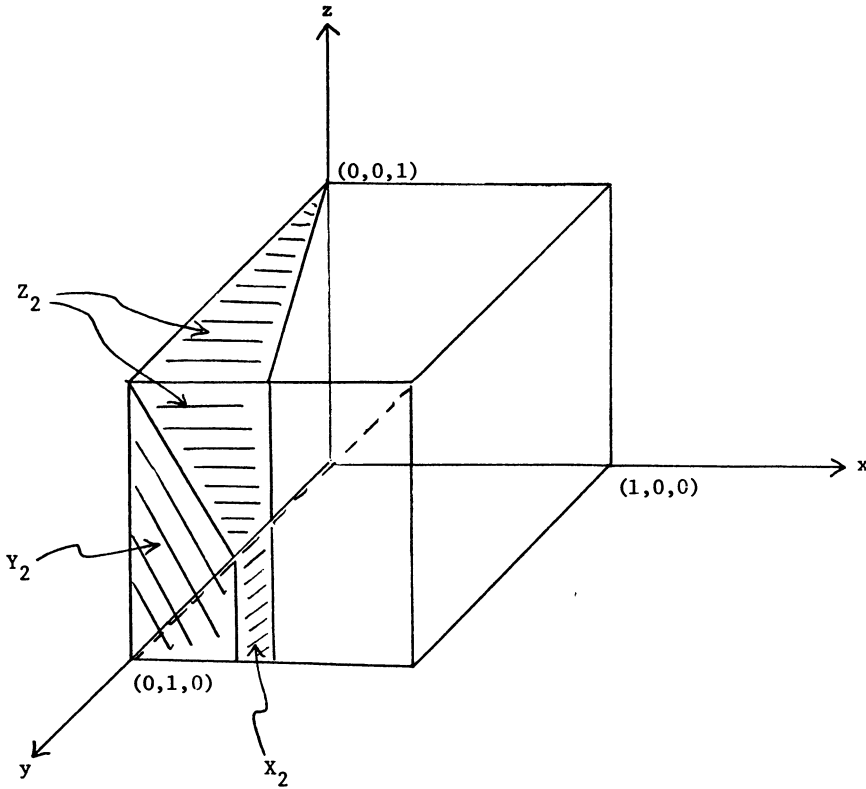


FIGURE 1

We wish to follow the situation of Section 3. Evidently we cannot have a close analogue of 3.2 and 3.3 because the corresponding sets in R^3 actually overlap, and no $E_{ij}^*(F)$ covers an entire face F_r .

Wishing an analogy of 3.4 and 3.5 we eliminate sets $E_{ij}^*(F_t)$ with $t \neq i$. For example, $E_{12}^{(k+j)*}(F_3) \subset E_{12}^{k*}(F_3)$, and $E_{12}^{(k+j)*}(F_2) \subset E_{12}^{k*}(F_2)$ for all $j = 1, 2, \dots$, whereas 3.4 and 3.5 suggest we want the intersection to be a line or a point. (Incidentally, analogy with 3.5 also suggests that the desired sets should lie entirely on one face; on this score alone we would eliminate sets $E_{ij}^*(F_t)$ with $t \neq i, t \neq j$.)

Fortunately, the collection

$$T_1 = \{E_{ij}^{k*}(F_i) \mid k = 1, 2, \dots; i, j \in \{1, 2, 3\}, i \neq j\}$$

fits in nicely with our requirements: the members of T_1 are convex quadrilaterals with disjoint interior; T_1 partitions F , and T_1 induces the partition P_1 on six of the nine edges of F , and the partition $P_0 = \{0, 1\}$ on the remaining three edges of F .

There is another consideration here also: we wish eventually to have an algorithm, based on constructing successively finer partitions of F and "locating" a point of F in these successive partitions. Convergence will follow from an application of the Cantor

Intersection Theorem. An efficient algorithm then would arise by choosing the members of the partitions as small as possible (consistent with other requirements). It is obvious to the eye that the sets Y_k, Z_k are bigger than X_k . In fact, for all k , Y_k and Z_k always have an edge of length one, whereas all the edges of X_k have length $\leq 1/k$.

For the reasons outlined above, we shall use only the collection T_1 , in order to have a situation analogous to that described in Section 3. This amounts to restricting the domain of E_{ij}^* to the face F_i ; this restriction is understood in the remainder of this paper.

We emphasize at this point, that our aim is to generalize the work in Section 3 in as simple and natural a way as possible; thus we do not consider sets which might arise in other ways.

Before describing our algorithm, we first prove some results about the maps E_{ij}^* . These results demonstrate further analogies to the two-dimensional case, and also help to motivate the derivation of the algorithm.

4.1. LEMMA. Let $i, j \in \{1, 2, 3\}, i \neq j$.

(a) E_{ij}^{n*} is a continuous 1-1 function on F , and $E_{ij}^{n*}(F_i) \subset F_j$.

(b) If L is a line segment in F_i , then $E_{ij}^{n*}(L)$ is a line segment in F_j ;

(c) If K is a convex polygon with vertices X_1, X_2, \dots, X_k in F_i , then $E_{ij}^{n*}(K)$ is a convex polygon in F_j with vertices $E_{ij}^{n*}(X_1), \dots, E_{ij}^{n*}(X_k)$.

Proof. (a) It is clear that E_{ij}^{n*} is continuous and 1-1 on F . Suppose, e.g. that $(i, j) = (1, 2)$ and suppose $(1, y, z) \in F_1$. Then

$$E_{12}^{n*}(1, y, z) = \left(\frac{1}{y+n}, 1, \frac{z}{y+n} \right) \in F_2,$$

so that $E_{12}^{n*}(F_1) \subset F_2$; similarly for all other choices of $i, j \in \{1, 2, 3\}, i \neq j$.

(b) Suppose again that $(i, j) = (1, 2)$ and that

$$L = \{(1, y, z) \mid 0 \leq y, z \leq 1; ay + bz + c = 0\}$$

is a line segment in F_1 . Then

$$E_{12}^{k*}(L) = \left\{ \left(\frac{1}{y+k}, 1, \frac{z}{y+k} \right) \mid (1, y, z) \in L \right\} \subset F_2.$$

The equation $ay + bz + c = 0$ may be rewritten

$$a(y+k) + bz + (c-ak) = 0,$$

and putting $x' = (y+k)^{-1}$ and $z' = z(y+k)^{-1}$, it is clear that the points $(x', 1, z')$ in $E_{12}^{k*}(L)$ satisfy the equation

$$(c-ak)x' + bz' + a = 0,$$

that is, these points are collinear. Now from part (a) (continuity) $E_{12}^{k*}(L)$ is a line segment in F_2 ; similarly for the other choices of (i, j) .

Part (c) follows immediately from (b).

We next state

4.2. THEOREM. *The sets $E_{ij}^{k*}(F_i)$ are convex quadrilaterals with disjoint interiors which partition F . In addition, we have*

(a) $E_{ij}^{k*}(F_i) \cap E_{ij}^{n*}(F_i) = \emptyset$ unless $|n - k| \leq 1$. If $|n - k| = 1$, the intersection is a line segment in F_j , parallel to one edge of F_j .

(b) $E_{ij}^{k*}(F_i) \cap E_{ij}^{n*}(F_r) = \emptyset$ unless $|n - k| \leq 1$. If $|n - k| = 1$, the intersection is a line segment on a diagonal of F_j ; if $|n - k| = 0$, the intersection is a point on a diagonal of F_j .

(c) $E_{ij}^*(F_i) \cap E_{ji}^*(F_j)$ is on the common edge of F_i and F_j .

(d) $E_{ij}^*(F_i) \cap E_{rs}^*(F_r) = \{(1, 1, 1)\}$ if $(r, s) \neq (j, i)$.

(e) In all other cases, $E_{ij}^{k*}(F_i) \cap E_{rs}^{n*}(F_r) = \emptyset$.

Proof. The proof is computational, using the fact that

$$E_{12}^{k*}(F_1) = \left\{ \left(\frac{1}{k+y}, 1, \frac{z}{k+y} \right) \mid (1, y, z) \in F_1 \right\}$$

and the analogous descriptions for the other $E_{ij}^{k*}(F_i)$.

We give here a sketch of the x -face F_1 partitioned by the sets $E_{21}^{k*}(F_2)$ and $E_{31}^{k*}(F_3)$.

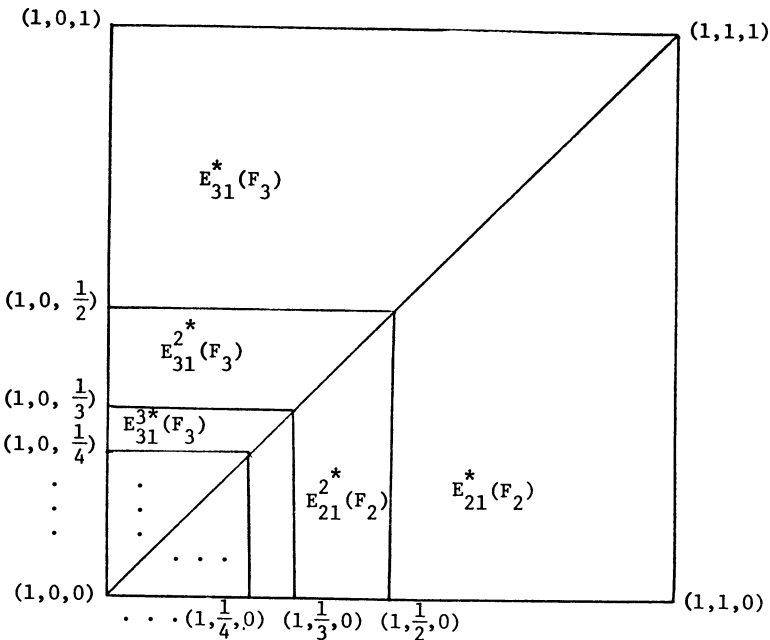


FIGURE 2

If an edge of F intersects a coordinate plane, call it an outer edge; otherwise call it an inner edge. We have

4.3. COROLLARY. *The partition of Theorem 4.2 induces the partition $P_1 = \{1/n \mid n = 1, 2, \dots\}$ on the outer edges of F , and the partition $P_0 = \{0, 1\}$ on the inner edges.*

Let

$$T_1 = \{E_{ij}^{k*}(F_i) \mid i, j \in \{1, 2, 3\}, i \neq j; k = 1, 2, \dots\}.$$

We shall define inductively, in analogy with the work in Section 2, successive partitions T_k such that T_k refines T_{k-1} , etc.

4.4. *Definition.* For each $k = 2, 3, \dots$ define a collection T_k as follows:

$$T_k = \left\{ E_{ij}^{n*}(X) \mid \begin{array}{l} i, j \in \{1, 2, 3\}; i \neq j, \\ n = 1, 2, \dots; X \subset F_i, X \in T_{k-1} \end{array} \right\}.$$

Thus, for example,

$$T_2 = \{(E_{ij}^n \cdot E_{ki}^t)^*(F_k)\}.$$

For convenience, we state

4.5. **THEOREM.** For each $k = 2, 3, \dots$, the collection T_k is a proper refinement of T_{k-1} . T_k partitions F into convex quadrilaterals with disjoint interiors.

Proof. The proof is immediate from Lemma 4.1 and the definition of T_k .

Derivation of the Algorithm. Let x_0, y_0 be nonnegative real numbers. Put v_0 equal to the intersection of the line through $(0, 0, 0)$ and $(1, x_0, y_0)$ with F . Since T_1 partitions F , we must have v_0 in one of the members of T_1 , say $v_0 \in (E_{i_1 j_1}^1)^*(F_{i_1})$. Then we may write

$$v_0 = (E_{i_1 j_1}^1)^*(v_1); \quad v_1 \in F_{i_1}.$$

Now $v_1 = (E_{i_2 j_1}^2)^*(v_2)$ for some $v_2 \in F_{i_2}$, $v_2 = (E_{i_3 j_2}^3)^*(v_3)$ for some $v_3 \in F_{i_3}$, and so on. The algorithm consists essentially of computing the entries of the successive v_i and keeping track of the $E_{i_k j_k}^k$.

5. Convergence. Recall that the diameter of a (convex) quadrilateral in the plane is the length of its longest diagonal. Our procedure in this section may be roughly described as follows: Given a set X in the partitioning set T_k . Then X is a convex quadrilateral contained in one of the faces F_1, F_2, F_3 . Now the set $\{Y \in T_{k+2} \mid Y \subset X\}$ partitions the set X , and it is a comparatively simple matter (though tedious) to show that if $Y \in T_{k+2}$ and $Y \subset X$, then the diameter of Y satisfies $\text{diam } Y < (2/3) \text{diam } X$.

We begin with some preliminary lemmas.

5.1. **LEMMA.** Let $X \in T_k$ and suppose that $X \subset F_i$ ($i \in \{1, 2, 3\}$). Then there exists a matrix A such that

$$A = E_{i_1 j_1}^{n_1} E_{i_2 j_2}^{n_2} \cdots E_{i_k j_k}^{n_k},$$

where the n_r are positive integers, $j_r = i_{r-1}$ ($r = 2, \dots, k$), $j_1 = i$, and $A^*(F_{i_k}) = X$.

Proof. The proof follows immediately from the definition of T_k . If $X \in T_k$, then $X = E_{i_1 j_1}^{n_1}(U)$ for some $U \in T_{k-1}$, $U \subset F_{i_1}$; and $X \subset F_i$ implies $j_1 = i$. Now $U = E_{i_2 j_2}^{n_2}(W)$ for some $W \in T_{k-2}$, $W \subset F_{i_2}$. Then $U \subset F_{i_1}$ implies $j_2 = i_1$, and so on.

5.2. **LEMMA.** If $X \in T_k$, then X is a convex quadrilateral contained in one of the faces F_i ($i \in \{1, 2, 3\}$). If we write $X = A^*(F_{i_k})$ where A is the matrix defined

in Lemma 5.1, then letting E_1, E_2, E_3, E_4 be the edges of face F_{i_k} and D_1, D_2 its two diagonals, we have that the edges of X are given by $A^*(E_i)$ ($i = 1, 2, 3, 4$) and its two diagonals by $A^*(D_1)$ and $A^*(D_2)$.

Proof. This follows from Lemma 4.1(b) and (c).

5.3. LEMMA. If $X \in T_k$ and $Y \in T_{k+2}$, $Y \subset X$, and we write $X = A^*(F_{i_k})$ as in Lemma 4.2, then we have

$$Y = (AE_{i_t}^{n_{k+1}} E_{i_{t+1} i_t}^{n_{k+2}})^*(F_{i_{t+1}})$$

for some $i_t, i_{t+1} \in \{1, 2, 3\}$; and n_{k+1}, n_{k+2} positive integers; that is,

$$Y = A^*(U)$$

for some $U \in T_2$.

Proof. The proof is immediate from the definition of T_{k+2} .

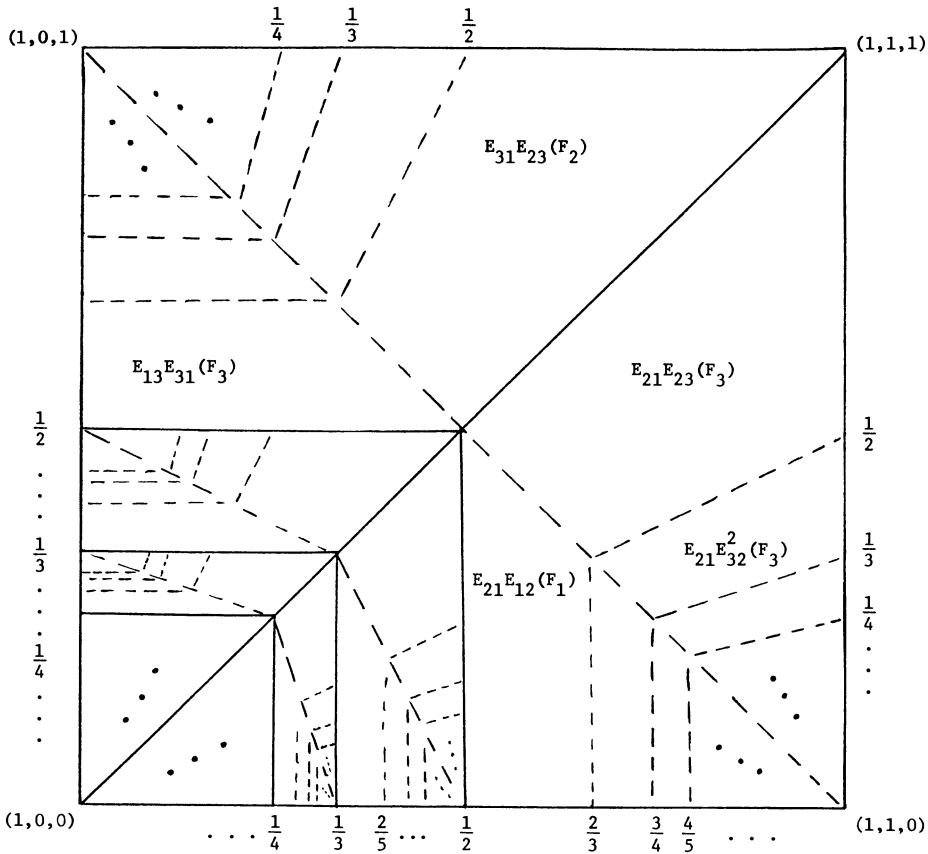


FIGURE 3. T_2 on the x -face

We append here a sketch of the x -face as partitioned by T_2 (Figure 3). The solid lines surround members of T_1 , members of T_2 have at least two edges on the dotted lines. (Cf. Figure 2.)

Suppose now that $X = A^*(F_{i_k})$ as in Lemma 5.1, and for convenience suppose $X \subset F_1$ and $i_k = 2$; $X = A^*(F_2)$. Now A is a matrix with nonnegative integer entries, say

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}.$$

The vertices of F_2 are: $(0, 1, 0)$, $(0, 1, 1)$, $(1, 1, 0)$, $(1, 1, 1)$ and so the vertices of $A^*(F_2)$ are points in F_1 given by:

$$\begin{aligned} v_1 &= A^*(0, 1, 0) = \left(1, \frac{b_2}{b_1}, \frac{b_3}{b_1}\right), \\ v_2 &= A^*(0, 1, 1) = \left(1, \frac{b_2 + c_2}{b_1 + c_1}, \frac{b_3 + c_3}{b_1 + c_1}\right), \\ v_3 &= A^*(1, 1, 0) = \left(1, \frac{a_2 + b_2}{a_1 + b_1}, \frac{a_3 + b_3}{a_1 + b_1}\right), \\ v_4 &= A^*(1, 1, 1) = \left(1, \frac{a_2 + b_2 + c_2}{a_1 + b_1 + c_1}, \frac{a_3 + b_3 + c_3}{a_1 + b_1 + c_1}\right). \end{aligned}$$

We remind the reader that the definition of A assures that we do have $b_2 < b_1$, $b_3 < b_1$, $b_2 + c_2 < b_1 + c_1$, etc. so that the second and third coordinates of v_i are between 0 and 1.

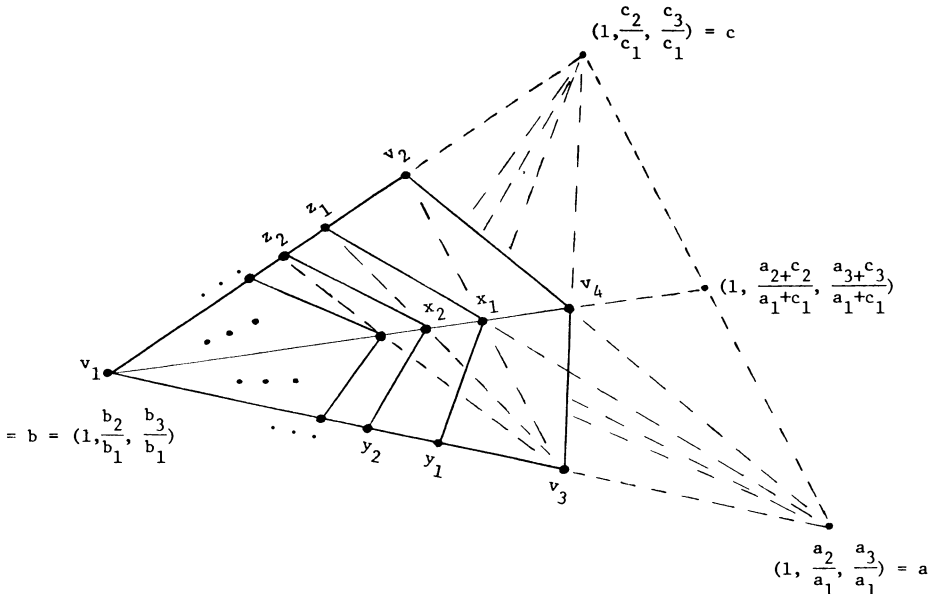


FIGURE 4

In Figure 4, we sketch the set $X = A^*(F_{i_k})$ as partitioned into the sets $Y = A^*(U)$ for $U \in T_2$, $U \subset F_{i_k}$. As an aid in sketching we assume $a_2, a_3 < a_1$ and $c_2, c_3 < c_1$ so that in fact $A^*(F) \subset F_1$.

The points x_1, x_2, \dots are given by $A^*(1/n, 1, 1/n)$; either the points z_1, z_2, \dots are $A^*(1/n, 1, 0)$ and the points y_1, y_2, \dots are $A^*(0, 1, 1/n)$; or vice versa. The heavy lines give the outlines of the sets $Y = A^*(U)$ for $U \in T_1$.

For $A^*(U)$, $U \in T_2$, the picture becomes

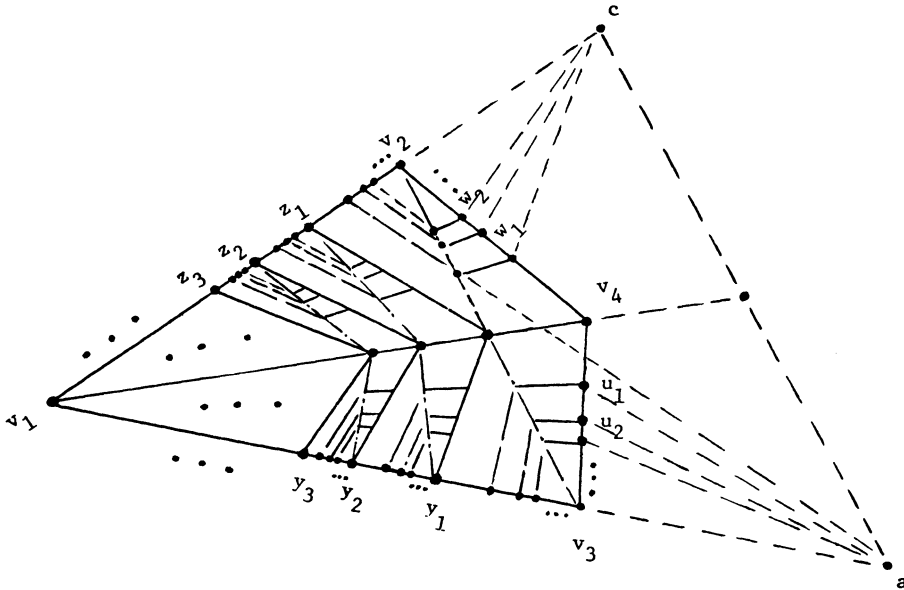


FIGURE 5

The new points w_i are the intersections of the lines through c and y_i with the edge v_2v_4 and similarly for the points u_i .

We now state

5.5. THEOREM. *Let $X = A^*(F_{i_k})$ as in Lemma 5.1, and let $Y = A^*(U)$ where $U \in T_2$, $U \subset F_{i_k}$. Then $\text{diam } Y < (2/3) \text{diam } X$.*

Proof. The idea of the proof is geometrically quite simple. If $Y = A^*(U)$, then Y is one of the quadrilaterals outlined in Figure 5. If L is an edge or a diagonal of Y , it is indicated in Figure 5 that Y lies on a line segment L' in X , extending from one edge or vertex of X to another edge or vertex of X , and it appears that L is always strictly shorter than the corresponding line L' .

Without loss of generality, we may assume $X \subset F_1$ and $i_k = 2$, that is, $A^*(F_2) \subset F_1$. Let $U \in T_2$, $U \subset F_2$. It is clear from Figure 3 that either

- (a) U has two sides parallel to an edge of F_2 and to each other,
- (b) U does not have two sides parallel.

In case (a) we have either $U = (E_{12}^t E_{21}^n)^*(F_2)$ or $U = (E_{32}^t E_{23}^n)^*(F_2)$, while in case (b) we have either $U = (E_{12}^t E_{31}^n)^*(F_3)$ or $U = (E_{32}^t E_{13}^n)^*(F_1)$.

For case (a), consider for example, U with vertices

$$u_1 = \left(\frac{1}{t + 1/(n + 1)}, 1, \frac{1/(n + 1)}{t + 1/(n + 1)} \right), \quad u_2 = \left(\frac{1}{t + 1/(n + 1)}, 1, 0 \right),$$

$$u_3 = \left(\frac{1}{t + 1/n}, 1, \frac{1/n}{t + 1/n} \right), \quad u_4 = \left(\frac{1}{t + 1/n}, 1, 0 \right).$$

Then the diagonals of U run from u_1 to u_4 and from u_2 to u_3 . Now the segment $\overline{u_1 u_4}$ lies on the segment L_1 from

$$p_1 = \left(\frac{1}{t + 1/n}, 1, 0 \right) \quad \text{to} \quad p_2 = \left(\frac{n + 1}{nt + 1}, 1, 1 \right)$$

and the segment $\overline{u_2 u_3}$ lies on the segment L_2 from

$$q_1 = \left(\frac{1}{t + 1/(n + 1)}, 1, 0 \right) \quad \text{to} \quad q_2 = \left(\frac{n/(n + 1)}{t + 1/(n + 1)}, 1, 1 \right).$$

Writing

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix},$$

we have, e.g.

$$A^*(u_1) = \left(1, \frac{(n + 1)a_2 + \{(n + 1)t + 1\}b_2 + c_2}{(n + 1)a_1 + \{(n + 1)t + 1\}b_1 + c_1}, \frac{(n + 1)a_3 + \{(n + 1)t + 1\}b_3 + c_3}{(n + 1)a_1 + \{(n + 1)t + 1\}b_1 + c_1} \right)$$

and, similarly, one may write out $A^*(u_2)$, $A^*(u_3)$, $A^*(u_4)$. (We have found it simpler to keep the numerators and denominators of the entries of $A^*(u_i)$ clear of fractions.)

Put

$$S_2 = (a_2 b_1 - a_1 b_2) + (nt + 1)(c_2 b_1 - b_2 c_1) + n(a_1 c_2 - a_2 c_1),$$

$$S_3 = (a_3 b_1 - a_1 b_3) + (nt + 1)(c_3 b_1 - b_3 c_1) + n(a_1 c_3 - a_3 c_1),$$

$$R_2 = (a_2 b_1 - a_1 b_2) + (n + 1)(a_2 c_1 - s_1 c_2) + \{(n + 1)t + 1\}(b_2 c_1 - c_2 b_1),$$

$$R_3 = \{a_3 b_1 - a_1 b_3\} + (n + 1)(a_3 c_1 - a_1 c_3) + \{(n + 1)t + 1\}(b_3 c_1 - c_3 b_1).$$

One computes easily that

$$d(A^*(u_1), A^*(u_4)) = \frac{\sqrt{S_2^2 + S_3^2}}{((n + 1)a_1 + \{(n + 1)t + 1\}b_1 + c_1)(na_1 + (nt + 1)b_1)},$$

$$d(A^*(p_1), A^*(p_2)) = \frac{(nt + 1)\sqrt{S_2^2 + S_3^2}}{(na_1 + (nt + 1)b_1)((n + 1)a_1 + (nt + 1)b_1 + (nt + 1)c_1)},$$

$$d(A^*(u_2), A^*(u_3)) = \frac{\sqrt{R_2^2 + R_3^2}}{(na_1 + (nt + 1)b_1 + c_1)((n + 1)a_1 + \{(n + 1)t + 1\}b_1)}.$$

$$d(A^*(q_1), A^*(q_2)) = \frac{\sqrt{R_2^2 + R_3^2}}{((n+1)a_1 + \{(n+1)t+1\}b_1)(na_1 + \{t(n+1)+1\}b_1 + \{t(n+1)+1\}c_1)}.$$

Then

$$\begin{aligned} \frac{d(A^*(u_1), A^*(u_4))}{d(A^*(p_1), A^*(p_2))} &= \frac{(n+1)a_1 + (nt+1)b_1 + (nt+1)c_1}{(nt+1)[(n+1)a_1 + ((n+1)t+1)b_1 + c_1]} \\ &= \frac{1}{1 + \frac{nt[(n+1)a_1 + ((n+1)t+1)b_1]}{(n+1)a_1 + (nt+1)b_1 + (nt+1)c_1}} \\ &= \frac{1}{1 + \alpha}. \end{aligned}$$

We must now show $\alpha > \frac{1}{2}$. To do this we shall use the fact that A has the form described in Lemma 5.1. If $A = BE_{23}$, then column 2 of A is elementwise larger than column 3 of A ; then $b_1 > c_1$ and we have $\alpha > 1 > \frac{1}{2}$. If $A = E_{i_1 j_1}^{n_1} \cdots E_{i_k j_k}^{n_k}$ has all $r, j_r = 1$ or 2, then $c_1 = 0$, and again $b_1 > c_1$. If $A = BE_{13}E_{21} \cdots E_{21}$, then column 1 is greater than column 3; if $A = BE_{23}E_{12} \cdots E_{21}$, then column 2 is greater than column 3. Thus, either $c_1 = 0$, or one of a_1 and b_1 is greater than c_1 . In all cases $\alpha > \frac{1}{2}$.

A similar argument shows that

$$\frac{d(A^*(u_2), A^*(u_3))}{d(A^*(q_1), A^*(q_2))} < \frac{2}{3}$$

and again we must use the fact that our matrix A has either a_1 or b_1 greater than c_1 .

Case (b) is somewhat simpler than case (a), in that case (b) does not require any restrictions on A except that $A^*(F_2) \subset F_1$ and the entries of A are nonnegative integers. Let $U \in T_2$ fall under case (b). Then suppose U has vertices

$$\begin{aligned} u_1 &= \left(\frac{1}{t}, 1, \frac{1}{nt}\right), & u_2 &= \left(\frac{1}{t + \frac{1}{n}}, 1, \frac{1}{nt+1}\right), \\ u_3 &= \left(\frac{1}{t}, 1, \frac{1}{(n+1)t}\right), & u_4 &= \left(\frac{1}{t + \frac{1}{n+1}}, 1, \frac{1}{(n+1)t+1}\right). \end{aligned}$$

The diagonals of U are the segments $\overline{u_1 u_4}$ and $\overline{u_2 u_3}$.

The segment $\overline{u_1 u_4}$ lies on the line from $P_1 = (1/(t+1), 1, 0)$ to

$$P_2 = \begin{cases} \left(\frac{n+1}{t+1}, 1, 1\right), & n \leq t, \\ \left(1, 1, \frac{t}{n}\right), & n \geq t; \end{cases}$$

and the segment $\overline{u_2 u_3}$ lies on the line from Q_1 to Q_2 where

$$Q_1 = \left(1, 1, \frac{1}{n+1}\right), \quad Q_2 = \left(0, 1, \frac{1}{n+1}\right) \quad \text{if } t = 1,$$

$$Q_1 = \left(\frac{1}{t-1}, 1, 0\right), \quad Q_2 = \left(0, 1, \frac{1}{n+1}\right) \quad \text{if } t > 1.$$

The computation of the distances proceeds as before; when it comes time to take the ratios, the square roots cancel, and one has, e.g.

$$\frac{d(A^*(u_1), A^*(u_4))}{d(A^*(P_1), A^*(P_2))} = \frac{(a_1 + (t+1)b_1)[(n+1)a_1 + (t+1)b_1 + (t+1)c_1]}{(t+1)[na_1 + NTB_1 + c_1][a_1 + (t+1)b_1 + na_1 + NTB_1 + c_1]}$$

$$= \frac{1}{\frac{(t+1)[na_1 + NTB_1 + c_1]}{(n+1)a_1 + (t+1)b_1 + (t+1)c_1} \times \left(1 + \frac{na_1 + NTB_1 + c_1}{a_1 + (t+1)b_1}\right)}$$

$$= \frac{1}{\alpha \times \beta}.$$

It is clear that $\alpha > 1$ and that $\beta > 1 + 1/2 = 3/2$ so that $\alpha\beta > 3/2$ as desired.

Similarly, one shows that

$$\frac{d(A^*(u_2), A^*(u_3))}{d(A^*(Q_1), A^*(Q_2))} < \frac{2}{3}.$$

Similar computations may be used for any of the quadrilaterals U in case (a) or (b); the corresponding ratios are always $< 2/3$. This completes the proof.

Conjecture. If $X \in T_k$ and $Y \in T_{k+1}$, $Y \subset X$, then there is some fixed $\alpha < 1$ (perhaps $\alpha = 4/5$) such that $\text{diam } Y < \alpha \text{ diam } X$.

Open question: Find the best possible value of α for extension from T_k to T_{k+1} ; from T_k to T_{k+2} , etc.

Open question: The reader will recall that among continued fractions, the quantity

$$\alpha = 1 + \frac{1}{1 + \frac{1}{1 + \dots}} = \langle 1, 1, \dots, 1, 1, \dots \rangle = \frac{1 + \sqrt{5}}{2}$$

converges most slowly. Figure 3 (and pictures of T_k for $k > 2$) suggest strongly that something analogous happens in this algorithm. The largest sets in T_2 (having longest edges, longest diagonals, and largest area) are those of the form

$$(E_{ij}^1 E_{ri}^1)^*(F_r) \quad \text{with } r \neq j.$$

In T_3 , corresponding largest sets are of the form

$$(E_{i_1 j_1}^1 E_{i_2 i_1}^1 E_{i_3 i_2}^1)^*(F_{i_3}) \quad \text{with } i_2 \neq j_1, i_3 \neq i_1.$$

In fact, after 3 steps, the indices repeat, e.g.,

$$E_{12}E_{31}E_{23}E_{12}E_{31}E_{23} \cdots$$

One might expect that the matrix $A = E_{12}E_{31}E_{23}$ would be as interesting as, for $n = 2$:

$$E_{12}E_{21} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = B;$$

B has characteristic roots $1 + \alpha$, $1 + \beta$ ($\alpha = (1 + \sqrt{5})/2$, $\beta = (1 - \sqrt{5})/2$) and has $(1, \alpha)$ as a characteristic vector.

The author would like to thank the referee for his careful reading of the paper, and for helpful suggestions.

Department of Mathematics
University of North Carolina at Greensboro
Greensboro, North Carolina 27412

1. LEON BERNSTEIN, *The Jacobi-Perron Algorithm: Its Theory and Application*, Lecture Notes in Math., vol. 207, Springer-Verlag, Berlin and New York, 1971.
2. C. D. OLDS, *Continued Fractions*, Math. Assoc. Amer., New Mathematical Library, 1963.
3. HARRIS HANCOCK, *Development of the Minkowski Geometry of Numbers*, Dover, New York, 1964.