Some Facts About Continued Fractions That Should Be Better Known

J. O. Shallit

University of Waterloo

I. Introduction.

In this report I will give proofs of some simple theorems concerning continued fractions that are known to the *cognoscenti*, but for which proofs in the literature seem to be missing, incomplete, or hard to locate. In particular, I will give two proofs of the following "folk theorem": if θ is an irrational number whose continued fraction has bounded partial quotients, then any non-trivial linear fractional transformation of θ also has bounded partial quotients. The second proof is of interest because it uses the connection between continued fractions and finite automata first enunciated by G. N. Raney [R].

I will assume that the reader knows basic facts about continued fractions, at the level of [HW, Chapter X].

Note to the reader: It is intended that this report will eventually form a part of a longer article with the same name, written in collaboration with A. J. van der Poorten.

II. Some Notation.

In this report, θ denotes an irrational number. We let

$$\theta = [a_0, a_1, a_2, \dots]$$

be the simple continued fraction expansion of θ. We write

$$\frac{p_n}{q_n} = [a_0, a_1, \dots, a_n],$$

the n-th convergent. We define

$$a'_n = [a_n, a_{n+1}, a_{n+2}, \dots],$$

the n-th complete quotient. We also define

$$\|\theta\| = \min(\theta - \lfloor \theta \rfloor, \lceil \theta \rceil - \theta),$$

Research supported in part by a grant from NSERC .

the distance between θ and the nearest integer. Finally, we define

$$K(\theta) = \sup_{k \geq 1} \; a_k,$$

the largest partial quotient in the continued fraction for θ. If $K(\theta) < \infty$, we say θ has *bounded partial quotients.*

III. Badly approximable numbers

In this section, we obtain a precise relation between $K(\theta)$ and the degree to which θ can be approximated by rationals.

Theorem 1.

Let $r \geq 1$, and suppose $q\|q\theta\| \geq \frac{1}{r}$ for all $q \geq 1$. Then $K(\theta) < r$.

Proof.

Let p_n/q_n be a convergent to θ with $n \geq 1$. Since

$$\left| \frac{p_n}{q_n} - \theta \right| < \frac{1}{a_{n+1} q_n^2},$$

(see, e.g. [HW, Thm. 171]), we see

$$q_n |p_n - q_n \theta| < \frac{1}{a_{n+1}}.$$

Now clearly $\|q_n\theta\| \leq |p_n - q_n\theta|$, so

$$\frac{1}{r} \leq q_n \|q_n\theta\| \leq q_n |p_n - q_n\theta| < \frac{1}{a_{n+1}},$$

and thus $a_{n+1} < r$ for all $n \geq 0$, and hence $K(\theta) < r$. ■

Theorem 2.

Let $r \geq 1$ and suppose $K(\theta) \leq r$. Then

$$q\|q\theta\| \geq \frac{1}{r+2}$$

for all $q \geq 1$.

Proof.

We prove the contrapositive. Suppose there exists a $q \geq 1$ such that

$$q\|q\theta\| < \frac{1}{r+2}.$$

Let p be an integer such that $\|q\theta\| = |q\theta - p|$. Then $q|q\theta - p| < \frac{1}{r+2}$. Thus

$$\left|\theta - \frac{p}{q}\right| < \frac{1}{(r+2)q^2} < \frac{1}{2q^2},$$

from which it follows (see, e.g. [HW, Thm. 184]) that p/q is a convergent to θ, say $p/q = p_n/q_n$. Thus $p = ap_n, q = aq_n$ for some integer $a \geq 1$.

Now from [HW, Thm. 163], we have

$$|q_n\theta - p_n| = \frac{1}{a'_{n+1}q_n + q_{n-1}}.$$

Hence

$$\begin{aligned}\frac{1}{r+2} > q\|q\theta\| &= q|q\theta - p| \\ &= aq_n|aq_n\theta - ap_n| \\ &\geq q_n|q_n\theta - p_n| \\ &\geq \frac{1}{a'_{n+1} + \frac{q_{n-1}}{q_n}} \\ &\geq \frac{1}{(a_{n+1}+1)+1}.\end{aligned}$$

Thus we conclude $a_{n+1} > r$ and so $K(\theta) > r$. ■

A weaker form of Theorem 2, with a worse constant, was given by Hardy and Wright [HW, Thm. 187].

Theorems 1 and 2 above were essentially proved by W. M. Schmidt [S, pp. 22-23], but the proof provided is a little vague in spots. One can also deduce the theorems by filling in the details in [B, p. 47].

IV. Bounded partial quotients and linear fractional transformations

Definition.

We say θ is of type $< r$ if $q\|q\theta\| \geq \frac{1}{r}$ for all integers $q \geq 1$.

Theorem 3.

Let a, b be integers with $a \geq 1$, $|b| \geq 1$. If θ is of type $< M$, then $\frac{a}{b}\theta$ is of type $< |ab|M$.

Proof.

We prove the contrapositive. Assume

$$q\|q\frac{a}{b}\theta\| < \frac{1}{|ab|M}$$

for some $q \geq 1$.

Now there exists an integer p such that $\|q\frac{a}{b}\theta\| = |q\frac{a}{b}\theta - p|$. Thus

$$q\left|q\frac{a}{b}\theta - p\right| < \frac{1}{|ab|M},$$

and so, multiplying by $|ab|$, we get

$$qa|qa\theta - pb| < \frac{1}{M}.$$

Hence $qa\|qa\theta\| < \frac{1}{M}$. ■

Corollary 4.

$K(\frac{a}{b}\theta) < |ab|(K(\theta) + 2)$.

Theorem 5.

Let a, b be integers with $|a| \geq 1$, $b \geq 1$. If θ is of type $< M$, then $\theta + \frac{a}{b}$ is of type $< b^2 M$.

Proof.

We prove the contrapositive. Assume that

$$q\|q(\theta + \frac{a}{b})\| < \frac{1}{b^2 M}$$

for some $q \geq 1$.

Then there exists an integer p such that

$$\|q(\theta + \frac{a}{b})\| = |q(\theta + \frac{a}{b}) - p|.$$

Thus

$$q\left|q(\theta + \frac{a}{b}) - p\right| < \frac{1}{b^2 M}$$

and so, multiplying by b^2, we get

$$qb|qb\theta + qa - pb| < \frac{1}{M}.$$

Hence $qb\|qb\theta\| < \frac{1}{M}$. ■

Corollary 6.

$$K(\theta + \frac{a}{b}) < b^2(K(\theta) + 2).$$

I have not been able to find these results given explicitly in the literature. Theorems similar to Theorems 3 and 5 were given by Cusick and Mendès France [CMF]. Instead of studying $\sup_{q \geq 1} q\|q\theta\|$, they studied $\limsup_{q \to \infty} q\|q\theta\|$, which is somewhat more natural. Also see Perron [P].

Chowla [C] proved in 1931 that $K(\frac{a}{b}\theta) < 2ab(K(\theta) + 1)^3$, a bound much weaker than that obtained above.

Theorem 7.

$$K(\frac{1}{\theta}) \le \begin{cases} K(\theta), & \textit{if } 0 < \theta < 1 \\ \max(K(\theta), \lfloor\theta\rfloor), & \textit{if } \theta > 1 \\ K(\theta)+1, & \textit{if } -1 < \theta < 0 \\ \max(K(\theta)+2, -\lfloor\theta\rfloor - 2), & \textit{if } \theta < -1. \end{cases}$$

Proof.

As in [K, Ex. 4.5.3.10], we see that

(a) If $0 < \theta < 1$, then $\theta = [0, a_1, a_2, \ldots]$ and $1/\theta = [a_1, a_2, \ldots]$.

(b) If $\theta > 1$, then $\theta = [a_0, a_1, \ldots]$ and $1/\theta = [0, a_0, a_1, \ldots]$.

(c) If $-1/2 < \theta < 0$, then $\theta = [-1, 1, a_2, a_3, a_4]$ and $1/\theta = [-(a_2+2), 1, a_3 - 1, a_4, \ldots]$. (Note: this collapses to $[-(a_2+2), a_4+1, a_5, \ldots]$ if $a_3 = 1$.)

(d) If $-1 < \theta < -1/2$, then $\theta = [-1, a_1, a_2, a_3, \ldots]$, where $a_1 \ge 2$, and $1/\theta = [-2, 1, a_1 - 2, a_2, \ldots]$. (Note: this collapses to $[-2, a_2+1, a_3, \ldots]$ if $a_1 = 2$.)

(e) If $\theta < -1$, then $\theta = [a_0, a_1, a_2, \ldots]$, where $a_0 \le -2$, and $1/\theta = [-1, 1, -(a_0 + 2), 1, a_1 - 1, a_2, \ldots]$. (Note: this collapses to $[-1, 2, a_1 - 1, a_2, \ldots]$ if $a_0 = -2$ and $a_1 \ge$; to $[-1, 1, -(a_0+2), a_2+1, a_3, \ldots]$ if $a_1 = 1$ and $a_0 \le -3$, and to $[-1, a_2+2, a_3, \ldots]$ if $a_0 = -2$ and $a_1 = 1$.) ■

Theorem 8.

Let a, b, c, d be integers with $ad - bc \ne 0$. Then θ has bounded partial quotients iff $\frac{a\theta+b}{c\theta+d}$ has bounded partial quotients.

Proof.

$\Rightarrow$: If $c = 0$, this follows from Corollaries 4 and 6. If $c \ne 0$, then

$$\frac{a\theta+b}{c\theta+d} = \frac{b - \frac{ad}{c}}{c\theta+d} + \frac{a}{c},$$

and the result follows from Corollaries 4 and 6.

$\Leftarrow$: Let $\tau = \frac{a\theta+b}{c\theta+d}$. Then $\theta = \frac{d\tau-b}{-c\tau+a}$, and the result follows from the argument above. ■

I do not know any proof of Theorem 8 in the literature.

V. Another proof using Raney's theorem.

In this section, I show how to obtain Theorem 8 *directly* from the continued fraction expansion of θ, using a theorem of Raney [R]. This idea was suggested to me by J. C. Lagarias. For those who are familiar with formal languages, the proof will recall the proof of the so-called "pumping lemma' for regular sets; see [HU].

Recall the LR-expansion of an irrational number θ. If

$$\theta = [a_0, a_1, a_2, \ldots],$$

then its LR-expansion is

$$R^{a_0} L^{a_1} R^{a_2} L^{a_3} \cdots .$$

The letters L and R are shorthand for the matrices

$$L = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Raney proved that the LR-expansion of $\tau = \frac{a\theta+b}{c\theta+d}$ can be deduced from that of θ with the aid of a finite-state transducer. The transitions of this transducer correspond to certain products of matrices. When we write an expression such as

$$ARL = LR^3 B, \tag{1}$$

we mean that this transducer, in state A, accepts the string RL as input and outputs LR^3, and then changes to state B. The expression (1) can also be viewed simply as an identity on 2×2 matrices; e.g. where

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that all state matrices are invertible.

In what follows, we regard an expression such as LRL both as a string of length 3, and a certain 2×2 matrix representing a product of the matrices L and R. In particular, we use $|W|$ to denote the length, or number of symbols, in the string W.

Second Proof of Theorem 8.

Consider the transducer mapping the LR-expansion of θ to that of τ. Let m denote the maximum number of R's in any string output by a transition. Let the transducer have s states.

Suppose θ has partial quotients bounded by B, and $ad - bc \neq 0$. Assume, contrary to what we want to prove, that $\tau = \frac{a\theta+b}{c\theta+d}$ has unbounded partial quotients. Then its LR-expansion must contain arbitrarily long strings of R's or L's (not necessarily both). Without loss of generality, assume it contains arbitrarily long strings of R's. Thus we may choose a substring in the LR-expansion of τ of at least $B(m+3)(s+1)$ consecutive R's. Partition this string into $s+1$ groups of $B(m+3)$ R's, as follows:

$$\underbrace{\overbrace{RR\cdots R}^{B(m+3)} \; \overbrace{RR\cdots R}^{B(m+3)} \cdots \overbrace{RR\cdots R}^{B(m+3)}}_{s+1 \text{ groups}}$$

Consider the first group, and the corresponding part of the LR-expansion of θ that is transduced to get this string of $B(m+3)$ R's. Partition it into sections according to the states $S_1, S_2, \ldots, S_{k+1}$ encountered in the transduction:

$$\underbrace{RR\cdots R}\ {}^{S_1}\ \underbrace{\overset{W_1}{RR\cdots R}}_{i_1}\ {}^{S_2}\ \underbrace{\overset{W_2}{RR\cdots R}}_{i_2}\ {}^{S_3}\ \cdots\ \underbrace{\overset{W_k}{RR\cdots R}}_{i_k}\ {}^{S_{k+1}}\ \underbrace{RR\cdots R}$$

From the definition of the transducer, we have

$$\begin{aligned} S_1 W_1 &= R^{i_1} S_2 \\ S_2 W_2 &= R^{i_2} S_3 \\ &\vdots \\ S_k W_k &= R^{i_k} S_{k+1}. \end{aligned}$$

I claim that the words W_j cannot consist of all L's or all R's. For if they did, then since

$$|W_1 W_2 \cdots W_k| \geq B(m+1),$$

the LR-expansion of θ would contain at least $B\frac{m+1}{m}$ consecutive L's or R's, a contradiction.

In the same manner, we can list the first state of the transducer encountered in each of the $s+1$ groups. Since the transducer has precisely s distinct states, at least one state (call it B_1) must be repeated. Thus we have

$$\begin{aligned} B_1 X_1 &= R^{j_1} B_2 \\ B_2 X_2 &= R^{j_2} B_3 \\ &\vdots \\ B_{s+1} X_{s+1} &= R^{j_{s+1}} B_1, \end{aligned}$$

and the string $X_1 X_2 \cdots X_{s+1}$ contains at least one L and one R. Thus we see

$$\begin{aligned} B_1 X_1 X_2 \cdots X_{s+1} &= R^{j_1} B_2 X_2 \cdots X_{s+1} \\ &= R^{j_1} R^{j_2} B_3 X_3 \cdots X_{s+1} \\ &= \quad \cdots \quad = R^{j_1} R^{j_2} \cdots R^{j_{s+1}} B_1. \end{aligned}$$

Let $X = X_1 X_2 \cdots X_{s+1}$ and $j = j_1 + j_2 + \cdots + j_{s+1}$. Considered as a string, X contains at least one L and one R. Therefore, considered as a matrix, all of X's entries are ≥ 1. On the other hand,

$$R^j = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}.$$

Letting

$$B_1 = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

we have

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Hence

$$\begin{aligned} gx_{11} + hx_{21} &= g \\ hx_{12} + hx_{22} &= h. \end{aligned}$$

Since $x_{11}, x_{12}, x_{21}, x_{22} \geq 1$, we have $g = h = 0$. But then $\det(B_1) = 0$, which contradicts the fact that all the transition matrices are invertible. This contradiction (essentially) completes the proof.

I say "essentially" because there is one small point that remains to be cleared up: Raney's transducer does not work for arbitrary matrices, but only for the so-called "doubly-balanced" ones. As Raney shows, however, the general linear fractional transformation can be mapped into the doubly-balanced case by changing a finite number of terms at the beginning of the LR-expansion for θ. Clearly this does not change the (supposed) unboundedness of the partial quotients for τ. *Now* we are really done with the proof. ■

References

[B]. A. Baker, *A Concise Introduction to the Theory of Numbers*, Cambridge University Press, 1984.

[C]. S. D. Chowla, *Some problems of diophantine approximation (I)*, Math. Zeitschrift **33** (1931), 544–563.

[CMF]. T. W. Cusick and M. Mendès France, *The Lagrange spectrum of a set*, Acta Arith. **34** (1979), 287–293.

[HW]. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 1985.

[HU]. J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, 1979.

[K]. D. E. Knuth, *The Art of Computer Programming, V. II (Seminumerical Algorithms)*, Addison-Wesley, 1981.

[P]. O. Perron, *Über die Approximation irrationaler Zahlen durch rationale*, Sitz. Heidelberg. Akad. Wiss. **XII A** *(4. Abhundlung)* (1921), 3–17.

[R]. G. N. Raney, *On continued fractions and finite automata*, Math. Annalen **206** (1973), 265-283.

[S]. W. Schmidt, *Diophantine Approximation*, Lecture Notes in Mathematics, vol. 785, Springer-Verlag, 1980.

J. O. Shallit
Department of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

E-mail: shallit@graceland.waterloo.edu