CONTINUED FRACTIONS WITH PARTIAL QUOTIENTS BOUNDED IN AVERAGE

Joshua N. Cooper

Department of Mathematics, University of South Carolina,
LeConte College, 1523 Greene Street, Columbia, SC 29208, USA
e-mail: cooper@math.sc.edu

(Submitted September 2004-Final Revision August 2006)

ABSTRACT

We ask, for which n does there exists a k, $1 \leq k < n$ and $(k, n) = 1$, so that k/n has a continued fraction whose partial quotients are bounded in average by a constant B? This question is intimately connected with several other well-known problems, and we provide a lower bound in the case of $B = 2$. The proof, which is completely elementary, involves a simple "shifting" argument, the Catalan numbers, and the solution to a linear recurrence.

1. INTRODUCTION

An important question in the theory of quasirandomness, uniform distribution of points, and diophantine approximation is the following: For which $n \in \mathbb{Z}$ is it true that there exists an integer k, $1 \leq k < n$ and $(k, n) = 1$, so that k/n has a continued fraction whose partial quotients are bounded in average by a constant B? That is, if we write $k/n = [0; a_1, a_2, \ldots, a_m]$, we wish to find k so that

$$t^{-1} \sum_{i=1}^{t} a_i \leq B$$

for all t with $1 \leq t \leq m$. Denote by $\mathcal{F}(B)$ the set of all n for which such a k exists. These sets are discussed at length in [2] and the related matter of partial quotients bounded *uniformly* by a constant appears as an integral part of [6]. This latter question is closely connected with Zaremba's Conjecture ([8]), which states that such a k exists for all $n > 1$ if we take $B = 5$.

Define the *continuant* $K(a_1, a_2, \ldots, a_m)$ to be the denominator of the continued fraction $k/n = [0; a_1, a_2, \ldots, a_m]$. In [3], it is proven that, if $S_n(B)$ is the number of sequences $\mathbf{a}= (a_1, \ldots, a_m)$ bounded uniformly by B with $K(\mathbf{a}) \leq n$ and $H(B)$ is the Hausdorff dimension of the set of continued fractions with partial quotients bounded uniformly by B, then

$$\lim_{n\to\infty} \frac{\log(S_n(B))}{\log n} = 2H(B).$$

Then, in [4], $H(2)$ is calculated with a great deal of accuracy: $H(2) \approx 0.53128$. Therefore, $S_n(2)$, and thus the number of p/q with $q \leq n$ whose partial quotients are bounded by 2, is $n^{1.0625\ldots+o(1)}$. (This improves the previous best known lower bound, $n^{\approx 1.017}$ computed in [3], slightly.)

Define $\bar{S}_n(B)$ to be the number of sequences $\mathbf{a}= (a_1, \ldots, a_m)$ with partial quotients bounded *in average* by B so that $K(\mathbf{a}) \leq n$. Clearly, $\bar{S}_n(B) \geq S_n(B)$, so $\bar{S}_n(2) \gg n^{1.0625}$. In the next section, we prove something much stronger, however – an exponent of ≈ 1.5728394 –

thus providing a lower bound in the first nontrivial case. Section 3 discusses the implications for the density of $\mathcal{F}(2)$ and a few open problems.

2. THE PROOF

Theorem 1: *For any* $\epsilon > 0$, $\bar{S}_n(2) \gg n^{2\log 2/\log(1+\sqrt{2})-\epsilon}$.

Proof: The proof consists of two parts: computing the number of positive sequences of length m bounded in average by 2, and then computing the smallest possible m so that $K(a_1,\dots,a_m) > n$ and the a_i are bounded in average by 2.

First, we wish to know how many sequences $(a_1,\dots,a_m)$ there are with $a_j \geq 1$ for each $j \in [m]$ and $\sum_{j=1}^r a_j \leq 2r$ for each $r \in [m]$. Call this number $T(m)$. By writing $b_j = a_j - 1$, we could equivalently ask for sequences $(b_1,\dots,b_m)$ with $b_j \geq 0$ for each $j \in [m]$ and $\sum_{j=1}^r b_j \leq r$ for each $r \in [m]$. This is precisely the number of lattice paths from $(0,0)$ to (m,m) which do not cross the line $y = x$, and so $T(m)$ is the m^{th} Catalan number, or $(m+1)^{-1}\binom{2m}{m} = 4^{m(1-o(1))}$.

In the following lemmas, we show that $K(a_1,\dots,a_m) \leq n$ if $m \leq \log n(1-o(1))/\log(1+\sqrt{2})$. Therefore, setting m as large as possible, we have at least

$$4^{\log n(1-o(1))/\log(1+\sqrt{2})} = n^{2\log 2/\log(1+\sqrt{2})-o(1)}$$

sequences with partial quotients bounded in average by 2 and continuant $\leq n$. □

We must show that the size of a continuant with partial quotients bounded in average by B is at most the largest size of a continuant with partial quotients bounded by B.

Lemma 2: *If the sequence* $(a_1,\dots,a_m)$ *of positive integers is bounded in average by* $B > 1$, *then* $K(a_1,\dots,a_m) \leq K(\underbrace{B,\dots,B}_{m})$.

Proof: We prove the Lemma by a "shifting" argument. That is, we perform induction on the size of the entry a_j such that $a_j > B$ and j is as small as possible. If $\mathbf{a}= (a_1,\dots,a_m)$ contains no $a_t > B$, we are done, because increasing the partial quotients can only increase the continuant. If there is some $a_t > B$, let $t \geq 2$ be the smallest such index. We consider two cases: (i) $a_t \geq B+2$ or $a_{t-1} < B$, and (ii) $a_t = B+1$, $a_k = B$ for $s \leq k \leq t-1$ for some $2 \leq s \leq t-1$, and $a_{s-1} < B$. (Clearly, $\mathbf{a} \neq (B,B,\dots,B,B+1,a_{t+1},\dots,a_m)$, since this sequence is not bounded in average by B. Therefore we may assume $s \geq 2$.)

Case (i):

Let $\mathbf{b} = (b_1,\dots,b_m) = (a_1,\dots,a_{t-1}+1,a_t-1,\dots,a_m)$. We show that $K(\mathbf{b}) > K(\mathbf{a})$. First, note that

$$\sum_{j=1}^r b_j = \begin{cases} \sum_{j=1}^r a_j & \text{if } r \neq t-1 \\ 1+\sum_{j=1}^{t-1} a_j & \text{if } r = t-1. \end{cases}.$$

Since $a_t \geq B+1$, $\sum_{j=1}^{t-1} a_j \leq tB-B-1$, so $1+\sum_{j=1}^{t-1} a_j \leq (t-1)B$, and $\mathbf{b}$ is bounded in average by B. Second, note that it suffices to consider the case of $t = m$, since, if $K(b_1,\dots,b_j) > K(a_1,\dots,a_j)$ for $1 \leq j \leq t$, then $K(\mathbf{b}) > K(\mathbf{a})$. (That is, $K(\cdot)$ is monotone increasing.)

Let $q_j = K(a_1, \dots, a_j)$ and $q'_j = K(b_1, \dots, b_j)$. (We use the convention that $q_j = 0$ when $j < 0$ and $q_0 = 1$.) Clearly, $q_j = q'_j$ if $j < t-1$. When $j = t-1$, we have $q'_{t-1} > q_{t-1}$ by monotonicity. When $j = t$,

$$q_t = a_t q_{t-1} + q_{t-2} = a_t(a_{t-1}q_{t-2} + q_{t-3}) + q_{t-2} = (a_t a_{t-1} + 1)q_{t-2} + a_t q_{t-3},$$

and

$$\begin{aligned} q'_t &= (b_t b_{t-1} + 1)q'_{t-2} + b_t q'_{t-3} \\ &= ((a_t - 1)(a_{t-1} + 1) + 1)q_{t-2} + (a_t - 1)q_{t-3} \\ &= q_t + q_{t-2}(a_t - a_{t-1} - 1) - q_{t-3}. \end{aligned}$$

Since $a_t \geq a_{t-1} + 2$ and $q_{t-2} > q_{t-3}$, we have

$$q'_t \geq q_t + q_{t-2} - q_{t-3} > q_t.$$

Case (ii):

Now, assume that $a_t = B + 1$, $a_k = B$ for $s \leq k \leq t-1$ for some $2 \leq s \leq t-1$, and $a_{s-1} < B$. Then define $\mathbf{b} = (b_1, \dots, b_m)$ by letting $b_j = a_j$ if $j \neq s-1$ and $j \neq t$; $b_{s-1} = a_{s-1} + 1$; and $b_t = a_t - 1$. Again, we may assume that $t = m$. Then

$$\sum_{j=1}^{r} b_j = \begin{cases} \sum_{j=1}^{r} a_j & \text{if } r = t \text{ or } r < s-1 \\ 1 + \sum_{j=1}^{r} a_j & \text{if } s-1 \leq r \leq t-1. \end{cases}$$

For any r such that $s-1 \leq r \leq t-1$,

$$\sum_{j=1}^{r} a_j = \sum_{j=1}^{t} a_j - \sum_{j=r+1}^{t} a_j \leq Bt - (B(t-r-1) + (B+1)) \leq Br - 1.$$

Therefore, $\sum_{j=1}^{r} b_j \leq Br$ for all $r \in [t]$, and we may conclude that $\mathbf{b}$ is bounded in average by B.

Define F_k as follows: $F_0 = 0$, $F_1 = 1$, and, for $k > 1$, $F_k = BF_{k-1} + F_{k-2}$. Then it is easy to see by induction that

$$K(\underbrace{B, \dots, B}_{k}, x) = F_{k+1}x + F_k.$$

Also,

$$K(y, c_1, \dots, c_r) = yK(c_1, \dots, c_r) + K(c_2, \dots, c_r). \tag{1}$$

Taking $k = t - s$, we deduce

$$K(a_{s-1}, \dots, a_t) = a_{s-1}((B+1)F_{k+1} + F_k) + (B+1)F_k + F_{k-1},$$

and

$$\begin{aligned} K(b_{s-1}, \dots, b_t) &= (a_{s-1} + 1)(BF_{k+1} + F_k) + BF_k + F_{k-1} \\ &= K(a_{s-1}, \dots, a_t) + (B - a_{s-1})F_{k+1} \\ &\geq K(a_{s-1}, \dots, a_t) + F_{k+1}. \end{aligned}$$

If $s = 2$, we are done. Otherwise, we use that

$$\begin{aligned} K(b_{s-2}, \dots, b_t) &= a_{s-2}K(b_{s-1}, \dots, b_t) + K(b_s, \dots, b_t) \\ &\geq a_{s-2}K(a_{s-1}, \dots, a_t) + F_{k+1} + b_tK(b_s, \dots, b_{t-1}) + K(b_s, \dots, b_{t-2}) \\ &= a_{s-2}K(a_{s-1}, \dots, a_t) + F_{k+1} + K(a_s, \dots, a_t) - K(a_s, \dots, a_{t-1}) \\ &= K(a_{s-2}, \dots, a_t). \end{aligned}$$

Now, inductive application of (1) to the continuants $K(b_{s-j}, \dots, b_t)$, $3 \leq j \leq s-1$, yields $K(\mathbf{b}) \geq K(\mathbf{a})$, since $a_{s-j} = b_{s-j}$ in this range.

By repeating cases (i) and (ii) as appropriate, we will eventually reach a sequence of partial quotients bounded by B, and at each stage we never decrease the corresponding continuant. The result therefore follows. $\square$

It remains to find a bound on $K(B, \dots, B)$.

Lemma 3: *If* $B \geq 1$, $K(\underbrace{B, \dots, B}_{m}) \leq \left(\frac{1}{2}(B + \sqrt{B^2+4})\right)^{m+1}$.

Proof: We proceed by induction. The case $m = 0$ is trivial. Suppose it is true for all $m < M$. Then, by (1),

$$\begin{aligned} K(\underbrace{B, \dots, B}_{M}) &= BK(\underbrace{B, \dots, B}_{M-1}) + K(\underbrace{B, \dots, B}_{M-2}) \\ &\leq B\left(\frac{1}{2}(B + \sqrt{B^2+4})\right)^{M} + \left(\frac{1}{2}(B + \sqrt{B^2+4})\right)^{M-1} \\ &\leq \left(\frac{1}{2}(B + \sqrt{B^2+4})\right)^{M-1}\left(\frac{1}{2}B^2 + \frac{1}{2}B\sqrt{B^2+4} + 1\right) \\ &= \left(\frac{1}{2}(B + \sqrt{B^2+4})\right)^{M+1}. \quad \square \end{aligned}$$

3. THE DENSITY OF $\mathcal{F}(2)$

Corollary 4: *There is a constant* C *and a subset* S *of the positive integers such that* $\log|S \cap [n]|/\log n \geq \log 2/\log(1+\sqrt{2}) - o(1) \approx 0.786$ *so that, for each* $n \in S$*, there exists a* $k \in [n]$, $(k, n) = 1$ *so that* k/n *has partial quotients bounded in average by* 2.

Proof: Let U be the set of all reduced fractions p/q, $1 < p < q$, whose partial quotients $\mathbf{a}= (a_1, a_2, \dots, a_m)$ are bounded in average by 2 and such that $\mathbf{a}' = (a_2, \dots, a_m)$ is bounded in average by 2. The number of such $\mathbf{a}$ with $K(\mathbf{a}) \leq n$ is at least twice the number of sequences $\mathbf{a}' = (a_2, \dots, a_m)$ bounded in average by 2 with $K(\mathbf{a}') \leq n/3$, because, if $[\mathbf{a}'] = p/q$, then $K(\mathbf{a}) = a_1q+p \leq 3K(\mathbf{a}') \leq n$. (The fact that $\mathbf{a}'$ is bounded in average by 2 implies that $[1, \mathbf{a}']$

and $[2, \mathbf{a}']$ are also.) Then, since every rational has at most two representations as a continued fraction, the number of elements of U whose denominator is $\leq n$ is at least $\bar{S}_{n/3}(2)$, which is at least $n^{2\log 2/\log(1+\sqrt{2})-o(1)}$. Let S be the set of denominators of fractions appearing in U. If $p/q = [\mathbf{a}]$ is in U, then $[\mathbf{a}'] = (q - a_1 p)/p$, so p is the continuant of a sequence whose partial quotients are bounded in average by 2. Therefore, $\bar{S}_{n/3}(2) \leq |S \cap [n]|^2$, and we may conclude that $\log|S \cap [n]|/\log n \geq \log 2/\log(1+\sqrt{2}) - o(1)$. $\square$

Attempts by the author to find a generalization of the above result to $\mathcal{F}(B)$ by applying much more careful counting arguments when $B > 2$ have failed thus far. It would also be interesting to (i), calculate the Hausdorff dimension of the set of reals in $[0, 1)$ whose partial quotients are bounded in average by B, and (ii), draw a connection, similar to that of the "uniform" case, between this quantity and the asymptotic density of $\mathcal{F}(B)$.

REFERENCES

[1] J. N. Cooper. "Quasirandom Permutations." *J. Combin. Theory Ser. A* **106.1** (2004): 123-143.

[2] J. N. Cooper. *Survey of Quasirandomness in Number Theory*, 2003, preprint.

[3] T. W. Cusick. "Continuants With Bounded Digits, III." *Monatsh. Math.* **99.2** (1985): 105-109.

[4] D. Hensley. "A Polynomial Time Algorithm for the Hausdorff Dimension of Continued Fraction Cantor Sets." *J. Number Theory* **58.1** (1996): 9-45.

[5] G. Larcher. "On the Distribution of Sequences Connected with Good Lattice Points." *Monatsh. Math.* **101** (1986): 135-150.

[6] H. Niederreiter. "Quasi-Monte Carlo Methods and Pseudo-random Numbers." *Bull. Amer. Math. Soc.* **84** (1978): 957-1041.

[7] G. Ramharter. "Some Metrical Properties of Continued Fractions." *Mathematika* **30.1** (1983): 117-132.

[8] S. K. Zaremba, ed. "Applications of Number Theory to Numerical Analysis." *Proceedings of the Symposium at the Centre for Research in Mathematics*, University of Montréal, Academic Press, New York-London, (1972).

AMS Classification Numbers: 11K50; 11K38

✠ ✠ ✠