JOURNAL OF NUMBER THEORY 35, 150–156 (1990)

Length and Denominators of Egyptian Fractions, III

GÉRALD TENENBAUM

Département de Mathématiques, Université de Nancy I, BP 239, 54506 Vandœuvre, France

AND

HISASHI YOKOTA

Department of Mathematics, Hiroshima Institute of Technology, Itsukaichi, Hiroshima, Japan

Communicated by H. Zassenhaus

Received October 31, 1988; revised July 2, 1989

We show that for large N every rational number $a/N \in]0, 1[$ has an egyptian fraction expansion

$$\frac{a}{N} = \sum_{j=1}^{r} \frac{1}{n_j},$$

where $r \leqslant (1 + o(1)) \log N/\log_2 N$ and $n_r \leqslant 4N \log^2 N \log_2 N$. This is essentially best possible.

1. INTRODUCTION

Let a, N be positive integers such that $a < N$. We call an "egyptian fraction expansion" of a/N an equality of the type

$$\frac{a}{N} = \sum_{j=1}^{r} \frac{1}{n_j}, \tag{1}$$

where $0 < n_1 < n_2 < \cdots < n_r$. Ideally, one would like an algorithm which yields a short expansion where the denominators stay small. A quantitative measure of the quality, in the above sense, of a given expansion, is given by comparison of r and n_r to

$$L(N) := \max_{1 \leqslant a < N} \min_{(1)} r, \qquad D(N) := \max_{1 \leqslant a < N} \min_{(1)} n_r,$$

0022-314X/90 $3.00

where the minima range over all expansions of type (1). It is now known (cf. [2, 5]) that

$$\log_2 N \ll L(N) \ll \sqrt{\log N} \tag{2}$$

holds uniformly for $N \geqslant 3$ and (see [1, 9]) that

$$P \frac{\log P \log_2 P}{\prod_{j=4}^{k} \log_j P \cdot (\log_{k+1} P)^2} \leqslant D(P) \leqslant P \log P (\log_2 P)^5 \tag{3}$$

holds, for any given $k \geqslant 4$, and all large primes P. Here and in the sequel, we let $\log_j$ denote the jth fold iterated logarithm. The upper bound of (3) also provides upper bounds for $D(N)$ since this function of N is submultiplicative—see [6].

Besides the question of sharpening (2) and (3) one of the remaining problems in this area is to construct an algorithm which yields good bounds for $L(N)$ and $D(N)$ simultaneously. In [7], the second author showed that, for a large prime P, there is no algorithm which yields both

$$L(P) \leqslant \frac{c \log P}{\log_2 P} \quad \text{and} \quad D(P) \leqslant P(\log P)^{1+1/c-\varepsilon}. \tag{4}$$

But in [8], he developed an algorithm that gives

$$L(N) \leqslant \frac{\lambda \log N}{\log_2 N} \quad \text{and} \quad D(N) \leqslant N(\log N)^{2+\delta}, \tag{5}$$

where λ and δ are functions of N such that $\lambda \to 2$ and $\delta \to 0$ as $N \to \infty$. This is close to (4) with $c = 1$ but the upper bound for $L(N)$ is about twice as large as predicted by (4).

Thus the optimal bounds corresponding to $c = 1$ will follow if one can obtain $\lambda = \lambda_N \to 1$. As noticed in [8], this problem can be reduced to that of finding an increasing sequence $\{s_j : j = 1, 2, ...\}$ such that, for large k, any integer s in the range $1 \leqslant s < S_k := \prod_{j=1}^{k} s_j$ can be written as a sum of $o(k)$ distinct divisors of S_k. The case $s_j = j$ was raised by Erdős in [3]. In this article, we answer positively this question (including Erdős' example—cf. Lemma 4) and derive the following result.

THEOREM. *Let $\varepsilon > 0$. There exists a real number $N_0 = N_0(\varepsilon)$ such that for all a, N, $N > N_0$, $1 \leqslant a < N$, a/N has an egyptian fraction expansion* (1) *with*

$$r \leqslant (1+\varepsilon) \frac{\log N}{\log_2 N}, \qquad n_r \leqslant 4N(\log N)^2 \log_2 N. \tag{6}$$

As we remarked earlier, this is best possible in the sense that one cannot replace ε by $-\varepsilon$ in the upper bound for r.

2. LEMMATA

The letter p, with or without subscript, denotes exclusively a prime number. The jth prime is written as p_j. $S := \{\sigma_j : j \geqslant 1\}$ is the increasing sequence of all positive integers of the form p^{2^i}, $i \geqslant 0$.

LEMMA 1. *For large k and all N in the range $\prod_{j=1}^{k-1} p_j \leqslant N \leqslant \prod_{j=1}^{k} p_j$, we have*

$$p_k \leqslant \log N \left(1 + \frac{1}{\log_2 N}\right), \qquad k \leqslant \frac{\log N}{\log_2 N}\left(1 + \frac{2}{\log_2 N}\right).$$

This is a straightforward consequence of the prime number theorem.

LEMMA 2. *There exists a constant n_0 such that, for all $n \geqslant n_0$ and all r in the range $(1 - 2/\sqrt{\sigma_n}) \prod_{j=1}^{n} \sigma_j \leqslant r < 2 \prod_{j=1}^{n} \sigma_j$, there are distinct integers d_i $(1 \leqslant i \leqslant m)$ with the property that $d_i \mid \prod_{j=1}^{n} \sigma_j$, $d_i \geqslant \prod_{j=1}^{n} \sigma_j / 2\sigma_n^2 \log \sigma_n$ for all i, and $r = \sum_{i=1}^{m} d_i$, where $m \leqslant \pi(\sigma_n) + 2\pi(\sqrt{\sigma_n}) \log \sigma_n$.*

This is Lemma 2.7 of [8].

In the next two lemmata, we use N_k to denote either $k!$ or $\prod_{j=1}^{k} p_j$. The same reasoning would in fact apply to many other similar sequences.

LEMMA 3. *Let $0 < \varepsilon < \frac{1}{2}$, and put $\varepsilon_k := \exp\{-(\log k)^{3/2 - \varepsilon}\}$ $(k \geqslant 1)$. There exists a constant $k_0 = k_0(\varepsilon)$ such that, for $k \geqslant k_0$ and $\sqrt{N_k} \leqslant z \leqslant \sqrt{N_{k+1}}$, every interval $[z, z(1 + \varepsilon_k)]$ contains at least one divisor of N_k.*

This is an immediate consequence of Theorem 3 of [4].

LEMMA 4. *With the above notation, there exists a constant $k_1 = k_1(\varepsilon)$ such that, for $k \geqslant k_1$, every integer s, $1 \leqslant s \leqslant N_k$, may be written as a sum of at most $k/(\log k)^{1/2 - \varepsilon}$ distinct divisors of N_k.*

Proof. We are going to construct a strictly decreasing sequence $d_1 > d_2 > \cdots > d_m$ of divisors of N_k such that $s = \sum_{i=1}^{m} d_i$. We put $s_0 = s$, $s_j = s - \sum_{i=1}^{j} d_i$ $(j \geqslant 1)$. Let $Z = \sqrt{N_{k_0+1}}$. If $s > N_k/Z$ we use slightly different arguments to obtain successively, for suitable integers $0 < b < h < l < m$, $s_b \leqslant N_k/Z$, $s_h \leqslant \sqrt{N_k}$, $s_l \leqslant Z$, and finally $s_m = 0$. Thus, there is no loss of generality to suppose that $s > N_k/Z$: if this is not so, we simply skip the steps which bring s_j in its actual range.

Hence we start with $s_0 > N_k/Z$. We may plainly suppose that k_1 is so large that for suitable integers $u + 1, ..., u + b \leqslant k_1$, we have

$$p_{u+b+1} > Z \qquad \text{and} \qquad \sum_{j=1}^{b} \frac{1}{p_{u+j}} \leqslant \frac{s}{N_k} < \sum_{j=1}^{b+1} \frac{1}{p_{u+j}}.$$

We then set $d_j = N_k/p_{u+j}$ $(1 \leqslant j \leqslant b)$ and this realizes our first step. We note that $b \ll_\varepsilon 1$.

Put $\beta = \frac{1}{2}(3-\varepsilon)$. We shall prove that one can obtain $s_h \leqslant \sqrt{N_k}$ with $h \ll k(\log k)^{1-\beta}$. Let m_1 be the unique integer such that $\sqrt{N_{m_1}} \leqslant N_k/s_b < \sqrt{N_{m_1+1}}$. Then $k_0(\varepsilon) \leqslant m_1 \leqslant k$, and we deduce from Lemma 3 that there exists a t_1, $t_1 | N_{m_1} | N_k$, such that

$$(1-\alpha_1)t_1 \leqslant N_k/s_b \leqslant t_1$$

with $\alpha_1 := \varepsilon_{m_1}$. We set $d_{b+1} = N_k/t_1$, and obtain

$$0 \leqslant s_{b+1} \leqslant \alpha_1 s_b.$$

We check that $d_{b+1} \leqslant s_b < N_k/p_{u+b+1} < d_b$. If $s_{b+1} \leqslant \sqrt{N_k}$, we put $h = b+1$, otherwise we repeat the application of Lemma 3 and produce a $t_2 \mid N_k$ such that, with $d_{b+2} := N_k/t_2$,

$$0 \leqslant s_{b+2} = s_{b+1} - d_{b+2} \leqslant \alpha_2 s_{b+1} \leqslant \alpha_1 \alpha_2 s_b,$$

with $\alpha_2 = \varepsilon_{m_2}$, for some m_2 such that $\sqrt{N_{m_2+1}} > N_k/s_{b+1} > 1/\alpha_1$. Since $\log_2 \sqrt{N_{m_2+1}} \leqslant 2 \log m_2$ for sufficiently large k_0, we infer that

$$\alpha_2 \leqslant \exp\left\{-\left(\frac{1}{2}\log_2\left(\frac{1}{\alpha_1}\right)\right)^{\beta}\right\}.$$

Moreover we still have $d_{b+2} < d_{b+1}$ since $s_{b+1} - d_{b+1} \leqslant -(1-2\alpha_1)\, s_b < 0 \leqslant s_{b+1} - d_{b+2}$. Iterating the procedure, we eventually obtain $s_h \leqslant \sqrt{N_k}$. We estimate h by using the inequality

$$s_j \leqslant \gamma_j s_b \qquad (b < j \leqslant h),$$

where $\gamma_j = \alpha_1 \cdots \alpha_j$ satisfies

$$\gamma_{j+1} \leqslant \gamma_j \exp\left\{-\left(\frac{1}{2}\log_2\left(\frac{1}{\gamma_j}\right)\right)^{\beta}\right\}.$$

A simple computation yields $\log(1/\gamma_j) \gg j(\log j)^{\beta}$ and since $s_b \leqslant N_k \leqslant k^k$, we readily obtain the required estimate for h.

We now show similarly that $s_l \leqslant Z$ holds for some $l \ll k(\log k)^{1-\beta}$. Let q_1 be defined by $\sqrt{N_{q_1}} \leqslant s_h \leqslant \sqrt{N_{q_1+1}}$. Plainly $q_1 \leqslant k$. If $q_1 \leqslant k_0$, there is nothing to do; otherwise we apply Lemma 3 to the effect that

$$0 \leqslant s_{h+1} = s_h - d_{h+1} \leqslant \varepsilon_{q_1} s_h$$

for some divisor d_{h+1} of $N_{q_1} \mid N_k$. As before, we have $d_{h+1} < d_h$ because $s_h - d_h < 0$. Moreover, since $\log_2 N_{q_1} \leqslant 2 \log q_1$ for suitable k_0, we get

$$\log s_{h+1} \leqslant \log s_h - \left(\frac{1}{2} \log_2 s_h\right)^{\beta}.$$

Iterating, we obtain $s_l \leqslant Z$ in $l - h \ll k(\log k)^{1-\beta}$ steps, as required.

It remains to show that s_l is representable as a sum of distinct divisors of N_k, all less than d_l. For this purpose we use the well known fact (see, e.g., [4, p. 7]), that the ratio of two consecutive divisors of N_k does not exceed 2. (This may be easily shown by induction using $p_{j+1} \leqslant 2p_1 \cdots p_j$. We omit the details.) Let d_{l+1} be the largest divisor of N_k not exceeding s_l. Then $s_{l+1} = s_l - d_{l+1} \leqslant d_{l+1}$ and equality is only possible if s_l is itself a divisor of N_k—in which case the desired representation is trivially found. If $s_l < d_{l+1}$, we iterate and obviously use only divisors less than d_{l+1}. In a finite number of steps we must obtain $s_m = 0$. This completes the proof of Lemma 4.

3. Proof of the Theorem

In this section we put $N_k := \prod_{j=1}^{k} p_j$. Let k be such that $N_{k-1} \leqslant N < N_k$. If $N \mid N_k$, then

$$\frac{a}{N} = \frac{b}{N_k},$$

where $b = aN_k/N < N_k$. By Lemma 4, we have a representation

$$b = \sum_{i=1}^{m} d_i,$$

where $d_i \mid N_k$, $m \leqslant k/(\log k)^{1/2-\varepsilon}$. Thus, the expansion $a/N = \sum_{i=1}^{m} d_i/N_k$ is of type (1) with

$$r = m \ll \log N/(\log_2 N)^{3/2-\varepsilon}$$

and

$$n_r \leqslant N_k \ll N \log N,$$

by Lemma 1. This plainly sufficient.

We may now restrict ourselves to the case when $N \nmid N_k$. We also suppose that $a \geqslant 2$, since $a=1$ is trivial. From the euclidian algorithm, we obtain $(a-1) N_k = Ns + t$, $0 \leqslant t < N$, whence

$$aN_k = Ns + r, \qquad N_k \leqslant r < 2N_k. \tag{7}$$

Moreover $0 \leqslant s < N_k$, and Lemma 4 implies that there are divisors d_i $(1 \leqslant i \leqslant m_1)$ of N_k such that

$$\frac{s}{N_k} = \sum_{i=1}^{m_1} \frac{d_i}{N_k} = \sum_{i=1}^{m_1} \frac{1}{e_i} \qquad \text{(say)} \tag{8}$$

with $m_1 \leqslant k/(\log k)^{1/2-\varepsilon}$.

Let n be such that $\sigma_n = p_k$. Then $N_k \mid \prod_{j=1}^{n} \sigma_j$, and we may write r/N_k as $r^*/\prod_{j=1}^{n} \sigma_j$. From (7), it follows that $\prod_{j=1}^{n} \sigma_j \leqslant r^* < 2\prod_{j=1}^{n} \sigma_j$, and we apply Lemma 2 to obtain

$$\frac{r^*}{\prod_{j=1}^{n} \sigma_j} = \sum_{i=1}^{m_2} \frac{d_i^*}{\prod_{j=1}^{n} \sigma_j} = \sum_{i=1}^{m_1} \frac{1}{e_i^*} \qquad \text{(say)},$$

where $e_i^* \leqslant 2\sigma_n^2 \log \sigma_n \leqslant 3(\log N_k)^2 \log_2 N_k$ and $m_2 \leqslant \pi(\sigma_n) + 2\pi(\sqrt{\sigma_n}) \log \sigma_n$ $\leqslant (1+\varepsilon/2)k$. Thus we have finally shown that

$$\frac{a}{N} = \frac{aN_k}{NN_k} = \frac{s}{N_k} + \frac{r}{NN_k} = \sum_{i=1}^{m_1} \frac{1}{e_i} + \sum_{i=1}^{m_2} \frac{1}{Ne_i^*},$$

where the e_i and Ne_j^* are distinct. Indeed $N \nmid e_i$ since $N \nmid N_k$ and $e_i \mid N_k$. This is an expansion of type (1) with

$$r \leqslant m_1 + m_2 \leqslant k \left\{ 1 + (\log k)^{-1/2+\varepsilon} + \frac{1}{2}\varepsilon \right\} \leqslant (1+\varepsilon) \frac{\log N}{\log_2 N}$$

and

$$n_r \leqslant N_k + 3N(\log N_k)^2 \log_2 N_k \leqslant 4N(\log N)^2 \log_2 N.$$

This completes the proof of our theorem.

REFERENCES

1. M. N. BLEICHER AND P. ERDŐS, Denominators of Egyptian fractions, *J. Number Theory* **8** (1976), 157–168.
2. P. ERDŐS, The solution in whole number of the equation: $1/x_1 + 1/x_2 + \cdots + 1/x_n = a/b$, *Mat. Lapok* **1** (1950), 192–210.

3. P. ERDŐS AND R. L. GRAHAM, "Old and New Problems in Combinatorial Number Theory," L'enseignement Mathématique, Université de Genève, Imprimerie Kundig, Genève, 1980.
4. G. TENENBAUM, Sur un problème extrémal en arithmétique, *Ann. Inst. Fourier* **37**, No. 2 (1987), 1–18.
5. M. D. VOSE, Egyptian fractions, *Bull. London Math. Soc.* **17** (1985), 21–24.
6. H. YOKOTA, On a conjecture of M. N. Bleicher and P. Erdős, *J. Number Theory* **24** (1986), 89–94.
7. H. YOKOTA, Lentgh and denominators of Egyptian fractions, *J. Number Theory* **24** (1986), 249–258.
8. H. YOKOTA, Lentgh and denominators of Egyptian fractions, II, *J. Number Theory* **28** (1988), 272–282.
9. H. YOKOTA, On a problem of Bleicher and Erdős, *J. Number Theory* **30** (1988), 198–207.