PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 117, Number 4, April 1993

UPPER BOUND OF $\sum 1/(a_i \log a_i)$ FOR PRIMITIVE SEQUENCES

PAUL ERDÖS AND ZHENXIANG ZHANG

(Communicated by William W. Adams)

ABSTRACT. A sequence $A = \{a_i\}$ of positive integers $a_1 < a_2 < \cdots$ is said to be primitive if no term of A divides any other. The senior author conjectures that, for any primitive sequence A,

$$\sum_{a \le n,\, a \in A} \frac{1}{a \log a} \le \sum_{p \le n} \frac{1}{p \log p} \quad \text{for } n > 1,$$

where p is a variable prime. In our two previous papers we partially proved this conjecture. The main result of this paper is: for any primitive sequence A,

$$\sum_{a \in A} \frac{1}{a \log a} < 1.84.$$

We also give a necessary and sufficient condition for this conjecture, i.e.,

$$\sum_{b \in B} \frac{1}{b \log b} \le \sum \frac{1}{p \log p}$$

for any primitive sequence B.

1. INTRODUCTION

A sequence $A = \{a_i\}$ of positive integers $a_1 < a_2 < \cdots$ is said to be primitive if no term of A divides another [3, 4]. We denote by p_m the mth prime and by p a variable prime. We define the degree of an integer a, denoted by $\Omega(a)$, to be the number of prime factors of a, counted with multiplicity. The degree of a primitive sequence A, denoted by $d^\circ(A)$, is defined as the maximum degree of its terms. We take $d^\circ(A) = 0$ if $A = \{1\}$ or $\varnothing$.

We define $f(A) = \sum_{a \in A} 1/(a \log a)$. We take $f(A) = 0$ if $d^\circ(A) = 0$. Erdös [1] (cf. [4]) proved that there exists an absolute constant C such that $f(A) < C$ for any primitive sequence, using the inequality

$$\sum_{a \in A} \frac{1}{a} \prod_{p \le P(a)} \left(1 - \frac{1}{p}\right) \le 1, \tag{1}$$

where $P(a)$ is the largest prime factor of a. It is significant to get a reasonable upper bound for $f(A)$. Using (1) we first got $f(A) < 2.886$; then Robin

Received by the editors January 16, 1991 and, in revised form, July 9, 1991.
1991 *Mathematics Subject Classification*. Primary 11B83.
Key words and phrases. Primitive sequences.

improved it to 2.77 (neither is published). It seems to us that we could not do better by only using (1).

Erdös [2] has asked if the inequality

$$\sum_{a\le n,\, a\in A} \frac{1}{a\log a} \le \sum_{p\le n} \frac{1}{p\log p} \quad \text{for } n>1 \tag{2}$$

is always true for any A. Zhang [7] proved that if A is primitive with $d^\circ(a)\le 4$ then (2) is true. For a given primitive sequence A and $m\ge 1$, let $A_m = \{a : a\in A,\ \text{all prime factors of } a \text{ are } \ge p_m\}$ and $A'_m = \{a\colon a\in A_m,\ p_m|a\}$. Zhang [8] proved that (2) is true for any primitive sequence A with the property that each A_m is "homogeneous." That is, for each m, there is some integer s_m such that either $A_m=\varnothing$ or $\Omega(a)=s_m$ for all $a\in A_m$.

In this paper we prove the following two theorems.

Theorem 1. *$f(A)<1.84$ for any primitive sequence $A=\{a_i\}$.*

Theorem 2. *The necessary and sufficient condition for (2) being true for any primitive sequence A is that $f(B)\le \sum 1/(p\log p)$ for any primitive sequence B.*

The basic idea for proving Theorem 1 is the same as that used in [7, 8]. That is, we consider the least prime factors of terms of A. For a given primitive sequence A and $m\ge 1$, with A_m and A'_m already defined, let $A''_m=\{a/p_m : a\in A'_m\}$. Clearly all of them are primitive. We study upper bounds F_m of $f(A_m)$. To find F_m we can suppose A_m is finite by Lemma 3. At first we find F_m for $m\ge N+1$, where $N=10^5$, by induction on $d^\circ(A_m)$. It is based on the fact that

$$f(A_m)=\sum_{i\ge m} f(A'_i) \quad\text{and}\quad f(A'_i)\le \frac{f(A''_i)}{p_i}.$$

Then we find F_m from F_{m+1} for $m=N,\ N-1,\dots,2,1$. It is based on the fact that

$$f(A_m)=f(A_{m+1})+f(A'_m) \quad\text{and}\quad f(A'_m)\le f(A''_m)/p_m.$$

At last

$$f(A)=f(A_1)\le F_1<1.84.$$

2. Proofs of the theorems

We have three constants: $N=10^5$, $\alpha=1.0072629$, and $k=0.0072847$. We need five lemmas.

Lemma 1. *We have $p_n\ge n(\log n+\log\log n-\alpha)$ for $n\ge 2$* [5].

Lemma 2. *We have*

$$\sum_{i\ge m}\frac{1}{p_i(\log i+\log\log i-k)} < \frac{1}{\log m+\log\log m-k}$$

for $m\ge N+1$.

Proof. Put $h(m)=\sum_{i\ge m} 1/\{p_i(\log i+\log\log i-k)\}$,

$g(m)=1/(\log m+\log\log m-k)$,

$c = (1 + 1/\log N)/N$,
$H(y) = (y + \log y - \alpha)(y + \log y - k)$, and
$G(y) = (y + \log y + c - k)^2$.

Then by Lemma 1,

$$h(m+1) < \int_m^\infty \frac{dx}{x(\log x + \log\log x - \alpha)(\log x + \log\log x - k)} = \int_{\log m}^\infty \frac{dy}{H(y)},$$

$$\begin{aligned} g(m+1) &> \frac{1}{\log m + 1/m + \log(\log m + 1/m) - k} \\ &> \frac{1}{\log m + \log\log m + (1 + 1/\log m)/m - k} \\ &\geq \frac{1}{\log m + \log\log m + c - k} \\ &= \int_{y=\log m}^\infty \frac{d(y + \log y + c - k)}{G(y)} = \int_{\log m}^\infty \frac{1 + 1/y}{G(y)}\, dy \quad \text{for } m \geq N. \end{aligned}$$

Since $k - \alpha - 2c + 1 > 10^{-8} > 0$,

$$\begin{aligned} H(y)(1 + 1/y) - G(y) &= (k - \alpha - 2c + 1)y + (k - \alpha - 2c + 2)\log y \\ &\quad + (\log y - \alpha)(\log y - k)/y + \alpha k - \alpha - k - (c - k)^2 \\ &> \log y + \alpha k - \alpha - k - (c - k)^2 > 1.4 > 0 \end{aligned}$$

for $y \geq \log m \geq \log N$; thus $h(m+1) < g(m+1)$ for $m \geq N$, i.e., $h(m) < g(m)$ for $m \geq N + 1$.

Lemma 3. *If c is a constant such that $f(S) \leq c$ for any finite subsequence S of a given primitive sequence $A = \{a_i\}$, then $f(A) \leq c$.*

Proof. Put $c_m = \sum_{1 \leq i \leq m} 1/(a_i \log a_i)$. Then $c_m \leq c$ and $c_m < c_{m+1}$ for $m \geq 1$; thus $f(A) = \lim_{m\to\infty} c_m \leq c$.

Lemma 4. *We have $f(A_m) \leq 1/(\log m + \log\log m - k)$ for $m \geq N + 1$ for any given primitive sequence $A = \{a_i\}$.*

Proof. By Lemma 3, we suppose A_m is finite. Induction on $d^\circ(A_m)$. If $d^\circ(A_m) \leq 1$, then

$$\begin{aligned} f(A_m) &\leq \sum_{i \geq m} \frac{1}{p_i \log p_i} < \sum_{i \geq m} \frac{1}{p_i(\log i + \log\log i - k)} \\ &< \frac{1}{\log m + \log\log m - k} \end{aligned}$$

by Lemmas 1 and 2.

Now suppose $d^\circ(A_m) = s > 1$. Since $A_m = \bigcup_{i \geq m} A_i'$ is disjoint, $f(A_m) = \sum_{i \geq m} f(A_i')$. If $d^\circ(A_i') \leq 1$ then

$$f(A_i') \leq \frac{1}{p_i \log p_i} < \frac{1}{p_i(\log i + \log\log i - k)}.$$

If $d^\circ(A_i') > 1$ then

$$f(A_i') < \frac{f(A_i'')}{p_i} < \frac{1}{p_i(\log i + \log\log i - k)}$$

since $d^\circ(A_i'') = d^\circ(A_i') - 1 < s$. Thus by Lemma 2,

$$\begin{aligned} f(A_m) = \sum_{i \geq m} f(A_i') &\leq \sum_{i \geq m} \frac{1}{p_i(\log i + \log\log i - k)} \\ &< \frac{1}{\log m + \log\log m - k}. \end{aligned}$$

Lemma 5. *If $f(A_{m+1}) \leq F_{m+1}$ for any primitive sequence A and*

$$F_m = \max\{F_{m+1}/(1 - 1/p_m),\ F_{m+1} + 1/(p_m \log p_m)\},$$

then $f(A_m) \leq F_m$ for any primitive sequence A.

Proof. Since $A_m = A_{m+1} \cup A_m'$ is disjoint,

$$f(A_m) = f(A_{m+1}) + f(A_m') \leq F_{m+1} + f(A_m').$$

By Lemma 3, we can suppose A_m is finite. Induction on $d^\circ(A_m)$. If $d^\circ(A_m) \leq 1$ or $d^\circ(A_m') \leq 1$ then

$$f(A_m') \leq 1/(p_m \log p_m)$$

and

$$f(A_m) \leq F_{m+1} + 1/(p_m \log p_m) \leq F_m.$$

Now suppose $d^\circ(A_m) = s > 1$ and $d^\circ(A_m') > 1$. Then $d^\circ(A_m'') < s$; and thus

$$f(A_m'') \leq F_m \quad \text{and} \quad f(A_m') < f(A_m'')/p_m \leq F_m/p_m.$$

Therefore,

$$f(A_m) \leq F_{m+1} + f(A_m') \leq F_{m+1} + F_m/p_m \leq F_m(1 - 1/p_m) + F_m/p_m = F_m.$$

Theorem 1. *$f(A) < 1.84$ for any primitive sequence A.*

Proof. Put $F_{N+1} = 1/\{\log(N+1) + \log\log(N+1) - k\}$. Then

$$f(A_{N+1}) \leq F_{N+1}$$

by Lemma 4. Let

$$F_m = \max\{F_{m+1}/(1 - 1/p_m),\ F_{m+1} + 1/(p_m \log p_m)\}.$$

Then

$$f(A_m) \leq F_m \quad \text{for } m = N, N-1, \ldots, 2, 1$$

by Lemma 5. Thus $f(A) = f(A_1) \leq F_1 < 1.84$. This completes the proof.

Theorem 2. *The necessary and sufficient condition for* (2) *to be true for any primitive sequence A is that $f(B) \leq \sum 1/(p \log p)$ for any primitive sequence B.*

Proof. (Necessity) If there exists a primitive sequence B such that $f(B) > \sum 1/(p \log p)$, then there exists an integer n such that

$$\sum_{b \in B,\ b \leq n} \frac{1}{b \log b} > \sum \frac{1}{p \log p},$$

thus

$$\sum_{b \in B,\ b \le n} \frac{1}{b \log b} > \sum_{p \le n} \frac{1}{p \log p};$$

this contradicts (2).

(Sufficiency) Suppose there exists a primitive sequence $A = \{a_i\}$ and an integer n such that

$$\sum_{a \in A,\ a \le n} \frac{1}{a \log a} > \sum_{p \le n} \frac{1}{p \log p}.$$

Let $B = \{a : a \in A,\ a \le n\} \cup \{p : p > n\}$. Then B is primitive and $f(B) > \sum 1/(p \log p)$. This is a contradiction.

Remark. From Lemma 1 and the inequality

$$p_n \le n(\log n + \log\log n - 0.5) \quad \text{for } n > 20 \text{ [6]}$$

we can get $1.63 < \sum 1/(p \log p) < 1.64$. From this fact and the above theorems we can see how far we are now from the complete proof of the conjecture.

References

1. P. Erdös, *Note on sequences of integers no one of which is divisible by any other*, J. London Math. Soc. (2) **10** (1935), 126–128.
2. ——, Seminar at the University of Limoges, 1988.
3. P. Erdös, A. Sarközi, and E. Szemerédi, *On divisibility properties of sequences of integers*, Colloq. Math. Soc. Janos Bolyai, vol. 2, Debrecen, Hungary, 1968.
4. H. Halberstam and K. F. Roth, *Sequences*, Chapter V, Oxford Univ. Press, 1966.
5. Guy Robin, *Estimation de la fonction de Tchebychef θ sur le k-ième nombre premier et grandes valeurs de la fonction $\omega(n)$ nombre de diviseurs premiers de n*, Acta Arith. **52** (1983), 367–389.
6. J. Barkley Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), 64–94.
7. Zhang Zhenxiang, *On a conjecture of Erdös on the sum $\sum_{p \le n} 1/(p \log p)$*, J. Number Theory **39** (1991), 14–17.
8. ——, *On a problem of Erdös concerning primitive sequences*, submitted.

Mathematical Institute, Hungarian Academy of Sciences, Budapest, Hungary

Department of Mathematics, Anhui Normal University, 241000 Wuhu, Anhui, People's Republic of China