DISCRETE ANALYSIS, 2020:16, 13 pp.
www.discreteanalysisjournal.com

A generalization of primitive sets and a conjecture of Erdős

Tsz Ho Chan Jared Duker Lichtman Carl Pomerance

Received 30 March 2020; Published 23 September 2020

Abstract: A set of integers greater than 1 is *primitive* if no element divides another. Erdős proved in 1935 that the sum of $1/(n \log n)$ for n running over a primitive set A is universally bounded over all choices for A. In 1988 he asked if this universal bound is attained by the set of prime numbers. We answer the Erdős question in the affirmative for 2-primitive sets. Here a set is 2-*primitive* if no element divides the product of 2 other elements.

Key words and phrases: primitive set, primitive sequence

1 Introduction and Statement of results

A set of integers greater than 1 is called primitive if no element divides any other. Erdős [4] showed that there is a constant K such that for any primitive set A,

$$f(A) := \sum_{n \in A} \frac{1}{n \log n} \leq K.$$

Further, in 1988 he asked if $f(A)$ is maximized by the primes $A = \mathbb{P}$. This is now referred to as the Erdős conjecture for primitive sets:

$$\text{For } A \text{ primitive, we have } \; f(A) \leq f(\mathbb{P}) = \sum_{p \in \mathbb{P}} \frac{1}{p \log p} =: C = 1.636616\cdots,$$

where $\mathbb{P}$ is the set of prime numbers. By a simple argument, the Erdős conjecture is equivalent to the assertion that $f(A) \leq f(\mathcal{P}(A))$ for any primitive set A, where $\mathcal{P}(A)$ denotes the set of primes dividing some member of A.

Recently, the second and third authors [9] proved that

 DOI: 10.19086/da.17290

Theorem 1. *For any primitive set A,*

$$f(A) < e^{\gamma} = 1.781072\cdots$$

where $\gamma = 0.5772\cdots$ is the Euler-Mascheroni constant. Further, if A does not contain a multiple of 8, then

$$f(A) \le f(\mathcal{P}(A)) + 2.37 \times 10^{-7}.$$

One can strengthen the notion of primitivity as follows. We say that a set A of integers greater than 1 with $|A| \ge k+1$ is *k-primitive* if no element divides the product of k distinct other elements. Note that k-primitive implies j-primitive for all $k \ge j \ge 1$.

In 1938, Erdős [6] first studied the maximal cardinality of 2-primitive sets in an interval. The first author together with Győri and Sárközy [3] extended it to all k and it was subsequently improved in [2] and [10]. While the full conjecture remains out of reach, we prove the Erdős conjecture for 2-primitive sets (and hence k-primitive for all $k \ge 2$).

Theorem 2. *For any 2-primitive set A,*

$$f(A) \le f(\mathcal{P}(A)).$$

An immediate consequence is the following

Corollary 1. *If A is a primitive set with $f(A) > f(\mathcal{P}(A))$, then there exist three elements $a, b, c \in A$ with $a \mid bc$.*

On the other hand, the primes are not optimal among primitive sets with respect to logarithmic density. Indeed, Erdős, Sárközy, and Szemerédi [8] obtained the best possible upper bound

$$\sum_{\substack{n \in A \\ n \le x}} \frac{1}{n} \le \left(\frac{1}{\sqrt{2\pi}} + o(1) \right) \frac{\log x}{\sqrt{\log\log x}}$$

for any primitive set A, while Erdős [7] showed that

$$\sum_{\substack{n \in A' \\ n \le x}} \frac{1}{n} \ge \left(\frac{1}{\sqrt{2\pi}} + o(1) \right) \frac{\log x}{\sqrt{\log\log x}}$$

where A' is the set of positive integers $a \le x$ with $\Omega(a) = [\log\log x]$. (Here, $\Omega(a)$ is the number of prime factors of a, counted with multiplicity.) By contrast, the primes satisfy

$$\sum_{p \le x} \frac{1}{p} = \log\log x + O(1).$$

Nevertheless, one may wonder if the primes still maximize the logarithmic density among 2-primitive sets. Indeed, we prove this to be the case.

Proposition 1. *For all $x \geq 2$ and any 2-primitive set A,*

$$\sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n} \leq \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p}$$

We use this to deduce Theorem 2.

Proof of Theorem 2 given Proposition 1. By Proposition 1, we have $F(x) \geq 0$ for all $x \geq 2$, where

$$F(x) := \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p} - \sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n}$$

Then by partial summation,

$$\sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p \log p} - \sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n \log n} = \frac{F(x)}{\log x} + \int_{2^-}^{x} \frac{F(u)}{u \log^2 u} du \geq 0.$$

Hence taking $x \to \infty$ gives $f(\mathcal{P}(A)) \geq f(A)$ as desired. □

In light of Proposition 1, it is natural to ask if there exists an exponent $\lambda < 1$ for which

$$\sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n^\lambda} \leq \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p^\lambda} \tag{1.1}$$

holds for all 2-primitive A, $x \geq 2$. Banks and Martin [1] settled the question in the setting of 1-primitive sets, proving (1.1) holds for all primitive A if and only if

$$\lambda \geq \tau_1 := 1.1403659 \cdots,$$

where $t = \tau_1$ is the unique real solution to the equation

$$\sum_{\mathbb{P}} p^{-t} = 1 + \Big(1 - \sum_{\mathbb{P}} p^{-2t}\Big)^{1/2}.$$

The fact that τ_1 is markedly larger than 1 gives some indication as to why the full Erdős conjecture remains open.

In the setting of 2-primitive sets, we extend the range of valid exponents λ.

Theorem 3. *For any $\lambda \geq 0.7983$, $x \geq 2$, and any 2-primitive set A,*

$$\sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n^\lambda} \leq \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p^\lambda} \tag{1.2}$$

We remark it suffices to verify Theorem 3 with $\lambda = 0.7983$. Indeed, suppose that $F_\lambda(x) \geq 0$ for all $x \geq 2$, where

$$F_t(x) = \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} p^{-t} - \sum_{\substack{n \in A \\ n \leq x}} n^{-t}.$$

Then, by partial summation, for any $t > \lambda$,

$$F_t(x) = x^{\lambda - t} F_\lambda(x) + (t - \lambda) \int_2^x u^{\lambda - t - 1} F_\lambda(u)\, du \geq 0.$$

Hence we may define the **critical exponent** τ_2 for 2-primitive sets, as the infimum over all λ for which (1.2) holds. Thus, Theorem 3 implies that $\tau_2 \leq 0.7983$.

We also note that Theorem 3 with $\lambda = 1$ gives us Proposition 1. However, Theorem 3 does not hold for every positive value of λ. Indeed, in [6], Erdős showed that there is a 2-primitive set A in $[1,x]$ of cardinality $\pi(x) - \pi(x^{1/3}) + cx^{2/3}/(\log x)^2$. It consists of primes in $(x^{1/3}, x]$ and a subset of $\{p_1 p_2 p_3 : p_i \text{ are primes } \leq x^{1/3}\}$ where the triples $\{p_1, p_2, p_3\}$ form a Steiner triple system. Thus, by the prime number theorem,

$$\sum_{n \in A} \frac{1}{n^\lambda} \geq \sum_{x^{1/3} < p \leq x} \frac{1}{p^\lambda} + \frac{cx^{2/3}}{(\log x)^2} \frac{1}{x^\lambda} > \sum_{p \leq x} \frac{1}{p^\lambda}$$

when $\lambda < 0.5$ and x is sufficiently large. Hence the above argument and Theorem 3 together imply that the critical exponent lies in the interval

$$0.5 \leq \tau_2 \leq 0.7983. \tag{1.3}$$

In a sequel paper, we shall address the question of critical exponents for k-primitive sets, with $k \geq 3$.

2 Combinatorial Lemmas

Before proving Theorem 3, we need lemmas in counting the maximal number of elements in a k-primitive set.

We first recall the following famous result due to Erdős and Szekeres [5], whose proof we provide for completeness.

Lemma 1 (Erdős–Szekeres). *A sequence of $(r-1)(s-1)+1$ real numbers has either a monotonic nondecreasing subsequence of length r or a monotonic nonincreasing subsequence of length s.*

Proof. Say the sequence is $a_1, a_2, \ldots, a_n$, where $n = (r-1)(s-1)+1$. For each a_i consider the ordered pair (b_i, c_i), where b_i is the length of the longest nondecreasing subsequence ending at a_i and c_i is the length of the longest nonincreasing subsequence ending at a_i. Then no two pairs (b_i, c_i) and (b_j, c_j) can be equal, so for at least one choice of i we have $b_i \geq r$ or $c_i \geq s$. □

We next bound the size of a k-primitive set based on the number of prime factors used to generate its elements.

Lemma 2. *For $k \geq 2$, suppose A is a k-primitive set and $T \subset A$ with $|\mathcal{P}(T)| = n$. If $n \leq k$, then $|T| \leq n$. If $n = k+1$, then $|T| \leq n+1$. Further, for $k = 2$, $n = 4$ we have $|T| \leq 19$.*

Proof. We may assume that $|T| \geq n$. Let $\mathcal{P}(T) = \{q_1, \ldots, q_n\}$ and write each $t = \prod_i q_i^{e_i} \in T$ as an exponent vector $\vec{v} = (e_1, \ldots, e_n)$. Define the notation $\vec{v} \geq 0$ if $e_i \geq 0$ for all i, and define $\vec{v} \leq \vec{w}$ if $\vec{w} - \vec{v} \geq 0$. Take $\vec{v}_1$ with maximal entry e_1 among T. Then take $\vec{v}_2$ with maximal e_2 among the remaining vectors, and similarly define $\vec{v}_3, \ldots, \vec{v}_n$. Thus, the chosen vectors are distinct.

Case $n \leq k$: If $|T| \geq n+1$ then T has some vector $\vec{v} \neq \vec{v}_i$ for all i. But then $\vec{v} \leq \vec{v}_1 + \cdots + \vec{v}_n$. This implies that T, and hence A, is not n-primitive, and since $n \leq k$, it implies that A is not k-primitive, a contradiction. Hence we cannot have $|T| \geq n+1$ when $n \leq k$.

Case $n = k+1$: If $|T| \geq n+2$ then T has vectors $\vec{w}_1 \neq \vec{w}_2$ with $\vec{w}_j \notin \{\vec{v}_1, \ldots \vec{v}_n\}$ for $j = 1, 2$. Write $\vec{w}_j = (f_1^{(j)}, \ldots, f_n^{(j)})$. By the pigeonhole principle, we may assume

$$f_i^{(1)} \leq f_i^{(2)}$$

for at least $n/2$ values of i, say $i = 1, \ldots, \lceil n/2 \rceil$. Thus, we deduce

$$\vec{w}_1 \leq \vec{w}_2 + \vec{v}_{\lceil n/2 \rceil + 1} + \cdots + \vec{v}_n$$

contradicting T as k-primitive, since $1 + \lfloor n/2 \rfloor = 1 + \lfloor (k+1)/2 \rfloor \leq k$.

Now say $k = 2$, $n = 4$. Suppose there are 20 members in T with corresponding vectors

$$\vec{w}_i := (e_{i,1}, e_{i,2}, e_{i,3}, e_{i_4}) \text{ for } 1 \leq i \leq 20.$$

Since A is 2-primitive, so is T. Without loss of generality, say $\vec{w}_{18}$ has maximal first coordinate, $\vec{w}_{19} \neq \vec{w}_{18}$ has maximal second coordinate among the remaining 19 vectors, and $\vec{w}_{20}$ has maximal third coordinate among the remaining 18 vectors with $\vec{w}_{20} \neq \vec{w}_{18}, \vec{w}_{19}$. Arrange the remaining 17 vectors in ascending order of their first coordinate (i.e., $e_{1,1} \leq e_{2,1} \leq \ldots \leq e_{17,1}$). By Lemma 1, there is a monotonic sequence of length 5 among the $e_{i,2}$'s. Without loss of generality, say $e_{1,2}$, $e_{2,2}$, $e_{3,2}$, $e_{4,2}$, $e_{5,2}$ form such a sequence.

Case 1: $e_{1,2} \leq e_{2,2} \leq e_{3,2} \leq e_{4,2} \leq e_{5,2}$. Consider the numbers $e_{i,3}$ for $i = 1, \ldots, 5$. By Lemma 1, there is a monotonic sequence of length 3 among the $e_{i,3}$'s, without loss of generality, say it is $e_{1,3}, e_{2,3}, e_{3,3}$. If $e_{1,3} \leq e_{2,3} \leq e_{3,3}$, this forces $e_{2,4} > e_{1,4} + e_{3,4}$ for otherwise $\vec{w}_2 \leq \vec{w}_1 + \vec{w}_3$, contradicting T being 2-primitive. But this implies that $\vec{w}_1 \leq \vec{w}_2$ which contradicts T being primitive. Hence, we must have $e_{1,3} \geq e_{2,3} \geq e_{3,3}$. Again, this forces $e_{2,4} > e_{1,4} + e_{3,4}$, which in turn implies that $\vec{w}_1 \leq \vec{w}_2 + \vec{w}_{20}$, again a contradiction.

Case 2: $e_{1,2} \geq e_{2,2} \geq e_{3,2} \geq e_{4,2} \geq e_{5,2}$. By Lemma 1, there is a monotonic sequence of length 3 among the $e_{i,3}$'s, without loss of generality, say it is $e_{1,3}, e_{2,3}, e_{3,3}$. If $e_{1,3} \leq e_{2,3} \leq e_{3,3}$, then again this forces $e_{2,4} > e_{1,4} + e_{3,4}$. But then $\vec{w}_1 \leq \vec{w}_2 + \vec{w}_{19}$. Hence, we must have $e_{1,3} \geq e_{2,3} \geq e_{3,3}$. This forces $e_{2,4} > e_{1,4} + e_{3,4}$. But then $\vec{w}_3 \leq \vec{w}_2 + \vec{w}_{18}$, again a contradiction.

Therefore, there can be at most 19 members in T. □

Remark 2.1. It is not clear if the number "19" in Lemma 2 is optimal. We will not need it here, but by similar methods one can prove that if T is a 2-primitive set of positive integers with $|\mathcal{P}(T)| = n \geq 3$, then $|T| \leq 9^{2^{n-3}}$.

3 Proof of Theorem 3

Let $A \subset (1,x]$ be a 2-primitive set. Let $0.79 \leq \lambda < 1$ be a parameter to be defined later. First, we partition A into primes S and composites T. Note S and $\mathcal{P}(T)$ are disjoint since A is primitive. For a prime p, define

$$T_p := \{t \in T : p \mid t\}.$$

If some prime $p \in \mathcal{P}(T)$ satisfies

$$\sum_{t\in T_p} \frac{1}{t^\lambda} \leq \frac{1}{p^\lambda}, \tag{3.1}$$

then we replace the members of T_p with the prime p (i.e., redefine $A = (T \backslash T_p) \cup \{p\}$). This would make $\sum_{T_p} t^{-\lambda}$ at least as big while keeping A 2-primitive. Repeat the process with each prime $p \in \mathcal{P}(T)$ until no such prime satisfies (3.1). If $T = \emptyset$ after doing this, then $A = S$ consists of primes so Proposition 1 follows. Otherwise $T \neq \emptyset$, so we may assume

$$\sum_{t\in T_p} \frac{1}{t^\lambda} > \frac{1}{p^\lambda} \qquad \text{for all} \quad p \in \mathcal{P}(T). \tag{3.2}$$

Consider the set

$$D := \{t/p : t \in T,\ p \mid t\} \tag{3.3}$$

We record some useful properties of T and D.

Lemma 3. *Let T be a 2-primitive set for which (3.2) holds and let D be as in (3.3).*

(i) For each $p \in \mathcal{P}(T)$, T_p has at least 3 elements.

(ii) The map sending ordered pairs (t, p) with $t \in T$ and $p \mid t$ to $t/p \in D$ is injective.

(iii) Each $t \in T$ has at least 3 prime factors (counted with multiplicity).

(iv) D is a primitive set of composite numbers.

Proof. (i) For $p \in \mathcal{P}(T)$, (3.2) implies that

$$\sum_{t\in T_p} \frac{1}{(t/p)^\lambda} > 1 > 2^{-0.79} + 3^{-0.79},$$

Thus (i) follows, since $t/p \in \mathbb{Z}_{>1}$ for all $t \in T_p$.

(ii) If not, then $t_1/p_1 = t_2/p_2$ for some t_1, t_2, $p_1 \mid t_1$, and $p_2 \mid t_2$. If $t_1 \neq t_2$, by (i) there exists some $p_1 k \in T_{p_1}$ other than t_1, t_2. But then $t_1 = (t_1/p_1)p_1 = (t_2/p_2)p_1 \mid t_2(p_1 k)$, which contradicts T as 2-primitive. Hence $t_1 = t_2$, which forces $p_1 = p_2$.

(iii) If not, say $t = pq$. Since T_p, T_q each have at least 3 elements, there are some pm and qn other than $t \in T$. But then, $t = pq \mid (pm)(qn)$ which contradicts T as 2-primitive. (This argument holds whether or not $p \neq q$.)

(iv) If not, then $(t/p) \mid (t_1/p_1)$ for some $t, t_1 \in T$, $p \mid t$, $p_1 \mid t_1$, and $t/p \neq t_1/p_1$. If $p_1 = p$, then $t \mid t_1$ which contradicts T as primitive. And if $p_1 \neq p$, then there is some $pl \in T_p$ other than t and t_1. This implies $t \mid t_1 \cdot pl$, and since $t \neq t_1$ (otherwise $p = p_1$), we have a contradiction to T being 2-primitive. Thus D is primitive, and also composite by (iii). □

For Theorem 3, we must show

$$\sum_{t \in T} \frac{1}{t^\lambda} - \sum_{p \in \mathcal{P}(T)} \frac{1}{p^\lambda} < 0. \tag{3.4}$$

Suppose $\mathcal{P}(T)$ consists of primes $q_1 < q_2 < \cdots < q_r$. Let $2 = p_1 < p_2 < \cdots < p_r$ be the first r primes in $\mathbb{P}$. We are going to modify the set T by the following process. First, if each $q_i = p_i$, we let T stand as it is. Otherwise, let i be the smallest index such that $q_i > p_i$. Then $q_j = p_j$ for all $j < i$ and we have $p_i \nmid t$ for all $t \in T$. Then replace each $t \in T_{q_i}$ with $p_i/q_i \cdot t$. This keeps T as 2-primitive, and by (3.2),

$$0 < \sum_{t \in T_{q_i}} \frac{1}{t^\lambda} - \frac{1}{q_i^\lambda} < \frac{q_i^\lambda}{p_i^\lambda}\left(\sum_{t \in T_{q_i}} \frac{1}{t^\lambda} - \frac{1}{q_i^\lambda}\right) = \sum_{t \in T_{q_i}} \frac{1}{(p_i/q_i \cdot t)^\lambda} - \frac{1}{p_i^\lambda}.$$

So replacing each $t \in T_{q_i}$ with $p_i/q_i \cdot t$ preserves (3.2). We repeat this process for each i with $q_i > p_i$ and in the end we have $\mathcal{P}(T) = \{p_1, p_2, \ldots, p_r\}$. By showing (3.4) for this T it would follow that (3.2) fails for some p_i, and this contradiction would prove the theorem.

We have reduced Theorem 3 to the following.

Theorem 3.1. *Suppose $\lambda \geq 0.7983$ and T is a 2-primitive set of composite numbers satisfying (3.2) with $\mathcal{P}(T) = \mathbb{P} \cap (1, Y]$ for some Y. Then*

$$\sum_{t \in T} \frac{1}{t^\lambda} - \sum_{p \leq Y} \frac{1}{p^\lambda} < 0. \tag{3.5}$$

Our goal now is to prove Theorem 3.1. For a parameter $0 < \theta < 1$ to be chosen later, we define λ as

$$\lambda = \tau(1-\theta), \text{ where } \tau = 1.140366. \tag{3.6}$$

First consider those $t \in T$ with greatest prime factor $P(t) \geq t^\theta$. Then $t^{1-\theta} \geq t/P(t)$ and so $t^{-\lambda} \leq (t/P(t))^{-\lambda/(1-\theta)} = (t/P(t))^{-\tau}$. Hence

$$\sum_{\substack{t \in T \\ P(t) \geq t^\theta}} t^{-\lambda} \leq \sum_{\substack{t \in T \\ P(t) \geq t^\theta}} \left(\frac{t}{P(t)}\right)^{-\tau} \leq \sum_{p \leq Y} p^{-\tau} \tag{3.7}$$

by (1.1), since $\{t/P(t) : t \in T\} \subset D$ is primitive by part (iii) of Lemma 3.

For a positive integer t, we consider the following unique factorization

$$t = m(t)M(t)$$

into positive integers $m(t) \leq M(t)$ with ratio $M(t)/m(t)$ minimal. Let

$$\mathcal{M}(T) = \{m(t) : t \in T\} \cup \{M(t) : t \in T\}.$$

We need two lemmas.

Lemma 4. *For any* 2*-primitive set* T*, consider the graph on the integers with edges* $\{t,m(t)\}$ *and* $\{t,M(t)\}$ *for* $t \in T$*, where if* $m(t) = M(t)$*, there is just one edge containing* t*. This graph contains a matching from* T *into* $\mathcal{M}(T)$*.*

Proof. Let $t \in T$. If $m(t) \notin \{m(t'),M(t')\}$ for all other $t' \in T$, then we can match t with $m(t)$. So assume $m(t) \in \{m(t'),M(t')\}$ for some other $t' \in T$. Then $M(t) \notin \{m(t''),M(t'')\}$ for all $t'' \in T$ with $t'' \neq t,t'$, since otherwise $t \mid t't''$, contradicting T being 2-primitive.

If $m(t) < M(t)$, then 2-primitive implies $M(t) \notin \{m(t'),M(t')\}$ so we can match t with $M(t)$.

Otherwise $m(t) = M(t)$, which means $t = m(t)^2$. Then $t' \neq t$ forces $m(t') < M(t')$, so we make define $m' = t'/m(t)$ (that is m' is the singleton in $\{m(t'),M(t')\} \setminus \{m(t)\}$). We would like to match t' with m' instead of $m(t)$, freeing up $m(t)$ to be matched with t. So suppose this is blocked by some t'' different from t' (and necessarily different from t) with $m' \in \{m(t''),M(t'')\}$. But then $t' \mid tt''$, a violation of 2-primitivity. Thus, the matching can be completed. □

Lemma 5. *Suppose* $0 < \theta < 1/3$ *and that* T *is* 2*-primitive with* $P(t) < t^{\theta}$ *for each* $t \in T$*. Let* $N(z) = |T \cap [2,z]|$*. Then, with* q *running over primes in the interval* $I := [z^{(1+\theta)/4}, z^{(1+\theta)/2})$*, we have*

$$N(z) < z^{(1+\theta)/2} - \sum_{q \in I} \left\lfloor \frac{z^{(1+\theta)/2}}{q} \right\rfloor$$

Proof. By Lemma 4, it suffices to bound $|\mathcal{M}(T \cap [2,z])|$. We first show that $\mathcal{M}(T \cap [2,z]) \subset [1, z^{(1+\theta)/2})$. Let $t \in T$ with $t \leq z$. Say $t = q_1 q_2 ... q_r$ where the primes q_i are written in nondecreasing order. Let $d = q_1 q_2 ... q_i$ be maximal with $d \leq t^{(1-\theta)/2}$. Then $d' = dq_{i+1}$ satisfies $t^{(1-\theta)/2} < d' < t^{(1+\theta)/2}$. Also, $d'' = t/d'$ satisfies the same double inequality. Thus,

$$t^{(1-\theta)/2} < m(t) \leq M(t) < t^{(1+\theta)/2} \leq z^{(1+\theta)/2}.$$

We further note that the members m of $\mathcal{M}(T \cap [2,z])$ satisfy $P(m) < z^{\theta}$, since m divides some member of $T \cap [2,z]$ and every t in that set has $P(t) < z^{\theta}$. In particular, m is not divisible by any prime $q \geq z^{\theta}$. Note that if $\theta < 1/3$, then $\theta < (1+\theta)/4$. So, m is not divisible by any prime in the interval I. Since no integer below $z^{(1+\theta)/2}$ is divisible by 2 primes from I, the lemma follows. □

Set

$$T^p = \{t \in T : P(t) = p\},$$

so that $T^p \subset T_p$. We have the following variant of Lemma 5.

Lemma 6. *For any* 2*-primitive set* T *and prime* p*, let* $N_p(z)$ *denote the number of members* t *of* T^p *with* $t \leq z$*. With* q *running over the primes in* $I_p := (\max\{p, z^{1/4}\}, z^{1/2})$*, we have*

$$N_p(z) \leq z^{1/2} - \sum_{q \in I_p} \left\lfloor \frac{z^{1/2}}{q} \right\rfloor.$$

Proof. Note that if T is 2-primitive, so too is $T^p/p = \{t/p : t \in T^p\}$. Thus, we may apply Lemma 4 to obtain a matching from T^p/p into $\mathcal{M}(T^p/p)$. The prime factors of each element $t/p \in T^p/p$ are at most p, so following the proof of Lemma 5, we have $m(t/p), M(t/p) \in [t^{1/2}/p, t^{1/2})$. The lemma then follows in the same way as Lemma 5. □

Lemma 7. *For $x \geq 2$ we have*

$$\sum_{\substack{x^{1/2}<q<x \\ q \text{ prime}}} \left\lfloor \frac{x}{q} \right\rfloor \geq \left(\log 2 - \frac{1.25}{\log x} - \frac{2.5}{(\log x)^2} \right) x.$$

Proof. First suppose that $x \geq 286^2$. We have the sum is at least

$$\sum_{x^{1/2}<q<x} \frac{x}{q} - \pi(x).$$

From [11, (3.7)], we have that $\pi(x) < 1.25x/\log x$ and from [11, (3.17)] that

$$\sum_{q<x} \frac{1}{q} > \log\log x + B - \frac{1}{2(\log x)^2},$$

where B is the Mertens constant. Further, from [11, (3.18)],

$$\sum_{q \leq x^{1/2}} \frac{1}{q} < \log\log x^{1/2} + B + \frac{1}{2(\log x^{1/2})^2} = \log\log x - \log 2 + B + \frac{2}{(\log x)^2}.$$

This proves the lemma in the range $x \geq 286^2$ and direct calculation shows that it holds in the wider range $x \geq 2$. □

We shall find it useful to use the following asymptotically weaker estimates in small cases. The proof follows by checking values of $x \leq 3213$ after which Lemma 7 is stronger.

Corollary 2. *For $x \geq 185$, we have $\sum_{q \in (x^{1/2}, x]} \lfloor x/q \rfloor > 0.5x$. For $x \geq 67$, we have $\sum_{q \in (x^{1/2}, x]} \lfloor x/q \rfloor > 0.45x$.*

Let

$$\theta = 0.3, \quad \lambda = 0.7982562, \quad \nu = 1/\theta = 10/3. \tag{3.8}$$

For each prime p, let

$$S_p = \sum_{\substack{t \in T \\ P(t)=p<t^\theta}} \frac{1}{t^\lambda}.$$

With (3.7) it will suffice to prove Theorem 3.1 if we show under its hypotheses that for each $Y \geq 2$,

$$\sum_{p \leq Y} S_p \leq \sum_{p \leq Y} \left(\frac{1}{p^\lambda} - \frac{1}{p^\tau} \right). \tag{3.9}$$

3.1 Small primes, $Y \le 37$

We are going to estimate S_p for various small primes p. Take $t \in T$ with $P(t) < t^\theta$. If $t \le q^\nu$ for a prime q, then $P(t) < (q^\nu)^\theta = q$. If $q = 3$, we see there can be at most one such t; that is, T can contain at most one power of 2. The values of $t \le 5^\nu$ are supported on $\{2,3\}$, so by Lemma 2 with $k = n = 2$ we see that there are at most 2 such members of T. Similarly, Lemma 2 with $k = 2$, $n = 3$ shows that T has at most 4 members below 7^ν, and with $k = 2, n = 4$, T has at most 19 members below 11^ν. Since members t of T with $P(t) < t^\theta$ have at least $\lceil \nu \rceil = 4$ prime factors (counted with multiplicity), we have

$$\begin{aligned} S_2 &\le \frac{1}{2^{4\lambda}} < 0.1093463, \\ S_2 + S_3 &< 0.1093463 + \frac{2-1}{3^{\nu\lambda}} < 0.1631052, \\ S_2 + S_3 + S_5 &< 0.1631052 + \frac{4-2}{5^{\nu\lambda}} < 0.1907220, \\ S_2 + S_3 + S_5 + S_7 &< 0.1907220 + \frac{19-4}{7^{\nu\lambda}} < 0.2753295. \end{aligned} \tag{3.10}$$

Computing $\sum_{p \le Y}(1/p^\lambda - 1/p^\tau)$ directly for $Y = 2,3,5,7$ gives lower bounds

$$0.121399,\ 0.251741,\ 0.368904,\ 0.471733,$$

respectively. Thus we observe $\sum_{p \le Y} S_p < \sum_{p \le Y}(1/p^\lambda - 1/p^\tau)$, so by (3.9), Theorem 3.1 holds when $Y = 2,3,5,7$, respectively.

Now consider $11 \le p \le 37$. By partial summation, we have the equality

$$S_p = \int_{p^\nu}^{\infty} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz, \tag{3.11}$$

noting that the integral converges, since $N_p(z) \le z^{(1+\theta)/2}$ by Lemma 5.

We use Lemmas 6 and 7 to get the upper estimates for $N_p(z)$:

$$N_p(z) \le \lfloor \sqrt{z} \rfloor - \sum_{\max(p, z^{1/4}) < q \le \sqrt{z}} \Big\lfloor \frac{\sqrt{z}}{q} \Big\rfloor, \tag{3.12}$$

$$N_p(z) \le \sqrt{z}\Big(1 - \log 2 + \frac{2.5}{\log z} + \frac{10}{(\log z)^2}\Big), \text{ when } p \le z^{1/4}. \tag{3.13}$$

We split the integral in (3.11) at p^4. In the first range when $z < p^4$, we bound the contribution to (3.11) by splitting up into intervals $[m^2, (m+1)^2]$ and using (3.12), which gives

$$\begin{aligned} S_p' := \int_{p^\nu}^{p^4} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz \ &\le \sum_{m_0 < m < p^2} \int_{m^2}^{(m+1)^2} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz \ + \int_{p^\nu}^{(m_0+1)^2} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz \\ &\le \sum_{m_0 < m < p^2} \Big(\frac{1}{m^{2\lambda}} - \frac{1}{(m+1)^{2\lambda}}\Big)\Big(m - \sum_{p < q \le m} \lfloor \frac{m}{q} \rfloor\Big) \\ &\quad + \Big(\frac{1}{p^{\nu\lambda}} - \frac{1}{(m_0+1)^{2\lambda}}\Big)\Big(m_0 - \sum_{p < q \le p^{\nu/2}} \lfloor \frac{m_0}{q} \rfloor\Big) \end{aligned} \tag{3.14}$$

where $m_0 = \lfloor p^{\nu/2} \rfloor$.

For the second range when $z \geq p^4$, we use (3.13) when $z \geq 3213^2$ and for smaller values of z we use Corollary 2. That is,

$$S_p'' := \int_{p^4}^{\infty} \frac{\lambda}{z^{1+\lambda}} N_p(z)\,dz \;\leq \int_{3213^2}^{\infty} \frac{\lambda}{z^{\lambda+1/2}} \Big(1 - \log 2 + \frac{2.5}{\log z} + \frac{10}{(\log z)^2}\Big)\,dz$$
$$+\; 0.5 \int_{\max(p^4,185^2)}^{3213^2} \frac{\lambda}{z^{1/2+\lambda}}\,dz \;+\; 0.55 \int_{p^4}^{\max(p^4,185^2)} \frac{\lambda}{z^{1/2+\lambda}}\,dz.$$

Denote the integrals

$$f(y) := \int_y^{\infty} \frac{\lambda}{z^{\lambda+1/2}}\,dz$$
$$g(y) := \int_y^{\infty} \frac{\lambda}{z^{\lambda+1/2}} \Big(1 - \log 2 + \frac{2.5}{\log z} + \frac{10}{(\log z)^2}\Big)\,dz.$$

So we obtain

$$\begin{aligned} S_p'' \leq\; & (1-\log 2) f(3213^2) + g(3213^2) \\ & +\; 0.5[f(\max(p^4,185^2)) - f(3213^2)] \;+\; 0.55[f(p^4) - f(\max(p^4,185^2))] \\ =\; & (0.5 - \log 2) f(3213^2) + g(3213^2) \;-\; 0.05 f(\max(p^4,185^2)) \;+\; 0.55 f(p^4). \end{aligned} \tag{3.15}$$

Using the estimates in (3.14), (3.15), we bound $S_p = S_p' + S_p''$ by the following.

p	$S_p \leq$	$\sum_{q \leq p} S_q \leq$	$\sum_{q \leq p} (q^{-\lambda} - q^{-\tau}) \geq$
11	0.13259	0.40792	0.55427
13	0.11241	0.52033	0.62966
17	0.08382	0.60415	0.69432
19	0.07601	0.68016	0.75484
23	0.06194	0.74210	0.80868
29	0.04757	0.78967	0.85521
31	0.04501	0.83468	0.89978
37	0.03680	0.87148	0.93950

Note that the first entry in the third column is found by adding S_{11} to the estimate in (3.10). Since the entries in the fourth column exceed the entries in the third column, (3.9) implies Theorem 3.1 for $Y \leq 37$.

3.2 Large primes, $Y \geq 41$

Now assume that $Y \geq 41$. We have via partial summation that

$$\sum_{\substack{t \in T \\ P(t) < t^\theta}} \frac{1}{t^\lambda} = \sum_{p \leq 7} S_p + \sum_{11 \leq p \leq 23} \int_{p^\nu}^{29^\nu} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz + \int_{29^\nu}^{\infty} \frac{\lambda}{z^{1+\lambda}} N(z)\, dz.$$

(As before, the last integral converges.) From (3.10) the S_p terms contribute at most 0.27533. Using Lemmas 5, 6, and 7, and Corollary 2, we similarly obtain

$$\sum_{\substack{t \in T \\ P(t) < t^\theta}} \frac{1}{t^\lambda}$$
$$< 0.27533 + 0.08455 + 0.06576 + 0.03756 + 0.02953 + 0.01487 + 0.45614 = 0.96374,$$

where the second to the sixth terms correspond to the five finite integrals, and the last term is our estimate for the tail integral. We also note that

$$\sum_{p \leq Y} \left(\frac{1}{p^\lambda} - \frac{1}{p^\tau} \right) \geq \sum_{p \leq 41} \left(\frac{1}{p^\lambda} - \frac{1}{p^\tau} \right) > 0.97661.$$

Since this estimate exceeds the prior one, this gives Theorem 3.1 with $\lambda = 0.7982562$.

Acknowledgments

The authors would like to acknowledge the University of Memphis for hosting the 2019 Erdős Lecture Series, during which the initial ideas for this paper were conceived. The second author was supported by a Clarendon Scholarship at the University of Oxford. The third author is grateful for the hospitality of Santa Clara University and was funded by their Paul R. and Virginia P. Halmos Endowed Professorship in Mathematics and Computer Science. The authors thank the referee for a careful reading.

References

[1] W.D. Banks and G. Martin, Optimal primitive sets with restricted primes, *Integers* **13** (2013), #A69, 10 pp. 3

[2] T.H. Chan, On sets of integers, none of which divides the product of k others, *European J. Comb.* **32** (2011), 443–447. 2

[3] T.H. Chan, E. Győri and A. Sárközy, On a problem of Erdős on integers, none of which divides the product of k others, *European J. Comb.* **31** (2010), 260–269. 2

[4] P. Erdős, Note on sequences of integers no one of which is divisible by any other, *J. London Math. Soc.* **10** (1935), 126–128. 1

[5] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470. 4

[6] P. Erdős, On sequences of integers no one of which divides the product of two others and on some related problems, *Tomsk. Gos. Univ. Ucen. Zap.* **2** (1938), 74–82. 2, 4

[7] P. Erdős, On the integers having exactly k prime factors, *Annals of Math.* **49** (1948), 53–66. 2

[8] P. Erdős, A. Sárközy and E. Szemerédi, On an extremal problem concerning primitive sequences, J. London Math. Soc. **42** (1967), 484-488. 2

[9] J.D. Lichtman and C. Pomerance, The Erdős conjecture for primitive sets, *Proc. Amer. Math. Soc.*, Series B **6** (2019), 1–14. 1

[10] P.P. Pach and Cs. Sándor, Multiplicative bases and an Erdős problem, *Combinatorica* **38** (2018), 1175–1203. 2

[11] J.B. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.* **6** (1962), 64–94. 9

AUTHORS

Tsz Ho Chan
Kennesaw State University
Kennesaw, GA 30144
thchan6174@gmail.com
https://www.memphis.edu/msci/people/tchan.php

Jared Duker Lichtman
University of Oxford
Oxford, OX2 6GG, UK
jared.d.lichtman@gmail.com
https://www.maths.ox.ac.uk/people/jared.lichtman

Carl Pomerance
Dartmouth College
Hanover, NH 03755
carl.pomerance@dartmouth.edu
https://math.dartmouth.edu/~carlp/