Joint Mathematics Meetings
Baltimore, MD, January 15–19, 2019

The Erdős conjecture on primitive sets

Carl Pomerance

Dartmouth College (emeritus)

University of Georgia (emeritus)

Joint work with **Jared Duker Lichtman**

A subset $\mathcal{A}$ of the positive integers is said to be *primitive* if no member of $\mathcal{A}$ divides another.

Some examples:

1. The set $\mathcal{P}$ of prime numbers.

2. The set $\mathbb{N}_k = \{n : \Omega(n) = k\}$, where $\Omega(n)$ is the number of prime factors of n counted with repetition.

3. The squarefree members of $\mathbb{N}_k$.

4. The set $(x, 2x] \cap \mathbb{N}$.

5. With σ the sum-of-divisors function, the set

$$\mathcal{A} = \{n \in \mathbb{N} : \sigma(n)/n \ge 2,\ \ \sigma(d)/d < 2 \text{ for all } d \mid n, d < n\}.$$

The last example is the set of *primitive nondeficient numbers*. They have a finite reciprocal sum (Erdős, 1934), so the nondeficient numbers (i.e., $\sigma(n)/n \ge 2$) have an asymptotic density. We now know (Kobayashi, 2010) that this density is between 0.2476171 and 0.2476475. And we know (Lichtman, 2018) that the sum of reciprocals of primitive nondeficient numbers is between 0.34842 and 0.37937.

One might guess that a primitive set always has asymptotic density 0. It's true for our 5 examples, and more generally, it's true that the lower asymptotic density of a primitive set must be 0. Somewhat counter-intuitively, the upper asymptotic density need not be zero!

Here's a construction. The set of integers in the interval $(x, 2x]$ is primitive; let $D(x)$ be the asymptotic density of the set of multiples. We know after Besicovitch (1934) that $\liminf D(x) = 0$. In fact, after work of Erdős, Tenenbaum, and Ford, we now know that

$$D(x) \asymp \frac{1}{(\log x)^c (\log\log x)^{3/2}}, \quad c = 1 - \frac{1 + \log\log 2}{\log 2} = 0.08607\ldots.$$

But already, just from Besicovitch's result, we can construct primitive sets with upper asymptotic density arbitrarily close to 1/2. Namely, choose a very briskly increasing sequence $x_1, x_2, \ldots$ with $D(x_1)$ very small and $D(x_j) \downarrow 0$. Take all numbers in $(x_1, 2x_1]$, all numbers in $(x_2, 2x_2]$ not divisible by any number in $(x_1, 2x_1]$, all numbers in $(x_3, 2x_3]$ not divisible by any number previously chosen, etc.

This result is best possible: The upper asymptotic density of a primitive set is always $< 1/2$.

From now on we only consider primitive sets $\mathcal{A} \neq \{1\}$.

Theorem (Erdős, 1935). *If $\mathcal{A}$ is a primitive set, then*

$$f(\mathcal{A}) := \sum_{a \in \mathcal{A}} \frac{1}{a \log a} < \infty.$$

In fact, $f(\mathcal{A})$ is uniformly bounded as $\mathcal{A}$ varies over primitive sets.

With $\mathcal{P}$ the set of primes, let $C = f(\mathcal{P}) = 1.63661632336\ldots$, the calculation done by Cohen.

Conjecture (Erdős, 1988). *For $\mathcal{A}$ primitive, $f(\mathcal{A}) \le C$.*

Conjecture (Erdős, 1988). *For $\mathcal{A}$ primitive,* $f(\mathcal{A}) \le C = 1.63661632336\ldots$, *where* $f(\mathcal{A}) = \sum_{a \in \mathcal{A}} 1/(a \log a)$.

What do we know about $f(\mathcal{A})$?

Erdős, Zhang (unpublished): $f(\mathcal{A}) < 2.886$.

Robin (unpublished): $f(\mathcal{A}) < 2.77$.

Erdős, Zhang (1993): $f(\mathcal{A}) < 1.84$.

The first two results used the original Erdős argument, but the 1993 paper used a new argument.

Let $\mathcal{P}(\mathcal{A})$ denote the set of primes which divide some member of $\mathcal{A}$. Note that the Erdős conjecture is equivalent to the assertion that

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

for all primitive sets $\mathcal{A}$. Indeed, if $f(\mathcal{A}) > f(\mathcal{P}(\mathcal{A}))$ for some primitive set $\mathcal{A}$, let $\mathcal{A}'$ be $\mathcal{A}$ together with every prime not in $\mathcal{P}(\mathcal{A})$, so that $\mathcal{A}'$ is primitive, $\mathcal{P}(\mathcal{A}') = \mathcal{P}$, and $f(\mathcal{A}') > f(\mathcal{P})$.

Recall that $\mathbb{N}_k = \{n : \Omega(n) = k\}$.

Zhang, (1991): *$f(\mathcal{A}) \le C$ if each $a \in \mathcal{A}$ has $\Omega(a) \le 4$.*

Zhang, (1993): *For each $k \ge 2$, $f(\mathbb{N}_k) < f(\mathbb{N}_1) = C$.*

Banks, Martin, (2013): *If $\sum_{p \in \mathcal{P}(\mathcal{A})} 1/p < 1.7401\ldots$, then $f(\mathcal{A}) \le f(\mathcal{P}(\mathcal{A}))$.*

Banks, Martin, (2013): **Conjecture**: $f(\mathbb{N}_1) > f(\mathbb{N}_2) > f(\mathbb{N}_3) \ldots$.

Bayless, Kinlaw, Klyve, (2018): $f(\mathbb{N}_2) > f(\mathbb{N}_3)$.

Lichtman, P, (2018). *For $\mathcal{A}$ primitive,*

- $f(\mathcal{A}) < e^{\gamma} = 1.78109\ldots$.

- *If no member of $\mathcal{A}$ is divisible by 8, then* $f(\mathcal{A}) < f(\mathcal{P}(\mathcal{A})) + 2.37 \times 10^{-7}$.

- *There is an absolute constant $c > 0$ such that $f(\mathbb{N}_k) > c$ for all k.*

- *Assuming RH and LI, there is a set of primes $\mathcal{P}_0$ of relative lower logarithmic density ≥ 0.995 such that $f(\mathcal{A}) \leq f(\mathcal{P}(\mathcal{A}))$ when $\mathcal{P}(\mathcal{A}) \subset \mathcal{P}_0$. Unconditionally, $\mathcal{P}_0$ contains all of the odd primes up to* $\exp(10^6)$.

Note: The relative lower logarithmic density of a set of primes $\mathcal{P}_0$ is

$$\liminf_{x\to\infty} \frac{1}{\log\log x} \sum_{\substack{p\in\mathcal{P}_0\\ p\le x}} \frac{1}{p}.$$

In a new paper **Lichtman, Martin, & P** show that $\mathcal{P}_0$ has relative lower asymptotic density at least $0.99999973\ldots$.

Notation: For an integer $a \ge 2$, let

$$p(a) := \min\{p\in\mathcal{P} : p \mid a\},$$
$$P(a) := \max\{p\in\mathcal{P} : p \mid a\}.$$

A slightly cleaned up version of the 1935 Erdős argument:

Let $S_a = \{ba : p(b) \ge P(a)\}$. The asymptotic density of S_a is

$$\delta(S_a) = \frac{1}{a} \prod_{p<P(a)} \left(1 - \frac{1}{p}\right).$$

Moreover the sets S_a, as a varies over a primitive set $\mathcal{A}$, are pairwise disjoint. So

$$\sum_{a\in\mathcal{A}} \frac{1}{a} \prod_{p<P(a)} \left(1 - \frac{1}{p}\right) = \sum_{a\in\mathcal{A}} \delta(S_a) = \delta\left(\bigcup_{a\in\mathcal{A}} S_a\right) \le 1.$$

But

$$\frac{1}{a} \prod_{p<P(a)} \left(1 - \frac{1}{p}\right) \gg \frac{1}{a \log P(a)} \ge \frac{1}{a \log a},$$

so that $f(\mathcal{A}) \ll 1$.

To do a little better, we should be more careful with the step where we say $\prod_{p<P(a)}(1-1/p) \gg 1/\log(P(a)) \geq 1/\log a$.

Note that as $x \to \infty$, we have $\prod_{p<x}(1-1/p) \sim 1/(e^\gamma \log x)$. Also, if a is composite, then $a \geq 2P(a)$.

Lemma. *For $x \geq 2$, we have $1/\log(2x) < e^\gamma \prod_{p<x}(1-1/p)$.*

Conclude: If a is composite, then $1/\log a < e^\gamma \prod_{p<P(a)}(1-1/p)$. So, if every member of $\mathcal{A}$ is composite, then $f(\mathcal{A}) < e^\gamma$, since

$$f(\mathcal{A}) = \sum_{a\in\mathcal{A}} \frac{1}{a\log a} < e^\gamma \sum_{a\in\mathcal{A}} \frac{1}{a} \prod_{p<P(a)} \left(1-\frac{1}{p}\right) = e^\gamma \sum_{a\in\mathcal{A}} \delta(S_a) \leq e^\gamma.$$

With a non-strict inequality this much was proved by **Clark** (1995). But we show $f(\mathcal{A}) < e^\gamma$ for *all* primitive sets $\mathcal{A}$.

One path we followed was suggested by the work of Erdős and Zhang: partition $\mathcal{A}$ by the least prime factor of the elements: Let $\mathcal{A}(q) = \{a \in \mathcal{A} : p(a) = q\}$. We claim that

$$f(\mathcal{A}(q)) < \frac{e^\gamma}{q} \prod_{p<q} \left(1 - \frac{1}{p}\right) \text{ if } q \notin \mathcal{A}(q),$$
$$f(\mathcal{A}(q)) = \frac{1}{q \log q} \text{ if } q \in \mathcal{A}(q).$$

The second assertion is obvious since $q \in \mathcal{A}(q)$ implies $\mathcal{A}(q) = \{q\}$. For the first assertion, note that

$$\bigcup_{a \in \mathcal{A}(q)} S_a \subset \{ba : a \in \mathcal{A}(q), p(b) \geq q\} \subset \{n : p(n) = q\},$$

so that

$$\sum_{a \in \mathcal{A}(q)} \delta(S_a) \leq \frac{1}{q} \prod_{p<q} \left(1 - \frac{1}{p}\right).$$

But in this case every $a \in \mathcal{A}(q)$ is composite, and we've seen then that

$$\frac{1}{a \log a} < e^{\gamma} \delta(S_a),$$

so the claim above holds, i.e., if $q \notin \mathcal{A}(q)$,

$$f(\mathcal{A}(q)) = \sum_{a \in \mathcal{A}(q)} \frac{1}{a \log a} < e^{\gamma} \sum_{a \in \mathcal{A}(q)} \delta(S_a) \le \frac{e^{\gamma}}{q} \prod_{p<q} \left(1 - \frac{1}{p}\right).$$

Say a prime q is **Mertens** if

$$e^{\gamma} \prod_{p<q} \left(1 - \frac{1}{p}\right) \le \frac{1}{\log q}.$$

So, if q is Mertens, then $f(\mathcal{A}(q)) \le 1/(q \log q)$ regardless if $q \in \mathcal{A}(q)$ or $q \notin \mathcal{A}(q)$. Thus, if every prime in $\mathcal{P}(\mathcal{A})$ is Mertens, then

$$f(\mathcal{A}) = \sum_{q \in \mathcal{P}(\mathcal{A})} f(\mathcal{A}(q)) \le \sum_{q \in \mathcal{P}(\mathcal{A})} \frac{1}{q \log q} = f(\mathcal{P}(\mathcal{A})).$$

A prime q is **Mertens** if $e^\gamma \prod_{p<q}(1-1/p) \le 1/\log q$. And, if every prime in $\mathcal{P}(\mathcal{A})$ is Mertens, then $f(\mathcal{A}) \le f(\mathcal{P}(\mathcal{A}))$. That is, the Erdős conjecture is true for sets supported on the Mertens primes.

Let's try it out. Is 2 Mertens?

$$\prod_{p<2}\left(1-\frac{1}{p}\right) = 1, \qquad \frac{1}{e^\gamma \log 2} = 0.81001\ldots.$$

So, 2 is not Mertens. ☹

A prime q is **Mertens** if $e^\gamma \prod_{p<q}(1-1/p) \le 1/\log q$. And, if every prime in $\mathcal{P}(\mathcal{A})$ is Mertens, then $f(\mathcal{A}) \le f(\mathcal{P}(\mathcal{A}))$. That is, the Erdős conjecture is true for sets supported on the Mertens primes.

Let's try it out. Is 2 Mertens?

$$\prod_{p<2}\left(1-\frac{1}{p}\right) = 1, \qquad \frac{1}{e^\gamma \log 2} = 0.81001\ldots.$$

So, 2 is not Mertens. ☹

But every odd prime up to p_{10^8} is indeed Mertens. ☺

Theorem (Lamzouri, 2016). *Assuming RH and LI, the set of real numbers x with $e^\gamma \prod_{p\le x}(1-1/p) < 1/\log x$ has logarithmic density* $0.99999973\ldots$.

Corollary (Lichtman, Martin, P, 2018). *Assuming RH and LI, the set of Mertens primes has relative logarithmic density* 0.99999973... .

Note that 0.99999973... is the exact same log density that **Rubinstein & Sarnak** found in their famous 1994 paper for the set of x where $\mathrm{li}(x) > \pi(x)$, on assumption of RH and LI, though the two sets are not the same. (For technical reasons, the density of the Mertens race and the density of the π/li race turn out to be the same number.)

When we started investigating primitive sets we had no idea that we would find a connection to "Chebyshev's bias" and "prime number races".

We could push the verification of Mertens primes beyond the 10^8th prime, but instead, we found an alternative that allows us to push things *much* higher. As mentioned earlier:

Theorem (Lichtman, P, 2018). *The Erdős conjecture holds for primitive sets* $\mathcal{A}$ *supported on the odd primes up to* e^{10^6}.

For details and our other results, see our papers posted on arXiv or on my home page:

J. D. Lichtman and C. Pomerance, *The Erdős conjecture for primitive sets*, 2018.

J. D. Lichtman, G. Martin, and C. Pomerance, *Primes in prime number races*, 2018.

Thank you