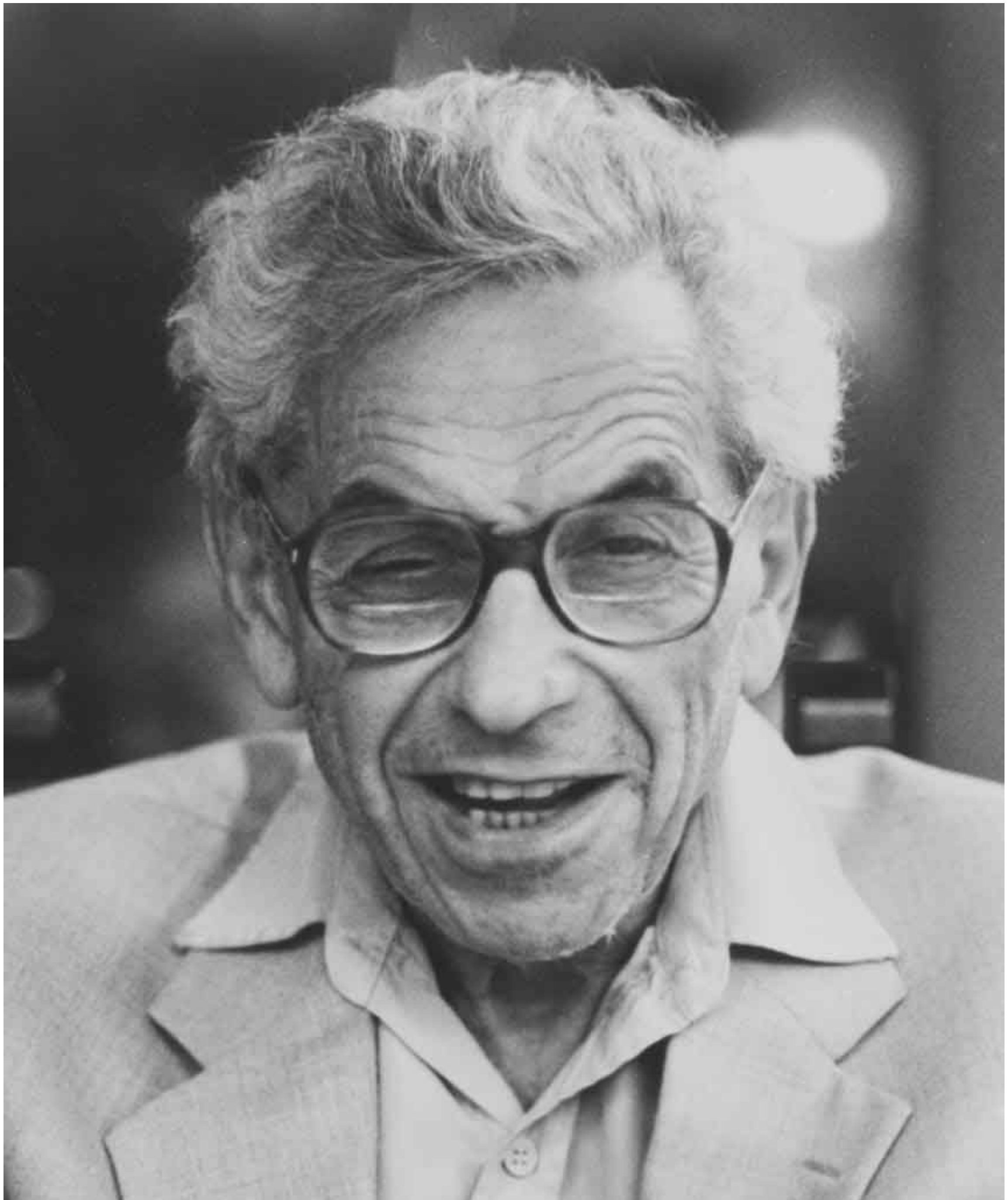
A 1935 Erdős paper on prime numbers and Euler's function

Carl Pomerance, Dartmouth College

with

Florian Luca, UNAM, Morelia



ON THE NORMAL NUMBER OF PRIME FACTORS OF $p-1$ AND SOME RELATED PROBLEMS CONCERNING EULER'S ϕ-FUNCTION

By PAUL ERDŐS (*Manchester*)

[Received 13 November 1934]

THIS paper is concerned with some problems considered by Hardy and Ramanujan, Titchmarsh, and Pillai. Suppose we are given a set M of positive integers m. Let $N(n)$ denote the number of m in the interval $(0, n)$. By saying that the normal number of prime factors of a number m is $B(n)$, we mean that, as $n \to \infty$, there are only $o[N(n)]$ of the m $(\leq n)$ for which the number of prime factors does

Hardy & Ramanujan, **1917**: *The normal number of prime divisors of n is* $\log\log n$.

That is, for each fixed $\epsilon > 0$, the set of n with

$$|\omega(n) - \log\log n| > \epsilon \log\log n$$

has asymptotic density 0. Here, $\omega(n)$ is the number of prime divisors of n. The same is true for $\Omega(n) - \log\log n$, where $\Omega(n)$ is the number of prime power divisors of n.

Turán, **1934**: A beautiful "probabilistic" proof of the Hardy–Ramanujan theorem.

Erdős, **1935**: *The normal number of prime divisors of* $p-1$*, where* p *is prime, is* $\log\log p$*.*

Erdős could not adapt the slick Turán proof; rather he used the older Hardy–Ramanujan proof together with Brun's (sieve) method.

As an application:

$$\sum_{p\le x} \tau(p-1) \ge x/(\log x)^{1-\log 2+o(1)},$$

where τ is the divisor function. Titchmarsh, in 1930, had exponent $1/2$ in the denominator.

That was a straightforward application. Next came a typically Erdős application. *What can one say about the range of Euler's function φ?* If $V(x)$ denotes the number of Euler values in $[1, x]$, then since φ is 1-to-1 on the primes, we have $V(x) \geq \pi(x) \sim x/\log x$.

Pillai, **1929**: $V(x) \ll x/(\log x)^{(\log 2)/e}$.

As the principal application of the normal order of $\omega(p-1)$:

Erdős, **1935**: $V(x) = x/(\log x)^{1+o(1)}$.

Using either the result of Pillai or Erdős one has that there are values of φ with arbitrarily many preimages. In particular, there is some $c > 0$ such that for all large x, below x there is a number with more than $(\log x)^c$ preimages.

Thus, the following seems completely unexpected!

Erdős, **1935**: *There is some $c > 0$ such that for all large x, below x there is a number with more than x^c preimages under φ.*

What have we learned since 1935?

One of the first applications of the Bombieri–Vinogradov inequality was a proof that

$$\sum_{p \le x} \tau(p-1) \sim Cx,$$

for a certain positive constant C, which thus solved the Titchmarsh divisor problem. (Solved earlier by Linnik using his "dispersion method".)

I believe we still don't know the aysmptotic order of $\sum_{p \le x} \tau_3(p-1)$, where $\tau_3(n)$ is the number of ordered factorizations of n into 3 factors.

Concerning $\omega(p-1)$, we know after Barban, Vinogradov, & Levin that we have an Erdős–Kac-type theorem. Namely the relative density of those primes p with

$$\omega(p-1) \le \log\log p + u(\log\log p)^{1/2}$$

is $G(u)$ (the Gaussian distribution).

For $V(x)$, the number of Euler values in $[1, x]$, we now know after papers of Erdős & Hall, Maier & Pomerance, and Ford, the true order of magnitude of $V(x)$. It is

$$\frac{x}{\log x} \exp\left(c_1(\log_3 x - \log_4 x)^2 + c_2 \log_3 x + c_3 \log_4 x\right)$$

for certain explicit constants c_1, c_2, c_3. We still do not have an asymptotic formula for $V(x)$, nor do we know that the number of Euler values in $[1, x]$ is asymptotically equal to the number of them in $[x, 2x]$.

For popular values, after work of Wooldridge, Pomerance, Fouvry & Grupp, Balog, Friedlander, Baker & Harman, we now know that there are numbers below x with more than $x^{0.7067}$ Euler preimages.

This problem is connected to the distribution of Carmichael numbers in that improvements in the popular-value result are likely to lead to improvements in the lower bound in the distribution of Carmichael numbers.

Sketch of the Erdős proof on the range of φ:

• If $\varphi(n) \le x$, then $n \le X := cx \log\log x$.

• For K large enough, we may assume $\omega(n) \ge (1/K)\log\log x$.

• Primes p with $\omega(p-1) \le 40K$ are rare, so n may be assumed to be divisible by at least $(1/(2K))\log\log x$ primes q with $\omega(q-1) > 40K$.

• Thus, but for $O(x/(\log x)^{1-\epsilon})$ values $\varphi(n) \le x$, we have $\Omega(\varphi(n)) > 20\log\log x$. But there are very few such integers.

Let φ_k be the k-fold iterate of φ. What can one say about the range of φ_k?

The function φ_k was studied by Pillai: how many iterations to get to 1? For example, 31, 32, 33, 34, 35, 36, and 37 each take 5 iterations, but 38 takes only 4. Also studied by Shapiro and Erdős, Granville, Pomerance, & Spiro.

Using the Bateman–Horn conjecture, one can show that

$$V_k(x) \gg_k x/(\log x)^k$$

for each k, where $V_k(x)$ denotes the number of values of φ_k in $[1, x]$.

Indeed, consider primes p where $p - 1 = 2q$ with q prime, $q - 1 = 2r$, with r prime, etc.

Erdős & Hall, **1977**:

$$V_2(x) \ll \frac{x}{(\log x)^2} \exp\left(\frac{c \log_2 x \log_4 x}{\log_3 x}\right).$$

In addition they claimed they were able to prove that $V_2(x) \gg x/(\log x)^\alpha$ for some $\alpha > 2$.

Luca & Pomerance, **2009**:

$$\frac{x}{(\log x)^2} \ll V_2(x) \ll \frac{x}{(\log x)^2} \exp\left(37(\log_2 x \log_3 x)^{1/2}\right).$$

The upper bound generalizes:

$$V_k(x) \ll \frac{x}{(\log x)^k} \exp\left(13k^{3/2}(\log_2 x \log_3 x)^{1/2}\right)$$

uniformly for each positive integer k.

Sketch of proof:

The lower bound for $V_2(x)$ uses Chen's theorem and Brun's method.

For the upper bound on $V_k(x)$:

- Write $n = pm$ with p the largest prime factor of n, assume $n \le X := x(c \log\log x)^k$.

- We can assume $\Omega(\varphi_k(n)) \le 2.9k \log\log x$, so $\Omega(\varphi_k(p)) \le 2.9k \log\log x$ and $\Omega(\varphi_k(m)) \le 2.9k \log\log x$. Thus, $\Omega(\varphi(m)) \le 3k \log\log x$.

- Use the large sieve to get an upper bound for the number of such primes $p \le X/m$. Then use the old Erdős strategy to get an upper bound on $\sum 1/m$ over $m \le X$ with $\Omega(\varphi(m)) \le 3k \log\log x$.

The range of the Carmichael function λ

First discussed by Gauss, $\lambda(n)$ is the *exponent* of the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^{\times}$. It is the smallest positive integer such that $a^{\lambda(n)} \equiv 1 \pmod{n}$ for all a coprime to n.

So, for coprime m, n, $\lambda(mn) = \mathrm{lcm}[\lambda(m), \lambda(n)]$. And $\lambda(p^j) = \varphi(p^j)$ for prime powers p^j, except when $p = 2, j \geq 3$ in which case $\lambda(2^j) = 2^{j-2}$.

Being so similar to φ, one might expect similar results about the range of λ. But one big headache appears: while $\varphi(n) \leq x$ implies $n \ll x \log\log x$ (in fact, the number of n with $\varphi(n) \leq x$ is $\sim cx$, a result of Bateman), there can be extraordinarily huge numbers n with $\lambda(n) \leq x$.

How huge? Try $\exp(x^{c/\log\log x})$.

In addition, there are $\gg x^{(\log x)^2}$ numbers n with $\lambda(n) \le x$.

So, with so many chances to hit numbers in $[1, x]$, it is not even clear that $V_\lambda(x)$, the number of λ-values in $[1, x]$, is $o(x)$.

But it is true; it follows from a lemma in

Erdős & Wagstaff, **1980**: *Let d_B be the upper density of those numbers n divisible by some $p-1$ with $p > B$ prime. Then $\lim_{B\to\infty} d_B = 0$.*

Note that if $\lambda(n)$ is not divisible by any $p-1$ with $p > B$ prime, then n is not divisible by any prime $p > B$, so $\lambda(n)$ has no prime factors $> B$. So, there are few such integers.

This argument was first outlined by Erdős, Pomerance, & Schmutz (1991) claiming that $V_\lambda(x) \ll x/(\log x)^c$ for some $c > 0$.

Friedlander & Luca, (2007):

$$V_\lambda(x) \le x/(\log x)^{1-(e/2)\log 2+o(1)},$$

where $1-(e/2)\log 2 = 0.05791\ldots$.

Luca & Pomerance, (2009?):

$$V_\lambda(x) \le x/(\log x)^{1-(1+\log\log 2)/\log 2+o(1)},$$

where $1-(1+\log\log 2)/\log 2 = 0.08607\ldots$.

Towards a lower bound:

Clearly $V_\lambda(x) \geq \pi(x)$ (since as with φ, λ is 1-to-1 on the primes), so $V_\lambda(x) \gg x/\log x$.

Can we do better?

Banks, Friedlander, Luca, Pappalardi, & Shparlinski, (2006):

$$V_\lambda(x) \gg \frac{x}{\log x} \exp\left(c(\log\log\log x)^2\right).$$

So, what is your instinct? Is 1 the "correct" exponent on $\log x$, or is it some number smaller than 1?

Luca & Pomerance, (2009?):

$$V_\lambda(x) \gg \frac{x}{(\log x)^{3/5}}.$$

In fact, we show this for numbers $\lambda(n)$, where $n = pq$ with p, q primes. This seems counter-intuitive, since the number of integers $n = pq \le x$ is $\sim x \log\log x / \log x$. But recall, it is not $n \le x$ that we need, but $\lambda(n) \le x$.

So, let $R(x)$ be the number of triples a, b, d where $abd \le x$, $\gcd(a, b) = 1$, and $p = ad + 1, q = bd + 1$ are both prime. Then $\lambda(pq) = abd$, and except for possible overcounting, we have $V_\lambda(x) \ge R(x)$. But overcounting needs to be considered!

Assume $p \le \exp((\log x)^c)$ with c chosen appropriately. Say $\lambda(pq) = \lambda(p'q')$. If $R_1(x)$ is the number of times this happens with $q = q'$ and $R_2(x)$ is the number of times this happens with $q \ne q'$, then by Cauchy–Schwarz,

$$V_\lambda(x) \ge R(x)^2/(R_1(x) + R_2(x)).$$

In getting upper bounds for $R_1(x), R_2(x)$ we assume further that parameters a, b, d have close to a set number of prime divisors, where the settings are chosen to optimize the final estimate.