ON THE RANGE OF THE ITERATED EULER FUNCTION

Florian Luca

Instituto de Matemáticas, Universidad Nacional Autónoma de México, C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

Carl Pomerance

Department of Mathematics, Dartmouth College, Hanover, NH 03755-3551, USA
carl.pomerance@dartmouth.edu

Received: , Accepted: , Published:

Abstract

For a positive integer $k$ let $\phi_k$ be the $k$-fold composition of the Euler function $\phi$. In this paper, we study the size of the set $\{\phi_k(n) \leq x\}$ as $x$ tends to infinity.

1 Introduction

Let $\phi$ be Euler’s function. For a positive integer $k$, let $\phi_k$ be the $k$-fold composition of $\phi$. In this paper, we study the range $\mathcal{V}_k$ of $\phi_k$. For a positive real number $x$ we put

$$\mathcal{V}_k(x) = \{\phi_k(n) \leq x\}.$$

In 1935, Erdős [7] showed that $\#\mathcal{V}_1(x) = x/(\log x)^{1+o(1)}$. (Stronger estimates are known for $\#\mathcal{V}_1(x)$, see [10], [17].) In 1977, Erdős and Hall [8] considered the more general problem of estimating $\#\mathcal{V}_k(x)$, suggesting that it is $x/(\log x)^{k+o(1)}$ for each fixed integer $k \geq 1$. They were able to prove that

$$\#\mathcal{V}_2(x) \lesssim \frac{x}{(\log x)^2 + o(1)},$$

and in fact, they were able to establish a somewhat more explicit form for this inequality. Our first result is the following general upper bound on $\#\mathcal{V}_k(x)$ which is uniform in $k$.

Theorem 1. The estimate

$$\#\mathcal{V}_k(x) \leq \frac{x}{(\log x)^k} \exp\left(13k^{3/2}(\log \log x \log \log \log x)^{1/2}\right)$$  \hspace{1cm} (1)
holds uniformly in $k \geq 1$ once $x$ is sufficiently large.

As a corollary we have, when $x \to \infty$,
$$\#V_k(x) \leq \frac{x}{(\log x)^{k+o(1)}}$$
when $k = o((\log \log x/\log \log \log x)^{1/3})$, and
$$\#V_k(x) \leq \frac{x}{(\log x)^{(1+o(1))k}}$$
when $k = o(\log \log x/\log \log \log x)$. Note that (1) is somewhat stronger than the explicit upper bound in [8] for the case $k = 2$.

Let $k \geq 1$ be fixed. Let $m > 2$ be such that $m, 2m + 1, \cdots, 2^{k-1}m + 2^{k-1} - 1$ are all prime numbers. Then $\phi_k(2^{k-1}m + 2^{k-1} - 1) = m - 1$. The quantitative version of the Prime $k$-tuples Conjecture of Bateman and Horn [2] implies that the number of such values $m \leq x$ should be $\geq c_k x/(\log x)^k$ for $x$ sufficiently large, where $c_k > 0$ is a constant depending on $k$. Thus, we see that up to the factor of size $(\log x)^{o(1)}$ appearing on the right hand side of estimate (1), it is likely that $\#V_k(x) = x/(\log x)^{k+o(1)}$ holds when $k$ is fixed as $x \to \infty$, thus verifying the surmise of Erdős and Hall.

Next, we prove a lower bound on $\#V_2(x)$ comparable to the one predicted by the above heuristic construction.

**Theorem 2.** There exists an absolute constant $c_2 > 0$ such that the inequality
$$\#V_2(x) \geq c_2 \frac{x}{(\log x)^2}$$
holds for all $x \geq 2$.

In [8], Erdős and Hall assert that they were able to prove such a lower bound with the exponent 2 replaced by any larger real number.

In the last section we study the integers that are in every $V_k$ and we also discuss analogous problems for Carmichael’s universal exponent function $\lambda(n)$.

In what follows, we use the Vinogradov symbols $\gg$ and $\ll$ and the Landau symbols $O$ and $o$ with their usual meaning. The constants and convergence implied by them might depend on some other parameters such as $k$, $K$, $\varepsilon$, etc. We use $p$ and $q$ with or without subscripts for prime numbers. We use $\omega(n)$ for the number of distinct prime factors of $n$, $\Omega(n)$ for the number of prime power divisors ($> 1$) of $n$, $p(n)$ and $P(n)$ for the smallest and largest prime divisors of $n$, respectively, and $v_2(n)$ for the exponent of 2 in the factorization of $n$. We write $\log_1 x = \max\{1, \log x\}$, and for $k \geq 2$ we put $\log_k x$ for the $k$-fold iterate of the function $\log_1$ evaluated at $x$. For a subset $\mathcal{A}$ of positive integers and a positive real number $x$ we write $\mathcal{A}(x)$ for the set $\mathcal{A} \cap [1, x]$. 
2 The proof of Theorem 1

Let \( x \) be large. By a result of Pillai [18], we may assume that \( k \leq \log x / \log 2 \), since otherwise \( V_k(x) = \{1\} \). Furthermore, we may in fact assume that \( k \leq 10^{-2} \log_2 x / \log_3 x \), since otherwise the upper bound on \#\( V_k(x) \) appearing in estimate (1) exceeds \( x \). We may also assume that \( n \geq x/(\log x)^k \), since otherwise there are at most \( x/(\log x)^k \) possibilities for \( n \), and, in particular, at most \( x/(\log x)^k \) possibilities for \( \phi_k(n) \) also.

By the minimal order of the Euler function, there exists a constant \( c_0 > 0 \) such that the inequality \( \phi(m)/m \geq c_0 m / \log \log m \) holds for all \( m \geq 3 \). From this it is easy to prove by induction on \( k \) that if \( x \) is sufficiently large and \( \phi_k(n) \leq n \), then \( n \leq x(2c_0 \log_2 x)^k \) for all \( k \) in our stated range. Let \( X := x(\log_2 x)^{2k} \), so that for large \( x \), we may assume that \( n \leq X \).

Let \( y = x^{1/(\log_2 x)^2} \) and write \( n = pm \), where \( p = P(n) \). By familiar estimates (see, for example, [3]), the number of \( n \leq X \) such that \( p \leq y \) is at most, for large \( x \),

\[
\frac{X}{(\log x)^{\log_2 x}} = \frac{x(\log_2 x)^{2k}}{(\log x)^{\log_2 x}} \leq \frac{x}{(\log x)^k},
\]

so we need only deal with the case \( p > y \). Assume that \( \Omega(\phi_k(n)) \geq 2.9k \log_2 x \). Lemma 13 in [15] shows that the number of such possibilities for \( \phi_k(n) \leq x \) is

\[
\ll \frac{kx \log x \log_2 x}{2^{2.9k \log_2 x}} \leq \frac{x(\log_2 x)^2}{(\log x)^{2.9k \log 2 - 1}} \ll \frac{x}{(\log x)^k}
\]

for all \( k \) in our range. It follows that we may assume that

\[
\Omega(\phi_k(n)) \leq 2.9k \log_2 x.
\]

It is easy to see that \( \Omega(\phi(a)) \geq \Omega(a) - 1 \) for every natural number \( a \). Thus, since \( \phi_k(m) \mid \phi_k(n) \), we have

\[
\Omega(\phi(m)) \leq 2.9k \log_2 x + k - 1 \leq 3k \log_2 x \tag{2}
\]

for all \( x \) sufficiently large.

Since also \( \phi_k(p) \mid \phi_k(n) \), we may assume that

\[
\Omega(\phi_k(p)) \leq 2.9k \log_2 x.
\]

Since \( p > y \), we have \( \log_2 p > \log_2 x - 2 \log_3 x \), so that \( \Omega(\phi_k(p)) \leq 3k \log_2 p \) for \( x \) large. Since \( p \leq X/m \), we thus have, in the notation of Lemma 4 below, that \( p \in \mathcal{A}_{k,3k}(X/m) \), and that result shows that the number of such possibilities is at most

\[
\#\mathcal{A}_{k,3k}(X/m) \leq \frac{X}{m(\log(X/m))^{k}} \exp \left(3k(6k \log_2 X \log_3 X)^{1/2} + 3k^2 \log_3 X \right).
\]

Observe further that with our bound on \( k \),

\[
3k(6k \log_2 X \log_3 X)^{1/2} + 3k^2 \log_3 X
= k^{3/2}(\log_3 X) \left(3(6 \log_2 X / \log_3 X)^{1/2} + 3k^{1/2} \right)
\leq k^{3/2}(\log_2 X \log_3 X)^{1/2}(3\sqrt{6} + 3/10).
\]
Since $3\sqrt{6} + 3/10 < 7.7$, it thus follows that if we put

$$U(x) = \exp(7.7k^{3/2}(\log_2 x \log_3 x)^{1/2}),$$

then for large $x$,

$$\#A_{k,3k}(X/m) \leq \frac{xU(x)(\log_2 x)^{2k}}{m(\log y)^k} \leq \frac{xU(x)(\log_2 x)^{4k}}{m(\log x)^k}$$

uniformly in $m$ and $k$. Thus, the number of such possibilities for $n \leq X$ is

$$\leq \frac{xU(x)(\log_2 x)^{4k}}{(\log x)^k} \sum_{m \in M} \frac{1}{m},$$

where $M$ is the set of all possible values of $m$. Such $m$ satisfy, in particular, the inequality (2). Lemma 3 below shows that if $x$ is sufficiently large then

$$\sum_{m \in M} \frac{1}{m} \leq \exp \left(2.9(3k \log_2 X \log_3 X)^{1/2} \right),$$

which together with the fact that $2.9\sqrt{3} < 5.1$ and the previous estimate shows that the count on the set of our $n \leq X$ is

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

for large values of $x$. We thus finish the proof of Theorem 1 and it remains to prove Lemmas 3 and 4.

**Lemma 3.** Let $x$ be large, $K$ be any positive integer and let $\mathcal{N}(K, x)$ denote the set of natural numbers $n \leq x$ with $\Omega(\phi(n)) \leq K \log_2 x$. Then

$$\sum_{n \in \mathcal{N}(K,x)} \frac{1}{n} \leq \exp(2.9(K \log_2 x \log_3 x)^{1/2})$$

holds for large values of $x$ uniformly in $K$.

**Proof.** We assume that $K \leq \log_2 x / \log_3 x$ since otherwise the right hand side above exceeds $(\log x)^{2.9}$, while the left hand side is at most $\log x + O(1)$, so the desired inequality holds anyway.

Let $z$ be a parameter that we will choose shortly. For each integer $n \leq x$ write $n = n_0n_1$, where each prime $q | n_0$ has $\Omega(q - 1) < \log z$ and each prime $q | n_1$ has $\Omega(q - 1) \geq \log z$. For $n \in \mathcal{N}(K, x)$ we have that $\Omega(n_1) \leq K \log_2 x / \log z$. Let $\mathcal{N}_0(x)$ denote the set of numbers $n_0 \leq x$ divisible only by primes $q$ with $\Omega(q - 1) < \log z$ and let $\mathcal{N}_1(x)$ denote the set of numbers $n_1 \leq x$ with $\Omega(n_1) \leq K \log_2 x / \log z$. We thus have

$$\sum_{n \in \mathcal{N}(K,x)} \frac{1}{n} \leq \left( \sum_{n_0 \in \mathcal{N}_0(x)} \frac{1}{n_0} \right) \left( \sum_{n_1 \in \mathcal{N}_1(x)} \frac{1}{n_1} \right).$$  \hfill (3)
Note that
\[
\sum_{n_0 \in \mathcal{N}_0(x)} \frac{1}{n_0} \leq \sum_{j=0}^{\infty} \frac{1}{j!} \left( \sum_{q \leq x} \frac{1}{q} + \frac{1}{q^2} + \cdots \right)^j = \exp \left( \sum_{q \leq x} \frac{1}{q-1} \Omega(q-1) \log z \right).
\]

It follows from Erdős [7] that there is some \( c > 0 \) such that the number of primes \( q \leq t \) with \( \omega(q-1) \leq \frac{1}{2} \log_2 q \) is \( O(t/(\log t)^{1+c}) \). Since \( \omega(q-1) \leq \Omega(q-1) \), the same \( O \)-estimate holds for the distribution of primes \( q \) with \( \Omega(q-1) \leq \frac{1}{2} \log_2 q \). In particular the sum of their reciprocals is convergent, so that
\[
\sum_{e^{\alpha^2} < q \leq x} \frac{1}{q-1} \leq \sum_{e^{\alpha^2} < q} \frac{1}{q-1} \ll 1.
\]

Thus,\[
\sum_{q \leq x} \frac{1}{q-1} \leq \sum_{q \leq e^{\alpha^2}} \frac{1}{q-1} + \sum_{e^{\alpha^2} < q \leq x} \frac{1}{q-1} \leq 2 \log z + O(1),
\]
and so
\[
\sum_{n_0 \in \mathcal{N}_0(x)} \frac{1}{n_0} \ll z^2. \tag{4}
\]

For the sum over \( \mathcal{N}_1(x) \), we have
\[
\sum_{n_1 \in \mathcal{N}_1(x)} \frac{1}{n_1} \leq \sum_{j \leq K \log_2 x / \log z} \frac{1}{j!} \left( \sum_{q \leq x} \frac{1}{q-1} \right)^j \leq \sum_{j \leq K \log_2 x / \log z} \frac{1}{j!} \left( \log_2 x + O(1) \right)^j.
\]

We choose \( z = \exp((\frac{1}{2} K \log_2 x \log_3 x)^{1/2}) \). Observe that the inequalities
\[
K \log_2 x / \log z = (2K \log_2 x / \log_3 x)^{1/2} < 2^{1/2} \log_2 x / \log_3 x < \log_2 x
\]
hold for large values of \( x \). Thus,
\[
\sum_{n_1 \in \mathcal{N}_1(x)} \frac{1}{n_1} \ll (2 \log_2 x)^{K \log_2 x / \log z}. \tag{5}
\]
Putting (4) and (5) into (3) and using the fact that \(2\sqrt{2} < 2.9\), we have
\[
\sum_{n \in \mathbb{N}} \frac{1}{n} \leq \exp(2.9(K \log_2 x \log_3 x)^{1/2})
\]
for all sufficiently large \(x\). This proves the lemma.

**Remark 1.** The above proof uses ideas from Erdős [7] and is also similar to Lemma 4 in Luca [14].

**Lemma 4.** Let \(k, K\) be positive integers not exceeding \(\frac{1}{2} \log_2 x\). Put
\[
\mathcal{A}_{k,K} = \{ p : \Omega(\phi_k(p)) \leq K \log_2 p \}.
\]
We have
\[
\#\mathcal{A}_{k,K}(x) \leq \frac{x}{(\log x)^k} \exp \left( 3k(2K \log_2 x \log_3 x)^{1/2} + 3k^2 \log_3 x \right)
\]
for all sufficiently large values of \(x\), independent of the choices of \(k, K\).

**Proof.** When \(k = 1\), this trivially follows from the Prime Number Theorem. We assume that \(k > 1\). We let \(p \in \mathcal{A}_{k,K}(x)\) and assume that \(p \geq \frac{x}{(\log x)^k}\) because there are only \(\pi(x/(\log x)^k) \leq x/(\log x)^k\) primes \(p\) failing this condition. Let \(p_0 = p\) and write
\[
p_0 - 1 = p_1 m_1; \\
p_1 - 1 = p_2 m_2; \\
\vdots \\
p_{k-2} - 1 = p_{k-1} m_{k-1},
\]
where \(p_i = P(p_{i-1} - 1)\) for all \(i = 1, \ldots, k-1\). Since \(\Omega(\phi(n)) \geq \Omega(n) - 1\), we have that
\[
\Omega(p_{i-1} - 1) \leq \Omega(\phi_i(p)) \leq \Omega(\phi_k(p)) + k \leq 2K \log_2 x
\]
for all \(i = 1, 2, \ldots, k-1\) if \(x\) is sufficiently large. In particular
\[
p_i \geq p_{i-1}^{1/(2K \log_2 x)} \geq p_{i-1}^{1/(\log_2 x)^2},
\]
so that for \(x\) sufficiently large we have
\[
p_i \geq p_0^{1/(\log_2 x)^{2i}} \geq y_i := \frac{1}{2} x^{1/(\log_2 x)^{2i}}
\]
for \(i = 1, 2, \ldots, k-1\).

Consider the \(k\) linear functions \(L_j(x) = A_j x + B_j\) for \(j = k, k-1, \ldots, 1\) given by \(L_k(x) = x\) and
\[
L_{k-1}(x) = m_{k-1} x + 1 \\
L_{k-2}(x) = m_{k-2} m_{k-1} x + m_{k-2} + 1 \\
\vdots \\
L_1(x) = m_1 \cdots m_{k-1} x + (m_1 \cdots m_{k-2} + m_1 \cdots m_{k-3} + \cdots + m_1 + 1).
\]
Note that \( p_{k-1} \leq x/(m_1 \cdots m_{k-1}) \) is such that \( L_j(p_{k-1}) \) is a prime for all \( j = 1, \ldots, k \). If some \((A_i, B_i) > 1\), then there is at most one prime \( p_{k-1} \) for which all of \( L_j(p_{k-1}) \) are prime. Further, since \( 0 = B_k < B_{k-1} < \cdots < B_1 \), it follows that if some \( A_jB_i = A_iB_j \) for some \( 0 \leq j < i \leq k - 1 \), then \( 1 < A_j/A_i \mid B_i \) so that \((A_i, B_i) > 1\). Thus, we may assume that each \( A_jB_i - A_iB_j \neq 0 \). The following result allows us to use something like a traditional sieve upper bound for prime \( k \)-tuples, where it is not assumed that \( k \) is bounded. Note that a stronger form of this lemma will appear in [11].

**Lemma 5.** Let \( L_i(n) = A_in + B_i \) be linear functions for \( i = 1, \ldots, k \) with integer coefficients such that each \( A_i > 0 \), each \((A_i, B_i) = 1\), and

\[
E := A_1 \cdots A_k \prod_{1 \leq j < i \leq k} (A_jB_i - A_iB_j)
\]

is nonzero. Put \( F(n) = \prod_{i=1}^{k} L_i(n) \) and for each \( p \) let \( \rho(p) \) be the number of congruence classes \( n \mod p \) such that \( F(n) \equiv 0 \pmod{p} \). Assume that for each \( p \), we have \( \rho(p) < p \). If \( N \geq 2 \) and \( k \leq \log N / (10 \log_2 N)^2 \), then the number of \( n \leq N \) such that each \( L_i(n) \) is prime is at most

\[
(ck \log_2 k)^k \left( \frac{\Delta}{\phi(\Delta)} \right)^k \frac{N(\log_2 N)^k}{(\log N)^k},
\]

where \( c \) is an absolute constant and \( \Delta \) is the product of the distinct primes \( p \mid E \) with \( p > k \).

**Proof.** We may assume that \( N \) is large since the constant \( c \) may be adjusted for smaller values. Let \( Z \) denote the number of \( n \leq N \) with each \( L_i(n) \) prime. We first show

\[
Z \leq N \prod_{k < p \leq N^{1/(100k \log_2 N)}} \left( 1 - \frac{\rho(p)}{p} \right) + O \left( \frac{N}{(\log N)^{10k}} \right), \tag{6}
\]

For the proof, let \( \rho(m) \) be the number of solutions \( n \mod m \) of the congruence \( F(n) \equiv 0 \pmod{m} \). Clearly, \( \rho \) is a multiplicative function. Put \( N_1 = N^{1/(100k \log_2 N)} \). Noting that \( \rho(p) \leq k \), it follows that \( \rho(d) \leq k^{\omega(d)} \) holds for all squarefree positive integers \( d \). Taking \( M \) to be the first even integer exceeding \( 10k \log_2 N \), we get, by the Principle of Inclusion and Exclusion and the Bonferroni upper-bound inequality, that

\[
Z \leq N^{1/2} + \sum_{k < p(d) \leq P(d) \leq N_1} \left( \frac{N\mu(d)\rho(d)}{d} + O(k^{\omega(d)}) \right)
\]

\[
\leq N \prod_{k < p \leq N_1} \left( 1 - \frac{\rho(p)}{p} \right)
\]

\[
+ O \left( \frac{N^{1/2} + \sum_{d : P(d) \leq N_1, \omega(d) \leq M} k^{\omega(d)} + N \sum_{d : \mu(d) \neq 0, P(d) \leq N_1, \omega(d) > M} \frac{k^{\omega(d)}}{d}}{d} \right).
\]
It remains to look at the $O$-terms. For the first sum, we have that
\[ k^{\omega(d)} \leq k^{10k \log_2 N + 2} = \exp((10k \log_2 N + 2) \log k) < N^{1/9} \]
for all large values of $N$ uniformly in our range for $k$. The number of possibilities for $d$ is $\leq N_1^{M} \leq N^{(10k \log_2 N + 2)/(100k \log_2 N)} < N^{1/9}$ for large values of $N$. Hence, the first sum is $< N^{2/9}$. The second one is
\[
\leq N \sum_{j > M} \left( \frac{e k \log_2 N + O(k)}{j} \right)^j \leq N \sum_{j > M} \left( \frac{e}{j} \right)^j \leq N \leq \frac{N}{(\log N)^{10k}}
\]
for large values of $N$. Note that in our range for $k$, this last error estimate dominates the other two. Thus, we have (6).

To finish the proof of the lemma, we estimate the main term in (6). We have
\[
\log \left( \prod_{k < p \leq N_1} \left( 1 - \frac{\rho(p)}{p} \right) \right) \leq - \sum_{k < p \leq N_1} \frac{\rho(p)}{p} \leq - \sum_{k < p \leq N_1} \frac{k}{p} + \sum_{\Delta} \frac{k}{p}
\]
\[= -k \log_2 N_1 + k \log_2 k - k \sum_{\Delta} \log(1 - 1/p) + O(k).\]
Since the last sum above is $-\log(\Delta/\phi(\Delta))$ and $\log_2 N_1 = \log_2 N - \log_3 N - \log_1 k + O(1)$, the main term in (6) is at most
\[
(ck \log_1 k)^k \left( \frac{\Delta}{\phi(\Delta)} \right)^k \frac{N (\log_2 N)^k}{(\log N)^k}
\]
for some absolute constant $c$. Thus, by adjusting the constant $c$ if necessary, we have the lemma. \qed

We apply Lemma 5 to our system of linear functions with $N = x/(m_1 \ldots m_{k-1}) \geq y_{k-1}$. Thus, the number of choices for $p_{k-1} \leq N$ with each $L_i(p_{k-1})$ prime is at most
\[
\frac{x (\log \log x)^k}{m_1 \ldots m_{k-1} (\log y_{k-1})^k} \left( \frac{\Delta}{\phi(\Delta)} \right)^k (k \log k)^k.
\]
We need an estimate for $\Delta/\phi(\Delta)$. For this, note that each $A_j B_i$ in our setting is at most $x^2$, so that $\Delta \leq x^{O(k^2)}$, therefore by the minimal order of $\phi$, we have
\[
\Delta/\phi(\Delta) \ll \log_1 k + \log_2 x \ll \log_2 x. \tag{7}
\]
With our choice for $y_{k-1}$, our upper bound for $k$ in the lemma, and the estimate (7), our count for the number of choices for $p_{k-1}$ is now at most
\[
\frac{x}{m_1 \ldots m_{k-1} (\log x)^k} \exp(3k^2 \log_3 x),
\]
for \( x \) sufficiently large.

Observe that \( \Omega(\phi_{k-j}(m_j)) \leq K \log x \) holds for all \( j = 1, \ldots, k-1 \), so that \( \Omega(\phi(m_j)) \leq 2K \log \log x \) for each \( j = 1, \ldots, k-1 \) if \( x \) is sufficiently large. It then follows, by Lemma 3, that summing up over all possibilities for \( m_1, \ldots, m_{k-1} \) (positive integers \( m \leq x \) such that \( \Omega(\phi(m)) \leq 2K \log \log x \)), we have

\[
\#A_{k,K}(x) \leq \frac{x \exp(3k^2 \log x)}{(\log x)^k} \left( \sum_{1 \leq m \leq x} \frac{1}{\Omega(\phi(m)) \leq 2K \log \log x} \right)^{k-1} \\
\leq \frac{x}{(\log x)^k} \exp \left( 3k(2K \log_2 x \log_3 x)^{1/2} + 3k^2 \log_3 x \right)
\]

once \( x \) is large. This completes the proof of Lemma 4.

\[\square\]

3 The proof of Theorem 2

Here, we use the following theorem essentially due to Chen [5, 6].

**Lemma 6.** There exists \( x_0 \) such that if \( x > x_0 \) the interval \([x/2, x]\) contains \( \gg x/(\log x)^2 \) primes \( p \) such that \((p-1)/2\) is either prime or a product of two primes each of them exceeding \( x^{1/10} \).

Let

\[ C_1(x) = \{ p \in [x/2, x] : (p-1)/2 \text{ is prime} \} \]

and let

\[ C_2(x) = \{ p \in [x/2, x] : (p-1)/2 = q_1q_2, q_i > x^{1/10} \text{ is prime for } i = 1, 2 \} \]

We distinguish two cases.

**Case 1.** \( \#C_1(x) \geq \#C_2(x) \).

In this case, for large \( x \), \( \phi_2(p) = (p-3)/2 \) is injective when restricted to \( C_1(x) \). Hence,

\[ \#V_2(x) \geq \#C_1(x) \gg \frac{x}{(\log x)^2} \],

where the last inequality follows from Lemma 6.

**Case 2.** \( \#C_1(x) < \#C_2(x) \).

Let \( p \in C_2(x) \) and write \( p-1 = 2q_1q_2 \), where \( x^{1/10} < q_1 \leq q_2 \). Put \( y = \exp((\log x)^{4/5}) \). Let \( C_3(x) \) be the subset of \( C_2(x) \) such that \( q_1 > x^{1/2}/y \). Since \( q_1q_2 < x \), we get that
$q_2 < x/q_1 < x^{1/2}y$. We find an upper bound on $\#C_3(x)$. Let $q_1 \in [x^{1/2}/y, x^{1/2}]$ be a fixed prime. By Brun’s sieve, the number of primes $q_2 \leq x/q_1$ such that $2q_1q_2 + 1$ is a prime is

$$x \ll \frac{x}{\phi(q_1)(\log(x/q_1))^2} \ll \frac{x}{q_1(\log x)^2}.$$ 

Summing the above bound for all $q_1 \in [x^{1/2}/y, x^{1/2}]$, we get that

$$\#C_3(x) \ll \frac{x}{(\log x)^2} \sum_{x^{1/2}/y \leq q_1 \leq x^{1/2}} \frac{1}{q_1} \ll \frac{x}{(\log x)^2} \cdot \frac{\log y}{\log x} = \frac{x}{(\log x)^{11/5}} = o(\#C_2(x))$$

as $x \to \infty$, where the last estimate follows again from Lemma 6.

We now look at primes $p \in C_2(x) \setminus C_3(x)$ and we let $C_4(x)$ be the set of such primes with the property that $\phi_2(p) = \phi_2(p')$ for some $p' \neq p$ also in $C_2(x) \setminus C_3(x)$. Writing $p - 1 = 2q_1q_2$ and $p' - 1 = 2q_1'q_2'$, we have $(q_1 - 1)(q_2 - 1) = (q_1' - 1)(q_2' - 1)$. Fix $q_1$ and $q_1'$. If $q_1 = q_1'$, we then get that $q_2 = q_2'$, therefore $p = p'$, which is false. So, $q_1 \neq q_1'$ and they are both $< x^{1/2}/y$. Let $D = \gcd(q_1 - 1, q_1' - 1)$. Then the equation

$$(q_1 - 1)(q_2 - 1) = (q_1' - 1)(q_2' - 1)$$

can be rewritten as

$$q_2 \left(\frac{q_1 - 1}{D}\right) + q_1' - q_1 = q_2' \left(\frac{q_1' - 1}{D}\right).$$

Let $A = (q_1 - 1)/D$, $B = (q_1' - q_1)/D$, $C = (q_1' - 1)/D$. Then $q_2A + B = Cq_2'$ and $A$ and $C$ are coprime. This puts $q_2$ into a fixed class modulo $C$, namely the congruence class of $\mod{-BA^{-1}}$ modulo $C$. Let this class be $C_0$, where $1 \leq C_0 \leq C - 1$. Then $q_2 = C\ell + C_0$ for some $\ell \geq 0$. We have $q_2 \leq x/q_1$, therefore $\ell \leq x/(q_1C)$. To count such $\ell$’s for a given choice of $q_1, q_1'$, note that

$$C\ell + C_0 = q_2, \quad 2q_1C\ell + 2q_1C_0 + 1 = 2q_1q_2 + 1 = p,$$

$$A\ell + \frac{AC_0 + B}{C} = q_2', \quad 2q_1'A\ell + 2q_1' \left(\frac{AC_0 + B}{C}\right) + 1 = 2q_1'q_2' + 1 = p'$$

are all four prime numbers. By the Brun sieve (it is easy to see that since $B \neq 0$, the four forms above satisfy the hypothesis from the Brun sieve for large $x$), it follows that if we put

$$\Delta = 2q_1q_1'AC_0(2q_1C_0 + 1)(AC_0 + B)(2q_1'(AC_0 + B)/C + 1),$$

then the number of $\ell \leq x/(q_1C)$ with the above property is bounded by

$$x \ll \frac{x}{(q_1C)(\log(x/q_1C))^4} \left(\frac{\Delta}{\phi(\Delta)}\right)^4 \ll \frac{xD(\log \log x)^4}{q_1q_1'(\log y)^4} = \frac{xD(\log \log x)^4}{q_1q_1'(\log x)^{16/5}}.$$
by the minimal order of the Euler function. Keeping now $D$ fixed and summing the above inequality over all pairs of primes $q_1, q'_1 \leq x^{1/2}$ which are congruent to 1 modulo $D$ we get, by the Brun-Titchmarsh theorem, that the number of such primes $p$ once $D$ is fixed is

$$\ll \frac{xD(\log \log x)^4}{(\log x)^{16/5}} \left( \sum_{1 \leq q \leq x^{1/2}} \frac{1}{q} \right)^2 \ll \frac{xD(\log \log x)^6}{\phi(D)^2(\log x)^{16/5}} \ll \frac{x(\log \log x)^8}{D(\log x)^{16/5}},$$

where we again used the minimal order of the Euler function. Summing up over all the values for $D$, we finally get that

$$\# C_4(x) \ll \frac{x(\log \log x)^8}{(\log x)^{16/5}} \sum_{D \leq x^{1/2}} \frac{1}{D} \ll \frac{x(\log \log x)^8}{(\log x)^{11/5}} = o(\# C_2(x))$$

as $x \to \infty$. Thus, putting $C_5(x) = C_2(x) \setminus (C_3(x) \cup C_4(x))$, we have, by the above calculations and Lemma 6, that $\# C_5(x) \gg x/(\log x)^2$. Certainly, $\phi$ is injective when restricted to $C_5(x)$. This takes care of the desired lower bound.

4 Further problems

Observe that $V_k \subseteq V_{k-1}$ for all $k \geq 2$. Put $V_\infty = \cap_{k \geq 1} V_k$. The following result, which was conjectured by A. Chakrabarti [4], characterizes $V_\infty$.

**Theorem 7.** The set $V_\infty$ is equal to the set of positive integers $n$ whose largest squarefree divisor is 1, 2, or 6.

**Proof.** It is clear that such numbers $n$ are in $V_\infty$, since if the largest squarefree divisor of $n$ is 1 or 2, then $\phi_k(2^k n) = n$ for every $k$, while if the largest squarefree divisor of $n$ is 6, then $\phi_k(3^k n) = n$.

Suppose that $n \in V_\infty$. There is thus a sequence $n = n_0, n_1, n_2, \ldots$ such that $\phi(n_i) = n_{i-1}$ for each $i \geq 1$. Note that $v_2(\phi(m)) \geq v_2(m)$ for $m$ not a power of 2. In addition, if we have equality, then $m = 2^c p^b$ where $b, c$ are positive and $p$ is a prime that is 3 (mod 4). Assume that $n_0$ is not a power of 2, so that

$$v_2(n_0) \geq v_2(n_1) \geq \cdots.$$  

Thus, starting at some point, say $n_k$, we have equality; that is,

$$v_2(n_k) = v_2(n_{k+1}) = \cdots.$$  

Thus, for $i \geq 1$ we have

$$n_{k+i} = 2^i p_i^{b_i}, \ p_i \equiv 3 \pmod{4}.$$
We may assume that all $p_i > 3$ for otherwise the theorem holds. If some $b_i > 1$, then $n_{k+i-1} = \varphi(n_{k+i})$ is divisible by two different odd primes, namely $p_i$ and an odd prime factor of $p_i - 1$. Thus, we may assume that each $b_i = 1$ for $i \geq 2$. We have

$$n_{k+i} = 2^i p_i, \quad i \geq 2, \quad p_i = 2p_{i-1} + 1, \quad i \geq 2.$$ 

We can solve this last recurrence, getting

$$p_i = 2^{i-1}(p_1 + 1) - 1, \quad i \geq 2.$$ 

But note then since $2^{p_1 - 1} \equiv 1 \pmod{p_1}$, we have

$$p_{p_1} \equiv (p_1 + 1) - 1 \equiv 0 \pmod{p_1}.$$ 

Thus, $p_{p_1}$ cannot be prime, a contradiction which proves the theorem.

**Remark 2.** Note that the numbers $n$ with largest squarefree divisor 1, 2, or 6 are precisely those $n$ with $\phi(n) \mid n$. Note too that from the counting function up to $x$ of the integers whose largest squarefree factor is 1, 2, or 6, we have

$$\# V_\infty(x) = \frac{1}{\log 3 \log 4} (\log x)^2 + O(\log x). \quad (8)$$

It is possible to use the proof of Theorem 7 to show that there is a number $k = k(n)$ such that if $n \in V_k$, then the largest squarefree divisor of $n$ is 1, 2, or 6. That is, if $n$ is not of this form, not only does there not exist an infinite “reverse Euler chain” starting at $n$, there also cannot exist arbitrarily long finite reverse Euler chains starting at $n$. It is an interesting question to estimate $k(n)$; in [11] it is shown on the generalized Riemann hypothesis that $k(n) \ll \log n$ for $n > 1$.

Let $\lambda(n)$ be the Carmichael function of $n$, or the universal exponent modulo $n$. This is the largest possible multiplicative order of invertible elements modulo $n$. For $k \geq 1$ let $\lambda_k(n)$ be the $k$-fold iterate of $\lambda$ evaluated at $n$. It would be interesting to study $\mathcal{L}_k = \{\lambda^{(k)}(n)\}$. For $k = 1$, an upper bound of the shape $\# \mathcal{L}_1(x) \ll x/(\log x)^{c_1}$ with an inexplicit positive constant $c_1$ was outlined in [9], and an actual numerical value for $c_1$ was established in [12]. Trivially, $\# \mathcal{L}_1(x) \gg x/\log x$. A slightly stronger lower bound appears in [1]. Stronger upper and lower bounds on $\# \mathcal{L}_1(x)$ will appear in [16]. While $\# \mathcal{L}_k(x)$ seems difficult to study for larger values of $k$, it is easy to see that the method of the present paper shows that uniformly for $x$ large,

$$\# \{\lambda_k(n) : n \leq x\} \leq \frac{x}{(\log x)^k} \exp \left(16k^{3/2}(\log_2 x \log_3 x)^{1/2}\right). \quad (9)$$

Indeed, to see this, assume in the notation of the proof of Theorem 1, that $n = pm \leq x$, and that $p > y$. Further, we may assume that $\lambda_k(n) \geq x/(\log x)^k$, since there are at most $x/(\log x)^k$ positive integers failing this condition. We assume that $\Omega(\lambda_k(n)) \leq 2.9k \log x$,
since otherwise Lemma 13 in [15] tells us again that there are at most \( O(x/(\log x)^k) \) possibilities for the number of such positive integers \( \lambda_k(n) \). We now note that \( \lambda_k(n) \mid \phi_k(n) \) and that \( \phi_k(n)/\lambda_k(n) \leq (\log x)^k \). Hence,

\[
\Omega(\phi_k(n)) = \Omega(\lambda_k(n)) + \Omega(\phi_k(n)/\lambda_k(n)) \\
\leq 2.9k \log \log x + \left( \frac{k}{\log 2} \right) \log \log x < 4.5k \log \log x.
\]

In particular, both \( \Omega(\phi_k(p)) \) and \( \Omega(\phi_k(m)) \) are at most \( 4.5k \log \log x \). The argument from the end of the proof of Theorem 1 combined with the fact that \( 3\sqrt{9} + 3/10 + 2.9\sqrt{4.5} < 16 \) shows that the number of possibilities for such \( n \leq x \) is at most what is shown in the right hand side of inequality (9). The conditional argument from the introduction suggests that \( c_k x/(\log x)^k \) should be a lower bound on the cardinality of the above set.

Finally we remark that if \( n \) has the property that \( \lambda(n) \mid n \), then \( n \) is in every set \( L_k \), as is easy to see. It is not clear if the converse holds; for example, is \( n = 10 \) in every \( L_k \)? It is not so easy to find values of \( \lambda \) that are not values of \( \lambda_2 \), but in fact, one can use Brun’s method to show most shifted primes \( p - 1 \) have this property. By using the basic argument at the end of [7] plus the latest results on the distribution of primes \( p \) with \( P(p - 1) \) small, one can prove that for large \( x \) there are at least \( x^{0.7067} \) numbers \( n \leq x \) with \( \lambda(n) \mid n \). Thus, there are at least this many numbers \( n \leq x \) which are in every \( L_k \), a result which stands in stark contrast to (8).

References


Florian Luca
Instituto de Matemáticas
Universidad Nacional Autonoma de México
C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

Carl Pomerance
Department of Mathematics
Dartmouth College
Hanover, NH 03755–3551, USA
carl.pomerance@dartmouth.edu