SOME REMARKS ON EULER'S \( \phi \) FUNCTION AND SOME RELATED PROBLEMS

PAUL ERDÖS

The function \( \phi(n) \) is defined to be the number of integers relatively prime to \( n \), and \( \phi(n) = n \cdot \prod_{p|n}(1 - p^{-1}) \).

In a previous paper\(^1\) I proved the following results:

1. The number of integers \( m \leq n \) for which \( \phi(x) = m \) has a solution is \( o(n \lfloor \log n \rfloor^{\epsilon-1}) \) for every \( \epsilon > 0 \).
2. There exist infinitely many integers \( m \leq n \) such that the equation \( \phi(x) = m \) has more than \( m^c \) solutions for some \( c > 0 \).

In the present note we are going to prove that the number of integers \( m \leq n \) for which \( \phi(x) = m \) has a solution is greater than \( cn(\log n)^{-1} \log \log n \).

By the same method we could prove that the number of integers \( m \leq n \) for which \( \phi(x) = m \) has a solution is greater than \( n(\log n)^{-1}(\log \log n)^k \) for every \( k \). The proof of the sharper result follows the same lines, but is much more complicated. If we denote by \( f(n) \) the number of integers \( m \leq n \) for which \( \phi(x) = m \) has a solution we have the inequalities

\[
n(\log n)^{-1}(\log \log n)^k < f(n) < n(\log n)^{\epsilon-1}.
\]

By more complicated arguments the upper and lower limits could be improved, but to determine the exact order of \( f(n) \) seems difficult.

Also Turán and I proved some time ago that the number of integers \( m \leq n \) for which \( \phi(m) \leq n \) is \( cn + o(n) \). We shall give this proof, and also discuss some related questions:

**Lemma 1.** Let \( a < e, \ b < n, \ a \neq b, \ e = (\log \log n)^{-100} \). Then the number of solutions \( N_n(a, b) \) of

\[
(p - 1)a = (q - 1)b, \quad p \leq na^{-1}, \quad q \leq nb^{-1},
\]

\( p, q \) primes, does not exceed

\[
\frac{(a, b)}{ab} \frac{n}{(\log n)^2} (\log \log n)^{30}.
\]

**Proof.** Put \( (a, b) = d \). Then we have \( p \equiv 1 \mod bd^{-1} \). Also \( (p - 1)ab^{-1} + 1 = q \) is a prime. We can assume that both \( p \) and \( q \) in (1) are greater

than \( n^{1/2} \), for the exceptional values of \( p \) and \( q \) give only \( 2n^{1/2} \) solutions of (1). Let \( r < n^\delta \), where \( \delta = (\log \log n)^{-10} \), be a prime. If \( p \) is a solution of (1) it must satisfy the following conditions
\[
\begin{align*}
p &\equiv 1 \mod bd^{-1}, \\
p &< na^{-1}, \\
p &\not\equiv 0 \mod r, \\
p &\not\equiv (-ba^{-1} + 1) \mod r.
\end{align*}
\]
If \( r \) is not a divisor of \( a(a-b) \) the excluded two residues are different. Thus we obtain by Brun's argument\(^2\)
\[
N_n(a, b) < 2n^{1/2} + c_1nd(ab)^{-1} \prod_{r \mid a(a-b)} (1 - 2r^{-1}),
\]
where \( r \) runs through the primes less than \( n^\delta \). Now it is well known that\(^3\)
\[
\prod_{r \leq x} (1 - 2r^{-1}) < c_2 (\log x)^{-2}, \quad \prod_{r > x} (1 - 2r^{-1}) > c_4 (\log \log x)^{-2}.
\]
Hence
\[
N_n(a, b) < 2n^{1/2} + c_4nd(ab)^{-1}(\log \log n)^{10}(\log n)^{-2}
\]
\[
< nd(ab)^{-1}(\log \log n)^{10}(\log n)^{-2},
\]
which completes the proof.

**Lemma 2.** \( \sum (p - 1)^{-1} < (\log \log n)^{20}d^{-1} \) if this sum is extended over all \( p < n^\delta \) for which \( p \equiv 1 \mod d \).

Clearly (summing over the indicated \( p \))
\[
\sum p^{-1} \leq d^{-1} \sum' x^{-1},
\]
where the dash indicates that the summation is extended over the \( x \) for which \( x < nd^{-1} \) and \( xd + 1 \) is a prime. Let \( y < nd^{-1} \); first we estimate the number of these \( x \leq y \leq n \). Let \( r < y^\delta \) (\( \delta = (\log \log n)^{-10} \)) be a prime; if \( (r, d) = 1 \) then \( x \not\equiv -d^{-1} \mod r \). Brun's method\(^4\) gives that the number of these \( x \leq y \) is less than
\[
cy \prod (1 - r^{-1}) < cy(\log y)^{-1}(\log \log y)^{10} \log \log d,
\]
where the product is extended over the \( r \) which satisfy \( r < y^\delta \), \( (r, d) = 1 \). Thus a simple argument gives
\[
\sum' x^{-1} < c \sum_{x < y} (\log \log x)^{10}(\log \log d)(x \log x)^{-1} < (\log \log n)^{10},
\]
which proves the lemma.

\(^3\) Hardy-Wright, *Theory of numbers*.
\(^4\) Landau, ibid.
Lemma 3. The number $A(n)$ of integers $m$ of the form $m=pq$, where $pq \leq n$, \( p, q \) primes, \( p > q, q < n^* \), equals

\[ n(\log \log n)(\log n)^{-1} + o([n(\log \log n)(\log n)^{-1}]) = \pi_2(n) + o(\pi_2(n)). \]

Remark. Thus the number of integers satisfying (3) is asymptotically equal to the number \( \pi_2(n) \) of integers which are less than \( n \) and have 2 prime factors.\(^6\)

The number of integers satisfying (3) is clearly not less than

\[
\sum (\pi(nq^{-1}) - n^*) = \sum nq^{-1}(\log (nq^{-1}))^{-1} - n^{2*} \\
+ \sum o(nq^{-1}[\log (nq^{-1})]^{-1}) \\
= n(\log \log n)(\log n)^{-1} + o(n(\log \log n)(\log n)^{-1})
\]

(here \( \pi(n) \) denotes the number of primes, and the sums are taken over \( q < n^* \), since \( \sum q^{-1} = \log_2 n + \log e + o(1) \) and \( \log (nq^{-1}) \) is asymptotic to \( \log n \) for \( q < n^* \). (The sum \( \sum q^{-1} \) is for \( q < n^* \).)

Theorem. The number \( f(n) \) of different integers \( m \) of the form \( m=\phi(pr) \) where \( p, r \) are primes and \( pr \leq n \) equals

\[ n(\log \log n)(\log n)^{-1} + o(n(\log \log n)(\log n)^{-1}) = \pi_2(n) + o(\pi_2(n)). \]

Denote by \( B(n) \) the number of solutions of \((p-1)(r-1) = (q-1)(s-1)\), where \( p, q, r, s \) are primes, with \( pq, rs < n \) and \( s, r < n^* \). Clearly

\[ f(n) \geq A(n) - B(n). \]

We have by Lemma 1 (the following sum being for \( r, s < n^* \))

\[
B(n) = \sum N_n(r - 1, s - 1) \\
< n(\log \log n)^{20}(\log n)^{-1} \sum (r - 1, s - 1)(r - 1)^{-1}(s - 1)^{-1}.
\]

Put \( (r-1, s-1) = d. \) Then

\[
B_n < n(\log n)^{-2}(\log \log n)^{20} \sum d(q - 1)^{-1}(s - 1)^{-1},
\]

where the first sum is for \( d < n^* \) and the second for \( r \equiv s \equiv 1 \) mod \( d \), with \( r, s < n^* \). By Lemma 2 we have, summing over the same \( r \) and \( s \),

\[
\sum (r - 1)^{-1}(s - 1)^{-1} < (\log \log n)^{40}d^{-2}.
\]

\(^6\) Denote by \( \pi_k(n) \) the number of integers having \( k \) different prime factors. Landau proves (\textit{Verteilung der Primzahlen}, vol. 1, pp. 208–213) that \( \pi_k(n) \sim (n/\log n)(\log \log n)^{k-1}/(k-1)! \). The same asymptotic formula holds if \( \pi_k(n) \) denotes the number of integers having \( k \) prime factors, multiple factors counted multiply. (Landau, ibid.)
Hence
\[ B(n) = c\gamma n (\log n)^{-1} (\log \log n)^{\gamma} = o(n (\log n)^{-1}). \]
Hence by Lemma 3
\[ f(n) \geq n (\log \log n) (\log n)^{-1} - o(n (\log n)^{-1}), \]
which completes the proof. (Clearly \( f(n) < \pi_2(n) < (1 + \epsilon)n (\log \log n) \cdot (\log n)^{-1} \). Our result shows that the number of different integers not greater than \( n \) of the form \((p-1)(q-1)\) is asymptotic to the total number of integers not greater than \( n \) of the form \((p-1)(q-1)\). Nevertheless there exist integers \( m \) such that \((p-1)(q-1) = m\) has arbitrarily many solutions.\(^6\)

By similar but more complicated methods we can prove:
The number of integers not greater than \( n \) of the form
\[ \prod_{i=1}^{k} (p_i - 1) = \phi(p_1, \ldots, p_k) \quad (p_i \text{ primes}) \]
is greater than
\[ cn (\log \log n)^{k-1} [(k - 1)! \log n]^{-1} = c\pi_k(n) + o(\pi_k(n)) \]
(\(\pi_k(n)\) denotes the number of integers not greater than \( n \) having exactly \( k \) prime factors). The constant \( c \) depends on \( k \) and tends to 0 as \( k \to \infty \). For \( k \geq 3 \), \( c < 1 \). We omit the proof of these results.

**Theorem.** The number \( M(n) \) of integers for which \( \phi(m) \leq n \) equals \( cn + o(n) \).

Denote by \( f(x) \) the density of integers for which \( m/\phi(m) \geq x \). It is well known that this density exists.\(^7\) We are going to prove that
\[ c = 1 + \int_{1}^{\infty} f(x)dx. \]
First we have to show that \( \int_{1}^{\infty} f(x)dx \) exists. Since \( f(x) \) is nondecreasing it will suffice to show that for large \( r \), \( f(r) < cr^{-2} \). We have
\[ \sum_{m=1}^{n} (m/\phi(m))^2 = \sum_{m=1}^{n} \prod_{p|m} (1 + p^{-1} + \cdots)^2 < \sum_{m=1}^{n} \prod_{p|m} (1 + 5p^{-1}) \]
\[ = \sum_{m=1}^{n} \sum_{d|m} \mu(d)d^{-15\ast(d)} < n \sum_{d=1}^{\infty} 5d^{-3} < cn. \]

---


PAUL ERDÖS

Hence

$$\lim n^{-1} \sum_{m=1}^{n} (m/\phi(m))^2 < c$$

and this shows $f(r) < c r^{-2}$.

Let $k$ be a large number. Consider the integers $m$ satisfying $nuk^{-1} \leq m < n(u+1)k^{-1}$, $u \geq k$. We clearly have

$$\limsup M(n)/n < 1 + k^{-1} \sum_{u=k}^{\infty} f(uk^{-1})$$

and

$$\liminf M(n)/n > 1 + k^{-1} \sum_{u=k}^{\infty} f((u+1)k^{-1}).$$

(If $uk^{-1} \leq m \leq (u+1)k^{-1}$ and $m/\phi(m) \geq (u+1)k^{-1}$, $\phi(m) < n$ and if $m/\phi(m) < uk^{-1}$, $\phi(m) > n$.) If $k \to \infty$ both sums tend to $\int_{1}^{\infty} f(x)dx$, thus

$$\lim M(n)/n = 1 + \int_{1}^{\infty} f(x)dx$$

which completes the proof.

Let $\sigma(m)$ be the sum of the divisors of $m$. By the same methods as used before we can prove the following results:

1. The number of integers $m$ for which $\sigma(m) \leq n$ is $cn + o(n)$.

2. Denote by $g(m)$ the number of integers $m \leq n$ for which $\sigma(x) = m$ is solvable. Then $n(\log n)^{-1}(\log \log n)^{k} < g(n) < n(\log n)^{-1}(\log n)^{k}$.

It seems likely that there exist integers $m$ such that the equation $\phi(x) = m$ has more than $m^{1+\epsilon}$ solutions, and also that there exist, for every $k$, consecutive integers $n$, $n+1$, $\ldots$, $n+k-1$ such that $\phi(n) = \phi(n+1) \cdots \phi(n+k-1)$.* We can make analogous conjectures for $\sigma(n)$. It also would seem likely that there are infinitely many pairs of integers $x$ and $y$ with $\sigma(x) = \sigma(y) = x+y$, that is, there are infinitely many friendly numbers, but these conjectures seem intractable at present.

One final remark: Let $\psi(n) \geq 0$ be a multiplicative function which has a distribution function.* $f(x)$ denotes the density of integers with $\psi(n) \geq x$. Denote by $M(n)$ the number of integers for which $n\psi(n) \leq n$. Then $\lim M(n)/n$ always exists since it can be shown that $\int_{1}^{\infty} f(x)dx$ always exists. The proof is the same as in the case of $\phi(n)$.

* It is known that there exists a number $n < 10000$ such that $\phi(n) = \phi(n+1) = \phi(n+2)$, but I do not remember $n$ and cannot trace the reference.

* The necessary and sufficient condition for the existence of the distribution function is given by Erdős-Wintner, Amer. J. Math. vol. 61 (1939) pp. 713-721.