SOME REMARKS ON EULER'S ϕ 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 $m\leqq n$ for which $\phi(x)=m$ has a solution is $o(n[\log n]^{\epsilon-1})$ for every $\epsilon>0$.

(2) There exist infinitely many integers $m\leqq n$ 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 $m\leqq n$ 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 $m\leqq n$ 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 $m\leqq n$ 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 $m\leqq n$ for which $\phi(m)\leqq n$ is $cn+o(n)$. We shall give this proof, and also discuss some related questions:

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

$$(1)\qquad (p-1)a=(q-1)b,\qquad p\leqq na^{-1},\qquad q\leqq nb^{-1},$$

p, q primes, does not exceed

$$(2)\qquad \frac{(a, b)}{ab}\,\frac{n}{(\log n)^2}(\log\log n)^{80}.$$

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

Received by the editors February 9, 1945.

[1] *On the normal number of prime factors of $p-1$*, Quart. J. Math. Oxford Ser. vol. 6 (1935) pp. 205–213.

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

$$p \equiv 1 \bmod bd^{-1}, \qquad p < na^{-1},$$
$$p \not\equiv 0 \bmod r, \qquad p \not\equiv (-ba^{-1}+1) \bmod r.$$

If r is not a divisor of $a(a-b)$ the excluded two residues are different. Thus we obtain by Brun's argument[2]

$$N_n(a, b) < 2n^{1/2} + c_1 nd(ab)^{-1} \prod_{r \nmid a(a-b)} (1 - 2r^{-1}),$$

where r runs through the primes less than n^{δ}.

Now it is well known that[3]

$$\prod_{r \leqq x} (1 - 2r^{-1}) < c_2(\log x)^{-2}, \qquad \prod_{r \mid x} (1 - 2r^{-1}) > c_3(\log\log x)^{-2}.$$

Hence

$$N_n(a, b) < 2n^{1/2} + c_4 nd(ab)^{-1}(\log\log n)^{22}(\log n)^{-2}$$
$$< nd(ab)^{-1}(\log\log n)^{30}(\log n)^{-2},$$

which completes the proof.

LEMMA 2. *$\sum(p-1)^{-1}<(\log\log n)^{20}d^{-1}$ if this sum is extended over all $p<n^{\epsilon}$ for which $p\equiv 1 \bmod d$.*

Clearly (summing over the indicated p)

$$\sum p^{-1} \leqq d^{-1} \sum{}' x^{-1},$$

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 $x\leqq y\leqq n$. Let $r<y^{\delta}$ $(\delta=(\log\log n)^{-10})$ be a prime; if $(r, d)=1$ then $x\not\equiv -d^{-1} \bmod r$. Brun's method[4] gives that the number of these $x\leqq y$ 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

$$\sum{}' x^{-1} < c \sum_{z<n} (\log\log z)^{10}(\log\log d)(z \log z)^{-1} < (\log\log n)^{20},$$

which proves the lemma.

[2] Landau, *Vorlesungen über Zahlentheorie*, vol. 1, p. 71.
[3] Hardy-Wright, *Theory of numbers*.
[4] Landau, ibid.

LEMMA 3. *The number $A(n)$ of integers m of the form $m=pq$, where*

$$pq \leqq n, \tag{3}$$

p, q primes, $p>q$, $q<n^{\epsilon}$, 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.[5]

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

$$\begin{aligned}\sum (\pi(nq^{-1}) - n^{\epsilon}) &= \sum nq^{-1}(\log (nq^{-1}))^{-1} - n^{2\epsilon} \\ &\quad + \sum o(nq^{-1}[\log (nq^{-1})]^{-1}) \\ &= n(\log\log n)(\log n)^{-1} + o(n(\log\log n)(\log n)^{-1})\end{aligned}$$

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

THEOREM. *The number $f(n)$ of different integers m of the form $m=\phi(pr)$ where p, r are primes and $pr \leqq n$ 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, $r<n^{\epsilon}$. Clearly

$$f(n) \geqq A(n) - B(n).$$

We have by Lemma 1 (the following sum being for r, $s<n^{\epsilon}$)

$$\begin{aligned}B(n) &= \sum N_n(r-1, s-1) \\ &< n(\log\log n)^{30}(\log n)^{-1} \sum (r-1, s-1)(r-1)^{-1}(s-1)^{-1}.\end{aligned}$$

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

$$B_n < n(\log n)^{-2}(\log\log n)^{30} \sum\sum d(q-1)^{-1}(s-1)^{-1},$$

where the first sum is for $d<n^{\epsilon}$ and the second for $r\equiv s\equiv 1 \bmod d$, with r, $s<n^{\epsilon}$. By Lemma 2 we have, summing over the same r and s,

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

[5] Denote by $\pi_k(n)$ the number of integers having k different prime factors. Landau proves (*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

$$B(n) = c\epsilon n(\log n)^{-1}(\log\log n)^{70} = o(n(\log n)^{-1}).$$

Hence by Lemma 3

$$f(n) \geqq n(\log\log n)(\log n)^{-1} - o(n(\log n)^{-1}),$$

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

$$\prod_{i=1}^{k}(p_i - 1) = \phi(p_1, \cdots, p_k) \qquad (p_i \text{ primes})$$

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 $k \geqq 3$, $c<1$. We omit the proof of these results.

THEOREM. *The number $M(n)$ of integers for which $\phi(m) \leqq n$ equals $cn+o(n)$.*

Denote by $f(x)$ the density of integers for which $m/\phi(m) \geqq x$. 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

$$\sum_{m=1}^{n}(m/\phi(m))^2 = \sum_{m=1}^{n}\prod_{p|m}(1+p^{-1}+\cdots)^2 < \sum_{m=1}^{n}\prod_{p|m}(1+5p^{-1})$$

$$= \sum_{m=1}^{n}\sum_{d|m}\mu(d)d^{-1}5^{\nu(d)} < n\sum_{d=1}^{\infty}5d^{-2} < cn.$$

[6] P. Erdös, *On the totient of the product of two primes*, Quart. J. Math. Oxford Ser. vol. 7 (1936) pp. 227–229.

[7] Schönberg, Math. Zeit. vol. 28 (1928) pp. 171–199.

Hence

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

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

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

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

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

(If $uk^{-1} \leqq m \leqq (u+1)k^{-1}$ and $m/\phi(m) \geqq (u+1)k^{-1}$, $\phi(m) < n$ 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 $\sigma(m) \leqq n$ is $cn + o(n)$.

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

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 $n, n+1, \cdots, n+k-1$ such that $\phi(n) = \phi(n+1) \cdots \phi(n+k-1)$.[8] 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 $\psi(n) \geqq 0$ be a multiplicative function which has a distribution function.[9] $f(x)$ denotes the density of integers with $\psi(n) \geqq x$. Denote by $M(n)$ the number of integers for which $n\psi(n) \leqq n$. Then $\lim M(n)/n$ always exists since it can be shown that $\int_0^\infty f(x)dx$ always exists. The proof is the same as in the case of $\phi(n)$.

UNIVERSITY OF MICHIGAN

[8] 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.

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