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 $(\leqslant n)$ for which the number of prime factors does not lie between $(1 \pm \epsilon)B(n)$ for arbitrarily small positive ϵ.

We use throughout the following notation: $N(M, n)$ denotes the number of integers not exceeding n in the set M; $d(n)$ is the number of divisors of n; $\mu = \log n$, $\nu = \log\log n$; p, p_1, p_1',... are prime numbers, and C_1, C_2,... denote positive constants independent of n, m.

In the first part, I prove that, if M is the set $p-1$, and so $N(n) \sim n/\mu$, then $B(n) = \nu$. I use the method of Brun and also that employed by Hardy and Ramanujan* in their proof that, when M is the set of all natural numbers, $B(n) = \nu$. I then apply my result to a problem of Titchmarsh† who showed that, if

$$S = \sum_{p \leqslant n} d(p-1),$$

(i) $S < Cn$, by Brun's method;

(ii) $S = \Omega\left(\frac{n}{\sqrt{\mu}}\right)$ by analytical methods;

(iii) $S = C_1 n + o(n)$ by assuming the Riemann hypothesis.

As my result means that, for almost all p not exceeding n, i.e. except for $o(n/\log n)$ of the p, $p-1$ has more than $(1-\epsilon)\nu$ prime

* Hardy-Ramanujan, *Quart. J. of Math.* 48 (1917), 76–92. See also S. Ramanujan, *Collected Papers*, 262–75. Recently P. Turán gave a very simple proof of this theorem, but the application of his method seems to be impossible here. *J. of London Math. Soc.* 9 (1934), 274–76.

† E. C. Titchmarsh, *Rend. del Circ. Mat. di Palermo*, 54 (1930), 414–19.

factors, it is obvious that

$$S > \frac{n}{2\mu} 2^{(1-\epsilon)\nu},$$

since $N(p, n) > \frac{1}{2}n/\mu$. This result is better than (ii) and is obtained in a more elementary way.

In the second part I deal with Euler's function $\phi(n)$. I consider first the number $N(M, n)$ where the set M denotes now the integers which can be expressed as the ϕ of another integer. S. S. Pillai* found that

$$N(M, n) < \frac{C_2 n}{\mu^{(\log 2)/e}}.$$

I deduce from the first part that

$$N(M, n) < \frac{n}{\mu^{1-\epsilon}}$$

for every positive ϵ and every n exceeding some $n(\epsilon)$. I can prove by Brun's method that

$$N(M, n) > C_3 \frac{n}{\mu} \log \nu.$$

In the third part I examine how often an integer m can be represented as the ϕ of another integer. S. S. Pillai showed that integers m exist with at least $C_4(\log m)^{(\log 2)/e}$ representations. I replace this number by m^{C_5} by using Brun's method.

1. We shall presently evaluate $N(M, n)$ for a certain set M. It will suffice to deal only with the m satisfying the following two conditions:

(i) the greatest prime factor of m is greater than $n^{1/20\nu}$;

(ii) the greatest prime factor occurs to the first power only.

For we have

LEMMA 1. *The number of m (and in fact of all positive integers not exceeding n) which do not satisfy both the conditions* (i), (ii) *is* $o(n\mu^{-2})$.

We divide the integers not exceeding n which do not satisfy (i) into two classes N_1, N_2 in number, putting in the first those which have at most 10ν different prime factors. As the $\{\mu/(\log 2)\}$th power of any prime less than $n^{1/20\nu}$ is greater than n, we have

$$N_1 < \left\{\left(1 + \frac{\mu}{\log 2}\right) n^{1/20\nu}\right\}^{10\nu} = n^{\frac{1}{2}}\left(1 + \frac{\mu}{\log 2}\right)^{10\nu} = o\left(\frac{n}{\mu^2}\right).$$

* I have seen this in an American periodical that I cannot now trace.

The integers m of the second class have more than 10ν different prime factors; and so $d(m) > 2^{10\nu}$. But

$$\sum_{l=1}^{n} d(l) = O(n\mu)$$

and so
$$N_2 = O\left(\frac{n\mu}{2^{10\nu}}\right) = o\left(\frac{n}{\mu^2}\right),$$

since
$$\frac{\mu^3}{2^{10\nu}} = e^{(3-10\log 2)\nu} = o(1).$$

Hence
$$N_1+N_2 = o\left(\frac{n}{\mu^2}\right).$$

In dealing with the integers not satisfying (ii) we may, from the first part, suppose that their greatest prime factor exceeds $n^{1/(20\nu)}$. Hence these integers are divisible by a square exceeding $n^{1/(10\nu)}$ and so their number is less than

$$\sum_{l^2>n^{1/(10\nu)}} \frac{n}{l^2} = O\left(\frac{n}{n^{1/(20\nu)}}\right) = o\left(\frac{n}{\mu^2}\right).$$

This proves the lemma.

We now require the following result which is an immediate consequence of Brun's* method.

If a is a given integer and $\phi_n(a)$ denotes $N(p,n)$ where $(p-1)/a$ is a prime, then

$$\begin{aligned}\phi_n(a) &< C_6\frac{n}{a}\prod_{\substack{p<n/a\\ p>2}}\left(1-\frac{2}{p}\right)\prod_{p|a}\left(1-\frac{1}{p}\right)\Big/\prod_{\substack{p|a\\ p>2}}\left(1-\frac{2}{p}\right)\\ &< C_7\frac{n}{a(\log n/a)^2}\prod_{p|a}\left(1-\frac{1}{p}\right)\Big/\prod_{\substack{p|a\\ p>2}}\left(1-\frac{2}{p}\right)\\ &< C_8\frac{n\nu^2}{a(\log n/a)^2}, \qquad (1)\end{aligned}$$

since
$$\prod_{\substack{p|a\\ p>2}}\left(1-\frac{2}{p}\right) > \frac{C_9}{(\log\log a)^2}$$

follows easily from Landau's result $\phi(a) > C_{10}\,a/(\log\log a)$.

Denote the positive integers containing exactly k different prime

* V. Brun, *Vidensk. selsk. skrifter, Mat.-Naturw. Kl.* (Kristiania), 3 (1920), and *Comptes rendus*, 168 (1919), 544–6. See also *Bull. Soc. Math.* (2) 43 (1914), 1–9.

factors by $a_1^{(k)}, a_2^{(k)},\ldots$ and put $f_n(k) = N(p,n)$ where p is such that $p-1$ equals one of the $a_i^{(k)}$. We prove that

$$f_n(k) \leqslant \sum_{a_i^{(k-1)}=1}^{n^{1-1/(20\nu)}} \phi_n(a_i^{(k-1)}) + o\left(\frac{n}{\mu^2}\right). \tag{2}$$

For let us write down the $f_n(k)$ primes p not exceeding n for which

$$p-1 = a_i^{(k)} = q_1^{\alpha_1} q_2^{\alpha_2} \ldots q_k^{\alpha_k},$$

where the q's are primes and $q_1 < q_2 \ldots < q_k$. By Lemma 1 we need only consider the cases given by $q_k > n^{1/(20\nu)}$, $\alpha_k = 1$. Consider also the primes p' such that

$$p'-1 = qa_i^{(k-1)},$$

where q is a prime and $a_i^{(k-1)} < n^{1-1/(20\nu)}$. The inequality (2) will be proved, if every p occurs among the p', and this is obviously the case, since, for given p, we may choose

$$a_i^{(k-1)} = q_1^{\alpha_1} q_2^{\alpha_2} \ldots q_{k-1}^{\alpha_{k-1}} < n^{1-1/(20\nu)},$$

since $q_k > n^{1/(20\nu)}$. Thus (2) is established.

From (1), (2), we have

$$f_n(k) < C_{11} \sum_{a_i^{(k-1)}=1}^{n^{1-1/(20\nu)}} n\nu^2 \Big/ a_i^{(k-1)} \left(\log \frac{n}{a_i^{(k-1)}}\right)^2 + o\left(\frac{n}{\mu^2}\right)$$

$$\leqslant C_{12} \frac{n\nu^4}{\mu^2} \sum_{a_i^{(k-1)}=1}^{n} \frac{1}{a_i^{(k-1)}} + o\left(\frac{n}{\mu^2}\right). \tag{3}$$

Now
$$\sum_{a_i^{(k-1)}=1}^{n} \frac{1}{a_i^{(k-1)}} < \frac{\left(\sum_{p_i \leqslant n} \sum_{\alpha=1}^{\infty} 1/p_i^{\alpha}\right)^{k-1}}{(k-1)!} \leqslant \frac{(\nu + C_{13})^{k-1}}{(k-1)!}; \tag{4}$$

so
$$f_n(k) < \frac{C_{12}\, n(\nu + C_{13})^{k+3}}{(k-1)!\, \mu^2} + o\left(\frac{n}{\mu^2}\right), \tag{5}$$

or say
$$f_n(k) = B_k + o\left(\frac{n}{\mu^2}\right).$$

Applying the method used by Hardy and Ramanujan to prove that almost all integers have ν different prime factors, we now prove our theorem that ν is also the normal number of prime factors of $p-1$. We have to show that

$$\sum_{k<\nu(1-\epsilon)} f_n(k) + \sum_{k>\nu(1+\epsilon)} f_n(k) = O\left(\frac{n}{\mu^{1+\delta}}\right).$$

It suffices to deal with the case $k < \nu(1-\epsilon)$, since the sum for $k > \nu(1+\epsilon)$ follows in a similar way. Now

$$\sum_{k=1}^{\infty} B_k = \frac{C_{12}\, n(\nu+C_{13})^4}{\mu^2} \sum_{k=1}^{\infty} \frac{(\nu+C_{13})^{k-1}}{(k-1)!}$$
$$< C_{14}\frac{n\nu^4}{\mu^2} e^{\nu+C_{13}}$$
$$< C_{15}\, n\frac{\nu^4}{\mu}.$$

Clearly $B_1 < B_2 < \ldots < B_{[(1-\epsilon)\nu]}$, for $\nu > \nu(\epsilon)$. Also

$$\frac{B_{[\nu(1-\frac{1}{2}\epsilon)]}}{B_{[\nu(1-\epsilon)]}} = \frac{(\nu+C_{13})^{[\nu(1-\frac{1}{2}\epsilon)]-[\nu(1-\epsilon)]}}{\{[\nu(1-\frac{1}{2}\epsilon)]-1\}\{[\nu(1-\frac{1}{2}\epsilon)]-2\}\ldots[\nu(1-\epsilon)]}$$
$$> \frac{(\nu+C_{13})^{\frac{1}{2}\epsilon\nu-1}}{\nu(1-\frac{1}{2}\epsilon)\{\nu(1-\frac{1}{2}\epsilon)-1\}\ldots\{\nu(1-\epsilon)+1\}}$$
$$> \frac{1}{(\nu+C_{13})}\left\{\frac{\nu+C_{13}}{\nu(1-\frac{1}{2}\epsilon)}\right\}^{\frac{1}{2}\epsilon\nu}$$
$$> \frac{1}{(\nu+C_{13})}(1+\tfrac{1}{2}\epsilon)^{\frac{1}{2}\epsilon\nu}, > \nu^5\mu^\delta \text{ for sufficiently small } \delta.$$

Hence

$$\sum_{k=1}^{\nu(1-\epsilon)} B_k < \nu B_{[\nu(1-\epsilon)]} < \frac{B_{[\nu(1-\frac{1}{2}\epsilon)]}}{\nu^4\mu^\delta} < \sum_{k=1}^{\infty}\frac{B_k}{\mu^\delta\nu^4} < \frac{C_{15}\, n}{\mu^{1+\delta}} = O\left(\frac{n}{\mu^{1+\delta}}\right).$$

Also
$$\sum_{k=1}^{\nu} o\left(\frac{n}{\mu^2}\right) = o\left(\frac{n}{\mu^{1+\delta}}\right).$$

Thus
$$\sum_{k<\nu(1-\epsilon)} f_n(k) < \sum_{k=1}^{\nu(1-\epsilon)} \left\{B_k + o\left(\frac{n}{\mu^2}\right)\right\} = O\left(\frac{n}{\mu^{1+\delta}}\right),$$

the required result.

By similar but perhaps a little more complicated arguments, we can show that the same result holds when multiple factors are counted multiply, i.e. when a prime power q^α dividing $p-1$ is reckoned as α factors instead of 1.

2. We prove the

THEOREM. $N(M,n) = o(n\mu^{\epsilon-1})$ *for all positive* ϵ, *where the set* M *are the integers which can be expressed in the form* $\phi(x)$.

The proof depends upon the result, due to Hardy and Ramanujan,

$$N(m_k, n) < C_{16}\frac{n(\log\log n + C_{17})^{k-1}}{(k-1)!\log n}, \tag{6}$$

where m_k denotes the integers having k different prime factors. Since

$$\phi(x) > C_{10} \frac{x}{\log\log x},$$

clearly $\phi(x) > n$ if $x > C_{18}n\nu$. Hence it will suffice to prove that there are only $o(n\mu^{\epsilon-1})$ different values in the set $\phi(1)$, $\phi(2)$,..., $\phi([C_{18}n\nu])$.

Consider first the integers not exceeding $C_{18}n\nu$ which have less than ν/k different prime factors where k is for the moment arbitrary. On replacing n, k in (6) by $C_{18}n\nu$, ν/k respectively, and noting that $k! > (k/e)^k$, we prove easily that their number is $o(n\mu^{1-\epsilon})$ for every ϵ if $k > k(\epsilon)$, say, independent of n, and so they need not be dealt with any further.

We have still to consider the integers which have more than ν/k different prime factors. Denote now by p, q respectively the primes such that $p-1$ has respectively less than and not less than $40k+1$ different prime factors. From (5), we deduce that, for sufficiently large n,

$$N(p,n) < \frac{C_{12}\,40kn\nu(\nu+C_{13})^{40k+3}}{\mu^2} + O\left(\frac{n}{\mu^2}\right) < \frac{n}{\mu^{\frac{3}{2}}}.$$

Hence $\sum_p p^{-1}$ converges, since

$$\sum_p p^{-1} = \sum_{n=1}^{\infty} \frac{N(p,n)-N(p,n-1)}{n} = \sum_{n=1}^{\infty} N(p,n)\left(\frac{1}{n}-\frac{1}{n+1}\right)$$
$$= \sum_{n=1}^{\infty} O\left(\frac{1}{n\mu^{\frac{3}{2}}}\right).$$

We now divide the integers having more than ν/k different prime factors into two classes M_1, M_2, putting in the first those divisible by at least $\frac{1}{2}\nu/k$ of the p and in the second class the remainder, say the b's, which of course are divisible by at least $\frac{1}{2}\nu/k$ of the q. The integers m_1 are divisible by an integer a (say) composed of exactly $[\frac{1}{2}\nu/k]$ of the p. Hence

$$N(m_1, C_{18}n\nu) < C_{18}n\nu \sum_a 1/a$$
$$< \frac{C_{18}n\nu\left(\sum_{\alpha_i \geq 1, p} \frac{1}{p^{\alpha_i}}\right)^{[\frac{1}{2}\nu/k]}}{[\frac{1}{2}\nu/k]!}$$

$$< \frac{C_{18}\, n\nu A^{[\frac{1}{2}\nu/k]}}{[\frac{1}{2}\nu/k]\,!} = o\left(\frac{n}{\mu}\right),$$

where $\sum_{\alpha_i \geqslant 1, p} 1/p^{\alpha_i}$ converges to A, say.

We now deal with the b's. Clearly $\phi(b)$ has more than $(\frac{1}{2}\nu/k)40k$, i.e. 20ν prime factors, p^{α} now reckoning as α factors. The integers having more than 20ν prime factors are now divided into two sets of which the first includes the integers whose square-free part has more than 10ν prime factors. Each of these integers has more than $2^{10\nu}$ divisors and so, since

$$\sum_{n=1}^{x} d(n) \sim x \log x,$$

their number is less than

$$\frac{C_{18}\, n\nu \log n\nu}{2^{10\nu}} = o\left(\frac{n}{\mu}\right).$$

The second set includes the integers whose square-free part has not more than 10ν prime factors, and so their quadratic part has at least 10ν prime factors. An integer, however, whose quadratic part is s is divisible by a square exceeding $s^{\frac{2}{3}}$, as is easily seen by putting $s = p_1'^{\alpha_1} p_2'^{\alpha_2} \ldots$ $(\alpha_i > 1)$. Hence the number of the integers of the second set is less than

$$C_{18}\, n\nu \sum_{k^2 > 2^{20\nu/3}} \frac{1}{k^2} = O\left(\frac{n\nu}{2^{10\nu/3}}\right) = o\left(\frac{n}{\mu}\right),$$

since $2^{10\nu/3} > \mu^2$.

Hence there are only $o(n/\mu)$ different values for $\phi(b)$ and so the theorem is proved.

3. We require three lemmas.

Lemma 2. *$N(m, n) = o(n^{\epsilon})$ for every positive ϵ, if m is a number whose greatest prime factor is less than μ.*

Every integer can be expressed in one and only one way as a product of an rth power $(r > 1)$, and an integer not divisible by any rth power. Denote by m_r an integer free from rth-power divisors, whose greatest prime factor is less than μ. Then

$$N(m_r, n) < r^{C_{19}\,\mu/\nu},$$

since the number of primes less than μ is less than $C_{19}\,\mu/\nu$. Hence

$$N(m, n) < n^{1/r} r^{C_{19}\,\mu/\nu}.$$

But r is arbitrary and can be taken so large that

$$N(m,n) = o(n^{\epsilon}).$$

Let ρ be any fixed number such that $0 < \rho < 1$. Then from the prime-number theorem

$$N(p, \mu^{1+\rho}) > C_{20}\mu^{1+\rho}/\nu.$$

We now prove

LEMMA 3. *The number of square-free integers not exceeding n composed of $[C_{21}\mu^{1+\rho}/\nu]+1$ arbitrarily given primes not exceeding $\mu^{1+\rho}$, where $C_{21} < C_{20}$, is $\Omega(n^{\sigma})$ $(0 < \sigma < \frac{1}{2}\rho)$.*

For consider the square-free integers composed of the given primes and having $[\mu/(1+\rho)\nu]$ factors. These are all less than n, since

$$(\mu^{1+\rho})^{\mu/(1+\rho)\nu} = n,$$

and their number is the binomial coefficient

$$\binom{\left[\dfrac{C_{21}\mu^{1+\rho}}{\nu}+1\right]}{\left[\dfrac{\mu}{(1+\rho)\nu}\right]}.$$

Since $\binom{n}{k} \geqslant \left(\dfrac{n}{k}\right)^k$, this coefficient is greater than

$$\{C_{21}\mu^{\rho}(1+\rho)\}^{[\mu/(1+\rho)\nu]} > C_{21}{}^{[\mu/(1+\rho)\nu]}(\mu^{\rho})^{\mu/2\nu}$$
$$> C_{21}{}^{[\mu/(1+\rho)\nu]}n^{\frac{1}{2}\rho} = \Omega(n^{\sigma}) \qquad (0 < \sigma < \tfrac{1}{2}\rho).$$

LEMMA 4. *We can find a positive ρ so small that there are more than $C_{22}\mu^{1+\rho}/\nu$ primes p not exceeding $\mu^{1+\rho}$ such that $p-1$ is composed of primes all less than μ.*

If $p-1$ has a prime factor q not less than μ, then

$$p-1 = aq, \qquad a \leqslant \mu^{\rho}.$$

By (1) the number of values of p not exceeding $\mu^{1+\rho}$ and satisfying this equation for given a is less than

$$C_{23}\mu^{1+\rho}\prod_{p|a}\left(1-\frac{1}{p}\right)\Big/ a\left(\log\frac{\mu^{1+\rho}}{a}\right)^2 \prod_{\substack{p|a\\ p\neq 2}}\left(1-\frac{2}{p}\right)$$
$$< C_{24}\mu^{1+\rho}\prod_{p|a}\left(1-\frac{1}{p}\right)\Big/ a\nu^2 \prod_{\substack{p|a\\ p\neq 2}}\left(1-\frac{2}{p}\right).$$

The sum in a

$$< \frac{C_{24}\mu^{1+\rho}}{\nu^2}\sum_{a=2}^{\mu^\rho}\prod_{p|a}\left(1-\frac{1}{p}\right)\Big/ a\prod_{\substack{p|a\\ p\neq 2}}\left(1-\frac{2}{p}\right)$$

$$< \frac{C_{24}\mu^{1+\rho}}{\nu^2}\sum_{a=2}^{\mu^\rho}\prod_{p|a}\left\{1+O\left(\frac{1}{p^2}\right)\right\}\Big/ a\prod_{\substack{p|a\\ p\neq 2}}\left(1-\frac{1}{p}\right)$$

$$< \frac{C_{25}\mu^{1+\rho}}{\nu^2}\sum_{a=2}^{\mu^\rho}\frac{1}{\phi(a)},$$

since $\prod_p \left\{1+O\left(\frac{1}{p^2}\right)\right\}$ converges.

Since*
$$\sum_{a=1}^{x}\frac{1}{\phi(a)} = \frac{315}{2\pi^4}\log x + o(\log x),$$

the sum in a is less than $C_{26}\rho\mu^{1+\rho}/\nu$,

where C_{26} is independent of ρ. This proves the lemma since the number of primes not exceeding $\mu^{1+\rho}$ is greater than $C_{20}\mu^{1+\rho}\nu^{-1}$ and

$$(C_{20}-C_{26}\rho)\frac{\mu^{1+\rho}}{\nu} > \frac{C_{27}\mu^{1+\rho}}{\nu}$$

for sufficiently small ρ.

We now proceed to our main theorem. We consider the square-free integers not exceeding n composed of the primes in Lemma 4. By Lemma 3 there are $\Omega(n^\sigma)$ of them. Clearly the ϕ of all these integers is divisible only by primes less than μ. By Lemma 2 these ϕ have only $o(n^\epsilon)$ different values. Hence, if we choose ϵ less than $\frac{1}{2}\sigma$, we have an integer m not exceeding n which can be represented $\Omega(n^{\sigma-\epsilon})$ $[> \Omega(n^{\frac{1}{2}\sigma})]$ times as the ϕ of another integer. Since $n \geqslant m$, the number of these representations is greater than m^{C_5} where $C_5 > \frac{1}{2}\sigma$, as was stated in the introduction.

* E. Landau, *Göttinger Nachr.* (1900), 177–86.