Can. J. Math., Vol. XXXIX, No. 5, 1987, pp. 1107-1122

ON THE SIEVE OF ERATOSTHENES

M. RAM MURTY AND N. SARADHA

1. Introduction. Let $\nu(n)$ denote the number of distinct prime factors of a natural number n. A classical theorem of Hardy and Ramanujan states that the normal order of $\nu(n)$ is log log n. That is, given any $\epsilon > 0$, the number of natural numbers not exceeding x which fail to satisfy the inequality

$$(1) \quad |\nu(n) - \log\log n| < \epsilon \log\log n$$

is $o(x)$ as $x \to \infty$. A very simple proof of this was subsequently given by Turán. He showed that

$$(2) \quad \sum_{n \leqq x} (\nu(n) - \log\log n)^2 = x \log\log x + O(x).$$

Then, (1) is an immediate consequence of (2). The methods of proof were subsequently generalised to treat other additive functions. The historical developments which ultimately led to the celebrated Erdös-Kac theorem are described in the monographs of Elliott [**1**].

In 1935, Erdös [**2**] proved that the number of primes p not exceeding x which fail to satisfy the inequality

$$(3) \quad |\nu(p - 1) - \log\log p| < \epsilon \log\log p$$

is $o(x/\log x)$, for any given $\epsilon > 0$. His main tool to establish (3) was Brun's sieve. If ϕ denotes the Euler ϕ function, then recently, it was established in [**4**] and [**3**], that $\nu(\phi(n))$ has normal order $1/2(\log\log n)^2$. In both papers, the Bombieri-Vinogradov theorem on primes in arithmetic progressions was invoked. In [**4**], it was shown that

$$(4) \quad \sum_{n \leqq x} \left(\nu(\phi(n)) - \frac{1}{2}(\log\log n)^2\right)^2 = o(x(\log\log x)^4)$$

as $x \to \infty$. This theorem was a consequence of a general result concerning prime divisors of various sequences, with special reference to Fourier coefficients of modular forms. By utilizing a general theorem of Kubilius and Shapiro [**3**], Erdös and Pomerance proved in [**3**] that if we let $G(x, u)$ denote the number of $n \leqq x$ satisfying the inequality

Received March 5, 1986 and in revised form January 7, 1987. The research of the first author was supported in part by NSERC grant U0237.

$$(5) \qquad \nu(\phi(n)) < \frac{1}{2}(\log\log x)^2 + \frac{u}{\sqrt{3}}(\log\log x)^{3/2},$$

then

$$(6) \qquad \lim_{x\to\infty} \frac{G(x, u)}{x} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt.$$

The purpose of this paper is to establish (3) and (4) utilising nothing more than the sieve of Eratosthenes. In the course of the proof, we do not need any information on primes in arithmetic progressions, nor do we use the elementary estimate of Tchebycheff on the number of primes less than x.

2. The sieve of Eratosthenes. Let A be any set of natural numbers $\leqq x$ and let P be a set of primes. To each prime p, let there be $\omega(p)$ distinguished residue classes (mod p). Let A_p denote the set of elements of A lying in at least one of these distinguished residue classes (mod p) and set for any natural number d,

$$A_d = \cap_{p|d} A_p.$$

As usual, we denote by $S(A, P, z)$ the number of elements of

$$A \setminus \cup_{p\leqq z, p\in P} A_p.$$

Let

$$\omega(d) = \prod_{p|d} \omega(p)$$

for each squarefree number d. We suppose that there is an X such that

$$|A_d| = \frac{X\omega(d)}{d} + R_d.$$

We also set

$$P(z) = \prod_{p\leqq z, p\in P} p.$$

μ, as usual, will denote the Möbius function and we set

$$W(z) = \prod_{p\leqq z, p\in P} \left(1 - \frac{\omega(p)}{p}\right).$$

We shall further suppose that there is a constant $C > 0$ such that $|A_d| = 0$ for $d > Cx$. For convenience, we write the iterated logarithms:

$$\log_m x = \log(\log_{m-1} x), \quad \log_1 x = \log x.$$

THEOREM 1. *Let A be a set of natural numbers $\leqq x$, satisfying*

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

(i) $|R_d| = O(\omega(d))$

(ii) $$\sum_{p \leqq z, p \in P} \frac{\omega(p) \log p}{p} \leqq \kappa \log z + O(1).$$

Then

$$S(A, P, z) = XW(z) + O\left(x(\log z)^{\kappa+1} \exp\left(-\frac{\log x}{\log z}\right)\right).$$

We will require the following lemmas.

LEMMA 1. *Let*

$$F_\omega(t, z) = \sum_{d \leqq t, d \mid P(z)} \omega(d).$$

Then,

$$F_\omega(t, z) = O\left(t(\log z)^\kappa \exp\left(-\frac{\log t}{\log z}\right)\right).$$

Proof. We utilise a classical method of Rankin. Clearly, for any $\delta > 0$, we have

$$F_\omega(t, z) \leqq \sum_{d \mid P(z)} \omega(d)(t/d)^\delta.$$

As ω is multiplicative, we deduce that

$$F_\omega(t, z) \leqq \exp\left(\delta \log t + \sum_{p \leqq z, p \in P} \omega(p) p^{-\delta}\right),$$

on applying the elementary inequality $1 + x \leqq e^x$. Setting $\delta = 1 - \eta$ and utilising the inequality $e^x \leqq 1 + xe^x$, which is valid for $x \geqq 0$, we find

$$F_\omega(t, z) \leqq t \exp\left(-\eta \log t + \sum_{p \leqq z, p \in P} \frac{\omega(p)}{p} + \eta z^n \sum_{p \leqq z, p \in P} \frac{\omega(p) \log p}{p}\right).$$

By condition (ii) in Theorem 1, we find by partial summation that

$$\sum_{p \leqq z, p \in P} \frac{\omega(p)}{p} \leqq \kappa \log_2 z + O(1).$$

Thus,

$$F_\omega(t, z) \ll t \exp(-\eta \log t + \kappa \log_2 z + \kappa\eta(\log z) z^\eta).$$

Choosing $\eta = (\log z)^{-1}$, gives the desired result.

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

Remark. Let $G(x, z)$ denote the number of $n \leqq x$ all of whose prime factors are $\leqq z$. A similar method yields the estimate

$$G(x, z) \ll x(\log z) \exp\left(-\frac{\log x}{\log z}\right).$$

LEMMA 2.

$$\sum_{d|P(z), d>Cx} \frac{\omega(d)}{d} = O\left((\log z)^{\kappa+1} \exp\left(-\frac{\log x}{\log z}\right)\right).$$

Proof. Clearly,

$$\sum_{d|P(z), d>Cx} \frac{\omega(d)}{d} \ll \int_{Cx}^{\infty} \frac{F_\omega(t, z)}{t^2} dt$$

and the result now easily follows from Lemma 1.

Proof of Theorem 1. We have by the inclusion-exclusion principle,

$$S(A, P, z) = \sum_{d|P(z), d \leqq Cx} \mu(d) |A_d|$$

$$= \sum_{d|P(z), d \leqq Cx} \mu(d) \frac{X\omega(d)}{d} + O(F_\omega(Cx, z))$$

by utilising (i). Thus, by Lemmas 1 and 2, we find easily

$$S(A, P, z) = XW(z) + O\left(x(\log z)^{\kappa+1} \exp\left(-\frac{\log x}{\log z}\right)\right)$$

as desired.

COROLLARY 1. *Let* $\pi(x)$ *denote the number of primes* $p \leqq x$. *Then*,

$$\pi(x) \ll \frac{x}{\log x} \log \log x.$$

Proof. Set

$$\log z = \frac{\log x}{3 \log \log x}$$

in Theorem 1.

Remark. The elementary reasoning used above actually yields a better result than stated, but to keep the exposition simple we have chosen our parameters in the simplest possible manner. More precisely, if

$$2\kappa z \log z > \log x,$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

then

$$(\dagger)\quad S(A, P, z) = XW(z) + O\left(x(\log z)^{\kappa+1} \exp\left(-\frac{\log x}{2 \log z} \log\left(\frac{\log x}{\log z}\right)\right)\right).$$

This is obtained by setting

$$\eta = \frac{1}{\log z} \log\left(\frac{\log x}{2\kappa \log z}\right) < 1$$

in Lemmas 1 and 2. This would then yield the result

$$\pi(x) \ll \frac{x \log_2 x}{\log x \log_3 x}$$

for the choice

$$\log z = \frac{\log x}{\log_2 x} \log_3 x.$$

This bound for $\pi(x)$ is better than the estimate derived from the elementary Brun's sieve.

COROLLARY 2. *For any natural number $m < x$, denote by $N(x, m)$ the number of solutions of*

$$p - 1 = qm,$$

where p and q are prime numbers $\leqq x$. Then for some absolute constant $B > 0$, we have

$$N(x, m) \leqq \frac{Bx(\log_2(x/m))^2}{\phi(m)(\log(x/m))^2}.$$

Proof. Let A denote the set of natural numbers $\leqq x/m$. Clearly, $N(x, m) - z^2$ is less than the number of solutions of $b - 1 = am$, $a \in A$, where a and b are free of prime factors $\leqq z$. Thus we count the number of $a \in A$ such that a and $am + 1$ are free of prime factors $\leqq z$. Let P denote the set of primes $p \leqq z$. Then for each prime $p \in P$, $(p, m) = 1$, we have $\omega(p) = 2$. If $p|m$, then $\omega(p) = 1$. For each $d|P(z)$,

$$|A_d| = \frac{\omega(d)x}{md} + R_d,$$

where $|R_d| \leqq \omega(d)$. Moreover, $|A_d| = 0$ for $d > x/m$. Thus, with $X = x/m$, we apply Theorem 1 and obtain the desired result with

$$\log z = \frac{\log(x/m)}{5 \log_2(x/m)}.$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

COROLLARY 3. *Let $E > 0$ and let $\pi(x, q)$ denote the number of primes $p \leqq x$, $p \equiv 1 \pmod q$. Then*

$$\pi(x, q) \ll \frac{x \log_2 x}{\phi(q) \log x}$$

uniformly for $q \leqq \log^E x$, where the implied constant depends only on E.

Proof. Let A consist of $n \leqq x$, $n \equiv 1 \pmod q$ and let P be the set of primes $p \leqq z$ where

$$\log z = \frac{\log x}{(E + 3) \log \log x}.$$

Then,

$$S(A, P, z) \ll \frac{x \log_2 x}{\phi(q) \log x} + O\left(\frac{x}{\log^{E+1} x}\right).$$

But

$$\pi(x, q) \leqq z + S(A, P, z) \ll \frac{x \log_2 x}{\phi(q) \log x}.$$

COROLLARY 4. *Uniformly for $q \leqq \log^E x$,*

$$\sum_{\substack{p \leqq x \\ p \equiv 1 (\mathrm{mod} q)}} \frac{1}{p} \ll \frac{(\log \log x)^2}{\phi(q)}.$$

Proof. This follows easily from Corollary 3 and partial summation.

Remark. Both Corollaries 3 and 4 can be refined in view of our earlier remark concerning the choice of z. Indeed, it is clear that (†) implies that

$$\pi(x, q) \ll \frac{x \log_2 x}{\phi(q)(\log x) \log_3 x}$$

uniformly for $q \leqq \log^E x$. This would yield the corresponding result in Corollary 4:

$$\sum_{\substack{p \leqq x \\ p \equiv 1 (\mathrm{mod} q)}} \frac{1}{p} \ll \frac{(\log \log x)^2}{\phi(q) \log_3 x},$$

uniformly for $q \leqq \log^E x$. We will need this slight improvement of Corollary 4 in Section 4 to prove that $\nu(\phi(n))$ has normal order

$$\frac{1}{2}(\log \log n)^2.$$

It is interesting that crude estimates suffice for establishing the normal order of $\Omega(\phi(n))$.

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

3. The normal order of $\nu(p-1)$. In this section, we shall show that the normal number of prime factors of $p-1$ is $\log\log p$. In this, we essentially follow Erdös [**2**], except in the use of Brun's sieve. Since we need a more refined version of his result, we state the results explicitly.

LEMMA 3. *For $k \geqq 1$, let $f_k(x)$ denote the number of primes $p \leqq x$ such that $\nu(p-1) = k$. Then, for any $E > 0$,*

$$f_k(x) < \frac{Bx}{(k-1)!(\log x)^2}(\log_2 x + D)^{k+4} + O\left(\frac{x}{(\log x)^E}\right),$$

where B and D are constants and the constant implied by the O symbol does not depend on k.

Remark. Let $\Omega(n)$ denote the number of prime factors of n counted with multiplicity. Then, it will be apparent from the proof that the number of primes $p \leqq x$ such that $\Omega(p-1) = k$ satisfies a similar estimate.

Proof. Let M_1 denote the set of all primes $p \leqq x$ such that all of the prime factors of $p-1$ are less than

$$y = \exp\left(\frac{\log x}{(E+1)\log\log x}\right).$$

Then, by the remark after Lemma 1, the number of such primes is

$$O\left(\frac{x}{(\log x)^E}\right).$$

Thus, we can suppose that some prime factor of $p-1$ is greater than y. If we let M_2 denote the set of all primes $p \leqq x$ such that the square of the largest prime factor of $p-1$ divides $p-1$, and $p \notin M_1$, then the cardinality of M_2 is bounded by

$$\sum_{m>y} \frac{x}{m^2} = O\left(\frac{x}{(\log x)^E}\right).$$

Now let $p \notin M_1 \cup M_2$ and suppose that $\nu(p-1) = k$, then

$$p - 1 = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k},$$

where $p_1 < p_2 < \ldots < p_k$, $\alpha_k = 1$. Thus,

$$p - 1 = p_k n_k,$$

where n_k is a natural number $< x/y$ such that $\nu(n_k) = k-1$. For a fixed n_k, the number of primes $p < x$, satisfying $p - 1 = qn_k$, is by Corollary 2,

$$\frac{Bx(\log\log(x/n_k))^2}{\phi(n_k)(\log(x/n_k))^2},$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

and therefore,

$$f_k(x) \ll \frac{x}{(\log x)^E} + Bx \sum_{\substack{n<x/y \\ \nu(n)=k-1}} \frac{(\log_2(x/n))^2}{\phi(n)\log^2(x/n)}$$

$$\ll \frac{x}{(\log x)^E} + \frac{x(\log_2 x)^4}{\log^2 x} \sum_{\substack{n<x/y \\ \nu(n)=k-1}} \frac{1}{\phi(n)}.$$

Since

$$\sum_{\substack{n<z \\ \nu(n)=k-1}} \frac{1}{\phi(n)} < \frac{1}{(k-1)!}\left(\sum_{p<z} \frac{1}{p-1} + c_1\right)^{k-1}$$

for some constant c_1, it follows by elementary estimates that

$$f_k(x) \ll \frac{x(\log_2 x + c_1)^{k+3}}{(k-1)!\log^2 x} + \frac{x}{(\log x)^E},$$

where the implied constant does not depend on k. This completes the proof of Lemma 3.

LEMMA 4. *For* $\delta > 1$,

(i) $$\sum_{k>\delta a} \frac{a^k}{k!} \leqq \delta^{-\delta a} e^{\delta a}$$

and for $\delta < 1$,

(ii) $$\sum_{k<\delta a} \frac{a^k}{k!} \leqq \delta^{-\delta a} e^{\delta a}(\delta a).$$

Proof. We have

$$\sum_{k>\delta a} \frac{a^k}{k!} = \sum_{k>\delta a} \frac{(\delta a)^k}{k!}\delta^{-k} \leqq \delta^{-\delta a} e^{\delta a}$$

for $\delta > 1$. This proves (i). To prove (ii), we utilise the elementary inequality

$$\log k! > \int_1^k \log t dt > k\log k - k$$

and the fact that $a^k/k!$ is an increasing function of k for $k < \delta a$ when $\delta < 1$, to deduce that

$$\sum_{k<\delta a} \frac{a^k}{k!} \leqq (\delta a)\delta^{-\delta a} e^{\delta a}$$

which establishes (ii).

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

THEOREM 2. *Let* $0 < \epsilon < 1$. *The number of primes* $p \leqq x$ *which fail to satisfy the inequality*

$$|\nu(p-1) - \log\log p| < \epsilon \log\log p$$

is

$$O\left(\frac{x}{(\log x)^{1+\epsilon^2}}\right).$$

Proof. We need to estimate the sums

$$\sum_{k>(1+\epsilon)\log_2 x} f_k(x)$$

and

$$\sum_{k<(1-\epsilon)\log_2 x} f_k(x).$$

To estimate the first sum, we first note that $f_k(x) = 0$ if $k > \log x$, since $\nu(n) < \log n$. Thus, by Lemmas 3 and 4, we obtain

$$\sum_{k>(1+\epsilon)\log_2 x} f_k(x) \ll \frac{x(\log_2 x)^4}{\log^2 x} \sum_{k>(1+\epsilon)\log_2 x} \frac{(\log_2 x + c_1)^{k-1}}{(k-1)!}$$

$$\ll \frac{x(\log_2 x)^4}{(\log x)^{1+\epsilon^2}}.$$

The second sum is estimated similarly.

In order to deduce that the normal order of $\nu(p-1)$ is $\log\log p$, we need to show that

$$\pi(x) \gg \frac{x}{\log x}.$$

We have the following elementary result.

LEMMA 5.

$$\pi(x) \gg \frac{x}{\log x}.$$

Proof. Utilising

$$\sum_{p\leqq x} \frac{1}{p} = \log\log x + c_1 + O\left(\frac{1}{\log x}\right),$$

for some constant c_1, we deduce that for some $\eta > 0$, the inequality

$$\sum_{\eta x<p<x} \frac{1}{p} \gg \frac{1}{\log x}$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

holds. Thus, we obtain the inequality

$$\frac{\pi(x)}{\eta x} \geqq \sum_{\eta x < p < x} \frac{1}{p} \gg \frac{1}{\log x}$$

which gives us the desired inequality.

Hence, Theorem 1 and Lemma 5 imply that for almost all primes, the number of prime factors of $p-1$ is $\log\log p$. We can proceed to obtain an estimate for the variance. As this is routine and straightforward, we suppress the details. The method is illustrated in the following lemma which we need in the next section.

LEMMA 6. *As* $x \to \infty$,

(i) $$\sum_{p \leqq x} \nu(p-1) \sim \pi(x) \log\log x,$$

(ii) $$\sum_{p \leqq x} \nu^2(p-1) \ll \pi(x)(\log\log x)^2.$$

Proof. We have

$$\sum_{p \leqq x} \nu(p-1) \ll (1+\epsilon)\pi(x)\log\log x + \sum_{\substack{p \leqq x \\ \nu(p-1) > (1+\epsilon)\log\log p}} \nu(p-1). \tag{7}$$

The latter sum is by the Cauchy-Schwartz inequality, Theorem 2, and Lemma 5,

$$\leqq \left(\frac{\pi(x)}{\log^{\delta} x}\right)^{1/2} \left(\sum_{p \leqq x} \nu^2(p-1)\right)^{1/2}.$$

Moreover,

$$\sum_{p \leqq x} \nu^2(p-1) \ll \pi(x)(\log\log x)^2 + \sum_{\substack{p \leqq x \\ \nu(p-1) > 2K\log\log x}} \nu^2(p-1).$$

For K sufficiently large, we deduce by Lemmas 3 and 4 that

$$\sum_{k > 2K\log\log x} f_k(x) \ll \frac{x}{\log^K x}$$

so that

$$\sum_{\substack{p \leqq x \\ \nu(p-1) > 2K\log\log x}} \nu^2(p-1) \ll \frac{x}{\log^{K-2} x}$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

since $\nu(p-1) = O(\log p)$. Again, we have used the fact that $f_k(x) = 0$ if $k > \log x$. Thus,

$$\sum_{p \leqq x} \nu^2(p-1) \ll \pi(x)(\log\log x)^2.$$

This proves (ii). Therefore, the second sum in (7) is found to be

$$o(\pi(x)\log\log x)$$

as $x \to \infty$. This completes the proof of Lemma 6, since the lower bound asymptotic formula is similarly deduced.

Let $\Omega(n)$ denote the total number of prime factors of n, counted with multiplicity, defined earlier. It is clear that the above methods yield a corresponding result for $\Omega(p-1)$. We state this for future reference.

LEMMA 7. *As* $x \to \infty$,

(i) $$\sum_{p \leqq x} \Omega(p-1) \sim \pi(x)\log\log x,$$

(ii) $$\sum_{p \leqq x} \Omega^2(p-1) \ll \pi(x)(\log\log x)^2.$$

4. The normal order of $\nu(\phi(n))$. We begin by showing that $\Omega(\phi(n))$ has normal order $1/2(\log\log n)^2$.

LEMMA 8.

$$\sum_{p \leqq x} \frac{\Omega(p-1)}{p} \sim \frac{1}{2}(\log\log x)^2.$$

Proof. By Lemma 7, the sum is by partial summation,

$$\sim \int_2^x \frac{\pi(t)\log\log t}{t^2}\,dt.$$

But

$$\sum_{p \leqq x} \frac{\log\log p}{p} \sim \int_2^x \frac{\pi(t)\log\log t}{t^2}\,dt.$$

Moreover,

$$\sum_{p \leqq x} \frac{\log\log p}{p} \sim \frac{1}{2}(\log\log x)^2,$$

which is derived by partial summation from the elementary result

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

$$\sum_{p \leqq x} \frac{1}{p} = \log \log x + c_1 + O\left(\frac{1}{\log x}\right).$$

The desired result now follows immediately.

THEOREM 3. *As* $x \to \infty$,

$$\sum_{n \leqq x} (\Omega(\phi(n)) - \frac{1}{2}(\log \log x)^2)^2 = o(x(\log \log x)^4).$$

Proof. Let us define

$$f(n) = \sum_{p|n} \Omega(p - 1).$$

Then,

$$f(n) \leqq \Omega(\phi(n)) \leqq f(n) + \Omega(n).$$

In view of (2), it therefore suffices to show that

$$\sum_{n \leqq x} \left(f(n) - \frac{1}{2}(\log \log x)^2\right)^2 = o(x(\log \log x)^4),$$

since the inequality

$$(a + b)^2 \leqq 2(a^2 + b^2)$$

implies that

$$\sum_{n \leqq x} (\Omega(\phi(n)) - \frac{1}{2}(\log \log x)^2)^2$$

$$\ll \sum_{n \leqq x} \left(f(n) - \frac{1}{2}(\log \log x)^2\right)^2 + \sum_{n \leqq x} (\Omega(\phi(n)) - f(n))^2.$$

To this end, we note that

$$\sum_{n \leqq x} f(n) = x \sum_{p \leqq x} \frac{\Omega(p - 1)}{p} + O\left(\sum_{p \leqq x} \Omega(p - 1)\right).$$

Since

$$\sum_{p \leqq x} \Omega(p - 1) \ll \sum_{n \leqq x} \Omega(n) \ll x \log \log x,$$

by elementary estimates, we deduce from Lemma 8 that

$$\sum_{n \leqq x} f(n) \sim \frac{1}{2} x(\log \log x)^2.$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

Also,

$$\sum_{n\leqq x} f^2(n) = \sum_{n\leqq x} \sum_{p,q|n} \Omega(p-1)\Omega(q-1),$$

where p and q denote primes. Then,

$$\sum_{n\leqq x} f^2(n) \leqq x \sum_{pq\leqq x} \frac{\Omega(p-1)\Omega(q-1)}{pq} + x \sum_{p\leqq x} \frac{\Omega^2(p-1)}{p}.$$

Noting that

$$\left(\sum_{p\leqq\sqrt{x}} \frac{\Omega(p-1)}{p}\right)^2 \leqq \sum_{pq\leqq x} \frac{\Omega(p-1)\Omega(q-1)}{pq} \leqq \left(\sum_{p\leqq x} \frac{\Omega(p-1)}{p}\right)^2$$

we deduce from Lemma 8 that

$$\sum_{pq\leqq x} \frac{\Omega(p-1)\Omega(q-1)}{pq} \sim \frac{1}{4}(\log\log x)^4.$$

Moreover, by Lemma 7 (ii) and partial summation, we deduce that

$$\sum_{p\leqq x} \frac{\Omega^2(p-1)}{p} \ll \int_2^x \frac{\pi(t)(\log\log t)^2}{t^2}\,dt$$

$$\sim \sum_{p\leqq x} \frac{(\log\log p)^2}{p} \ll (\log\log x)^3.$$

It therefore follows from the above that

$$\sum_{n\leqq x} f^2(n) \leqq \frac{1}{4}x(\log\log x)^4 + o(x(\log\log x)^4)$$

as $x \to \infty$. The assertion of the theorem follows immediately from this.

We can now prove:

THEOREM 4. *$\nu(\phi(n))$ has normal order $1/2(\log\log n)^2$.*

Proof. We first note that $\nu(\phi(n)) \leqq \Omega(\phi(n))$ and so

$$\nu(\phi(n)) \leqq (1+\epsilon)\frac{1}{2}(\log\log n)^2$$

for almost all n. To establish the corresponding lower bound, we first note that if

$$\Omega(\phi(n)) - \nu(\phi(n)) \geqq 1,$$

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

then there must be a prime q such that $q^2|\phi(n)$. If $q > y$, the number of $n \leqq x$ such that $q^2|\phi(n)$ is

$$(8) \quad \leqq \sum_{q>y} \frac{x}{q^3} + \sum_{q>y} \sum_{\substack{p \leqq x \\ q^2|p-1}} \frac{x}{p} + \sum_{q>y} \sum_{\substack{p \leqq x \\ q|p-1, q|p'-1}} \frac{x}{pp'}.$$

The first sum is clearly bounded by

$$\frac{x}{y^2}.$$

To handle the second sum, we write it as

$$(9) \quad \sum_{q>\log^E x} + \sum_{y<q<\log^E x}$$

and handle the penultimate sum in a trivial way:

$$\sum_{q>\log^E x} \sum_{\substack{p \leqq x \\ q^2|p-1}} \frac{x}{p} \ll \sum_{q>\log^E x} \sum_{t \leqq x} \frac{x}{q^2 t} \ll \frac{x}{\log^{E-1} x}.$$

For the last sum in (9), we use Corollary 4 of Theorem 1 to obtain that the two sums in (9) are

$$\ll \frac{x}{\log^{E-1} x} + \frac{x(\log\log x)^2}{y}.$$

Similarly, the third sum in (8) is seen to be

$$\ll \frac{x}{\log^{E-2} x} + \frac{x(\log\log x)^4}{y}.$$

Choosing $E = 3$ and $y = (\log\log x)^5$, we deduce that for almost all $n \leqq x$, $q^2|\phi(n)$ implies that $q \leqq y$. Therefore, if we define $\Omega_y(n)$ to be the number of prime powers q^a dividing n such that $q < y$, then we have proved that for almost all n,

$$\nu(\phi(n)) \geqq \Omega(\phi(n)) - \Omega(n) - \sum_{p|n} \Omega_y(p-1).$$

Since $\Omega(n)$ has normal order $\log\log n$, to prove the theorem, it suffices to show that for $\eta > 0$,

$$\sum_{p|n} \Omega_y(p-1) \leqq \eta(\log\log n)^2$$

for almost all n. To this end, we have

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

$$\sum_{n\leqq x}\sum_{p\mid n}\Omega_y(p-1) \leqq x\sum_{p\leqq x}\frac{\Omega_y(p-1)}{p}$$

$$\ll x\sum_{q\leqq y}\sum_{\substack{p\leqq x\\ q^a\mid p-1}}\frac{1}{p}$$

$$\ll x\sum_{\substack{q<y\\ q^a<\log^3 x}}\sum_{\substack{p\leqq x\\ q^a\mid p-1}}\frac{1}{p}+x\sum_{\substack{q<y\\ q^a\geqq\log^3 x}}\sum_{\substack{p\leqq x\\ q^a\mid p-1}}\frac{1}{p}.$$

The last sum is estimated in a trivial way, as before, and is seen to be $O(x)$. By the remark after Corollary 4, the penultimate sum is seen to be

$$\ll x\frac{(\log\log x)^2}{\log_3 x}\log_4 x.$$

Thus, the inequality

$$\sum_{p\mid n}\Omega_y(p-1) > \eta(\log\log n)^2$$

can hold for at most $o(x)$ numbers $n \leqq x$. Hence, for almost all n,

$$\Omega(\phi(n)) - \eta(\log\log n)^2 \leqq \nu(\phi(n)) \leqq \Omega(\phi(n)).$$

Theorem 3 now completes the proof.

5. Concluding remarks. In proving (6), Erdös and Pomerance [**3**] utilise a general theorem of Kubilius and Shapiro (see [**1**]). This theorem states that for any strongly additive real-valued function f, (that is, $f(p^a) = f(p)$),

$$\lim_{x\to\infty} G_f(x, u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-t^2/2}dt$$

where

$$G_f(x, u) = \text{card}\{n \leqq x : f(n) - A(x) \leqq uB(x)\},$$

$$A(x) = \sum_{p\leqq x}\frac{f(p)}{p},$$

$$B(x)^2 = \sum_{p\leqq x}\frac{f^2(p)}{p},$$

provided that for each $\epsilon > 0$,

https://doi.org/10.4153/CJM-1987-056-8 Published online by Cambridge University Press

$$\sum_{\substack{p \leqq x \\ f(p) > \epsilon B(x)}} \frac{f^2(p)}{p} = o(B^2(x))$$

as $x \to \infty$. It is possible to verify the above condition utilising only the sieve of Eratosthenes (Theorem 1) following the methods of the previous sections. This would then give a proof of (6) independent of the Bombieri-Vinogradov theorem.

It would be highly interesting to derive the "modular analogues" of these results. For instance, in [**4**], it was proved that

$$\sum_{\substack{p \leqq x \\ \tau(p) \neq 0}} (\nu(\tau(p)) - \log\log p)^2 \ll \frac{x \log_2 x}{\log x},$$

where τ denotes the Ramanujan τ function, assuming a non-abelian analogue of the Bombieri-Vinogradov theorem. This latter hypothesis is a consequence of the generalised Riemann hypothesis for the Dedekind zeta functions. The "modular analogue" was established in [**5**]. Therefore, it would be of exceeding importance if the methods of this paper could be extended to treat Fourier coefficients of modular forms.

Acknowledgements. We would like to thank E. Fouvry and the referee for their comments on an earlier version of this paper.

References

1. P. D. T. A. Elliott, *Probabilistic number theory I and II* (Springer Verlag, New York, 1980).

2. P. Erdös, *On the normal number of prime factors of p − 1 and some related problems concerning the Euler's φ function*, Quarterly Journal of Mathematics *6* (1935), 205-213.

3. P. Erdös and C. Pomerance, *On the normal number of prime factors of φ(n)*, Rocky Mountain Journal *15* (1985), 343-352.

4. M. Ram Murty and V. Kumar Murty, *Prime divisors of Fourier coefficients of modular forms*, Duke Math. Journal *51* (1984), 57-76.

5. ——— *An analogue of the Erdös-Kac theorem for Fourier coefficients of modular forms*, Indian Journal of Pure and App. Math. *15* (1984), 1090-1101.

McGill University,
Montréal, Québec;
Concordia University,
Montréal, Québec