ACTA ARITHMETICA
LXXXVI.2 (1998)

On the number of elements with maximal order in the multiplicative group modulo n

by

Shuguang Li (Athens, Ga.)

1. Introduction. A *primitive root modulo the prime* p is any integer a coprime to p such that its exponent modulo p is $p-1$. There are totally $\phi(p-1)$ primitive roots modulo p in $[1,p]$, where ϕ is Euler's function. It is a natural problem to consider the fraction $\phi(p-1)/(p-1)$, which is the proportion of non-zero residues mod p which are primitive roots.

Trivially, we have $0 < \phi(p-1)/(p-1) \le 1/2$. For any real numbers $x \ge 2$ and u, let

$$D_\pi(x,u) = \frac{1}{\pi(x)} \sum_{\substack{p\le x \\ \phi(p-1)/(p-1)\le u}} 1,$$

where $\pi(x)$ is the number of primes up to x. In 1974, P. D. T. A. Elliott [2] proved that the limit

$$\lim_{x\to\infty} D_\pi(x,u) = D_\pi(u)$$

exists for all real numbers u. The function $D_\pi(u)$ is continuous and is strictly increasing on the interval $[0,1/2]$. I. J. Schoenberg [12] had earlier considered the distribution problem for $\phi(n)/n$. He proved the existence of $\lim_{x\to\infty} x^{-1}\sum_{n\le x,\,\phi(n)/n\le u} 1$ for any real number u, and the continuity of the limit as a function of u.

The concept of primitive root modulo a prime can be generalized. This was done by R. D. Carmichael [1]. He defined a "primitive λ-root modulo n" as any integer coprime to n and having the maximal exponent modulo n. Thus a primitive root for a prime p is a primitive λ-root modulo p. He also found the following properties (see [1, 7]) of the maximal exponent modulo n, denoted by $\lambda(n)$:

(i) $\lambda(p^e) = \phi(p^e)$ *for all primes* p *and* $e \ge 1$ *except* $p = 2$ *and* $e > 2$ *in which case we have* $\lambda(2^e) = \phi(2^e)/2 = 2^{e-2}$.

1991 *Mathematics Subject Classification*: 11A27, 11B05, 11N36, 11N37.

(ii) $\lambda(n) = \operatorname{lcm}_{p^e \| n}\{\lambda(p^e)\}$.

(iii) *Let a be an integer coprime to n and let $l_a(n)$ be the exponent of a modulo n. Then $l_a(n) \mid \lambda(n)$.*

Thus, in this notation, an integer a coprime to n is a primitive λ-root modulo n if and only if $l_a(n) = \lambda(n)$. In [3] one can find results concerning the size of the function $\lambda(n)$.

Let $R(n)$ be the number of primitive λ-roots modulo n in $[1, n]$. In this paper we investigate the distribution of the values of $R(n)/\phi(n)$, in analogy with that of $\phi(p-1)/(p-1)$ considered by Elliott. That is, if we define

$$D(x,u) = \frac{1}{x} \sum_{\substack{n \le x \\ R(n)/\phi(n) \le u}} 1$$

for any real numbers $x > 0$ and u, what can we say about $D(x,u)$? In particular, does $\lim_{x\to\infty} D(x,u)$ exist for all u?

Note that the fraction $R(n)/\phi(n)$ represents the proportion of residue classes mod n that are primitive λ-roots to the total number of residue classes coprime to n. We trivially have $0 < R(n)/\phi(n) \le 1$. That this inequality is nearly best possible is contained in the following result.

THEOREM 1. *We have*

$$\overline{\lim_{n\to\infty}}\, R(n)/\phi(n) = 1, \qquad \underline{\lim_{n\to\infty}}\, R(n)/(\phi(n)/\ln\ln n) = e^{-\gamma},$$

where γ is Euler's constant.

The function $D(x,u)$ is only interesting for $0 < u < 1$. Perhaps surprisingly, we show that there are values of $u \in (0,1)$ where $\lim_{x\to\infty} D(x,u)$ does not exist. To attack the question we wish we could take a more natural approach, say by working with the first moment of $R(n)/\phi(n)$ or $R(n)/n$. However these functions are not multiplicative as the function $\phi(n)/n$ is, and thus the methods used by Elliott and Schoenberg do not appear to work.

Thanks to Pomerance's suggestion, we turn our attention to the first moment of $\ln(R(n)/\phi(n))$ instead. Though this function is also not multiplicative, it can be approximated by a comparatively simple sum over prime factors of $\lambda(n)$. Adopting the notation $\ln_k x$, suggested by John Selfridge, for the k-fold iteration of the natural logarithm of x, we will prove our principal results as indicated in the next two theorems.

THEOREM 2. *The maximal order of the function*

$$\frac{1}{x} \sum_{n \le x} |\ln(R(n)/\phi(n))|$$

is less than or equal to $c \ln_5 x$ *for some constant* $c > 0$, *and greater than or equal to* $c' \ln_6 x$ *for another constant* $c' > 0$. *On the other hand, the minimal order of the function is less than or equal to a constant.*

As a consequence of Theorem 2 we have

THEOREM 3. *There is a positive constant* c *and an unbounded set of numbers* x *such that for each* u *in* $(0,1)$, *we have* $D(x,u) \le c/|\ln u|$. *On the other hand, there are positive constants* δ, b *and an unbounded set of numbers* x *with* $D(x,(\ln_5 x)^{-b}) \ge \delta$. *Thus for some positive constant* u_0, $\lim_{x\to\infty} D(x,u)$ *does not exist for all* u *with* $0 < u < u_0$.

It follows immediately from Theorem 3 that the sequence of distribution functions $D(n,u)$ does not converge weakly. Suppose that it does. Let $D(u)$ be the distribution function to which $D(n,u)$ converges weakly [4]. Then $D(u)$ is discontinuous in $(0,u_0)$ by Theorem 3, which contradicts the fact that the set of discontinuities of $D(u)$ is countable—a well-known property of monotone functions.

In a forthcoming paper [8] we will prove, by a more complicated argument, that the constant δ in Theorem 3 can be taken as anything less than 1. This result will be shown to be relevant to the study of the integers n for which a fixed integer a is a primitive λ-root, in analogy with Artin's conjecture for primes.

The author would like to take this opportunity to express his heartfelt thanks to Carl Pomerance for patient advice concerning problems in this paper and remarkable comments. Without Pomerance's help this project would not have survived. The author is indebted to Andrew Granville and Vsevolod Lev for their suggestions. The author would also like to thank the referee for a comment regarding Theorem 2.1.

2. The closed form for $R(n)$ and a few properties of $R(n)$. Throughout this paper we always use p and q to represent primes and k, m, n to represent natural numbers. We give an explicit formula for $R(n)$ in the next theorem, to which a different approach can be found in [9]. Then we will study some deeper features of the function.

THEOREM 2.1. *Let* $\Delta_q(n) := \#\{\text{prime } p : p^e \parallel n \text{ and } q^v \mid \lambda(p^e)\}$ *for a prime* q *with* $q^v \parallel \lambda(n)$, *except the case* $2^3 \parallel n$ *and* $2 \parallel \lambda(n)$, *when* $\Delta_2(n) := 1 + \#\{\text{prime } p : p \mid n\}$. *Then*

$$R(n) = \phi(n) \prod_{q \mid \lambda(n)} \left(1 - \frac{1}{q^{\Delta_q(n)}}\right).$$

Proof. For any integer $m \ge 1$ let $N(m)$ be the number of elements of order m in $(\mathbb{Z}/n\mathbb{Z})^*$. We claim that $N(m)$ is a multiplicative function.

Suppose $(k,l)=1$. Clearly the product of two elements, of orders k and l respectively, has order kl. Conversely, every element of order kl can be written uniquely as a product of an element of order k and an element of order l. Thus $N(kl)=N(k)N(l)$.

As a consequence we have $N(\lambda(n))=\prod_{q^v\|\lambda(n)}N(q^v)$. To compute $N(q^v)$ it is convenient to factor $(\mathbb{Z}/n\mathbb{Z})^*$ into a direct sum of cyclic groups.

Let C_m denote a cyclic group of order m. Let $n=p_1^{e_1}\dots p_r^{e_r}\cdot 2^e$ where $p_1,\dots,p_r$ are distinct odd primes and $e_1,\dots,e_r$ are positive integers. By the Chinese remainder theorem we have

$$(\mathbb{Z}/n\mathbb{Z})^*\cong C_{\lambda(p_1^{e_1})}\oplus\dots\oplus C_{\lambda(p_r^{e_r})}\oplus(\mathbb{Z}/2^e\mathbb{Z})^*.$$

If $e=0$ the last summand drops off. Otherwise

$$(\mathbb{Z}/2^e\mathbb{Z})^*\cong\begin{cases}C_{\lambda(2^e)} & \text{if } e=1 \text{ or } 2,\\ C_{\lambda(2^e)}\oplus C_2 & \text{if } e\ge 3.\end{cases}$$

We have factored $(\mathbb{Z}/n\mathbb{Z})^*$ into a direct sum of $r+r'$ cyclic groups, where $r'=0$, 1 or 2. Let n_i be the order of the ith summand, so that

$$(\mathbb{Z}/n\mathbb{Z})^*\cong C_{n_1}\oplus C_{n_2}\oplus\dots\oplus C_{n_{r+r'}}.$$

The number of solutions to $x^m=1$ in C_k is (m,k), so the number of solutions to $x^{q^v}=1$ in $(\mathbb{Z}/n\mathbb{Z})^*$ is given by $\prod_{i=1}^{r+r'}(q^v,n_i)$. Similarly $x^{q^{v-1}}=1$ has $\prod_{i=1}^{r+r'}(q^{v-1},n_i)$ solutions in $(\mathbb{Z}/n\mathbb{Z})^*$. Thus

$$N(q^v)=\prod_{i=1}^{r+r'}(q^v,n_i)-\prod_{i=1}^{r+r'}(q^{v-1},n_i).$$

Note that the second product is $q^{-\Delta_q(n)}$ times the first product, since

$$(q^{v-1},n_i)=\begin{cases}(q^v,n_i) & \text{if } q^v\nmid n_i,\\ q^{-1}(q^v,n_i) & \text{if } q^v\mid n_i.\end{cases}$$

Thus, $N(q^v)=(1-q^{-\Delta_q(n)})\prod_{i=1}^{r+r'}(q^v,n_i)$ and so

$$N(\lambda(n))=\prod_{q^v\|\lambda(n)}N(q^v)=\prod_{q\mid\lambda(n)}\left(1-\frac{1}{q^{\Delta_q(n)}}\right)\cdot\prod_{q^v\|\lambda(n)}\prod_{i=1}^{r+r'}(q^v,n_i).$$

But for the double product we have

$$\prod_{q^v\|\lambda(n)}\prod_{i=1}^{r+r'}(q^v,n_i)=\prod_{i=1}^{r+r'}\prod_{q^v\|\lambda(n)}(q^v,n_i)=\prod_{i=1}^{r+r'}(\lambda(n),n_i)=\prod_{i=1}^{r+r'}n_i=\phi(n),$$

which completes the proof.

We note that $\Delta_q(n)\ge 0$ and equality holds if and only if $\lambda(n)$ is not divisible by q. Let us define the function $r(n):=R(n)/\phi(n)$. Then $r(n)=$

$\prod_{\text{prime } q}(1-1/q^{\Delta_q(n)})$. The word "prime" will be suppressed as we always use q to denote a prime. We now prove Theorem 1 in the introduction.

Proof of Theorem 1. It is trivial that $r(n) \le 1$. Let us prove that $\overline{\lim}_{n\to\infty} r(n) = 1$. Let x be any large number and ε be a small positive constant. Let $\mathcal{B} = \{\text{primes } p \le x : \gcd(p-1, P(x^\varepsilon)) = 1 \text{ and } p \equiv 3 \bmod 4\}$, where $P(z) = \prod_{2<p\le z} p$ for any z. Then there is a positive constant δ such that for all sufficiently large x we have

$$\#\mathcal{B} \ge \delta \frac{x}{(\ln x)^2}.$$

This result can be obtained by applying Theorem 7.4 of [5] to sieve the set $\mathcal{A} = \{p-1 : p \le x \text{ and } p \equiv 3 \bmod 4\}$ with the set $\mathcal{P} = \{\text{primes } p > 2\}$, taking $\kappa = 1$, $\alpha = 1/2$ and $z = x^\varepsilon$. Here we can choose any $\varepsilon < 1/4$.

If $p \in \mathcal{B}$ and q is a prime factor of $p-1$ other than 2, then $q > x^\varepsilon$. But $p \le x$. Thus $p-1$ has at most $1/\varepsilon$ odd prime factors, counting multiplicity. Choose $[\ln x]$ such primes $p_i \in \mathcal{B}$, $i = 1, \dots, [\ln x]$. Let $n_x = \prod_{i=1}^{[\ln x]} p_i$. Then by definition of $r(n)$,

$$r(n_x) = \prod_{q|\lambda(n_x)} \left(1 - \frac{1}{q^{\Delta_q(n_x)}}\right) \ge \left(1 - \frac{1}{2^{[\ln x]}}\right)\left(1 - \frac{1}{x^\varepsilon}\right)^{[\ln x]/\varepsilon},$$

which has limit 1 as x goes to infinity. Thus we proved $\overline{\lim}_{n\to\infty} r(n) = 1$.

To see the lower bound, first note that

$$r(n) \ge \prod_{q|\lambda(n)} \left(1 - \frac{1}{q}\right) \ge \prod_{q\le N(n)} \left(1 - \frac{1}{q}\right),$$

where $N(n)$ is chosen to be the least number such that the product of the primes up to $N(n)$ is greater than or equal to $\lambda(n)$. From prime number theory, $N(n) = (1+o(1)) \ln \lambda(n) \le (1+o(1)) \ln n$. Thus, from Mertens' theorem, $r(n) \ge (e^{-\gamma} + o(1))/\ln_2 n$.

We claim this inequality is sharp. For any given large x let $m = \prod_{q\le \ln x} q$. Then from prime number theory again, we have $\ln m \sim \ln x$. Thus $x^{1/2} \le m \le x^2$ if x is sufficiently large. By Linnik's theorem, there is a prime p such that $p \equiv 1 \bmod m$ and $p \le m^c$ for some absolute constant c. With this choice of p we have $x^{1/2} \le p \le x^{2c}$. Thus by Mertens' theorem,

$$r(p) = \prod_{q|p-1} \left(1 - \frac{1}{q}\right) \le \prod_{q\le \ln x} \left(1 - \frac{1}{q}\right) = \frac{1+o(1)}{e^\gamma \ln_2 x} = \frac{1+o(1)}{e^\gamma \ln_2 p}.$$

Send x to infinity to see that there are infinitely many such primes. So we proved our claim and our theorem.

Our goal is to study the distribution of values of the function $r(n)$. As we noted before, we can write

$$r(n)=\prod_q\Big(1-\frac{1}{q^{\Delta_q(n)}}\Big)=\exp\Big(-\sum_{q:\,\Delta_q(n)=1}\frac{1}{q}+O(1)\Big),$$

where the $O(1)$ is less than or equal to zero. It is convenient to introduce the function

$$f(n):=\sum_{q:\,\Delta_q(n)=1}\frac{1}{q}.$$

Thus, $r(n)\le e^{-f(n)}\le cr(n)$ for an absolute constant $c\ge 1$. We see that the distribution of values of $r(n)$ is dominated by its counterpart of $f(n)$. It is important to understand the behavior of the function $f(n)$. Our strategy is to study the first moment of $f(n)$, the sum $\sum_{n\le x}f(n)$. We note that

$$\sum_{n\le x}f(n)=\sum_{n\le x}\sum_{q:\,\Delta_q(n)=1}\frac{1}{q}=\sum_{q\le x}\frac{1}{q}\sum_{\substack{n\le x\\ \Delta_q(n)=1}}1, \tag{1}$$

where $q\le x$ because $\Delta_q(n)=1$ implies that $q\,|\,\lambda(n)\le n$. We close this section by showing that the contribution to this sum from the terms with $q>\ln_2 x$ is negligible.

Theorem 2.2. *As $x\to\infty$,*

$$\sum_{n\le x}\sum_{\substack{q\mid\lambda(n)\\ q>\ln_2 x}}\frac{1}{q}=O\Big(\frac{x}{\ln_3 x}\Big).$$

Let us mention the following lemma before proving Theorem 2.2.

Lemma 2.3 (see [10, 11]). *For any integer $k\ge 2$ and any $x\ge 2$,*

$$\sum_{\substack{p\le x\\ p\equiv 1 \bmod k}}\frac{1}{p}=\frac{\ln_2 x}{\phi(k)}+O\Big(\frac{\ln k}{\phi(k)}\Big)$$

where the implied constant is uniform.

Proof (of Theorem 2.2). Notice that for a prime $q\,|\,\lambda(n)$, either $q^2\,|\,n$ or $q\,|\,p-1$ for some prime $p\,|\,n$. Note that

$$\sum_{q>\ln_2 x}\frac{1}{q}\sum_{\substack{n\le x\\ q^2\mid n}}1\le x\sum_{q>\ln_2 x}\frac{1}{q^3}\ll\frac{x}{(\ln_2 x)^2\ln_3 x}.$$

Thus

$$\sum_{n\le x}\sum_{\substack{q\mid\lambda(n)\\ q>\ln_2 x}}\frac{1}{q}=\sum_{\ln_2 x<q\le x}\frac{1}{q}\sum_{\substack{p\le x\\ p\equiv 1 \bmod q}}\sum_{\substack{n\le x\\ p\mid n}}1+O\Big(\frac{x}{\ln_2^2 x}\Big). \tag{2}$$

We have

$$\sum_{q>y}\frac{1}{q^2}=\frac{1}{y\ln y}(1+o(1)) \quad\text{and}\quad \sum_{q>y}\frac{\ln q}{q^2}=\frac{1}{y}(1+o(1)).$$

Using these facts and Lemma 2.3 we have

$$\sum_{q>\ln_2 x}\frac{1}{q}\sum_{\substack{p\le x\\ p\equiv 1 \bmod q}}\sum_{\substack{n\le x\\ p\mid n}}1\le x\sum_{q>\ln_2 x}\frac{1}{q}\sum_{\substack{p\le x\\ p\equiv 1 \bmod q}}\frac{1}{p}$$
$$=x\sum_{q>\ln_2 x}\frac{1}{q}\Big(\frac{\ln_2 x}{q-1}+O\Big(\frac{\ln q}{q-1}\Big)\Big)=O\Big(\frac{x}{\ln_3 x}\Big).$$

Our theorem follows by substituting the above estimate in (2).

3. The first moment of $f(n)$. When $q\le \ln_2 x$ the inner sum of (1) has the following bounds.

THEOREM 3.1. *There exist positive constants c_1 and c_2 so that, for all large numbers x and any prime $q\le \ln_2 x$, we have*

$$\frac{c_2 x}{q^{1-\{\ln_3 x/\ln q\}}}\le \sum_{\substack{n\le x\\ \Delta_q(n)=1}}1\le \frac{c_1 x}{q^{\|\ln_3 x/\ln q\|}},$$

where $\{y\}$ denotes the fractional part of the real number y and $\|y\|$ the minimal distance from y to the integers, that is, $\|y\|=\min_{n\in\mathbb{Z}}\{|y-n|\}$.

Before proving Theorem 3.1 let us look at its consequences.

THEOREM 3.2. *For the positive constant c_1 in Theorem 3.1 we have*

$$\sum_{n\le x}f(n)\le 2c_1 x\sum_{q\le\ln_2 x}\frac{1}{q^{1+\|\ln_3 x/\ln q\|}}$$

for all large x.

Proof. The theorem follows immediately by combining Theorems 2.2 and 3.1, and the observation that $2^{1+\|\ln_3 x/\ln 2\|}\le 2\sqrt{2}$ in (1).

Although we can bound the first moment of $f(n)$ from below in a similar way, we are not able to use these estimates to get the correct order of magnitude. We would need this to show there are many values of n for which $f(n)$ is large. We get around this problem by introducing a smaller

function $\widetilde{f}(n)$ for which we are able to find the correct order of magnitude for its average order. Let

$$\widetilde{f}(n) := \begin{cases} \displaystyle\sum_{q\le \ln_4 n,\ \Delta_q(n)=1} \frac{1}{q} & \text{if } n > e^e, \\ 0 & \text{otherwise.} \end{cases}$$

Note that every term in the sum for $\widetilde{f}(n)$ is also in the sum for $f(n)$, so $f(n) \ge \widetilde{f}(n)$. For the new function we have

THEOREM 3.3. *For the positive constant c_2 in Theorem* 3.1,

$$\sum_{n\le x} \widetilde{f}(n) \ge \frac{1}{2} c_2 x \sum_{q\le \ln_4 x} \frac{1}{q^{2-\{\ln_3 x/\ln q\}}}$$

for all large x.

P r o o f. The proof follows almost immediately from Theorem 3.1. One just needs to check that the contribution from pairs n, q with $\ln_4 n < q \le \ln_4 x$ is negligible.

Let us turn to the proof of Theorem 3.1. We need some preparations.

LEMMA 3.4. *We have*

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1}} 1 = \sum_{k\ge 1} \sum_{\substack{p\le x^{1/2}\\ q^k \| p-1}} \sum_{\substack{m\le x/p\\ (m, P_{q^k}(x/p))=1}} 1 + O\Big(\frac{x\ln q}{q}\Big),$$

where the error is non-negative.

P r o o f. We will work out the above formula for $\sum_{n\le x,\ \Delta_q(n)=1} 1$ in such a way that detailed explanations for the formulae marked by numbers to their left are supplied in the subsequent discussion:

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1}} 1 = \sum_{k\ge 1} \sum_{\substack{n\le x\\ q^k \| \lambda(n)\\ \Delta_q(n)=1}} 1$$

$$= \sum_{k\ge 1} \sum_{\substack{p\le x\\ q^k \| p-1}} \sum_{r\ge 1} \sum_{\substack{m\le x/p^r\\ (m, P_{q^k}(x/p^r))=1}} 1 + O\Big(\frac{x}{q^2}\Big) \tag{3}$$

$$= \sum_{k\ge 1} \sum_{\substack{p\le x\\ q^k \| p-1}} \sum_{\substack{m\le x/p\\ (m, P_{q^k}(x/p))=1}} 1 + O\Big(\frac{x}{q^2}\Big) \tag{4}$$

$$= \sum_{k\ge 1} \sum_{\substack{p\le x^{1/2}\\ q^k \| p-1}} \sum_{\substack{m\le x/p\\ (m, P_{q^k}(x/p))=1}} 1 + O\Big(\frac{x\ln q}{q}\Big). \tag{5}$$

To see why the equality of (3) is true, let us first notice that $q^k \parallel \lambda(n)$ implies either $q^{k+1} \mid n$ or $q^k \parallel p-1$ for some prime $p \mid n$. Since the 4-fold sum in (3) counts the exact number of the positive integers $n \le x$ with $\Delta_q(n) = 1$ and a prime factor p for which $q^k \parallel p-1$ where $q^k \parallel \lambda(n)$, the difference between this 4-fold sum in (3) and the sum just before it is bounded by

$$\sum_{\substack{n\le x\\ q^2\mid n}} 1 \le \frac{x}{q^2}.$$

The difference between the 4-fold sum in (3) and the sum in (4) is

$$\le \sum_{k\ge1} \sum_{\substack{p\le x\\ q^k\parallel p-1}} \sum_{r\ge2} \frac{x}{p^r} \ll x \sum_{k\ge1} \sum_{\substack{p\le x\\ q^k\parallel p-1}} \frac{1}{p^2} \ll x \sum_{k\ge1} \frac{1}{q^{2k}} = \frac{x}{q^2-1}.$$

By Lemma 2.3, the difference between the sum in (4) and the sum in (5) is bounded by

$$\sum_{k\ge1} \sum_{\substack{x^{1/2}<p\le x\\ q^k\parallel p-1}} \frac{x}{p} = x \sum_{k\ge1} \left(\frac{\ln_2 x - \ln_2 x^{1/2}}{q^k} + O\left(\frac{k\ln q}{q^k}\right)\right) \ll \frac{x\ln q}{q}.$$

It is easy to notice that the errors in (3), (4) and (5) are all non-negative. Therefore we proved Lemma 3.4.

For the sake of convenience we introduce a few notations of sieve methods. Let $\mathcal{A}$ be the set of positive integers up to x. Let $\mathcal{P}_{q^k}$ be the set of primes congruent to 1 modulo q^k. Let

$$P_{q^k}(z) := \prod_{\substack{p\le z\\ p\equiv 1 \bmod q^k}} p \quad \text{for any } z \le x.$$

Then

$$S(\mathcal{A}, \mathcal{P}_{q^k}, y) := \sum_{\substack{n\in\mathcal{A}\\ (n,\, P_{q^k}(y))=1}} 1 \quad \text{and} \quad W(z) := \prod_{\substack{p\le z\\ p\in\mathcal{P}_{q^k}}} \left(1-\frac{1}{p}\right).$$

Lemma 3.5. (i) *With the notations introduced above we have*

$$S(\mathcal{A}, \mathcal{P}_{q^k}, x) \ll xW(x)$$

uniformly for all q, k and x.

(ii) *Let $z = \exp(\ln x/\ln_2 x)$. As $x \to \infty$ we have*

$$S(\mathcal{A}, \mathcal{P}_{q^k}, z) = x\left(1 + O\left(\frac{1}{\ln x}\right)\right)W(z)$$

uniformly for all q and k.

P r o o f. See Theorems 2.2 and 7.2 in [5].

LEMMA 3.6. *There are absolute positive constants c and x_0 so that*

$$S(\mathcal{A}, \mathcal{P}_{q^k}, x) \ge cx,$$

provided that $q^k \ge \ln_2 x$ and $x \ge x_0$.

P r o o f. Let $z = \exp(\ln x/\ln_2 x)$. Write

$$S(\mathcal{A}, \mathcal{P}_{q^k}, x) = S(\mathcal{A}, \mathcal{P}_{q^k}, z) - E$$

where E is the number of positive integers $n \le x$ such that $\gcd(n, P_{q^k}(z)) = 1$ and $\gcd(n, P_{q^k}(x)/P_{q^k}(z)) > 1$. Then by the condition $q^k \ge \ln_2 x$ and Lemma 2.3,

$$\begin{aligned} E &\le \sum_{\substack{z<p\le x\\ p\in\mathcal{P}_{q^k}}} \sum_{\substack{n\le x\\ p\mid n}} 1 \le \sum_{\substack{z<p\le x\\ p\in\mathcal{P}_{q^k}}} \frac{x}{p} \\ &= x\Big(\frac{\ln_2 x - \ln_2 z}{q^k(1-1/q)} + O\Big(\frac{k\ln q}{q^k}\Big)\Big) = O\Big(\frac{x\ln_3 x}{\ln_2 x}\Big), \end{aligned}$$

where the implied constant is independent of q and k. By Lemma 3.5(ii) we have

$$S(\mathcal{A}, \mathcal{P}_{q^k}, z) = xW(z)(1+o(1))$$

uniformly as $x \to \infty$, where

$$\begin{aligned} W(z) &= \prod_{\substack{p\le z\\ p\in\mathcal{P}_{q^k}}} \Big(1-\frac{1}{p}\Big) = \exp\Big(-\sum_{\substack{p\le z\\ p\in\mathcal{P}_{q^k}}} \frac{1}{p} + O\Big(\frac{1}{q^k}\Big)\Big) \\ &= \exp\Big(-\frac{\ln_2 z}{q^k(1-q^{-1})} + O\Big(\frac{k\ln q}{q^k}\Big)\Big), \end{aligned}$$

by Lemma 2.3 again. But $\ln_2 z = \ln_2 x - \ln_3 x$ and $q^k \ge \ln_2 x$. Thus $W(z) \ge c'$ for some constant $c' > 0$ independent of q and k. Therefore, if x is sufficiently large,

$$S(\mathcal{A}, \mathcal{P}_{q^k}, x) \ge c'x\,(1+o(1)) - o(x) \ge cx$$

for some constant c with $0 < c < c'$. This ends the proof.

Proof of Theorem 3.1. First we will show that

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1}} 1 \le \frac{c_1 x}{q^{\|\ln_3 x/\ln q\|}}$$

for some positive constant c_1. First we claim that

$$(6)\qquad \sum_{\substack{n\le x\\ \Delta_q(n)=1}} 1 \ll x\sum_{k\ge 1}\frac{\ln_2 x}{q^k}\exp\Big(-\frac{\ln_2 x}{q^{k-1}(q-1)}\Big)+O\Big(\frac{x\ln q}{q}\Big).$$

To see this let us look at the innermost sum in Lemma 3.4. Noting that $p\le x^{1/2}$, by Lemma 3.5 and Lemma 2.3, we have

$$\begin{aligned}\sum_{\substack{m\le x/p\\ (m,P_{q^k}(x/p))=1}} 1 &\ll \frac{x}{p}\prod_{\substack{\text{prime } l\le x/p\\ l\equiv 1 \bmod q^k}}\Big(1-\frac{1}{l}\Big)=\frac{x}{p}\exp\Big(-\sum_{\substack{l\le x/p\\ l\equiv 1 \bmod q^k}}\frac{1}{l}+O\Big(\frac{1}{q^{2k}}\Big)\Big)\\ &=\frac{x}{p}\exp\Big(-\frac{\ln_2(x/p)}{q^{k-1}(q-1)}+O\Big(\frac{k\ln q}{q^k}\Big)\Big)\ll\frac{x}{p}\exp\Big(-\frac{\ln_2 x}{q^{k-1}(q-1)}\Big),\end{aligned}$$

uniformly in q and p. Now put this result into the sum in Lemma 3.4 and use Lemma 2.3 again to find that

$$\begin{aligned}\sum_{\substack{n\le x\\ \Delta_q(n)=1}} 1 &\ll x\sum_{k\ge 1}\Big(\frac{\ln_2 x^{1/2}}{q^k}+O\Big(\frac{k\ln q}{q^k}\Big)\Big)\exp\Big(-\frac{\ln_2 x}{q^{k-1}(q-1)}\Big)+O\Big(\frac{x\ln q}{q}\Big).\end{aligned}$$

Since $\sum_{k\ge 1} k\ln q/q^k = O(\ln q/q)$, we have (6).

Next we divide the sum on the right side of (6) into two sums, according to whether $q^k > \ln_2 x$ or $q^k \le \ln_2 x$. Let M be the minimal integer so that $q^M > \ln_2 x$. Then $M = \ln_3 x/\ln q - \{\ln_3 x/\ln q\} + 1$, which equals $[\ln_3 x/\ln q]+1$. Then

$$\begin{aligned}\sum_{q^k>\ln_2 x}\frac{\ln_2 x}{q^k}\exp\Big(-\frac{\ln_2 x}{q^k(1-1/q)}\Big) &\le \sum_{q^k>\ln_2 x}\frac{\ln_2 x}{q^k}=\frac{\ln_2 x}{q^M(1-1/q)}\le\frac{2\ln_2 x}{q^M}=\frac{2}{q^{1-\{\ln_3 x/\ln q\}}}.\end{aligned}$$

Let L be the maximal integer so that $q^L\le \ln_2 x$. Then

$$L=\Big[\frac{\ln_3 x}{\ln q}\Big]=\frac{\ln_3 x}{\ln q}-\Big\{\frac{\ln_3 x}{\ln q}\Big\}.$$

Noting that $k\ge 1$, we have

$$\sum_{q^k\le \ln_2 x}\frac{\ln_2 x}{q^k}\exp\Big(-\frac{\ln_2 x}{q^k(1-1/q)}\Big)\le \sum_{q^k\le \ln_2 x}\frac{\dfrac{\ln_2 x}{q^k}}{\dfrac{1}{2!}\Big(\dfrac{\ln_2 x}{q^k(1-1/q)}\Big)^2}$$
$$=\frac{2(q-1)}{q\ln_2 x}(q^L-1)<\frac{2q^L}{\ln_2 x}$$
$$=\frac{2}{q^{\{\ln_3 x/\ln q\}}}.$$

Put these results back in (6). We have

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1}}1\ll\frac{x}{q^{\|\ln_3 x/\ln q\|}}+O\Big(\frac{x\ln q}{q}\Big)\ll\frac{x}{q^{\|\ln_3 x/\ln q\|}},$$

since $\|\ln_3 x/\ln q\|\le 1/2$. Thus we proved one half of Theorem 3.1.

Secondly we will prove that

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1}}1\ge\frac{c_2 x}{q^{1-\{\ln_3 x/\ln q\}}}$$

for some positive constant c_2 independent of q. By Lemma 3.4, noting that the error in it is non-negative, we have

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1}}1\ge\sum_{\substack{k\text{ with}\\ q^k>\ln_2 x}}\sum_{\substack{p\le x^{1/2}\\ q^k\|p-1}}\sum_{\substack{m\le x/p\\ (m,P_{q^k}(x/p))=1}}1$$

$$\ge\sum_{\substack{k\text{ with}\\ q^k>\ln_2 x}}\sum_{\substack{p\le x^{1/2}\\ q^k\|p-1}}c\frac{x}{p}\tag{7}$$

$$=cx\sum_{\substack{k\text{ with}\\ q^k>\ln_2 x}}\Big(\frac{\ln_2 x^{1/2}}{q^k}+O\Big(\frac{k\ln q}{q^k}\Big)\Big)\tag{8}$$

$$\ge cx\Big(\frac{\ln_2 x}{q^M}+O\Big(\frac{M\ln q}{q^M}\Big)\Big),\tag{9}$$

where M is the smallest integer so that $q^M>\ln_2 x$. We obtain (7) and (8) by using Lemmas 3.6 and 2.3, respectively. So what is left to be explained is (9). Notice that the sum in (8) can be written as

$$\sum_{k\ge M}\frac{\ln_2 x}{q^k}+O\Big(\sum_{k\ge M}\frac{k\ln q}{q^k}\Big),$$

so that (9) follows.

Since $M = \ln_3 x/\ln q - \{\ln_3 x/\ln q\} + 1$ and $q \le \ln_2 x$, we have

$$\frac{M \ln q}{q^M} \ll \frac{(\ln_3 x/\ln q)\ln q}{q^{\ln_3 x/\ln q - \{\ln_3 x/\ln q\}+1}} = \frac{\ln_3 x}{\ln_2 x} q^{\{\ln_3 x/\ln q\}-1}.$$

Therefore, choosing $c_2 = c/2$, we have

$$\begin{aligned}\sum_{\substack{n\le x\\ \Delta_q(n)=1}} 1 &\ge cx\left(\frac{\ln_2 x}{q^M} + O\left(\frac{M\ln q}{q^M}\right)\right)\\ &= cxq^{\{\ln_3 x/\ln q\}-1}(1+o(1)) \ge c_2 x q^{\{\ln_3 x/\ln q\}-1},\end{aligned}$$

for x sufficiently large. We have proved Theorem 3.1.

4. Two crucial series. As one can see from Theorems 3.2 and 3.3, it is necessary to understand the series

$$\sum_{q\le T} \frac{1}{q^{1+\|\ln T/\ln q\|}} \quad \text{and} \quad \sum_{q\le \ln_2 T} \frac{1}{q^{2-\{\ln T/\ln q\}}}$$

in the course of estimating the first moments of $f(n)$ and $\widetilde{f}(n)$. This section is dedicated to the study of some features of the two series.

THEOREM 4.1. *For $T \ge e^{e^e}$, we have*

$$\sum_{q\le T} \frac{1}{q^{1+\|\ln T/\ln q\|}} \ll \ln_3 T.$$

First let us mention the following well known fact in prime number theory.

LEMMA 4.2. *For all $x \ge 2$,*

$$\sum_{p\le x} \frac{1}{p} = \ln\ln x + c_3 + O(\exp(-c_4 (\ln x)^{1/2})),$$

where c_3 and $c_4 > 0$ are constants.

Proof (of Theorem 4.1). Write the series as the sum of s_1 and s_2 as follows:

$$\sum_{q\le Q} \frac{1}{q^{1+\|\ln T/\ln q\|}} + \sum_{Q<q\le T} \frac{1}{q^{1+\|\ln T/\ln q\|}} = s_1 + s_2,$$

where Q will be determined later. We will use the trivial estimate $s_1 = O(\ln_2 Q)$. For s_2, let us write

$$\begin{aligned}s_2 &= \sum_{\substack{Q<q\le T\\ \|\ln T/\ln q\| > \ln A/\ln q}} \frac{1}{q^{1+\|\ln T/\ln q\|}} + \sum_{\substack{Q<q\le T\\ \|\ln T/\ln q\| \le \ln A/\ln q}} \frac{1}{q^{1+\|\ln T/\ln q\|}}\\ &= s_2^{(1)} + s_2^{(2)},\end{aligned}$$

where $A > e$ will be chosen so that $\ln A/\ln Q < 1/2$ but the exact value for A will be determined later. Clearly $s_2^{(1)} = O\big(\frac{\ln_2 T}{A}\big)$ while

$$s_2^{(2)} = \sum_{\substack{Q<q\le T\\ |k-\ln T/\ln q|\le \ln A/\ln q}} \frac{1}{q^{1+|k-\ln T/\ln q|}},$$

where there is at most one integer $k \ge 1$ for each prime q as our A satisfies $\ln A/\ln Q < 1/2$. Thus, by Lemma 4.2,

$$s_2^{(2)} \le \sum_{k\le(\ln T+\ln A)/\ln Q} \sum_{(T/A)^{1/k}\le q\le (AT)^{1/k}} \frac{1}{q}$$

$$= \sum_{k\le(\ln T+\ln A)/\ln Q} \bigg(\ln_2(AT)^{1/k} - \ln_2\left(\frac{T}{A}\right)^{1/k} + O\bigg(\exp\bigg(-c_4\sqrt{\frac{1}{k}\ln\frac{T}{A}}\bigg)\bigg)\bigg).$$

Since $\ln A/\ln Q < 1/2$ and $Q \le T$, we have $\ln A/\ln T < 1/2$, so that

$$\begin{aligned}\ln_2(AT)^{1/k} - \ln_2\left(\frac{T}{A}\right)^{1/k} &= \ln\frac{\ln T+\ln A}{\ln T-\ln A} = \ln\left(1+\frac{2\ln A}{\ln T-\ln A}\right)\\ &< \ln\left(1+\frac{4\ln A}{\ln T}\right) < \frac{4\ln A}{\ln T}\end{aligned}$$

and

$$\frac{1}{k}\ln\frac{T}{A} \ge \frac{\ln T-\ln A}{\ln T+\ln A}\ln Q \ge \frac{\ln Q}{3}.$$

Hence

$$s_2^{(2)} \le \frac{\ln T+\ln A}{\ln Q}\bigg(\frac{4\ln A}{\ln T} + O\bigg(\exp\bigg(-\frac{c_4}{\sqrt{3}}(\ln Q)^{1/2}\bigg)\bigg)\bigg).$$

Now choose Q so that $\ln Q = (3/c_4^2)(\ln_2 T)^2$, and choose $A = (\ln_2 T)^2$. If T is sufficiently large, we have $Q \le T$ and $\ln A/\ln Q < 1/2$. We thus have

$$s_2^{(2)} \le \frac{\ln T+\ln A}{\ln Q}\bigg(\frac{4\ln A}{\ln T} + O\bigg(\frac{1}{\ln T}\bigg)\bigg) = O\bigg(\frac{\ln_3 T}{(\ln_2 T)^2}\bigg).$$

At the same time,

$$s_2^{(1)} = O\bigg(\frac{\ln_2 T}{A}\bigg) = O\bigg(\frac{1}{\ln_2 T}\bigg) \quad \text{and so} \quad s_2 = O\bigg(\frac{1}{\ln_2 T}\bigg).$$

Then

$$\sum_{q\le T}\frac{1}{q^{1+\|\ln T/\ln q\|}} = s_1 + s_2 = O(\ln_2 Q) = O(\ln_3 T).$$

This concludes the proof of Theorem 4.1.

Theorem 4.1 gives us an upper bound for the series mentioned at the beginning of the section. Next let us investigate the normal value of the series.

We note that

$$\begin{aligned}\sum_{q\le T}\frac{1}{q^{1+\|\ln T/\ln q\|}} &= \sum_{\substack{q\le T\\ \|\ln T/\ln q\|<\ln_2 q/\ln q}}\frac{1}{q^{1+\|\ln T/\ln q\|}}\\ &\quad+\sum_{\substack{q\le T\\ \|\ln T/\ln q\|\ge\ln_2 q/\ln q}}\frac{1}{q^{1+\|\ln T/\ln q\|}}\\ &\le \sum_{\substack{q\le T\\ \|\ln T/\ln q\|<\ln_2 q/\ln q}}\frac{1}{q}+\sum_{\substack{q\le T\\ \|\ln T/\ln q\|\ge\ln_2 q/\ln q}}\frac{1}{q\ln q}\\ &= \sum_{\substack{q\le T\\ \|\ln T/\ln q\|<\ln_2 q/\ln q}}\frac{1}{q}+O(1)\end{aligned}$$

where the last equality follows from the fact that $\sum 1/(p\ln p)<\infty$, p running over primes. This suggests considering the average value of the following function. Let us define

$$g(t):=\sum_{\substack{q\le e^t\\ \|t/\ln q\|<\ln_2 q/\ln q}}\frac{1}{q}$$

for all $t\ge 0$. Thus the above argument shows that

$$\sum_{q\le e^t}\frac{1}{q^{1+\|t/\ln q\|}}\le g(t)+O(1). \tag{10}$$

Lemma 4.3. *There is a positive constant c_5 so that, for all $y>0$,*

$$\sum_{k=1}^{[y]} g(k)\le c_5 y.$$

Proof. By definition,

$$\sum_{k=1}^{[y]} g(k)=\sum_{1\le k\le y}\ \sum_{\substack{q\le e^k\\ \|k/\ln q\|<\ln_2 q/\ln q}}\frac{1}{q}\le\sum_{q\le e^y}\frac{1}{q}\ \sum_{\substack{1\le k\le y\\ \|k/\ln q\|<\ln_2 q/\ln q}}1.$$

Since $\ln_2 q/\ln q<1/2$, if $\|k/\ln q\|<\ln_2 q/\ln q$ then there is a unique integer l with

$$l-\frac{\ln_2 q}{\ln q}<\frac{k}{\ln q}<l+\frac{\ln_2 q}{\ln q}.$$

For the fixed integer l there are at most $2\ln_2 q + 1$ integers k so that

$$\left\| \frac{k}{\ln q} \right\| = \left| \frac{k}{\ln q} - l \right| < \frac{\ln_2 q}{\ln q}.$$

Therefore

$$\sum_{k=1}^{[y]} g(k) \le \sum_{q \le e^y} \frac{1}{q} \sum_{l=0}^{\lceil y/\ln q \rceil} (2\ln_2 q + 1) \ll y \sum_{q \le e^y} \frac{\ln_2 q}{q \ln q} \ll y,$$

as the last sum is bounded. Thus we proved Lemma 4.3.

THEOREM 4.4. *There is a positive number c_6 so that, for all $y \ge 1$,*

$$\sum_{k=1}^{[y]} \sum_{q \le e^k} \frac{1}{q^{1+\|k/\ln q\|}} \le c_6 y.$$

P r o o f. This follows from (10) and Lemma 4.3.

DEFINITION. Let S be a set of natural numbers. If $\lim_{x\to\infty} 1/x\#\{n \le x : n \in S\}$ exists, we call it the *density* of S. Otherwise we call the corresponding upper or lower limit the *upper* or *lower* density of S, respectively.

As a corollary of Theorem 4.4 we have

THEOREM 4.5. *Let c_6 be the same constant as in Theorem* 4.4, *and let $b > c_6$ be any number. Then the set S_b of $k \in \mathbb{N}$ with*

$$\sum_{q \le e^k} \frac{1}{q^{1+\|k/\ln q\|}} \le b$$

has positive upper density.

P r o o f. Suppose that S_b has density zero. Then $\mathbb{N} \setminus S_b$ has density one. On the other hand, for any $y > 0$,

$$\sum_{k \le y} \sum_{q \le e^k} \frac{1}{q^{1+\|k/\ln q\|}} \ge b \sum_{\substack{k \le y \\ k \notin S_b}} 1.$$

Then by Theorem 4.4, we have

$$c_6 \ge b \frac{1}{y} \sum_{\substack{k \le y \\ k \notin S_b}} 1.$$

Send y to infinity, we have $c_6 \ge b$, a contradiction. Therefore S_b has positive upper density. We are done.

In the rest of the section we will consider the second series mentioned at the beginning of this section. Our attention will focus on how big the series could be when T is large. The next lemma is an elementary result.

LEMMA 4.6. *If $q \geq 5$ and $\|\ln 2T/\ln q\| < \ln 2/\ln q$, then*

$$\frac{1}{q^{2-\{\ln T/\ln q\}}} \geq \frac{1}{4q}.$$

For a given T let Q be the largest prime less than or equal to $\ln_2 T$ such that all primes q with $5 \leq q \leq Q$ satisfy

$$\left\| \frac{\ln 2T}{\ln q} \right\| \leq \frac{\ln 2}{\ln q}.$$

We want to know how large Q can be as a function of T. By the following result on simultaneous Diophantine approximation we can say something about this question.

LEMMA 4.7. *Let $\xi_1, \ldots, \xi_l$ be any l real numbers. Then, for any integer $N > 1$, there exists a positive integer $m \leq N^l$ such that $\|m\xi_i\| < 1/N$ for all $i = 1, \ldots, l$.*

P r o o f. See the proof of Theorem 200 in [6].

Let Q be any large prime. Consider the l irrationals $1/\ln 5$, $1/\ln 7$, ..., $1/\ln Q$ where $l = \pi(Q) - 2$. Let $N = \lceil \ln Q/\ln 2 \rceil$. Then by Lemma 4.7 there exists an integer m, $1 \leq m \leq N^l$, so that every prime q with $5 \leq q \leq Q$ satisfies $\|m/\ln q\| < 1/N \leq \ln 2/\ln Q$. Since $\|m/\ln 5\| < \ln 2/\ln Q$ and $\ln 5$ is irrational it follows that $m = m(Q) \to \infty$ as $Q \to \infty$.

By the definition of N we have $N = (\ln Q/\ln 2) + O(1)$ when Q is sufficiently large. But $l = \pi(Q) - 2$, so for Q sufficiently large by the prime number theorem we have $Q > l \ln N \geq \ln m$. Now choose T such that $2T = e^m$. Clearly $m > \ln T$. Thus $Q > \ln_2 T$. With this choice of T all primes q with $5 \leq q \leq \ln_2 T$ satisfy $\|\ln 2T/\ln q\| < \ln 2/\ln q$. Thus by Lemma 4.6,

$$\begin{aligned}\sum_{q \leq \ln_2 T} \frac{1}{q^{2-\{\ln T/\ln q\}}} &\geq \sum_{\substack{5 \leq q \leq \ln_2 T \\ \|\ln 2T/\ln q\| < \ln 2/\ln q}} \frac{1}{q^{2-\{\ln T/\ln q\}}} \\ &\geq \sum_{5 \leq q \leq \ln_2 T} \frac{1}{4q} = \frac{\ln_4 T}{4} + O(1).\end{aligned}$$

Therefore we have proved the following theorem.

THEOREM 4.8. *There is an unbounded set of numbers T for which*

$$\sum_{q \leq \ln_2 T} \frac{1}{q^{2-\{\ln T/\ln q\}}} \geq \frac{1}{5} \ln_4 T.$$

5. Proofs of Theorems 2 and 3. Recall the function $r(n) = R(n)/\phi(n)$ from Section 2. We can restate Theorem 2 in the following form. These estimates are also bounds for the first moment of $f(n)$ and that of $\widetilde{f}(n)$.

THEOREM 5.1. *There exist positive constants c_7, c_8 and c_9 so that*

(i) *for all sufficiently large x we have*

$$\sum_{n\le x} |\ln r(n)| \le c_7 x \ln_5 x;$$

(ii) *there is an unbounded set of numbers x for which*

$$\sum_{n\le x} |\ln r(n)| \ge c_8 x \ln_6 x;$$

(iii) *there is an unbounded set of numbers x for which*

$$\sum_{n\le x} |\ln r(n)| \le c_9 x.$$

P r o o f. By the definition of $f(n)$ in Section 2 we have $f(n) = -\ln r(n) + O(1)$, where the $O(1)$ is non-positive. Since $0 < r(n) \le 1$, we have

$$\sum_{n\le x} f(n) = \sum_{n\le x} |\ln r(n)| + O(x),$$

where the $O(x)$ is non-positive. Noticing that $f(n) \ge \widetilde{f}(n)$, we can get Theorem 5.1 by applying Theorems 3.2, 3.3, 4.1, 4.5 and 4.8.

COROLLARY 5.2. *Let c_9 be the constant in Theorem* 5.1. *There is an unbounded set of numbers x such that*

$$D(x,u) \le c_9/|\ln u|$$

for all u with $0 < u < 1$.

P r o o f. By definition,

$$\sum_{\substack{n\le x\\ r(n)\le u}} |\ln r(n)| \ge x|\ln u| D(x,u)$$

for all u with $0 < u < 1$. Thus, the corollary follows from Theorem 5.1(iii).

COROLLARY 5.3. *There is a positive constant c_{10} with the property that, for any positive constant $b < c_{10}$, the set $\mathcal{S}_b = \{n \in \mathbb{N} : r(n) \le (\ln_5 n)^{-b}\}$ has positive upper density.*

Consider the set

$$\mathcal{S}'_b = \{n \in \mathbb{N} : \widetilde{f}(n) \ge b \ln_6 n\}.$$

Since $\ln r(n) = -f(n) + O(1)$ with the $O(1)$ being non-positive and $f(n) \le \widetilde{f}(n)$, we have $\mathcal{S}'_b \subseteq \mathcal{S}_b$. It is enough to show that there is a constant c_{10} such that, for all b with $0 < b < c_{10}$, $\mathcal{S}'_b$ has a positive upper density.

From the definition of $\widetilde{f}(n)$ we trivially have $\widetilde{f}(n) \le 2\ln_6 n$ when n is large enough. Thus we have, for x sufficiently large,

$$\sum_{n\le x}\widetilde{f}(n) = \sum_{\substack{n\le x\\ n\notin\mathcal{S}'_b}}\widetilde{f}(n) + \sum_{\substack{n\le x\\ n\in\mathcal{S}'_b}}\widetilde{f}(n) \le \sum_{\substack{n\le x\\ n\notin\mathcal{S}'_b}} b\ln_6 x + 2\ln_6 x \sum_{\substack{n\le x\\ n\in\mathcal{S}'_b}} 1.$$

On the other hand, we choose $c_{10} = c_2/10$ where c_2 is the constant in Theorem 3.1. Then, by Theorems 3.3 and 4.8, there exists an unbounded set of real numbers x for which

$$\sum_{n\le x}\widetilde{f}(n) \ge c_{10}x\ln_6 x.$$

Let b be any constant with $0 < b < c_{10}$. Then combining the above we have for such numbers x,

$$c_{10} \le \frac{b}{x}\sum_{\substack{n\le x\\ n\notin\mathcal{S}'_b}} 1 + \frac{2}{x}\sum_{\substack{n\le x\\ n\in\mathcal{S}'_b}} 1 \le b + \frac{2}{x}\sum_{\substack{n\le x\\ n\in\mathcal{S}'_b}} 1.$$

Thus the upper density of $\mathcal{S}'_b$ is at least $(c_{10}-b)/2$, which is positive. Thus we have proved the corollary.

COROLLARY 5.4. *There exist positive constants δ, b and an unbounded set of numbers x with $D(x, (\ln_5 x)^{-b}) \ge \delta$.*

Proof. This follows immediately from the definition of $D(x,u)$ and from Corollary 5.3.

THEOREM 5.5. *There exists a positive number u_0 such that, for each $u \in (0, u_0)$, the function $D(x,u)$ does not have a limit as $x \to \infty$. Thus the function $r(n)$ does not have distribution function.*

Proof. This is a corollary of Corollaries 5.2 and 5.4.

Note that Theorem 3 in the introduction follows from Corollaries 5.2, 5.4 and Theorem 5.5.

References

[1] R. D. Carmichael, *The Theory of Numbers*, Wiley, New York, 1914.

[2] P. D. T. A. Elliott, *On the limiting distribution of $f(p+1)$ for non-negative additive functions*, Acta Math. 132 (1974), 53–75.

[3] P. Erdős, C. Pomerance and E. Schmutz, *Carmichael's lambda function*, Acta Arith. 58 (1991), 363–385.

[4] J. Galambos, *Advanced Probability Theory*, 2nd ed., Dekker, New York, 1995.

[5] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, New York, 1974.

[6] G. H. Hardy and B. M. Wright, *An Introduction to the Theory of Numbers*, Oxford Univ. Press, London, 1960.
[7] W. J. LeVeque, *Topics in Number Theory*, Vol. I, Addison-Wesley, Reading, Mass., 1956.
[8] S. Li, *Artin's conjecture on average for composite moduli*, preprint.
[9] G. Martin, *The least prime primitive root and the shifted sieve*, Acta Arith. 80 (1997), 277–288.
[10] K. K. Norton, *On the number of restricted prime factors of an integer I*, Illinois J. Math. 20 (1976), 681–705.
[11] C. Pomerance, *On the distribution of amicable numbers*, J. Reine Angew. Math. 293/294 (1977), 217–222.
[12] I. J. Schoenberg, *On asymptotic distributions of arithmetical functions*, Trans. Amer. Math. Soc. 39 (1936), 315–330.

Department of Mathematics
University of Georgia
Athens, Georgia 30602
U.S.A.
E-mail: sli@math.uga.edu

Received on 31.1.1997
and in revised form on 15.4.1998 (3123)