ACTA ARITHMETICA
XLIX (1988)

On the number of distinct values of Euler's φ-function

by

HELMUT MAIER* and CARL POMERANCE* (Athens, Ga.)

Dedicated to Paul Erdős on his 75th birthday

1. Introduction. Let $V(x)$ denote the number of distinct values of Euler's φ-function not exceeding x. The subject of estimating $V(x)$ has a long history dating back to 1929 when S. S. Pillai [5] showed that

$$V(x) \ll \frac{x}{(\log x)^{(\log 2)/e}}.$$

In 1935, Erdős [1] improved this result, getting $V(x) = x/(\log x)^{1+o(1)}$. Subsequent papers have dealt with the nature of the factor $(\log x)^{o(1)}$ in this formula. In Erdős and Hall [2], [3] it was shown that

$$(1.1)\qquad \frac{x}{\log x}\exp\{c_1(\log\log\log x)^2\} \ll V(x) \ll \frac{x}{\log x}\exp\{c_2(\log\log x)^{1/2}\}$$

for certain positive constants c_1, c_2. Recently, in [6], the second author showed that the lower bound in (1.1) is close to the truth about $V(x)$. Namely, there is a positive constant c_3 such that

$$(1.2)\qquad V(x) \ll \frac{x}{\log x}\exp\{c_3(\log\log\log x)^2\} \quad \text{for all large } x.$$

In fact, it was shown in [6] that (1.2) holds for all

$$c_3 > (2-2\log(e-1))^{-1} = 1.090096128\ldots$$

and the first inequality in (1.1) holds for all

$$c_1 < (2\log 2 - 2\log(2+\log 2-\sqrt{4\log 2+\log^2 2}))^{-1} = 0.617122930\ldots$$

---

* Research supported in part by NSF grants.

In this paper we eliminate the gap between c_1 and c_3, showing that there is a constant C such that the first inequality in (1.1) is true for all $c_1 < C$ and (1.2) is true for all $c_3 > C$. To properly state this result, we must define the number C.

Let

$$(1.3)\qquad F(x) = \sum_{n=1}^{\infty} a_n x^n, \qquad a_n = (n+1)\log(n+1) - n\log n - 1.$$

Since $a_n \sim \log n$, $a_n > 0$, we have that $F(x)$ is defined and strictly increasing on $[0, 1)$, $F(0) = 0$ and $F(x) \to \infty$ as $x \to 1^-$. Thus there is a unique number $c_0 \in (0, 1)$ defined by the equation

$$(1.4)\qquad F(c_0) = 1.$$

Thus $c_0 = 0.54259859\ldots$ (Thanks are due to J. P. Massias for this calculation.) Let

$$(1.5)\qquad C = \frac{1}{2|\log c_0|} = 0.81781465\ldots$$

In this paper we shall prove the following result.

THEOREM. *With C given by (1.5) we have*

$$V(x) = \frac{x}{\log x} \exp\{(C + o(1))(\log\log\log x)^2\}.$$

The proof is divided into two parts, one for the lower bound and one for the upper bound. As with [2] and [6], the key tool employed is the result from [1] that the shifted prime $p-1$ normally has $\log\log p$ prime factors. However, here we use the somewhat finer result that the normal number of primes dividing $p-1$ that lie between $\exp\{(\log x)^a\}$ and $\exp\{(\log x)^b\}$ is $(b-a)\log\log x$. A more precise description of the result we use on shifted primes is given in the next section.

With this finer tool on the normal number of prime factors of a shifted prime at out disposal, the upper bound in the theorem follows by the same argument as in [6]. The lower bound argument is basically new, but has some of the same flavor as the argument in [3] and its improvement in [6].

We shall use the notation $P(n)$ to denote the largest prime factor of n for $n > 1$ and we shall let $P(1) = 1$. Throughout, the letters p, q will always denote primes. The constants $c_1, c_2, \ldots$ will always be absolute, that is, they will not depend on any parameters.

2. Normal primes. If $T_1 < T$ let $\Omega(n, T_1, T)$ denote the total number of prime factors of n, counted with multiplicity, that lie in the interval $(T_1, T]$. If $\delta > 0$, $S > 1$, we say an integer n is *δ,S-normal* if $\Omega(n, 1, S) < 2\log\log(10S)$ and for each pair T_1, T with $S \leqslant T_1 < T \leqslant n$ we have

$$(2.1)\qquad |\Omega(n, T_1, T) - (\log\log T - \log\log T_1)| < \delta \log\log T.$$

Then from standard arguments one can get the following result.

PROPOSITION 2.1. *Given $\delta > 0$ arbitrary, there are constants $c = c(\delta) > 0$, $\varepsilon = \varepsilon(\delta) > 0$ such that the number of $n \leqslant x$ which are not δ,S-normal is at most $cx/(\log S)^\varepsilon$ uniformly for all $1 < S < x$.*

We shall say a prime p is δ,S-normal if $p-1$ is δ,S-normal. Then using Proposition 2.1 together with Brun's method one can prove a similar result for primes which are not normal. We suppress the details, but they are similar to the arguments in [1].

PROPOSITION 2.2. *Given $\delta > 0$ arbitrary, there are constants $c = c(\delta) > 0$, $\varepsilon = \varepsilon(\delta) > 0$ (not necessarily the same constants as in Proposition 2.1) such that the number of primes $p \leqslant x$ which are not δ,S-normal is at most*

$$\frac{cx}{(\log S)^\varepsilon \log x}$$

uniformly for all $1 < S < x$.

In the definition of normal prime it should be noted that there is no special reason to use $p-1$. In fact the same result would hold for $p+a$ where a is any fixed non-zero integer. For the Euler φ-function though, the case $a = -1$ is the one of interest.

3. The lower bound. Let α be an arbitrary, but fixed number with $1/2 < \alpha < c_0$ where c_0 is given by (1.4). Let δ be an arbitrary, but fixed number with $0 < \delta < (1-\alpha)/2$. Later, we shall let δ be sufficiently small depending on the choice of α. Let x be large, where "large" may depend on the choice of α, δ. Most of the following functions will depend on x. For $k = 0, 1, 2, \ldots$ let

$$z_k = \exp\{(\log x)^{\alpha^k}\}, \qquad w_k = \exp\{(\log x)^{(1-\delta)\alpha^k}\},$$

and let P_k denote the set of δ, $\log x$ — normal primes in the interval $[w_k, z_k]$.

Let

$$L = \left[\frac{1-\delta}{|\log\alpha|}\log\log\log x\right] + 1$$

and let

$$A = \{n\colon x/2 \leqslant n \leqslant x,\ n = p_0 p_1 \ldots p_L \text{ where each } p_k \in P_k\}.$$

Then $V(x) \geqslant \# B$ where

$$B = \{\varphi(n)\colon n \in A\}.$$

The main outline of the argument is to first estimate $\# A$ and then show $\# B \sim \# A$, that is, φ is usually one-to-one when restricted to A. The latter step will only hold if δ is sufficiently small depending on the choice of α. In our first lemma we obtain the estimate for $\#A$.

LEMMA 3.1. *There is a positive constant* c_4 *such that*

$$\#A \geqslant \frac{x}{\log x}\exp\left\{\frac{1-\delta^2}{2|\log\alpha|}(\log\log\log x)^2 + c_4(\log\delta)\log\log\log x\right\}.$$

Proof. If $m = p_1 \ldots p_L$ where each $p_k \in P_k$, then $m \leqslant z_1^2$, so that the number of $p_0 \in P_0$ with $x/2m \leqslant p_0 \leqslant x/m$ is $\gg x/(m\log x)$, using Proposition 2.2. Thus it remains to estimate $\sum 1/m$. Again using Proposition 2.2, $\sum 1/m$ is at least (for large x)

$$\prod_{k=1}^{L}(\log\log z_k - \log\log w_k - 1)$$

$$= \prod_{k=1}^{L}(\delta\alpha^k\log\log x - 1) > (\log\log x)^L(\delta/2)^L\alpha^{L^2/2}$$

$$> \exp\left\{\frac{1-\delta^2}{2|\log\alpha|}(\log\log\log x)^2 + c_4(\log\delta)\log\log\log x\right\},$$

from which our result follows.

If we now show $\#B \sim \#A$, then the lower bound in the theorem will immediately follow from Lemma 3.1 since we may let δ be arbitrarily close to 0 and α arbitrarily close to c_0. To show $\#B \sim \#A$ we shall use a kind of second-moment argument. Indeed, let

$$B_2 = \{(n_1, n_2) \in A \times A:\ \varphi(n_1) = \varphi(n_2),\ n_1 \neq n_2\}.$$

Then evidently

$$\begin{aligned}(3.1)\qquad \#B &\geqslant \#\{n \in A:\ \text{if } \varphi(n) = \varphi(n') \text{ for some } n' \in A, \text{ then } n = n'\}\\ &= \#A - \#\{n \in A:\ (n, n') \in B_2 \text{ for some } n'\}\\ &\geqslant \#A - \#B_2,\end{aligned}$$

so that it will suffice to show that $\#B_2 = o(\#A)$. To do this we shall first consider the smaller set

$$B_2^* = \{(n_1, n_2) \in B_2:\ \gcd(n_1, n_2) = 1\}$$

and then show how the argument for B_2^* can be extended to all of B_2.

Note that if $(n_1, n_2) \in B_2^*$, then there are primes $p_k, q_k \in P_k$, $p_k \neq q_k$, for $k = 0, \ldots, L$ with

$$n_1 = p_0 \ldots p_L, \qquad n_2 = q_0 \ldots q_L.$$

Thus the number $m = \varphi(n_1) = \varphi(n_2)$ has the two factorizations

$$(p_0 - 1)\ldots(p_L - 1) = (q_0 - 1)\ldots(q_L - 1).$$

Moreover, because of the way n_1, n_2 are formed, m has a predictable (and abnormal) number of primes in the various intervals $(z_k, z_{k-1}]$. Also note that if we discard all the primes in m above z_k, then the resulting divisor of m has two induced factorizations obtained from discarding all primes above z_k in the various $p_i - 1$, $q_i - 1$. We shall estimate B_2^* by counting the number of pairs of factorizations of a truncated m and how this number grows when we move from k to $k-1$.

Before this procedure is made precise, we record a fact about the prime factors of members of B.

LEMMA 3.2. *If* $m \in B$, *then for each* $k = 1, \ldots, L$ *we have*

$$|\Omega(m, z_k, z_{k-1}) - k(\alpha^{k-1} - \alpha^k)\log\log x| < \delta(k+1)\alpha^{k-1}\log\log x.$$

Proof. This follows immediately from the definition of the sets of primes $P_0, \ldots, P_L$.

Note that $\Omega(m, z_k, z_{k-1})$ is about k times the expected amount for a random integer m.

We now begin the description of the details needed for considering dual factorizations of truncated members of B. Let

$$f(m, y_1, y_2) = \prod_{\substack{p^a \| m\\ y_1 < p \leqslant y_2}} p^a$$

and let $f(m, y) = f(m, 1, y)$. For each pair $(n_1, n_2) \in B_2^*$ we consider the $2L+3$-tuple

$$\sigma(n_1, n_2) = (\varphi(n_1);\ p_0 - 1, \ldots, p_L - 1;\ q_0 - 1, \ldots, q_L - 1)$$

where $n_1 = p_0 \ldots p_L$, $n_2 = q_0 \ldots q_L$ and $p_k, q_k \in P_k$ for each k. Let

$$\sigma_k(n_1, n_2) = \big(f(\varphi(n_1), z_k);\ f(p_0 - 1, z_k), \ldots, f(q_L - 1, z_k)\big),$$

$$\tau_k(n_1, n_2) = \big(f(\varphi(n_1), z_k, z_{k-1});\ f(p_0 - 1, z_k, z_{k-1}), \ldots, f(q_L - 1, z_k, z_{k-1})\big).$$

If $\sigma = (b_1, \ldots, b_t)$, $\tau = (d_1, \ldots, d_t)$, let $\sigma\tau = (b_1 d_1, \ldots, b_t d_t)$. Thus for $k = 1, \ldots, L$ we have

$$(3.2)\qquad \sigma_{k-1}(n_1, n_2) = \sigma_k(n_1, n_2)\,\tau_k(n_1, n_2).$$

Let S_k denote the set of $\sigma_k(n_1, n_2)$ for $(n_1, n_2) \in B_2^*$ and let T_k denote the corresponding set of $\tau_k(n_1, n_2)$. Clearly there is a one-to-one correspondence between B_2^* and S_0, so $\#B_2^* = \#S_0$. But then from (3.2),

$$(3.3)\qquad \#B_2^* = \#S_0 = \sum_{\sigma \in S_1}\ \sum_{\substack{\tau \in T_1\\ \sigma\tau \in S_0}} 1.$$

We now estimate the inner sum in (3.3). We shall use the following notation: if σ is a t-tuple, then $\sigma^{(j)}$ is the jth coordinate of σ.

LEMMA 3.3. *For each* $\sigma \in S_1$,

$$\sum_{\substack{\tau \in T_1 \\ \sigma\tau \in S_0}} 1 \leqslant \frac{x}{\sigma^{(1)} (\log x)^{2+\alpha-\delta}}$$

for all large x *depending only on the choice of* α, δ.

Proof. Say $\sigma^{(1)} = m$, $\sigma^{(2)} = u$, $\sigma^{(L+3)} = v$. If $\tau \in T_1$ is such that $\sigma\tau \in S_0$, then $\tau^{(1)} = t$ has the following properties:

(i) $tm \leqslant x$,

(ii) every prime in t is in $(z_1, z_0]$,

(iii) $tu+1$, $tv+1$ are unequal primes.

By Brun's method (see Halberstam–Richert [4]) the number of t satisfying (i), (ii), (iii) is at most

$$\frac{c_5 x/m}{\log z_1 \log^2(x/m)} (\log\log x)^2 \tag{3.4}$$

for some absolute constant c_5. From Lemma 3.2 and a simple calculation, we have

$$m < \exp\{(\log x)^{\alpha+\delta}\},$$

so that $\log(x/m) \sim \log x$. Also note that $\log z_1 = (\log x)^\alpha$. Thus the lemma now follows from (3.4) and the observation that for each t satisfying (i), (ii), (iii), there is at most one $\tau \in T_1$ with $\sigma\tau \in S_0$ and $\tau^{(1)} = t$.

We now see that from (3.3) and Lemma 3.3, an estimate for $\# B_2^*$ may be obtained by estimating $\sum 1/\sigma^{(1)}$ where σ runs over S_1. This is done in L steps where each step is similar to Lemma 3.3. The general step is treated in the following result.

LEMMA 3.4. *If* $2 \leqslant k \leqslant L$ *and* $\sigma \in S_k$, *then*

$$\sum_{\substack{\tau \in T_k \\ \sigma\tau \in S_{k-1}}} \frac{1}{\tau^{(1)}} \leqslant (\log x)^{(\alpha^{k-1}-\alpha^k)(k\log k+k) - 2\alpha^{k-1} + c_6 \delta \alpha^k k \log k}$$

for some absolute constant c_6.

Proof. Note that any $\tau \in T_k$ is of the form

$$\tau = (t; f_0, \ldots, f_{k-1}, 1, \ldots, 1; g_0, \ldots, g_{k-1}, 1, \ldots, 1)$$

where each string of 1's is of length $L+1-k$. Thus $f_{k-1} = \tau^{(k+1)}$, $g_{k-1} = \tau^{(L+k+2)}$. The entries f_{k-1}, g_{k-1} are of particular interest since if $\sigma\tau \in S_{k-1}$, then

$$f_{k-1}\sigma^{(k+1)}+1, \quad g_{k-1}\sigma^{(L+k+2)}+1 \text{ are unequal primes.} \tag{3.5}$$

In addition the entries of τ satisfy

$$t = f_0 \ldots f_{k-1} = g_0 \ldots g_{k-1}. \tag{3.6}$$

We partition T_k into 2 classes $T_{k,1}$, $T_{k,2}$ where in $T_{k,1}$ we have $P(f_{k-1}) \neq P(g_{k-1})$ and in $T_{k,2}$ we have $P(f_{k-1}) = P(g_{k-1})$. In either case, however, since the primes in n_1, n_2 are $\delta, \log x$-normal, it follows that

$$P(f_{k-1}),\ P(g_{k-1}) > \exp\{(\log w_{k-1})^{1-\delta}\} > \exp\{(\log x)^{(1-2\delta)\alpha^{k-1}}\}. \tag{3.7}$$

We consider now the contribution to the sum in the lemma from the members of $T_{k,1}$. If $\tau \in T_{k,1}$ and $\sigma\tau \in S_{k-1}$, write

$$\tau^{(1)} = \tau_0^{(1)} Q_1 Q_2$$

where $Q_1 = P(f_{k-1})$, $Q_2 = P(g_{k-1})$. We now fix all information about τ except the choice of Q_1, Q_2. If $a_k = \sigma^{(k+1)} f_{k-1}/Q_1$, $b_k = \sigma^{(L+k+2)} g_{k-1}/Q_2$, then a_k, b_k are thus assumed fixed. From (3.5), (3.7), the primes Q_1, Q_2 satisfy

(i) $a_k Q_1 + 1$, $b_k Q_2 + 1$ are primes,

(ii) $Q_1, Q_2 > \exp\{(\log x)^{(1-2\delta)\alpha^{k-1}}\}$.

Thus by Brun's method we have

$$\sum_{\text{(i),(ii)}} \frac{1}{Q_1 Q_2} = \sum \frac{1}{Q_1} \sum \frac{1}{Q_2} \leqslant \frac{c_5^2 (\log\log x)^2}{(\log x)^{2(1-2\delta)\alpha^{k-1}}}. \tag{3.8}$$

It thus remains to estimate $\sum 1/\tau_0^{(1)}$ for $\tau \in T_{k,1}$ with $\sigma\tau \in S_{k-1}$. We do this by first counting the number of such τ with $\tau_0^{(1)} = t_0$ and then estimating $\sum 1/t_0$. From Lemma 3.2, if $j = \Omega(\tau_0^{(1)})$ for some $\tau \in T_{k,1}$, then

$$|j - k(\alpha^{k-1} - \alpha^k)\log\log x| < \delta(k+1)\alpha^{k-1}\log\log x + 2. \tag{3.9}$$

From (3.6), the number of τ corresponding to a given t_0 is at most the number of dual factorizations of t_0 into k factors each. This is at most

$$k^{2\Omega(t_0)}. \tag{3.10}$$

We are now ready to estimate $\sum 1/t_0$. Since every prime in each t_0 exceeds z_k, the contribution to $\sum 1/t_0$ for those t_0 divisible by the square of a prime is at most $(\log x)/z_k$. The contribution to $\sum 1/t_0$ from the remaining (square-free) t_0 is at most

$$\sum_j{}' \frac{1}{j!} \Big(\sum_{z_k < p \leqslant z_{k-1}} \frac{1}{p} \Big)^j \leqslant \sum_j{}' \frac{s_k^j}{j!}$$

where $\sum'$ denotes a sum over j satisfying (3.9) and where

$$s_k = (\alpha^{k-1} - \alpha^k)\log\log x + 1.$$

Therefore, with (3.10) we have

$$\sum_{\substack{\tau \in T_{k,1} \\ \sigma\tau \in S_{k-1}}} \frac{1}{\tau_0^{(1)}} \leqslant \sum_j{}' k^{2j} \Big(\frac{s_k^j}{j!} + \frac{\log x}{z_k} \Big).$$

The expression $k^{2j}s_k^j/j!$ reaches its maximal value when $j = k^2 s_k$, but from (3.9) we are considering much smaller values of j, namely $j \approx k s_k$. Thus the sum is dominated by its largest term which is given by the largest j satisfying (3.9). A calculation then gives

$$\sum_{\substack{\tau \in T_{k,1} \\ \sigma\tau \in S_{k-1}}} \frac{1}{\tau_0^{(1)}} \leqslant (\log x)^{(\alpha^{k-1}-\alpha^k)(k\log k+k)+c_7\delta\alpha^k k\log k}$$

for some absolute constant c_7. From this estimate and (3.8) we have

$$\sum_{\substack{\tau \in T_{k,1} \\ \sigma\tau \in S_{k-1}}} \frac{1}{\tau^{(1)}} \leqslant (\log x)^{(\alpha^{k-1}-\alpha^k)(k\log k+k)-2\alpha^{k-1}+c_8\delta\alpha^k k\log k} \tag{3.11}$$

for some absolute constant c_8.

The estimate for $T_{k,2}$ goes exactly the same way except now we let $\tau_0^{(1)} = \tau^{(1)}/Q$ where $Q = P(f_{k-1}) = P(g_{k-1})$. The prime Q satisfies

(i′) $a_k Q+1$, $b_k Q+1$ are unequal primes,
(ii′) $Q > \exp\{(\log x)^{(1-2\delta)\alpha^{k-1}}\}$

in analogy to (i), (ii). As in (3.8), the sieve gives

$$\sum_{(\mathrm{i}'),(\mathrm{ii}')} \frac{1}{Q} \leqslant \frac{c_5(\log\log x)^2}{(\log x)^{2(1-2\delta)\alpha^{k-1}}}.$$

Thus an estimate of the same form as (3.11) also holds for $T_{k,2}$. We thus have the lemma.

Using (3.3), Lemma 3.3, and then sequentially Lemma 3.4 for $k = 2, \ldots, L$, we get

$$\# B_2^* \leqslant x(\log x)^{-2-\alpha+\delta+\sum_{k=2}^{L} b_k} \sum_{\sigma \in S_L} \frac{1}{\sigma^{(1)}} \tag{3.12}$$

where $b_k = (\alpha^{k-1}-\alpha^k)(k\log k+k)-2\alpha^{k-1}+c_6\delta\alpha^k k\log k$. It remains to estimate the two sums in (3.12).

Note that

$$-2-\alpha+\sum_{k=2}^{L}\left((\alpha^{k-1}-\alpha^k)(k\log k+k)-2\alpha^{k-1}\right) < -2+\sum_{k=1}^{L-1} a_k\alpha^k < -2+F(\alpha)$$

where the a_k's and F are given by (1.3). Thus there is some absolute constant c_9 with

$$-2-\alpha+\delta+\sum_{k=2}^{L} b_k \leqslant -2+F(\alpha)+c_9\delta. \tag{3.13}$$

Since $F(\alpha) < 1$, we may choose δ sufficiently small so that

$$-2+F(\alpha)+c_9\delta < -1. \tag{3.14}$$

For $\sum 1/\sigma^{(1)}$ in (3.12) note that if $\sigma \in S_L$, then

$$P(\sigma^{(1)}) \leqslant z_L < \exp\exp\{(\log\log x)^\delta\},$$

so that if s runs over integers of the form $\sigma^{(1)}$ for some $\sigma \in S_L$, then

$$\sum 1/s \ll \exp\{(\log\log x)^\delta\}.$$

Also note that from Proposition 2.2, if $\sigma \in S_L$, then

$$\Omega(\sigma^{(1)}) < (L+1)(1+2\delta)(\log\log x)^\delta.$$

Thus the number of $\sigma \in S_L$ with $\sigma^{(1)} = s$, a fixed integer, is at most

$$(L+1)^{2\Omega(s)} \leqslant \exp\{c_{10}(\log\log x)^\delta \log\log\log x \log\log\log\log x\}$$

for some absolute constant c_{10}. Thus

$$\sum_{\sigma \in S_L} \frac{1}{\sigma^{(1)}} < \exp\{(\log\log x)^{2\delta}\} \tag{3.15}$$

for large x.

Thus from (3.12)–(3.15), there is some $\varepsilon > 0$ such that

$$\# B_2^* \leqslant \frac{x}{(\log x)^{1+\varepsilon}}. \tag{3.16}$$

This estimate easily gives $\# B_2^* = o(\# A)$ and we are now nearly done with our proof of the lower bound in the theorem. It remains to show how the argument for B_2^* can be extended to all of B_2.

If $m|n$ for some $n \in A$, let

$$B_2(m) = \{(n_1, n_2) \in B_2;\ \gcd(n_1, n_2) = m\}.$$

Thus $B_2(1) = B_2^*$. If $(n_1, n_2) \in B_2(m)$, then

$$n_1 = p_{j_0}, \ldots, p_{j_T} m, \qquad n_2 = q_{j_0}, \ldots, q_{j_T} m$$

where $0 \leqslant j_0 < \ldots < j_T \leqslant L$ and $p_{j_k}, q_{j_k} \in P_{j_k}$, $p_{j_k} \neq q_{j_k}$ for each k. Tracing through the above estimate of B_2^* we arrive at

$$\# B_2(m) \leqslant \frac{x}{m}(\log x)^{-2\alpha^{j_0}-\alpha^{j_1}+\delta+\sum_{k=2}^{T} b_k'} \sum_{\sigma \in S_{L(m)}} \frac{1}{\sigma^{(1)}}$$

in analogy to (3.12), where

$$b_k' = (\alpha^{j_k-1}-\alpha^{j_k})(k\log k+k)-2\alpha^{j_k-1}+c_6\delta\alpha^{j_k-1}k\log k$$

and where $S_L(m)$ is defined in analogy to S_L. Since $b_k' \leqslant b_k$, we quickly arrive at

$$\#B_2(m) \leqslant \frac{x}{m(\log x)^{\alpha^{j_0}(1+\varepsilon)}} \tag{3.17}$$

for some fixed $\varepsilon > 0$, all possible values of m, and all $x \geqslant x_0(\alpha, \delta)$. Notice that j_0 is a function of m — it is the least number in $\{0, \ldots, L\}$ such that m has no prime factor in P_{j_0}. Suppose $(n_1, n_2) \in B_2(m)$. Then $P(n_1/m) \in P_{j_0}$ and in fact $n_1/m < z_{j_0}^2$, so that

$$x/m \leqslant 2n_1/m < 2z_{j_0}^2.$$

Taking logs, we get

$$\log(x/m) \ll (\log x)^{\alpha^{j_0}},$$

so that from (3.17), we get

$$\#B_2(m) \ll \frac{x}{m(\log(x/m))^{1+\varepsilon}} \tag{3.18}$$

uniformly.

We now sum (3.18) over all m which have at most one prime in each P_k and no other primes. As in Lemma 3.1, the number of such $m \leqslant z$ is at most

$$\frac{z}{\log z}\exp\{C(\log\log\log z)^2\}.$$

Also note that $m \leqslant x/w_L$ and that $w_L > \exp\exp\{(1-\delta)(\log\log x)^\delta\}$. With these observations and using partial summation (cf. the proof of Lemma 2 in [6]), we get

$$\#B_2 = \sum_m \#B_2(m) \ll \frac{x\log\log x}{\log x}\cdot\frac{\exp\{C(\log\log\log x)^2\}}{\exp\{\varepsilon(1-\delta)(\log\log x)^\delta\}} = o\left(\frac{x}{\log x}\right) = o(\#A).$$

This estimate then completes our proof of the lower bound.

4. The upper bound. In this section we indicate how the upper bound in the theorem can be proved. We follow the argument in [6], but with a few changes.

As in [6], define $w(x)$ by the equation

$$V(x) = \frac{x}{\log x} w(x)$$

and let

$$W(x) = \sup_{2 \leqslant y \leqslant x} w(y).$$

Let β be an arbitrary, but fixed number with $c_0 < \beta < 1$ where c_0 is given by (1.4) above. Let $k \geqslant 3$ denote a fixed large integer. Let $\delta > 0$ be fixed and small. How large k must be and how small δ must be shall be described later, but this shall depend on the choice of β.

The following lemmas are in analogy to Lemmas 1–3 in [6].

LEMMA 4.1. *The number $V_1(x)$ of distinct values of $\varphi(n) \leqslant x$ such that either* (i) $\Omega(n) \leqslant k+1$, (ii) $n \leqslant x/\log x$, (iii) $d^2 | n$ *for some* $d \geqslant \log x$, *or* (iv) $d^2 | \varphi(n)$ *for some* $d \geqslant \log x$ *is* $o(V(x))$.

The proof is the same as for Lemma 1 in [6].

LEMMA 4.2. *The number $V_2(x)$ of distinct values of $\varphi(n) \leqslant x$ such that n is divisible by a prime $p > \exp\{(\log x)^{\beta^k}\}$ which is not δ,$\log x$-normal is* $o\left(\frac{x}{\log x} W(x)\right)$.

The proof follows from the same method as in [6], but now we use Proposition 2.2 above.

For $i \geqslant 1$, let $P_i(n)$ denote the i-th largest prime factor of n if $\Omega(n) \geqslant i$. Otherwise, let $P_i(n) = 1$.

LEMMA 4.3. *The number $V_3(x)$ of distinct values of $\varphi(n) \leqslant x$ such that*

$$P_i(n) > \exp\{(\log x)^{\beta^{i-1}}/\log\log x\} \quad \textit{for} \quad i = 1, \ldots, k \tag{4.1}$$

is $o\left(\frac{x}{\log x} W(x)\right)$.

Proof. For $j = 0, 1, \ldots, k$, let $y_j = \exp\{(\log x)^{\beta^j}\}$. From Lemma 4.2, we may assume that the primes $P_1(n), \ldots, P_k(n)$ are all δ, $\log x$-normal primes. Thus if n satisfies (4.1) and $l_j = \Omega(\varphi(n), y_j, y_{j-1})$, we have

$$|l_j - j(\beta^{j-1} - \beta^j)\log\log x| < (j+1)\delta\beta^{j-1}\log\log x \tag{4.2}$$

for $j = 1, \ldots, k$ and all large x. From Lemma 4.1, we may assume that all prime factors of $\varphi(n)$ above y_k appear to exactly the first power. Let L_j denote the set of l which satisfy (4.2) for a given j. Let M denote the set of m such that m is square-free, m is only divisible by primes in $(y_k, y_1]$ and $\Omega(m, y_j, y_{j-1}) \in L_j$ for $j = 2, \ldots, k$. Thus if $m \in M$ we have

$$m \leqslant y_1^{O(\log\log x)} = x^{o(1)}. \tag{4.3}$$

If $\varphi(n) \leqslant x$ for some n satisfying (4.1), then but for $o\left(\frac{x}{\log x} W(x)\right)$ exceptions, we may assume $\varphi(n) = mt$ where $m \in M$ and t is not divisible by any prime in $(y_k, y_1]$. From (4.3) and the fundamental lemma of the combinatorial sieve ([4], Th. 2.5), the number of such $t \leqslant x/m$ is uniformly at most

$$2\frac{x}{m}\cdot\frac{\log y_k}{\log y_1} = 2\frac{x}{m}(\log x)^{\beta^k - \beta}$$

for large x. (The constant 2 can be replaced by any number larger than 1.) Thus for large x,

$$V_3(x) \leqslant 2x(\log x)^{\beta^k-\beta} \sum_{m\in M} \frac{1}{m} + o\left(\frac{x}{\log x} W(x)\right). \tag{4.4}$$

Now

$$\sum_{m\in M} \frac{1}{m} \leqslant \prod_{j=2}^{k} \sum_{l\in L_j} \frac{1}{l!}\left(\sum_{y_j<p\leqslant y_{j-1}} \frac{1}{p}\right)^l$$
$$\leqslant \prod_{j=2}^{k} (\log x)^{-(\beta^{j-1}-\beta^j)(j\log j-j)+c_{11}\delta\beta^{j-1}j\log(j+1)}$$
$$\leqslant (\log x)^{-(\sum_{j=2}^{k}(\beta^{j-1}-\beta^j)(j\log j-j))+c_{12}\delta},$$

where c_{11}, c_{12} are absolute constants. Thus

$$(\log x)^{\beta^k-\beta} \sum_{m\in M} \frac{1}{m} \leqslant (\log x)^{-(\sum_{j=1}^{k-1} a_j\beta^j)+\beta^k(k\log k-k+1)+c_{12}\delta}, \tag{4.5}$$

where the a_j's are given by (1.3).

The result (4.5) instructs us how to choose k, δ. We choose k so large and $\delta>0$ so small that

$$-(\sum_{j=1}^{k-1} a_j\beta^j)+\beta^k(k\log k-k+1)+c_{12}\delta < -1.$$

This can be done since

$$F(\beta) = \sum_{j=1}^{\infty} a_j\beta^j > 1.$$

Thus the lemma follows from (4.4) and (4.5).

The remainder of the proof follows exactly the same way as in [6]. We thus obtain

$$V(x) \leqslant \frac{x}{\log x}\exp\left\{\frac{(\log\log\log x)^2}{2|\log\beta|}+O(\log\log\log x)\right\}.$$

Since the value of β chosen above can be arbitrarily close to c_0, we obtain the upper bound in the theorem, concluding our proof.

5. Further questions. Almost surely $V(x)$ is a smooth function. For example, Erdős and Hall have asked if $V(2x)/V(x)\to 2$ holds. Our proof does show

$$V(2x)-V(x) = \frac{x}{\log x}\exp\{(C+o(1))(\log\log\log x)^2\},$$

but this is too weak to establish $V(2x)/V(x)\to 2$. Perhaps there is some other attack on expressions such as $V(2x)/V(x)$.

Concerning our proof of the lower bound, we assumed every member of A exceeds $x/2$. Thus if $A' = \{2n\colon n\in A\}$, then every member of A' is between x and $2x$ and $\varphi(2n) = \varphi(n) \leqslant x$ for $2n\in A'$. It thus follows that if $V'(x)$ denotes the cardinality of $\{\varphi(n)\leqslant x\colon n>x\}$, then

$$V'(x) = \frac{x}{\log x}\exp\{(C+o(1))(\log\log\log x)^2\}.$$

A harder question is to estimate $V_0(x)$, the cardinality of

$$\{\varphi(n)\leqslant x\colon n>x\}-\{\varphi(n)\colon n\leqslant x\}.$$

Our methods do not give a good lower bound for $V_0(x)$.

Our proof works equally well for values of the sum of the divisors function $\sigma(n)$ and for many similar arithmetic functions. In fact the same estimate can be obtained for the number of distinct integers which are products of the members of $\{p+a\colon p$ is prime, $a\in S\}$, where S is any finite set of non-zero integers.

References

[1] P. Erdős, *On the normal number of prime factors of $p-1$ and some related questions concerning Euler's φ-function*, Quart. J. Math. (Oxford Ser.) 6 (1935), pp. 205–213.
[2] P. Erdős and R. R. Hall, *On the values of Euler's φ-function*, Acta Arith. 22 (1973), pp. 201–206.
[3] — — *Distinct values of Euler's φ-function*, Mathematika 23 (1976), pp. 1–3.
[4] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London 1974.
[5] S. S. Pillai, *On some functions connected with $\varphi(n)$*, Bull. Amer. Math. Soc. 35 (1929), pp. 832–836.
[6] C. Pomerance, *On the distribution of the values of Euler's function*, Acta Arith. 47 (1986), pp. 63–70.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
Athens, GA 30602
USA

Received on 18.4.1986 (1622)