ELECTRONIC RESEARCH ANNOUNCEMENTS
OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 4, Pages 27–34 (April 27, 1998)
S 1079-6762(98)00043-2

THE DISTRIBUTION OF TOTIENTS

KEVIN FORD

(Communicated by Hugh Montgomery)

ABSTRACT. This paper is an announcement of many new results concerning the set of totients, i.e. the set of values taken by Euler's ϕ-function. The main functions studied are $V(x)$, the number of totients not exceeding x, $A(m)$, the number of solutions of $\phi(x) = m$ (the "multiplicity" of m), and $V_k(x)$, the number of $m \le x$ with $A(m) = k$. The first of the main results of the paper is a determination of the true order of $V(x)$. It is also shown that for each $k \ge 1$, if there is a totient with multiplicity k, then $V_k(x) \gg V(x)$. We further show that every multiplicity $k \ge 2$ is possible, settling an old conjecture of Sierpiński. An older conjecture of Carmichael states that no totient has multiplicity 1. This remains an open problem, but some progress can be reported. In particular, the results stated above imply that if there is one counterexample, then a positive proportion of all totients are counterexamples. Determining the order of $V(x)$ and $V_k(x)$ also provides a description of the "normal" multiplicative structure of totients. This takes the form of bounds on the sizes of the prime factors of a pre-image of a typical totient. One corollary is that the normal number of prime factors of a totient $\le x$ is $c \log\log x$, where $c \approx 2.186$. Lastly, similar results are proved for the set of values taken by a general multiplicative arithmetic function, such as the sum of divisors function, whose behavior is similar to that of Euler's function.

Let $\mathcal{V}$ denote the set of values taken by Euler's ϕ-function (totients), i.e.

$$\mathcal{V} = \{1, 2, 4, 6, 8, 10, 12, 16, 18, 20, 22, 24, 28, 30, \dots\}.$$

Let

$$\begin{aligned} \mathcal{V}(x) &= \mathcal{V} \cap [1, x], \\ V(x) &= |\mathcal{V}(x)|, \\ A(m) &= |\{n : \phi(n) = m\}|, \\ V_k(x) &= |\{m \le x : A(m) = k\}|. \end{aligned} \tag{1}$$

We will refer to $A(m)$ as the multiplicity of m. This paper summarizes recent work on the following questions.

1. What is the order of $V(x)$?
2. What is the order of $V_k(x)$ when the multiplicity k is possible?
3. What multiplicities are possible?
4. What is the normal multiplicative structure of totients?

Received by the editors August 13, 1997.
1991 *Mathematics Subject Classification.* Primary 11A25, 11N64; Secondary 11N35.
Key words and phrases. Euler's function, totients, distributions, Carmichael's conjecture, Sierpiński's conjecture.

1. The Prime Number Theorem implies $V(x) \gg x/\log x$, since $\phi(p) = p-1$ for primes p. Pillai [Pi] gave the first non-trivial bound on $V(x)$, namely

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

Using sieve methods, Erdös [E1] improved this to

$$V(x) \ll_\varepsilon \frac{x}{(\log x)^{1-\varepsilon}}$$

and later in [E2] showed that

$$V(x) \gg \frac{x \log_2 x}{\log x}.$$

Here and throughout this paper $\log_k x$ denotes the kth iterate of the logarithm. Further sharpening of the upper and lower bounds were found by Erdös and Hall [EH1], [EH2], Pomerance [P1], and Maier and Pomerance [MP]. The last result is

$$V(x) = \frac{x}{\log x} \exp\{(C+o(1))(\log_3 x)^2\}, \tag{2}$$

where C is a constant defined as follows. Let

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

Since $a_n \sim \log n$ and $a_n > 0$, it follows 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 ϱ such that

$$F(\varrho) = 1 \qquad (\varrho = 0.542598586098471021959\ldots)$$

and we set

$$C = \frac{1}{2|\log \varrho|} = 0.81781464640083632231\ldots.$$

Our first major result is a determination of the true order of $V(x)$. First define

$$\begin{aligned} D &= 2C(1 + \log F'(\varrho) - \log(2C)) - 3/2 \\ &= 2.17696874355941032173\ldots. \end{aligned}$$

Theorem 1 ([F1, Theorem 1]). *We have*

$$V(x) = \frac{x}{\log x} \exp\{C(\log_3 x - \log_4 x)^2 + D\log_3 x - (D + 1/2 - 2C)\log_4 x + O(1)\}.$$

2. Erdös [E3] showed by sieve methods that if $A(m) = k$, then for most primes p, $A(m(p-1)) = k$. If the multiplicity k is possible, it follows immediately that $V_k(x) \gg x/\log x$. Applying the machinery used to prove Theorem 1, we show that for each k, either $V_k(x) = 0$ for all x, or $V_k(x)$ is the same order as $V(x)$.

Theorem 2 ([F1, Theorem 2]). *If there is a number d with $A(d) = k$, then*

$$V_k(x) \gg_\varepsilon d^{-1-\varepsilon} V(x) \qquad (x \ge x_0(k)).$$

In other words, a positive fraction of totients have multiplicity k if the multiplicity k is possible. This suggests that the multiplicity of "most" totients is bounded.

Theorem 3 ([F1, Theorem 3]). *We have*

$$\frac{|\{m \in \mathcal{V}(x) : A(m) \ge N\}|}{V(x)} = \sum_{k \ge N} \frac{V_k(x)}{V(x)} \ll N^{-1} \exp\{O(\sqrt{\log N})\}.$$

3. In 1907, Carmichael [C1] announced that for every m, the equation $\phi(x) = m$ has either no solutions x or at least two solutions. In other words, no totient can have multiplicity 1. His proof of this assertion was flawed, however, but he did show in [C2] that no number $m < 10^{37}$ has multiplicity 1, and conjectured that no such m exists (this is now known as Carmichael's Conjecture). Improved lower bounds for a possible counterexample have been found by Klee [K], Masai and Valette [MV], and recently by Schlafly and Wagon [SW], who showed that a counterexample must exceed $10^{10,000,000}$. Carmichael's Conjecture, however, remains an open problem.

An immediate and important corollary of Theorems 1 and 2 is

Theorem 4 ([F1, Theorem 5]). *We have*

$$\limsup_{x \to \infty} \frac{V_1(x)}{V(x)} < 1.$$

Furthermore, Carmichael's Conjecture is equivalent to the bound

$$\liminf_{x \to \infty} \frac{V_1(x)}{V(x)} = 0.$$

Although this is a long way from proving Carmichael's Conjecture, Theorem 4 shows that the set of counterexamples cannot be a "thin" subset of $\mathcal{V}$. Either there are no counterexamples or a positive fraction of totients are counterexamples.

The basis for the computations of lower bounds for a possible counterexample is a lemma of Carmichael, generalized by Klee, which allows one to show that if $A(m) = 1$, then x must be divisible by the squares of many primes. Using the method outlined in [SW] and modern computer hardware, we push the lower bound for a counterexample to an aesthetically pleasing bound.

Theorem 5 ([F1, Theorem 6]). *If $A(m) = 1$, then m exceeds $10^{10^{10}}$.*

Aside from computations, the only other non-trivial result concerning the behavior of $V_1(x)$ is the bound

$$\liminf_{x \to \infty} \frac{V_1(x)}{V(x)} \le \frac{1}{2},$$

proved by elementary methods in an unpublished note of Pomerance (see [P2]). A modification of his argument combined with the computations giving Theorem 5 yields

Theorem 6 ([F1, Theorem 7]). *We have*

$$\liminf_{x \to \infty} \frac{V_1(x)}{V(x)} \le 10^{-5,000,000,000}.$$

In the 1950's, Sierpiński conjectured that all multiplicities $k \ge 2$ are possible (see [S1] and [E3]), and in 1961, Schinzel [S2] deduced this conjecture from his well-known Hypothesis H. Schinzel's Hypothesis H [SS], a generalization of Dickson's Prime k-tuples Conjecture [D], states that any set of polynomials $F_1(n), \ldots, F_k(n)$, subject to natural restrictions, are simultaneously prime for infinitely many n. Using the modern theory of "almost primes" (numbers with few prime factors) together with an iterative approach, we provide an unconditional proof of Sierpiński's Conjecture.

Theorem 7 ([F2, Theorem 1]). *For each $k \ge 2$, there is a number d with $A(d) = k$.*

Therefore, by Theorem 2, $V_k(x) \gg V(x)$ for $k \ge 2$. The computations leading to Theorems 5 and 6 motivate another classification of totients. Let $V(x;k)$ be the number of totients up to x, all of whose pre-images are divisible by k. A corollary of the proof of Theorem 2 is

Theorem 8 ([F1, Theorem 8]). *If d is a totient, all of whose pre-images are divisible by k, then*

$$V(x;k) \gg_\varepsilon d^{-1-\varepsilon} V(x).$$

Thus, for each k, either $V(x;k) = 0$ for all x or $V(x;k) \gg_k V(x)$.

It is natural to ask for which k do there exist totients, all of whose pre-images are divisible by k. A short search reveals examples for each $k \le 11$ except $k = 6$ and $k = 10$. For $k = 2, 4$ and 8, take $d = 2^{18} \cdot 257$, for $k = 3$ or 9 take $d = 54 = 2 \cdot 3^3$, for $k = 5$ take $d = 12500 = 4 \cdot 5^5$, for $k = 7$, take $d = 294 = 6 \cdot 7^2$ and for $k = 11$, take $d = 110$. It appears that there might not be any totient, all of whose pre-images are divisible by 6, but I cannot prove this.

Conjecture. *There is no totient, all of whose pre-images are divisible by 6.*

In particular, any totient with a unique pre-image must have that pre-image divisible by 6, so the non-existance of such numbers implies Carmichael's Conjecture.

4. Establishing Theorems 1 and 2 requires a determination of what a "normal" totient looks like. This will initially take the form of a series of linear inequalities in the prime factors of a pre-image of a totient. An analysis of these inequalities reveals the normal sizes of the prime factors of a pre-image of a typical totient.

To state our results, we first define

$$L_0 = L_0(x) = [2C(\log_3 x - \log_4 x)].$$

In a simplified form, we show that for all but $o(V(x))$ totients $m \le x$, every pre-image n satisfies

$$\log_2 q_i(n) \sim \varrho^i (1 - i/L_0) \log_2 x \qquad (0 \le i \le L_0), \tag{3}$$

where $q_i(n)$ denotes the $(i+1)$st largest prime factor of n. Let $\omega(m)$ denote the number of distinct prime factors of m and let $\Omega(m)$ denote the number of prime factors of m counted with multiplicity. Then we have

Theorem 9 ([F1, Theorem 11]). *Suppose $g(x)$ is an increasing function of x satisfying $g(x) = o(\log_3 x)$. For a given x, set $L_0 = L_0(x)$ and $\beta_i = \varrho^i(1 - i/L_0)$ for $0 \le i \le L_0$. Then the number of totients $m \le x$ with a pre-image n not satisfying*

$$\left| \frac{\log_2 q_i(n)}{\beta_i \log_2 x} - 1 \right| \le \sqrt{\frac{i}{L_0(L_0 - i)}} \log(L_0 - i) \log g(x) \qquad (1 \le i \le L_0 - g(x)) \tag{4}$$

and

$$L_0(x) - g(x) \le \omega(n) \le \Omega(n) \le L_0(x) + 30 g(x) \varrho^{-g(x)}$$

is

$$\ll V(x)\left(e^{-\log^2 g(x)} + e^{-\frac{1}{4}\log_4 x \log g(x)}\right).$$

In essence, Theorem 9 says that the set of $n \le x$ having about $L_0(x)$ prime factors distributed according to (4) generates almost all totients. It also says that for most totients, all of its pre-images are "virtually" square-free. The function $g(x)$ need not tend to infinity. Notice that the intervals in (4) are not only disjoint, but the gaps between them are rather large. In particular, this "discreteness phenomenon" means that for most totients $m \le x$, no pre-image n has any prime factors in the intervals

$$0.999 \ge \frac{\log_2 p}{\log_2 x} \ge 0.543, \quad 0.542 \ge \frac{\log_2 p}{\log_2 x} \ge 0.295, \text{ etc.}$$

This should be compared to the distribution of the prime factors of a normal integer $n \le x$ (e.g. Theorem 12 of [HT]).

We also deduce the normal order of $\Omega(m)$ and $\omega(m)$ for totients m. If each prime $q_i(n)$ of a pre-image n is "normal" and (3) holds, then $\Omega(m)$ should be about

$$(1+\varrho+\varrho^2+\cdots)\log_2 x = \frac{\log_2 x}{1-\varrho}.$$

Theorem 10 ([F1, Theorem 12]). *Suppose* $\varepsilon = \varepsilon(x)$ *satisfies* $0 \le \varepsilon \le 0.8$. *Then*

$$\#\left\{m \in \mathcal{V}(x) : \left|\frac{\Omega(m)}{\log_2 x} - \frac{1}{1-\varrho}\right| \ge \varepsilon\right\} \ll V(x)\exp\{-K\varepsilon\log_3 x + O(\sqrt{\varepsilon \log_3 x})\},$$

where

$$K = \frac{2Ca_1(1-\varrho)}{1-(1+a_1)\varrho} = 1.166277\ldots.$$

Consequently, if $g(x) \to \infty$ *as slowly as desired, then almost all totients* $m \le x$ *satisfy*

$$\left|\frac{\Omega(m)}{\log_2 x} - \frac{1}{1-\varrho}\right| \le \frac{g(x)}{\log_3 x}.$$

Moreover, the theorem holds with $\Omega(m)$ *replaced by* $\omega(m)$.

Corollary 11 ([F1, Corollary 13]). *If either* $h(m) = \omega(m)$ *or* $h(m) = \Omega(m)$, *then*

$$\sum_{m \in \mathcal{V}(x)} h(m) = \frac{V(x)\log_2 x}{1-\varrho}\left(1 + O\left(\frac{1}{\log_3 x}\right)\right).$$

By contrast, Erdös and Pomerance [EP] showed that the average of $\Omega(\phi(n))$, where the average is taken over all $n \le x$, is $\frac{1}{2}(\log_2 x)^2 + O((\log_2 x)^{3/2})$.

The details of the proofs of these results are extremely complex and require very delicate estimating, but the central ideas are fairly simple. First, for most integers m, the prime divisors of m are "nicely distributed", meaning the number of prime factors of m lying between a and b is about $\log_2 b - \log_2 a$. This is a more precise version of the classical result of Hardy and Ramanujan [HR] that most numbers m have about $\log_2 m$ prime factors. Take an integer n with prime factorization $p_0 p_1 \cdots$, where for simplicity we assume n is square-free, and $p_0 > p_1 > \cdots$. By sieve methods it can be shown that for most primes p, the prime divisors of $p-1$ have the same "nice" distribution. If $p_0, p_1, \ldots$ are such "normal" primes, it follows that $\phi(n) = (p_0 - 1)(p_1 - 1)\cdots$ has about $\log_2 n - \log_2 p_1$ prime factors in $[p_1, n]$, about $2(\log_2 p_1 - \log_2 p_2)$ prime factors in $[p_2, p_1]$, and in general, $\phi(n)$ will have $k(\log_2 p_{k-1} - \log_2 p_k)$ prime factors in $[p_k, p_{k-1}]$. That is, n has k times as many prime factors in the interval $[p_k, p_{k-1}]$ as does a "normal" integer of its size. If n has many "large" prime divisors, then the prime factors of $m = \phi(n)$ will be much denser than normal, and the number, N_1, of such integers m will be "small". On the other hand, the number, N_2, of integers n with relatively few "large" prime factors is also "small". Our objective then is to precisely define these concepts of "large" and "small" so as to minimize $N_1 + N_2$.

The argument in [MP] is based on the heuristic that a normal totient is generated from a number n satisfying

$$\log_2 q_i(n) \approx \varrho^i \log_2 x \tag{5}$$

for each i (compare with (3)). As an alternative to this heuristic, assuming all prime factors of a pre-image n of a totient are normal leads to consideration of a series of inequalities among the prime factors of n. Specifically, let x be large, $x/2 < n \le x$ and let $q_0 \ge q_1 \ge \cdots$ denote the prime factors of n. Fix L and map n to the point $(x_1, \ldots, x_L)$, where $x_i = \log_2 q_i / \log_2 x$ if $i < \Omega(n)$, and $q_i \ge 3$, $x_i = 0$ otherwise. Consider the following inequalities:

$$\begin{aligned}
0 \le x_L \le \cdots &\le x_1 \le 1, \\
a_1 x_1 + \cdots + a_L x_L &\le 1, \\
a_1 x_2 + \cdots + a_{L-1} x_L &\le x_1, \\
a_1 x_3 + \cdots + a_{L-2} x_L &\le x_2, \\
&\;\;\vdots \\
a_1 x_{L-1} + a_2 x_L &\le x_{L-2}.
\end{aligned}$$

We show that such n generate "most" totients and then reduce the problem of counting such n to the problem of finding the volume of the region in $\mathbb{R}^L$ defined by these inequalities, denoted $\mathcal{S}_L$. What we show is essentially

$$V(x) \approx \frac{x}{\log x} \max_L \mathrm{Vol}(\mathcal{S}_L)(\log_2 x)^L. \tag{6}$$

With the estimate

$$\mathrm{Vol}(\mathcal{S}_L) \approx \frac{\varrho^{L(L+3)/2}}{L!}(F'(\varrho))^L,$$

the maximum in (6) occurs at about $L = L_0(x)$ and Theorem 1 follows. Careful analysis of these inequalities reveals that the bulk of $\mathcal{S}_L$ is concentrated near the point $(\beta_1, \ldots, \beta_L)$, where $\beta_i = (1 - \frac{i}{L})\varrho^i$. It follows that "most" of the integers n mapping into $\mathcal{S}_L$ satisfy (3). Thus, the heuristic (5) gives numbers n whose smaller prime factors are too large. In particular, with $L = L_0(x)$, we have $\log_2 p_{L-1} \approx \frac{1}{L}\varrho^{L-1}\log_2 x \approx 1$, an important observation for determining the true order of $V(x)$.

Lastly, most of these results may be easily extended to more general multiplicative arithmetic functions such as $\sigma(n)$, the sum of divisors function. Defining analogous functions $\mathcal{V}_f$, $V_f(x)$, $V_{f,k}(x)$ and $A_f(m)$, we prove

Theorem 12 ([F1, Theorem 14]). *Suppose $f : \mathbb{N} \to \mathbb{N}$ is a multiplicative function satisfying*

$$\{f(p) - p : p \text{ prime}\} \text{ is a finite set not containing } 0,$$
$$\sum_{\substack{h \ge 16 \\ h \text{ square-full}}} \frac{\varepsilon(h)}{f(h)} \ll 1, \qquad \varepsilon(h) = \exp\{\log_2 h (\log_3 h)^{20}\}.$$

Then the analogs of Theorems 1–3, 7 and 9–11 hold with $f(n)$ replacing $\phi(n)$, with the exception of the dependence on d in Theorems 2 and 3, which may be different.

The possible set of multiplicities for each f, however, depends on the particular function, since the values of $f(p^k)$ for $k \ge 2$ play a more important role. In particular, the proof of Theorem 7 relies on the property that $\phi(p^2) = p\phi(p)$ for primes p. However, when $f = \sigma$, the analog of Theorem 7 has been proved ([FK, Theorem 1]) and the method extends to many other functions satisfying the hypotheses of Theorem 12 for which $A_f(m) = 1$ has a solution.

References

[C1] R. D. Carmichael, *On Euler's ϕ-function*, Bull. Amer. Math. Soc. **13** (1907), 241–243.

[C2] ———, *Note on Euler's ϕ-function*, Bull. Amer. Math. Soc. **28** (1922), 109–110.

[D] L. E. Dickson, *A new extension of Dirichlet's theorem on prime numbers*, Messenger of Math. **33** (1904), 155–161.

[E1] P. Erdös, *On the normal number of prime factors of $p-1$ and some related problems concerning Euler's ϕ-function*, Quart. J. Math. Oxford (1935), 205–213.

[E2] ———, *Some remarks on Euler's ϕ-function and some related problems*, Bull. Amer. Math. Soc. **51** (1945), 540–544. MR **7**:49f

[E3] ———, *Some remarks on Euler's ϕ-function*, Acta Arith. vol 4 (1958), 10–19. MR **22**:1539

[EH1] P. Erdös and R. R. Hall, *On the values of Euler's ϕ-function*, Acta Arith. **22** (1973), 201-206. MR **53**:13143

[EH2] ———, *Distinct values of Euler's ϕ-function*, Mathematika **23** (1976), 1–3. MR **54**:2603

[EP] P. Erdös and C. Pomerance, *On the normal number of prime factors of $\phi(n)$*, Rocky Mountain J. of Math. **15** (1985), 343–352. MR **87e**:11112

[F1] K. Ford, *The distribution of totients*, The Ramanujan J. **2** (1998), no. 1-2 (to appear).

[F2] ———, *The number of solutions of $\phi(x) = m$*, Annals of Math. (to appear).

[FK] K. Ford and S. Konyagin, *On a conjecture of Sierpiński concerning the sum of divisors function*, Proceedings of the International Number Theory Conference dedicated to Professor Andrzej Schinzel, Zakopane, Poland (to appear).

[HT] R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge University Press, 1988. MR **90a**:11107

[HR] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quart. J. Math. **48** (1917), 76–92.

[K] V. Klee, *On a conjecture of Carmichael*, Bull. Amer. Math. Soc. **53** (1947), 1183–1186. MR **9**:269d

[MP] H. Maier and C. Pomerance, *On the number of distinct values of Euler's ϕ-function*, Acta Arith. **49** (1988), 263–275. MR **89d**:11083

[MV] P. Masai and A. Valette, *A lower bound for a counterexample to Carmichael's Conjecture*, Boll. Un. Mat. Ital. A (6) **1** (1982), 313–316. MR **84b**:10008

[Pi] S. Pillai, *On some functions connected with $\phi(n)$*, Bull. Amer. Math. Soc. **35** (1929), 832–836.

[P1] C. Pomerance, *On the distribution of the values of Euler's function*, Acta Arith. **47** (1986), 63–70. MR **88b**:11060

[P2] ———, *Problem 6671*, Amer. Math. Monthly **98** (1991), 862.

[S1] A. Schinzel, *Sur l'équation $\phi(x) = m$*, Elem. Math. **11** (1956), 75–78. MR **18**:194c

[S2] ———, *Remarks on the paper "Sur certaines hypothèses concernant les nombres premiers"*, Acta Arith. **7** (1961/62), 1–8. MR **24**:A70

[SS] A. Schinzel and W. Sierpiński, *Sur certaines hypothèses concernant les nombres premiers*, Acta Arith. **4** (1958), 185–208. MR **21:**4936

[SW] A. Schlafly and S. Wagon, *Carmichael's conjecture on the Euler function is valid below* $10^{10,000,000}$, Math. Comp. **63** (1994), 415–419. MR **94i:**11008

Department of Mathematics, University of Texas at Austin, Austin, TX 78712
E-mail address: `ford@math.utexas.edu`