Reprinted from

Analytic Number Theory

Proceedings of a Conference in Honor of Paul T. Bateman

Edited by
Bruce C. Berndt Harold G. Diamond
Heini Halberstam Adolf Hildebrand

Printed in the United States of America

1990

Birkhäuser
Boston • Basel • Berlin

On the Normal Behavior of the Iterates Of some Arithmetic Functions

P. ERDÖS, A. GRANVILLE[1], C. POMERANCE[2], AND C. SPIRO

Dedicated to our friend, colleague and teacher, Paul Bateman

Abstract

Let $\varphi_1(n) = \varphi(n)$ where φ is Euler's function, let $\varphi_2(n) = \varphi(\varphi(n))$, etc. We prove several theorems about the normal order of $\varphi_k(n)$ and state some open problems. In particular, we show that the normal order of $\varphi_k(n)/\varphi_{k+1}(n)$ is $ke^\gamma \log\log\log n$ where γ is Euler's constant. We also show that there is some positive constant c such that for all n, but for a set of asymptotic density 0 , there is some k with $\varphi_k(n)$ divisible by every prime up to $(\log n)^c$. With $k(n)$ the first subscript k with $\varphi_k(n) = 1$, we show, conditional on a certain form of the Elliott-Halberstam conjecture, that there is some positive constant α such that $k(n)$ has normal order $\alpha \log n$. Let $s(n) = \sigma(n) - n$ where σ is the sum of the divisors function, let $s_2(n) = s(s(n))$, etc. We prove that $s_2(n)/s(n) = s(n)/n + o(1)$ on a set of asymptotic density 1 and conjecture the same is true for $s_{k+1}(n)/s_k(n)$ for any fixed k .

§1. Introduction

Let $\varphi(n) = \varphi_1(n)$ denote Euler's phi-function and if $\varphi_{k-1}(n)$ has already been defined, let $\varphi_k(n) = \varphi(\varphi_{k-1}(n))$. If $n > 1$, then $n > \varphi(n)$. Thus the sequence $n, \varphi_1(n), \varphi_2(n), \cdots$ is strictly decreasing until it reaches 1 when it becomes constant. Let $k(n) = k$ be the least number such that $\varphi_k(n) = 1$. Further, let $k(1) = 1$.

[1] Supported in part by an NSERC grant
[2] Supported in part by an NSF grant

Note that if $n = 2^j$, then $k(n) = j = (\log n)/\log 2$. Also if $n = 2 \cdot 3^j$, then $k(n) = j + 1 = \lceil (\log n)/\log 3 \rceil$ where $\lceil x \rceil$ denotes the least integer $\geq x$. It turns out that these two examples essentially demonstrate the extreme behavior of $k(n)$, for as Pillai [14] showed in 1929,

$$\lceil (\log n)/\log 3 \rceil \leq k(n) \leq \lceil (\log n)/\log 2 \rceil \tag{1.1}$$

for all n. Further, by considering numbers n of the form $2^a 3^b$ it is easy to see that the set of numbers of the form $k(n)/\log n$ is dense in $[1/\log 3, 1/\log 2]$. What is still in doubt about $k(n)$ is its average and normal behavior. We conjecture that there is some constant α such that $k(n) \sim \alpha \log n$ on a set of asymptotic density 1. If this is true, then (1.1) immediately would imply that

$$\frac{1}{x} \sum_{n \leq x} k(n) \sim \alpha \log x.$$

The function $k(n)$ possesses more algebraic structure than is immediately apparent from its definition. Shapiro [16] has shown that the function $g(n) := k(n) - 1$ is additive and in fact satisfies the stronger relation

$$g(mn) = g(m) + g(n) + \epsilon_{(m,n)}$$

for all natural numbers m, n where $\epsilon_{(m,n)}$ is 0 unless (m, n) is even in which case $\epsilon_{(m,n)} = 1$.

Let $F(n)$ denote the number of even terms of the sequence

$$n, \ \varphi(n), \ \varphi_2(n), \ \cdots.$$

Then $F(n) = k(n)$ for n even and $F(n) = k(n) - 1$ for n odd. It is not hard to show (we leave this for the reader) that the function $F(n)$ is completely additive; that is, $F(mn) = F(m) + F(n)$ for all natural numbers m, n. Note that $F(2) = 1$ and for p an odd prime, $F(p) = F(p-1)$. So in fact, we have an alternative definition of F that does not have anything to do with iterating the phi-function. Namely, F is the completely additive function which is defined inductively on the primes as follows:

$$F(p) = \begin{cases} 1, & \text{if } p = 2 \\ F(p-1), & \text{if } p > 2. \end{cases}$$

Thus our conjectures on the normal and average orders of $k(n)$ can be equivalently put in terms of the normal and average orders of the function $F(n)$. Using this translation of the problem, we are able to prove

these conjectures conditionally on a certain form of the Elliott-Halberstam conjecture. This conjecture states that for any A,

$$\sum_{k\le Q}\max_{(a,k)=1}\max_{x'\le x}\left|\pi(x';k,a)-\frac{\pi(x')}{\varphi(k)}\right|\ll_A \frac{x}{\log^A x}, \tag{1.2}$$

where $\pi(x;k,a)$ denotes the number of primes $p\le x$ with $p\equiv a \pmod k$, where $\pi(x)$ is the number of primes $p\le x$, and where Q is some function of the form $x^{1-o(1)}$. This conjecture for $Q=x/\log^B x$, which was the original conjecture of Elliott and Halberstam, was recently disproved in [4], while in [5] the conjecture is disproved for $Q=x/\exp\{c(\log\log x)^2/\log\log\log x\}$ for some positive constant c. But presumably, if $Q=x^{1-\epsilon(x)}$ and $\epsilon(x)$ tends to 0 slowly enough, then (1.2) holds. In section 2 below, we show that $F(n)$ (and thus $k(n)$) possesses normal and average order $\alpha\log n$ provided (1.2) holds for $Q=x^{1-\epsilon(x)}$ with $\epsilon(x)=(\log\log x)^{-2}$. Further we can weaken (1.2) by deleting the double max (letting $x'=x$ and $a=1$), by restricting k to integers with at most two prime factors, and taking $A=2$.

Short of proving our conjecture on the normal order of $k(n)$ unconditionally, there are still many interesting questions about the normal behavior of the functions $\varphi_k(n)$. In 1928, Schoenberg [15] showed that $n/\varphi(n)$ has a distribution function. That is, $D_\varphi(u)$, defined as the asymptotic density of the set of n with $n/\varphi(n)\le u$, exists for every u. In addition, $D_\varphi(u)$ is continuous and strictly increasing on $[1,\infty)$, with asymptotic limit 1.

It turns out that the situation for the higher iterates of φ is much simpler. We show below that the normal order of $\varphi_k(n)/\varphi_{k+1}(n)$ is

$$ke^\gamma\log\log\log n,$$

where γ is Euler's constant, for each fixed $k\ge 1$. In fact, this result continues to hold true if k is allowed to tend to infinity at a modest rate. (For fixed k, this result was stated without proof in [7].)

As a corollary, we have that the set

$$\{n : n/\varphi_{k+1}(n)\le uk!e^{k\gamma}(\log\log\log n)^k\}$$

has asymptotic density $D_\varphi(u)$ for every integer $k\ge 0$ and for every real number u.

It is well known that the maximal order of $n/\varphi(n)$ is $e^\gamma\log\log n$ but that very few integers n have $n/\varphi(n)$ this order of magnitude. We show below the existence of a positive constant c such that

$$\max_k \varphi_k(n)/\varphi_{k+1}(n) > c\log\log n$$

holds for a set of n of asymptotic density 1 . In fact a stronger result is true. We show the existence of a positive constant c' such that the set of n for which there is a k with $\varphi_k(n)$ divisible by every prime up to $(\log n)^{c'}$ has asymptotic density 1 .

The following two conjectures are perhaps tractable, but so far have resisted our efforts. We define

$$\Phi(n) = n \prod_{k \geq 1} \varphi_k(n).$$

Conjecture 1. *For each prime p, let $N(x,p)$ denote the number of $n \leq x$ with $p \mid \Phi(n)$. Then for every $\epsilon > 0$, $N(x,p) = o(x)$ uniformly in the region $p > (\log x)^{1+\epsilon}$ and $N(x,p) \sim x$ uniformly in the region $p < (\log x)^{1-\epsilon}$.*

Conjecture 2. *For each $\epsilon > 0$, the upper asymptotic density of the set of n with the property that the largest prime factor of $\varphi_k(n)$ exceeds n^{ϵ} tends to 0 as $k \to \infty$.*

Concerning Conjecture 1, we show below that for every n , the number of distinct prime factors of $\Phi(n)$ is at most $\lceil (\log n)/\log 2 \rceil$. Thus for each $\epsilon > 0$ and all $x \geq x_0(\epsilon)$, there is **no** $n \leq x$ with $\Phi(n)$ divisible by every prime $p \leq (\log x)^{1+\epsilon}$. However, we not only cannot prove the first assertion in Conjecture 1 for every $\epsilon > 0$, we cannot prove it for any specific choice of ϵ , even for very large choices. From our theorem mentioned above on $\varphi_k(n)$ being divisible by every prime up to $(\log n)^{c'}$, it follows that if $0 < c < c'$, then $N(x,p) \sim x$ uniformly for $p < (\log x)^{c}$. The second assertion in Conjecture 1 has both stronger and weaker versions that may be worth stating. The stronger version is that for each $\epsilon > 0$, there is a set $S_\epsilon(x)$ of integers $n \leq x$ of cardinality $o_\epsilon(x)$ such that if $n \leq x$, $n \notin S_\epsilon(x)$, then $\Phi(n)$ is divisible by every prime $p \leq (\log x)^{1-\epsilon}$. From the above mentioned theorem, this is true for all $\epsilon < 1 - c'$. The weaker version is that

$$\sum_{\substack{p < \log n \\ p \nmid \Phi(n)}} 1/p = o(1)$$

on a set of n of asymptotic density 1. Perhaps this is tractable. Note that from the above comments, we have

$$\sum_{\substack{p > \log n \\ p \mid \Phi(n)}} 1/p \to 0 \quad \text{as } n \to \infty.$$

By using sieve methods, we can prove Conjecture 2 for $\epsilon > 2/3$. We do not give the proof here.

The sum of the divisors function $\sigma(n)$ resembles in many ways Euler's function $\varphi(n)$. Yet it seems very difficult to prove anything non-trivial about the sequence of k-fold iterates $\sigma_k(n)$. For example, consider the following statements:

(i) for every $n > 1$, $\sigma_{k+1}(n)/\sigma_k(n) \to 1$ as $k \to \infty$;

(ii) for every $n > 1$, $\sigma_{k+1}(n)/\sigma_k(n) \to \infty$ as $k \to \infty$;

(iii) for every $n > 1$, $\sigma_k(n)^{1/k} \to \infty$ as $k \to \infty$;

(iv) for every $n > 1$, there is some k with $n \mid \sigma_k(n)$;

(v) for every n, $m > 1$, there is some k with $m \mid \sigma_k(n)$;

(vi) for every n, $m > 1$, there are some k, ℓ, with $\sigma_k(m) = \sigma_\ell(n)$.

We can neither prove nor disprove any of these statements.

Let $s(n) = \sigma(n) - n$ and let $s_k(n)$ be the k-fold iterate of s at n. In [8], the first author stated the following: For each $\epsilon > 0$ and k , the set of n with

$$\left| \frac{s(n)}{n} - \frac{s_{j+1}(n)}{s_j(n)} \right| < \epsilon \quad \text{for } j = 1, 2, \cdots, k$$

has asymptotic density 1 . This result is "half proved" in [8]. Namely, it is shown that the set of n with

$$\frac{s_{j+1}(n)}{s_j(n)} > \frac{s(n)}{n} - \epsilon \quad \text{for } j = 1, 2, \cdots, k$$

has asymptotic density 1 . The other half of the statement is claimed, but no argument is given. The first author now wishes to retract this claim and state the following as an open problem.

Conjecture 3. *For each $\epsilon > 0$ and k, the set of n with*

$$\frac{s_{j+1}(n)}{s_j(n)} < \frac{s(n)}{n} + \epsilon \quad \text{for } j = 1, , \cdots, k$$

has asymptotic density 1 .

In section 5 we give a proof of Conjecture 3 in the case $k = 1$. We also show that the full Conjecture 3 would be implied by the following conjecture.

Conjecture 4. *If $\mathcal{A}$ is a set of natural numbers of positive upper density, then $s(\mathcal{A}) = \{s(n) : n \in \mathcal{A}\}$ also has positive upper density.*

Note that it is possible for $s(\mathcal{A})$ to have positive density when $\mathcal{A}$ has density 0 . For example, if $p \neq q$ are primes, then $s(pq) = p + q + 1$.

While the set of integers of the form pq has asymptotic density 0 , the set of integers of the form $p+q+1$ with p, q distinct primes has asymptotic density $1/2$. This follows from work on the "exceptional set" in Goldbach's conjecture. In fact, a more complicated version of this idea gives that the set of $s_k(pq)$ has lower asymptotic density at least $1/2$ for any fixed k. We show this in section 5.

Suppose for every K there is a number C_K such that for any m there are at most C_K numbers $n \leq Km$ with $s(n) = m$. We are not sure whether we believe this hypothesis and in fact it may be possible to disprove it. We note though that it implies Conjecture 4.

In some sense, the paper [8] was motivated by a problem of H. W. Lenstra, Jr. [12] to show that for each k , there is an n with

$$n < s(n) < s_2(n) < \cdots < s_k(n). \tag{1.2}$$

Let α be the asymptotic density of the set of n with $n < s(n)$. Then $\alpha > 0$ and the correct half of [8] shows that for each k, (1.2) holds for a set of n of asymptotic density. That is, if the first inequality in (1.2) holds, then almost certainly all of the inequalities in (1.2) hold. Thus [8] provides a very strong solution to Lenstra's problem. The third author wishes to acknowledge a conversation with Lenstra in which the difficulty in the proof of the other half of [8] was discovered.

In [9], the first and third authors prove a theorem on the normal number of prime factors of $\varphi(n)$. Abdelhakim Smati has pointed out to us an error in the proof of Lemma 2.2 in this paper and another minor error. We correct these errors below in the last section.

Throughout the paper the letters p, q, r will always denote primes.

§2. The average and normal order of $F(n)$

Most of the results in this section are conditional on certain suitably strong versions of the Elliott-Halberstam conjecture. Before we state our results we define a few terms.

Definition. *We say a positive, continuous function $\epsilon(x)$ defined on $(1,\infty)$ is acceptable if*

(i) $\epsilon(x)\log x$ is eventually increasing and $\to \infty$ as $x \to \infty$;

(ii) for some $\delta > 0$, $\epsilon(x)(\log\log x)^{1+\delta}$ is eventually decreasing.

Some examples of acceptable functions are

$$\epsilon(x) = (\log\log 3x)^{-2},$$
$$\epsilon(x) = (\log x)^{-1/2},$$
$$\epsilon(x) = \exp\left((\log\log 3x)^{1/2}\right)/\log x.$$

Consider the two statements:

$$\sum_{p \le x^{1-\epsilon(x)}} \left| \pi(x;p,1) - \frac{\pi(x)}{p-1} \right| \ll \epsilon(x)\pi(x), \tag{A_ϵ}$$

$$\sum_{\substack{m \le x^{1-\epsilon(x)} \\ \Omega(m) \le 2}} \left| \pi(x;m,1) - \frac{\pi(x)}{\varphi(m)} \right| \ll \epsilon(x)\pi(x). \tag{B_ϵ}$$

Here the function $\Omega(m)$ counts the total number of prime factors of m with multiplicity, so that the statement B_ϵ implies the statement A_ϵ.

We now state the principal results of this section. Please note that if $\epsilon(x)$ is an acceptable function, then $\epsilon(x)\log\log x = o(1)$.

Theorem 2.1. *If A_ϵ holds for some acceptable function $\epsilon(x)$, then there is some positive constant α such that*

$$\frac{1}{x}\sum_{n \le x} F(n) = \alpha \log x + O(\epsilon(x)\log x \log\log x). \tag{2.1}$$

Theorem 2.2. *If B_ϵ holds for some acceptable function $\epsilon(x)$ and if α is the constant of Theorem 2.1, then*

$$\frac{1}{x}\sum_{n \le x} (F(n) - \alpha \log n)^2 \ll \epsilon(x)\log^2 x \log\log x.$$

In particular, $F(n)$ has normal order $\alpha \log n$.

Corollary 2.3. *If $\epsilon(x)$ is an acceptable function of the form $(\log x)^{-1+o(1)}$ and if B_ϵ holds, then for each $\delta > 0$, the set of n with*

$$|F(n) - \alpha \log n| < (\log n)^{1/2+\delta}$$

has asymptotic density 1.

The implied constants in Theorems 2.1, 2.2 depend, respectively, on the implied constants in A_ϵ, B_ϵ and on which specific function $\epsilon(x)$ is used. Thus if one had A_ϵ with an explicit constant for some explicit $\epsilon(x)$, say $\epsilon(x) = (\log\log 3x)^{-2}$, then the constant α would be effectively computable.

We begin the proof of Theorem 2.1 with an unconditional result.

Lemma 2.4. *For any function $\epsilon(x)$ with $x^{1/2} \le x^{1-\epsilon(x)} \le (1-\delta)x$ for x large and $\delta > 0$ some constant, we have*

$$\frac{1}{\pi(x)}\sum_{p\le x} F(p) - \sum_{p\le x}\frac{F(p)}{p} \ll \epsilon(x)\log x + \frac{\log^2 x}{x}\sum_{p\le x^{1-\epsilon(x)}}\left|\pi(x;p,1) - \frac{\pi(x)}{p-1}\right|.$$

Proof: From the definition of F we have

$$\begin{aligned}\sum_{p\le x} F(p) &= 1 + \sum_{3\le p\le x} F(p) = 1 + \sum_{3\le p\le x} F(p-1)\\ &= 1 + \sum_{3\le p\le x}\sum_{q^a \mid p-1} F(q) = 1 + \sum_{q^a\le x} F(q)\pi(x;q^a,1).\end{aligned}$$

Thus

$$\sum_{p\le x} F(p) - \pi(x)\sum_{p\le x}\frac{F(p)}{p} = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3, \tag{2.3}$$

where

$$\begin{aligned}\sum\nolimits_1 &= 1 + \sum_{\substack{p^a\le x\\ a\ge 2}} F(p)\pi(x;p^a,1),\\ \sum\nolimits_2 &= \sum_{p\le x^{1-\epsilon(x)}} F(p)\left(\pi(x;p,1) - \frac{\pi(x)}{p}\right),\\ \sum\nolimits_3 &= \sum_{x^{1-\epsilon(x)}<p\le x} F(p)\left(\pi(x;p,1) - \frac{\pi(x)}{p}\right).\end{aligned}$$

We have (using $F(p) \ll \log p$)

$$\begin{aligned}\sum\nolimits_1 &\ll 1 + \sum_{\substack{p^a\le x^{1/3}\\ a\ge 2}} (\log p)\left(\pi(x;p^a,1) - \frac{\pi(x)}{\varphi(p^a)}\right)\\ &\quad + \sum_{\substack{p^a\le x^{1/3}\\ a\ge 2}} (\log p)\frac{\pi(x)}{\varphi(p^a)} + \sum_{\substack{x^{1/3}<p^a\le x\\ a\ge 2}} (\log p)\pi(x;p^a,1)\\ &\ll \frac{x}{\log^2 x} + \frac{x}{\log x} + \sum_{\substack{p^a>x^{1/3}\\ a\ge 2}}\frac{x\log p}{p^a} \ll \frac{x}{\log x},\end{aligned} \tag{2.4}$$

where we used the Bombieri-Vinogradov theorem for the first sum over $p^a \le x^{1/3}$. In addition, we have

$$\sum\nolimits_2 \ll \log x \sum_{p\le x^{1-\epsilon(x)}}\left|\pi(x;p,1) - \frac{\pi(x)}{p-1}\right| + \frac{x}{\log x}. \tag{2.5}$$

For $\sum_3$, we have

$$\sum\nolimits_3 \ll \log x \sum_{x^{1-\epsilon(x)}<p\le x} \pi(x;p,1) + \pi(x) \sum_{x^{1-\epsilon(x)}<p\le x} \frac{\log p}{p}$$
$$\ll \log x \sum_{x^{1-\epsilon(x)}<p\le x} \pi(x;p,1) + \pi(x)\epsilon(x)\log x. \tag{2.6}$$

We estimate the sum on the right of (2.6) using Brun's method as follows:

$$\sum_{x^{1-\epsilon(x)}<p\le x} \pi(x;p,1) = \sum_{x^{1-\epsilon(x)}<p\le x} \sum_{\substack{q\le x \\ q\equiv 1 \pmod p}} 1$$
$$\le \sum_{m<x^{\epsilon(x)}} \sum_{\substack{p\le x/m \\ pm+1 \text{ is prime}}} 1 \ll \sum_{m<x^{\epsilon(x)}} \frac{m}{\varphi(m)} \frac{x/m}{\log^2(x/m)}$$
$$\ll \frac{x}{\log^2 x} \sum_{m<x^{\epsilon(x)}} \frac{1}{\varphi(m)} \ll \frac{\epsilon(x)x}{\log x}. \tag{2.7}$$

Putting this estimate in (2.6) and assembling (2.3)-(2.6), we obtain the lemma.

Corollary 2.5. *If $\epsilon(x)$ is some function that satisfies the hypothesis of Lemma 2.4 and if A_ϵ holds, then*

$$\frac{1}{\pi(x)} \sum_{p\le x} F(p) - \sum_{p\le x} \frac{F(p)}{p} \ll \epsilon(x)\log x.$$

Proof of Theorem 2.1: We unconditionally have

$$\frac{1}{x} \sum_{n\le x} F(n) = \sum_{p\le x} \frac{F(p)}{p} + O(1). \tag{2.8}$$

Indeed, using $F(p) \ll \log p$, we have

$$\frac{1}{x} \sum_{n\le x} F(n) = \frac{1}{x} \sum_{n\le x} \sum_{p^a|n} F(p) = \frac{1}{x} \sum_{p^a\le x} F(p) \left[\frac{x}{p^a}\right]$$
$$= \sum_{p^a\le x} \frac{F(p)}{p^a} + O(1) = \sum_{p\le x} \frac{F(p)}{p} + O(1).$$

Corollary 2.5 gives a (conditional) connection between $\sum_{p\le x} F(p)/p$ and $\sum_{p\le x} F(p)$. There is another (unconditional) connection which comes from partial summation. Let

$$R(x) := \frac{1}{x} \sum_{p\le x} F(p).$$

Then

$$\sum_{p\le x}\frac{F(p)}{p} = R(x) + \int_2^x \frac{1}{t}R(t)\,dt. \tag{2.9}$$

Assume now that $\epsilon(x)$ is an acceptable function and that A_ϵ holds. Then from Corollary 2.5 and (2.9) we have

$$R(x)\log x = R(x) + \int_2^x \frac{1}{t}R(t)\,dt + O(\epsilon(x)\log x),$$

so that

$$R(x) = \frac{1}{\log x}\int_2^x \frac{1}{t}R(t)\,dt + O(\epsilon(x)), \tag{2.10}$$

using $R(x) \ll 1$.

Let

$$V(x) := \frac{1}{\log x}\int_2^x \frac{1}{t}R(t)\,dt.$$

Since $R(x)$ is continuous but for a discrete set of jump discontinuities it follows that $V(x)$ is continuous, differentiable where $R(x)$ is continuous and satisfies

$$V(x) = \int_2^x V'(t)\,dt. \tag{2.11}$$

But at points where $R(x)$ is continuous , we have

$$\begin{aligned} V'(x) &= \frac{R(x)}{x\log x} - \frac{1}{x\log^2 x}\int_2^x \frac{1}{t}R(t)\,dt \\ &= \frac{1}{x\log x}(R(x) - V(x)) \ll \frac{\epsilon(x)}{x\log x}, \end{aligned} \tag{2.12}$$

by (2.10).

Note that by the definition of acceptable function, we have

$$\int_2^x \frac{\epsilon(t)}{t\log t}\,dt < \infty.$$

Thus by (2.11) and (2.12), we have that

$$\alpha := \int_2^\infty V'(t)\,dt = \lim_{x\to\infty} V(x)$$

exists and is positive.

We define now

$$\bar{\epsilon}(x) := \int_x^\infty \frac{\epsilon(t)}{t\log t}\,dt \tag{2.13}$$

and note that from the definition of acceptable function we have

$$\begin{aligned}\epsilon(x) &= \int_x^\infty \frac{\epsilon(x)\log x}{t\log^2 t}\,dt \le \int_x^\infty \frac{\epsilon(t)\log t}{t\log^2 t}\,dt = \bar{\epsilon}(x)\\ &= \int_x^\infty \frac{\epsilon(t)(\log\log t)^{1+\delta}}{t\log t(\log\log t)^{1+\delta}}\,dt \le \int_x^\infty \frac{\epsilon(x)(\log\log x)^{1+\delta}}{t\log t(\log\log t)^{1+\delta}}\,dt\\ &= \frac{1}{\delta}\epsilon(x)\log\log x \qquad (2.14)\end{aligned}$$

for some $\delta > 0$ and all sufficiently large x. From (2.10)-(2.14), we have

$$R(x) = \alpha + O(\bar{\epsilon}(x)).$$

Putting this estimate and (2.14) into (2.9) gives

$$\sum_{p\le x}\frac{F(p)}{p} = \alpha\log x + O(\bar{\epsilon}(x)\log x) = \alpha\log x + O(\epsilon(x)\log x\log\log x). \quad (2.15)$$

Thus (2.1) follows from (2.8) and (2.15).

We now turn our attention to the proof of Theorem 2.2. We shall prove this result by Turán's method (see Elliott [3], vol. II, p.112). In particular, let

$$E(x) := \frac{1}{x}\sum_{n\le x} F(n).$$

Thus (2.2) follows directly from (2.1) and the assertion

$$\frac{1}{x}\sum_{n\le x}(F(n)-E(x))^2 \ll \epsilon(x)\log^2 x\log\log x. \qquad (2.16)$$

But

$$\begin{aligned}\frac{1}{x}\sum_{n\le x}(F(n)-E(x))^2 &= \frac{1}{x}\sum_{n\le x}F(n)^2 - \frac{2E(x)}{x}\sum_{n\le x}F(n) + \frac{[x]}{x}E(x)^2\\ &= \frac{1}{x}\sum_{n\le x}F(n)^2 - E(x)^2 + O(\frac{\log^2 x}{x}).\end{aligned}$$

Thus (2.16) follows from (2.1) and the assertion

$$\frac{1}{x}\sum_{n\le x}F(n)^2 = \alpha^2\log^2 x + O(\epsilon(x)\log^2 x\log\log x). \qquad (2.17)$$

We have thus reduced Theorem 2.2 to proving (2.17) (under the hypothesis of Theorem 2.2) .

We now turn to the sum in (2.17) . We have

$$\sum_{n\le x} F(n)^2 = \sum_{n\le x}\left(\sum_{p^a \mid n} F(p)\right)^2 = \sum_{p^a,q^b\le x} F(p)F(q) \sum_{\substack{n\le x \\ p^a\mid n,\ q^b\mid n}} 1$$

$$= \sum_{p\le x} F(p)^2\left[\frac{x}{p}\right] + 2\sum_{\substack{pq\le x \\ p<q}} F(p)F(q)\left[\frac{x}{pq}\right] + \sum\nolimits_1 + \sum\nolimits_2, \quad (2.18)$$

where

$$\sum\nolimits_1 = \sum_{\substack{p^a,p^b\le x \\ a+b\ge 3}} F(p)^2\left[\frac{x}{p^{\max\{a,b\}}}\right]$$

$$\ll x\sum_{\substack{p^a \\ a\ge 2}} \frac{F(p)^2}{p^a} \sum_{b\le(\log x)/\log p} 1$$

$$\ll x(\log x)\sum_{\substack{p^a \\ a\ge 2}} \frac{\log p}{p^a} \ll x\log x$$

and

$$\sum\nolimits_2 = 2\sum_{\substack{p^aq^b\le x \\ p<q \\ a+b\ge 3}} F(p)F(q)\left[\frac{x}{p^aq^b}\right]$$

$$\ll x\sum_{p^a\le x}\frac{F(p)}{p^a}\sum_{\substack{q^b \\ b\ge 2}}\frac{F(q)}{q^b} \ll x\sum_{p^a\le x}\frac{F(p)}{p^a} \ll x\log x.$$

Further note that removing the brackets on the right of (2.18) introduces an error of at most $O(x\log x)$. Thus

$$\frac{1}{x}\sum_{n\le x} F(n)^2 = \sum_{p\le x}\frac{F(p)^2}{p} + 2\sum_{\substack{pq\le x \\ p<q}}\frac{F(p)F(q)}{pq} + O(\log x). \quad (2.19)$$

We thus will have (2.17) and Theorem 2.2 from (2.14), (2.19) and the following result.

Proposition 2.6. *Under the hypothesis of Theorem 2.2 we have*

$$\sum_{p\le x}\frac{F(p)^2}{p}=\frac{1}{2}\alpha^2\log^2 x+O(\epsilon(x)\log^2 x\log\log x), \tag{2.20}$$

$$2\sum_{\substack{pq\le x\\ p<q}}\frac{F(p)F(q)}{pq}=\frac{1}{2}\alpha^2\log^2 x+O(\bar{\epsilon}(x)\log^2 x) \tag{2.21}$$

where $\bar{\epsilon}(x)$ is defined in (2.13).

Proof: We begin with the proof of (2.21) which is easier and actually used in the proof of (2.20). First note that from $\epsilon(x)\le\bar{\epsilon}(x)$ for large x (see (2.14)), we have

$$\frac{d}{dx}(\bar{\epsilon}(x)\log x)=\frac{\bar{\epsilon}(x)-\epsilon(x)}{x}\ge 0$$

for large x , so that $\bar{\epsilon}(x)\log x$ is eventually increasing. Thus from (2.15) we have

$$\begin{aligned}\sum_{\substack{pq\le x\\ p<q}}\frac{F(p)F(q)}{pq}&=\sum_{p\le\sqrt{x}}\frac{F(q)}{q}\sum_{p<q\le x/p}\frac{F(p)}{p}\\ &=\sum_{p\le\sqrt{x}}\frac{F(p)}{p}\left(\alpha\log\frac{x}{p}-\alpha\log p+O\left(\bar{\epsilon}\left(\frac{x}{p}\right)\log\left(\frac{x}{p}\right)\right)\right.\\ &\qquad\left.+O\left(\bar{\epsilon}(p)\log p\right)\right)\\ &=\alpha\log x\sum_{p\le\sqrt{x}}\frac{F(p)}{p}-2\alpha\sum_{p\le\sqrt{x}}\frac{F(p)\log p}{p}\\ &\qquad+O\left(\bar{\epsilon}(x)\log x\sum_{p\le\sqrt{x}}\frac{F(p)}{p}\right)\\ &=2\alpha\int_2^{\sqrt{x}}\frac{1}{t}\sum_{p\le t}\frac{F(p)}{p}\,dt+O\left(\bar{\epsilon}(x)\log^2 x\right).\end{aligned} \tag{2.22}$$

By (2.15), the integral is

$$\begin{aligned}&2\alpha\int_2^{\sqrt{x}}\frac{\alpha\log t}{t}\,dt+O\left(\int_2^{\sqrt{x}}\frac{\bar{\epsilon}(t)\log t}{t}\,dt\right)\\ &\qquad=\alpha^2\log^2\sqrt{x}+O\left(\bar{\epsilon}(x)\log x\int_2^{\sqrt{x}}\frac{dt}{t}\right)\\ &\qquad=\frac{1}{4}\alpha^2\log^2 x+O\left(\bar{\epsilon}(x)\log^2 x\right),\end{aligned}$$

so (2.22) gives (2.21).

We now turn to the proof of (2.20). By partial summation, we have

$$\sum_{p\le x}\frac{F(p)^2}{p}=\frac{1}{x}\sum_{p\le x}F(p)^2+\int_2^x\frac{1}{t^2}\sum_{p\le t}F(p)^2\,dt. \tag{2.23}$$

We now expand $\sum_{p\le x}F(p)^2$. We have

$$\begin{aligned}\sum_{p\le x}F(p)^2&=1+\sum_{3\le p\le x}F(p-1)^2=1+\sum_{3\le p\le x}\left(\sum_{q^a\mid p-1}F(q)\right)^2\\&=1+\sum_{p^a,q^b\le x}F(p)F(q)\pi(x;[p^a,q^b],1)\\&=1+\sum_{p\le x}F(p)^2\pi(x;p,1)+2\sum_{\substack{pq\le x\\p<q}}F(p)F(q)\pi(x;pq,1)\\&\quad+\sum_{\substack{p^a,q^b\le x\\a+b\ge 3}}F(p)F(q)\pi(x;[p^a,q^b],1).\end{aligned} \tag{2.24}$$

We have

$$\begin{aligned}&\sum_{p\le x}F(p)^2\pi(x;p,1)\\&=\sum_{p\le x}F(p)^2\frac{\pi(x)}{p-1}+\sum_{p\le x}F(p)^2\left(\pi(x;p,1)-\frac{\pi(x)}{p-1}\right)\\&=\pi(x)\sum_{p\le x}\frac{F(p)^2}{p-1}+O\left(\log^2 x\sum_{p\le x^{1-\epsilon(x)}}\left|\pi(x;p,1)-\frac{\pi(x)}{p-1}\right|\right)\\&\quad+O\left(\log^2 x\sum_{x^{1-\epsilon(x)}<p\le x}\pi(x;p,1)\right)+O\left(\pi(x)\sum_{x^{1-\epsilon(x)}<p\le x}\frac{\log^2 p}{p}\right)\\&=\frac{x}{\log x}\sum_{p\le x}\frac{F(p)^2}{p}+O\left(\epsilon(x)x\log x\right),\end{aligned} \tag{2.25}$$

using hypothesis A_ϵ and (2.7). Next, we have using hypothesis B_ϵ,

$$\sum_{\substack{pq\le x\\p<q}}F(p)F(q)\pi(x;pq,1)$$

$$= \pi(x) \sum_{\substack{pq \le x \\ p<q}} \frac{F(p)F(q)}{\varphi(pq)} + O\left(\log^2 x \sum_{pq \le x^{1-\epsilon(x)}} \left| \pi(x;, pq, 1) - \frac{\pi(x)}{\varphi(pq)} \right| \right)$$

$$+ O\left(\log x \sum_{\substack{x^{1-\epsilon(x)} < pq \le x \\ p<q}} (\log p)\pi(x; pq, 1) \right)$$

$$+ O\left(\pi(x) \sum_{x^{1-\epsilon(x)} < pq \le x} \frac{\log p \log q}{pq} \right)$$

$$= \frac{x}{\log x} \sum_{\substack{pq \le x \\ p<q}} \frac{F(p)F(q)}{pq} + O\left(\frac{x}{\log x} \sum_{\substack{pq \le x \\ p<q}} \frac{\log p \log q}{p^2 q} \right) + O\left(\epsilon(x) x \log x\right)$$

$$+ O\left(\log x \sum_{m < x^{\epsilon(x)}} \sum_{p \le \sqrt{\frac{x}{m}}} \log p \sum_{\substack{q \le \frac{x}{pm} \\ qpm+1 \text{ prime}}} 1 \right)$$

$$+ O\left(\frac{x}{\log x} \sum_{p \le \sqrt{x}} \frac{\log p}{p} \sum_{\frac{x^{1-\epsilon(x)}}{p} < q \le \frac{x}{p}} \frac{\log q}{q} \right)$$

$$= \frac{x}{\log x} \sum_{\substack{pq \le x \\ p<q}} \frac{F(p)F(q)}{pq} + O\left(\epsilon(x) x \log x\right) + O\left(\frac{x}{\log x} \sum_{m < x^{\epsilon(x)}} \sum_{p \le \sqrt{\frac{x}{m}}} \frac{\log p}{p\varphi(m)} \right)$$

$$= \frac{x}{\log x} \sum_{\substack{pq \le x \\ p<q}} \frac{F(p)F(q)}{pq} + O\left(\epsilon(x) x \log x\right). \tag{2.26}$$

For the last term in (2.24) we have the estimate

$$\ll \sum_{\substack{p^a \le x \\ a \ge 2}} \log p \sum_{\substack{r \le x \\ r \equiv 1 \pmod{p^a}}} \sum_{q^b \mid r-1} \log q$$

$$= \sum_{\substack{p^a \le \sqrt{x} \\ a \ge 2}} \log p \sum_{\substack{r \le x \\ r \equiv 1 \pmod{p^a}}} \log r + \sum_{\substack{\sqrt{x} < p^a \le x \\ a \ge 2}} \log p \sum_{\substack{r \le x \\ r \equiv 1 \pmod{p^a}}} \log r$$

$$= x \sum_{\substack{p^a \le \sqrt{x} \\ a \ge 2}} \frac{\log p}{p^a} + x \log x \sum_{\substack{\sqrt{x} < p^a \le x \\ a \ge 2}} \frac{\log p}{p^a} \ll x,$$

using the Brun-Titchmarsh inequality. Putting this estimate, (2.25) and (2.26) into (2.24) we get

$$\frac{\log x}{x}\sum_{p\le x}F(p)^2=\sum_{p\le x}\frac{F(p)^2}{p}+2\sum_{\substack{pq\le x\\ p<q}}\frac{F(p)F(q)}{pq}+O\left(\epsilon(x)\log^2 x\right). \quad (2.27)$$

Now using this estimate with (2.14), (2.21) and (2.23), we get

$$\frac{\log x}{x}\sum_{p\le x}F(p)^2=\frac{1}{2}\alpha^2\log^2 x+\int_2^x\frac{1}{t^2}\sum_{p\le t}F(p)^2\,dt+O(\bar{\epsilon}(x)\log^2 x),$$

so that if

$$R_2(x):=\frac{1}{x}\sum_{p\le x}F(p)^2,$$

then we have

$$R_2(x)=\frac{1}{2}\alpha^2\log x+\frac{1}{\log x}\int_2^x\frac{1}{t}R_2(t)\,dt+O(\bar{\epsilon}(x)\log x). \quad (2.28)$$

Let

$$V_2(x):=\frac{1}{\log x}\int_2^x\frac{1}{t}R_2(t)\,dt.$$

As in the proof of Theorem 2.1, we have

$$V_2(x)=\int_2^x V_2'(t)\,dt \quad (2.29)$$

and

$$V_2'(x)=\frac{1}{x\log x}(R_2(x)-V_2(x))=\frac{\alpha^2}{2x}+O\left(\frac{\bar{\epsilon}(x)}{x}\right).$$

Thus from (2.29), we have

$$V_2(x)=\frac{1}{2}\alpha^2\log x+O\left(1+\int_2^x\frac{\bar{\epsilon}(t)}{t}\,dt\right).$$

But for large x

$$\begin{aligned}\int_2^x\frac{\bar{\epsilon}(t)}{t}\,dt&=\bar{\epsilon}(x)\log x-\int_2^x\bar{\epsilon}'(t)\log t\,dt\\&=\bar{\epsilon}(x)\log x+\int_2^x\frac{\epsilon(t)\log t}{t\log t}\,dt\\&\le\bar{\epsilon}(x)\log x+\epsilon(x)\log x\int_2^x\frac{dt}{t\log t}\\&=\bar{\epsilon}(x)\log x+\epsilon(x)\log x(\log\log x-\log\log 2)\\&\ll\epsilon(x)\log x\log\log x\end{aligned}$$

by (2.14), so that

$$V_2(x) = \frac{1}{2}\alpha^2 \log x + O(\epsilon(x) \log x \log\log x).$$

Thus from (2.28), we get

$$\frac{1}{x}\sum_{p\le x} F(p)^2 = \alpha^2 \log x + O(\epsilon(x) \log x \log\log x).$$

Finally, using this and (2.21) in (2.27) gives (2.20).

REMARKS: With a little more care, the right side of (2.2) can be replaced with

$$\bar{\epsilon}(x) \log^2 x + \log x \int_2^x \frac{\epsilon(t)}{t} dt \ .$$

For some choices of acceptable functions $\epsilon(x)$, this expression is $O(\epsilon(x) \log^2 x)$, which is smaller than the right side of (2.2) by a $\log\log x$ factor. For example, we would have this for $\epsilon(x) = (\log x)^{-\delta}$ for some fixed $\delta, 0 < \delta < 1$.

For each prime q , define a completely additive function $F_q(n)$ by inductively defining its values on the primes as follows:

$$F_q(p) = \begin{cases} 0, & \text{if } p < q \\ 1, & \text{if } p = q \\ F_q(p-1), & \text{if } p > q. \end{cases}$$

Thus $F_2(n) = F(n)$. The functions $F_q(n)$ have the following connection with the iterated phi-function:

$$F_q(n) = \#\{j \ge 0 : \ q \,|\, \varphi_j(n)\},$$

where we interpret $\varphi_0(n) = n$. We have already seen this for $q = 2$ in the Introduction.

Theorems 2.1 and 2.2 hold for the functions F_q for each q with corresponding constants α_q (with $\alpha_2 = \alpha$) , except that we are not sure that $\alpha_q > 0$ for $q > 2$. This, in fact, can be proved assuming hypothesis A_ϵ holds for $\epsilon(x) = (\log x)^{c-1}$ for some c with $0 < c < c_{10}$ where c_{10} is the constant of Theorem 4.5 below. Indeed, if $\varphi_j(n)$ is divisible by every prime up to $(\log n)^{c_{10}}$ and if n is large, then $\varphi_{j+1}(n)$ is divisible by q^k where

$$k > \frac{(\log n)^{c_{10}}}{q c_{10} \log\log n}.$$

Thus Theorem 4.5 implies $F_q(n) \gg (\log n)^{c_{10}}/\log\log n$ on a set of asymptotic density 1. However, this is incompatible with (2.1) if $\alpha_q = 0$, $\epsilon(x) = (\log x)^{c-1}$.

Let $v_p(n)$ denote the exponent on p in the prime factorization of n. Note that for any natural number m we have

$$v_p(\varphi(m)) - v_p(m) = \begin{cases} -1 + \sum_{q|m} v_p(q-1), & \text{if } v_p(m) > 0 \\ \sum_{q|m} v_p(q-1), & \text{if } v_p(m) = 0. \end{cases}$$

Let $k = k(n)$. Then

$$\begin{aligned} 0 = v_p(\varphi_k(n)) &= v_p(n) + \sum_{i=1}^{k} (v_p(\varphi_i(n)) - v_p(\varphi_{i-1}(n))) \\ &= v_p(n) - \sum_{\substack{i\ge 0 \\ v_p(\varphi_i(n))>0}} 1 + \sum_{i\ge 0} \sum_{q \mid \varphi_i(n)} v_p(q-1) \\ &= v_p(n) - F_p(n) + \sum_q v_p(q-1) F_q(n); \end{aligned}$$

that is, for every prime p and every natural number n, we have

$$F_p(n) = v_p(n) + \sum_q v_p(q-1) F_q(n), \tag{2.30}$$

where the sum is over all primes q.

We can generate another pretty identity involving the functions F_p via the elementary relation

$$\log m - \log \varphi(m) = \sum_{p \mid m} \log \frac{p}{p-1}.$$

We have

$$\begin{aligned} \log n = \sum_{i\ge 0} (\log \varphi_i(n) - \log \varphi_{i+1}(n)) &= \sum_{i\ge 0} \sum_{p \mid \varphi_i(n)} \log \frac{p}{p-1} \\ &= \sum_p F_p(n) \log \frac{p}{p-1}. \end{aligned} \tag{2.31}$$

Using (2.30) with $p = 2$ to eliminate $F_2(n)$ in (2.31), we have

$$\log n = v_2(n)\log 2 + F_3(n)\log 3 + F_5(n)\log 5 + F_7(n)\log \frac{7}{3} + \cdots \tag{2.32}$$

where the general term on the right is $F_p(n)\log\frac{p}{(p-1)_2}$ for $p \geq 3$ and where $(p-1)_2$ is the largest odd divisor of $p-1$. We can now use (2.30) to eliminate $F_3(n)$ in (2.32) and continuing, if we eliminate all $F_p(n)$ for $p \leq q$, we obtain the identity (valid for all n and q):

$$\log n = \sum_{p\leq q} v_p(n)\log p + \sum_{p>q} F_p(n)\log\frac{p}{(p-1)_q}, \tag{2.33}$$

where $(p-1)_q$ denotes the largest divisor of $p-1$ not divisible by any prime up to and including q.

Since for every $p > 2$ we have $(p-1)_q = 1$ for some $q < p$, a corollary of (2.33) is the theorem

$$F_p(n) \leq \log n/\log p \tag{2.34}$$

for all n and all $p > 2$. From (2.31), this inequality holds for $p = 2$ as well. Note that if $n = p^k$, then $F_p(n) = k = \log n/\log p$, so (2.34) is best possible.

Suppose now that A_ϵ holds for some acceptable function $\epsilon(x)$. Then each of the numbers α_p exists and an immediate corollary of (2.34) is that

$$\alpha_p \leq 1/\log p \tag{2.35}$$

for each p. In particular, $\lim_{p\to\infty}\alpha_p = 0$. Further, (2.30) implies that

$$\alpha_p \geq \sum_{q\leq p_0} v_p(q-1)\alpha_q$$

for any prime p_0. Letting $p_0 \to \infty$, we obtain

$$\alpha_p \geq \sum_{q} v_p(q-1)\alpha_q \tag{2.36}$$

for every p. The case $p = 2$ shows that $\sum\alpha_p$ converges. Similarly, using (2.31) and (2.33) we get

$$\begin{aligned} 1 &\geq \sum_p \alpha_p \log\frac{p}{p-1}, \\ 1 &\geq \sum_{p>q} \alpha_p \log\frac{p}{(p-1)_q} \end{aligned} \tag{2.37}$$

for every q. Thus if infinitely many p have $\alpha_p > 0$, we have strict inequality in (2.35) for every p.

Assume now that $0 < c < c_{10}$ and that A_ϵ holds for $\epsilon(x) = (\log x)^{c-1}$. We've seen that this then implies each $\alpha_p > 0$. Thus (2.36) and Dirichlet's

theorem on primes in an arithmetic progression imply we have $\alpha_p > \alpha_q$ for all primes p, q with $q \equiv 1 \pmod{p}$. We conjecture that we have $\alpha_p > \alpha_q$ whenever $q > p$.

We can prove that we have equality in the first statement in (2.37) as follows. By (2.34), we have

$$\frac{1}{\log n} \sum_{p > p_0} F_p(n) \log \frac{p}{p-1} \leq \sum_{p > p_0} \frac{\log(p/(p-1))}{\log p} \to 0 \text{ as } p_0 \to \infty. \quad (2.38)$$

But for any p_0 , we have by (2.31)

$$1 = \frac{1}{[x]} \sum_{n \leq x} \frac{1}{\log n} \sum_{p \leq p_0} F_p(n) \log \frac{p}{p-1} + \frac{1}{[x]} \sum_{n \leq x} \frac{1}{\log n} \sum_{p > p_0} F_p(n) \log \frac{p}{p-1}$$

$$= \sum_{p \leq p_0} \alpha_p \log \frac{p}{p-1} + o(1) + \frac{1}{[x]} \sum_{n \leq x} \frac{1}{\log n} \sum_{p > p_0} F_p(n) \log \frac{p}{p-1}$$

as $x \to \infty$. But from (2.38) we can make the last expression as small as we please uniformly for every x by taking p_0 large enough. Thus

$$1 = \sum_p \alpha_p \log \frac{p}{p-1}.$$

We conjecture we also have equality in (2.36) and in the second statement of (2.37).

§3. Results on the sum of the reciprocals of primes

From a theorem of Landau (for example, see Davenport [2], p. 94) there is a positive constant c_0 with the following property. Let $\mathcal{E}(c_0)$ denote the set of natural numbers n for which there is a real primitive character χ mod n for which $L(s, \chi)$ has a real root $\beta \geq 1 - c_0/\log n$. Then $1 \notin \mathcal{E}(c_0)$ and for any x there is at most one member n of $\mathcal{E}(c_0)$ between x and x^2 .

Lemma 3.1. *There are positive absolute constants $c_1 \leq 1, c_2 > 1$ such that if $n > 1$ is a natural number with n not divisible by any member of $\mathcal{E}(c_0)$, then*

$$\sum_{\substack{p \leq x \\ p \equiv 1 \pmod{n}}}{}' \frac{1}{p} \geq \frac{c_1}{\varphi(n)} (\log\log x - \log\log n)$$

for all $x \geq n^{c_2}$, where $\sum'$ signifies that the sum is over primes not in $\mathcal{E}(c_0)$.

Proof: This result follows from the proof of Linnik's theorem given in Section 6 of Bombieri [1]. In particular, from this proof, if c_2 is sufficiently large, then

$$\sum_{\substack{p \leq t \\ p \equiv 1 \pmod{n}}}{}' \log p > \frac{t}{2\varphi(n)}$$

for any $t \geq n^{c_2/2}$. Then if $x \geq n^{c_2}$,

$$\sum_{\substack{p \leq x \\ p \equiv 1 \pmod n}}{}' \frac{1}{p} \geq \int_n^x \frac{1}{t^2 \log t} \sum_{\substack{p \leq t \\ p \equiv 1 \pmod n}}{}' \log p \, dt$$

$$\geq \frac{1}{2\varphi(n)} \int_{n^{c_2/2}}^x \frac{dt}{t \log t}$$

$$= \frac{1}{2\varphi(n)} (\log\log x - \log\log(n^{c_2/2}))$$

$$\geq \frac{c_1}{\varphi(n)} (\log\log x - \log\log n)$$

for $c_1 \leq (\log 2)/(2 \log c_2)$.

Lemma 3.2. *Suppose $\mathcal{S}$ is a set of primes. For any x, let*

$$S_1 = \sum_{p \in \mathcal{S}} \sum_{\substack{q \leq x \\ q \equiv 1 \pmod p}}{}' \frac{1}{q}, \quad S_2 = \sum_{\substack{p, p' \in \mathcal{S} \\ p < p'}} \sum_{\substack{q \leq x \\ q \equiv 1 \pmod{pp'}}}{}' \frac{1}{q}.$$

If $q \leq x$ is prime, let a_q denote the number of prime factors of $q - 1$ that are in $\mathcal{S}$. If $S_1 > 0$, then

$$\sum_{\substack{q \leq x \\ a_q > 0}}{}' \frac{1}{q} \geq \frac{S_1^2}{2S_2 + S_1}.$$

Proof: This is just the Cauchy-Schwarz inequality. In fact,

$$S_1 = \sum_{q \leq x}{}' \frac{a_q}{q} = \sum_{q \leq x}{}' \frac{1}{\sqrt{q}} \cdot \frac{a_q}{\sqrt{q}} \leq \left(\sum_{\substack{q \leq x \\ a_q > 0}}{}' \frac{1}{q} \right)^{1/2} \left(\sum_{q \leq x}{}' \frac{a_q^2}{q} \right)^{1/2}$$

$$= \left(\sum_{\substack{q \leq x \\ a_q > 0}}{}' \frac{1}{q} \right)^{1/2} (2S_2 + S_1)^{1/2},$$

since $a_q^2 = 2\binom{a_q}{2} + a_q$.

Lemma 3.3. *Suppose $y \geq 3$ and $\mathcal{S}$ is a set of primes such that if $p \in \mathcal{S}$ then $p \leq y$ and $p \notin \mathcal{E}(c_0)$. There is an absolute positive constant c_3 such that if $x \geq y^{c_2}$, then*

$$\sum_{\substack{q \leq x \\ a_q > 0}}{}' \frac{1}{q} \geq \min\left\{ \frac{c_1^2}{16 c_3} \frac{(\log\log x - \log\log y)^2}{\log\log x}, \frac{c_1}{4} (\log\log x - \log\log y) \sum_{p \in \mathcal{S}} \frac{1}{p} \right\}$$

where a_q is defined in Lemma 3.2.

Proof: The lemma is clearly true if $2 \in \mathcal{S}$ or if $\mathcal{S} = \emptyset$, so assume $2 \notin \mathcal{S}$ and $\mathcal{S} \neq \emptyset$. Using the notation of Lemma 3.2, we have

$$S_1 \geq \frac{1}{2} c_1 (\log\log x - \log\log y) \sum_{p \in \mathcal{S}} \frac{1}{p}$$

from Lemma 3.1. Also, using partial summation and the Brun-Titchmarsh inequality we have for some absolute constant $c_3 \geq 1$,

$$\sum_{\substack{q \leq x \\ q \equiv 1 \pmod n}} \frac{1}{q} \leq \frac{c_3}{\varphi(n)} \log\log x \tag{3.1}$$

for any natural number n and any $x \geq 3$. Thus

$$S_2 \leq c_3 \log\log x \sum_{\substack{p,p' \in \mathcal{S} \\ p<p'}} \frac{1}{(p-1)(p'-1)} \leq c_3 \log\log x \left(\sum_{p \in \mathcal{S}} \frac{1}{p} \right)^2 =: S_2',$$

since $2 \notin \mathcal{S}$. Thus from Lemma 3.2, we have

$$\sum_{\substack{q \leq x \\ a_q > 0}}{}' \frac{1}{p} \geq \frac{S_1^2}{S_1 + 2S_2} \geq \frac{S_1^2}{S_1 + 2S_2'} \geq \min\left\{ \frac{S_1^2}{4S_2'}, \frac{1}{2} S_1 \right\}$$

and our conclusion follows.

If k, n are natural numbers, let

$$S_k'(x,n) = \sum_{\substack{p \leq x \\ n \mid \varphi_k(p)}}{}' \frac{1}{p}$$

where again the dash means that $p \notin \mathcal{E}(c_0)$.

Theorem 3.4. *There are absolute constants $0 < c_4, c_5, c_6 \leq 1$ such that for any A and $x \geq x_0(A)$ we have*

$$S_k'(x,n) \geq \min\left\{ c_4 \log\log x, \frac{1}{\varphi(n)} \left(\frac{c_5 \log\log x}{k} \right)^k \right\}$$

for all $n \le (\log x)^A$ *and* $k \le c_6 \log\log x$.

Proof: Fix an arbitrary number A and assume $n \le (\log x)^A$. If $y \ge \exp\left((\log x)^{1/3}\right)$, then by partial summation and the Siegel-Walfisz theorem, provided $x \ge x_0(A)$, we have

$$S_1'(y,n) \ge \int_{\exp\left((\log x)^{1/6}\right)}^{\exp\left((\log x)^{1/3}\right)} \frac{1}{t^2} \sum_{\substack{p \le t \\ p \equiv 1 \pmod n}} 1 \, dt \ge \frac{1}{7\varphi(n)} \log\log x. \tag{3.2}$$

By letting $y = x$ in (3.2) we have the theorem for $k = 1$.

Suppose now $k = 2$. Let $\mathcal{S}$ be the set of primes $p \le \exp((\log x)^{1/3})$ for which $p \equiv 1 \pmod n$ and $p \notin \mathcal{E}(c_0)$. Then in the notation of Lemma 3.2, we have

$$S_2'(x,n) \ge {\sum_{\substack{q \le x \\ a_q > 0}}}' \frac{1}{q}.$$

From Lemma 3.3 and (3.2) with $y = \exp((\log x)^{1/3})$ we have

$$S_2'(x,n) \ge \min\left\{ \frac{c_1^2}{36c_3} \log\log x, \frac{c_1}{42\varphi(n)} (\log\log x)^2 \right\},$$

which gives the theorem for $k = 2$.

Now let $k = 3$. Let

$$S_i' = S_i'\left(\exp((\log x)^{i/3}), n\right) \quad \text{for } i = 1, 2, 3.$$

Then from Lemma 3.3 we have

$$\begin{aligned} S_2' &\ge \min\left\{ \frac{c_1^2}{96c_3} \log\log x, \frac{c_1}{12} (\log\log x) S_1' \right\}, \\ S_3' &\ge \min\left\{ \frac{c_1^2}{144c_3} \log\log x, \frac{c_1}{12} (\log\log x) S_2' \right\} \\ &\ge \min\left\{ \frac{c_1^2}{144c_3} \log\log x, \frac{c_1^3}{1152c_3} (\log\log x)^2, \frac{c_1^2}{144} (\log\log x)^2 S_1' \right\}. \end{aligned}$$

Since $S_1' \ge \frac{1}{7\varphi(n)} \log\log x$ by (3.2), we have our theorem for $k = 3$.

Suppose now $k \ge 4$. Let

$$y_j = \exp\left((\log x)^{\frac{1}{3} + \frac{j}{3(k-3)}}\right) \quad \text{for } j = 0, 1, \cdots, k-3.$$

If c_6 is sufficiently small, then $k \le c_6 \log\log x$ implies that $y_j \ge y_{j-1}^{c_2}$ for $j = 1, \cdots, k-3$. Note that

$$\log\log y_j - \log\log y_{j-1} = \frac{1}{3(k-3)} \log\log x.$$

Thus from Lemma 3.3 we have for $j = 1, \cdots, k-3$

$$S'_{j+1}(y_j, n) \ge \min\left\{ \frac{c_1^2 \log\log x}{96c_3(k-3)^2},\ \frac{c_1 \log\log x}{12(k-3)} S'_j(y_{j-1}, n) \right\}. \tag{3.3}$$

The min is the first term if and only if

$$S'_j(y_{j-1}, n) \ge \frac{c_1}{8c_3(k-3)}. \tag{3.4}$$

We shall choose c_6 so small that we also have $c_6 \le c_1/12$. Then $k - 3 \le (c_1/12)\log\log x$, so that

$$\frac{c_1^2 \log\log x}{96c_3(k-3)^2} \ge \frac{c_1}{8c_3(k-3)}. \tag{3.5}$$

Thus if $0 < j < k-3$ and the min in (3.3) is the first term, then (3.5) implies that

$$S'_{j+1}(y_j, n) \ge \frac{c_1}{8c_3(k-3)}$$

and so (3.4) implies the same is true when j is replaced with $j+1$; i.e., the min in (3.3) is again the first term. Thus by iterating (3.3), we have

$$S'_{k-2}(y_{k-3}, n) \ge \min\left\{ \frac{c_1^2 \log\log x}{96c_3(k-3)^2},\ \left(\frac{c_1 \log\log x}{12(k-3)}\right)^{k-3} S'_1(y_0, n) \right\}. \tag{3.6}$$

Note that from (3.2) we have

$$S'_1(y_0, n) \ge \frac{1}{7\varphi(n)} \log\log x. \tag{3.7}$$

Note also that $y_{k-3} = \exp((\log x)^{2/3})$. Thus from Lemma 3.3, we have

$$S'_{k-1} := S'_{k-1}\left(\exp((\log x)^{5/6}),\ n\right) \ge$$
$$\min\left\{ \frac{c_1^2}{480c_3} \log\log x,\ \frac{c_1}{24}(\log\log x) S'_{k-2}(y_{k-3}, n) \right\},$$

$$S'_k(x,n) \geq \min\left\{\frac{c_1^2}{576c_3}\log\log x,\ \frac{c_1}{24}(\log\log x)S'_{k-1}\right\}.$$

Thus from (3.6) and (3.7)

$$S'_{k-1} \geq \min\left\{\frac{c_1^2}{480c_3}\log\log x, \frac{c_1^3(\log\log x)^2}{2304c_3(k-3)^2}, \left(\frac{c_1\log\log x}{12(k-3)}\right)^{k-2}\frac{\log\log x}{14\varphi(n)}\right\},$$

so that

$$S'_k(x,n) \geq \min\left\{\frac{c_1^2}{576c_3}\log\log x,\ \frac{c_1^3}{11520c_3}(\log\log x)^2,\ \frac{c_1^4(\log\log x)^3}{55296c_3(k-3)^2},\ \left(\frac{c_1\log\log x}{12(k-3)}\right)^{k-1}\frac{\log\log x}{28\varphi(n)}\right\}$$

Thus our theorem holds with

$$c_4 = \min\left\{\frac{c_1^2}{576c_3},\ \frac{c_1^4}{55296c_3c_6^2}\right\},\ c_5 = c_1/15$$

if $x \geq x_0$.

Theorem 3.5. *If c_3 is the constant in (3.1), we have*

$$\sum_{\substack{n\leq x\\ p\,|\,\varphi_k(n)}} 1 \leq \frac{x}{p}(2c_3\log\log x)^k$$

for every odd prime p, for every $k \geq 0$ and for all x with $\log\log x \geq 2/c_3$. (We define $\varphi_0(n) = n$.)

Proof: The theorem holds for $k = 0$ since $\varphi_0(n) = n$. Suppose $k \geq 0$ and the theorem holds for k. If $p \mid \varphi_{k+1}(n)$ then either $p^2 \mid \varphi_k(n)$ or there is some prime $q \mid \varphi_k(n)$ with $q \equiv 1 \pmod p$. Thus

$$\sum_{\substack{n\leq x\\ p\,|\,\varphi_{k+1}(n)}} 1 \leq \sum_{\substack{n\leq x\\ p^2\,|\,\varphi_k(n)}} 1 + \sum_{q\equiv 1 \pmod p}\ \sum_{\substack{n\leq x\\ q\,|\,\varphi_k(n)}} 1$$

$$\leq x(2c_3\log\log x)^k\left(\frac{1}{p} + \sum_{\substack{q\leq x\\ q\equiv 1 \pmod p}}\frac{1}{q}\right)$$

by the induction hypothesis, the fact that $p^2 \mid \varphi_k(n)$ implies $p|\varphi_k(n)$ and the observation that if $n \le x$ and $q \mid \varphi_k(n)$, then $q \le x$. Using (3.1) to estimate the remaining sum we have

$$\begin{aligned}\sum_{\substack{n\le x\\ p\mid\varphi_{k+1}(n)}} 1 &\le x(2c_3\log\log x)^k\left(\frac{1}{p}+\frac{1}{p-1}c_3\log\log x\right)\\ &\le \frac{x}{p}(2c_3\log\log x)^k\left(1+\frac{3}{2}c_3\log\log x\right)\\ &\le \frac{x}{p}(2c_3\log\log x)^{k+1}.\end{aligned}$$

REMARK: If we let $S_k(x,p)$ denote the sum of $1/q$ for primes $q \le x$ with $p \mid \varphi_k(q)$, then by essentially the same proof we have

$$S_k(x,p) \le \frac{1}{p}(2c_3\log\log x)^k$$

for the same set of p, k, x as in Theorem 3.5. Although this result will not be of use to us it is interesting to compare it with Theorem 3.4 in the case $n = p$.

§4. More on the iterated phi-function

Using the constants c_3, c_5 of the preceding section, let

$$\alpha_k(x) = (c_5k^{-1}\log\log x)^k, \quad \beta_k(x) = (2c_3\log\log x)^k.$$

Let

$$f_k(n,x) = \sum_{\substack{p\le(\log\log x)^k\\ p\nmid\varphi_k(n)}}\frac{1}{p} + \sum_{\substack{p>(\log\log x)^k\\ p\mid\varphi_k(n)}}\frac{1}{p}.$$

Thus if $n \le x$, $f_k(n,x)$ measures in some sense how far $\varphi_k(n)/\varphi_{k+1}(n)$ is from

$$\prod_{p\le(\log\log x)^k}(1-1/p)^{-1}.$$

Theorem 4.1. *There is an absolute constant c_7 such that*

$$\frac{1}{x}\sum_{n\le x} f_k(n,x) \le c_7(\log k)/(\log\log\log x - \log k)$$

for all $x \geq x_0$, $1 \leq k < \log\log x$.

Proof: We have $\alpha_k(x) \leq (\log\log x)^k \leq \beta_k(x)$ for all $k \geq 1$, $x \geq 3$. Thus

$$\begin{aligned}\sum_{n\leq x} f_k(n,x) &= \sum_{p\leq(\log\log x)^k} \frac{1}{p} \sum_{\substack{n\leq x \\ p\nmid \varphi_k(n)}} 1 + \sum_{p>(\log\log x)^k} \frac{1}{p} \sum_{\substack{n\leq x \\ p\mid \varphi_k(n)}} 1 \\ &\leq \sum_{p\leq\alpha_k(x)} \frac{1}{p} \sum_{\substack{n\leq x \\ p\nmid \varphi_k(n)}} 1 + \sum_{\alpha_k(x)<p\leq\beta_k(x)} \frac{1}{p} \sum_{n\leq x} 1 + \sum_{p>\beta_k(x)} \frac{1}{p} \sum_{\substack{n\leq x \\ p\mid \varphi_k(n)}} 1 \\ &= S_1 + S_2 + S_3, \text{ say.} \end{aligned} \tag{4.1}$$

If $p\nmid \varphi_k(n)$, then n is not divisible by any prime q with $p|\varphi_k(q)$. Thus by Brun's method (see Halberstam-Richert [11]) we have

$$\begin{aligned}\sum_{\substack{n\leq x \\ p\nmid \varphi_k(n)}} 1 &\ll x \prod_{\substack{q\leq x \\ p\mid \varphi_k(q)}} (1-\frac{1}{q}) \ll x \exp\left(-\sum_{\substack{q\leq x \\ p\mid \varphi_k(q)}} \frac{1}{q}\right) \\ &\leq x\exp(-S_k'(x,p))\end{aligned}$$

uniformly for all x, p, k, where $S_k'(x,p)$ is defined in section 3. Let

$$\alpha_k'(x) = \alpha_k(x)/(c_4\log\log x).$$

Thus by Theorem 3.4, there is an absolute constant c_8 such that

$$\sum_{\substack{n\leq x \\ p\nmid \varphi_k(n)}} 1 \leq \begin{cases} c_8 x e^{-\alpha_k(x)/(p-1)}, & \text{if } p > \alpha_k'(x)+1 \\ c_8 x(\log x)^{-c_4}, & \text{if } p \leq \alpha_k'(x)+1 \end{cases} \tag{4.2}$$

for all $x \geq x_0$, $p \leq (\log x)^2$, $k \leq c_6\log\log x$.

The theorem holds trivially if $k \gg \log\log x$, so assume $k/\log\log x \leq \min\{\frac{1}{2}c_5, c_6\}$. Since for any k, $\alpha_k(x) \leq (\log x)^{c_5/e}$, (4.2) implies

$$\begin{aligned} S_1 &= \sum_{p\leq\alpha_k(x)} \frac{1}{p} \sum_{\substack{n\leq x \\ p\nmid \varphi_k(n)}} 1 \\ &\leq c_8 x(\log x)^{-c_4} \sum_{p\leq\alpha_k'(x)+1} \frac{1}{p} + c_8 x \sum_{\alpha_k'(x)<p\leq\alpha_k(x)} \frac{1}{p} e^{-\alpha_k(x)/(p-1)} \\ &\ll x/\log\alpha_k(x) \ll x/(\log\log\log x - \log k). \end{aligned} \tag{4.3}$$

Since we are assuming $k \le \frac{1}{2}c_5 \log\log x$, we have

$$S_2 \le x \sum_{\alpha_k(x)<p\le\beta_k(x)} \frac{1}{p} = x\left(\log\log\beta_k(x) - \log\log\alpha_k(x) + O(1/\log\alpha_k(x))\right)$$
$$\ll x(\log k)/(\log\log\log x - \log k) \tag{4.4}$$

uniformly in k.

For S_3 we use Theorem 3.5 to estimate the inner sum. We have

$$S_3 \le x\beta_k(x) \sum_{p>\beta_k(x)} \frac{1}{p^2} \ll x/\log\beta_k(x) \le x/\log\log\log x.$$

Assembling this estimate, (4.1), (4.3) and (4.4) we have the theorem.

Theorem 4.2. *Let $\epsilon(x) > 0$ tend to 0 arbitrarily slowly as $x \to \infty$. If $k \le (\log\log x)^{\epsilon(x)}$, then the normal order of $\varphi_k(n)/\varphi_{k+1}(n)$ for $n \le x$ is $ke^\gamma \log\log\log x$.*

Proof: Let $\delta > 0$ be arbitrary. Let x be large and let $k \le (\log\log x)^{\epsilon(x)}$. From Theorem 4.1, the average value of $f_k(n,x)$ for $n \le x$ is $O(\epsilon(x))$. Thus if $x \ge x_0(\delta)$, $f_k(n,x) < \delta$ for at least $(1-\delta)x$ values of $n \le x$. But

$$\frac{\varphi_k(n)}{\varphi_{k+1}(n)} = \left(\prod_{p\le(\log\log x)^k}\left(1-\frac{1}{p}\right)^{-1}\right)\left(\prod_{\substack{p\le(\log\log x)^k\\ p\nmid\varphi_k(n)}}\left(1-\frac{1}{p}\right)\right)$$
$$\left(\prod_{\substack{p>(\log\log x)^k\\ p\mid\varphi_k(n)}}\left(1-\frac{1}{p}\right)^{-1}\right),$$

so that

$$\log\left(\frac{\varphi_k(n)}{\varphi_{k+1}(n)}\prod_{p\le(\log\log x)^k}\left(1-\frac{1}{p}\right)\right) \ll f_k(n,x).$$

Thus, for at least $(1-\delta)x$ values of $n \le x$

$$\frac{\varphi_k(n)}{\varphi_{k+1}(n)} = (1+O(\delta))ke^\gamma \log\log\log x.$$

Theorem 4.3. *Let $\epsilon(x) > 0$ tend to 0 arbitrarily slowly as $x \to \infty$. Then if $k \le \epsilon(x) \log\log\log x / \log\log\log\log x$, the normal order of $\varphi(n)/\varphi_{k+1}(n)$ for $n \le x$ is $k! e^{k\gamma} (\log\log\log x)^k$.*

Proof: From the proof of Theorem 4.2, the number of $n \le x$ for which

$$\left| \log \left(\frac{\varphi_j(n)}{\varphi_{j+1}(n)} (je^{\gamma} \log\log\log x)^{-1} \right) \right| \le \frac{\log j}{\log\log\log x}$$

fails is $O\left(\frac{x \log j}{\log\log\log x}\right)$ uniformly for any $j \le k$. Summing for $j = 1, \cdots, k$ we have that

$$\left| \log \left(\frac{\varphi(n)}{\varphi_k(n)} \left(k! e^{k\gamma} (\log\log\log x)^k\right)^{-1} \right) \right| \le \epsilon(x)$$

but for at most $O(\epsilon(x)x)$ integers $n \le x$. Since $\epsilon(x) \to 0$, we have our theorem.

Theorem 4.4. *There is an absolute constant $c_9 > 0$ such that if $1 \le k \le c_9 \log\log x$, then the number of $n \le x$ for which*

$$\frac{\varphi_k(n)}{\varphi_{k+1}(n)} > k(\log\log\log x - \log k) \tag{4.5}$$

fails is $O(xk^{-1}(\log\log\log x - \log k)^{-1})$. In particular

$$\max_k \frac{\varphi_k(n)}{\varphi_{k+1}(n)} \gg \log\log n$$

for a set of n of asymptotic density 1.

Proof: As in (4.3), if $c_9 > 0$ is small enough, then

$$\sum_{n \le x} \sum_{\substack{p \le \alpha_k(x) \\ p \nmid \varphi_k(n)}} \frac{1}{p} = \sum_{p \le \alpha_k(x)} \frac{1}{p} \sum_{\substack{n \le x \\ p \nmid \varphi_k(n)}} 1 \ll x / \log \alpha_k(x)$$
$$\ll xk^{-1}(\log\log\log x - \log k)^{-1}$$

uniformly for all $k \le c_9 \log\log x$. Thus but for at most $O(xk^{-1}(\log\log\log x - \log k)^{-1})$ exceptional values of $n \le x$, we have

$$\prod_{\substack{p \le \alpha_k(x) \\ p \nmid \varphi_k(n)}} \left(1 - \frac{1}{p}\right) \ge \frac{3}{4}.$$

For those values of n we have

$$\frac{\varphi_k(n)}{\varphi_{k+1}(n)} \geq \prod_{p \leq \alpha_k(x)} \left(1-\frac{1}{p}\right)^{-1} \prod_{\substack{p \leq \alpha_k(x) \\ p \nmid \varphi_k(n)}} \left(1-\frac{1}{p}\right)$$

$$\geq \frac{3}{4} e^{\gamma} \log \alpha_k(x) - 1$$

$$> k(\log\log\log x - \log k)$$

provided x is sufficiently large and $c_9 \leq \frac{1}{2}c_5$. This proves the theorem.

Theorem 4.5. *There is a positive absolute constant c_{10} such that the set of natural numbers n, for which there is some k with $\varphi_k(n)$ divisible by every prime up to $(\log n)^{c_{10}}$, has asymptotic density 1 .*

Proof: There is a positive absolute constant c_{11} such that if $k = [c_{11} \log\log x]$, then $\alpha_k(x) > (\log x)^{c_{11}}$. Then by (4.2) we have

$$\sum_{\substack{n \leq x \\ p \nmid \varphi_k(n)}} 1 \leq c_8 x (\log x)^{-c_4}$$

for all $x \geq x_0$, primes $p \leq (\log x)^{c_{11}/2}$, $k = [c_{11} \log\log x]$. Let $c_{10} = \min\{c_4/2, c_{11}/2\}$. Then

$$\sum_{p \leq (\log x)^{c_{10}}} \sum_{\substack{n \leq x \\ p \nmid \varphi_k(n)}} 1 < c_8 x (\log x)^{-c_4/2}.$$

Thus but for at most $c_8 x (\log x)^{-c_4/2}$ exceptional integers $n \leq x$ we have $p \mid \varphi_k(n)$ for every prime $p \leq (\log x)^{c_{10}}$ if $k = [c_{11} \log\log x]$ and $x \geq x_0$. This proves the theorem.

In contrast to Theorem 4.5 we give the following result. The proof is not an application of the theorems in section 3, but rather follows from the easy identity (2.30). Let $\nu(m)$ denote the number of distinct prime factors of m .

Theorem 4.6. *Let $\Phi(n) = n \prod_{k=1}^{\infty} \varphi_k(n)$. Then for all n, $\nu(\Phi(n)) \leq \lceil (\log n)/\log 2 \rceil$. In particular, for all n there is some prime $p \ll \log n \log\log n$ with $p \nmid \Phi(n)$.*

Proof: For any $n > 1$ we have, using (2.30) with $p = 2$ and (1.1),

$$\nu(\Phi(n)) = 1 + \sum_{\substack{q > 2 \\ q \mid \Phi(n)}} 1 \leq 1 + \sum_{q>2} F_q(n) \leq \begin{cases} 1 + F(n), & n \text{ odd} \\ F(n), & n \text{ even} \end{cases}$$

$$= k(n) \leq \lceil (\log n)/\log 2 \rceil.$$

§5. Aliquot sequences

Let $s(n) = \sigma(n) - n$, where σ is the sum of the divisors function. Let $s_1(n) = s(n)$, $s_2(n) = s(s_1(n))$, etc. What is now known as the Catalan-Dickson conjecture is that for any n , the "aliquot sequence" n, $s_1(n)$, $s_2(n)$, $\cdots$ eventually terminates at 0 or is eventually periodic. The least n for which this conjecture is in doubt is 276. Guy and Selfridge [10] instead conjecture that for infinitely many n the aliquot sequence beginning with n tends to ∞. The function $s(n)$ has been studied since antiquity when numbers were classified as perfect, abundant or deficient depending on whether $s(n) = n$, $s(n) > n$ or $s(n) < n$, respectively.

As discussed in the Introduction, the first author proved in [8] that for each $\epsilon > 0$ and k , the set of n for which

$$\frac{s_{j+1}(n)}{s_j(n)} > \frac{s(n)}{n} - \epsilon \text{ for } j = 1, \cdots, k \tag{5.1}$$

has asymptotic density 1 . Further, he claimed that similar methods would show that

$$\frac{s_{j+1}(n)}{s_j(n)} < \frac{s(n)}{n} + \epsilon \text{ for } j = 1, \cdots, k$$

for a set of n of asymptotic density 1. This claim of a proof is now retracted but we still remain convinced of the truth of this statement; it is our Conjecture 3 in section 1. We now give a proof of the case $k = 1$.

Theorem 5.1. *For each $\epsilon > 0$, the set of n with*

$$\frac{s_2(n)}{s(n)} < \frac{s(n)}{n} + \epsilon \tag{5.2}$$

has asymptotic density 1 .

Proof: Let $1 > \delta > 0$ be arbitrary. We shall show that for all large x, the number of $n \leq x$ for which (5.2) fails is at most $c\delta x$ for some absolute constant c.

Let $P(n)$ denote the largest prime factor of n. If $\eta > 0$ is sufficiently small, then the number of $n \leq x$ for which

$$P(n) > x^{\eta}, \quad P(n)^2 \nmid n \tag{5.3}$$

fails is at most δx for all large x . This result follows from either sieve methods or work of Dickman and others on the distribution of integers n with no large prime factors. Fix such a number η.

Since $\sum_{n \leq x} \sigma(n)/n \ll x$, there is a number B so large that the number of $n \leq x$ for which

$$\sigma(n)/n \leq B \tag{5.4}$$

fails is at most δx for all large x. Fix such a number B.

If $\alpha > 0$, say that an integer n is *α-primitive* if $s(n)/n \geq \alpha$ and if $d \mid n$, $d < n$, then $s(d)/d < \alpha$ (also called a primitive $(1+\alpha)$-abundant number).

Let α be a rational number with $0 < \alpha_1 < 1/2$, $\alpha_1 \leq \epsilon/4B$. Also let α_2 be a rational number with $0 < \alpha_2 < \alpha_1\eta/24$. Since α_1, α_2 are rational, it follows from the proof in [6] for the case $\alpha = 1$, that

$$\sum\nolimits^{(1)} 1/a < \infty, \quad \sum\nolimits^{(2)} 1/a < \infty,$$

where for $i = 1,\ 2$, $\sum^{(i)}$ denotes a sum over α_i-primitive numbers. Since $a/\varphi(a)$ is bounded if a is α_i-primitive, it follows that there is a number T so large that

$$\sum_{a \geq T}\nolimits^{(1)} 1/a < \delta, \quad \sum_{a \geq T}\nolimits^{(2)} 1/\varphi(a) < \delta\eta. \tag{5.5}$$

Also assume T is so large that

$$T > \frac{1}{\alpha_2} + 1, \quad \prod_{p \geq T} \frac{p^2}{p^2 - 1} < 1 + \frac{1}{2}\alpha_2. \tag{5.6}$$

If $n > 1$ is an integer, factor n as $n_1 n_2$ and $s(n)$ as $N_1 N_2$ where every prime factor of $n_1 N_1$ is less than T and every prime factor of $n_2 N_2$ is at least T. It follows from the work in [8] that but for a set of n of asymptotic density 0 , we have

$$n_1 = N_1. \tag{5.7}$$

The idea of the proof is that but for a set of n of asymptotic density 0 , the number n_1 is not too large, say $n_1 < (\log\log n)^{1/2}/\prod_{p<T} p$. For these n , there is almost certainly a prime $q \| n$ with

$$q \equiv -1 \pmod{n_1 \prod_{p<T} p}.$$

Then but for a set of n of asymptotic density 0 , we have $n_1 \prod_{p<T} p \mid \sigma(n)$. For these n we have $n_1 | s(n)$ and $(\prod_{p<T} p, s(n)/n_1) = 1$, i.e. (5.7) holds.

The number of $n \leq x$ with n_2 divisible by an α_1-primitive number a is at most

$$\sum_{(a, \prod_{p<T} p)=1}\nolimits^{(1)} \left[\frac{x}{a}\right] \leq x \sum_{a \geq T}\nolimits^{(1)} \frac{1}{a} < \delta x$$

by (5.5). Thus but for at most δx exceptional values of $n \leq x$, we have

$$\sigma(n_2)/n_2 < 1 + \alpha_1. \tag{5.8}$$

Suppose now that (5.2) fails for n. By adding 1 to both sides, we get

$$\frac{\sigma(s(n))}{s(n)} \geq \frac{\sigma(n)}{n} + \epsilon,$$

so that from (5.7) and (5.4)

$$\begin{aligned}\frac{\sigma(N_2)}{N_2} &\geq \frac{\sigma(N_2)/N_2}{\sigma(n_2)/n_2} = \frac{\sigma(s(n))/s(n)}{\sigma(n)/n} \\ &\geq 1 + \frac{\epsilon}{\sigma(n)/n} \geq 1 + \frac{\epsilon}{B} \geq 1 + 4\alpha_1.\end{aligned}$$

Factor N_2 as $N_3 N_4$ where every prime in N_3 also divides n and $(N_4, n) = 1$. If $N_3 = \prod p_i^{\beta_i}$, where $p_i \geq T$ are distinct primes and each $\beta_i \geq 1$, then

$$\begin{aligned}\frac{\sigma(N_3)}{N_3} &= \prod \frac{p_i - p_i^{-\beta_i}}{p_i - 1} < \prod \frac{p_i}{p_i - 1} \\ &\leq \left(\prod \frac{p_i}{p_i - 1} \cdot \frac{p_i}{p_i + 1}\right) \frac{\sigma(n_2)}{n_2}\end{aligned}$$

since each $p_i \mid n_2$. Then from (5.6) and (5.8) we have

$$\frac{\sigma(N_3)}{N_3} < \left(\prod_{p \geq T} \frac{p^2}{p^2 - 1}\right) \frac{\sigma(n_2)}{n_2} < \left(1 + \frac{1}{2}\alpha_2\right)(1 + \alpha_1) < 1 + 2\alpha_1.$$

Thus

$$\frac{\sigma(N_4)}{N_4} = \frac{\sigma(N_2)/N_2}{\sigma(N_3)/N_3} > \frac{1 + 4\alpha_1}{1 + 2\alpha_1} > 1 + \alpha_1,$$

so $s(n)$ is divisible by an α_1 -primitive number a_1 not divisible by any prime below T and with $(a_1, n) = 1$.

We now show that any α_1-primitive number a_1 which is not divisible by any primes below T must have an α_2-primitive divisor a_2 with $a_2 \leq a_1^{\eta/2}$. Indeed, let the distinct prime factors of a_1 be $q_1, \cdots, q_t$, where

$$T \leq q_1 < \cdots < q_t.$$

Let $a_0 = q_1 q_2 \cdots q_{[\eta t/2]}$. Then

$$a_0 \leq (q_1 \cdots q_t)^{[\eta t/2]/t} \leq a_1^{\eta/2},$$

so it is sufficient to show $\sigma(a_0)/a_0 \geq 1 + \alpha_2$, for this will guarantee it having an α_2-primitive divisor a_2.

Note that

$$[\eta t/2] \geq \eta t/3,$$

since if not, we have $t < 6/\eta$, which implies by (5.6)

$$\begin{aligned} 1+\alpha_1 \leq \frac{\sigma(a_1)}{a_1} &< \prod_{i\leq t} \frac{q_i}{q_i-1} \\ &< \left(1+\frac{1}{T-1}\right)^{6/\eta} < (1+\alpha_2)^{6/\eta} \\ &< \left(1+\frac{\alpha_1\eta}{24}\right)^{6/\eta} < 1+\frac{1}{2}\alpha_1, \end{aligned}$$

a contradiction. Thus from (5.6),

$$\begin{aligned} \frac{\sigma(a_0)}{a_0} = \prod_{i\leq[\frac{\eta t}{2}]} \frac{q_i+1}{q_i} &> \left(\prod_{i\leq[\frac{\eta t}{2}]} \frac{q_i}{q_i-1}\right) \prod_{p\geq T} \frac{p^2-1}{p^2} \\ &> \left(\prod_{i\leq t} \frac{q_i}{q_i-1}\right)^{[\eta t/2]/t} \left(1+\frac{1}{2}\alpha_2\right)^{-1} \\ &> (1+\alpha_1)^{\eta/3} \left(1+\frac{1}{2}\alpha_2\right)^{-1} \\ &> \left(1+\frac{24\alpha_2}{\eta}\right)^{\eta/3} \left(1+\frac{1}{2}\alpha_2\right)^{-1} > 1+\alpha_2. \end{aligned}$$

We have seen above, but for $O(\delta x)$ integers $n \leq x$, if $n \leq x$ does not satisfy (5.2), then $s(n)$ is divisible by an α_1-primitive number a_1 with $(a_1, n) = 1$ and a_1 not divisible by any prime below T and further that (5.3) holds. Thus such an n must have $s(n)$ divisible by an α_2-primitive number a_2 with $(a_2, n) = 1$, with a_2 not divisible by any prime below T and with

$$a_2 \leq a_1^{\eta/2} \leq s(n)^{\eta/2} < x^{2\eta/3}$$

for x large. For such an n, we factor it as mp where $p = P(n)$. From (5.3), $m < x^{1-\eta}$, $p \nmid m$. Consider the α_2-primitive number a_2 just discovered dividing $s(n)$. We have $s(n) = p(\sigma(m) - m) + \sigma(m)$, so that

$$p(\sigma(m) - m) \equiv -\sigma(m) \pmod{a_2}. \tag{5.9}$$

Since $(a_2, pm) = 1$ we have $(a_2, \sigma(m)) = 1$ so that there is a certain residue class $c(m, a_2) \mod a_2$ such that if p, m, a_2 satisfy (5.9), then $p \equiv c(m, a_2)$

mod a_2. Thus but for $O(\delta x)$ integers, the number of $n \leq x$ which do not satisfy (5.2) is at most

$$\begin{aligned}
&\sum_{x^{2\eta/3} \geq a_2 \geq T}^{(2)} \sum_{m < x^{1-\eta}} \sum_{\substack{p \leq x/m \\ p \equiv c(m,a) \pmod{a_2}}} 1 \\
&\ll \sum_{x^{2\eta/3} \geq a_2 \geq T}^{(2)} \sum_{m < x^{1-\eta}} \frac{x}{\varphi(a_2) m \log(x/a_2 m)} \\
&\ll \frac{x}{\eta} \sum_{a_2 \geq T} \frac{1}{\varphi(a_2)} < \delta x,
\end{aligned}$$

where we used the Brun-Titchmarsh theorem for the first inequality and (5.5) for the last.

Theorem 5.2. *Conjecture 4 implies Conjecture 3.*

Proof: Let k be a natural number. Let $T = T(n)$ tend to infinity very slowly, say $T(n)$ is the $3k$-fold iterated logarithm. For $j = 1, \cdots, k$, factor $s_j(n) = m_j n_j$ where every prime factor of m_j is less than T and every prime factor of n_j is at least T. We analogously factor $n = m_0 n_0$. In the same way as (5.7) is established, the set of n for which

$$m_0 = m_1 = \cdots = m_k \tag{5.10}$$

fails has asymptotic density 0 . Indeed, this is essentially established in [8].

By a simple averaging argument one can show that the set of n for which

$$\sum_{\substack{p \mid n \\ p \geq T}} \frac{1}{p-1} < \frac{1}{T}$$

fails has asymptotic density 0 . Indeed, the average value of the sum is $\sim (T \log T)^{-1}$. But

$$\log \frac{\sigma(n_0)}{n_0} < \log \left(\prod_{\substack{p \mid n \\ p \geq T}} \left(1 + \frac{1}{p-1}\right) \right) < \sum_{\substack{p \mid n \\ p \geq T}} \frac{1}{p-1}.$$

Thus, but for a set of n of asymptotic density 0 , we have

$$\sigma(n_j)/n_j < e^{1/T} \quad \text{for } j = 0, 1, \cdots, k, \tag{5.11}$$

using Conjecture 4 in the form: if $\mathcal{A}$ has an asymptotic density 0, then $s^{-1}(\mathcal{A})$ has asymptotic density 0.

By the same argument involved with (5.4), we have that the set of n for which

$$\frac{\sigma(m_0)}{m_0} < \log T \tag{5.12}$$

fails has asymptotic density 0 . Then from (5.10) and (5.11), for $j \le k$ we have

$$\begin{aligned}\frac{s_{j+1}(n)}{s_j(n)} - \frac{s(n)}{n} &= \frac{\sigma(m_0)}{m_0}\left(\frac{\sigma(n_j)}{n_j} - \frac{\sigma(n_0)}{n_0}\right) \\ &< (\log T)(e^{1/T} - 1) \ll (\log T)/T = o(1),\end{aligned}$$

which gives Conjecture 3.

REMARK. Note that (5.10), the case $j = 0$ of (5.11) (which does not require Conjecture 4) and (5.12) immediately give

$$\begin{aligned}\frac{s(n)}{n} - \frac{s_{j+1}(n)}{s_j(n)} &= \frac{\sigma(m_0)}{m_0}\left(\frac{\sigma(n_0)}{n_0} - \frac{\sigma(n_j)}{n_j}\right) \\ &< (\log T)(e^{1/T} - 1) = o(1).\end{aligned}$$

That is, (5.1) holds for all n, but for a set of asymptotic density 0, the principal result of [8].

Theorem 5.3. *Let $S_k(x)$ denote the number of odd numbers $m \le x$ not in the range of the function s_k. There is a positive number δ_0 such that*

$$S_k(x) \ll x^{1-\delta_0}$$

uniformly for all natural numbers k and $x > 0$.

Proof: Let $E(x, y)$ denote the number of odd integers $n \le x$ with $r(n) \le y$, where $r(n)$ is the number of representations of n in the form $1+p+q$ where $p < q$ are primes. Since

$$s(pq) = 1 + p + q,$$

it follows that for any $y \ge 0$

$$S_1(x) \le E(x, y). \tag{5.13}$$

We now prove that for any natural number k and any $y > 0$,

$$S_{k+1}(x) \le \frac{S_k(x^2)}{y} + E(x, y). \tag{5.14}$$

Let S_j denote the set of odd numbers not in the range of s_j. Suppose $n \in S_{k+1}$. Consider the $r(n)$ representations

$$n = 1 + p_i + q_i, \quad i = 1, \cdots, r(n)$$

where $p_i < q_i$ are primes. Then all of the numbers $p_i q_i$ are in S_k, for if $p_i q_i = s_k(m)$ for some m, then $n = s_{k+1}(m)$, contradicting $n \in S_{k+1}$. Note that the integers $p_i q_i$ are distinct and each $p_i q_i < n^2$. Moreover if p_j', q_j' are associated with n' and $n \neq n'$, then $p_i q_i \neq p_j' q_j'$. Thus

$$\begin{aligned} S_{k+1}(x) &= \#\{n \leq x : n \in S_{k+1}, r(n) \leq y\} + \#\{n \leq x : n \in S_{k+1}, r(n) > y\} \\ &\leq \#\{n \leq x : r(n) \leq y\} + y^{-1} \cdot \#\{m \leq x^2 : m \in S_k\} \\ &= E(x, y) + y^{-1} S_k(x^2), \end{aligned}$$

which is (5.14).

Next we show there is some $\delta_1 > 0$, $B > 0$ such that

$$E(x, y) \leq B y \log^{38} x \tag{5.15}$$

for all $x \geq 2$, $y \geq x^{1-\delta_1}$. This result follows from the proof in Montgomery and Vaughan [13]. To see this, let $E_0(x, y)$ denote the number of odd numbers n with $x/2 < n \leq x$ and $r(n) \leq y$. Then from the proof in [13], we have

$$E_0(x, x^{1-\frac{3}{2}\delta} \log^{-3} x) \ll x^{1-2\delta} \log^{35} x$$

uniformly for $\delta \leq \delta_0$ for some $\delta_0 > 0$. Let $z = x^{3\delta/2}$. Then for i such that $2^i \leq z$,

$$\begin{aligned} E_0\left(2^{-i}x, \frac{x}{z \log^3 x}\right) &\leq E_0\left(2^{-i}x, \frac{2^{-i}x}{2^{-i} z \log^3(2^{-i}x)}\right) \\ &\ll \frac{2^{-i}x}{(2^{-i}z)^{4/3}} \log^{35}(2^{-i}x) \leq 2^{i/3} \frac{x}{z^{4/3}} \log^{35} x. \end{aligned}$$

Let j be such that $2^j \leq z < 2^{j+1}$. Then

$$\begin{aligned} E\left(x, \frac{x}{z \log^3 x}\right) &= E\left(2^{-(j+1)}x, \frac{x}{z \log^3 x}\right) + \sum_{i=0}^{j} E_0\left(2^{-i}x, \frac{x}{z \log^3 x}\right) \\ &\ll 2^{-(j+1)}x + \sum_{i=0}^{j} 2^{i/3} \frac{x}{z^{4/3}} \log^{35} x \\ &\ll \frac{x}{z} \log^{35} x. \end{aligned}$$

Thus letting $y = xz^{-1}\log^{-3} x$, we have (5.15) for $y \geq x^{1-3\delta_0/2}\log^3 x$. Letting $\delta_1 = 5\delta_0/4$, we have (5.15) for $y \geq x^{1-\delta_1}$.

Suppose we know that for some specific $k \geq 1$, there is some constant $C(k) \geq B$ with

$$S_k(x) \leq C(k)x^{1-\delta_1}\log^{38} x \tag{5.16}$$

for all $x \geq e$. Then letting $y = 2^{19}(C(k)/B)^{1/2}x^{1-\delta_1}$ and using (5.14) and (5.15) we have

$$S_{k+1}(x) \leq C(k+1)x^{1-\delta_1}\log^{38} x,$$

where

$$C(k+1) := 2^{20}(C(k)B)^{1/2}. \tag{5.17}$$

Since we have (5.16) for $k = 1$ and $C(1) = B$ by (5.13) and (5.15), we thus have it for all k where $C(k)$ is inductively defined by (5.17). Note that $C(k) < 2^{40}B$ for all k. In addition, since $\delta_1 = 5\delta_0/4$, we have our theorem.

§6. Corrections for an earlier paper

In [9], the first and third authors considered the normal number of prime factors of $\varphi(n)$. The principal result is that this normal order is $\frac{1}{2}(\log\log n)^2$ and there is a Gaussian distribution with standard deviation $\frac{1}{\sqrt{3}}(\log\log n)^{3/2}$. It has been pointed out to us by Abdelhakim Smati that there is an error in the proof of Lemma 2.2 of this paper. We now give a (hopefully) correct proof of this result.

Let $\Omega_y(n)$ denote the number of prime factors $p \leq y$ of n counted with multiplicity. Lemma 2.1 of [9] gives the average order for $\Omega_y(p-1)$ for p prime:

$$\sum_{p\leq x}\Omega_y(p-1) = \frac{x\log\log y}{\log x} + O\left(\frac{x}{\log x}\right) \tag{6.1}$$

uniformly for $3 \leq y \leq x$. Lemma 2.2 estimates the square mean.

"Lemma 2.2". *If $3 \leq y \leq x$, then*

$$\sum_{p\leq x}\Omega_y(p-1)^2 = \frac{x(\log\log y)^2}{\log x} + O\left(\frac{x\log\log y}{\log x}\right)$$

where the implied constant is uniform.

Proof: Let u range over the integers with exactly 2 distinct prime factors, neither exceeding y. Then

$$\begin{aligned}\sum_{p\leq x}\Omega_y(p-1)^2 &= \sum_{p\leq x}\sum_{\substack{q^a\| p-1\\ q\leq y}} a^2 + 2\sum_{p\leq x}\sum_{u\mid p-1} 1\\ &= S_3 + S_4,\end{aligned}$$

say. (In [9], the expression for S_4 is wrong.)

As in [9], we get

$$S_3 = O\left(\frac{x \log\log y}{\log x}\right)$$

using (6.1) and the Brun-Titchmarsh inequality.

For S_4, we write

$$S_4 = S_{4,1} + S_{4,2}$$

where in $S_{4,1}$ neither prime power in u exceeds $x^{1/6}$ and is $S_{4,2}$ at least one prime power in u exceeds $x^{1/6}$. We have

$$\begin{aligned} S_{4,1} &= 2 \sum_{\substack{1<q^a,\, r^b \le x^{1/6} \\ q<r\le y}} \pi(x; q^a r^b, 1) \\ &= 2\mathrm{li}(x) \sum_{\substack{1<q^a,\, r^b \le x^{1/6} \\ q,r\le y}} \frac{1}{\varphi(q^a r^b)} + O\left(\frac{x}{\log^2 x}\right) \\ &= \frac{x(\log\log y)^2}{\log x} + O\left(\frac{x \log\log y}{\log x}\right) \end{aligned}$$

using the Bombieri-Vinogradov theorem and a simple calculation.

For $S_{4,2}$ we have

$$S_{4,2} \ll \sum_{p\le x} \Omega_y(p-1) \ll \frac{x \log\log y}{\log x},$$

using (6.1). This, together with our estimates for S_3 and $S_{4,1}$ completes the proof.

A. Smati also points out that the three cases on p. 350 of [9] for $p^2 \mid \varphi(n)$, $p > y$ (where now $y = (\log\log x)^2$) do not exhaust all possibilities. This is fixed by changing (i) to (i') $p^2 \mid n$. The number of $n \le x$ in this case is at most $\sum_{p>y} x/p^2 = o(x/y) = o(x)$.

We are grateful to A. Smati for pointing these difficulties out to us.

REFERENCES

[1] E. Bombieri, Le grand crible dans la théorie analytique des nombres, Astérisque **18** (1974), 1–87.

[2] H. Davenport, *Multiplicative Number Theory , 2nd edition*, Springer Verlag, New York, 1980.

[3] P.D.T.A. Elliott, *Probabilistic Number Theory , vols. I, II*, Springer Verlag, New York, 1980.

[4] J. Friedlander and A. Granville, Limitations to the equi-distribution of primes I, Ann. Math. **129** (1989), 363–382.
[5] J. Friedlander, A. Granville, A. Hildebrand and H. Maier, Oscillation theorems for primes in arithmetic progressions and for sifting functions, preprint..
[6] P. Erdös, On the density of the abundant numbers, J. London Math. Soc. **9** (1934), 278–282.
[7] P. Erdös, Some remarks on the iterates of the φ and σ functions, Colloq. Math. **17** (1967), 195–202.
[8] P. Erdös, On asymptotic properties of aliquot sequences, Math. Comp. **30** (1976), 641–645.
[9] P. Erdös and C. Pomerance, On the normal number of prime factors of $\varphi(n)$, Rocky Mountain Math. J. **15** (1985), 343–352.
[10] R. K. Guy and J. L. Selfridge, What drives an aliquot sequence?, Math. Comp. **29** (1975), 101–107; Corrigendum, Math. Comp. **34** (1980), 319–321.
[11] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
[12] H. W. Lenstra, Jr., Problem 6064, American Math. Monthly **82** (1975), p. 1016; Solution by the proposer, **84** (1977), p. 580.
[13] H. L. Montgomery and R. C. Vaughan, The exceptional set in Goldbach's problem, Acta Arith. **27** (1975), 353–370.
[14] S. S. Pillai, On a function connected with $\varphi(n)$, Bull. A.M.S. **35** (1929), 837-841.
[15] I. J. Schoenberg, Über die asymptotische Verteilung reeller Zahlen mod 1, Math. Z. **28** (1928), 171–200.
[16] H. Shapiro, An arithmetic function arising from the φ-function, American Math. Monthly **50** (1943), 18–30.

Paul Erdös
Mathematical Institute
Hungarian Academy of Sciences
Reáltanoda u. 13-15
Budapest, Hungary

Andrew Granville
School of Mathematics
Institute for Advanced Study
Princeton, NJ 08540

Carl Pomerance
Department of Mathematics
University of Georgia
Athens, Georgia 30602

Claudia Spiro
Department of Mathematics
M. I. T.
Cambridge, Massachusetts 02139