PACIFIC JOURNAL OF MATHEMATICS
Vol. 71, No. 2, 1977

ON AN ADDITIVE ARITHMETIC FUNCTION

K. ALLADI AND P. ERDÖS

We discuss in this paper arithmetic properties of the function $A(n) = \sum_{p^\alpha || n} \alpha p$. Asymptotic estimates of $A(n)$ reveal the connection between $A(n)$ and large prime factors of n. The distribution modulo 2 of $A(n)$ turns out to be an interesting study and congruences involving $A(n)$ are considered. Moreover the very intimate connection between $A(n)$ and the partition of integers into primes provides a natural motivation for its study.

0. Introduction. Let a positive integer n be expressed as a product of distinct primes in the canonical fashion $n = \prod_{i=1}^{r} p_i^{\alpha_i}$. Define a function $A(n) = \sum_{i=1}^{r} \alpha_i p_i$.

(i) The function $A(n)$ is not injective. In fact for a fixed integer m, the number of solutions in n to $A(n) = m$, is the number of partitions of m into primes.

(ii) $A(n)$ fluctuates in size appreciably. It is easily seen that $A(n) = n$ when n is a prime, while $A(n) = O(\log n)$ when n is a power of a small prime. Actually the "average order" of $A(n)$ turns out (as a corollary to Theorem 1.1) to be $\pi^2 n/6 \log n$. The term average order is defined below.

(iii) The function $A(n)$ is additive and one can expect it to take odd and even values with equal frequency.

The term "average order" calls for some explanation. We follow the usage in Hardy and Wright [6]. If $f(n)$ is a function defined on the positive integers we consider

$$F(x) = \sum_{n \leq x} f(n) .$$

Usually F can be expressed in terms of well behaved functions like polynomials or exponentials and the like. That is we seek an asymptotic estimate for F in terms of these functions. Then we seek a similar well behaved function g so that

$$F(x) = \sum_{n \leq x} f(n) \sim \sum_{n \leq x} g(n) .$$

The function g may be thought of as the average order of f. For instance if φ is the Euler function then

$$F(x) = \sum_{n \leq x} \varphi(n) = \frac{3x^2}{\pi^2} + O(x \log x) \sim \sum_{n \leq x} \frac{6n}{\pi^2}$$

so the average order of $\varphi(n)$ is $6n/\pi^2$.

It is surprising that the function $A(n)$ with such nice arithmetic properties has not been studied in detail. Besides the work of one of us (KA; [1]) some of the other references are [4], [7], [8], [9] and [10]. Of course the contents of this paper are different.

1. Average order estimates of $A(n)$. Here and in what follows the letter p (with or without subscript) shall always denote a prime.

So let $n = \prod_{i=1}^{r} p_i^{\alpha_i}$ and let $\Omega(n) = \sum_{i=1}^{r} \alpha_i$, $\omega(n) = r$. It is a well known result of Hardy and Ramanujan [6] that both $\Omega(n)$ and $\omega(n)$ have average order $\log \log n$, which tells us that generally the majority or prime factors occur only once. Applying this idea to $A(n)$ one expects it to have the same average order as $A^*(n) = \sum_{i=1}^{r} p_i$. In this sum it is natural to believe that the largest prime factor of $n(P_1(n)$ say) dominates the others so that $A(n)$ and $P_1(n)$ have the same average order. In fact this can be deduced as a corollary to Theorem 1.1 (where we prove much more) and the average order of $A(n)$ is $\pi^2 n/6 \log n$.

Let us assume without loss of generality that $p_1 < p_2 < \cdots < p_r$. Then let $P_1(n) = p_r$; $P_2(n) = P_1(n/P_1(n))$; $P_3(n) = P_1(n/P_1(n)P_2(n))$, etc., and in general

$$P_k(n) = \begin{cases} P_1\left(\dfrac{n}{P_1(n) \cdots P_{k-1}(n)}\right) & \text{for} \quad k \leqq \Omega(n) \\ 0 \quad \text{for} \quad k > \Omega(n) . \end{cases}$$

Thus $P_k(n)$ is the kth largest prime factor of n.

THEOREM 1.1. *For all integers $m \geqq 1$ we have*

$$\sum_{n \leqq x} P_m(n) \sim \sum_{n \leqq x} \{A(n) - P_1(n) - \cdots - P_{m-1}(n)\} \sim \frac{k_m x^{1+(1/m)}}{(\log x)^m}$$

where $k_m > 0$ is a constant depending only on m, and is a rational multiple of $\zeta(1 + 1/m)$ where ζ is the Riemann Zeta function.

LEMMA 1.2. *If $s > 1$ and x a large real number then*

$$\sum_{p \geqq x} \frac{1}{p^s} = \frac{1}{(s-1)x^{s-1}(\log x)} + O\left(\frac{1}{x^{s-1}(\log x)^2}\right) .$$

Proof. The proof of Lemma 1.2 is given by a simple direct method of using Stieltjes integrals, integration by parts and the prime number theorem in the form

$$(1.1) \qquad \pi(x) - li(x) = O\left(\frac{x}{\log^3 x}\right)$$

for all $\delta \geqq 2$. We have

$$
\begin{aligned}
\sum_{p \geqq x} \frac{1}{p^s} &= \int_{x-}^{\infty} \frac{d\pi(y)}{y^s} = \int_{x-}^{\infty} \frac{dy}{y^s \log y} + \int_{x-}^{\infty} \frac{d\{\pi(y) - li(y)\}}{y^s} \\
&= \int_{x-}^{\infty} \frac{dy}{y^s \log y} + O\Big(\frac{1}{x^{s-1} \log^\delta x}\Big) \qquad (1.2)\\
&\quad + \int_x^{\infty} \{\pi(y) - li(y)\} O\Big(\frac{1}{y^{s+1}}\Big) dy
\end{aligned}
$$

for $\delta \geqq 2$ because of (1.1). Now

$$
\begin{aligned}
\int_x^{\infty} \frac{\{\pi(y) - li(y)\}dy}{y^{s+1}} &= O\Big(\int_x^{\infty} \frac{dy}{y^s \log^\delta y}\Big) \qquad (1.3)\\
&= O\Big(\frac{1}{x^{s-1} \log^\delta x}\Big) .
\end{aligned}
$$

So (1.3) and (1.2) give

$$
\sum_{p \geqq x} \frac{1}{p^s} = \int_x^{\infty} \frac{dy}{y^s \log y} + O\Big(\frac{1}{x^{s-1} \log^2 x}\Big) . \qquad (1.4)
$$

But then

$$
\begin{aligned}
\int_x^{\infty} \frac{dy}{y^s \log y} &= \frac{1}{(s-1)x^{s-1} \log x} + O\Big(\int_x^{\infty} \frac{dy}{y^s \log^2 y}\Big) \qquad (1.5)\\
&= \frac{1}{(s-1)x^{s-1} \log x} + O\Big(\frac{1}{x^{s-1} \log^2 x}\Big) .
\end{aligned}
$$

Clearly (1.5) and (1.4) prove Lemma (1.2)

The above lemma establishes the following result which will be used often in the proof of Theorem 1.1.

LEMMA 1.3. *Let m be a positive integer and $s > 1, r \geqq 1$ be given real numbers. Then for x and z sufficiently large real numbers with $x^{1+(1/m)} < z < x^{3m}$ we have*

$$
\sum_{x \leqq p \leqq z/2} \frac{1}{p^s \log^r (z/p)} = \frac{1}{(s-1)x^{s-1} \log x \log^r (z/x)} + O\Big(\frac{\log\log x}{x^{s-1} \log^2 x \log^r (z/x)}\Big) .
$$

Proof. We break up the range of summation as

$$
\begin{aligned}
&\sum_{x \leqq p \leqq x \log^B x} \frac{1}{p^s \log^r (z/p)} + \sum_{x \log^B x < p \leqq z/2} \frac{1}{p^s \log^r (z/p)} \\
&= \sigma_1 + \sigma_2 \text{ respectively}
\end{aligned}
$$

where B for the moment is a constant not specified. Now in σ_1, $\log (z/p) = \log (z/x) + O(\log \log x)$ so that

$$\frac{1}{\log^r (z/p)} = \frac{1}{\log^r (z/x)} + O\Big(\frac{\log\log x}{\log x \cdot \log^r (z/x)}\Big)$$

because $\log (z/x)$ and $\log x$ are of the same order of magnitude. Now the above estimate, together with Lemma 1.2 gives

$$\sigma_1 = \frac{1}{(s-1)x^{s-1} \log x \log^r (z/x)} + O\Big(\frac{\log\log x}{x^{s-1}\log^2 x \log^r (z/x)}\Big).$$

To estimate σ_2 again apply Lemma 1.2 to get

$$\sigma_2 = O\Big(\sum_{x\log^B x<p} \frac{1}{p^s}\Big) = O\Big(\frac{1}{x^{s-1}\log x \cdot \log^{(s-1)B} x}\Big).$$

Comparing σ_1 and σ_2 we note that by a suitable choice of B Lemma 1.3 is true.

The crucial point in Lemma 1.3 is that by choice of z, $\log (z/x)$ and $\log x$ are of the same order of magnitude.

An argument similar to Lemma 1.2 yields the following:

LEMMA 1.4. *If* $s, r \geqq 0$, *then*

$$\sum_{p\leqq x} \frac{p^s}{(\log p)^r} = \frac{x^{s+1}}{(s+1)(\log x)^{r+1}} + O\Big(\frac{x^{s+1}}{\log^{r+2} x}\Big).$$

We omit the proof of Lemma 1.4, since it is similar to Lemma 1.2. Here we have to consider

$$\sum_{p\leqq x} \frac{p^s}{\log^r p} = \int_{2^-}^{x^+} \frac{y^s d\pi(y)}{\log^r y}$$

and compute just as we did in Lemma 1.2.

We now move on to the proof of Theorem 1.1. The proof involves complicated estimates in several places and we shall elaborate in detail the more important ones.

Proof of Theorem 1.1. We are first going to estimate $\sum_{n\leqq x} P_m(n)$. Let an integer n be written as $n = kp_1 \cdots p_m$, $p_1 \leqq p_2 \leqq \cdots \leqq p_m$, $P_1(k) \leqq p_1$, and let

$$k = P_m^*(n) = \frac{n}{P_1(n)\cdots P_m(n)}.$$

We keep $k = P_m^*(n)$ fixed and ask for those $n \leqq x$ for which $P_m^*(n) = k$. We sum $P_m(n)$ over these n and finally sum over k. Actually only small values of k will contribute to the principal term and large values will be treated separately.

So let k be small. The word "small" will be explained below.

Note that each p_i can range from $P_1(k)$ up to the minimum of p_{i+1} and $x/kp_m \cdots p_{i+1}$. So we shall break up the range of p_{i+1}, and discuss several cases, and in each of them we shall be able to decide without ambiguity which of p_{i+1} and $x/kp_m \cdots p_{i+1}$ is smaller, thereby determining the range of p_i.

Case 1. Let $p_m \leqq \sqrt[m]{x/k}$. Here the range of p_i is between $P_1(k)$ and p_{i+1} for $i = 1, 2, \cdots, m-1$.

Case 2. Now let $\sqrt[m]{x/k} < p_m \leqq x/k$. We have now several choices. First we make $p_{m-1} \leqq \sqrt[m-1]{x/kp_m}$. Then the p_i range from $P_1(k)$ to p_{i+1} for $i = 1, 2, \cdots, m-2$.

Case 3. Here $\sqrt[m]{x/k} < p_m \leqq x/k$ and $\sqrt[m-1]{x/kp_m} < p_{m-1} \leqq p_m$. Here we make $p_{m-2} \leqq \sqrt[m-2]{x/kp_m p_{m-1}}$ so that $p_i \leqq p_{i+1}$ for $i = 1, 2, \cdots, m-3$.

...

...

General case. We have $\sqrt[m]{x/k} < p_m \leqq x/k$, $\sqrt[m-1]{x/kp_m} < p_{m-1} \leqq p_m$, $\sqrt[m-2]{x/kp_m p_{m-1}} < p_{m-2} \leqq p_{m-1} \cdots \sqrt[i+1]{x/kp_m \cdots p_{i+2}} < p_{i+1} \leqq p_{i+2}$ with $p_i \leqq \sqrt[i]{x/kp_m \cdots p_{i+1}}$ so that $p_{i-1} \leqq p_i$, $p_{i-2} \leqq p_{i-1}$, $\cdots$, $p_2 \leqq p_3$ and $p_1 \leqq p_2$.

$\cdots$ etc.

So we have a total of m cases to consider. We sum these over $k \leqq x^\varepsilon$, $\varepsilon = 1/m(m+1)$ and one can check that the contribution of $P_1(k)$ to each summation is negligible and so we omit writing it. We elaborate this below.

$$(S_m) \qquad \sum_{k<x^\varepsilon} \sum_{p_m \leqq \sqrt[m]{x/k}} \sum_{p_{m-1} \leqq p_m} \cdots \sum_{p_2 \leqq p_3} \sum_{p_1 \leqq p_2} p_1$$

$$(S_{m-1}) \qquad \sum_{k<x^\varepsilon} \sum_{\sqrt[m]{x/k} < p_m \leqq x/k} \sum_{p_{m-1} \leqq \sqrt[m-1]{x/kp_m}} \sum_{p_{m-2} \leqq p_{m-1}} \cdots \sum_{p_2 \leqq p_3} \sum_{p_1 \leqq p_2} p_1$$

$$(S_{m-2}) \qquad \sum_{k<x^\varepsilon} \sum_{\sqrt[m]{x/k} < p_m \leqq x/k} \sum_{\sqrt[m-1]{x/kp_m} < p_{m-1} \leqq p_m} \sum_{p_{m-2} \leqq \sqrt[m-2]{x/kp_m p_{m-1}}} \sum_{p_{m-3} \leqq p_{m-2}} \cdots \sum_{p_2 \leqq p_3} \sum_{p_1 \leqq p_2} p_1$$

...

...

...

General term.

$$(S_i) \qquad \sum_{k<x^\varepsilon} \sum_{\sqrt[m]{x/k} < p_m \leqq x/k} \sum_{\sqrt[m-1]{x/kp_m} < p_{m-1} \leqq p_m} \sum_{\sqrt[m-2]{x/kp_m p_{m-1}} < p_{m-2} \leqq p_{m-1}} \cdots$$

$$\sum_{\sqrt[i+1]{x/kp_m \cdots p_{i+2}} < p_{i+1} \leqq p_{i+2}} \sum_{p_i \leqq \sqrt[i]{x/kp_m \cdots p_{i+1}}} \sum_{p_{i-1} \leqq p_i} \cdots \sum_{p_2 \leqq p_3} \sum_{p_1 \leqq p_2} p_1$$

...

...

Last term.

$$(S_1) \qquad \sum_{k < x^\varepsilon} \sum_{\sqrt[m]{x/k} < p_m \leqq x/k} \sum_{\sqrt[m-1]{x/kp_m} < p_{m-1} \leqq p_m} \sum_{\sqrt[m-2]{x/kp_m p_{m-1}} < p_{m-2} \leqq p_{m-1}} \cdots \sum_{\sqrt{x/kp_m \cdots p_3} < p_2 \leqq p_3} \sum_{p_1 \leqq x/kp_m \cdots p_2} p_1 \,.$$

We shall first obtain upper bound estimates for each (S_i). Our process will indicate how the terms grow and establishing the upper bound first makes explanation later simpler when we take up asymptotic estimates and need upper bound estimates for errors.

We know from Lemma 1.4 that $\sum_{p_1 \leqq p_2} p_1 = O(p_2^2/\log p_2)$. Now another application of Lemma 1.4 gives

$$\sum_{p_2 \leqq p_3} \sum_{p_1 \leqq p_2} p_1 = O\left(\frac{p_3^3}{\log^2 p_3}\right) .$$

Thus taking the first i summations in (S_i) gives a term

$$(1.6) \qquad O\left(\frac{x^{1+(1/i)}}{(kp_m \cdots p_{i+1})^{1+(1/i)} L(i,\, i,\, x/kp_m \cdots p_{i+1})}\right)$$
$$\text{where } L(i,\, j,\, x) = (\log \sqrt[i]{x})^j$$

We have to sum the term above over the variable p_{i+1} in the range $\sqrt[i+1]{x/kp_m \cdots p_{i+2}} < p_{i+1} \leqq p_{i+2}$. This is certainly less than if it is summed in the interval $\sqrt[i+1]{x/p_m \cdots p_{i+2}} < p_{i+1}$.

We are going to apply Lemma 1.3 with $z = x/kp_m \cdots p_{i+2}$ and x in lemma replaced by $\sqrt[i+1]{x/kp_m \cdots p_{i+2}}$ which we will denote for the moment by X. We can also assume $z > X \log^\delta X$, where $\delta > 0$ is a suitably chosen large positive constant so that Lemma 1.3 is applicable. For if $z < X \log^\delta X$ then we infer that

$$p_1 \leqq \frac{n}{kp_m \cdots p_{i+2}} \leqq \frac{x}{kp_m \cdots p_{i+2}} = O(\log^{\delta'} x)$$

so that the sum of $P_m(n)$ for $n \leqq x$ over n satisfying the above inequality is $O(x \log^{\delta'} x)$ which is certainly of lower order of magnitude compared to the leading term mentioned in the theorem. We shall meet this situation as we move left along each summation and so we assume that $p_1 > \log^\delta x$ for some $\delta > 0$, fixed and large, say $> m^3$. Now we apply Lemma 1.3 to infer from (1.6) that we get a term of size

$$(1.7)\qquad O\left(\frac{x^{1+(1/i+1)}}{(kp_m\cdots p_{i+2})^{1+(1/i+1)}L(i+1,\,i+1,\,x/kp_m\cdots p_{i+2})}\right).$$

Note that the term in (1.7) is just the term in (1.8) with i replaced by $i+1$. Thus making the first m summations of (S_i) gives

$$(1.8)\qquad O\left(\frac{x^{1+(1/m)}}{k^{1+(1/m)}L(m,\,m,\,x/k)}\right).$$

Obviously (1.8) summed over k gives $O(x^{1+(1/m)}/\log^m x)$.

Now we proceed to the asymptotic estimate. We shall see that the leading terms we get are exactly those mentioned above. But the error terms can be estimated just as we got upper bound estimates but there will be an extra factor of $\log x$ in the denominator, giving a sum of lower order compared to the leading term.

So by Lemma 1.4,

$$(1.9)\qquad \sum_{p_1\leqq p_2} p_1 = \frac{p_2^2}{2\log p_2} + O\left(\frac{p_2^2}{\log^2 p_2}\right).$$

Summing the term in (1.9) up to p_3, we have by Lemma 1.4

$$(1.10)\qquad \sum_{p_2\leqq p_3}\sum_{p_1\leqq p_2} p_1 = \frac{p_3^3}{6\log^2 p_3} + O\left(\frac{p_3^3}{\log^3 p_3}\right).$$

So it is now clear that making the first i summations as we did in (1.9) and (1.10) above, by repeating application of Lemma 1.4 we get

$$(1.11)\qquad \begin{aligned}&\frac{x^{1+(1/i)}}{(i+1)!\,(kp_m\cdots p_{i+1})^{1+(1/i)}L(i,\,i,\,x/kp_m\cdots p_{i+1})}\\ &\qquad + O\left(\frac{x^{1+(1/i)}\log\log x}{(kp_m\cdots p_{i+1})^{1+(1/i)}L(i,\,i+1,\,x/kp_m\cdots p_{i+1})}\right)\end{aligned}$$

Now the O-term in (1.11) has an extra factor of log in the denominator, compared to the leading term. So summing this the way we did (1.6) up to (1.8) we get a term or order $O(x^{1+(1/m)}\cdot\log\log x/\log^{m+1}x)$. So we can forget the error term in (1.11).

Now each summation after the ith summation in (S_i) is of the form $\sum_{A<p_j\leqq B}$ which we will interpret as $\sum_{A<p_j} - \sum_{B<p_j}$. There is no harm in writing it in this way, for each sum is actually a finite one because the p_j's occur in the denominator in the ith summation. Now we apply Lemma 1.3 to estimate the sum of the leading term in (1.11) over the $(i+1)$th summation. We have

$$\sum_{\sqrt[i+1]{x/kp_m\cdots p_{i+2}}<p_{i+1}\leq p_{i+2}} \frac{x^{1+(1/i)}}{(i+1)!\,(kp_m\cdots p_{i+1})^{1+(1/i)}L(i,\,i,\,x/kp_m\cdots p_{i+1})}$$

$$(1.12)\quad \begin{aligned} &= \frac{ix^{1+(1/i+1)}}{(i+1)!\,(kp_m\cdots p_{i+2})^{1+(1/i+1)}L(i+1,\,i+1,\,x/kp_m\cdots p_{i+2})} \\ &\quad + O\Big(\frac{x^{1+(1/i+1)}\cdot\log\log x}{(kp_m\cdots p_{i+2})^{1+(1/i+1)}L(i+1,\,i+2,\,x/kp_m\cdots p_{i+2})}\Big) \\ &\quad + \frac{(i)x^{1+(1/i)}}{(i+1)!\,(kp_m\cdots p_{i+3})^{1+(1/i)}p_{i+2}^{1+(2/i)}L(i+1,\,i+1,\,x/kp_m\cdots p_{i+2})}\,. \end{aligned}$$

Equation (1.12) needs some explanation. The first two terms on the right are obtained by considering $\sum_{A<p_j}$. As regards $\sum_{B<p_j}$ we distinguish two cases. The first is when $B>A\log^{\delta}A$ (δ sufficiently large, say $>m^2$). Now by Lemma 1.3 this sum is small compared to the former and there is no harm in writing it in the form of the third term on the right in (1.12). If $A<B<A\log^{\delta}A$, then the log term does not change appreciably and again Lemma 1.3 gives the third term on the right of (1.12) as the leading term with the error being absorbed in the O-term in (1.12). Note that the O-term in (1.12) again has an extra factor of log in the denominator which as mentined before is pulled through to give an error term $O(x^{1+1/m}\log\log x/\log^{m+1}x)$. So what we are essentially saying is that we can forget the error terms totally since (1.12) is the type of estimate we will meet as we proceed left along (S_i). As regards the leading terms, they will be of the form of the first term in (1.12) or the last term, depending whether we choose the left side bound which we call A, or the right side bound which we call B in each summation. However in the summation involving p_m, we have to take $\sum_{\sqrt[m]{x/k}<p_m<\infty}$ because $\sum_{x/k<p_m<\infty}$ is a summation over the null set since $kp_1\cdots p_m\leqq x$. So in the first j_1 summations from the ith one of (S_i), we choose the left limit, and in the next i_1 we choose the right one, and in the next j_2 we choose the left one and so on. Then the sign of the total summation is $(-1)^{i_1+i_2+\cdots}$. Note that j_1 could be zero. We elaborate this below and this is our final step. The vertical lines in (1.13) tell us where the changes in limit takes place, and the arrow indicates the first step where we change.

$$(1.13)\quad \begin{aligned} &\sum_{k<x^{\theta}}\ \sum_{\sqrt[m]{x/k}<p_m<\infty}\cdots\Big|\sum_{p_{i+j_1+i_1+1}<p_{i+j_1+i_1}<\infty}\cdots\sum_{p_{i+j_1+2}<p_{i+j_1+1}<\infty}\longleftarrow \\ &\Big|\sum_{\sqrt[i+j_1]{x/kp_m\cdots p_{i+j_1+1}}<p_{i+j_1}<\infty}\sum\cdots \\ &\qquad\sum_{\sqrt[i+1]{x/kp_m\cdots p_{i+2}}<p_{i+1}<\infty}\frac{x^{1+(1/i)}}{(i+1)!(kp_m\cdots p_{i+1})^{1+(1/i)}L(i,\,i,\,x/kp_m\cdots p_{i+1})} \end{aligned}$$

The first summation in (1.13) gives the first term on the right of

(1.12). What we are summing in (1.13) is the term in (1.11). In the process of going from (1.11) to (1.12) note that what has happened is that i has been replaced by $i+1$ for the variables and there is an extra factor of i. So making the first j_1 summations we get a term which is

$$(1.14)\quad \frac{i(i+1)\cdots(i+j_1-1)x^{1+(1/i+j_1)}}{(i+1)!(kp_m\cdots p_{i+j_1+1})^{1+(1/i+j_1)}L(i+j_1,\, i+j_1,\, x/kp_m\cdots p_{j+j_1+1})}\,.$$

We have to sum the term in (1.14) over the variable p_{i+j_1+1} in the summation indicated by the arrow above. Now by Lemma 1.3 we get a term of the type of the third term in (1.12) with i replaced by $i+j_1$. So we have

$$(1.15)\quad \frac{i(i+1)\cdots(i+j_1)x^{1+(1/i+j_1)}}{\begin{pmatrix}(i+1)!\,(kp_m\cdots p_{i+j_1+3})^{1+(1/i+j_1)}(p_{i+j_1+2})^{1+(2/i+j_1)}\\ \times L(i+j_1+1,\, i+j_1+1,\, x/kp_m\cdots p_{i+j_1+2})\end{pmatrix}}$$

The only thing we have to observe in (1.15) is that the exponent of p_{i+j_1+2} is $1+(2/i+j_1)$, and the exponent of x has not changed from (1.14) to (1.15). This affects the nature of the constant to appear in the numerator of (1.16) below. What we get after the next summation is

$$(1.16)\quad \frac{\left(\dfrac{i+j_1}{2}\right)(i)(i+1)\cdots(i+j_1)x^{1+(1/i+j_1)}}{\begin{pmatrix}(i+1)!\,(kp_m\cdots p_{i+j_1+4})^{1+(1/i+j_1)}(p_{i+j_1+3})^{1+(3/i+j_1)}\\ \times L(i+j_1+2,\, i+j_1+2,\, x/kp_m\cdots p_{i+j_1+3})\end{pmatrix}}\,.$$

Now (1.16) is the term in (1.15) with subscripts changed by 1 and change of constants. So going to the end of j_1+i_1 summations we get

$$(1.17)\quad \frac{\left(\dfrac{i+j_1}{2}\right)\cdot\left(\dfrac{i+j_1}{3}\right)\cdots\left(\dfrac{i+j_1}{i_1}\right)(i)(i+1)\cdots(i+j_1)x^{1+(1/i+j_1)}}{\begin{pmatrix}(i+1)!\,(kp_m\cdots p_{i+j_1+i_1+2})^{1+(1/i+j_1)}(p_{i+j_1+1})^{1+(i_1+1/i+j_1)}\\ \times L(i+j_1+i_1,\, i+j_1+i_1,\, x/kp_m\cdots p_{i+j_1+i_1+1})\end{pmatrix}}\,.$$

Now when we sum (1.17) we are doing it in the range $A < p < \infty$, where A is a left limit. If we show that this summation leads to a term similar to the one with which we started in (1.13) we are done. It is indeed remarkable that this happens. For again by Lemma 1.3 if we observe that

$$1+\frac{1}{i+j_1}-\frac{i_1+1}{(i+j_1)(i+j_1+i_1+1)}=1+\frac{1}{i+j_1+i_1+1}$$

we find that the exponent of x which had remained constant for these i_1 summations changes suitably to give a term

$$\frac{(\text{constant})\cdot x^{1+1/(i+j_1+i_1+1)}}{\left((kp_m \cdots p_{i+j_1+i_1+2})^{1+1/(i+j_1+i_1+1)} \times L(i+j_1+i_1+1, i+j_1+i_1+1, x/kp_m \cdots p_{i+j_1+i_1+2})\right)} . \tag{1.18}$$

Now the term in (1.18) is just the term in (1.13) with i replaced by $i + j_1 + i_1 + 1$. So after $i_1 + j_1 + 1$ steps we are back in the same situation. So everytime we choose a left bound in a summation, we are back to the situation with which we started. But in the last summation involuing p_m, we have to choose the left bound. So ultimately we get

$$\frac{c_0 x^{1+(1/m)}}{k^{1+(1/m)} \log^m (x/k)}$$

where c_0 is rational. Summing this over $k < x^\varepsilon$, using a method similar to σ_1 and σ_2 in Lemma 1.3, gives

$$\frac{c_0\zeta(1+(1/m))x^{1+(1/m)}}{(\log x)^m} .$$

Of course this is just one of the subcases of (S_i). Considering all the subcases of (S_i) we get c_i^* rational and

$$\frac{c_i^* \zeta\left(1+\dfrac{1}{m}\right)x^{1+(1/m)}}{(\log x)^m} .$$

Since the summations involve positive quantities we infer $c_i^* > 0$. So summing over all the i from 1 up to m, gives a positive rational c_m so that the contribution from (S_1) up to (S_m) is

$$\frac{c_m x^{1+(1/m)} \zeta\left(1+\dfrac{1}{m}\right)}{(\log x)^m} . \tag{1.19}$$

So this is the contribution for $k < x^\varepsilon$, $\varepsilon > 0$. If $k > x^\varepsilon$ then

$$\sum_{\substack{1 \leq n \leq x \\ k > x^\varepsilon}} P_m(n) = O\left(\sum_{n \leq x} n^{(1-\varepsilon)/m}\right) = O\left(\frac{x^{1+(1/m)}}{x^{\varepsilon/m}}\right) . \tag{1.20}$$

So (1.20) and (1.19) yield

$$\sum_{n \leq x} P_m(n) = \frac{c_m \zeta\left(1+\dfrac{1}{m}\right)x^{1+(1/m)}}{\log^m x} + O\left(\frac{x^{1+(1/m)} \log\log x}{\log^{m+1} x}\right) \tag{1.21}$$

as shown by our investigation of error terms. Our theorem will be proved if we show that

$$\sum_{n\leq x}\{A(n)-P_1(n)-P_2(n)-\cdots-P_m(n)\}=o\Big(\sum_{n\leq x}P_m(n)\Big)\,.$$

Observe that $\{A(n)-P_1(n)-P_2(n)-\cdots-P_m(n)\}=A(P_m^*(n))=A(k)$ and $A(k)\leqq\Omega(k)P_1(k)=\Omega(k)P_{m+1}(n)\leqq n^{1/(m+1)}\cdot\Omega(n)$. So

$$\begin{aligned}\sum_{n\leq x}\{A(n)-P_1(n)-P_2(n)-\cdots-P_m(n)\}&\leqq\sum_{n\leq x}n^{1/(m+1)}\Omega(n)\\&=O\Big(x^{1/(m+1)}\sum_{n\leq x}\Omega(n)\Big)\\&=O(x^{1+(1/(m+1)}\log\log x)\\&=o\Big(\sum_{n\leq x}P_m(n)\Big)\end{aligned}$$

because of (1.21). The proof of Theorem 1.1 is complete.

COROLLARY. *The average order of $A(n)$ is $\pi^2n/6\log n$.*

Proof. Set $m=1$ in Theorem 1.1. Then there is only one case to consider, namely $(S_m)=(S_1)$. So

$$\sum_{n\leq x}A(n)\sim\sum_{n\leq x}P_1(n)\sim\frac{\pi^2x^2}{12\log x}$$

which gives the corollary.

It is clear that $A(n)\geqq A^*(n)\geqq P_1(n)$ so that $A^*(n)$ also has average order $\pi^2n/6\log n$.

THEOREM 1.5. *The average order of $A(n)-A^*(n)$ is $\log\log n$. To be more precise*

$$\sum_{n\leq x}\{A(n)-A^*(n)\}=x\log\log x+O(x)\,.$$

Proof. It is not difficult to see that

$$\sum_{n\leq x}\{A(n)-A^*(n)\}=\sum_{p^2\leq x}p[x/p^2]+\sum_{p^3\leq x}p[x/p^3]+\cdots.$$

For if we write $A(n)-A^*(n)=\sum_{p^\alpha||n}(\alpha-1)p$, then p is counted $[x/p^2]$ more times, giving the first term. If $(\alpha-1)p=2p$ we count only $[x/p^3]$ more times and so on. Now

$$\begin{aligned}\sum_{p^2\leq x}p\frac{x}{p^2}&=\sum_{p\leq\sqrt{x}}\frac{x}{p}+O\Big(\sum_{p\leq\sqrt{x}}p\Big)\\&=x\log\log x+O(x)\end{aligned}$$

and

$$\sum_{p^i \leq x} p \frac{x}{p^i} = x \sum_{p^i \leq x} \frac{1}{p^{i-1}} = O\Big(\sum_{p \leq \sqrt[i]{x}} p \Big) .$$

So

$$\sum_{i \geq 3} \sum p \frac{x}{p^i} = O(x)$$

which proves Theorem 1.5.

THEOREM 1.6. *We have*

$$\sum_{n \leq x} \{A^*(n) - P_1(n) - P_2(n) - \cdots - P_{m-1}(n)\} \sim \sum_{n \leq x} P_m(n) \sim \frac{k_m x^{1+1/m}}{(\log x)^m} .$$

Proof. The theorem follows by combining Theorems 1.1 and 1.5.

THEOREM 1.7. *For any fixed integer M, the set of solutions to $A(n) - A^*(n) = M$ has a natural density > 0.* (Note: A sequence $\{a_n\}_{n=1}^{\infty}$ has a natural density $\delta(A)$ if $\lim_{n\to\infty} n/a_n = \delta(A)$.)

Proof. Let us define an integer n to be powerful if $n = \prod_{i=1}^{r} p_i^{\alpha_i}$, $\alpha_i \geq 2$, $i = 1, 2, \cdots, r$. The set S_n of integers of the form $n \cdot n'$, $(n, n') = 1$ and n' squarefree has natural density

$$(1.23)\quad \frac{1}{n} \prod_{p|n} \Big(1 - \frac{1}{p}\Big) \cdot \prod_{q \nmid n} \Big(1 - \frac{1}{q^2}\Big) = \frac{6}{n\pi^2} \prod_{p|n} \frac{\Big(1 - \dfrac{1}{p}\Big)}{\Big(1 - \dfrac{1}{p^2}\Big)} = \frac{6}{n\pi^2} \prod_{p} \Big(1 + \frac{1}{p}\Big)^{-1}$$
$$= \delta(S_n) .$$

Consider a partition of M into primes as $M = \sum_{i=1}^{r} \beta_i p_i$. Any integer n with $A(n) - A^*(n) = M$ is of the form $\prod_{i=1}^{r} p_i^{(\beta_i+1)} \prod_{j=1}^{s} q_j$, where q_j are primes different from p_i. Consider a particular partition π_j of M, as $M = \sum_{i=1}^{r} \beta_i p_i$ and the powerful integer $m_j = \prod_{i=1}^{r} p_i^{(\beta_i+1)}$. This partition generates a set of solutions which is the set of numbers of the form $m_j \cdot m'$, $(m_j, m') = 1$, m' square free. This set denoted by S_{m_j} has natural density $\delta(S_{m_j})$. Now the complete set of solutions is given by

$$\bigcup_{j=1}^{p(M)} S_{m_j}$$

where $p(M)$ is the number of partitions of M into primes. Thus

$$\delta\left(\bigcup_{j=1}^{p(M)} S_{m_j}\right) = \sum_{j=1}^{p(M)} \delta(S_{m_j}) \tag{1.24}$$

as $S_{m_i} \cap S_{m_j} = \emptyset$ if $i \neq j$. In fact because of (1.23) and (1.24) the density is a rational multiple of $1/\zeta(2) = 6/\pi^2$.

2. Congruences involving $A(n)$. We now recall some results in [1]. For any integer m, the number of solutions to $A(n) = m$ is the number of partitions of m into primes. Note that $A(n) = n$ if and only if n is a prime or $n = 4$, so that it would be of interest to study the congruence

$$n \equiv 0 (\operatorname{mod} A(n)) . \tag{2.1}$$

Call a solution to (2.1) non-trivial if $n \neq A(n)$ and let the nontrivial solutions be called "special numbers". It is worth noting that if m is fixed then the number of solutions to

$$n \equiv 0 (\operatorname{mod} A(n)) , \quad A(n) = m \tag{2.2}$$

is the number of partitions of $m - A(m)$ into primes. So the number of solutions to (2.2) is much less than the number of solutions to $A(n) = m$, generally, and so one expects that special numbers are rather rare. Let $\{l_n\}$ denote the sequence of special numbers. The following can be proved (see [1]).

(1) The sequence $\{l_n\}$ is infinite.

(2) $\lim_{n\to\infty} A(l_n)/l_n = 0$

(3) For any pair of integers a and b, the number of solutions to $l_n \equiv a (\operatorname{mod} b)$ is infinite.

(4) If $\pi(x, 2)$ represents the number of twin primes $\leqq x$ and $\pi(x, 2) \sim cx/\log^2 x$ then $\lim_{n\to\infty} l_n/l_{n+1} = 1$.

Denote by $\mathscr{L}(x)$ the number of $l_n \leqq x$. We obtain bounds for $\mathscr{L}(x)$.

THEOREM 2.1. *There exists a contant $c > 0$ so that for all $x \geqq e$*

$$\mathscr{L}(x) = O\left(\frac{x}{e^{c\sqrt{\log x \log\log x}}}\right) .$$

Proof. As before $P_1(n)$ denotes the largest prime factor of n. By a result of deBruijn (see [2], page 54, equation 1.6), if $\psi(x, y)$ is the number of solutions $\leqq x$ to $P_1(n) < y$ then

$$\psi(x, y) < c_1 x \log^2 y e^{-u\log u - \log\log u + c_2 u} \tag{2.3}$$

where $y = x^{1/u}$. Now if we set $u = \sqrt{\log x/\log\log x}$, then y is seen to be $e^{\sqrt{\log x \log\log x}}$. Also

$$\psi(x, y) < \frac{c_1 x \log x \log\log x}{e^{1/2\sqrt{\log x \log\log x} - c_3 u}} \tag{2.4}$$
$$= O\left(\frac{x}{e^{1/2\sqrt{\log x \log\log x}}}\right).$$

So we will restrict our attention to $P_1(n) > e^{\sqrt{\log x \log\log x}}$ for the number of n not satisfying this is given by (2.4). We also assume that if $\tau(n)$ is the number of divisors of n then

$$\tau(n) < e^{1/2\sqrt{\log x \log\log x}}. \tag{2.5}$$

For the number of integers not satisfying (2.5) is easily seen to be

$$O\left(\frac{x \log x}{e^{1/2\sqrt{\log x \log\log x}}}\right) \tag{2.6}$$

because $\sum \tau(n) = O(x \log x)$. So we confine ourselves to $n \leqq x$ satisfying (2.5) and $P_1(n) > e^{\sqrt{\log x \log\log x}}$. Let these numbers be denoted by the sequence $\{n_i\}$. Denote by t the following

$$\frac{n_i}{P_1(n_i)} = t \Longrightarrow A(n_i) = P_1(n_i) + A(t). \tag{2.7}$$

Clearly as $n_i \leqq x$ we have

$$t < xe^{-\sqrt{\log x \log\log x}}. \tag{2.8}$$

Let t for the moment be fixed. We have two possibilities arising out of (2.7).

Case 1. $A(t) \equiv 0 (\operatorname{mod} P_1(n_i))$.

Since t is fixed and we are seeking solutions to (2.7) it is clear that the $P_1(n_i)$ are distinct and divide $A(t)$. Also as we require special numbers, $t \neq 1$ and so $A(t) \neq 0$. Thus the number of solutions to Case 1 is at most $O(\log x)$, since $t \leqq x$.

Case 2. $A(t) \not\equiv 0 (\operatorname{mod} P_1(n_i))$.

Since we are interested in special numbers we require

$$n_i \equiv 0 (\operatorname{mod} A(n_i)) \equiv o(\operatorname{mod} P_1(n_i) + A(t)). \tag{2.9}$$

But Case 2 implies that $(A(t), P_1(n_i)) = 1$ which means (2.9) gives

$$\frac{n_i}{P_1(n_i)} = t \equiv 0 (\operatorname{mod} P_1(n_i) + A(t)). \tag{2.10}$$

Again $A(t) + P_1(n_i)$ are distinct when t is fixed, so that by our choice of n_i, by (2.5) the number of solutions to (2.10) is less than

$$e^{1/2\sqrt{\log x \log\log x}}.$$

Thus for fixed t, the number of solutions to (2.7) in special numbers is at most

$$\log x + e^{1/2\sqrt{\log x \log\log x}} = O(e^{1/2\sqrt{\log x \log\log x}}).$$

But by (2.8) we have an upper bound on the number of choices of t. Thus the $\{l_n\}$ among the n_i do not exceed

$$O\left(\frac{x}{e^{\sqrt{\log x \log\log x}}} e^{1/2\sqrt{\log x \log\log x}}\right) = O\left(\frac{x}{e^{1/2\sqrt{\log x \log\log x}}}\right). \tag{2.11}$$

But the number of integers not included among the $\{n_i\}$ is by (2.6) and (2.4)

$$O\left(\frac{x \log x}{e^{1/2\sqrt{\log x \log\log x}}}\right). \tag{2.12}$$

So (2.12) and (2.11) prove the theorem with any $c < 1/2$.

Now for a lower bound,

THEOREM 2.2. *There exists a constant $c' > 0$ so that*

$$\mathscr{L}(x) \gg \frac{x}{e^{c'\sqrt{\log x \log\log x}}}.$$

Proof. Let x be a large real number and define z and k as follows:

$$z = \frac{e^{\sqrt{\log x \log\log x}}}{(\log x)^{c_4}}; \quad k = \sqrt{\frac{\log x}{\log\log x}}$$

where $c_4 > 0$ is a constant to be determined soon. Consider the number of k-tuples of primes $\leqq z$ which clearly is $\binom{\pi(z)}{k}$. This can be easily seen to be greater than

$$\begin{aligned}\frac{(\pi(z)-k)^k}{k!} &> \frac{(\pi(z)-k)^k}{e^{k\log k}} > \left(\frac{e^{\sqrt{\log x \log\log x}}}{(\log x)^{c_4+\varepsilon}\sqrt{\log x \log\log x}}\right)^k \cdot \frac{1}{e^{k\log k}} \\ &> \frac{x}{e^{(c_4+1+\varepsilon)\sqrt{\log x \log\log x}}}\end{aligned} \tag{2.14}$$

for sufficiently large x. Now let the product of these primes define a sequence $\{u_j\}$, all $\leqq z^k$ which is seen to be

$$z^k = \frac{x}{e^{c_4\sqrt{\log x \log\log x}}}. \tag{2.15}$$

Let us put $c_5 = c_4 + 1 + \varepsilon$, $\varepsilon > 0$ arbitrary. So we have at least $xe^{-c_5\sqrt{\log x \log\log x}}$ distinct numbers $< z^k$ given in (2.15). Consider the product of the first r-primes $p_1 \cdots p_r$ such that it is just greater than zk^2. We shall produce a number L so that

$$A(u_j p_1 \cdots p_r \cdot L) = p_1 \cdots p_r$$

so that $u_j \cdot p_1 \cdots p_r \cdot L$ is a special number. Clearly we need

$$(2.16) \qquad A(L) = p_1 \cdots p_r - (p_1 + p_2 + \cdots + p_r) - A(u_j) .$$

By our choice of u_j, $A(u_j) \leqq zk$ and so the quantity in (2.16) is of order zk^2. If it is odd, use Vinogradov's theorem, and partition it into three primes, and take their product to get L. Otherwise subtract L and partition the rest into three primes and L is the product of these primes. In any case $L = O((zk^2)^3)$ and so $p_2 \cdots p_r \cdot L = O((zk^2)^5)$. Now we want $u_j \cdot p_1 \cdots p_r \cdot L \leqq x$. So choose $c_4 > 5$ so that by (2.15) the product $u_j \cdot p_1 \cdots p_r \cdot L \leqq x$. Now the number of repetitions among $u_j \cdot p_1 \cdots p_r \cdot L$ is at most $(zk^2)^4$. So they are at least

$$\frac{x}{e^{c_5\sqrt{\log x \log\log x}}} \cdot \frac{1}{(zk^2)^4} \gg \frac{x}{e^{(10+\varepsilon)\sqrt{\log x \log\log x}}}$$

special numbers $\leqq x$, by our choice of c_4 and z. So Theorem 2.2 holds with any $c'' > 10 + \varepsilon$.

REMARK. The problem discussed in this section can be worded differently. "How often can a sum of primes (not necessarily distinct) divide their product?" That is we want $\sum \alpha_i p_i$ to divide $\prod p_i^{\alpha_i}$ where each p_i has α_i repetitions. This is precisely of the problem of special numbers discussed above.

It might be true that $\mathscr{L}(x)$ is actually of the order

$$xe^{-c''(1+o(1))\sqrt{\log x \log\log x}} .$$

An asymptotic formula for $\mathscr{L}(x)$ seems very hard to obtain. The constants in Theorems 2.1 and 2.2 can be sharpened with more accurate computation, but our estimates indicate the method.

We conclude this section with a few interesting questions. Does the product of consecutive primes infinitely often determine special numbers? For instance $2\cdot3\cdot5 \equiv 0 (\bmod\, 2+3+5)$. Also $n = 2\cdot3\cdot5\cdot7\cdot11\cdot13\cdot17\cdot19$ is special. $A(n)$ is $77 = 7\cdot11$. Another example is $n = 2\cdot3\cdot5\cdot7\cdot11\cdot13\cdot17\cdot19\cdot23\cdot29\cdot31\cdot37\cdot41$ where $A(n) = 238 = 2\cdot7\cdot17$. We guess there are infinitely many such numbers!

It is easy to see that infinitely many special numbers are square free. For, take a prime p, and partition $p - 2$ into distinct primes $p_1 + p_2 + \cdots + p_r$. This is possible. Then $2\cdot p\cdot p_1\cdot p_2\cdot \cdots \cdot p_r$ is special for $A(2\cdot p\cdot p_1\cdot p_2\cdot \cdots \cdot p_r) = 2p$.

One can show that for sufficiently large composite numbers n, there exists m with $m \equiv 0 (\text{mod } A(m))$, $A(m) = n$ and m/n square free and prime to n. This follows from Vinogradov's theorem, and here we partition $n - A(n)$ into primes. It might be of interest to determine (besides the primes), all the other n for which this is not true.

3. **Distribution modulo 2.** First we shall show that $A(n)$ is uniformly distributed modulo 2, and the error is of the order of the sum of the Möbius function $M(x)$. Here we shall concentrate on the function $\alpha(n) = (-1)^{A(n)}$, which is easily seen to obey $\alpha(m\cdot n) = \alpha(m)\cdot\alpha(n) \forall m, n$. Thus for any complex number s, with $\operatorname{Re} s > 1$, we have

$$\sum_{n=1}^{\infty} \frac{\alpha(n)}{n^s} = \prod_p \left(1 - \frac{\alpha(p)}{p^s}\right)^{-1} = \frac{2^s + 1}{2^s - 1} \cdot \frac{\zeta(2s)}{\zeta(s)} . \tag{3.1}$$

Now as $s \to 1^+$, the right side of (3.1) tends to zero, and so it is natural to expect

$$\sum_{n=1}^{\infty} \frac{\alpha(n)}{n} = 0 . \tag{3.2}$$

We prove (3.2) in Theorem 3.2. But first we show that $A(n)$ is uniformly distributed modulo 2. This is expressed in

THEOREM 3.1. *There exists a constant $c_6 > 0$ so that*

$$\sum_{1 \leq n \leq x} \alpha(n) = O(xe^{-c_6\sqrt{\log x \log\log x}}) .$$

Proof. Consider the sum $a(n) = \sum_{d|n} \alpha(d)$. If $n = 2^{\alpha_0} \prod_{i=1}^{r} p_i^{\alpha_i}$ where p_i are odd, then

$$\begin{aligned} a(n) &= (\alpha_0 + 1) \prod_{i=1}^{r} (1 + \alpha(p_i) + \alpha(p_i^2) + \cdots + \alpha(p_i^{\alpha_i})) \\ &= (\alpha_0 + 1) \prod_{i=1}^{r} \frac{(-1)^{(\alpha_i+1)p_i} - 1}{(-)1^{p_i} - 1} . \end{aligned} \tag{3.3}$$

We infer from (3.3) that if any one of the α_i is odd, then $a(n) = 0$. Thus $a(n)$ is non-zero only over integers of the form $2^{\alpha_0}\cdot m^2$ where m is odd. Also $a(n) \geqq 0$. Clearly

$$\sum_{n \leqq x} a(n) \leqq \sum_{2^{\alpha_0} \leqq x} (\alpha_0 + 1)\sqrt{x/2^{\alpha_0}} = O(\sqrt{x}) \tag{3.4}$$

and

$$\sum_{n=1}^{\infty} \frac{a(n)}{n} = c_7 < \infty \; . \tag{3.5}$$

Now if μ is the Möbius function then

$$\begin{aligned} \sum_{1 \leq n \leq x} \alpha(n) &= \sum_{1 \leq n \leq x} \sum_{d|n} \mu(d) a\left(\frac{n}{d}\right) \\ &= \sum_{1 \leq d \leq \sqrt{x}} a(d) \sum_{1 \leq d' \leq x/d} \mu(d') + \sum_{1 \leq d < \sqrt{x}} \mu(d) \sum_{\sqrt{x} < d' \leq x'/d} a(d') \\ &= \sum_{1 \leq d \leq \sqrt{x}} a(d) M\left(\frac{x}{d}\right) + \sum_{1 \leq d \leq \sqrt{x}} O(\sqrt{x/d}) \\ &= \sum_{1 \leq d \leq \sqrt{x}} a(d) M\left(\frac{x}{d}\right) + O(x^{3/4}) \end{aligned} \tag{3.6}$$

by using (3.4). It is known from the investigation of the error term in the prime number theorem (see [3]) that there is a constant $c_8 > 0$ so that

$$M(x) = O(xe^{-c_8\sqrt{\log x \log\log x}}) \tag{3.7}$$

so that one inferes from (3.6), (3.7), and (3.5) that Theorem 3.1 is true.

Finally we prove

THEOREM 3.2. $\sum_{n=1}^{\infty} \alpha(n)/n = 0$.

Proof. As we have already remarked, $a(n)$ is nonzero only at values $n = 2^{\alpha_0} \cdot m^2$, where m is odd, and $a(n)$ here is $\alpha_0 + 1$. Thus

$$\begin{aligned} \sum_{n > x} \frac{a(n)}{n} &= \sum_{\alpha_0 = 0}^{\infty} \sum_{\substack{m > \sqrt{x/2^{\alpha_0}} \\ m \text{ odd}}} \frac{\alpha_0 + 1}{2^{\alpha_0} \cdot m^2} \\ &= \sum_{\alpha_0 = 0}^{\infty} \frac{\alpha_0 + 1}{2^{\alpha_0}} \cdot O(\sqrt{2^{\alpha_0}/x}) = O(1/\sqrt{x}) \end{aligned} \tag{3.8}$$

so that if we set

$$\chi(x) = x\left(\sum_{n > x} \frac{a(n)}{n}\right)$$

we infer from (3.8) that

$$\chi(x) = O(x^{1/2}) \; . \tag{3.9}$$

Also χ is of bounded variation on finite intervals. It follows that

$$\begin{aligned} x \sum_{d \leq x} \frac{\alpha(d)}{d} &= x \sum_{nm \leq x} \frac{\mu(m) a(n)}{mn} \\ &= x \sum_{m \leq x} \frac{\mu(m)}{m} \sum_{n \leq x/m} \frac{a(n)}{n} \; . \end{aligned} \tag{3.10}$$

Now by (3.5), (3.10) is rewritten as

$$x\sum_{d\leq x}\frac{\alpha(d)}{d} = c_7 x\sum_{m\leq x}\frac{\mu(m)}{m} - \sum_{m\leq x}\mu(m)\frac{x}{m}\sum_{n>x/m}\frac{a(n)}{n} \tag{3.11}$$
$$= c_7 x\sum_{m\leq x}\frac{\mu(m)}{m} - \sum_{m\leq x}\mu(m)\chi\left(\frac{x}{m}\right).$$

We can deduce Theorem 3.2 from (3.11), if we appeal to Axer's Theorem 267 in [5] stated below.

AXER'S THEOREM. *If $\{b_n\}$ is a sequence of real numbers with*

$$\sum_{n\leq x} b_n = o(x) \quad and \quad |b_n| = O(1)$$

and χ a function of bounded variation on finite intervals with $\chi(x) = O(x^\alpha)$ for some $0 < \alpha < 1$, then

$$\sum_{n\leq x} b_n\chi\left(\frac{x}{n}\right) = o(x) .$$

If we apply Axer's Theorem with $b_n = \mu(n)$, and observe that $\chi(x) = O(x^\alpha)$ with $\alpha = 1/2$ in (3.9) then because

$$\sum_{n=1}^{\infty}\frac{\mu(n)}{n} = 0 \tag{3.12}$$

we infer from (3.11) that Theorem 3.2 is true. For a proof of (3.12) see [5].

By slight variation of the proofs of the above theorems one can show that for some fixed integer N

$$\sum_{\substack{1\leq n\leq x\\(n,N)=1}}\alpha(n) = O(xe^{-c_9\sqrt{\log x\log\log x}})$$

and

$$\sum_{\substack{n=1\\(n,N)=1}}^{\infty}\frac{\alpha(n)}{n} = 0 .$$

REMARK. We would like to conclude by mentioning a few interesting problems connected with $A(n)$.

Let $f(n)$ be the smallest integer m so that $A(m) = n$. Consider a partition of n into primes, $n = p_1 + p_2 + \cdots$ where p_1 is the largest prime $\leq n$, $p_1 \neq n-1$, p_2 the largest prime $\leq n - p_1$, $p_2 \neq n - p_1 - 1$, and so on, and denote by $\mathrm{F}(n) = p_1\cdot p_2\cdot\,\cdots$. It appears at first sight that $f(n) = F(n)$ but this need not be so. In fact this does not happen quite often. For instance $f(6) = 8$, $F(6) = 9$. It would be of interest to consider the relative sizes of $f(n)$ and $F(n)$.

In this context we mention the following curious problem. Replace the primes above by squares. That is

$$G(n) = \min \quad \prod a_i^2; \; g(n) = \prod b_i^2 \quad \sum a_i^2 = n$$

where b_i^2 is the largest square $\leqq n$, and so on. It might be true that both $G(n)$ and $g(n)$ are both $< c \cdot n^2$ where c is a constant. In $G(n)$ above, we require that not more than three of the $a_i = 1$, for $3 = 1 + 1 + 1$ is the only decomposition of 3.

For more results on $A(n)$, see a forthcoming paper of Erdös and Pomerance where it is proved that the set of solutions to $A(n) = A(n+1)$ is of density zero. One could also consider equations involving $A(n)$ of similar type but these problems are in general difficult.

ACKNOWLEDGMENT. We are grateful to the referee for his comments and suggestions to improve the paper.

One of (K. A.) would like to express his gratitude to Professor E. G. Straus for the encouragement he gave and we would like to thank him for helping us in preparing the manuscript.

REFERENCES

1. K. Alladi, *A new logarithmic function*, Srinivasa Ramanujan Commemmoration Volume, Oxford Press, Madras, India, (1974) part II.

2. N. J. deBruijn, *On the number of positive integers $\leqq x$ free of prime factors $>y$*, Indag. Math., (1951), 50–60.

3. K. Chandrasekharan, *Arithmetical Functions*, Springer-Verlag, 1970, 50–65.

4. L. M. Chawla, *On a pair of Arithmetic functions*, J. Natur. Sci. and Math., **8** (1968), 263–269.

5. G. H. Hardy, *Divergent Series*, Oxford, Clarendon Press, 1949, 378–380.

6. G H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, Clarendon Press, 4-ed., 1968, 263, 354–358.

7. S. M. Kerawala, *On a pair of Arithmetic functions analogous to Chawla's pair*, J. Natur. Sci. and Math., **9** (1969), 97–103.

8. M. O. LeVan, *On the sum $F(x, r)$*, J. Natur. Sci. and Math., **11** (1971), 285–290.

9. J. MohanLal, *On the iterates of a number theoretic function*, Math. Comp., **23** (1969), 181–183, A-34-38.

10. N. J. Nicolas, *Sur-1-ordre maximum d'un element dans le groupe S_n des permutations*, Acta Arith., **14** (1958), 315–332.

Received October 29, 1975 and in revised form October 29, 1976.

UNIVERSITY OF CALIFORNIA
LOS ANGELES, CA 90024
AND
HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, HUNGARY