ON A PROBLEM OF CHOWLA AND SOME RELATED PROBLEMS

BY PAUL ERDÖS

[Communicated by Prof. G. H. HARDY]

[*Received* 21 March, *read* 26 October 1936]

Let $d(m)$ denote the number of divisors of the integer m. Chowla has conjectured that the integers for which $d(m+1) > d(m)$ have density $\frac{1}{2}$. In this paper I prove and generalize this conjecture. I prove in § 1 a corresponding result for a general class of functions $f(m)$, and in § 2 the result for $d(m)$ which is not included among the $f(m)$. I employ the method used in my paper: "On the density of some sequences of numbers." *

1. The functions $f(m)$ and $\phi(m)$ are called additive and multiplicative respectively if they are defined for non-negative integers m, and if, for $(m_1, m_2) = 1$,

$$f(m_1 m_2) = f(m_1) + f(m_2),$$
$$\phi(m_1 m_2) = \phi(m_1)\,\phi(m_2).$$

We suppose throughout that $f(m) \geqslant 0$, $\phi(m) \geqslant 1$.

If $\phi(m)$ is multiplicative, $\log \phi(m)$ is evidently additive, so that it will suffice to consider additive functions only.

We denote by $G(f, n)$ the number of integers $m \leqslant n$ for which $f(m+1) \geqslant f(m)$, and by $S(f, n)$ the number for which $f(m+1) \leqslant f(m)$. We suppose throughout that n is a sufficiently large integer and that the c's are absolute constants.

First we prove the following

THEOREM: *Let the additive function $f(m) \geqslant 0$ satisfy the following condition: $\sum \frac{f(p)}{p}$ converges when the summation is extended to all primes p. Then*

$$\lim_{n\to\infty} \frac{G(f,n)}{n} = \tfrac{1}{2}, \tag{1}$$

$$\lim_{n\to\infty} \frac{S(f,n)}{n} = \tfrac{1}{2}. \tag{2}$$

We prove that $\lim_{n\to\infty} \frac{G(f,n)}{S(f,n)} = 1$, and that the number of integers $m \leqslant n$ for which $f(m+1) = f(m)$ is $o(n)$, i.e. the number of integers belonging both to the set G and to the set S is $o(n)$.

* *Journal London Math. Soc.* 10 (1935), 120–125.

The method will be more intelligible if we consider first the special case in which $f(p^\alpha)=f(p)$ for any integral exponent α, so that

$$f(m)=\sum_{p\mid m} f(p).$$

Consider also the function $f_k(m)=\sum_{\substack{p\mid m\\ p\leqslant p_k}} f(p)$,

where p_k denotes the kth prime.

We show first that $$\lim_{n\to\infty}\frac{G(f_k,n)}{S(f_k,n)}=1.$$

Let us denote by $a_1, a_2, \ldots$ the square-free integers whose prime factors are all less than or equal to p_k, and by $a(m)$ the greatest a_i contained in m. Evidently

$$f_k(m)=f[a(m)].$$

By $\psi(n, a_i, a_j)$ we denote the number of integers $m\leqslant n$ such that $a(m)=a_i$, $a(m+1)=a_j$. Evidently $\psi(n, a_i, a_j)=0$ if $(a_i, a_j)\neq 1$.

We obtain $\psi(n, a_i, a_j)$ by taking all integers $m\leqslant n$ for which $a_i\mid m$ but $p\nmid m$ if $p\leqslant p_k$ unless $p\mid a_i$; and $a_j\mid(m+1)$ but $p\nmid(m+1)$ if $p\leqslant p_k$ unless $p\mid a_j$.

With these conditions we find by the sieve of Eratosthenes and omission of the square brackets

$$\frac{n}{a_i a_j}\prod_{\substack{p\leqslant p_k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)-2^{2k}<\psi(n, a_i, a_j)<\frac{n}{a_i a_j}\prod_{\substack{p\leqslant p_k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)+2^{2k};$$

and similarly (3)

$$\frac{n}{a_i a_j}\prod_{\substack{p\leqslant p_k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)-2^{2k}<\psi(n, a_j, a_i)<\frac{n}{a_i a_j}\prod_{\substack{p\leqslant p_k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)+2^{2k}.$$

From these $$\lim_{n\to\infty}\frac{\psi(n, a_i, a_j)}{\psi(n, a_j, a_i)}=1. \qquad (4)$$

Since $$G(f_k, n)=\sum_{f(a_i)\leqslant f(a_j)}\psi(n, a_i, a_j),$$

and similarly

$$S(f_k, n)=\sum_{f(a_i)\geqslant f(a_j)}\psi(n, a_i, a_j)=\sum_{f(a_i)\leqslant f(a_j)}\psi(n, a_j, a_i),$$

we have, by (4), $$\lim_{n\to\infty}\frac{G(f_k, n)}{S(f_k, n)}=1.$$

We now prove that, for every $\epsilon>0$, a k exists so great that, if $n\geqslant n(\epsilon)$,

$$|\,G(f,n)-G(f_k, n)\,|<\epsilon n, \qquad (5)$$

and similarly $$|\,S(f,n)-S(f_k, n)\,|<\epsilon n. \qquad (6)$$

From these $\lim_{n\to\infty}\frac{G(f,n)}{S(f,n)}=1$ follows immediately.

We require two lemmas.

LEMMA 1. *For every ϵ we can find a number δ such that, if ν is the number of integers $m \leqslant n$ for which $|f_k(m+1)-f_k(m)| \leqslant \delta$, then $\nu < \frac{1}{2}\epsilon n$ for $k > k(\epsilon)$.*

We have evidently
$$\nu = \sum_{\substack{a_i \quad a_j \\ |f(a_j)-f(a_i)| \leqslant \delta}} \sum \psi(n, a_i, a_j). \tag{7}$$

We now split the sum (7) into two parts $\sum_1$ and $\sum_2$, $\sum_1$ containing those a_i's and a_j's for which $\prod_{\substack{p \mid a_i a_j \\ p \neq 2}} \left(1-\frac{2}{p}\right) < e^{-1/\epsilon^2}$ and $\sum_2$ all the other a_i's and a_j's.

First we evaluate $\sum_1$.

$\sum_1$ is evidently less than or equal to the number μ of integers $m \leqslant n$ for which
$$g(m) = \prod_{\substack{p \mid m(m+1) \\ p \leqslant p_k \\ p \neq 2}} \left(1-\frac{2}{p}\right) < e^{-1/\epsilon^2}.$$

Consider now the product $\prod_{m=1}^{n} g(m) < e^{-\mu/\epsilon^2}$. The factor $1-\frac{2}{p}$ for given p occurs at most $\left[\frac{n}{p}\right]+\left[\frac{n+1}{p}\right] \leqslant \frac{2n}{p}$ times and so (really by Legendre's argument)
$$\prod_{m=1}^{n} g(m) \geqslant \prod_{\substack{p \leqslant p_k \\ p \neq 2}} \left(1-\frac{2}{p}\right)^{2n/p} = \prod_{\substack{p \leqslant p_k \\ p \neq 2}} \left\{\left(1-\frac{2}{p}\right)^{2/p}\right\}^{n} > \frac{1}{c_1^n}.$$

Thus
$$e^{-\mu/\epsilon^2} > \frac{1}{c_1^n},$$

hence
$$\textstyle\sum_1 \leqslant \mu < \epsilon^2 n \log c_1.$$

We now split $\sum_2$ into two parts $\sum_2'$ and $\sum_2''$, where $\sum_2'$ contains only those a_i's and a_j's for which $a_i a_j > p_k^{1/\epsilon^2}$.

$\sum_2'$ is less than or equal to the number ρ of the integers $m \leqslant n$ for which
$$A(m) = a(m)\, a(m+1) > p_k^{1/\epsilon^2}.$$

By Legendre's argument, we have
$$\prod_{m=1}^{n} A(m) \leqslant \prod_{p \leqslant p_k} p^{2n/p} = \exp\left(2n \sum_{p \leqslant p_k} \frac{\log p}{p}\right) < p_k^{2c_2 n},$$

since
$$\sum_{p \leqslant p_k} \frac{\log p}{p} < c_2 \log p_k.$$

Hence
$$p_k^{\rho/\epsilon^2} < p_k^{2c_2 n},$$

thus
$$\textstyle\sum_2' \leqslant \rho < 2c_2 n \epsilon^2.$$

Finally, we have to evaluate $\sum_2''$.

For the a_i's and a_j's occurring in $\sum_2''$, we have, from $\prod_{p \leqslant p_k} \left(1-\frac{2}{p}\right) < \frac{c_3}{(\log p_k)^2}$, omitting the terms for which $\prod_{\substack{p | a_i a_j \\ p \neq 2}} \left(1-\frac{2}{p}\right) < e^{-1/\epsilon^2}$,

$$\prod_{\substack{p \leqslant p_k \\ p \nmid a_i a_j}} \left(1-\frac{2}{p}\right) < \frac{c_3 e^{1/\epsilon^2}}{(\log p_k)^2}.$$

Hence from (3) since a_i, a_j can each at most take 2^k values,

$$\sum_2'' < \frac{c_3 n e^{1/\epsilon^2}}{(\log p_k)^2} \sum_{\substack{a_i \\ |f(a_j)-f(a_i)|<\delta}}' \sum_{a_j}' \frac{1}{a_i a_j} + 2^{4k} < c_4 \frac{n e^{1/\epsilon^2}}{(\log p_k)^2} \sum_{\substack{a_i \\ |f(a_j)-f(a_i)|<\delta}}' \sum_{a_j}' \frac{1}{a_i a_j}.$$

The dash in the summation formulae means that the summation runs only over the a_i's and a_j's for which $a_i a_j < p_k^{1/\epsilon^2}$ and $\prod_{\substack{p | a_i a_j \\ p \leqslant p_k}} \left(1-\frac{2}{p}\right) > e^{-1/\epsilon^2}$.

We now prove that

$$\sum_{\substack{a_i \\ |f(a_j)-f(a_i)|<\delta}}' \sum_{a_j}' \frac{n}{a_i a_j} < c_5 \epsilon^2 e^{-1/\epsilon^2} (\log p_k)^2. \tag{8}$$

First we estimate the sum $\sum_{|f(a_j)-f(a_i)|<\delta}' \frac{1}{a_j}$ for fixed a_i.

We obtain in exactly the same way as in Lemma 1 of my paper* "On the density of some sequences of numbers" that for $l > c_6$ the number of integers $m \leqslant l$ for which $|f(m)-f(a_i)| < \delta$ is less than $\epsilon^4 e^{-1/\epsilon^2} l$. Hence

$$\sum_{|f(a_j)-f(a_i)|<\delta}' \frac{1}{a_j} < 1 + \tfrac{1}{2} + \ldots + \frac{1}{c_6} + \epsilon^4 e^{-1/\epsilon^2} \left(\frac{1}{c_4+1} + \frac{1}{c_4+2} + \ldots + \frac{1}{[p_k^{1/\epsilon^2}]}\right)$$

$$< 2 \log c_6 + \epsilon^4 e^{-1/\epsilon^2} \frac{2 \log p_k}{\epsilon^2} < c_7 \epsilon^2 e^{-1/\epsilon^2} \log p_k.$$

Since $\sum_{a_i} \frac{1}{a_i} = \prod_{p \leqslant p_k} \left(1+\frac{1}{p}\right) < c_8 \log p_k$, (8) is proved. From (8), we have

$$\sum_2'' < c_9 \epsilon^2 n.$$

And finally
$$V = \sum_1 + \sum_2' + \sum_2'' < \epsilon^2 n (\log c_1 + 2c_2 + c_9) < \tfrac{1}{2} \epsilon n.$$

LEMMA 2. *There are at most $\frac{1}{2}\epsilon n$ integers $m \leqslant n$ for which at least one of the inequalities*

$$f(m) - f_k(m) > \delta, \quad f(m+1) - f_k(m+1) > \delta$$

holds for sufficiently large $k = k(\epsilon)$.

* The lemma asserts that for every ϵ we can find a δ such that the number of integers $m \leqslant n$ for which $c \leqslant f(m) \leqslant c+\delta$ is less than ϵn. See also my paper "On the density of some sequences of numbers, II", which will appear shortly in the *Journal of the London Math. Soc.*

For clearly

$$\sum_{m=1}^{n}\{f(m)-f_k(m)\}=\sum_{p=p_{k+1}}^{n}\left[\frac{n}{p}\right]f(p)<n\sum_{p=p_{k+1}}^{\infty}\frac{f(p)}{p}<\tfrac{1}{4}\epsilon\delta n,$$

since $\sum_p \frac{f(p)}{p}$ converges. Hence the lemma is proved.

We proceed to prove (5) and (6). It will be sufficient to prove that the number of integers $m \leqslant n$, for which $f_k(m+1)-f_k(m)$ is not of the same sign as $f(m+1)-f(m)$, is less than ϵn.

We split these integers into two classes. In the first class are those for which $|f_k(m+1)-f_k(m)| \leqslant \delta$. By Lemma 1, the number of these is less than $\frac{1}{2}\epsilon n$. For the integers of the second class $|f_k(m+1)-f_k(m)|>\delta$. For these, evidently one of the inequalities $f(m)-f_k(m)>\delta$, $f(m+1)-f_k(m+1)>\delta$ holds. Thus by Lemma 2 their number is also less than $\frac{1}{2}\epsilon n$, and so (5) is proved.

We now have to show that there are only $o(n)$ integers $m \leqslant n$ for which $f(m)=f(m+1)$.

The argument is exactly the same as the one above. We split the integers $m \leqslant n$ with $f(m)=f(m+1)$ into two classes, putting into the first those for which $|f_k(m+1)-f_k(m)| \leqslant \delta$. By Lemma 1, it follows that their number is less than $\frac{1}{2}\epsilon n$. For the integers of the second class $|f_k(m+1)-f_k(m)| \geqslant \delta$, so that one of the inequalities $f(m)-f_k(m)>\delta$, $f(m+1)-f_k(m+1)>\delta$ holds; hence, from Lemma 2, their number is less than $\frac{1}{2}\epsilon n$.

Hence the Theorem is completely proved for the special case $f(p^\alpha)=f(p)$. The transition to the general case when $f(p^\alpha)\neq f(p)$ is so simple that it will suffice to outline the proof. We define

$$f_k(m)=\sum_{p_i<p_k} f(p_i^{\alpha_i}), \quad \text{where} \quad p_i^{\alpha_i} \mid m,\ p_i^{\alpha_i+1} \nmid m.$$

Then the proof runs just as in the special case if we note that there are at most $c_{10}n/p_k$ integers $m \leqslant n$ divisible by a square greater than p_k, since

$$\sum_{l>p_k}\frac{n}{l^2}<\frac{c_{10}n}{p_k}.$$

We now take for $f(m)$ the functions $\frac{\sigma(m)}{m}$ and $\frac{m}{\phi(m)}$, where $\sigma(m)$ denotes the sum of the divisors of m and $\phi(m)$ denotes Euler's function. We can then deduce the theorem that the number of integers $m \leqslant n$, for which $\sigma(m+1)>\sigma(m)$, is asymptotically $\frac{1}{2}n$; the same is true for $\phi(m)$, since we can easily deduce from Lemmas 1 and 2 that there are only $o(n)$ integers $m \leqslant n$ for which the sign of $\frac{\sigma(m)}{m}-\frac{\sigma(m+1)}{m+1}$ is not the same as the sign of $\sigma(m)-\sigma(m+1)$.

The same theorem holds for the slightly more general case when $\sum_p \frac{f(p)}{p}$ does not converge but the primes can be split into two classes, q_1's and q_2's, so that each of the series $\sum_{q_1} \frac{f(q_1)}{q_1}$ and $\sum_{q_2} \frac{1}{q_2}$ converges and can be proved in a similar way.

2. Now we come to $d(m)$. Denote by $V(m)$ the number of the different prime factors of m. Denote by $G(V, n)$ and $S(V, n)$ the number of integers $m \leqslant n$, for which $V(m) \leqslant V(m+1)$ and $V(m) \geqslant V(m+1)$ respectively. We prove that

$$\lim_{n\to\infty} \frac{G(V,n)}{n} = \tfrac{1}{2} \tag{9}$$

and

$$\lim_{n\to\infty} \frac{S(V,n)}{n} = \tfrac{1}{2}. \tag{10}$$

If we use the method of § 1 without any modification, denoting by $V_k(m)$ the number of different primes not greater than p_k dividing m, we come to no result, since Lemma 2 breaks down. We must take k as a function of n, e.g. $k = n^{\frac{1}{(\log\log n)^3}}$.

We give the particulars of the proof only where it differs essentially from the argument used in § 1.

First we show that

$$\lim_{n\to\infty} \frac{G(V_k, n)}{S(V_k, n)} = 1. \tag{11}$$

Let us denote again by $a_1, a_2, \ldots, a_l$ the square-free integers whose only factors are primes not greater than k, and by $a(m)$ the greatest a_i contained in m. Evidently $V_k(m) = V[a(m)]$.

We may show exactly as in Lemma 1 of § 1 that the number of integers $m \leqslant n$, for which $a(m)\,a(m+1) > n^{\frac{1}{(\log\log n)^2}}$, is $o(n)$.

We consider now the number of the m's, for which $a(m)\,a(m+1) \leqslant n^{\frac{1}{(\log\log n)^2}}$.

We denote by $\psi(n, a_i, a_j)$ the number of integers $m \leqslant n$ such that $a(m) = a_i$ and $a(m+1) = a_j$.

We evaluate $\psi(n, a_i, a_j)$ by Brun's method. As in § 1, we obtain $\psi(n, a_i, a_j)$ by taking all integers $m \leqslant n$ for which $a_i \mid m$ but $p \nmid m$ if $p \leqslant k$ and $p \nmid a_i$; and $a_j \mid m+1$ but $p \nmid m+1$ if $p \leqslant k$ and $p \mid a_j$.

Let now $p_1, p_2, \ldots, p_l$ be any l primes not dividing $a_i a_j$. Denote by

$$\left[\frac{n \,.\, 2^l}{a_i a_j p_1 p_2 \cdots p_l}\right]$$

the number of integers $m \leqslant n$ for which $a_i a_j \mid m$ and

$$\begin{aligned} m &\equiv 0 \text{ or } -1 \pmod{p_1}, \\ m &\equiv 0 \text{ or } -1 \pmod{p_2}, \\ &\cdots\cdots\cdots\cdots \\ m &\equiv 0 \text{ or } -1 \pmod{p_l}. \end{aligned}$$

We evidently have

$$\frac{n\,.\,2^l}{a_i a_j p_1 p_2 \dots p_l} - 2^l \leqslant \left[\frac{n\,.\,2^l}{a_i a_j p_1 p_2 \dots p_l}\right]' \leqslant \frac{n\,.\,2^l}{a_i a_j p_1 p_2 \dots p_l} + 2^l.$$

Now, by the sieve of Eratosthenes, we obtain

$$\psi(n, a_i, a_j) = \left[\frac{n}{a_i a_j}\right] - \left[\frac{2n}{a_i a_j p_1}\right]' - \left[\frac{2n}{a_i a_j p_2}\right]' - \dots + \left[\frac{4n}{a_i a_j p_1 p_2}\right]' + \dots$$
$$+ (-1)^l \sum \left[\frac{2^l n}{a_i a_j p_1 p_2 \dots p_l}\right]' + \dots,$$

where the summation refers to all sets of l primes all less than k no two of which are equal, and no one of which divides $a_i a_j$.

We write
$$s_l = \sum \frac{2^l n}{a_i a_j p_1 p_2 \dots p_l},$$
and
$$s_l' = \sum \left[\frac{2^l n}{a_i a_j p_1 p_2 \dots p_l}\right]'.$$

Let $2t-1$ be the least odd integer greater than $10 \log\log n$, then, following Landau's argument (*Vorlesungen über Zahlentheorie*, 1, 75), we obtain

$$\sum_{l=1}^{2t-1} (-1)^l s_l' \leqslant \psi(n, a_i, a_j) < \sum_{l=1}^{2t} (-1)^l s_l'.$$

By omitting the square brackets on both sides, we get

$$\sum_{l=1}^{2t-1} (-1)^l s_l - 2^{10\log\log n+1}(1+k)^{10\log\log n+1} < \psi(n, a_i, a_j)$$
$$< \sum_{l=1}^{2t} (-1)^l s_l + 2^{10\log\log n+2}(1+k)^{10\log\log n+2},$$

since the number of terms in s_l is less than $\binom{k}{l}$ and

$$1 + \binom{k}{1} + \binom{k}{2} + \dots + \binom{k}{2t} < (1+k)^{2t}.$$

Now
$$\sum_{l > 10\log\log n} s_l < \frac{n}{a_i a_j} \sum_{l > 10\log\log n} \frac{2^l \left(\sum_{p_i \leqslant k} \frac{1}{p_i}\right)^l}{l!}$$
$$< \frac{n}{a_i a_j} \sum_{l > 10\log\log n} \frac{(2\log\log n)^l}{l!} < \frac{2n}{a_i a_j} \frac{(2\log\log n)^h}{h!},$$

h being the least integer exceeding $10\log\log n$. Since $1/h! < h^h e^{-h}$,

$$\sum_{l > 10\log\log n} s_l < \frac{2n}{a_i a_j} \frac{(2\log\log n)^h}{h^h} e^h < \frac{2n}{a_i a_j} \frac{(2\log\log n)^{10\log\log n+1} e^{10\log\log n+1}}{(10\log\log n)^{10\log\log n}}$$
$$= \frac{2n}{a_i a_j} \left(\frac{2e}{10}\right)^{10\log\log n} 2e\log\log n$$
$$< \frac{2n}{a_i a_j} \left(\frac{3}{5}\right)^{10\log\log n} 2e\log\log n < \frac{n}{a_i a_j (\log n)^3}.$$

Hence

$$\sum_{l}(-1)^l s_l - 2^{10\log\log n+1}(1+k)^{10\log\log n+1} - \frac{n}{a_i a_j (\log n)^3} < \psi(n, a_i, a_j)$$
$$< \sum_{l}(-1)^l s_l + 2^{10\log\log n+2}(1+k)^{10\log\log n+2} + \frac{n}{a_i a_j (\log n)^3},$$

where the summation refers to all possible values of l, and so the sum is finite since there are only a finite number of primes not exceeding k. But

$$\sum_{l}(-1)^l s_l = \frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right) > \frac{n}{a_i a_j}\frac{1}{(\log n)^2}, \tag{12}$$

and, since $$a_i a_j < n^{\frac{1}{\log\log n}} \quad\text{and}\quad k < n^{\frac{1}{(\log\log n)^3}},$$

we have

$$\frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right) - \frac{2n}{a_i a_j(\log n)^3} < \psi(n, a_i, a_j) < \frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right) + \frac{2n}{a_i a_j(\log n)^3}.$$

Thus from (12)

$$\frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)\left(1-\frac{2}{\log n}\right) < \psi(n, a_i, a_j) < \frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)\left(1+\frac{2}{\log n}\right). \tag{13}$$

Similarly

$$\frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)\left(1-\frac{2}{\log n}\right) < \psi(n, a_j, a_i) < \frac{n}{a_i a_j}\prod_{\substack{p<k\\ p\nmid a_i a_j}}\left(1-\frac{2}{p}\right)\left(1+\frac{2}{\log n}\right). \tag{14}$$

Hence finally $$1-\frac{4}{\log n+2} < \frac{\psi(n, a_i, a_j)}{\psi(n, a_j, a_i)} < 1+\frac{4}{\log n-2}. \tag{15}$$

Now $$G(V_k, n) = \sum_{\substack{f(a_i)\leqslant f(a_j)\\ a_i a_j < n^{\frac{1}{\log\log n}}}} \psi(n, a_i, a_j) + o(n),$$

and $$S(V_k, n) = \sum_{\substack{f(a_i)\leqslant f(a_j)\\ a_i a_j < n^{\frac{1}{\log\log n}}}} \psi(n, a_j, a_i) + o(n).$$

Thus from (15) $$\lim_{n\to\infty}\frac{G(V_k, n)}{S(V_k, n)} = 1,$$

and so (11) is proved.

Now similarly, as in (5), we prove that

$$|\,G(V, n) - G(V_k, n)\,| < \epsilon n, \tag{16}$$

and $$|\,S(V, n) - S(V_k, n)\,| < \epsilon n, \tag{17}$$

by the aid of two lemmas.

LEMMA 3. *The number ν' of integers $m \leqslant n$ for which*

$$|V_k(m+1) - V_k(m)| < (\log\log\log n)^4$$

is $o(n)$.

We have, as in § 1,

$$\nu' = \sum_{\substack{a_i \quad a_j \\ |V(a_j)-V(a_i)|<(\log\log\log n)^4}} \sum \psi(n, a_i, a_j). \tag{18}$$

From (13) we obtain

$$\psi(n, a_i, a_j) < \frac{c_{11}n}{a_i a_j} \prod_{\substack{p\leqslant k \\ p\nmid a_i a_j}} \left(1 - \frac{2}{p}\right). \tag{19}$$

We detail the proof of Lemma 3, since this is the most complicated part of § 2; nevertheless it will be seen that it is very similar to the proof of Lemma 1 in § 1.

We split the sum (18) into two parts $\sum_1$ and $\sum_2$, $\sum_1$ containing only those a_i's and a_j's for which $\prod_{\substack{p|a_i a_j \\ p\neq 2}} \left(1-\frac{2}{p}\right) < \frac{1}{\log\log\log n}$ and $\sum_2$ all the other a_i's and a_j's.

$\sum_1$ evidently does not exceed the number μ' of integers $m \leqslant n$, for which

$$g(m) = \prod_{\substack{p|m(m+1) \\ p\leqslant k \\ p\neq 2}} \left(1-\frac{2}{p}\right) < \frac{1}{\log\log\log n}.$$

Now, by Legendre's argument,

$$\prod_{m=1}^{n} g(m) \geqslant \prod_{\substack{p\leqslant k \\ p\neq 2}} \left(1-\frac{2}{p}\right)^{2n/p} = \prod_{\substack{p\leqslant k \\ p\neq 2}} \left[\left(1-\frac{2}{p}\right)^{2/p}\right]^n.$$

Thus

$$\frac{1}{(\log\log\log n)^{\mu'}} > \frac{1}{c_{12}^n},$$

hence

$$\sum\nolimits_1 \leqslant \mu' < \frac{n\log c_{12}}{\log\log\log\log n} = o(n).$$

Finally we evaluate $\sum_2$.

For the a_i's and a_j's occurring in $\sum_2$, we have, since $\prod_{p\leqslant k}\left(1-\frac{2}{p}\right) < \frac{c_{13}}{(\log k)^2}$, from the definition of $\sum_2$,

$$\prod_{\substack{p\leqslant k \\ p\nmid a_i a_j}} \left(1-\frac{2}{p}\right) < \frac{c_{13}\log\log\log n}{(\log k)^2} = \frac{c_{13}\log\log\log n\,(\log\log n)^6}{(\log n)^2}.$$

Hence, from (18) and (19),

$$\sum\nolimits_2 < \frac{c_{13}\log\log\log n\,(\log\log n)^6}{(\log n)^2} \sum_{\substack{a_i \quad a_j \\ |V(a_j)-V(a_i)|<(\log\log\log n)^4}} \sum \frac{n}{a_i a_j}. \tag{20}$$

First we estimate the sum $\sum_{|V(a_j)-V(a_i)|<(\log\log\log n)^4} \frac{1}{a_j}$ for fixed a_i.

We require the sum of the reciprocal value R of the a's which have exactly v prime factors.

We evidently have

$$R < \frac{\left(\sum_{p\leqslant k}\frac{1}{p}\right)^v}{v!} < \frac{(\log\log k + c_{14})^v}{v!} \leqslant \frac{(\log\log k + c_{14})^q}{q!},$$

where q denotes the greatest integer not exceeding $\log\log k + c_{14}$.

Further, by Stirling's formula,

$$R < c_{15}\frac{(\log\log k + c_{14})^q e^q}{q^q q^{\frac{1}{2}}} < c_{16}\frac{\log k}{(\log\log k)^{\frac{1}{2}}} < c_{16}\frac{\log n}{(\log\log n)^{7/2}},$$

since
$$\frac{(\log\log k + c_{14})^q}{q^q} < \left(1+\frac{c_{17}}{q}\right)^q < c_{18}.$$

Hence summing for v, which runs through $2(\log\log\log n)^4$ values, we get

$$\sum_{|V(a_j)-V(a_i)|<(\log\log\log n)^4} \frac{1}{a_j} < 2c_{16}\frac{\log n\,(\log\log\log n)^4}{(\log\log n)^{7/2}}.$$

Since
$$\sum_{a_i}\frac{1}{a_i} = \prod_{p\leqslant k}\left(1+\frac{1}{p}\right) < c_{19}\log k = \frac{c\log n}{(\log\log n)^3},$$

we have, on multiplying the two right-hand sides just above,

$$\sum_{a_i}\sum_{a_j}{}_{|V(a_j)-V(a_i)|<(\log\log\log n)^4} \frac{1}{a_i a_j} < c_{20}\frac{(\log n)^2(\log\log\log n)^4}{(\log\log n)^{13/2}}.$$

From this by (20)
$$\Sigma_2 = c_{21}\frac{(\log\log\log n)^5}{(\log\log n)^{1/2}} = o(n);$$

hence finally
$$\nu' = \Sigma_1 + \Sigma_2 = o(n).$$

LEMMA 4. *There are only $o(n)$ integers $m \leqslant n$, for which one of the inequalities $V(m) - V_k(m) > (\log\log\log n)^2$, $V(m+1) - V_k(m+1) > (\log\log\log n)^2$ holds.*

The proof runs parallel to that of Lemma 2 in the first part. Just as we obtained (5) and (6) from Lemmas 1 and 2, so we derive (16) and (17) from Lemmas 3 and 4.

From (11), (16) and (17), it follows that

$$\lim_{n\to\infty}\frac{G(V,n)}{S(V,n)} = 1.$$

By Lemmas 3 and 4 we can show, as in § 1, that there are only $o(n)$ integers $m \leqslant n$ for which $V(m) = V(m+1)$. From this we deduce

$$\lim_{n\to\infty}\frac{G(V,n)}{n} = \tfrac{1}{2}; \quad \lim_{n\to\infty}\frac{S(V,n)}{n} = \tfrac{1}{2}.$$

To obtain Chowla's conjecture, we need only prove that there are $o(n)$ integers $m \leqslant n$ for which $V(m+1)-V(m) \geqslant 0$ and $d(m+1)-d(m) \leqslant 0$, or $V(m+1)-V(m) \leqslant 0$ and $d(m+1)-d(m) \geqslant 0$. It will be sufficient to settle the first case.

First we observe that it is easy to obtain from Lemmas 3 and 4 that for almost all integers* $\quad m \leqslant n, \quad |V(m+1)-V(m)| > (\log\log\log n)^2.$

We now split the integers for which both $V(m+1)-V(m) \geqslant 0$ and $d(m+1)-d(m) \leqslant 0$ into two classes, putting in the first those for which

$$V(m+1)-V(m) < (\log\log\log n)^2,$$

and in the second those for which

$$V(m+1)-V(m) \geqslant (\log\log\log n)^2.$$

The number of the integers of the first class is $o(n)$, by the remark above.

The integers of the second class satisfy $\dfrac{2^{V(m+1)}}{2^{V(m)}} \geqslant 2^{(\log\log\log n)^2}$, and since $d(m+1) \geqslant 2^{V(m+1)}$ we have $d(m) \geqslant 2^{V(m)}\, 2^{(\log\log\log n)^2}$. Put $m = AB^2$, where A is square-free. We have $d(m) \leqslant d(A)\, d(B^2) = 2^{V(m)}\, d(B^2)$, so that $d(B^2) \geqslant 2^{(\log\log\log n)^2}$ and hence $B^2 \geqslant 2^{(\log\log\log n)^2}$. Thus m is divisible by a square not less than $2^{(\log\log\log n)^2}$, so that the number of integers of the second class is less than equal to

$$\sum_{l^2 > 2^{(\log\log\log n)^2}} \frac{n}{l^2} = o(n).$$

Hence the result.

* More generally we can prove the following theorem. Let $X(n)$ be an arbitrary function with $\lim_{n\to\infty} X(n) = \infty$. Then, for almost all integers $m \leqslant n$,

$$\frac{\log\log n}{X(n)} < |V(m+1)-V(m)| < \log\log n X(n).$$

The first inequality may be proved by similar but stronger lemmas than Lemmas 3 and 4. The second inequality has been proved by P. Turán as follows:

$$\begin{aligned}\sum_{m=1}^{n} (V(m+1)-V(m))^2 &= \sum_{m=1}^{n} \{(V(m+1)-\log\log n)-(V(m)-\log\log n)\}^2 \\ &\leqslant 2\left[\sum_{m=1}^{n} (V(m)-\log\log n)^2 + \sum_{m=1}^{n} (V(m+1)-\log\log n)^2\right] \\ &= O(n \log\log n),\end{aligned}$$

which immediately establishes the result.