Indian Journal of Mathematics, Vol. 32, *No.* 3, *Sept.* 1990
Prof. Hansraj Gupta Memorial Volume

ON A THEOREM OF BESICOVITCH: VALUES OF ARITHMETIC FUNCTIONS THAT DIVIDE THEIR ARGUMENTS

By
PAUL ERDÖS AND CARL POMERANCE*
(*Received March* 30, 1990)

Suppose $g(n)$ tends monotonically to infinity and $g(n)/n$ tends to zero. If f is an integer-valued arithmetic function with normal order g, then the set of n such that $f(n)$ divides n has asymptotic density zero. More generally, the set of n with a divisor between $g(n)$ and $2\,g(n)$ has asymptotic density zero.

1. *Introduction and statement of results*

Recently, Cooper and Kennedy [2] considered the following problem. Let f be an arithmetic function with positive integer values. That is, $f: N \to N$. What can be said about the set of n with $f(n)|n$?

THEOREM (Cooper, Kennedy). *Let* $f: N \to N$, *let* $\mu(x)$ *denote the mean of* $f(1), \ldots, f([x])$ *and let* $\sigma(x)$ *denote the standard deviation. If* $\mu(x) \to \infty$ *and* $\sigma(x)/\mu(x) \to 0$, *then the set of* n *with* $f(n)|n$ *has asymptotic density* 0.

Cooper and Kennedy give several examples to illustrate their theorem including $f(n)=\omega(n)$, the number of distinct prime factors of n, and $f(n)=s_g\,(n)$, the sum of the base g digits of n.

In this note we first give a short proof of a result that is very similar to the Cooper, Kennedy theorem. In particular, it is good for every example they consider and a few for which their theorem is not strong enough. Next we give a somewhat more complicated proof of a stronger theorem.

*Supported, in part, by an NSF grant.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11K65; Secondary 11B83.

Recall that we say an arithmetic function f has normal order g if for each $\epsilon>0$, the set of n with

$$|f(n)-g(n)|>\epsilon\, g(n)$$

has asymptotic density 0. We first prove

THEOREM 1. *Suppose $g(n)$ is monotone, $g(n) \to \infty$ and $g(n)/n \to 0$. If f has normal order g, then the set of n with $f(n)|n$ has asymptotic density* 0.

Since $f(n)=\omega(n)$ has normal order loglog n, a result of Hardy and Ramanujan, it follows that $\omega(n)|n$ only on a set of asymptotic density 0. The same result holds for $f(n)=\Omega(n)$, the number of prime factors of n counted with multiplicity. (This was stated as an open problem by Cooper and Kennedy. However, $\Omega(n)$ also satisfies the hypothesis of their theorem.)

The question of the set of n for which $\omega(n)|n$ actually was already considered by Spiro [9] who showed the number of such $n\leq x$ is $\sim x/\log\log x$. She also obtained the same result for $\Omega(n)$.

It is clear that $s_g(n)$ has normal order $c_g \log n$ where $c_g=(g-1)/(2\log g)$. Thus the set of n with $s_g(n)|n$ has asymptotic density 0.

Another example of a function satisfying the hypothesis of Theorem 1 is $f(n)=\pi(n)$, the number of primes up to n. From the prime number theorem, $\pi(n)\sim n/\log n$, so that $\pi(n)$ has normal order $n/\log n$. In fact, $\pi(n)$ does not satisfy the hypothesis of the Cooper, Kennedy theorem. However, the fact that $\pi(n)$ almost never divides n is really rather trivial and does not require our Theorem 1. Indeed, $\pi(n)$ stays constant on long intervals, so from this one can see that $\pi(n)|n$ is rare. It is amusing to note that $\pi(n)|n$ does hold infinitely often-in fact, the only properties of $\pi(n)$ used to show this are that $\pi(n)$ is integer valued, monotone and $o(n)$.

Our Theorem 2 majorizes Theorem 1, but the proof is much harder.

THEOREM 2. *Suppose $g(n)$ is monotone, $g(n)\to\infty$ and $g(n)/n \to 0$. Then the set of n with a divisor between $g(n)$ and $2\,g(n)$ has asymptotic density* 0.

This theorem is very reminiscent of an old result of Besicovitch [1] that says that if d_T is the asymptotic density of the set of n with a divisor in $(T, 2T]$, then $\liminf d_T=0$. Later, in [3], the first author improved this result to $\lim d_T=0$. Our Theorem 2 follows from the same circle of ideas used in [3].

It is not too easy to find applications of Theorem 2 for naturally occurring arithmetic functions $f(n)$ for which Theorem 1 is not strong enough. One example is the following which appeared in Pomerance and Stone [8]. Let $D(n)$ denote the largest divisor of n in the interval $[1, \sqrt{n}]$. Then from Theorem 2 it easily follows that there is a set S of asymptotic density 1 such that $D(n)/\sqrt{n} \to 0$ as $n \to \infty, n \in S$. Of course, the function $\sqrt{n}$ can be replaced with any function $g(n)$ satisfying the hypotheses of Theorem 2.

Consider the function $S(n)=\Sigma_{p|n}P$, the sum of the prime factors of n. Our theorems do not immediately answer the question of how often $S(n)|n$. However, using the ideas of this paper and of [5], it can be shown that the set of n with $S(n)|n$ has asymptotic density 0. The idea is to first restrict to integers n with $P(n)<S(n)<2P(n)$, where $P(n)$ is the largest prime factor of n (see [5]). We now count how many such n have $S(n)|n$ and $n \in (x/2, x]$. Write such an n as mp, where $p=P(n)$. Then $S(n)|m$, so that m has a divisor between $x/(2m)$ and $2x/m$. The ideas of this paper can now complete the proof that the number of such n is $o(x)$.

Finally we remark that the Besicovitch-Erdös result was greatly improved by the first author in [4] (also see [6], Ch. V, Sec. 7 and [7], Ch. 2). Namely, it is shown that if $\epsilon_T \to 0$ arbitrarily slowly, then the set of integers with a divisor between T and $T^{1+\epsilon_T}$ has asymptotic density that tends to 0 as $T \to \infty$. It remains to be seen if there is an analog of this result along the lines of Theorem 2, but almost certainly some sort of stronger theorem could be proved using these ideas.

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

2. *Proof of Theorem* 1

Let $1>\epsilon>0$ be fixed, but arbitrary. It will be sufficient to show that there is some absolute constant c such that for all large x the number $N=N(x)$ of $n \in (x, 2x]$ with $f(n)|n$ is at most $c\,\epsilon\, x$. In fact, we shall show that $c=10$ will do.

If x is sufficiently large, then the number of $n \in (x, 2x]$ such that

$$(1) \qquad (1+\epsilon)^{-1}g(n)<f(n)<(1+\epsilon)g(n)$$

fails is at most ϵx. Indeed, this follows immediately from the definition of normal order.

Let k_0 be such that

$$(1+\epsilon)^{k_0-1} \leqslant \frac{g(2x)}{g(x)} < (1+\epsilon)^{k_0}.$$

For each $k=1, \ldots, k_0$ the set

$$\{n \in (x, 2x]: (1+\epsilon)^{k-1} g(x) \leqslant g(n) < (1+\epsilon)^k g(x)\}$$

is an interval of consecutive integers-say it is the set of integers in $(a_k, b_k]$, where these k_0 intervals partition $(x, 2x]$.

Let N_k denote the number of $n \in (a_k, b_k]$ such that (1) holds and $f(n)|n$. But if $n \in (a_k, b_k]$ and (1) holds, then

$$f(n) \in ((1+\epsilon)^{k-2} g(x), (1+\epsilon)^{k+1} g(x)). \tag{2}$$

Let I_k denote the set of integers in the interval in (2). Note that

$$|I_k| \leqslant 1 + (1+\epsilon)^{k+1} g(x) - (1+\epsilon)^{k-2} g(x).$$

Thus

$$N_k \leqslant \sum_{d \in I_k} \sum_{\substack{n \in (a_k, b_k] \\ d|n}} 1 = \sum_{d \in I_k} \left(\left[\frac{b_k}{d}\right] - \left[\frac{a_k}{d}\right] \right)$$

$$\leqslant (b_k - a_k) \sum_{d \in I_k} \frac{1}{d} + \sum_{d \in I_k} 1$$

$$\leqslant \frac{b_k - a_k}{(1+\epsilon)^{k-2} g(x)} |I_k| + |I_k|$$

$$\leqslant ((1+\epsilon)^3 - 1)(b_k - a_k) + \frac{b_k - a_k}{(1+\epsilon)^{k-2} g(x)} + |I_k|$$

$$\leqslant 7\epsilon (b_k - a_k) + \frac{1+\epsilon}{g(x)} (b_k - a_k) + |I_k|.$$

We conclude that the number of $n \in (x, 2x]$ for which (1) holds and $f(n)|n$ is at most

$$\sum_{k=1}^{k_0} N_k \leqslant 7\epsilon x + \frac{1+\epsilon}{g(x)} x + 3(1+\epsilon)^{k_0+1} g(x)$$

$$\leqslant 8\epsilon x + 3(1+\epsilon)^2 g(2x) \leqslant 9\epsilon x$$

for x sufficiently large. Together with the number of n for which (1) fails, we have $N \leqslant 10\epsilon x$.

3. *Proof of Theorem 2*

We first establish a lemma that will be used heavily in the proof of Theorem 2. The proof is essentially that of the Turán-Kubilius inequality

LEMMA. *Let $\Omega_z(n)$ denote the number of prime and prime power factors of n that do not exceed z. If $z \geqslant 3$ and $0<C<D$, then the number of integers $n \in (C, D]$ with $|\Omega_z(n)$-loglog $z| \geqslant \frac{1}{3}$ loglog z is uniformly*

$$\ll \frac{D-C}{\log\log z}+\frac{z}{\log z \log\log z}.$$

Proof. Let

$$E=E(z, C, D)=\sum_{C<n\leqslant D} \Omega_z(n),\quad B=B(z, C, D)=\sum_{C<n\leqslant D} \Omega_z(n)^2.$$

Then

$$E=\sum_{C<n\leqslant D}\ \sum_{\substack{p^a\leqslant z\\ p^a|n}} 1=\sum_{p^a\leqslant z}\ \sum_{\substack{C<n\leqslant D\\ p^a|n}} 1=\sum_{p^a\leqslant z}\left(\left[\frac{D}{p^a}\right]-\left[\frac{C}{p^a}\right]\right)$$

$$=(D-C)\sum_{p^a\leqslant z}\frac{1}{p^a}+O\Big(\sum_{p^a\leqslant z} 1\Big)$$

$$=(D-C)\log\log z+O(D-C)+O(z/\log z).$$

Also, we similarly have

$$B=\sum_{C<n\leqslant D}\Big(\sum_{\substack{p^a\leqslant z\\ p^a|n}} 1\Big)\Big(\sum_{\substack{q^b\leqslant z\\ q^b|n}} 1\Big)$$

$$=\sum_{\substack{p^a,\, q^b\leqslant z\\ p\neq q}}\left(\left[\frac{D}{p^aq^b}\right]-\left[\frac{C}{p^aq^b}\right]\right)+2\sum_{\substack{p^b\leqslant z\\ 0<a<b}}\left(\left[\frac{D}{p^b}\right]-\left[\frac{C}{p^b}\right]\right)$$

$$+\sum_{p^a\leqslant z}\left(\left[\frac{D}{p^a}\right]-\left[\frac{C}{p^a}\right]\right)$$

$$=(D-C)(\log\log z)^2+O((D-C)\log\log z)+O\left(\frac{z\log\log z}{\log z}\right),$$

using

$$\sum_{p^a,\, q^b \leqslant z} \frac{1}{p^a q^b} = (\log\log z)^2 + O(\log\log z).$$

Thus

$$\sum_{C<n\leqslant D} (\Omega_z(n) - \log\log z)^2 = B - 2E \log\log z + ([D]-[C])(\log\log z)^2$$

$$\ll (D-C) \log\log z + \frac{z \log\log z}{\log z}.$$

The lemma follows immediately from this estimate.

Proof of Theorem 2. For each natural number x, let $N=N(x)$ denote the number of $n \in (x, 2x]$ with a divisor in $(g(n), 2g(n)]$. It will suffice to show that $N=o(x)$.

Let k_0 be such that

$$2^{k_0-1} \leqslant g(2x)/g(x) < 2^{k_0}.$$

For each $k=1, 2, \ldots, k_0$, the set

$$\{n \in (x, 2x]:\ 2^{k-1}g(x) \leqslant g(n) < 2^k g(x)\}$$

is an interval of consecutive integers-say it is the set of integers in $(A_k, B_k]$ where these k_0 intervals partition the interval $(x, 2x]$. Let

$$z = z(x) = \min\left\{g(x), \frac{x}{g(2x)}\right\}.$$

We now divide the integers $n \in (x, 2x]$ with a divisor $d \in (g(n), 2g(n)]$ into four, possibly overlapping classes:

(i) n has a divisor $d \in (g(n), 2g(n)]$ with $\Omega_z(d) \leqslant \frac{2}{3} \log\log z$,

(ii) n has a divisor $d \in (g(n), 2g(n)]$ with $\Omega_z(n/d) \leqslant \frac{2}{3} \log\log z$,

(iii) n has a prime power divisor p^a with $a>1$, $p^a>z$,

(iv) $\Omega_z(n) \geqslant \frac{4}{3} \log\log z$.

To see that these classes exhaust all possibilities, suppose $n \in (x, 2x]$ has a divisor $d \in (g(n), 2g(n)]$ with

$$\Omega_z(d) > \frac{2}{3} \operatorname{loglog} z \quad \text{and} \quad \Omega_z(n/d) > \frac{2}{3} \operatorname{loglog} z.$$

If also n is not counted in class (iii), then

$$\Omega_z(n) = \Omega_z(d) + \Omega_z(n/d) > \frac{4}{3} \operatorname{loglog} z,$$

so that n is counted in class (iv).

Let $N_{(i)}, N_{(ii)}, N_{(iii)}, N_{(iv)}$ denote the number of n counted in each of the four classes (i), (ii), (iii), (iv), respectively.

Let $N_{(i),k}$ denote the contribution to $N_{(i)}$ from those $n \in (A_k, B_k]$. Let $I_k = (2^{k-1}g(x), 2^{k+1}g(x)]$ and let Σ' denote a sum over integers m with $\Omega_z(m) \leq \frac{2}{3} \operatorname{loglog} z$. Then by the Lemma,

$$\begin{aligned}
N_{(i),k} &\leq \sum_{d \in I_k}{}' \left(\left[\frac{B_k}{d}\right] - \left[\frac{A_k}{d}\right] \right) \\
&\leq \left(\frac{B_k - A_k}{2^{k-1}g(x)} + 1 \right) \sum_{d \in I_k}{}' 1 \\
&\ll \left(\frac{B_k - A_k}{2^{k-1}g(x)} + 1 \right) \left(\frac{2^{k+1}g(x) - 2^{k-1}g(x)}{\operatorname{loglog} z} + \frac{z}{\log z \operatorname{loglog} z} \right) \\
&\ll \frac{B_k - A_k}{\operatorname{loglog} z} + \frac{2^k g(x)}{\operatorname{loglog} z} + \left(\frac{B_k - A_k}{g(x)} + 1 \right) \frac{z}{\log z \operatorname{loglog} z}.
\end{aligned}$$

Thus

$$\begin{aligned}
N_{(i)} = \sum_{k=1}^{k_0} N_{(i),k} &\ll \frac{x}{\operatorname{loglog} z} + \frac{g(2x)}{\operatorname{loglog} z} + \left(\frac{x}{g(x)} + k_0 \right) \frac{z}{\log z \operatorname{loglog} z} \\
&\ll \frac{x}{\operatorname{loglog} z},
\end{aligned}$$

since $g(x) = o(x)$, $z \leqslant g(x)$, $z \leqslant \sqrt{x}$ and $k_0 \ll \log x$.

Next, let $N_{(ii),k}$ denote the contribution to $N_{(ii)}$ from those $n \in (A_k, B_k]$. Let

$$J_k = \left(\frac{A_k}{2^{k+1}g(x)}, \frac{B_k}{2^{k-1}g(x)} \right).$$

Thus if n is counted by $N_{(ii),k}$ then n has a divisor $m \in J_k$ with $\Omega_z(m) \leqslant \frac{2}{3} \log\log z$. Thus, as above,

$$N_{(ii),k} \leqslant \sum_{m \in J_k}' \left(\left[\frac{B_k}{m}\right] - \left[\frac{A_k}{m}\right] \right)$$

$$\leqslant \left(\frac{B_k - A_k}{A_k/(2^{k+1}g(x))} + 1 \right) \sum_{m \in J_k}' 1$$

$$\ll \left(\frac{B_k - A_k}{A_k/(2^{k+1}g(x))} + 1 \right) \left(\frac{B_k/(2^{k-1}g(x)) - A_k/(2^{k+1}g(x))}{\log\log z} + \frac{z}{\log z \log\log z} \right)$$

$$\ll \frac{B_k - A_k}{\log\log z} + \frac{x}{2^k g(x) \log\log z} + \frac{2^k g(x)\, z}{\log z \log\log z},$$

using $x \leqslant A_k \leqslant B_k \leqslant 2x$. Thus

$$N_{(ii)} = \sum_{k=1}^{k_0} N_{(ii),k} \ll \frac{x}{\log\log z} + \frac{g(2x)\, z}{\log z \log\log z} \ll \frac{x}{\log\log z}.$$

By an elementary argument,

$$N_{(iii)} \leqslant \sum_{\substack{p^a > z \\ a > 1}} \frac{2x}{p^a} \ll \frac{x}{\sqrt{z}}.$$

Finally, by the Lemma we have

$$N_{(iv)} \ll \frac{x}{\log\log z}.$$

Putting together our estimates, we have

$$N \leqslant N_{(i)} + N_{(ii)} + N_{(iii)} + N_{(iv)} \ll \frac{x}{\log\log z}.$$

Since $z = z(x) \to \infty$ as $x \to \infty$, we have the thorem.

REFERENCES

1.. A. S. Besicovitch, "On the density of certain sequences of integers", *Math. Ann.*, **110 (1934)**, 336-341.

2. C. N. Cooper and R. E. Kennedy, "Chebyshev's inequality and natural density", *Amer. Math. Monthly*, 96 (1989), 118-124.

3. P. Erdős, "Note on sequences of integers no one of which is divisible by any other", *J. London Math. Soc.*, 10 (1935), 126-128.

4. P. Erdős, "Generalization of a theorem of Besicovitch", *J. London. Math. Soc.*, 11 (1936), 92-98.

5. P. Erdős and C. Pomerance, "On the largest prime factors of n and $n+1$", *Aequationes Math.*, 17 (1978), 311-321.

6. H. Halberstam and K. F. Roth, *Sequences*, Springer-Verlag, New York, 1983.

7. R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge University Press, Cambridge, 1988.

8. C. Pomerance and D. R. Stone, "Solution to problem 342 (proposed by O. Bintz)", *College Math. J.*, 20 (1989), 164-166.

9. C. A. Spiro, The frequency with which an integral-valued, prime-independent, multiplicative or additive function of n divides a polynomial function of n, dissertation at Univ. Illinois at Urbana-Champaign, (1981), 180 pp.

Mathematical Institute of the
Hungarian Academy of Sciences
Realtanoda u. 13-15
H-1053 Budapest
Hungary

Department of Mathematics
University of Georgia
Athens, Georgia 30602
U. S. A.