ON THE SCHNIRELMANN AND ASYMPTOTIC DENSITIES

OF SETS OF NON-MULTIPLES¹

Paul Erdös Mathematical Institute Hungarian Academy of Sciences Budapest, Hungary

Carole B. Lacampagne Mathematics Department University of Michigan-Flint Flint, MI 48502

Carl Pomerance² Department of Mathematics University of Georgia Athens, GA 30602

J. L. Selfridge Mathematics Department Northern Illinois University DeKalb, IL 60115

\$1. Introduction.

If A is a set of natural numbers, let A(x) denote the number of members of A not exceeding x. The asymptotic density of A is

$$\delta A: = \lim_{x \to \infty} \frac{1}{x} A(x) ,$$

should this limit exist. The Schnirelmann density of A is

$$\sigma A: = \inf_{n \ge 1} \frac{1}{n} A(n) ,$$

where n denotes an integer. Of course the Schnirelmann density always exists. If δA exists, then clearly $\sigma A \leq \delta A$.

Often the Schnirelmann density gives little information about A . For example, if $1 \notin A$, then $\sigma A = 0$. One interest in Schnirelmann

 2 Research supported in part by an NSF grant.

¹ This paper contains original research and will not be published elsewhere.

density comes from the theorems of Schnirelmann and Mann [4] that if $\sigma A > 0$, then there is some integer h such that every natural number is a sum of h or fewer members of A. In fact, any $h \ge 1/\sigma A$ will do.

In this note we will be concerned with the case when A = A(S) is the set of natural numbers not divisible by any number in the set S. We shall abbreviate $\delta A(S)$, $\sigma A(S)$, A(S)(x), respectively, by $\delta(S)$, $\sigma(S)$, $S\langle x \rangle$. Most of our considerations will be when S is a set of primes. In this case (or more generally, whenever the members of S are pairwise coprime), $\delta(S)$ must exist; in fact,

$$\delta(S) = \prod_{s \in S} (1 - \frac{1}{s}) .$$

Whenever $\delta(S)$ exists, we define the discrepancy, D(S) , of the set S as the difference

$$D(S) = \delta(S) - \sigma(S) .$$

Among the problems we consider are the maximal possible discrepancy as S ranges over (i) sets of primes, (ii) sets of pairwise coprime integers, (iii) sets of integers for which D(S) is defined. We also consider when a set S of primes with positive discrepancy can be embedded in a larger set of primes S' with zero discrepancy, but the same Schnirelmann density. We also state several problems we have not been able to solve.

Our principal results are listed below.

<u>Theorem 1</u>. For finite sets of natural numbers S, D(S) = 0 if and only if the least member of S divides every member of S.

Theorem 2. sup{D(S):S a finite set of natural numbers} = 1.

Let **P** denote the set of prime numbers. Say a set $S \subset \mathbf{P}$ is <u>minimal</u> if there is some N with $\sigma(S) = \frac{1}{N} S \langle N \rangle$ and $S \subset [1,N]$. Note that if $S \subset \mathbf{P}$ and D(S) > 0, then there is some $S_0 \subset S$ with S_0 minimal, $\sigma(S_0) = \sigma(S)$, and $D(S_0) \ge D(S)$. Let max S denote the largest element of the set S. Let

Clearly we have $D' \leq D \leq D_0$. We have

Theorem 3. D' = 0.285025....

Theorem 4. $D_0 \le e^{-1} = 0.367879...$

<u>Theorem 5</u>. If $S \subset \mathbb{P}$, then there is a set S' with $S \subset S' \subset \mathbb{P}$ such that

- (i) $\sigma(S') = \sigma(S)$
- (ii) D(S') = 0.

§2. Proofs of Theorems 1-5.

<u>Theorem 1</u>. For finite sets of natural numbers S, D(S) = 0 if and only if the least member of S divides every member of S.

<u>Proof</u>. Let m denote the least member of the finite set of natural numbers S . If m divides every member of S , then

 $\delta(S) = \sigma(S) = 1 - \frac{1}{m} , \text{ so } D(S) = 0 .$

Now suppose $\,m\,$ does not divide every member of S , so that $1 < m < N \ , \ where \ N \ denotes \ the \ least \ common \ multiple \ of \ the \ members \ of \ S \ . \ Thus$

$$\delta(S) = \frac{1}{N} S \langle N \rangle \langle \frac{m-1}{m} \rangle.$$

Note that

$$S\langle N \rangle - S\langle N-m \rangle = m-1$$

since no number strictly between N-m and N is divisible by a member of S . Therefore

$$\sigma(S) \leq \frac{1}{N-m} S \langle N-m \rangle = \frac{S \langle N \rangle - (m-1)}{N-m} \langle \frac{S \langle N \rangle}{N} = \delta(S) ,$$

so that D(S) > 0.

Theorem 2. $\sup{D(S):S}$ a finite set of natural numbers} = 1.

<u>Proof</u>. For $n \ge 3$, let S_n denote the set of natural numbers in the interval $[[n/\log n], n]$. Evidently,

$$\sigma(S_n) \leq \frac{1}{n} S_n \langle n \rangle \leq \frac{1}{\log n} ,$$

so that $\lim \sigma(S_n) = 0$. On the other hand, from Erdös [2], $\lim \delta(S_n) = 1$. Therefore $\lim D(S_n) = 1$ and the theorem is proved.

Theorem 3. D' = 0.285025... .

<u>Proof</u>. We shall use the following corollary of the main result in Hildebrand [5].

Theorem (Hildebrand). Let

$$G(x,K) = \min\{\frac{1}{x} \ S \le x >: \ S \subset [1,x] \cap \mathbf{P} \ , \ \sum_{p \in S} \frac{1}{p} \le K\} \ .$$

Then there are positive constants γ , δ such that

$$G(x,K) = \rho(e^K)(1 + O((\log x)^{-\gamma}))$$

uniformly for $1 \leq K \leq \delta \log \log x$.

In this result, ρ denotes the Dickman-de Bruijn function defined by the conditions:

- (i) ρ is continuous on $[0,\infty)$,
- (ii) $\rho(u) = 1$ for $u \in [0,1]$,
- (iii) $\rho'(u) = -\rho(u-1)/u$ for u > 1.

It is known that $\rho(u)$ is positive, non-increasing, and that $\rho(u) \rightarrow 0$ as $u \rightarrow \infty$. From the defining properties of ρ it is easy to see that

(2.1)
$$\rho(u) = 1 - \log u \text{ for } 1 \le u \le 2$$
.

$$f(u): = 1/u - \rho(u)$$
.

Note that f(1) = 0, $f(u) \rightarrow 0$ as $u \rightarrow \infty$, and that (from (2.1)) $f(2) = \log 2 - 1/2 > 0$. Therefore the maximum value of f(u) on $[1,\infty)$ occurs at some finite point $u_0 > 1$. We now show that

(2.2)
$$u_0 = 2.9329475..., f(u_0) = 0.285025...$$

Indeed,

$$f'(u) = \frac{-1}{u^2} - \rho'(u) = \frac{-1 + u\rho(u-1)}{u^2} .$$

Let $g(u) = -1+u\rho(u-1)$, so that sign f' = sign g . For $1 \le u \le 2$, g(u) = -1 + u > 0. For u > 2,

$$g'(u) = \rho(u-1) + u\rho'(u-1) = \rho(u-1) - \frac{u}{u-1}\rho(u-2)$$

< $\rho(u-1) - \rho(u-2) \le 0$,

so that g(u) is decreasing on $[2,\infty)$. From (2.1), on the interval [2,3],

(2.3)
$$g(u) = -1 + u - u \log(u-1)$$

Thus g(2) = 1, $g(3) = 2 - 3 \log 2 < 0$. We conclude that $u_0 \in [2,3]$ and by Newton's method applied to (2.3) we find the value of u_0 claimed in (2.2). To compute $f(u_0)$, we use

$$f(u_{o}) = \frac{1}{u_{o}} - \rho(u_{o}) = \frac{1}{u_{o}} - \rho(2) - \int_{2}^{u_{o}} \rho'(t) dt$$
$$= \frac{1}{u_{o}} - \rho(2) + \int_{2}^{u_{o}} \frac{\rho(t-1)}{t} dt$$
$$= \frac{1}{u_{o}} - 1 + \log 2 + \int_{2}^{u_{o}} \frac{1 - \log(t-1)}{t} dt ,$$

and integrate numerically.

With these preliminaries aside, we can now prove Theorem 3. Let S(x,y) denote $[x,y] \cap P$ and let $u \ge 1$ be fixed. Then from Mertens' theorem,

Let

$$\delta(S(y,y^{u})) \sim \frac{1}{u}$$
 as $y \to \infty$.

Also from Dickman's theorem [1] ,

$$\sigma(S(y,y^u)) \leq y^{-u}S(y,y^u) \langle y^u \rangle \sim \rho(u)$$
 as $y \rightarrow \infty$.

Therefore

D'
$$\geq \lim \sup D(S(y,y^u)) \geq 1/u - \rho(u) = f(u)$$
,
 $y \rightarrow \infty$

so that $D' \ge f(u_0)$.

Thus to complete the proof of Theorem 3, we need only show the reverse inequality. For this we shall use Hildebrand's theorem quoted above. Let $S \subset \mathbf{P}$ be minimal with $\delta(S) \ge 1/4$ and let N be such that $\sigma(S) = \frac{1}{N} S \langle N \rangle$, max $S \le N$. We have

(2.4)
$$1/4 \le \delta(S) = \prod_{p \in S} (1 - \frac{1}{p}) \le \exp(-\sum_{p \in S} \frac{1}{p})$$

so that

$$\sum_{p \in S} \frac{1}{p} \le \log 4$$

If N is large enough, Hildebrand's theorem is applicable with x = N , K = $\sum_{p \in S} \frac{1}{p}$, giving

(2.5)
$$\sigma(S) \ge G(N, \sum_{p \in S} \frac{1}{p}) = \rho(\exp(\sum_{p \in S} \frac{1}{p}))(1 + O((\log N)^{-\gamma})) .$$

Therefore, from (2.4) and (2.5),

$$D(S) = \delta(S) - \sigma(S) \leq \exp(-\sum_{p \in S} \frac{1}{p}) - \rho(\exp(\sum_{p \in S} \frac{1}{p}))(1 + o(1)) ,$$

where the "o(1)" tends to 0 as $N \rightarrow \infty$. We conclude that

$$D' \leq \max\{\frac{1}{4}, \frac{1}{u_0} - \rho(u_0)\} = f(u_0)$$
,

which completes the proof of the theorem.

<u>Theorem 4</u>. $D_0 \le e^{-1} = 0.367879...$.

<u>Proof</u>. Let S be any set of pairwise coprime natural numbers. If $\delta(S) < e^{-1}$, then certainly $D(S) < e^{-1}$, so assume that $\delta(S) \ge e^{-1}$. We also may assume that D(S) > 0 so that there is some N with $\sigma(S) = \frac{1}{N} S \langle N \rangle$. Thus

$$\sigma(S) \ge 1 - \frac{1}{N} \sum_{m \in S} \left[\frac{N}{m} \right] \ge 1 - \sum_{m \in S} \frac{1}{m} .$$

Also,

$$e^{-1} \leq \delta(S) = \prod_{m \in S} (1-1/m) \leq \exp\{-\sum_{m \in S} \frac{1}{m}\},$$

so that

$$D(S) \leq \exp\{-\sum_{m \in S} \frac{1}{m}\} - 1 + \sum_{m \in S} \frac{1}{m}, \quad 0 \leq \sum_{m \in S} \frac{1}{m} \leq 1.$$

But the maximum value of $e^{-u} - 1 + u$ on [0,1] is at u = 1 which gives the value e^{-1} . This completes the proof of the theorem.

<u>Theorem 5</u>. If $S \subset P$, then there is a set S' with $S \subset S' \subset P$ such that

(i) $\sigma(S') = \sigma(S)$ (ii) D(S') = 0.

Proof. We first show the following lemma.

Lemma. Let $S_1 \subset S_2 \subset \ldots \subset \mathbb{P}$ and let $S = \bigcup S_i$. Then $\sigma(S) = \lim_{i \to \infty} \sigma(S_i)$.

<u>Proof</u>. We have $\sigma(S_1) \ge \sigma(S_2) \ge \ldots$, so the limit exists, call it σ . Since $\sigma(S_1) \ge \sigma(S)$ for each i, we have $\sigma \ge \sigma(S)$. Suppose $\sigma \ge \sigma(S)$. Then there is some N such that $\sigma \ge \frac{1}{N} S \langle N \rangle$. But $S \langle N \rangle = S_1 \langle N \rangle$ for all sufficiently large i. Thus $\sigma \ge \sigma(S_1)$ for some i, a contradiction.

We now turn to the proof of the theorem. Say $S \subset P$. We may suppose D(S) > 0, for otherwise let S' = S. Consider the set

$$\mathbf{F} = \{\mathbf{T}: \mathbf{S} \subset \mathbf{T} \subset \mathbf{P} , \ \sigma(\mathbf{S}) = \sigma(\mathbf{T})\} .$$

Let S' be any maximal element of F. (By the lemma, F has maximal elements.) It remains to show that D(S') = 0. Suppose not, so that D(S') > 0.

There exists some N so that if n > N, then

$$\frac{1}{n} \ \texttt{S'} \langle \texttt{n} \rangle \ \flat \ \sigma(\texttt{S'}) \ + \frac{1}{2} \ \texttt{D}(\texttt{S'}) \ .$$

Let p be any prime with

$$p \notin S'$$
, $p > max\{2/D(S'), N\}$.

We claim that S' U {p} is in F, contradicting the maximal choice of S' and thus proving the theorem. It will be sufficient to show that $\sigma(S') = \sigma(S' \cup \{p\})$. Say M is such that

$$\sigma(S' \cup \{p\}) = \frac{1}{M} (S' \cup \{p\}) \langle M \rangle .$$

Suppose $M \ge p$. Then

$$\sigma(S') < \frac{1}{M} S' \langle M \rangle - \frac{1}{2} D(S') < \frac{1}{M} S' \langle M \rangle - \frac{1}{p}$$
$$\leq \frac{1}{M} (S' \cup \{p\}) \langle M \rangle = \sigma(S' \cup \{p\}) \leq \sigma(S') .$$

a contradiction. Thus M < p and so evidently $\sigma(S' \cup \{p\}) = \sigma(S')$.

§3. Further problems.

Is D' = D? If not, then clearly there is some finite set of primes S with D = D(S). Perhaps a candidate for such a set S can be found numerically, but we have had no luck. We examined many sets S of primes and the largest discrepancy calculated was ≈ 0.245712 achieved for S the set of primes in the interval [19,12487]. Although it is not clear that a set of primes S with maximal discrepancy (should such a set S exist) must consist of all the primes in an interval, it was solely over such sets that we searched. As above, let S(a,b) denote the set of primes in the interval [a,b]. For each prime p , let b_p denote the prime which maximizes $D(S(p,b_p))$. From the proof of Theorem 3,

$$b_p = p^{u_0+o(1)}$$
, $D(S(p,b_p)) \rightarrow f(u_0)$ as $p \rightarrow \infty$,

where u_0 , $f(u_0)$ are given by (2.2). Must the convergence to $f(u_0)$ be from below? If not, then D' < D. For several small values of p we have computed candidate values of b_p and have approximated $D(S(p,b_p))$:

р 	b p	D(S(p,b _p))	log b _p /log p
2	13	0.129308	3.7004
3	113	0.173985	4.3031
5	719	0.204992	4.0871
7	1861	0.223270	3.8691
11	4759	0.235227	3.5313

Perhaps an example can be found to show $D'\, <\, D_O$, but we have not investigated this.

It is clear from the proof of Theorem 5 that actually a more general result is provable. Namely, given a set of primes S, then for any x with $\sigma(S) \leq x \leq \delta(S)$, there is a set of primes $S_X \supset S$ with $\sigma(S_x) = \sigma(S)$ and $\delta(S_x) = x$. Consider the set of points in \mathbb{R}^2

$$A = \{(\sigma(S), \delta(S)): S \subset \mathbf{P}\}.$$

It is clear that for each a, $0 \le a \le 1$, we have $(a,a) \in A$. Indeed, if S is a maximal set of primes with $\sigma(S) \ge a$ (from the lemma in the proof of Theorem 5, S exists), then $\sigma(S) = a$. For if $\sigma(S) \ge a$ and if p is any prime larger than $(\sigma(S) - a)^{-1}$ and not in S, then

$$\sigma(S \cup \{p\}) \ge \sigma(S) - \frac{1}{p} > a ,$$

contradicting S maximal. Also by the proof of Theorem 5, D(S) = 0, so $(a,a) \in A$.

From the above paragraph, for each $a, 0 \le a \le 1$, the set of b with $(a,b) \in A$ is an interval $[a,b_a)$ or $[a,b_a]$. Is the interval always closed? This would follow by showing that whenever

 $S_1, S_2, ...$ are minimal sets of primes with $\sigma(S_1) = a$, then $\sup{\delta(S_1)} = \delta(S_{1_0})$ for some i_0 .

The function $a \rightarrow b_a$ is continuous only at 0 and 1. Indeed, if a is irrational, $0 \le a \le 1$, then $b_a = a$. But for any $0 \le a \le 1$,

Indeed, let u be such that p(u) = a. Then from the methods of the proof of Theorem 3,

$$\lim_{x \to \infty} \sigma(S(x, x^{u})) = a, \qquad \lim_{x \to \infty} \delta(S(x, x^{u})) = u^{-1}$$

so that there is a sequence $t_n \rightarrow a$ with $b_{t_n} \rightarrow u^{-1} > a$.

What can be said about the set of a such that $(a,b) \in A$?

Is it always dense in the interval [ab,b] where

$$a_h = \inf\{\sigma(S):\delta(S) = b, S \subset P\}$$
?

If so, then D' = D. To see this, suppose D' < D and so D = D(S)for some $S \subset \mathbf{P}$. Let $b = \delta(S)$, so that $a_b = \sigma(S)$. Let $\{S_n\}$ be a sequence of sets of primes with $\delta(S_n) = b$ and $\sigma(S_n) \downarrow a_b$. Let $S_n^{\circ} \subset S_n$ be a minimal set of primes with $\sigma(S_n^{\circ}) = \sigma(S_n)$. Then $\delta(S_n^{\circ}) \ge b$, so that

$$\liminf_{n \to \infty} D(S_n^0) \ge b - a_b = D.$$

But since the sets S_n^o are mutually distinct, we must have $\{\max S_n^o\}$ unbounded, so that $D' \ge \liminf D(S_n^o) \ge D$, a contradiction.

Consider the function $b \rightarrow a_b$. It is not so hard to see that a_b is monotone non-decreasing and that $a_b < b$ whenever 0 < b < 1. Moreover, by monotonicity $b \rightarrow a_b$ is continuous at all but at most countably many b. Is it continuous at all b?

For $S \subset \mathbb{P}$ and D(S) > 0, can the set of N for which $\sigma(s) = \frac{1}{N} S\langle N \rangle$ be arbitrarily large? Let $N_1(S)$, $N_2(S)$, respectively, be the smallest, largest N with $\sigma(S) = \frac{1}{N} S\langle N \rangle$. Can $N_2(S)/N_1(S)$ be

arbitrarily large? Probably these questions can be answered in the affirmative by starting with a finite set of primes and adding some large primes so as to have the Schnirelmann density attained again. These questions seem to be much harder for minimal sets of primes. We can show that $N_1(S)/max S$ can be arbitrarily large. Indeed, if $S = [x, 2x] \cap \mathbf{P}$, then $N_1(S) \gg x^{1+c}$ for some c > 0. For if $N < x^2$, then

$$\frac{1}{N} S \langle N \rangle = 1 - \sum_{p \in S} \frac{1}{p} + \frac{1}{N} \sum_{p \in S} \left\{ \frac{N}{p} \right\}$$

where { } denotes the fractional part. Thus if $\sigma(S) = \frac{1}{N} S\langle N \rangle$ for some $N < x^2$, then $\frac{1}{N} \sum \{N/p\}$ is minimal for N in this range. It is easy to show that $\sum \{N/p\} x/\log x$ uniformly for $x \le N \le x^{1+c}$ using Hoheisel-type results on the distribution of primes in short intervals. Probably $N_1(S) \gg x^{2-\varepsilon}$ holds for each $\varepsilon > 0$. On the other hand it is easy to show $N_2(S) << x^{2+\varepsilon}$ for every $\varepsilon > 0$. In general, if S is the interval of primes $[x,y] \cap P$ and $|S| \ge 2$, then probably for $y \le x^2$ we have $N_1(S) = x^{2+o}(1) = N_2(S)$ and for $y > x^2$ we have $N_1(S) = y^{1+o}(1) = N_2(S)$. We can show that if $S = \{p,q\}$ where p > q are primes, then $N_1(S) \ge p(q-1)/(q-p)$, so that if $q - p < \log p$ (which can be arranged infinitely often), then $N_1(S) \ge p^2/\log p$.

Is it true that if $S \subset P$ is minimal and max S = p, then

 $\frac{1}{p} S \langle p \rangle - \sigma(S) \to 0 \quad \text{as} \quad p \to \infty$

or even

$$\frac{1}{p} S\langle p \rangle \sim \sigma(S) \text{ as } p \to \infty ?$$

These fail if $S \not\subset P$. For example, if S is the set of integers in (n,2n], then

$$\frac{1}{2n}$$
 S<2n> = $\frac{1}{2}$, $\frac{1}{4n}$ S<4n> ~ $\frac{11}{24}$.

The set of $\delta(S)$ where $S \subset \mathbf{P}$, S finite, is exactly the set of rationals of the form $\phi(n)/n$ where ϕ denotes Euler's function. There is a very simple decision procedure for membership in this set. What can be said about the set of $\sigma(S)$ where $S \subset \mathbf{P}$, S finite? Can any rational in (0,1] be shown to be not in this set? What if S is allowed to be any finite set of pairwise coprime integers? Is the membership problem for the set of $\delta(S)$ still decidable? Finally, what if S is any finite set of integers? Does every rational $r \in (0,1]$ satisfy $r = \sigma(S_1) = \delta(S_2)$ for some finite sets of integers S_1 , S_2 ?

Another line of research is to consider gaps between the elements in the set A(S) for various choices of S. This question and similar questions are studied in the papers [3], [6], [7], [8].

Finally we record the following old problem of the ageless first author. If S is a set of natural numbers, let

$$S'(x) = x - S(x)$$

for every natural number x. Thus $S'\langle x \rangle$ is the number of integers up to x divisible by some member of S. There are easy examples where $\frac{1}{x}$ S' $\langle x \rangle$ is very large and then drops drastically. For example, from the proof of Theorem 2,

$$\frac{1}{n} S'_{n} \langle n \rangle \approx 1 , \quad \lim_{x \to \infty} \frac{1}{x} S'_{n} \langle x \rangle \approx 0$$

for n large. The question is if the reverse can happen. That is, can $\frac{1}{n}$ S'(n) be small for some $n \ge \max S$, but $\frac{1}{x}$ S'(x) is large for some x > n? To quantify this question, we ask if it is always true that

$$\frac{1}{x}$$
 S' $\langle x \rangle < \frac{2}{n}$ S' $\langle n \rangle$

for any finite set of natural numbers S and $x > n \ge max S$? That "2" cannot be replaced by a smaller number can be seen by looking at the case S = {k}, n = 2k - 1, x = 2k.

References

- K. Dickman, On the frequency of numbers containing prime factors of a certain relative magnitude, Ark. Mat. Astr. Fys. 22(1930), No. A 10.
- P. Erdös, A generalization of a theorem of Besicovitch, J. London Math. Soc. 11(1936), 92-98.
- P. Erdös, On the difference of consecutive terms of a sequence defined by divisibility properties, Acta Arith. 12(1966/67), 175-182.

- 4. H. Halberstam and K.F. Roth, <u>Sequences</u>, Springer-Verlag, New York, 1983, Ch. 1.
- 5. A. Hildebrand, Quantitative mean value theorem for non-negative multiplicative functions, II, to appear.
- M.J. Narlikar, On a theorem of Erdös and Szemerédi, Hardy-Ramanujan J. 3(1980), 41-47.
- M.J. Narlikar and K. Ramachandra, Contributions to the Erdös-Szemerédi theory of sieved integers, Acta Arith. 38(1980/81), 157-165.
- E. Szemerédi, On the difference of consecutive terms of sequences defined by divisibility properties. II, Acta Arith. 25(1973/74), 191-197.