Generalization of some theorems on sets of multiples and primitive sequences

by

W. KLOTZ (Cunsthal)

1. Introduction. The main results of this paper are generalizations of a theorem of Besicovitch on primitive sequences and of a theorem of Davenport and Erdős on sets of multiples. For these theorems and a survey of related results we refer to the final chapter of Halberstam and Roth [3].

By a system $\sigma$ we mean a non-empty set of finite, non-empty sets of positive integers. The system $\sigma$ is called homogeneous, if for each $n \in \mathbb{N}$ (set of positive integers)

$$S \in \sigma \text{ implies } nS = \{ns : s \in S \} \in \sigma.$$  

The set $A \subset \mathbb{N}$ is said to be $\sigma$-free, if it does not contain a subset belonging to $\sigma$. For a given homogeneous system $\sigma$ we discuss the question of the 'greatest possible density' a $\sigma$-free set may have. We investigate natural densities and logarithmic densities of $\sigma$-free sets.

The author wishes to acknowledge his debt to the valuable advice of Professor E. Wirsing.

2. Natural densities of $\sigma$-free sets. First we introduce some notations. For real numbers $\alpha, \beta$ we define the interval $[\alpha, \beta] = \{n \in \mathbb{N}, \alpha \leq n \leq \beta\}$. If $A$ is a finite set, then $|A|$ denotes the number of elements in $A$. The counting function of $A \subset \mathbb{N}$ is $A(n) = \{A \cap [1, n]\}$. The limit $d(A) = \lim_{n \to \infty} A(n)/n$, if it exists, is called the natural density of $A$. The lower and upper natural densities $d(A)$ and $\bar{d}(A)$ are defined by the liminf and limsup of the same expression. The system $\sigma$ is characterized by

$$\tau_{\sigma}(n) = \max \{A(n) : A \text{ $\sigma$-free}\},$$

$$\tau_{\sigma} = \liminf_{n \to \infty} \tau_{\sigma}(n)/n, \quad \tau_{\sigma} = \limsup_{n \to \infty} \tau_{\sigma}(n)/n.$$  

If $\tau_{\sigma} = \tau_{\sigma}$ let $\tau_{\sigma} = \tau_{\sigma} = \tau_{\sigma}$. Furthermore, we define

$$d(\sigma) = \sup \{d(A) : A \text{ $\sigma$-free}\}, \quad \bar{d}(\sigma) = \sup \{\bar{d}(A) : A \text{ $\sigma$-free}\}.$$  

If $\bar{d}(\sigma)$ and $d(\sigma)$ coincide, the common value is denoted by $d(\sigma)$.
Every system $\sigma_0$ generates a homogeneous system $\sigma$,
$$
\sigma = N\sigma_0 = \{ T : T = nS, n \in N, S \in \sigma_0 \}.
$$
The investigation of a homogeneous system is facilitated by a small generating system.

**Theorem 1.** Suppose that the homogeneous system $\sigma$ is generated by
$$
\sigma_0 = \{ S_1, S_2, \ldots \}.
$$
Let $a_1, a_2, \ldots, a_k$ be coprime integers greater than 1 and
$$
U = \{ u : u = a_1^{e_1}a_2^{e_2} \cdots a_k^{e_k}, e_i \in \{ 0 \} \cup N \}.
$$
If $S_i \subset U$ for each $i$ then $r_\sigma(n)$ exists and
$$
r_\sigma(n) = r_\sigma(n) + O(\log^2 n).
$$
Here $r_\sigma$ is less than 1. There is a $\sigma$-free set $A$ with $d(A) = r_\sigma$.

**Proof.** Denote by $V$ the sequence of positive integers which are not a multiple of any of the numbers $a_j$. It is well-known that
$$
V(n) = \frac{n}{d(V)} + O(1), \text{ where } d(V) = \prod_{j=1}^{k} \left( 1 - \frac{1}{a_j} \right).
$$
We have
$$
\sum_{u \in U} \frac{1}{u} = \frac{1}{d(V)} \quad \text{and} \quad U(n) = \left( 1 + \frac{\log n}{\log 2} \right)^{k},
$$
Every positive integer has a unique representation of the form $n \nu$, $u \in U$, $u \in V$. Therefore, it follows from (1) and (2)
$$
n = \sum_{u \in U(n)} V\left( \frac{n}{u} \right) = n\frac{d(V)}{d(U(n))} + O\left( U(n) \right),
$$
$$
\sum_{u > n} \frac{1}{u} = O\left( \frac{U(n)}{n} \right) = O\left( \frac{\log^k n}{n} \right),
$$
where summation is taken over the numbers $u \in U$.

If we define
$$
r_\sigma'(n) = \max\{ A(n) : A \subset U, A \text{ $\sigma$-free} \},
$$
and if $R$ denotes the unique subset of $U$ having the counting function
$$
R(n) = r_\sigma'(n),
$$
then
$$
r_\sigma(n) = \sum_{u \in R} r_\sigma'(\frac{n}{u}) = \sum_{u \in R} \frac{n}{u}.
$$
Thus, by (1),
$$
r_\sigma(n) = \sum_{r \in R} V\left( \frac{n}{r} \right) = n\frac{d(V)}{d(n)} - \frac{n}{d(V)} \sum_{r > n} \frac{1}{r} + O(R(n)),
$$
where summation is over $r \in R$. By $R \subset U$, it now follows from (2) and (3) that
$$
r_\sigma(n) = r_\sigma(n) + O(\log^k n), \quad \text{where} \quad \tau_\sigma = \frac{d(V)}{d(n)} \sum_{r \in R} \frac{1}{r}.
$$
Since $\sigma$ and the sets $S_i$ are non-empty by definition, $R$ is a proper subset of $U$. We have
$$
\sum_{r \in R} \frac{1}{r} < \sum_{u \in U} \frac{1}{u} = \frac{1}{d(V)},
$$
whence $\tau_\sigma < 1$. The existence of a $\sigma$-free set $A$ with $d(A) = r_\sigma$ is ensured by the following lemma.

**Lemma 1.** Suppose that the homogeneous system $\sigma$ is generated by
$$
\sigma_0 = \{ S_1, S_2, \ldots \}.
$$
If $M = \{ x : x = \max S_i, i \in \sigma_0 \}$ has natural density 0, then there is a $\sigma$-free set $A$ with $d(A) = \tau_\sigma$.

**Proof.** Let $e_j (j = 1, 2, \ldots)$ be positive numbers satisfying $0 < e_j < 1$ and lim $e_j = 0$. There is a sequence of integers $x_j$ starting with $x_0 = 0$ and having the following properties for $j > 0$.

(a) $x_j > \frac{1}{e_j} x_{j-1},$

(b) $\tau_\sigma(x_j) > (\tau_\sigma - e_j) x_j,$

(c) if $T_{j-1} = \{ m \in M : n \in [1, x_{j-1}] \}$ then $T_{j-1}(x_j) < e_j x_j.$

Let $A_j$ be a $\sigma$-free set in $[1, x_j]$ with $A_j > (e_j - e_j) x_j$. Using the notation $B^{-1}O = \{ x : x \in B, x \in O \}$ we define
$$
A_j = A_j^{-1}\{ [1, x_{j-1}] \cup T_{j-1} \}, \quad A = \bigcap_{j=1}^{\infty} A_j.
$$
The sets $A_j$ are disjoint and $\sigma$-free. From (a), (b), (c) we obtain
$$
A(x_j) \geq A_j(x_j) > (\tau_\sigma - e_j) x_j,
$$
hence $d(A) \geq \tau_\sigma$. Assume now that $A$ contains a set $nS_i$. Let $d = \min S_i$ and $D = \max S_i$. Since the sets $A_j$ are $\sigma$-free, we must have
$$
n d A_h, \quad n D A_q, \quad k < q.
$$
From $n \leq x_k \leq x_{k-1}$ and $D \in M$ follows $n D A_{q-1}$, which contradicts the definition of $A_q$. Therefore, $A$ is $\sigma$-free.

**Lemma 2.** Suppose that the homogeneous system $\sigma$ is generated by
$$
\sigma_0 = \{ S_1, S_2, \ldots \}.
$$
Let $d_j = \min S_j$ and $D_j = \max S_j$. If $\lim_{j \to \infty} d_j/D_j = 0$ then $\tau_\sigma$ exists.

**Proof.** Denote by $\sigma_j$ the homogeneous system generated by $\{ S_1, S_2, \ldots, S_j \}$. By Theorem 1, the density $\tau_{\sigma_j}$ exists. Moreover, lim $\tau_{\sigma_j} = \tau_\sigma$. 

2 - Acta Arithmetica XXXII, 1
exists, because \( r_\delta \geq r_{\delta, +1} \). If for positive \( \varepsilon \) the integer \( j \) is chosen so large that \( d_\delta / B_j < \varepsilon \) for each \( k > j \) then

\[
\tau_j(n) - \varepsilon n \leq \tau_n(n) \leq \tau_j(n), \quad \tau_j - \varepsilon \leq \tau_n \leq \tau_j.
\]

For \( j \to \infty \) and \( \varepsilon \to 0 \) we obtain \( \tau_\delta = \tau_\varepsilon = \tau \).

We are now going to state the announced generalization of a theorem of Besicovitch ([3], p. 257) on primitive sequences. We denote by \( NG \) the set of multiplicative \( \{ n \in \mathbb{N}, \gcd \neq 0 \} \).

**Theorem 2.** Let the homogeneous system \( \sigma \) be generated by \( \sigma_0 = \{ S_1, S_2, \ldots \} \). Suppose that there is a sequence \( G = \{ g_1, g_2, \ldots \} \) of positive integers satisfying

1. \( S_i \cap G = \emptyset \) for each \( j \),
2. \( \lim_{j \to \infty} \delta(NG_0) = 0 \) if \( G_0 = \{ g_1, g_2, \ldots \} \).

Then \( \tau_\sigma \) exists and \( \delta(\sigma) = \tau_\sigma \). Furthermore, \( \tau_\sigma = 0 \) is equivalent to \( (1) \epsilon \sigma \), and \( 1 + \epsilon \) implies \( \delta(\sigma) > 0 \).

Proof. We make use of the following lemma which is easily deduced from an inequality of Behrend ([3], p. 263).

**Lemma 3 (Erdös [2]).** If \( 1 + \epsilon \) and \( \lim_{j \to \infty} \delta(NG_j) = 0 \) then \( \delta(NG) \) exists and is less than 1.

We note that \( N \setminus NG \) is \( \epsilon \) free by (1). So Lemma 3 implies \( \delta(\sigma) > 0 \) if \( 1 + \epsilon \). Now suppose \( 1 + \epsilon \) and \( G' = G \setminus \{ 1 \} \) then

\[
( N \setminus NG') \cap [n/2, n]
\]

is \( \epsilon \) free for each \( n \in \mathbb{N} \) if \( 1 + \epsilon \). In this case Lemma 3 implies \( \tau_\sigma > 0 \).

It remains to prove the existence of \( \tau_\sigma \) and \( \delta(\sigma) = \tau_\sigma \). By (ii), we may assume that \( G \) is finite, \( G = \{ g_1, \ldots, g_l \} \). Then the existence of \( \tau_\sigma \) follows either from Theorem 1 or Lemma 2.

**Lemma 4.** If \( m \) is any positive real number then

\[
\lim_{n \to \infty} \delta(N[x/m, x]) = 0.
\]

This is an immediate consequence of a theorem of Erdös ([3], p. 268). To construct a \( \epsilon \) free set \( A \) with \( \delta(A) \geq \tau_\sigma - \varepsilon \) \( (0 < \varepsilon < 1) \), we choose

\[
m = \frac{3}{\varepsilon} g_1, \quad \varepsilon_j = \left\lceil \frac{\varepsilon}{3} \right\rceil (j = 0, 1, 2, \ldots).
\]

There is a sequence of integers \( x_j \) starting with \( x_0 = 0 \) and having the following properties for \( j > 0 \):

1. \( x_j > m x_{j-1} \),
2. \( x_j(x_j) > \tau_\sigma - \frac{\varepsilon}{3} x_j \),
3. if \( B_j = \bigcup_{x_j/m, x_j} \), then \( \delta(B_j) < \varepsilon_j \) and \( B_{j-1} \) \( (x) < \varepsilon_{j-1} x \) for each \( x \geq x_j \).

Let \( A_j \) be a \( \sigma \)-free set in \( [1, a_j] \) with \( A_j > (\tau_\sigma - \varepsilon) x_j \). Define

\[
A_j = A_j^c \setminus \left( \bigcup_{x \in [1, \varepsilon x_j]} \left( \bigcup_{x \in B_j} \right) \right), \quad A = \bigcup_{j=1}^\infty A_j.
\]

The sets \( A_j \) are disjoint and \( \sigma \)-free. From (a), (b), (c) we obtain

\[
A(x_j) \geq A_j(x_j) > (\tau_\sigma - \varepsilon) x_j > (\tau_\sigma - \varepsilon) \delta(\sigma),
\]

hence \( \delta(\sigma) \geq \tau_\sigma - \varepsilon \). To prove that \( A \) is \( \sigma \)-free, assume that \( A \) contains a set of the form \( nS_k, S_k \in \sigma \). Let \( d = \min S_k \) and \( D = \max S_k \). Since the sets \( A_j \) are \( \sigma \)-free, we must have

\[
\text{nd} \in A_k, \quad \text{nd} \in A_k, \quad k < q.
\]

By (i), \( S_k \) contains a number \( g \in G \). Thus \( ng \in A \) and, by (a), (5), and (7),

\[
ng = nd \leq aw, \quad w = \varepsilon, \quad \delta(\sigma) < \delta(\sigma) < \delta(\sigma).
\]

Now (6) implies \( ng \in A_k \). So we have

\[
x_j \delta(\sigma) < \delta(\sigma) \leq x_j \delta(\sigma) \leq x_j \delta(\sigma) \leq x_j \delta(\sigma).
\]

Therefore, \( nd \in B_k \), which contradicts the definition of \( A_k \).

**3. Logarithmic densities of \( \sigma \)-free sets.** For a homogeneous system \( \sigma \) and the natural density \( \delta(\sigma) \), we need not exist. Example 2 below shows that even for a finitely generated system \( \delta(\sigma) \) may be less than \( \delta(\sigma) \). More uniform results are obtained by considering logarithmic densities. We introduce the following logarithmic notions in analogy to the corresponding notions on natural density.

The logarithmic counting function of \( A \subset \mathbb{N} \) is \( A^n(\varepsilon) = \sum_{n=1}^{\infty} a(\varepsilon) \) (summation over \( a \in \mathbb{A} \)). The limit \( A(\varepsilon) = \lim \sum_{n=1}^{\infty} a(\varepsilon) \) is called the logarithmic density of \( A \). The lower and upper logarithmic densities \( \delta(\sigma) \) and \( \delta(\sigma) \) are defined by the limitinf and limsup of the same expression. Let

\[
\lambda_\sigma(n) = \max \{ A^n(\varepsilon) : A \text{ \( \sigma \)-free} \},
\]

\[
\lambda_\sigma = \lim \inf_{n \to \infty} \lambda_\sigma(n), \quad \lambda_\sigma = \lim \sup_{n \to \infty} \lambda_\sigma(n).
\]

If \( \lambda_\sigma = \lambda_\sigma \), put \( \lambda_\sigma = \lambda_\sigma \). Define

\[
\delta(\sigma) = \sup \{ \delta(\sigma) : A \text{ \( \sigma \)-free} \}, \quad \delta(\sigma) = \sup \{ \delta(\sigma) : A \text{ \( \sigma \)-free} \},
\]

If \( \delta(\sigma) = \delta(\sigma) \), denote the common value by \( \delta(\sigma) \).
We believe that on very general conditions for a homogeneous system \( \delta(\sigma) \) and \( \lambda_\sigma \) exist and coincide.

**Theorem 3.** Suppose that the homogeneous system \( \sigma \) is generated by \( \sigma_0 = \{S_1, S_2, \ldots \} \). Let \( \alpha_1, \alpha_2, \ldots, \alpha_n \) be coprime integers greater than 1 and
\[
U = \{ u : u = \alpha_1^{r_1} \alpha_2^{r_2} \ldots \alpha_n^{r_n}, r_i \in \{0 \} \cup N \}.
\]
If \( S_i \subseteq U \) for each \( i \) then \( \lambda_\sigma \) exists and
\[
\lambda_\sigma(n) = \lambda_\sigma \log n + O(\log \log n).
\]
Furthermore, \( \delta(\sigma) \) exists and \( \bar{d}(\sigma) = \delta(\sigma) = \lambda_\sigma \). If \( k = 1 \) there is a \( \sigma \)-free set \( A \) satisfying
\[
A^*(n) = \lambda_\sigma(n) \quad \text{and} \quad A(n) = \lambda_\sigma n + O(\log n).
\]

Proof. This proof is similar to that of Theorem 1. If we denote by \( V \) the sequence of positive integers which are not a multiple of any of the numbers \( a_j \) then
\[
V^*(n) = \bar{d}(V) \log n + O(1), \quad \text{where} \quad \bar{d}(V) = \prod_{j=1}^{k} \left(1 - \frac{1}{a_j}\right).
\]

Define
\[
\lambda_\sigma^V(n) = \max \{ A^*(n) : A \subseteq U, A \ \sigma \text{-free} \}.
\]
By (3), the limit \( \lim_{n \to \infty} \lambda_\sigma^V(n) = A \) exists and
\[
(9) \quad \lambda_\sigma^V(n) = a + O \left( \frac{\log^k n}{n} \right).
\]
Now we have, by (8) and (9),
\[
(10) \quad \lambda_\sigma(n) = \bar{d}(V) \log n + \left( \sum_{v \in V} \frac{1}{\lambda_\sigma^V \left( \frac{n}{v} \right)} \right) + O \left( \sum_{v \in V} \frac{\log \log(n/v)v}{v} \right),
\]
where \( v \in V \). Let \( m = \log^k n \) and \( f(y) = \log^k y \). Assuming that \( f(y) \) is strictly decreasing for \( y \geq m \), we obtain, by (8),
\[
\left( \sum_{v \in V \cap n / v} \frac{1}{\lambda_\sigma^V \left( \frac{n}{v} \right)} \right) + \left( \sum_{n / m \in V \cap n / v} \frac{1}{f \left( \frac{n}{v} \right)} \right) = O \left( f(m) \log n + \log m \right).
\]
Hence, by (10),
\[
(11) \quad \lambda_\sigma(n) = \lambda_\sigma \log n + O(\log \log n), \quad \text{where} \quad \lambda_\sigma = \bar{d}(V).
\]

To prove \( \bar{d}(\sigma) = \delta(\sigma) = \lambda_\sigma \), we construct a \( \sigma \)-free set \( A \) with natural density \( \lambda_\sigma \) greater than \( \lambda_\sigma - \epsilon \). Let \( U' \) be a finite \( \sigma \)-free subset of \( U \) satisfying
\[
\sum_{u \in U'} \frac{1}{u} > a - \epsilon.
\]
If \( A = U' V \) then \( A \) is \( \sigma \)-free and
\[
A(n) = \sum_{u \in U'} V \left( \frac{n}{u} \right) = n \bar{d}(V) \sum_{u \in U'} \frac{1}{u} + O(1).
\]
Therefore, \( \bar{d}(A) \) exists and \( \bar{d}(A) > (a - \epsilon) \bar{d}(V) > \lambda_\sigma - \epsilon \).

If \( h = 1 \) let \( a = a_1 \) and \( U_j = \{a^1, a^1, \ldots, a^j\} \), \( U_{-1} = \emptyset \). Define \( S \subseteq U \) by the following property:
\[
a' \in S \text{ if and only if } (S \cap U_{j-1}) \cup \{a'\} \text{ is } \sigma \text{-free } (j \geq 0, \ldots).
\]
Let \( S_j = S \cap U_j \). We prove by induction that \( S_j \) is the only \( \sigma \)-free set in \( U_j \) with \( S_j = S_j^*(a') \). This is certainly true for \( j = 0 \). Let it be true for \( j-1 \) \( (j \geq 1) \). Suppose now that \( M \) is a \( \sigma \)-free subset of \( U_j \), \( M \cap U_{j-1} \neq S_{j-1} \), then
\[
M^*(a') \leq \lambda_\sigma^V(a'^{j-1}) - \frac{1}{a'^{j-1}} + \frac{1}{a^j} < S_j^*(a').
\]
Hence \( S_j^*(a') = \lambda_\sigma^V(a') \) for \( j = 0, 1, \ldots \). If \( A = S V \) then \( A \) is \( \sigma \)-free, \( A^*(n) = \lambda_\sigma(n) \), and
\[
A(n) = \sum_{v \in S} V \left( \frac{n}{v} \right) = \lambda_\sigma n + O(\log n), \quad \text{where} \quad \lambda_\sigma = \bar{d}(V) \sum_{v \in S} \frac{1}{v}.
\]

**Lemma 5.** Let the homogeneous system \( \sigma \) be generated by \( \sigma_0 = \{S_1, S_2, \ldots \} \).
Let \( G = \{g_1, g_2, \ldots \} \) be a sequence of positive integers, \( G_j = \{g_1, g_2, \ldots, g_j\} \), and \( \sigma_j \) the homogeneous system generated by \( \{S_1, \ldots, S_j\} \). Suppose
(i) \( \lim_{j \to \infty} \delta(NG_j) = 0 \);
(ii) \( \delta(\sigma_j) \) and \( \lambda_\sigma \) exist and \( \delta(\sigma_j) = \lambda_\sigma \) for each \( j \geq N \).

Then \( \delta(\sigma) \) and \( \lambda_\sigma \) exist and \( \delta(\sigma) = \lambda_\sigma = \lim_{j \to \infty} \delta(\sigma_j) \). If, in addition to (i) and (ii), \( \lim_{j \to \infty} \bar{d}(NG_j) = 0 \) and \( \bar{d}(\sigma_j) = \delta(\sigma_j) = \lambda_\sigma \), for each \( j \geq N \) then
\[
\bar{d}(\sigma) = \delta(\sigma) = \lambda_\sigma.
\]

Proof. Since \( \sigma_1 \subseteq \sigma_2 \subseteq \ldots \subseteq \sigma_j \), the limit \( \lim_{j \to \infty} \lambda_\sigma_j = \lambda \) exists and \( \lambda_\sigma \leq \lambda \).
Let \( j \geq N \) and \( \lim_{j \to \infty} \lambda_j = \lambda_j \). By (ii), there is a \( \sigma_j \)-free set \( A_j \) with \( \delta(A_j) > \lambda_\sigma - \epsilon_j \). The set \( A_j = A_j \cap NG_j \) is \( \sigma \)-free and \( \delta(A_j) > \lambda_\sigma - \epsilon_j - \delta(NG_j) \). For \( j \to \infty \) follows \( \delta(\sigma) \geq \lambda \). Hence \( \delta(\sigma) \) and \( \lambda_\sigma \) exist and \( \delta(\sigma) = \lambda_\sigma \).

If \( \bar{d}(\sigma_j) = \delta(\sigma_j) = \lambda_\sigma \) and \( \lim_{j \to \infty} \bar{d}(NG_j) = 0 \) then we may demand \( \bar{d}(A_j) > \lambda_\sigma - \epsilon_j \). Now we have \( \bar{d}(A_j) > \lambda_\sigma - \epsilon_j - \bar{d}(NG_j) \), and for \( j \to \infty \) follows the final part of Lemma 5.
If \( A = \{ a_1, a_2, \ldots \} \) is a sequence of positive integers let \( A_j = \{ a_j, a_{j+1}, \ldots \} \) and \( A_j^c = A \setminus A_j \). It has been proved by Davenport and Erdős ([3], p. 258) that the logarithmic density \( \delta(NA) \) exists and
\[
\delta(NA) = \delta(NA) = \lim_{f \to \infty} \delta(NA_f).
\]
Note that
\[
(12) \quad \lim_{f \to \infty} \delta(NA_j \setminus NA_j^c) = 0.
\]

**Lemma 6.** Suppose that the homogeneous system \( \sigma \) is generated by \( \sigma' = \{ S_1, S_2, \ldots, S_q \} \). Then for any homogeneous subsystem \( \sigma' \subset \sigma \) the densities \( \delta(\sigma) \) and \( \lambda_\sigma \) exist and coincide.

**Proof.** Any homogeneous subsystem \( \sigma' \) is of the form:

\[
\sigma = \{ S : S = a_1 S_i, 1 \leq i \leq q, 1 \leq k < \infty \}, \quad a_1 < \cdots < a_q < \infty.
\]

Let \( A_1 = \{ a_1, a_2, \ldots \} \) and \( A_q = \{ a_q, a_{q+1}, \ldots \} \). According to (12), for \( \varepsilon > 0 \) the number \( j \) can be chosen so large that
\[
\delta(NA_j \setminus NA_j^c) < \varepsilon / g \quad \text{for each} \quad i = 1, \ldots, q.
\]

Denote by \( \sigma_j \) the homogeneous system generated by
\[
\{ S : S = a_1 S_i, 1 \leq i \leq q, 1 \leq k < \varepsilon \}.
\]

By Theorem 3, \( \delta(\sigma_j) \) and \( \lambda_{\sigma_j} \) exist and \( \delta(\sigma_j) = \lambda_{\sigma_j} \). Hence there is a \( \sigma_j \)-free set \( H_j \) with \( \delta(H_j) > \lambda_{\sigma_j} \). If \( \varepsilon < \varepsilon / g \), the set
\[
H_j = H_j \setminus \bigcup_{i=1}^{q} t_i(NA_j \setminus NA_j^c)
\]
is \( \sigma \)-free and, by (13),
\[
\delta(\sigma) \geq \delta(H_j) > \lambda_{\sigma_j} - 2\varepsilon.
\]

Since \( \sigma_1 < \sigma_2 < \cdots < \sigma \) the limit \( \lim_{f \to \infty} \lambda_{\sigma_j} = \lambda \) exists and \( \lambda_\sigma < \lambda \). Now, on letting \( j \to \infty \) and \( \varepsilon \to 0 \) in (14), we see that \( \delta(\sigma) \) and \( \lambda_\sigma \) exist and \( \delta(\sigma) = \lambda_\sigma = \lambda \).

Finally, we are going to extend Lemma 6 by Lemma 5. Let \( G = \{ g_1, g_2, \ldots \} \) and \( G_j = \{ g_j, g_{j+1}, \ldots \} \). We shall say that \( G \) has property \( P \) if \( \lim \delta(NG_j) = 0 \).

**Theorem 4.** Suppose that the homogeneous system \( \sigma \) is generated by \( \sigma' = \{ S_1, S_2, \ldots \} \). Let \( G \) be a sequence with property \( P \) and \( M_j = \bigcup \{ S : S \in \sigma', S \cap NG_j = \emptyset \} \). If each set \( M_j \) has property \( P \), then for any homogeneous subsystem \( \sigma' \subset \sigma \) the densities \( \delta(\sigma) \) and \( \lambda_{\sigma} \) exist and coincide.

**Proof.** If we denote by \( \sigma_j \) and \( \sigma'_j \) the homogeneous systems generated by
\[
\{ S : S \in \sigma', S \cap NG_j = \emptyset \} \quad \text{and} \quad \{ S : S \in \sigma'_j, S \cap NG_j = \emptyset \}
\]
then \( \sigma_j \subset \sigma'_j \). Suppose
\[
(\sigma_j) \text{ and } \lambda_{\sigma'} \text{ exist and } \delta(\sigma_j) = \lambda_{\sigma'} \text{ for each } j \in \mathbb{N}, \text{ then the existence of } \delta(\sigma) \text{ and } \lambda_{\sigma}, \text{ and } \delta(\sigma) = \lambda_{\sigma}, \text{ follow from Lemma 5. If each set } M_j \text{ is finite, then } (a) \text{ is true by Lemma 6.}
\]

(b) Theorem 4 is true, if each set \( M_j \) is finite. In the general case (a) follows from (b) applied to \( \sigma'_j \).

**4. Examples**

**Example 1.** Let \( \sigma \) consist of the solutions in positive integers of the equation
\[
(15) \quad a_1 + a_2 + \ldots + a_{2k} = (y_1 + y_2 + \ldots + y_{2k}).
\]

Clearly, the interval \( [n/2, n] \) is \( \sigma \)-free. Therefore, \( \tau_\sigma \geq 1/2 \). We prove
\[
\delta(\sigma) = 1/2, \quad \text{where } r = \min \{ x : x \in \mathbb{N}, x \equiv 2k \}.
\]

Obviously, the congruence class 1 modulo \( r \) is \( \sigma \)-free. Hence \( \delta(\sigma) \geq 1/2 \). Let \( A \subset N \) be \( \sigma \)-free. By equating some of the variables in (15) it follows that the equation
\[
(16) \quad a_1 + a_2 + \ldots + a_{2k} = (y_1 + y_2 + \ldots + y_{2k})
\]
has no solution in \( A \), if \( j \) divides \( 2k \), thus especially for \( j = 1, \ldots, r - 1 \). For \( x_i = y_1, \ldots, x_j = y_{j-1} \) the last equation becomes
\[
(17) \quad a_1 + \ldots + a_{j-1} + 2y_j \quad (j = 2, 3, \ldots, r - 1).
\]

By (16), \( a_1 + a_2 = 2(y_1 + y_2) \) has no solution in \( A \). For \( a_1 = a_2 \) this means that
\[
(18) \quad a_1 = y_1 + y_2
\]
is also unsolvable in \( A \). Let \( a \in A \). Substituting \( y_1 = y_2 = \ldots = y_j = a \) in (17) and \( y_j = a \) in (18) we see that none of the equations
\[
(19) \quad x_1 = y_1 + ja \quad (j = 1, 2, \ldots, r - 1)
\]
has a solution in \( A \). Hence \( \delta(A) \leq 1/2 \).

It would be interesting to know whether the logarithmic density \( \delta(\sigma) \) exists for every homogeneous system defined by a linear equation.

**Example 2.** We construct a finitely generated homogeneous system \( \sigma \) with \( \delta(\sigma) < \delta(\sigma) \). Suppose that \( a \) is a positive integer not equal to 1. Let \( \sigma \) consist of all 3-term geometric progressions of ratio \( a, a^2, a^3 \) or \( a^4 \). This system is generated by
\[
\{1, a, a^2\}, \quad \{1, a^2, a^3\}, \quad \{1, a^3, a^4\}, \quad \{1, a^4, a^5\}
\]
We determine \( \delta(\sigma) = \tau_\sigma \) and \( \delta(\sigma) = \lambda_\sigma \) according to the considerations to Theorem 1 and Theorem 3. By (4) and (11), we have
\[
(20) \quad \delta(\sigma) = \delta(V) \sum_{n \in \mathbb{N}} 1/r, \quad \delta(\sigma) = \alpha\delta(V),
\]
where \( R \) is the set satisfying \( R(n) = \tau^*_U(n) \) and \( \alpha = \lim_{n \to \infty} \tau^*_U(n) \). We determine \( \tau^*_U(n) = S^*(n) \) as indicated in the final part of the proof of Theorem 5. Thus we obtain, by (19),

\[
\overline{d}(\sigma) = \left( 1 - \frac{1}{a} \right) \left( 1 + \frac{1}{a} + \frac{1}{a^2} + \frac{1}{a^3} + \frac{1}{a^4} + \cdots \right),
\]

\[
\overline{d}(\sigma) = \left( 1 - \frac{1}{a} \right) \left( 1 + \frac{1}{a} + \frac{1}{a^2} + \frac{1}{a^3} + \cdots \right) \sum_{\nu \in \mathbb{N}} a^{\nu p} < \overline{d}(\sigma).
\]

Example 3. Denote by \( \mathcal{C} = \{ c_1, c_2, \ldots \} \) the sequence of integers greater than 1, which are a product of at most \( k \) primes (multiple factors counted multiply). Define \( \mathcal{S}_j = \{ c, \sigma \} \), and let \( \sigma \) be the homogeneous system generated by \( \sigma_0 = \{ \delta_1, \delta_2, \ldots \} \).

By Lemma 1 and Lemma 2, \( \tau_0 \), and there is a \( \sigma \)-free set \( A \)

with \( \overline{d}(A) = \tau_0 \). Since \( \eta_0 \geq 2 \) for each \( j \in \mathbb{N} \), we have \( \tau_0 \geq \frac{1}{2} \). Let us prove

\[
\delta(\sigma) = \lambda(\sigma) = \frac{1}{k+1}.
\]

Suppose that \( A \) is a \( \sigma \)-free set in \([1, n]\) satisfying \( A^*(n) = \lambda(n) \). We sketchily follow the words of Halberstam and Roth ([1], pp. 246-249) for a proof of Behrend’s theorem on primitive sequences.

\[
A^*(n) = \lambda(n) = \frac{1}{n} \sum_{u \in \mathbb{N}} r(u) + O(1),
\]

where \( r(u) \) is the number of divisors of \( u \) belonging to \( A \). Let \( u \) be a product of \( s(u) \) primes. According to de Bruijn, Tengbergen, and Kruijswijk [1], the set of divisors of \( u \) can be completely divided into \( \left( \frac{s(u)}{[s(u)]/2} \right) \) disjoint symmetrical chains. A symmetrical chain of \( m \) divisors cannot contain more than \( \frac{m}{k+1} + 1 \) numbers of \( A \). Therefore, if \( d(u) \) is the number of divisors of \( u \),

\[
r(u) \leq \frac{d(u)}{k+1} + \left( \frac{s(u)}{[s(u)]/2} \right)
\]

and, by (21),

\[
\lambda(n) \leq \frac{1}{n} \sum_{u \in \mathbb{N}} d(u) + O \left( \frac{1}{n} \sum_{u \in \mathbb{N}} d(u) \right),
\]

\[
\lambda(n) \leq \frac{1}{n} \sum_{u \in \mathbb{N}} d(u) + O \left( \frac{\log n}{(\log \log n)^{1/2}} \right),
\]

(22)

On the other hand, if \( A = \{ a : \sigma > 1, s(a) = 1 \mod (k+1) \} \) then \( A \) is \( \sigma \)-free, and it follows as before that

\[
A^*(n) \geq \frac{\log n}{k+1} + O \left( \frac{\log n}{(\log \log n)^{1/2}} \right).
\]

By (22) and (23), we obtain (20). Note that the constants involved in the \( O \)-estimates of (22) and (23) can be chosen independent of \( k \).

Example 4. Let \( \sigma \) consist of all \( n \)-term geometric progressions \( (n \geq 3, \text{ rational ratio}) \). Systems of this kind have been investigated by Rankin [4] and by Riddell [5]. The system \( \sigma \) is generated by

\[
\sigma_0 = \{ a \in A : a = a^{p-1}, a^{p-2}b, \ldots, b^{p-1} \}, a < b, (a, b) = 1 \}.
\]

Let \( \mathcal{G} = \{ a^{p-1}, a^{p-2}b, \ldots, b^{p-1} \} \). Since \( \sum \frac{1}{a^{n-1}} \) converges, we have

\[
\lim_{f \to \infty} \mathcal{N}(\mathcal{G}) = 0.
\]

By Theorem 2, \( \tau_0 \) exists, and from Lemma 1 follows the existence of a \( \sigma \)-free set \( A \) with \( \overline{d}(A) = \tau_0 \). Lemma 5 and Theorem 3 ensure the existence of \( \delta(\sigma) \) and \( \lambda(\sigma) \). Moreover, \( \overline{d}(\sigma) = \delta(\sigma) = \lambda(\sigma) \).

Suppose that \( B \subset \{ e \} \times \mathbb{N} \) is a set which does not contain an \( n \)-term arithmetic progression. Let \( A \) consist of those positive integers which have in their unique prime factorization only exponents belonging to \( \mathcal{E} \).

Then \( A \) is \( \sigma \)-free, \( d(A) \), exists, and

\[
\overline{d}(\sigma) \geq \overline{d}(A) = \prod_{p} \left( \frac{1}{1 - \frac{1}{p}} \right) \sum_{n \in \mathbb{N}} \frac{1}{p^n}
\]

where the product is taken over all primes. As in the proof of Theorem 3, it follows by induction that \( \sum \frac{1}{a^{n-1}} \) is maximal if and only if \( B \) is identical with the set \( E_n \) defined by the following property:

\[
r \neq E_n \text{ if and only if } (E_n \cap \{0, r-1\}) \cup \{ r \} \text{ does not contain an } n \text{-term arithmetic progression } (r \in \{0\} \times \mathbb{N}).
\]

The estimates of Rankin and Riddell obtained by (24) can be improved for \( n \geq 4 \), because they use a set \( B \neq B_n \). If \( n \) is a prime, then it follows from a paper of Scheidig [6] that \( E_n \) consists of the nonnegative integers, which have no digit \( n-1 \), when they are expressed in the scale of \( n \). In this case we have

\[
\overline{d}(\sigma) \geq \prod_{p} \left( \frac{1}{1 - \frac{1}{p}} \right) \sum_{n \in \mathbb{N}} \left( \frac{1}{p^{n-1}} \right) \geq \prod_{p} \left( \frac{1}{1 - \frac{1}{p}} \right) \sum_{n \in \mathbb{N}} \frac{1}{p^{n-1}}
\]

\[
= \prod_{p} \left( \frac{1}{1 - \frac{1}{p}} \right) \sum_{n \in \mathbb{N}} \frac{1}{\zeta(n-1)} \zeta(n)^{-1}.
\]
Some remarks on Goldbach’s problem

by

AKIO FUJII (Tokyo)

In this paper we shall prove by a modification of Chen’s work ([3]) that every sufficiently large even integer \( x \) is written as a sum of a prime and a natural number which has at most one prime factor less than \( \sqrt[100]{x/100} \).

1. Let \( x \) be a large even integer. Let \( G_2(x) \) be the number of primes \( p \leq x \) such that \( x - p \) has at most two prime factors. Chen ([3]) has proved that

\[
G_2(x) \geq \frac{0.67x \phi(x^2)}{(\log x)^2}, \quad \text{where} \quad \phi(x^2) = \prod_{p \leq x} \frac{p - 1}{p - 1},
\]

In fact, if we put \( G_2(x, I) \) — the number of primes \( p \leq x \) such that \( x - p \) is a prime or \( x - p = r_1r_2 \) with primes \( r_1 \) and \( r_2 \) satisfying \( r_1I \) and \( p_1 \leq p_2 \), for a subset \( I \) of \( (1, \sqrt{\phi(x^2)}) \), he has proved that \( G_2(x, I) \geq 0.67x \phi(x^2)/(\log x)^2 \). (Further Halberstam [6] or [7] has shown that 0.67 can be replaced by 0.689.) Now we wish to maximize \( I \) in \( (1, \sqrt{\phi(x^2)}) \) such that \( G_2(x, I) \geq Ax \phi(x^2)/(\log x)^2 \), where \( A \) is some positive absolute constant. To study this we use the following mean value theorem which is similar to Bombieri’s one.

**Lemma 0.** Assume that \( M + N < x^{1/4} \). For an arbitrarily large constant \( A \), there exist positive constants \( B = B(A) \) and \( E = E(A) \) such that if \( M \geq (\log x)^A \), and \( b(m) \leq (\log x)^2 \), with some positive constant \( C \) for any \( m \) in \( M < m < M + N \), then

\[
\sum_{d \leq x^{1/3}(\log x)^2} \max_{1 < n \leq M+N} \frac{1}{d} \left( \frac{A(n) - 1}{\phi(d)} \right) \frac{y}{m} \leq \frac{x}{(\log x)^A},
\]

where \( 0 \) is an arbitrarily given positive number, \( n = am^*(d) \) means \( n = am^*(d) \), and \( m^* \) satisfies \( nm^* = 1(d) \).