Arithmetic Progressions in Lacunary Sets

T. C. Brown and A. R. Freedman


Abstract

We make some observations concerning the conjecture of Erdős that if the sum of the reciprocals of a set $A$ of positive integers diverges, then $A$ contains arbitrarily long arithmetic progressions. We show, for example, that one can assume without loss of generality that $A$ is lacunary. We also show that several special cases of the conjecture are true.

1 Introduction

The now famous theorem of Szemerédi [7] is often stated:

(a) *If the density of a set $A$ of natural numbers is positive, then $A$ contains arbitrarily long arithmetic progressions.*

Let us call a set $A$ of natural numbers $k$-good if $A$ contains a $k$-term arithmetic progression. Call $A$ $\omega$-good if $A$ is $k$-good for all $k \geq 1$. We define four density functions as follows: For a set $A$ and natural numbers $m, n$, let $A[m, n]$ be the cardinality of the set $A \cap \{m, m + 1, m + 2, \ldots, n\}$. Then define

$$\delta(A) = \liminf_n \frac{A[1, n]}{n},$$

$$\overline{\delta}(A) = \limsup_n \frac{A[1, n]}{n},$$

$$u(A) = \liminf_n \min_{m \geq 0} \frac{A[m + 1, m + n]}{n}$$

and

$$\overline{u}(A) = \limsup_n \max_{m \geq 0} \frac{A[m + 1, m + n]}{n}.$$ 

It can be seen that the limits in the definitions of $u$ and $\overline{u}$ always exist. These four “asymptotic” set functions are called the lower and upper “ordinary” and the lower and upper “uniform” density of the set $A$ respectively. They are related by

$$u(A) \leq \delta(A) \leq \overline{\delta}(A) \leq \overline{u}(A)$$

for any set $A$. 

1
Szemerédi actually proved:

(b) If \( \overline{\sigma}(A) > 0 \), then \( A \) is \( \omega \)-good. Hence we also have

(c) If \( \overline{\sigma}(A) > 0 \), then \( A \) is \( \omega \)-good.

In fact, Szemerédi proved the following “finite” result (which we state in a general form to be used later):

(d) Let \( \varepsilon > 0 \) and \( k \in \mathbb{N} = \{1, 2, 3, \ldots \} \). Then there exists an \( n_0 \in \mathbb{N} \) such that if \( P \) is any arithmetic progression of length \( |P| \geq n_0 \) and \( A \subseteq P \) with \( |A| \geq \varepsilon |P| \), then \( A \) is \( k \)-good.

It is not hard to prove (without assuming the truth of any of the statements) that (b), (c), and (d) are equivalent.

Erdős has conjectured that the following stronger statement holds:

(e) If \( A \subseteq \mathbb{N} \) and \( \sum_{a \in A} \frac{1}{a} = \infty \), then \( A \) is \( \omega \)-good.

By \( \sum_{a \in A} \frac{1}{a} \) we mean of course \( \sum_{a \in A} \left( \frac{1}{a} \right) \). The proof (or disproof) of (e) is, at present, out of sight. In fact, it has not even been proved that \( \sum_{a \in A} \frac{1}{a} = \infty \) implies that \( A \) is \( 3 \)-good (compare Roth [6]). That (e) \( \Rightarrow \) (c) can be seen as follows: If \( \sigma(A) = \varepsilon > 0 \), then there exists a sequence of natural numbers \( 0 = n_0 < n_1 < n_2 < \cdots \), such that, for each \( i \),

\[
\frac{A[1, n_i]}{n_i} > \frac{\varepsilon}{2} \quad \text{and} \quad \frac{n_{i-1}}{n_i} < \frac{\varepsilon}{4}.
\]

Then

\[
\sum_{a \in A} \frac{1}{a} \geq \sum_{a \in A} \frac{1}{a} \geq \sum_{i=1}^{k} \frac{A[n_{i-1} + 1, n_i]}{n_i} \geq \sum_{i=1}^{k} \frac{A[1, n_i] - n_{i-1}}{n_i}
\]

\[
\geq k \left( \frac{\varepsilon}{2} - \frac{\varepsilon}{4} \right) = \frac{k \varepsilon}{4} \to \infty \quad (k \to \infty)
\]

and so \( \sum_{a \in A} \frac{1}{a} = \infty \). Assuming (e), it follows that \( A \) is \( \omega \)-good.

Hence Erdős’ conjecture is indeed stronger than Szemerédi’s theorem. Note also that Erdős’ conjecture, if true, would immediately answer in the affirmative the long-standing question of whether or not the primes are \( \omega \)-good.

In the next section we make some observations regarding this conjecture, and we show that several special cases of the conjecture are true.

Other observations can be found in Gerver [3, 4] and Wagstaff [8].

2 Main results

(2.1). First we consider the “finite form” of Erdős’ conjecture.

**Theorem 1.** Fix \( k \), and assume that for all sets \( A \subseteq \mathbb{N} \), if \( \sum_{a \in A} \frac{1}{a} = \infty \) then \( A \) is \( k \)-good. Under this assumption, there exists \( T \) such that if \( \sum_{a \in A} \frac{1}{a} > T \), then \( A \) is \( k \)-good.

(Gerver [3] has this result under the stronger hypothesis that if \( \sum_{a \in A} \frac{1}{a} = \infty \) then \( A \) is \( (k+1) \)-good.)
Proof. We may assume $k \geq 3$. Suppose the theorem is false. We will construct a set $A$ such that $\sum_A(1/a) = \infty$ and $A$ is not $k$-good. Choose a finite set $A_0$ such that $A_0$ is not $k$-good and $\sum_A(1/a) > 1$. Let $p_1$ be prime, $p_1 > 2 \max A_0$, and choose a finite subset $A_1$ of $\{t p_1 : t \geq 1\}$ such that $A_1$ is not $k$-good and $\sum_{A_1}(1/a) > 1$. Let $p_2$ be prime, $p_2 > 2 \max A_1$, and choose a finite subset $A_2$ of $\{t p_2 : t \geq 1\}$ such that $A_2$ is not $k$-good and $\sum_{A_2}(1/a) > 1$. Continuing in this way, we obtain finite sets $A_0, A_1, \ldots$ such that for each $i \geq 0, A_i$ is not $k$-good, $\min A_{i+1} \geq p_{i+1} > 2 \max A_i$, each elements of $A_{i+1}$ is a multiple of $p_{i+1}$, and $\sum_{A_i}(1/a) > 1$.

Let $A = \bigcup A_i$. It is clear that $\sum_A(1/a) = \infty$. To show that $A$ is not $k$-good, it suffices to show that every 3-term arithmetic progression contained in $A$ must be contained in a single set $A_i$.

To this end, suppose that $x < y < z$, with $x, y, z \in A$ and $z - y = y - x$. Let $y \in A_i$. Then $z \in A_i$ also, since otherwise $z - y \geq \min A_{i+1} - \max A_i > \max A_i > y - x$. Thus $y, z \in A_i \subseteq \{t p_i : t \geq 1\}$. Hence $x$ is divisible by $p_i$, so $x \geq p_i > \max A_{i-1}$, and $x \in A_i$. This finishes the proof of Theorem 1.

Corollary 1. The following statement is equivalent to statement (e):

(f) For each $k \in N$, there exists $T \in N$ such that if $\sum_A(1/a) > T$, then $A$ is $k$-good.

We state next a lemma which will be useful later.

Lemma 1. Let $F_1, F_2, \ldots$ be a sequence of finite subsets of $N$ such that for each $i$, $F_i$ is not $k$-good and $\min F_{i+1} \geq 2 \max F_i$. Then $F = \bigcup F_i$ is not $(k+1)$-good.

(The proof of Lemma 1 is contained in the proof of Theorem 1 above).

(2.2). Now we define an increasing sequence, $a_1 < a_2 < a_3 < \cdots$, of natural numbers to be lacunary if $d_n = a_{n+1} - a_n \to \infty$ as $n \to \infty$ and to be $M$-lacunary if, furthermore, $d_n \leq d_{n+1}$ for all $n$. We shall think of such a sequence simultaneously as a sequence and as a subset of $N$. Any lacunary sequence $A$ has $\pi(A) = 0$ (see [2]), so that Szemerédi’s theorem does not apply.

A subsequence of a lacunary sequence is lacunary, but the corresponding statement, unfortunately, does not hold for $M$-lacunary sequences. It is known that if the real series $\sum t_i$ is not absolutely convergent, then there exists a lacunary sequence $B$ such that $\sum_{i \in B} t_i$ diverges (see Freedman and Sember [2]). It follows that if $A \subseteq N$ and $\sum_A(1/a) = \infty$, then there exists a lacunary sequence $B \subseteq A$ such that $\sum_B(1/b) = \infty$. Thus we have the following.

Theorem 2. The following statement is equivalent to statement (e).

(g) If $A$ is a lacunary sequence and $\sum_A(1/a) = \infty$, then $A$ is $\omega$-good.

Hence we need only investigate lacunary sequences when contemplating the Erdős conjecture. It can also be shown that if $\sum t_i = \infty$ and $t_i \geq 0$ for all $i$, then there exists an $M$-lacunary sequence $B$ such that $\sum_{i \in B} t_i = \infty$. (We omit the rather cumbersome proof of this statement.) But notice that this does not imply that statement (h) below is equivalent to statement (e)! This is too bad—because we now prove (h).

Theorem 3. The following statement is true:

(h) If $A$ is $M$-lacunary and $\sum_A(1/a) = \infty$, then $A$ is $\omega$-good.

Proof. Let $A = \{a_1 < a_2 < a_3 < \cdots\}$ be an $M$-lacunary sequence with infinite reciprocal sum. Assume there is a $k$ such that $d_i < d_{i+k}$ for each $i$, where $d_n = a_{n+1} - a_n$, $n \geq 1$. We show that $a_{i+jk} \geq j^2/2$ for
all $i \geq 1$, $j \geq 0$. Indeed,
\[
\begin{align*}
a_{i+jk} &= a_i + d_i + d_{i+1} + \cdots + d_{i+jk-1} \\
&\geq d_i + d_{i+k} + d_{i+2k} + \cdots + d_{i+(j-1)k} \\
&> 1 + 2 + \cdots + j > j^2/2.
\end{align*}
\]

(\text{Note that to obtain the first inequality we have merely omitted some terms from the sum.}) But then
\[
\sum_{i=1}^{\infty} \frac{1}{a_i} = \sum_{i=0}^{\infty} \frac{1}{a_{i+1}} + \sum_{i=0}^{\infty} \frac{1}{a_{2i+1}} + \cdots + \sum_{i=0}^{\infty} \frac{1}{a_{i+k}+jk} \\
\leq k(1 + \sum_{j=1}^{\infty} \frac{2}{j^2}) < \infty, \quad \text{a contradiction.}
\]

Hence, for each $k$, there is an $i$ such that $d_i = d_{i+k}$, whence $a_i, a_{i+1}, \ldots, a_{i+k+1}$ are in arithmetic progression and $A$ is $\omega$-good. \qed

The following is an immediate corollary.

\textbf{Corollary 2.} If $A$ is a finite union of $M$-lacunary sets and $\sum_{A}(1/a) = \infty$, then $A$ is $\omega$-good.

(2.3). We now use some slightly expanded arguments to show that statement (g) holds for some special sequences which are not $M$-lacunary (but are nearly so).

\textbf{Theorem 4.} Let $A = \{a_1 < a_2 < a_3 < \cdots\}$ be any set. Suppose there are intervals $I_n = [s_n, t_n]$ with $t_n < s_{n+1}$ such that

$$\frac{1}{\sum_{n=1}^{\infty} \sqrt{a_n}} < \infty, \quad \sum_{k \in [\cup I_n]} \frac{1}{a_k} = \infty.$$ 

Suppose further that for each $n$, $d_k \leq d_{k+1}$ if $s_n \leq k < t_n$. Then $A$ is $\omega$-good.

\textbf{Proof.} We will arrive at a contradiction if we assume that there is a $K \in N$, such that $d_i < d_{i+K}$ whenever $i, i + K$ belong to the same interval $I_j$. Then, for any $K$, we have that there exists an $i$ such that $d_i = d_{i+1} = \cdots = d_{i+K}$ so that $a_i, a_{i+1}, \ldots, a_{i+K+1}$ are in arithmetic progression.

To get the required contradiction we proceed as follows: If $n, n + K, n + 2K, \ldots, n + cK \in I_i$, then

$$\frac{1}{a_n} + \frac{1}{a_{n+K}} + \frac{1}{a_{n+2K}} + \cdots + \frac{1}{a_{n+cK}} \\
\leq \frac{1}{a_n} + \frac{1}{a_n + d_n} + \frac{1}{a_n + d_n + d_{n+K}} + \cdots \\
+ \frac{1}{a_n + d_n + d_{n+K} + \cdots + d_{n+(c-1)K}} \\
< \sum_{j=0}^{\infty} \frac{1}{a_n + (j^2/2)} < \frac{b}{\sqrt{a_n}} \leq \frac{b}{\sqrt{a_n}} \quad (b \text{ constant}).$$

Hence,

$$\sum_{k \in I_n} \frac{1}{a_k} < \frac{Kb}{\sqrt{a_n}} \quad \text{and} \quad \sum_{k \in [\cup I_n]} \frac{1}{a_k} < Kb \sum_{i=1}^{\infty} \frac{1}{\sqrt{a_n}} < \infty,$$
definitions are members in $f$ $L$.

Proof. Let $A = \{a_1 < a_2 < a_3 < \cdots \}$ be a set. Suppose $I_n = [s_n, t_n]$ are intervals with $t_n < s_{n+1}$ such that $d_i \leq d_{i+1}$ if $s_n \leq i < t_n$ and $d_{n-1} < d_{n+1}$. Then, if $\sum_{i \in S} (1/\alpha_i) = \infty$, $A$ is $\omega$-good.

(2.4). We now define new density functions $\lambda$ and $\overline{\lambda}$ in terms of lacunary sequences: For all sets $A$, let $\overline{\lambda}(A) = 0$ if $A$ is finite or a finite union of lacunary sequences and otherwise let $\overline{\lambda}(A) = 1$. Define $\lambda(A) = 1 - \overline{\lambda}(N - A)$. These densities, taking only 0, 1 values, may seem a little odd. The definition could be improved so that $\lambda$ becomes “continuous” and has the correct value on an (infinite) arithmetic progression etc. However, this would not suit our purposes any better. One can prove that for any $A \subseteq N$,

$$\lambda(A) \leq \underline{\lambda}(A) \leq \overline{\lambda}(A) \leq \pi(A)$$

and so, in analogy to Szemerédi’s theorem it is natural to ask about the arithmetic progressions in $A$ if $\overline{\lambda}(A) > 0$.

Theorem 6. There exists a set $A$ such that $\overline{\lambda}(A) > 0$ and $A$ is not $\omega$-good.

Proof. Let $B_i = \{1!, 2!, \ldots, i!\}$. $B_i$ is not 3-good. Let $(H_i)$ be the sequence of sets

$$(B_1, B_1, B_2, B_3, B_1, B_2, B_3, B_1, B_2, B_3, B_1, \ldots).$$

Let $f_i$ be an increasing sequence of integers such that $f_i = 0$ and

$$\min(f_{i+1} + H_{i+1}) \geq 2 \max(f_i + H_i)$$

and define $A = \bigcup_i (f_i + H_i)$. By Lemma 1, $A$ is not 4-good. (By choosing $f_i$ sufficiently quickly increasing one can even make $A$ not 3-good.) Finally, $\overline{\lambda}(A) = 1$ since otherwise $A = L_1 \cup L_2 \cup \cdots \cup L_k$ where each $L_j$ is a lacunary sequence. Whenever $H_i = B_{k+1}$ we have $|f_i + H_i| \geq k$ and so some $L_j$ has at least two members in $f_i + H_i$. Hence we may find a fixed $j$ such that

$$|L_j \cap (f_i + B_{k+1})| \geq 2$$

for infinitely many $i$. Then $L_j$ has infinitely many differences $d_i < (k + 1)!$, and so $L_j$ is not lacunary. \(\square\)

(2.5). Let us consider “relative density”, that is, “the density of $A$ relative to $B$” where $A \subseteq B$. The definitions are

$$\underline{\lambda}(A|B) = \liminf_{i \to \infty} \frac{A[1, b_i]}{i}$$

and

$$\overline{\lambda}(A|B) = \limsup_{n \to \infty} \frac{A[b_{m+1}, b_{m+n}]}{n}.$$ 

$\underline{\lambda}(A|B)$ and $\overline{\lambda}(A|B)$ are obtained by replacing “inf” with “sup” and “min” with “max” respectively. One
can show, as before, for any \( A, B, A \subseteq B \), that
\[
\underline{d}(A|B) \leq \overline{d}(A|B) \leq \overline{\delta}(A|B) \leq \pi(A|B).
\]
Let \( B \) be \( M \)-lacunary and \( \sum b 1/b = \infty \). Then, by Theorem 3, \( B \) is \( \omega \)-good. We ask whether \( A \subseteq B \) and the relative density of \( A \) positive imply that \( A \) is also \( \omega \)-good. The answer is “yes” if \( \underline{d}(A|B) > 0 \) (Theorem 7), “no” if \( \overline{\delta}(A|B) > 0 \) (Theorem 8) and the question is open for \( \overline{d}(A|B) > 0 \).

**Theorem 7.** If \( B \) is \( M \)-lacunary, \( \sum b 1/b = \infty \), \( A \subseteq B \) and \( \underline{d}(A|B) > 0 \) then \( A \) is \( \omega \)-good.

**Proof.** By (the proof of) Theorem 3 there are arbitrarily large \( n, m \) such that
\[
P = \{b_{m+1}, b_{m+2}, \ldots, b_{m+n}\}
\]
is an arithmetic progression. By the definition of \( \underline{d}(A|B) \) we have \( |A \cap P| \geq \epsilon P \) where \( \epsilon = (1/2)\underline{d}(A|B) \) and \( |P| \) is arbitrarily large. Thus, by Szemerédi’s theorem (d) we have, for any \( k \), that \( |A \cap P| \) is \( k \)-good if \( |P| \) is sufficiently large. Hence \( A \) is \( \omega \)-good. \( \square \)

**Theorem 8.** There exists an \( M \)-lacunary sequence \( B \) with \( \sum b 1/b = \infty \) and an \( A \subseteq B \) with \( \overline{\delta}(A|B) > 0 \) (= 1 in fact) such that \( A \) is not \( 3 \)-good.

**Proof.** (leaving most of the details to the reader). Let \( F = \{1!, 2!, 3!, \ldots\} \), \( b_1 = 1 \) and define \( b_{n+1} = b_n + d_n \) where the \( d_n \)’s have the following properties: For all \( i, d_i \in F \) and \( d_i \leq d_{i+1} \). Furthermore, the set of natural numbers \( N \) can be partitioned into consecutive pairwise disjoint intervals \( J_1, J_2, J_3, \ldots \) such that if \( r \) is odd, then for \( i \in J_r \), \( d_i = d_{i+1} \) and \( \sum_{i \in J} 1/b_i \geq 1 \), and, if \( r \) is even, then, for \( i \in J_r \), \( d_i < d_{i+1} \), \( b_i > 2b_{i-1} - 1 \) and \( |J_r| > (\max J_{r-1})^2 \). Clearly \( B = \{b_1, b_2, \ldots\} \) is \( M \)-lacunary and \( \sum b 1/b = \infty \).

Let \( A = \{b_k : k \in \bigcup J_{2r}\} \). Then
\[
\overline{\delta}(A|B) \geq \lim_r \frac{|J_{2r}|}{|J_{2r}| + \max J_{2r-1}} = 1.
\]
One can also see that \( A \) is not \( 3 \)-good since \( a_t > 2a_{t-1} \) holds. \( \square \)

(2.6). Theorems 4, 5, and 7 notwithstanding, it seems to be difficult to generalize the notion of \( M \)-lacunary even slightly and still prove the corresponding case of the Erdős conjecture. In this connection let us define a lacunary sequence \( A \) to be \( M_k \)-lacunary (where \( k \geq 0 \)) if, for all \( i, j, i \leq j \), we have \( d_i \leq d_j + k \). Clearly, the \( M_0 \)-lacunary sequences are just the \( M \)-lacunary sequences. For no \( k \neq 0 \) are we able to prove that \( M_k \)-lacunary and \( \sum 1/a = \infty \) imply \( \omega \)-good. We can show if \( A \) is \( M_1 \)-or \( M_2 \)-lacunary with \( \sum 1/a = \infty \) then \( A \) is 3-good. We prove first a lemma which may have independent interest:

**Lemma 2.** If \( A = \{a_1 < a_2 < a_3 < \cdots\} \) is any subset of \( N \) and \( \sum_a 1/a = \infty \), then, for any \( t > 0 \), there exists an \( i \) such that \( d_{i+j} \leq d_i \) for \( j = 0, 1, \ldots, t \). (Of course, \( d_n = a_{n+1} - a_n \))

**Proof.** The method is familiar by now: Suppose there is a \( t \) such that, for each \( i \), there exists \( j \in [1, t] \) with \( d_i < d_{i+j} \). Then we can find a sequence \( (j_n) \) such that
\[
d_1 < d_{1+j_1} < d_{1+j_1+j_2} < \cdots (j_n \in [1, t]).
\]
It follows that
\[
\sum_{A} \frac{1}{a} \leq t \sum_{s=0}^{\infty} \frac{1}{a_{1+(1+s)}} \leq t \sum_{s=0}^{\infty} \frac{1}{a_{1+(1+2+\cdots+s)}} < \infty.
\]

**Theorem 9.** Let \( A \) be \( M_1 \)-or \( M_2 \)-lacunary and \( \sum_{A}(1/a) = \infty \). Then \( A \) is 3-good.

**Proof.** By the definition of \( M_k \)-lacunary and Lemma 2 we have: for any \( t > 0 \) there is an \( i \) such that
\[
d_i - e \leq d_{i+j} \leq d_i, \quad j = 0, 1, \ldots, t,
\]
where \( e = 1 \) or \( 2 \). Hence, in the sequence \( (d_i) \), we have arbitrarily long blocks where the \( d_i \) take on only two (in case \( e = 1 \)) or three (in case \( e = 2 \)) values. Such long blocks must contain two consecutive subblocks with identical composition (see Pleasants [5]). These two subblocks will determine three terms of the sequence \( A \) in arithmetic progression.

This last result suggests a conjecture which is related to van der Waerden’s theorem on arithmetic progressions and would immediately imply that \( M_k \)-lacunary with \( \sum_{A}(1/a) = \infty \) implies that \( A \) is 3-good.

**Conjecture.** Let \( x_i \) be a sequence of positive integers with \( 1 \leq x_i \leq K \). Then there are two consecutive intervals \( I, J \) of the same length, with \( \sum_{i \in I} x_i = \sum_{j \in J} x_j \). Equivalently, if \( a_1 < a_2 < \cdots \) satisfy \( a_{n+1} - a_n \leq K \), all \( n \), then there exist \( x < y < z \) such that \( x + z = 2y \) and \( a_x + a_z = 2a_y \).

**References**


