

## NEAR ARITHMETIC PROGRESSIONS IN SPARSE SETS

STEVEN C. LETH

(Communicated by Carl G. Jockusch, Jr.)

ABSTRACT. Nonstandard methods are used to obtain results in combinatorial number theory. The main technique is to use the standard part map to translate density properties of subsets of  ${}^*\mathbb{N}$  into Lebesgue measure properties on  $[0, 1]$ . This allows us to obtain a simple condition on a standard sequence  $A$  that guarantees the existence of intervals in arithmetic progression, all of which contain elements of  $A$  with various uniform density conditions.

### 1. INTRODUCTION

In 1936 Erdős and Turán conjectured that if the sequence  $A \subset \mathbb{N}$  has positive upper Banach Density, then  $A$  contains arbitrarily long arithmetic progressions. This was proved by Szemerédi in 1974 [21], and re-proved using completely different approaches by Furstenberg in 1977 [2, 3] and Gowers in 1999 [4]. A major question still unresolved, even with the recent proof by Green and Tao that the primes contain arbitrarily long arithmetic progressions [6], is a further conjecture of Erdős and Turán. They conjectured that if  $A = \langle a_n \rangle$  is a sequence of natural numbers with the property that  $\sum_{n=1}^{\infty} 1/a_n$  diverges, then  $A$  contains arbitrarily long arithmetic progressions [1].

In this paper we concentrate on conditions that imply that a sequence  $A$  contains intervals in arithmetic progression on which  $A$  is nonempty, and on which various density properties hold. We can obtain some results about the existence of such intervals for sequences that are much sparser than those in the conjecture above. We will then look at one example of a condition involving these concepts that would provide a positive solution to the Erdős-Turán conjecture.

In this paper we will make use of nonstandard methods. Some early work in the application of nonstandard analysis to problems in combinatorial number theory can be found in [16, 17] and [18]. Recently Renling Jin has obtained a number of strong standard results in combinatorial number theory through the use of nonstandard methods [8, 9, 10, 11, 12, 13, 14, 15]. The reader unfamiliar with the use of nonstandard methods will find excellent introductions to the subject in any of [7], [19] or [20].

The notation has been chosen with an effort to clarify the various objects used in the proof. The capital letters  $A$  and  $B$  will represent sequences, either internal

---

Received by the editors August 31, 2004 and, in revised form, October 14, 2004, November 7, 2004, and January 5, 2005.

2000 *Mathematics Subject Classification*. Primary 11B25, 11B05, 03H05, 03H15.

*Key words and phrases*. Nonstandard analysis, arithmetic progressions.

The author thanks the referee for identifying several errors in the original version.

sequences in  ${}^*\mathbb{N}$  or standard sequences in  $\mathbb{N}$ . The capital letters  $I$  and  $J$  are used for intervals, again either standard or nonstandard. All other capital letters represent nonstandard integers in  ${}^*\mathbb{N} - \mathbb{N}$ , i.e. nonstandard integers that are “infinite” in size. Lowercase letters represent standard natural or real numbers, or nonstandard natural numbers not necessarily known to be in  ${}^*\mathbb{N} - \mathbb{N}$ . Lowercase greek letters are used for either real numbers or as indexing natural numbers, and the use will be clear in the context.

## 2. THE MAIN THEOREM

**Definition 1.** Let  $A \subset \mathbb{N}$ , and let  $I = [a, b]$  be an interval in  $\mathbb{N}$ . We will write  $\mathbf{l}(I)$  for the length of  $I$  (i.e.  $l(I) = b - a + 1$ ), and we will write  $\delta(\mathbf{A}, \mathbf{I})$  or  $\delta(\mathbf{A}, [a, b])$  for the density of the set  $A$  on the interval  $I$ . Thus  $\delta(A, I) = \frac{|A \cap I|}{l(I)}$ .

**Definition 2.** Let  $t, d$  and  $w$  be in  $\mathbb{N}$ , and let  $\alpha \in \mathbb{R}$  with  $0 < \alpha < 1$ . For  $A \subset \mathbb{N}$  and  $I$  an interval in  $\mathbb{N}$  we say that  $A$  contains a  $t$ -termed  $\alpha$ -homogeneous cell of distance  $d$  and width  $w$  in  $I$  or simply a  $\langle t, \alpha, \mathbf{d}, \mathbf{w} \rangle$  cell in  $I$  iff there exists  $b \in I$  with  $b + (t - 1)d + w$  also in  $I$  such that for each  $\nu, \xi = 0, 1, 2, \dots, t - 1$ :

$$(2.1) \quad \delta(A, [b + \xi \cdot d, b + \xi \cdot d + w]) \geq (1 - \alpha) \delta(A, [b + \nu \cdot d, b + \nu \cdot d + w]) \geq (1 - \alpha)^2 \delta(A, I).$$

If each  $\delta(A, [b + \xi \cdot d, b + \xi \cdot d + w])$  is simply nonzero, i.e. the intervals are nonempty, then we say that  $A$  contains a  $\langle t, \mathbf{d}, \mathbf{w} \rangle$  cell.

For  $\beta > 0$  and  $0 \leq u \leq w$  we will say that a  $\langle t, \alpha, d, w \rangle$  cell is  $\mathbf{u}, \beta$  uniform if for each  $\nu = 0, 1, 2, \dots, t - 1$ , and all  $x$  such that  $u \leq x \leq w$ :

$$(2.2) \quad (1 - \beta) \delta(A, J_\nu) \leq \delta(A, [b + \nu \cdot d, b + \nu \cdot d + x]) \leq (1 + \beta) \delta(A, J_\nu),$$

where  $J_\nu$  denotes the interval  $[b + \nu \cdot d, b + \nu \cdot d + w]$ .

It is clear that an actual arithmetic progression of length  $t$  and distance  $d$  is an example of a  $\langle t, \alpha, w, d \rangle$  cell with  $w = 0$  and  $\alpha$  any nonnegative number. Furthermore, this cell is  $\mathbf{u}, \beta$  uniform for  $u = 0$  and any nonnegative  $\beta$ . We could view the existence of a  $\langle t, \alpha, d, w \rangle$  cell inside a sequence  $A$  as a weak form of an actual arithmetic progression inside  $A$ . These cells are “near” arithmetic progressions in some (perhaps very loose) sense, and intuitively are “nearer” to arithmetic progressions as the size of  $w$  decreases. In the results below  $w$  will be “small” in the sense that the ratio of  $w$  to  $d$  will be small compared to the ratio of  $d$  to the length of the interval  $I$ .

**Definition 3.** Let  $I$  be an interval in  $\mathbb{N}$ , and let  $A \subset I$ , with  $r > 1 \in \mathbb{R}$  and  $m \in \mathbb{N}$ . We say that  $A$  has the  $\mathbf{m}, \mathbf{r}$  density property on  $I$  iff for any interval  $J \subset I$ , if  $l(J) \geq \frac{l(I)}{m}$ , then  $\delta(A, J) \leq r \delta(A, I)$ .

Before proving the main theorem we need the following standard lemma. For  $E \subset \mathbb{R}$  we will write  $\lambda(E)$  for the Lebesgue measure of  $E$ .

**Lemma 1.** *If  $E$  is a subset of the reals with positive Lebesgue measure and  $t$  is any natural number, then there exists a real number  $s > 0$  such that for all  $d < s$ ,  $E$  contains a  $t$ -termed arithmetic progression with difference  $d$ , i.e. there exists  $b \in \mathbb{R}$  such that  $b + \nu d \in E$  for each  $\nu = 0, 1, 2, \dots, t - 1$ .*

*Proof.* Since  $E$  has positive Lebesgue measure, almost all points of  $E$  are points of density of  $E$ . Thus we may find  $x < y$  in  $\mathbb{R}$  such that  $\lambda(E \cap [x, y]) \geq (1 - 1/4t)(y - x)$ . Now choose  $s < (y - x)/2t$ . Then on the interval  $[x, (x + y)/2]$  the measure of any

$E - \nu d$  with  $\nu = 0, 1, 2, \dots, t - 1$ , and  $d < s$  is at least  $(y - x)/2 - (y - x)/(4t)$ , so that

$$\begin{aligned} & \lambda(E \cap (E - d) \cap (E - 2d) \cap \dots \cap (E - (t - 1)d) \cap [x, (x + y)/2]) \\ & \geq (y - x)/2 - (y - x)/4 = (y - x)/4 > 0. \end{aligned}$$

We may now let  $b$  be any element in  $E \cap (E - d) \cap (E - 2d) \cap \dots \cap (E - (t - 1)d) \cap [x, (x + y)/2]$ , and we obtain the desired result.  $\square$

Theorem 1 below is the main result of this paper. It gives a condition for the existence of “near” arithmetic progressions for any sequence on any interval  $I$  in which the density does not drastically increase as the size of the subinterval decreases. More specifically, it provides an absolute constant such that whenever the density of a sequence does not increase beyond a fixed ratio for any subinterval of size greater than the length of  $I$  divided by that fixed constant, then the sequence will contain a  $\langle t, \alpha, w, d \rangle$  cell with some relative “smallness” conditions for  $w$ .

**Theorem 1.** *Let  $h(x)$  be any increasing real valued function such that  $h(x) > 0$  whenever  $x > 0$ , and let  $g(x)$  be any real valued function which approaches infinity as  $x$  approaches infinity. For all real  $\alpha > 0$ ,  $r > 1$  and  $j, t \in \mathbb{N}$  there exists a standard natural number  $m$  such that for all  $n > m$ , whenever  $I$  is an interval of length  $n$  and any nonempty set  $A \subset I$  has the  $m, r$  density property on  $I$ , then  $A$  contains a  $u, \beta$  uniform  $\langle t, \alpha, d, w \rangle$  cell with  $\frac{u}{w} < h(\frac{w}{d})$ ,  $\frac{w}{d} < h(\frac{d}{n})$ ,  $\beta < h(\frac{d}{n})$  and  $\frac{n}{g(m)} < d < \frac{n}{j}$ . Furthermore, we may take  $w$  and  $d$  to be powers of 2.*

The theorem is strongest if we take  $h$  to be approaching 0 rapidly as  $x$  approaches 0, and  $g$  to be approaching infinity slowly as  $x$  approaches infinity. The  $m$  that we obtain will depend on  $h, g, \alpha, r, j$  and  $t$ , but will be independent of the sequence  $A$  or the interval  $I$ .

*Proof.* Suppose  $h(x), g(x), \alpha, j, r, t$  are given as in the statement above and that no such  $m$  exists. Since the result becomes stronger as  $\alpha$  decreases, we assume for convenience that  $\alpha < 1/2$ . By “overspill” there exists an  $M, N$  in  ${}^*\mathbb{N} - \mathbb{N}$  with  $M < N$  and a hyperfinite internal set  $A$  such that  $A$  has the  $M, r$  density property on an interval of length  $N$  but  $A$  contains no  $\langle t, \alpha, d, w \rangle$  cell on this interval with the required properties. Since the conditions are translation invariant, we may assume that the interval is  $[0, N - 1]$ . We now define a standard function  $f : [0, 1] \rightarrow [0, 1]$  by

$$f(x) = st \left( \frac{|A \cap [0, xN]|}{|A \cap [0, N]|} \right).$$

We first note that the function  $f(x)$  is increasing, and show that it satisfies a Lipschitz condition with constant  $r$ . To see this, note that for  $x < y$  in  $[0, 1]$  there exist hyperfinite  $L_1$  and  $L_2$  such that  $x = st(L_1/N)$  and  $y = st(L_2/N)$ , and  $L_2 - L_1$  is a noninfinitesimal multiple of  $N$ . Since  $A$  has the  $M, r$  density property on  $[0, N]$ , and since  $L_2 - L_1 > N/M$ ,

$$\delta(A, [L_1, L_2]) \leq r\delta(A, [0, N]).$$

This yields

$$\begin{aligned} f(y) - f(x) &= st \left( \frac{|A \cap [L_1, L_2]|}{|A \cap [0, N]|} \right) = st \left( \frac{|A \cap [L_1, L_2]|}{L_2 - L_1} \frac{L_2 - L_1}{|A \cap [0, N]|} \right) \\ &\leq r \, st \left( \delta(A, [0, N]) \frac{L_2 - L_1}{|A \cap [0, N]|} \right) = r \, st \left( \frac{L_2 - L_1}{N} \right) = r(y - x), \end{aligned}$$

and  $f$  satisfies a Lipschitz condition with constant  $r$  as desired. Thus, the function  $f$  is absolutely continuous, differentiable almost everywhere and equal to the integral of its derivative. Since  $f(1) = 1$ ,  $f(0) = 0$  and  $f$  is the integral of its derivative, it must be that the Lebesgue measure of  $\{x : f'(x) \geq (1 - \frac{\alpha}{4})\}$  is nonzero. Thus, there exists a real number  $c \geq 1$  such that the Lebesgue measure of the set

$$E = \left\{ x : c - \frac{\alpha}{4} \leq f'(x) \leq c \right\}$$

is nonzero. By the standard lemma above there exist  $B, D \in {}^*\mathbb{N} - \mathbb{N}$  noninfinitesimal to  $N$ , with  $D$  a power of 2,  $D < N/j$ , such that  $E$  contains a  $t$ -termed arithmetic progression of points  $b, b + d, b + 2d, \dots, b + (t - 1)d$ , where  $b = st(B/N)$  and  $d = st(D/N)$ . (A note on notation: it is helpful for clarity to use “ $d$ ” to denote the real number corresponding to the hyperfinite number “ $D$ ”, but it is “ $D$ ” that will act like the generic “ $d$ ” in the statement of the theorem.)

Since  $f$  is differentiable at all of the finitely many  $b + \nu d$  there exists a standard  $y > 0$  such that for all  $h$  with  $0 < h < y$  and all  $\nu = 0, \dots, t - 1$ ,

$$(2.3) \quad \left| f'(b + \nu d) - \frac{f(b + \nu d + h) - f(b + \nu d)}{h} \right| < \min \left\{ \frac{\alpha}{4}, h\left(\frac{d}{2}\right)/4 \right\}.$$

We now choose  $W, U \in {}^*\mathbb{N} - \mathbb{N}$  such that both  $W$  and  $U$  are powers of 2, with  $W/D < h(D/N)$ ,  $U/W < h(W/D)$ ,  $st(W/D) < y$ , and  $U$  noninfinitesimal to  $D$ . We note that in the above definitions we were able to take  $D, W$ , and  $U$  arbitrarily small but noninfinitesimal to  $N$ , allowing us to choose powers of 2 for these quantities, but that we do not have that freedom with the choice of  $B$ . Here we have also used the fact that  $h$  applied to noninfinitesimals is noninfinitesimal, which follows immediately from the conditions given on the function  $h$ . Finally, we let  $\beta = h(d/2)$ .

We wish to show that  $A$  contains a  $U, \beta$  uniform  $\langle t, \alpha, D, W \rangle$  cell with all the properties required in the theorem, contradicting our assumption about  $A$ . We first note that all the inequalities in the statement of the theorem hold by our choices of  $W, U, D$  and  $\beta$ . To see that  $A$  contains such a  $\langle t, \alpha, D, W \rangle$  cell, we begin by letting  $w = st(W/D)$ , and note that (2.3) above (using the inequality with  $\frac{\alpha}{4}$  on the right-hand side and with  $h = w$ ) gives us

$$f'(b + \nu d) - \frac{\alpha}{4} < \frac{f(b + \nu d + w) - f(b + \nu d)}{w} < f'(b + \nu d) + \frac{\alpha}{4}.$$

This and the fact that all  $b + \nu d$  are in  $E$  now yields

$$(2.4) \quad c - \frac{\alpha}{2} < \frac{f(b + \nu d + w) - f(b + \nu d)}{w} < c + \frac{\alpha}{4}.$$

The definition of the function  $f$  and elementary properties of the standard part function allow us to replace the middle quantity in the inequality (2.4) by

$$(2.5) \quad \begin{aligned} st\left(\frac{(|A \cap [B + \nu D, B + \nu D + W]|)}{(|A \cap [0, N]|)(W/N)}\right) &= \frac{st(|A \cap [B + \nu D, B + \nu D + W]|/W)}{st(|A \cap [0, N]|/N)} \\ &= \frac{\delta(A, [B + \nu D, B + \nu D + W])}{\delta(A, [0, N])} \end{aligned}$$

so that we obtain

$$\delta(A, [0, N])\left(c - \frac{\alpha}{2}\right) < \delta(A, [B + \nu D, B + \nu D + W]) < \delta(A, [0, N])\left(c + \frac{\alpha}{4}\right).$$

This immediately yields the right-hand inequality in (2.1). To see that the left-hand inequality of (2.1) holds, we use both of the inequalities above, once with  $\xi$  and once with  $\nu$ , to obtain

$$\begin{aligned} \delta(A, [B + \xi D, B + \xi D + W]) &> \delta(A, [0, N])\left(c - \frac{\alpha}{2}\right) \\ &> \frac{\left(c - \frac{\alpha}{2}\right)}{\left(c + \frac{\alpha}{4}\right)} \delta(A, [B + \nu D, B + \nu D + W]). \end{aligned}$$

Since  $\frac{c - \alpha/2}{c + \alpha/4} \geq \frac{1 - \alpha/2}{1 + \alpha/4} > 1 - \alpha$  we see that

$$\delta(A, [B + \xi D, B + \xi D + W]) > (1 - \alpha)\delta(A, [B + \nu D, B + \nu D + W])$$

so that condition (2.1) is satisfied with the quantities  $B, D$  and  $W$  in place of  $b, d$ , and  $w$ . To see that condition (2.2) holds, we fix  $\nu$  and let  $X$  be any hyperfinite number between  $U$  and  $W$ , and  $x = st(X/N)$ . Then inequality (2.3) with  $h(\frac{\beta}{2})/4 = \beta/4$  on the right-hand side yields each of the following inequalities:

$$\begin{aligned} -\frac{\beta}{4} &< \frac{f(b + \nu d + x) - f(b + \nu d)}{x} - f'(b + \nu d) < \frac{\beta}{4}, \\ -\frac{\beta}{4} &< f'(b + \nu d) - \frac{f(b + \nu d + w) - f(b + \nu d)}{w} < \frac{\beta}{4}. \end{aligned}$$

Adding these two inequalities, we obtain

$$-\frac{\beta}{2} < \frac{f(b + \nu d + x) - f(b + \nu d)}{x} - \frac{f(b + \nu d + w) - f(b + \nu d)}{w} < \frac{\beta}{2}.$$

Now, in exactly the same way that the middle quantity in (2.4) equals the expression given in (2.5), the above inequality becomes

$$-\frac{\beta}{2} < \frac{\delta(A, [B + \nu D, B + \nu D + X])}{\delta(A, [0, N])} - \frac{\delta(A, [B + \nu D, B + \nu D + W])}{\delta(A, [0, N])} < \frac{\beta}{2}.$$

Letting  $J = [B + \nu D, B + \nu D + W]$ , adding  $\frac{\delta(A, [B + \nu D, B + \nu D + W])}{\delta(A, [0, N])}$  to each term, multiplying by  $\delta(A, [0, N])$  and using the fact that  $\delta(A, J) \geq \delta(A, [0, N])(1 - \alpha)$  gives us

$$\begin{aligned} \delta(A, [0, N])\left(1 - \frac{\beta}{2(1 - \alpha)}\right) &< \delta(A, [B + \nu D, B + \nu D + X]) \\ &< \delta(A, [0, N])\left(1 + \frac{\beta}{2(1 - \alpha)}\right), \end{aligned}$$

and since  $\alpha < 1/2$ , so that  $\frac{\beta}{2(1 - \alpha)} < \beta$ , the above inequality shows that condition (1.2) holds. Thus,  $A$  contains a  $U, \beta$  uniform  $\langle t, \alpha, D, W \rangle$  cell with all the properties required in the theorem, and this contradiction completes the proof.  $\square$

**Definition 4.** For  $A$  a sequence of positive integers we define the **upper Banach Density** of  $A$  or **BD(A)** by

$$BD(A) = \inf_{x \in \mathbb{N} - \{0\}} \max_{a \in \mathbb{N}} \frac{|A \cap [a + 1, a + x]|}{x}.$$

Upper Banach density is often simply called Banach density, and is sometimes referred to in the literature as strong upper density.

**Corollary 1.** *Let  $h, g, \alpha, k, t$  be as in the previous corollary, and suppose that  $A$  is a sequence with  $BD(A) > 0$ . Then  $A$  contains a  $u, \beta$  uniform  $\langle t, \alpha, d, w \rangle$  cell with  $\frac{u}{w} < h(\frac{w}{d})$ ;  $\frac{w}{d} < h(\frac{d}{n})$ ,  $\beta < h(\frac{d}{n})$  and  $d > \frac{n}{g(m)}$ . Furthermore, we may take  $u, w$ , and  $d$  to be powers of 2.*

*Proof.* Let  $r \in \mathbb{R}$  be greater than  $1/BD(A)$ , and let  $m(r)$  be the  $m$  that exists by Theorem 1. Now find an interval  $I$  of length greater than  $m(r)$  such that the density of  $A$  on  $I$  is greater than  $1/r$ . Then  $A$  has the  $m, r$  density property on  $I$  so that the result follows from the theorem. □

In light of Szemerédi’s theorem the corollary above is very weak. The goal now is to look at deeper consequences of Theorem 1 in the hope of providing insight into questions about arithmetic progressions that are still unknown.

### 3. DENSITY CONDITIONS FOR NEAR ARITHMETIC PROGRESSIONS

We begin with a simple proposition needed for the next proof, followed by a result about relatively sparse sets.

**Proposition 1.** *Let  $m, n \in \mathbb{N}$ , with  $m < n$ , let  $A \subset \mathbb{N}$ , and let  $I$  be an interval of length  $n$ . Then there exists an interval  $J \subset I$  of length  $m$  such that  $\delta(A, J) > \delta(A, I)/2$ .*

*Proof.* We may translate so that the interval  $I = [0, n - 1]$ . It is clear that if  $m|n$ , then we can, in fact, find such a  $J$  for which  $\delta(A, J) \geq \delta(A, I)$ , so let  $n = km + r$ , with  $0 < r < m$ . If one of the intervals  $[\nu m, \nu m + m - 1]$  for  $\nu = 0, 1, \dots, k - 1$  is of density  $> \delta(A, I)/2$ , then we may let  $J$  equal that interval. Otherwise

$$|A \cap [km, km + r - 1]| = |A \cap I| - |A \cap [0, km - 1]| \geq n\delta(A, I) - km\delta(A, I)/2.$$

We may now let

$$J = [(k - 1)m + r, km + r - 1],$$

since

$$\begin{aligned} \delta(A, J) &= \frac{|A \cap J|}{m} \geq \frac{|A \cap [km, km + r - 1]|}{m} \geq \frac{n\delta(A, I) - km\delta(A, I)/2}{m} \\ &> k\delta(A, I) - k\delta(A, I)/2 = k\delta(A, I)/2 \geq \delta(A, I)/2. \end{aligned}$$

□

**Theorem 2.** *Let  $\alpha > 0$ , and  $t > 2 \in \mathbb{N}$  be given, let  $h(x)$  be any continuous real valued function such that  $h(x) > 0$  whenever  $x > 0$  and let  $A$  be a sequence in  $\mathbb{N}$  with the property that for all  $\varepsilon > 0$ ,  $|A \cap [0, n - 1]| > n^{1-\varepsilon}$  for sufficiently large  $n$ . Then for sufficiently large  $n$ ,  $A$  contains a  $u, \beta$  uniform  $\langle t, \alpha, d, w \rangle$  cell on  $[0, n - 1]$  with*

$$\frac{u}{w} < h\left(\frac{w}{d}\right), \quad \frac{w}{d} < h\left(\frac{\log d}{\log n}\right), \quad \text{and} \quad \beta < h\left(\frac{\log d}{\log n}\right).$$

*Proof.* Let  $N \in {}^*\mathbb{N} - \mathbb{N}$ , and let  $A$  be a subset of  $[0, N - 1]$  that contains no  $\langle t, \alpha, d, w \rangle$  cell on  $[0, N - 1]$  with the properties above. It suffices to show that  $|A| < N^{1-\varepsilon}$  for some standard  $\varepsilon > 0$ . Let the natural number  $m$  be the number that exists by Theorem 1 with  $r = 2$ , and with  $g$  the identity function and  $h$  as given here, i.e. for all  $n > m$ , whenever  $I$  is an interval of length  $n$  and any nonempty set contained in  $I$  has the  $m, 2$  density property on  $I$ , then the set contains a  $\langle t, \alpha, d, w \rangle$  cell with  $\frac{u}{w} < h(\frac{w}{d})$ ,  $\frac{w}{d} < h(\frac{d}{n})$ ,  $\beta < h(\frac{d}{n})$  and  $d > \frac{n}{m}$  (the theorem guarantees some additional conditions that we will not need here).

We now show that there exists no interval  $I \subset [1, N]$  with  $l(I) > \sqrt{N}$  such that  $A$  has the  $m, 2$  density property on  $I$ . If such an interval did exist, then by Theorem 1 and the definition of  $m$ ,  $A$  contains a  $u, \beta$  uniform  $\langle t, \alpha, d, w \rangle$  cell with

$$\frac{u}{w} < h\left(\frac{w}{d}\right), \frac{w}{d} < h\left(\frac{d}{l(I)}\right), \beta < h\left(\frac{d}{l(I)}\right) \text{ and } d > \frac{l(I)}{m}.$$

We now have

$$d > \frac{l(I)}{m} > \frac{\sqrt{N}}{m}$$

with  $m$  finite, so that

$$st\left(\frac{\log d}{\log N}\right) \geq st\left(\frac{1/2 \log N - \log m}{\log N}\right) = \frac{1}{2}.$$

But  $\frac{d}{l(I)}$  is at most  $\frac{1}{l}$ , as the entire cell is contained in  $I$ . Thus  $\frac{d}{l(I)} < \frac{\log d}{\log N}$  and since  $h$  is an increasing function, we see that the conditions

$$\frac{w}{d} < h\left(\frac{\log d}{\log N}\right) \text{ and } \beta < h\left(\frac{\log d}{\log N}\right)$$

are satisfied. But the condition that  $\frac{u}{w} < h(\frac{w}{d})$  is also satisfied, contradicting the definition of  $A$ .

Thus, every interval  $I$  in  $[0, N - 1]$  with  $l(I) \geq \sqrt{N}$  contains a subinterval  $J$  of length  $\left\lfloor \frac{l(I)}{m} \right\rfloor$  with  $\delta(A, J) \geq 2\delta(A, I)$ . Now let  $K$  be maximal such that  $m^{2K} \leq N$ , and using the proposition we may let  $I_0$  be a subinterval of  $[0, N - 1]$  of length  $m^{2K}$  such that  $\delta(A, I_0) \geq \delta(A, [0, N - 1])/2$ . We inductively define:

$$I_{i+1} = \text{a subinterval of } I_i \text{ of length } l(I_i)/m \text{ such that } \delta(A, I_{i+1}) \geq 2\delta(A, I_i).$$

Then  $I_K$  is an interval of length  $m^K \leq \sqrt{N} < m^{K+1}$  with  $\delta(A, I_K) \geq 2^K \delta(A, I_0)$ . This yields

$$\delta(A, [0, N - 1]) \leq 2\delta(A, I_0) \leq 2^{-(K-1)}\delta(A, I_K) \leq 2^{-(K-1)}$$

so that if we let  $\alpha = \frac{(K-1)\log 2}{2K \log m} - \frac{1}{K}$  and  $\varepsilon = st(\alpha/2)$  we have

$$\begin{aligned} |A \cap [0, N - 1]| &\leq 2^{-(K-1)}N \leq 2^{-(K-1)}m^{2K+2} = m^{2K+2-(K-1)\frac{\log 2}{\log m}} \\ &= (m^{2K})^{1-\alpha} \leq N^{1-\varepsilon}. \end{aligned}$$

Since  $\frac{1}{K}$  is infinitesimal,  $st(\alpha/2) = \frac{\log 2}{4 \log m}$ . Thus  $\varepsilon$  is greater than 0, completing the proof.  $\square$

Significant work still needs to be done regarding questions of this type, i.e. results that link the density of a sequence with the existence of cells of certain types. The interest in such results is underscored by the theorem and corollary below. They give one of many possible conditions that might provide an alternate means of

approaching the Erdős-Turán conjecture. We begin with a conjecture (the term conjecture is perhaps optimistic here).

**Conjecture 1.** *Suppose that  $t$  is given and  $c > 0$  is constant. Then there exists  $n_0$  such that for all  $n > n_0$ , whenever the sequence  $A$  satisfies*

$$|A \cap [0, n - 1]| > \frac{n}{(c \log n)^{2 \log \log n}},$$

*then  $A$  contains a  $\langle t, d, w \rangle$  cell on  $[0, n - 1]$  with  $\frac{w}{d} < \frac{d}{n}$ , where both  $w$  and  $d$  are powers of 2.*

In the above conjecture we recall that a  $\langle t, d, w \rangle$  cell is merely a collection of intervals on which  $A$  is nonempty in arithmetic progression, with width  $w$  and distance  $d$  between intervals. The proof of the theorem below underscores the desirability of having both  $w$  and  $d$  be powers of 2, a condition that has been in each result but that so far has not been used.

**Theorem 3.** *If the conjecture is true for fixed  $t$ , then for  $c$  constant and  $A \subset \mathbb{N}$ , if  $|A \cap [0, n - 1]| > \frac{cn}{(\log n)^2}$  for arbitrarily large  $n$ , then  $A$  contains arithmetic progressions of length  $t$ .*

*Proof.* For convenience we will use  $\log$  base 2 here. Let  $A$  satisfy  $|A \cap [0, n - 1]| > \frac{cn}{(\log n)^2}$  for infinitely many  $n$ . By adjusting the constant if necessary (and using  $2A$ ), we may assume that  $A$  contains no two consecutive numbers.

Let  $N = 2^{2^L}$ , where  $L \in {}^*\mathbb{N} - \mathbb{N}$  is such that there exists  $M$  with

$$2^{2^L} \leq M < 2^{2^{L+1}} \quad \text{and} \quad |{}^*A \cap [0, M - 1]| > \frac{cM}{(\log M)^2}.$$

Such an  $L$  exists by overspill.

Since  ${}^*A$  has density greater than  $\frac{c}{(\log M)^2}$  on an interval of length  $M$ , with  $N \leq M$ , then for each  $n \leq N$  there must be an interval  $I$  such that the density of  ${}^*A$  on  $I$  is greater than  $\frac{c}{2(\log M)^2}$ . But

$$\frac{c}{2(\log M)^2} \geq \frac{c}{2(2^{L+1})^2} = \frac{c/8}{(2^L)^2} = \frac{c/8}{(\log N)^2}.$$

We will define a subset  $B$  of  $[0, N - 1]$  with the property that  $|B \cap [0, N - 1]| > \frac{N}{(\sqrt{8/c} \log N)^{2 \log \log N}}$  by essentially using block copies of segments of  ${}^*A$ . More specifically, for  $0 \leq k < L$  we let  $[x_k, y_k]$  be an interval in  ${}^*\mathbb{N}$  of length  $2^{2^k}$  such that

$$\delta({}^*A, [x_k, y_k]) > \frac{c/8}{(\log N)^2},$$

and we let

$$i \in B_{k+1} \text{ iff } \left\lfloor \frac{i'}{(2^{2^L - 2^{k+1}})} \right\rfloor \in {}^*A - x_k, \text{ where } i' \text{ is the remainder of } i \text{ mod } (2^{2^L - 2^k}).$$

We note that  $\frac{2^{2^L - 2^k}}{2^{2^L - 2^{k+1}}} = 2^{2^k}$  so that the construction guarantees that at least a  $\frac{c/8}{(\log N)^2}$  portion of  $B_k$  always intersects with  $B_1 \cap \dots \cap B_{k-1}$ . Thus  $B_1 \cap \dots \cap B_k$  has cardinality at least  $\frac{(c/8)^k N}{(\log N)^{2^k}}$ . We now let

$$B = B_1 \cap \dots \cap B_L$$

so that

$$B \cap [0, N - 1] > \frac{(c/8)^L N}{(\log N)^{2L}} = \frac{N}{(\sqrt{8/c} \log N)^{2 \log \log N}}.$$

Since we are assuming that the conjecture is true,  $B$  contains a  $\langle t, d, w \rangle$  cell on  $[0, N - 1]$  with  $\frac{w}{d} < \frac{d}{n}$ , where both  $w$  and  $d$  are powers of 2. We show that this forces an actual arithmetic progression of length  $t$  in  ${}^*A$  and thus in  $A$  (by transfer).

To see this, we let  $k$  be such that  $2^{2^L - 2^{k+1}} \leq d < 2^{2^L - 2^k}$ . Then since  ${}^*A$  contains no two consecutive numbers, the  $\langle t, d, w \rangle$  cell on  $[0, N - 1]$  must be completely contained inside one of the blocks of length  $2^{2^L - 2^k}$ , i.e. inside some  $[\nu 2^{2^L - 2^k}, (\nu + 1)2^{2^L - 2^k} - 1]$ . But

$$w \leq \left(\frac{d}{N}\right) d < (2^{-2^k}) 2^{2^L - 2^k} = 2^{2^L - 2^{k+1}},$$

and since  $w$  and  $d$  are powers of 2

$$w | 2^{2^L - 2^{k+1}} \quad \text{and} \quad 2^{2^L - 2^{k+1}} | d,$$

so that there exist  $t$  intervals of length  $2^{2^L - 2^{k+1}}$  inside  $[\nu 2^{2^L - 2^k}, (\nu + 1)2^{2^L - 2^k} - 1]$  which contain elements of  $B$  and are in arithmetic progression. By the construction this means that  ${}^*A$  contains an arithmetic progression of length  $t$ , completing the proof.  $\square$

**Corollary 2.** *The Erdős-Turán conjecture follows from the above conjecture.*

*Proof.* Let  $A = \langle a_n \rangle$  (indexed in increasing order) and suppose that  $\sum_{n=1}^\infty 1/a_n$  diverges. We wish to show that there exist arbitrarily large  $n$  such that

$$|A \cap [0, n - 1]| > \frac{n}{(\log n)^2},$$

so that the result follows from the theorem above. If this is not the case, then there exists some  $n_0$  such that whenever  $n > n_0$ ,  $|A \cap [0, n - 1]| \leq \frac{n}{(\log n)^2}$ . Then for  $n > n_0$ , if  $a_n < \frac{1}{2}n(\log n)^2$ , then there are at least  $n$  elements in  $A$  less than  $\frac{1}{2}n(\log n)^2$ , i.e.

$$\left| A \cap \left[0, \frac{1}{2}n(\log n)^2 - 1\right] \right| \geq n.$$

But since  $\frac{1}{2}n(\log n)^2 > n_0$  (we assume  $n \geq 4$ ), we also have

$$\begin{aligned} \left| A \cap \left[0, \frac{1}{2}n(\log n)^2 - 1\right] \right| &\leq \frac{(1/2) n(\log n)^2}{(\log((1/2) n(\log n)^2))^2} \\ &= \frac{(1/2) n(\log n)^2}{(\log(1/2) + \log n + \log \log n)^2} \leq (1/2) n. \end{aligned}$$

This contradiction shows that for  $n > n_0$ ,  $a_n \geq \frac{1}{2}n(\log n)^2$ , so that

$$\sum_{n=n_0}^\infty \frac{1}{a_n} \leq 2 \sum_{n=n_0}^\infty \frac{1}{n(\log n)^2}$$

which converges, contradicting the assumption about  $A$ . Thus, there do exist arbitrarily large  $n$  such that  $|A \cap [0, n - 1]| > \frac{n}{(\log n)^2}$ , completing the proof.  $\square$

Clearly the conjecture would be a significant improvement over Corollary 1. It seems conceivable that the additional uniformity conditions in the corollary and in the main theorem might help with the task of improving the “smallness” conditions on  $w$  necessary to obtain the conjecture. The extent to which Theorem 2 can be improved is open, and it is possible that improvements in the “smallness” conditions might provide another possible means of approaching the Erdős-Turán conjecture. Improvements appear to be somewhat delicate, as it is possible to show that for all  $\alpha > 0$  there exist sequences  $A$  in  $\mathbb{N}$  with the property that for all  $\varepsilon > 0$ ,  $|A \cap [0, n - 1]| > n^{1-\varepsilon}$  for sufficiently large  $n$ , and yet  $A$  contains no  $\langle t, d, w \rangle$  cell satisfying  $\frac{w}{d} < \left(\frac{d}{n}\right)^\alpha$ . In fact, if for some  $\alpha$  every sequence with this property did contain such a cell, then by a similar technique as was used in the proof of Theorem 3, we could show that every sequence such that

$$|A \cap [0, n - 1]| > \frac{n}{e^{c\sqrt{\log n}}}$$

would contain arbitrarily long arithmetic progressions. But in [5] it is shown that there exist sets of this density without even a three-termed arithmetic progression.

*Remark 1.* Nonstandard methods are used in the proofs of all three theorems, but in very different ways. The proofs of Theorems 2 and 3 could be translated into standard proofs in a relatively straightforward way. Several times where we used the fact that one quantity was infinitesimal compared to another quantity, we would simply need to substitute a more careful limit argument. The proof of Theorem 1 is completely different. Here the use of the nonstandard model allows us to use deep results from analysis and measure theory in ways that are completely unavailable with a standard approach. No hint of how to obtain a standard proof exists within the proof given here.

#### REFERENCES

- [1] Erdős, P. and Turán, P. “On some sequences of integers”, *J. London Math. Soc.*, 1936, **11**, 261-264
- [2] Furstenberg, H., “Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions”, *J. Anal. Math. Soc.*, 1977, **31**, 204-256 MR0498471 (58:16583)
- [3] Furstenberg, H., *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981 MR0603625 (82j:28010)
- [4] Gowers, T., “A new proof of Szemerédi’s theorem”, *GAFSA*, 2001, **11**, 465-588 MR1844079 (2002k:11014)
- [5] Graham, R., Rothschild, B., and Spencer, J., *Ramsey Theory, 2nd ed*, Wiley, 1990 MR1044995 (90m:05003)
- [6] Green, B. and Tao, T., “The Primes Contain Arbitrarily Long Arithmetic Progressions.” Preprint. 8 Apr 2004. <http://arxiv.org/abs/math.NT/0404188>.
- [7] Henson, C.W., “Foundations of nonstandard analysis-A gentle introduction to nonstandard extensions, in *Nonstandard Analysis: Theory and Applications*, ed. by N.J. Cutland, C.W. Henson, and L. Arkeryd, Kluwer Academic Publishers, 1997 MR1603228 (99i:03085)
- [8] Jin, R., “The sunset phenomenon”, *Proc. of the Amer. Math. Soc.*, 2002, **130** no. 3, 855-861 MR1866042 (2002h:03149)
- [9] Jin, R., “Nonstandard methods for upper Banach Density Problems”, *J. of Number Theory*, 2000, **91**, 20-38 MR1869316 (2002h:11011)
- [10] Jin, R., “Applications of nonstandard analysis in additive number theory”, *Bull. of Symbolic Logic*, 2001, **6** no. 3, 20-38 MR1803637 (2001k:11262)
- [11] Jin, R., “Standardizing nonstandard methods for upper banach density problems”, *Unusual Applications of Number Theory*, ed. by Melvyn Nathanson, DIMACS series, vol. 64, 2004 MR2063207 (2005e:11008)

- [12] Jin, R., “The inverse problem for upper asymptotic density”, *Tran. of the Amer. Math. Soc.*, 2003, **355** no. 1, 57-78 MR1928077 (2004a:11006)
- [13] Jin, R., “The inverse problem for upper asymptotic density II”, to appear
- [14] Jin, R., “Solution to the inverse problem for upper asymptotic density”, to appear
- [15] Jin, R., “Freiman’s conjecture and nonstandard methods”, to appear
- [16] Leth, S., “Applications of nonstandard models and lebesgue measure to sequences of natural numbers”, *Tran. of the Amer. Math. Soc.*, 1988, **307** no. 2, 457-468 MR0940212 (89g:11012)
- [17] Leth, S., “Sequences in countable nonstandard models of the natural numbers”, *Studia Logica*, 1988, **XLVII** no. 3, 63-83 MR0999780 (90h:03054)
- [18] Leth, S., “Some nonstandard methods in combinatorial number theory”, *Studia Logica*, 1988, **XLVII** no. 3, 85-98 MR0999781 (90c:03067)
- [19] Lindstrom, T., “An invitation to nonstandard analysis”, in *Nonstandard Analysis and Its Application*, ed. by N. Cutland, Cambridge University Press, 1988 MR0971064
- [20] Loeb, P. and Wolff, M., *Nonstandard Analysis for the Working Mathematician*, Kluwer, 2000 MR1790871 (2001e:03006)
- [21] Szemerédi, E., “On sets of integers containing no  $k$  elements in arithmetic progression”, *Acta Arith.*, 1975, **27**, 199-245 MR0369312 (51:5547)

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF NORTHERN COLORADO, GREELEY,  
COLORADO 80639

*E-mail address:* `steven.leth@unco.edu`