Combinatorial Proof of Szemerédi's Theorem in Nonstandard Analysis[†] [‡]

Renling Jin[§]

College of Charleston
jinr@cofc.edu

Abstract

Following Terence Tao's interpretation of Szemerédi's original proof, we present a nonstandard proof of Szemerédi's Theorem.

1 Introduction

This article is basically a re-writing of Tao's notes [6] in the language of nonstandard analysis. The story of Szemerédi's Theorem starts with van der Waerden's Theorem. Let $\mathbb{N}$ be the set of all standard positive integers and $\mathbb{R}$ be the set of all standard real numbers.

Theorem 1.1 (van der Waerden, 1927) *Given any $k, n \in \mathbb{N}$, there exists a $W(k,n) \in \mathbb{N}$ such that if $\{U_1, U_2, \ldots, U_n\}$ is a partition of $\{1, 2, \ldots, W(k,n)\}$, then there is an $i \leq n$ such that U_i contains a k–term arithmetic progression.*

Conjecture 1.2 (P. Erdős and P. Turán, 1936) *If $X \subseteq \mathbb{N}$ has a positive upper density, then X contains a k–term arithmetic progression for every $k \in \mathbb{N}$.*

[†] *Mathematics Subject Classification 2020* Primary 11B25, Secondary 03H05

[‡] *Keywords:* arithmetic progression, Szemerédi's Theorem, nonstandard analysis

[§] *Acknowledgment:* This work was partially supported by a collaboration grant (ID: 513023) from Simons Foundation. The Author would like to thank the American Institute of Mathematics which sponsored the workshop *Nonstandard Methods in Additive Combinatorics*, where the author had an opportunity to attend Terence Tao's lectures and learn his interpretation of Szemerédi's original proof of Szemerédi's Theorem. The author would also like to thank Steven Leth and Isaac Goldbring for detecting some typos in previous version of the paper.

In 1953 K. F. Roth [5] gave a proof of the conjecture for $k = 3$ using harmonic analysis. In 1975, E. Szemerédi [1] gave a combinatorial proof of the full conjecture. Hence the conjecture is now called Szemerédi's Theorem. In 1977 H. Furstenberg [3] gave an ergodic proof of Szemerédi's Theorem. In 2001 T. Gowers [4] gave a harmonic proof of Szemerédi's Theorem with numerical information. All of these proofs are long and complicated.

In the workshop *Nonstandard methods in combinatorial number theory* sponsored by American Institute of Mathematics in San Jose, CA, August 2017, T. Tao gave a series of talks to explain the Szemerédi's original combinatorial proof and hope to simplify it so that it can be better understood. He believed that Szemerédi's combinatorial method should have a greater impact in combinatorics.

In the talks T. Tao challenged the audience to produce a nonstandard proof of Szemerédi's Theorem which is noticeably simpler and more transparent than Szemerédi's original proof. The current article is the product of Tao's challenge.

Because nonstandard analysis allows one to describe upper Banach density of a set X along unbounded arithmetic progressions by the density of ${}^*X$ in one arithmetic progression of hyperfinite length, and because one does not need to keep track of the interrelations between various parameters when these parameters are allowed to be hyperfinite, a nonstandard proof of Szemerédi's Theorem could be easier to read than the standard one for the reader with some nonstandard analysis or model theory background.

The paper is arranged in the following sections. In §2 we give a very brief introduction of nonstandard analysis enough for our purpose as well as notation used. In §3 we re-write the proofs of a so called mixing lemma based on a weak regularity lemma in a nonstandard setting. In §4 we present a proof of Roth's Theorem. In §5 we present a proof of Szemerédi's Theorem for $k = 4$ directly without going through induction. In §6 we re-write the proof of Szemerédi's Theorem for all k in a nonstandard setting.

2 Notation and Nonstandard Analysis

Let $V_0 := \mathbb{R}$ and $V_{n+1} = V_n \cup \mathcal{P}(V_n)$. By the standard model we mean the model $(V; \in)$ where

$$V(\mathbb{R}) = \bigcup_{n=0}^{\mathfrak{n}} V_n$$

for a fixed, standard, sufficiently large, positive integer $\mathfrak{n}$ together with the true membership relation $\in$ on V. Notice that all standard mathematical objects we are interested in this paper can be found in, say, V_{100}. Every $a \in \mathbb{R}$ is viewed as an urelement in $V(\mathbb{R})$, i.e., element without members in $V(\mathbb{R})$.

By a nonstandard model $({}^*V(\mathbb{R}); {}^*\!\in)$ we mean a proper elementary extension of $(V(\mathbb{R}); \in)$. Let $* : V(\mathbb{R}) \to {}^*V(\mathbb{R})$ be the elementary embedding. The element $*(A)$ is written as ${}^*A$ for every $A \in V(\mathbb{R})$. We can assume that $\mathbb{R} \subseteq {}^*\mathbb{R}$ and any $r \in {}^*\mathbb{R}$ is an urelement. We can also assume that ${}^*\!\in$ is the true membership relation among the elements in ${}^*V(\mathbb{R})$. For notational convenience we drop the symbol $*$ for ${}^*\!\in$, ${}^*\!\leq$, ${}^*\!+$, and ${}^*a$ when $a \in \mathbb{R}$.

A set A is called internal if $A \in {}^*V(\mathbb{R})$. Notice that if X is an infinite subset of $\mathbb{R}$, then $X \in \mathcal{P}({}^*\mathbb{R})$ but $X \notin {}^*\mathcal{P}(\mathbb{R})$. So X is not an internal subset of ${}^*\mathbb{R}$ but instead ${}^*X$ is an internal subset of ${}^*\mathbb{R}$.

An integer in ${}^*\mathbb{N} \smallsetminus \mathbb{N}$ is called hyperfinite. A hyperfinite integer is infinitely large in the standard point of view but finite in the nonstandard point of view.

An $r \in {}^*\mathbb{R}$ is an infinitesimal, denoted by $r \approx 0$, if $|r| < 1/n$ for every $n \in \mathbb{N}$. Let st be the standard part map, i.e., if $r \in {}^*\mathbb{R} \cap (-n, n)$ for some $n \in \mathbb{N}$ then $st(r)$ is the unique standard real number α such that $r - \alpha \approx 0$. Let $h_1, h_2 \in {}^*\mathbb{N}$. We say that h_1 is infinitesimally smaller than h_2, denoted by $h_1 \ll h_2$, if $h_1/h_2 \approx 0$. Hence h is hyperfinite if and only if $h \gg 1$.

Capital letters A, B, C, ... represent sets of integers except H, J, K, L, M, N which are reserved for hyperfinite integers. The letter k represents exclusively the length of the a.p. in Szemerédi's Theorem. All sets mentioned in this paper will be either standard or internal sets of integers. If $n \in {}^*\mathbb{N}$, let $[n] := \{1, 2, \ldots, n\}$. If $A \subseteq {}^*\mathbb{N}$ and $b \in {}^*\mathbb{N}$, let $bA = \{ba \mid a \in A\}$.

If A is an internal set, we write $\delta_H(A)$ for the quantity $|A|/H$ and $\mu_H(A) := st(\delta_H(A))$ where st is the standard part map in nonstandard analysis. Notice that δ_H is an internal function and μ_H is an external function. If $A \subseteq \Omega$ and $|\Omega| = H$, then $\mu_H(A)$ coincides with the Loeb measure of A in Ω. Our definition of $\mu_H(A)$ allows the value to be > 1 or even ∞ (in the standard sense). We often use δ_H in an internal argument. Other time use use μ_H to shorten an expression. For example, $\mu_H(A) > 0$ is a short form of $\delta_H(A) > 0$ and $\delta_H(A) \not\approx 0$.

We assume that the nonstandard model $({}^*V; \in)$ is countably saturated. As a consequence of the countable saturation we have that if $\{A_n \mid n \in \mathbb{N}\}$ is a nested sequence of non-empty internal sets, then $\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset$. So if $\{H_n \mid n \in \mathbb{N}\}$ is a decreasing sequence of hyperfinite integers, then there is a hyperfinite integer H which is a lower bound of all H_n's. Also, if $S_n \subseteq S_{n-1} \subseteq [N]$ and $\mu_N(S_n) = 1$ for all $n \in \mathbb{N}$, then there is an internal set $S \subseteq \bigcap_{n\in\mathbb{N}} S_n$ with $\mu_N(S) = 1$.

The abbreviation a.p. stands for "arithmetic progression" and k–a.p. stands for "k-term arithmetic progression." We write P and Q for a.p.'s of hyperfinite length, and p and q for a.p.'s of length k or other finite length. The length of an a.p. P can also be written as $|P|$. If p (or P) is an a.p., then $p(i)$ ($P(i)$) represents the i-th term of the a.p. We allow the common difference d of a k–a.p. to be any integer including the trivial case that $d = 0$.

If p and q are two k–a.p.'s, then $p \oplus q$ represents the k–a.p. $\{p(i) + q(i) \mid i = 1, 2, \ldots, k\}$.

Definition 2.1 *Let $A \subseteq {}^*\mathbb{N}$ and $|A| \gg 1$. The upper Banach density $BD(A)$ of A is defined by*

$$BD(A) := \sup\{\mu_H(A \cap [a, a+H)) \mid a \in {}^*\mathbb{N} \text{ and } H \gg 1\}.$$

The strong upper Banach density $SD(A)$ of A is defined by

$$SD(A) := \sup\{\mu_{|P|}(A \cap P) \mid P \text{ is an a.p. and } |P| \gg 1\}.$$

Let $\gamma > 0$. If $S \subseteq {}^*\mathbb{N}$ has $SD(S) \geq \gamma$ and $A \subseteq {}^*\mathbb{N}$, the strong upper Banach density of A relative to S and γ is defined by

$$SD_{S,\gamma}(A) := \sup\{\mu_{|P|}(A \cap P) \mid P \text{ is an a.p., } |P| \gg 1 \text{ and } \mu_{|P|}(S \cap P) \geq \gamma.\}.$$

Notice that $\sup(X)$ for $X \subseteq \mathbb{R}$ is the supreme of X in the standard sense. Notice also that if $A \subseteq \mathbb{N}$, then the upper Banach density of A coincides with $BD({}^*A)$ defined above. Similarly $SD({}^*A)$ is the upper Banach density of A along arithmetic progressions defined in [6].

Clearly, $0 \leq BD(A) \leq SD(A) \leq 1$ for each $A \subseteq {}^*\mathbb{N}$. On the other hand, if A is the set of all even numbers, then $BD(A) = 1/2$ and $SD(A) = 1$. It is also clear that if $S, B \subseteq [N]$ with $\mu_N(S) = \gamma$, then $SD_{S,\gamma}(B) \geq \mu_N(B)$.

Lemma 2.2 *Let $A \subseteq \mathbb{N}$ and $\alpha \in [0,1]$. Then $SD(A) \geq \alpha$ if and only if there exists an a.p. $P \subseteq {}^*\mathbb{N}$ with $|P| \gg 1$ such that $\mu_{|P|}({}^*A \cap P) \geq \alpha$.*

Proof The sufficiency is a consequence of the transfer principle and the necessity is a consequence of countable saturation. □

Lemma 2.3 *Let $N \gg 1$ and $A \subseteq [N]$ with $\mu_N(A) = BD(A) = \alpha$. Let H be such that $1 \ll H \leq N/2$. There exists an internal set $S \subseteq [N-H]$ such that $\mu_{N-H}(S) = 1$ and $\mu_H(A \cap (x + [H])) = \alpha$ for every $x \in S$.*

Proof If $st(H/N) > 0$, then $S = [N-H]$ works due to the supremality of α. So we can assume that $H \ll N$. Let $j \in {}^*\mathbb{N}$ and

$$S_j = \{x \in [N-H] \mid ||A \cap (x + [H])|/H - \alpha| < 1/j\}.$$

Then for each standard $j \in \mathbb{N}$, we have that S_j is internal and $\mu_{N-H}(S_j) = 1$ by a double counting argument. By countable saturation we can find a hyperfinite integer J such that $\mu_{N-H}(S_J) = 1$ and let $S = S_J$. □

3 Mixing Lemma

The following is a consequence of the weak regularity lemma from [2] presented in [6]. The weak regularity lemma is of course a consequence of the original regularity lemma by Szemerédi in [1]. Same as in [6] we do not prove the following lemma.

Lemma 3.1 *Let V, W be finite sets, let $\epsilon > 0$, and for each $w \in W$, let E_w be a subset of V. Then there exists a partition $V = V_1 \cup V_2 \cup \cdots \cup V_n$ with $n = O(b^{1/\epsilon})$ for some standard real $b > 1$, and real numbers $0 \leq c_{i,w} \leq 1$ for $i \leq n$ and $w \in W$ such that for any set $F \subseteq V$, one has*

$$\left| |F \cap E_w| - \sum_{i=1}^{n} c_{i,w} |F \cap V_i| \right| \leq \epsilon |V|$$

for all but $\epsilon|W|$ values of $w \in W$.

For the existence of n above, see [6, paragraph right below Lemma 1.3]).

Let $m', m'' \in {}^*\mathbb{N}$ and $W = W(m', m'') \in {}^*\mathbb{N}$ be the Van der Waerden number that if $[W]$ is partitioned into m' cells, one of these cells must contain an m''–a.p.

From Lemma 3.1, the following lemma called mixing lemma in [6] can be derived.

Lemma 3.2 (Mixing Lemma) *Let $N \gg 1$, $A \subseteq [N]$, $1 \ll H \leq N/2$, and $R \subseteq [N - H]$ be such that*

$$\mu_N(A) = SD(A) = \alpha > 0 \text{ and } \mu_H((x + [H]) \cap A) = \alpha$$

for every $x \in R$. Then

(i) For a set $E \subseteq [H]$ with $\mu_H(E) > 0$ and an a.p. $P \subseteq R$ of hyperfinite length, there is an $x \in P$ such that

$$\mu_H(A \cap (x + E)) \geq \alpha \mu_H(E);$$

(ii) Given an a.p. $P \subseteq R$ with $|P| = m \gg 1$, let $m' \gg 1$ be such that $W\left(3^{m'}, m'\right) \leq m$. For any internal partition $\{V_i \mid i \in [m']\}$ of $[H]$ there exists an m'–a.p. $P' \subseteq P$, a set $I \subseteq [m']$ with $\mu_H(U) = 1$ where $U = \bigcup\{V_i \mid i \in I\}$, and an infinitesimal $\epsilon > 0$ such that

$$|\delta_H(A \cap (x + V_i)) - \alpha \delta_H(V_i)| \leq \epsilon \delta_H(V_i)$$

for all $i \in I$ and all $x \in P'$;

(iii) Given an a.p. $P \subseteq R$ *with* $|P| = m \gg 1$ *and an internal collection of sets* $\{E_w \subseteq [H] \mid w \in W\}$ *with* $|W| \gg 1$ *and* $\mu_H(E_w) > 0$ *for every* $w \in W$*, there exists an* $x \in P$ *and* $T \subseteq W$ *such that* $\mu_{|W|}(T) = 1$ *and*

$$\mu_H(A \cap (x + E_w)) = \alpha \mu_H(E_w)$$

for every $w \in T$*.*

Proof The proof of Part (i) is a double counting argument.

Assume that (i) is not true. Let r_x be such that $\delta_H(A \cap (E + x)) = (\alpha - r_x)\delta_H(E)$. Then r_x must be positive non-infinitesimal for every $x \in P$. Since the function $x \mapsto r_x$ is internal, there is a positive non-infinitesimal number $r = \min\{r_x \mid x \in P\}$. Hence $\delta_H(A \cap (E+x)) \leq (\alpha - r)\delta_H(E)$. Since

$$\mu_{|P|}(A \cap (y + P)) \leq \alpha$$

for all $y \in [H]$ by the maximality of α and $P \subseteq S$ implies $\mu_H(A \cap (x + [H])) = \alpha$ for all $x \in P$, we have that

$$\mu_{|P|}(A \cap (y + P)) = \alpha$$

for μ_H–almost all $y \in [H]$. Hence

$$\begin{aligned}
\alpha \mu_H(E) &\approx \frac{1}{H} \sum_{y \in E} \frac{1}{|P|} \sum_{x \in P} \chi_A(x + y) \\
&= \frac{1}{|P|} \sum_{x \in P} \frac{1}{H} \sum_{y=1}^{H} \chi_{A \cap (E+x)}(x + y) \\
&\leq \frac{1}{|P|} \sum_{x \in P} (\alpha - r)\delta_H(E) = (\alpha - r)\delta_H(E) \\
&\approx (\alpha - st(r))\mu_H(E) < \alpha \mu_H(E),
\end{aligned}$$

which is absurd.

The proof of Part (ii) is similar to the proof of Part (ii) of the mixing lemma in [6] with a small twist in nonstandard analysis. To make the argument internal explicitly we use δ_H instead of μ_H. Fix $j \in \mathbb{N}$. For each $x \in P$

let

$$c_i^j(x) = \begin{cases} 1 & \text{if } \delta_H((x+V_i)\cap A) \geq \left(\alpha + \frac{1}{j}\right)\delta_H(V_i), \\ 0 & \text{if } \left(\alpha - \frac{1}{j}\right)\delta_H(V_i) < \delta_H((x+V_i)\cap A) < \left(\alpha + \frac{1}{j}\right)\delta_H(V_i), \\ -1 & \text{if } \delta_H((x+V_i)\cap A) \leq \left(\alpha - \frac{1}{j}\right)\delta_H(V_i). \end{cases}$$

and let $c^j : P \to \{-1,0,1\}^{[m']}$ be such that $c^j(x)(y) = c_y^j(x)$. By van der Waerden's Theorem there exists a m'–a.p. $P'_j \subseteq P$ such that $c^j(x) = c^j(x')$ for any $x, x' \in P'_j$. Let $I_j^+ = \{i \in [m'] \mid c^j(x)(i) = 1\}$, $I_j^- = \{i \in [m'] \mid c^j(x)(i) = -1\}$, and $I_j = [m'] \smallsetminus (I_j^+ \cup I_j^-)$ for every $x \in P'_j$. Let $U_j^+ = \bigcup\{V_i \mid i \in I_j^+\}$, $U_j^- = \bigcup\{V_i \mid i \in I_j^-\}$, and $U_j = [H] \smallsetminus (U_j^+ \cup U_j^-)$. Clearly, $\delta_H((x+U_j^-)\cap A) \leq (\alpha - 1/j)\delta_H(U_j^-)$ because U_j^- is a disjoint union of the V_i's for $i \in I_j^-$. Hence $\mu_H(U_j^-) = 0$ by (i). Notice that $\delta_H(A \cap (x+U_j^+)) \geq (\alpha+1/j)\delta_H(U_j^+)$. Since $\alpha \geq \mu_H(A \cap (x+U_j^+)) \geq (\alpha+1/j)\mu_H(U_j^+)$, we have that $\mu_H(U_j^+) < 1$, which implies $\mu_H(U_j) > 0$. If $\mu_H(U_j^+) > 0$, then $\mu_H(A \cap (x+U_j)) < \alpha\mu_H(U_j)$ for all $x \in P'_j$, which contradicts (i). Hence $\mu_H(U_j^+) = 0$ and $\mu_H(U_j) = 1$.

By countable saturation there is a $J \gg 1$ such that $|P'_J| \gg 1$ and $\mu_H(U_J) = 1$. The proof of (ii) is complete by setting $P' = P'_J$, $I = I_J$, and $U = U_J$. Notice that $1/J \approx 0$.

We now prove Part (iii) using Lemma 3.1. Choose a sufficiently large positive infinitesimal ϵ satisfying that there is a hyperfinite partition of $[H] = V_0 \cup V_1 \cup \cdots \cup V_{m'}$ and real numbers $0 \leq c_{i,w} \leq 1$ for each $i \in [m']$ and $w \in W$ such that $W(3^{m'}, m') \leq m$, and for any internal set $F \subseteq [H]$ there is a $T_F \subseteq W$ with $|W \smallsetminus T_F| \leq \epsilon|W|$ such that

$$\left| |F \cap E_w| - \sum_{i=1}^{m'} c_{i,w}|F \cap V_i| \right| \leq \epsilon H \tag{1}$$

for all $w \in T_F$. Notice that such ϵ exists because if ϵ is a standard positive number, then m' is in $\mathbb{N}$ (see [6, paragraph right below Lemma 1.3]). From (1) with $F := [H]$ we have

$$\left| |E_w| - \sum_{i=1}^{m'} c_{i,w}|V_i| \right| \leq \epsilon H \tag{2}$$

for all $w \in T_{[H]}$.

By (ii) we can find an $x \in P$, a positive infinitesimal ϵ_1, and $I \subseteq [m']$ where

$$I = \{i \in [m'] \mid |\delta_H((x + V_i) \cap A) - \alpha\delta_H(V_i)| < \epsilon_1\delta_H(V_i)\},$$

and $U = \bigcup\{V_i \mid i \in I\}$ with $\mu_H(U) = 1$. Let $I' = [m'] \setminus I$ and $U' = [H] \setminus U$. Then for each $w \in T := T_{[H]} \cap T_{(A-x)\cap[H]}$ we have

$$\begin{aligned}
&|\delta_H(A \cap (x + E_w)) - \alpha\delta_H(E_w)| \\
&\quad \leq \frac{1}{H}\left(\left||(A - x) \cap E_w| - \sum_{i\in[m']} c_{w,i}|(A - x) \cap V_i|\right|\right. \\
&\qquad + \left|\sum_{i\in[m']} c_{w,i}|A \cap (x + V_i)| - \sum_{i\in[m']} c_{w,i}\alpha|V_i|\right| \\
&\qquad \left.+ \left|\alpha \sum_{i\in[m']} c_{w,i}|V_i| - \alpha \sum_{i\in[m']} |E_w|\right|\right) \\
&\quad \leq \epsilon + \frac{1}{H}\sum_{i\in I} c_{w,i}\epsilon_1|V_i| + 2\delta_H(U') + \alpha\epsilon \\
&\quad \leq \epsilon + \epsilon_1\delta_H(U) + 2\delta_H(U') + \alpha\epsilon \approx 0.
\end{aligned}$$

Hence $\mu_H(A \cap (x + E_w)) = \alpha\mu_H(E_w)$ for all $w \in T$. Notice that $\mu_{|W|}(T) = 1$ because $\epsilon \approx 0$ and $\mu_{|W|}(T_{[H]}) = \mu_{|W|}(T_{[H]\cap(A-x)}) = 1$. □

4 Roth's Theorem

Theorem 4.1 (K. F. Roth, 1953) *If $U \subseteq \mathbb{N}$ and $SD(U) > 0$, then U contains nontrivial 3-term arithmetic progressions.*

Proof Let $\alpha = SD(U)$. Then $\alpha > 0$. Let $P \subseteq {}^*\mathbb{N}$ be an a.p. with $|P| \gg 1$ and $\mu_{|P|}({}^*U \cap P) = \alpha$. Without loss of generality we can assume that $P = [N] \cup \{0\}$. Let $A :={}^* U \cap [N]$. It suffices to find a 3–a.p. in A.

Let $H = \lfloor N/6 \rfloor$ and $S = [N - H] \cup \{0\}$. Notice that $\{0\} \cup (H + [H]) \cup (2H + 2[H]) \subseteq S$. For each $t \in [H]$ let

$$\mathcal{Q}_t = \{q \subseteq [H] \mid q \text{ is a 3–a.p.}, q(1) \in A \cap [H], \text{ and } q(3) = t\}$$

$$\text{and } E_t = \{q(2) \mid q \in \mathcal{Q}_t\}.$$

Notice that $\mu_H(E_t) = \alpha/2 > 0$ because $p(1) - t$ must be even and the density of A in an a.p. of difference 2 and length $\geq \lfloor N/16 \rfloor$ is also α. By (iii) of Lemma 3.2, there is an $l \in [H]$ and $T \subseteq [H]$ with $\mu_H(T) = 1$ such that

$$\mu_H(A \cap (H + l + E_t)) = \alpha^2/2$$

for all $t \in T$. Since $2H + 2l \in S$ and $\mu_H(T) = 1$, we have

$$\mu_H(A \cap (2H + 2l + T)) = \alpha > 0.$$

Let $t_0 \in T$ be such that $2H + 2l + t_0 \in A \cap (2H + 2l + T)$. Let $p_0 = \{0, H + l, 2H + 2l\}$ and $q_0 \in \mathcal{Q}_{t_0}$. Then $p_0 \oplus q_0$ is an 3–a.p. Clearly, $p_0(3) + q_0(3) = 2H + 2l + t_0 \in (2H + 2l + T) \cap A \subseteq A$, $p_0(2) + q_0(2) \in (H + l + E_{t_0}) \cap A \subseteq A$, and $p_0(1) + q_0(1) = q(1) \in A$ by the definition of E_{t_0}. □

Remark 4.2 *(i) Notice that there are about αH many choices of $t \in T \cap (A - 2H - 2l)$ and for each $t \in T \cap (A - 2H - 2l)$, there are about $\alpha H/2$ choices of $q \in \mathcal{Q}_t$.*

(ii) In the proof above we choose $p_0 = \{0, H + l, 2H + 2l\}$ instead of $\{0, h + l, 2h + 2l\}$ for $h < H$ because we want to make sure the 3–a.p. $p_0 \oplus q_0$ is nontrivial.

5 Szemerédi's Theorem for $k = 4$

We found that the proof of Szemerédi's Theorem for $k = 4$ does not require the Tower of Hanoi type induction ([6, Theorem 6.6] or Theorem 6.4 below). In this section we present a direct proof. This proof can be viewed as a warm-up for next section.

Lemma 5.1 *Let $N \gg 1$, $A \subseteq [N]$ be such that $\mu_N(A) = SD(A) = \alpha > 0$, and $H = \lfloor N/8 \rfloor$. There exists an interval $x_0 + [H] \subseteq [N]$, a set $T \subseteq x_0 + [H]$ with $\mu_H(T) = 1$, and*

$$\mathcal{P}_t := \{p \subseteq [N] \mid p \text{ is a 4–a.p.}, p(1), p(2) \in A, \text{ and } p(4) = t\}$$

such that $\mu_H(\mathcal{P}_t) = \alpha^2/3$ for each $t \in T$.

Proof The proof is almost the same as the proof of Theorem 4.1.

Let $S := [N-H] \cup \{0\}$. Notice that $\{0\} \cup (H+[H]) \cup (2H+2[H]) \cup (3H+3[H]) \subseteq S$.

For each $w \in [H]$ let

$$\mathcal{Q}_w := \{q \subseteq [H] \mid q \text{ is a 4–a.p., } q(1) \in A, \text{ and } q(4) = w\}$$

$$\text{and } E_w = \{q(2) \mid q \in \mathcal{Q}_w\}.$$

Clearly, $\mu_H(E_w) = \alpha/3$ because $|q(4) - q(1)|$ must be a multiple of 3. By (iii) of Lemma 3.2, there is an $l \in [H]$ and $W \subseteq [H]$ with $\mu_H(W) = 1$ such that

$$\mu_H(A \cap (H + l + E_w)) = \alpha\mu_H(E_w) = \alpha^2/3$$

for all $w \in W$. Now let $x_0 := 3H + 3l$ and $T := 3h + 3l + W \subseteq x_0 + [H]$. Let $p_0 := \{0, H+l, 2H+2l, 3H+3l\}$. For each $t = 3H + 3l + w \in T$, let

$$\mathcal{P}_t := \{p_0 \oplus q \mid q \in \mathcal{Q}_w \text{ and } H + l + q(2) \in (H + l + E_w) \cap A\}.$$

Then $\mu_H(\mathcal{P}_t) = \mu_H((H + l + E_w) \cap A) = \alpha^2/3$ and for each $p_0 \oplus q \in \mathcal{P}_t$ we have that $p_0(4) + q(4) = 3H + 3l + w = t$, $p_0(2) + q(2) = H + l + q(2) \in (H + l + E_w) \cap A \subseteq A$, and $p_0(1) + q(1) = q(1) \in [H] \cap A$ by the definition of $\mathcal{Q}_w$. □

Lemma 5.2 *Let $N \gg 1$, $B, S_\gamma \subseteq [N]$ be such that $B \subseteq S_\gamma$, $\mu_N(S_\gamma) = SD(S_\gamma) = \gamma > 11/12$, $\mu_N(B) = SD_{S_\gamma,\gamma}(B) = \beta > 0$. There exists an interval $x_0 + [\lfloor N/24 \rfloor] \subseteq [N]$ and a set $T \subseteq x_0 + [\lfloor N/24 \rfloor]$ with $\mu_{N/24}(T) \geq 1 - 12(1-\gamma)$, and a collection of 4–a.p.'s $\{p_x \mid x \in T\}$ such that $p_t(1), p_t(2) \in B$, $p_t(3), p_t(4) \in S_\gamma$, and $p_t(3) = t$ for each $t \in T$.*

Proof Let $H := [N/8]$. Notice that $\mu_H(S_\gamma \cap (x + [H])) = \gamma$ and $\mu_H(B \cap (x + [H])) = \beta$ for every $x \in [N-H] \cup \{0\}$. Let $\mathcal{Q}$ be the collection of all 4–a.p.'s in $[H]$. For each $w \in [\lfloor H/3 \rfloor, \lfloor 2H/3 \rfloor]$ let

$$\mathcal{Q}_w^3 := \{q \in \mathcal{Q} \mid q(1) \in B \text{ and } q(3) = w\}$$

$$\text{and } E_w^3 := \{q(2) \mid q \in \mathcal{Q}_w^3\}.$$

We have that $\mu_H(E^3_w) = \beta/2$. For each $w' \in [H]$ let

$$\mathcal{R}^i_{w'} := \{q \in \mathcal{Q} \mid q(1) \in B \text{ and } q(i) = w'\}$$

$$\text{and } F^i_{w'} := \{q(2) \mid q \in \mathcal{R}^i_{w'}\}$$

for $i = 3, 4$. Clearly, $\mu_H(F^i_{w'}) \leq \beta$.

By (iii) of Lemma 3.2, there is an $l \in [H]$, $W_3 \subseteq [\lfloor H/3 \rfloor, \lfloor 2H/3 \rfloor]$ with $\mu_H(W_3) = 1/3$, and $W^i \subseteq [H]$ with $\mu_H(W^i) = 1$ such that

$$\mu_H(B \cap (H + l + E^3_w)) = \frac{\beta^2}{2} \text{ and } \mu_H(B \cap (h + l + F^i_{w'})) \leq \beta^2$$

for all $w \in W_3$ and $w' \in W^i$ for $i = 3$ or 4. Clearly, $\mu_H(((i-1)H + (i-1)l + W^i) \cap S_\gamma) = \gamma$ for $i = 3$ or 4.

Let $T^3 := 2H + 2l + W_3$. For each $t = 2H + 2l + w \in T^3$ let

$$\mathcal{P}_t := \{p \subseteq [N] \mid p(1) \in B \cap [H], p(2) \in B \cap (H + l + E^3_w), p(3) = t\}$$

$$\text{and } \mathcal{P} := \bigcup_{i \in T^3} \mathcal{P}_t.$$

Notice that $\mu_H(\mathcal{P}_t) = \mu_H(B \cap (2H + 2l + E^3_w)) = \beta^2/2$ for each $t = 2H + 2l + w \in T^3$.

A 4–a.p. $p \in \mathcal{P}$ is called *good* if $p(i) \in S_\gamma \cap ((i-1)H + (i-1)l + [H])$ for $i = 3, 4$. Let $\mathcal{P}_g$ be the collection of all good 4–a.p.'s in $\mathcal{P}$. A 4–a.p. $p \in \mathcal{P}$ is *bad* if it is not good. Let $\mathcal{P}_b := \mathcal{P} \smallsetminus \mathcal{P}_g$. Let $T^3_g := \{p(3) \mid p \in \mathcal{P}_g\}$. Then $T^3_g \subseteq S_\gamma$. We show that $\mu_H(T^3_g) \geq \dfrac{1}{3} - 4(1-\gamma)$.

Notice that $\mathcal{P}_b \subseteq \bigcup_{i=3,4}\{p \in \mathcal{P} \mid p(1) \in B \cap [H], p(2) \in B \cap (h + l + [H]), p(i) \notin S_\gamma\}$. Hence

$$|\mathcal{P}_b| \leq \sum_{i=3}^{4} \sum_{w' \in [H] \smallsetminus (S_\gamma - (i-1)H - (i-1)l)} |F^i_{w'}|$$

$$\leq \sum_{i=3}^{4} \left(\sum_{w' \in [H] \smallsetminus W^i} |F^i_{w'}| + \sum_{w' \in W^i \smallsetminus (S_\gamma - (i-1)H - (i-1)l)} |F^i_{w'}| \right).$$

So $|\mathcal{P}_g| = |\mathcal{P}| - |\mathcal{P}_b|$

$$\geq \sum_{t\in T^3} |\mathcal{P}_t| - \sum_{i=3}^{4} \left(\sum_{w'\in[H]\smallsetminus W^i} |F^i_{w'}| + \sum_{w'\in W^i\smallsetminus(S_\gamma-(i-1)H-(i-1)l)} |F^i_{w'}| \right).$$

Hence we have

$$\mu_H(T^3_g)\cdot\frac{\beta^2}{2} = st\left(\frac{1}{H}\sum_{t\in T^3_g}\frac{1}{H}|\mathcal{P}_t|\right) \geq st\left(\frac{1}{H^2}|\mathcal{P}_g|\right) = st\left(\frac{1}{H^2}(|\mathcal{P}|-|\mathcal{P}_b|)\right)$$

$$\geq st\left(\frac{1}{H^2}\sum_{t\in T^3}|\mathcal{P}_t|\right)$$

$$-st\left(\frac{1}{H^2}\sum_{i=3}^{4}\left(\sum_{w'\in[H]\smallsetminus W^i}|F^i_{w'}| + \sum_{w'\in W^i\smallsetminus(S_\gamma-(i-1)H-(i-1)l)}|F^i_{w'}|\right)\right)$$

$$\geq \mu_H(T^3)\cdot\frac{\beta^2}{2} - 2(1-\gamma)\cdot\beta^2 = \left(\frac{1}{3}-4(1-\gamma)\right)\cdot\frac{\beta^2}{2},$$

which implies $\mu_H(T^3_g) \geq \dfrac{1}{3} - 4(1-\gamma)$. Hence $\mu_{N/24}(T^3_g) = 1-12(1-\gamma)$ because $H = \lfloor N/8 \rfloor$. Now the lemma is proven if we set $x_0 := 2H + 2l + \lfloor H/3 \rfloor - 1$, $T := T^3_g$, and choose one $p_t \in \mathcal{P}_g$ such that $P_t(3) = t$ for each $t \in T$. □

Remark 5.3 *The argument for showing $\mu_{N/24}(T^3_g) > 1 - 12(1-\gamma)$ is from [6, Page 34].*

Theorem 5.4 (E. Szemerédi, 1969) *If $U \subseteq \mathbb{N}$ and $SD(U) > 0$, then U contains nontrivial 4-term arithmetic progressions.*

Proof Let $N \gg 1$ and $A \subseteq [N]$ be such that $\mu_N(A) = SD(A) = \alpha > 0$. Same as in the beginning of the proof of Theorem 4.1, it suffices to find a 4–a.p. in A. For each $n, j \in \mathbb{N}$ let

$$S_{j,n} := \{x \in [N-n] \mid \mu_n((x+[n]) \cap A) \geq \alpha - 1/j\}.$$

Then $\lim_{n\to\infty} \mu_{N-n}(S_{j,n}) = 1$ by Lemma 2.3. So for all sufficiently large $n \in \mathbb{N}$ we assume that $\gamma_{j,n} := SD(S_{j,n}) > 11/12$. Let $R_{j,n}$ be an a.p. in $[N]$ with difference d and $|R_{j,n}| \gg 1$ such that $\mu_{|R_{j,n}|}(R_{j,n} \cap S_{j,n}) = \gamma_{j,n}$. For each $\tau \subseteq [n]$ let

$$B_\tau := \{x \in [R_{j,n}] \mid A \cap (x + [n]) = x + \tau\}.$$

Then there is a $\tau_{j,n}$ such that $\mu_{|R_{j,n}|}(B_{j,n}) = \beta_{j,n} > 0$ because n is finite where $B_{j,n} := B_{\tau_{j,n}}$. Let $P_{j,n} \subseteq R_{j,n}$ be an a.p. of difference d with $|P_{j,n}| = N' \gg 1$, $\mu_{N'}(P_{j,n} \cap S_{j,n}) = \gamma_{j,n}$, and $\mu_{N'}(P_{j,n} \cap B_{j,n}) = \beta_{j,n}$. Let $\varphi : P_{j,n} \to [N']$ be the affine map, i.e., $\varphi(x) = (x - \min P_{j,n})/d + 1$. Applying Lemma 5.2 to $[N']$ for $S' = \varphi((S_{j,n}) \cap P_{j,n})$, and $B' = \varphi(B_{j,n} \cap P_{j,n})$, and then pulling back through φ^{-1}, we obtain $x_0 + d\,[\lfloor |P_{j,n}|/24 \rfloor] \subseteq P_{j,n}$ and $T_{j,n} \subseteq x_0 + d\,[\lfloor |P_{j,n}|/24 \rfloor]$ with $\mu_{N'/24}(T_{j,n}) \geq 1 - 12(1 - \gamma_{j,n})$, and there exists a collection of 4–a.p.'s $\mathcal{P}_{j,n} = \{p_t \mid t \in T_{j,n}\}$ such that $p_t(1), p_t(2) \in B_{j,n} \cap P_{j,n}$, $p_t(3), p_t(4) \in S_{j,n} \cap P_{j,n}$, and $p_t(3) = t$ for each $t \in T_{j,n}$.

By countable saturation we can find fixed hyperfinite integer H and then J such that $\gamma := \gamma_{J,H} \approx 1$, $P := P_{J,H}$ with $|P| \gg 1$, $S := S_\gamma$, $B := B_{J,H} \subseteq S$, $T := T_{J,H}$, and $\mathcal{P}_{J,H} = \{p_t \mid t \in T\}$ such that $p_t(1), P_t(2) \in B$, $p_t(3), p_t(4) \in S$, and $p_t(3) = t$ for each $t \in T$.

Notice that $\mu_{N-H}(S) = 1$, $T \subseteq x_0 + d\,[\lfloor |P|/24 \rfloor]$, $\mu_{|P|/24}(T) = 1$, $\gamma \approx 1$, $x, y \in B$ implies $((x + [n]) \cap A) - x = ((y + [H]) \cap A) - y$, and $x \in S$ implies $\mu_H((x + [H]) \cap A) = \alpha$. It may be the case that $\mu_{|P|}(B) = 0$. But the existence of the collection $\mathcal{P}_{J,H} = \{p_x \mid x \in T\}$ is guaranteed by countable saturation.

Since $\mu_{N/24}(T) = 1$, we can find an a.p. of $P' \subseteq T$ of difference d with $|P'| \gg 1$. Let $\mathcal{P}' := \{p_t \in \mathcal{P}_{J,H} \mid t \in P'\}$. Notice that for each $p_t \in \mathcal{P}'$ we have that $p_t(1), p_t(2) \in B$, $p_t(3) = t \in S$, and $p_t(4) \in S$.

Let $\tau_0 := ((x + [H]) \cap A) - x$ for some $x \in B$. Then $\mu_H(\tau_0) = \alpha$ because $B \subseteq S$. By Lemma 5.1 with N being replaced by H, A being replaced by τ, we can find $x_0 + [\lfloor H/8 \rfloor] \subseteq [H]$, $T_Q \subseteq x_0 + [\lfloor H/8 \rfloor]$ with $\mu_H(T_Q) = 1/8$,

$$\mathcal{Q}_w := \{q \subseteq [H] \mid q(1), q(2) \in \tau_0, \text{ and } p(4) = w\},$$

and $E_w = \{q(3) \mid q \in \mathcal{Q}_w\}$ such that $\mu_H(E_w) = \alpha^2/24$ for each $w \in T_Q$.

By (iii) of Lemma 3.2 there is an $x' \in P'$ and $T'_Q \subseteq T_Q$ with $\mu_H(T'_Q) = 1/8$ such that $\mu_H((x' + E_w) \cap A) = \alpha\mu_H(E_w) = \alpha^3/24$ for each $w \in T'_Q$.

Fix $p_{x'} \in \mathcal{P}'$. Since $p_{x'}(4) \in S$, we have that $\mu_H((p_{x'}(4) + T'_Q) \cap A) = \alpha/8$. Hence there is a $w \in T'_Q$ such that $p_{x'}(4) + w \in A$. Let $q_w \in \mathcal{Q}_w$. Then

$p_{x'}(4)+q_w(4) = p_{x'}(4)+w \in A$. Notice that $p_{x'}(3)+q_w(3) \in (x+E_w)\cap A \subseteq A$. Notice also that $p_{x'}(1), p_{x'}(2) \in B$ imply $A \cap (p_{x'}(i) + [\lfloor H/8 \rfloor]) = p_{x'}(i) + \tau_0$ for $i = 1, 2$. Hence $p_{x'}(i) + q_w(i) \in p_{x'}(i) + \tau_0 \subseteq A$ for $i = 1, 2$. Therefore, $p_{x'}(i) + q_w(i)$ for $i = 1, 2, 3, 4$ is a nontrivial 4–a.p. in A. □

6 Szemerédi's Theorem for all k

Proposition 6.1 *Let $H \gg 1$, p^* be a k–a.p., $1 \leq i \leq k$, and*

$$C_H = [\lceil kH/(2k+1) \rceil,\ \lfloor (k+1)H/(2k+1) \rfloor]\,. \tag{3}$$

Then for every $t \in p^*(i) + C_H$, there is a collection of k–a.p.'s $\mathcal{P}^i_t$ with $|\mathcal{P}^i_t| = |C_H|$ such that

$$\mathcal{P}^i_t \subseteq \{p \mid p(i) = t \ \text{ and } \ p(j) \in p^*(j) + [H] \ \text{ for all } \ j \in [k]\}.$$

Furthermore, for each $j \leq k$, $j \neq i$, the set $\{p(j) \mid p \in \mathcal{P}^i_t\}$ is a nontrivial $|C_H|$–a.p.

Proof Notice that C_H is at the center of $[H]$ and $\mu_H(C_H) = 1/(2k+1)$.

Assume that $i < k$. For each $t' \in p^*(i+1) + C_H$ let $q_{t'}$ be a k–a.p. such that $q_{t'}(i) = t$ and $q_{t'}(i+1) = t'$. Then $q_{t'} \subseteq [H]$. Assume $i = k$. For each $t' \in C_H$ let $q_{t'}$ be a k–a.p. be such that $q_{t'}(k) = t$ and $q_{t'}(k-1) = t'$. Then $q_{t'} \subseteq [H]$. Since for different $t', t'' \in p^*(i+1) + C_H$ the k–a.p.'s $q_{t'}$ and $q_{t''}$ are different, there are $|C_H|$ such $q_{t'}$'s (including at most one trivial k–a.p. with difference 0). Set $\mathcal{P}^i_t = \{p^* \oplus q_{t'} \mid t' \in p^*(i+1) + C_H\}$. We have that $|\mathcal{P}^i_t| = |C_H|$. Since C_H is an $|C_H|$–a.p. of difference 1, $\{p(j) \mid p \in \mathcal{P}^i_t\}$ is a $|C_H|$–a.p. of difference $|i-j|$. □

Proposition 6.2 *Let $l_0 \in \mathbb{N}$. If $N \gg 1$ and $S \subseteq [N]$ with $\mu_N(S) > 1 - 1/l_0$, then S contains a sequence of at least l_0 consecutive numbers.*

Proof If every sequence of l_0 consecutive numbers contains one term not in S, then $\mu_N(S) \leq 1 - 1/l_0$. □

Definition 6.3 (of $C(k,\Omega)$) *Let $\Omega \subseteq [k]$ and let $C(k,\Omega)$ represents the following statement:*

> *For any $N \gg 1$ and $l_0 \in \mathbb{N}$, there is an $\epsilon > 0$ such that if $S \subseteq [N]$ with $\mu_N(S) > 1-\epsilon$, and a coloring $c : S \to \Theta$ with $|\Theta| \in \mathbb{N}$, there is a $\theta_0 \in \Theta$ and a collection of k–a.p.'s*
>
> $$\mathcal{P}_S = \{p_f \subseteq S \mid f \in [l_0]^\Omega\}$$
>
> *such that*
>
> *(i) $c(p_f(i)) = \theta_0$ for any $f \in [l_0]^\Omega$ and $i < \min\Omega$ (this condition is vacuously true if $\Omega = \emptyset$ or $\min\Omega = 1$),*
>
> *(ii) for each $k_0 \in [k]$ and $g \in [l_0]^{\Omega\cap[k_0]}$, there is a $\theta_g \in \Theta$ such that $c(p_f(k_0)) = \theta_g$ for every $f \in [l_0]^\Omega$ with $f \restriction (\Omega \cap [k_0]) = g$,*
>
> *(iii) for each $k_0 \in \Omega$ and $g \in [l_0]^{\Omega\smallsetminus\{k_0\}}$ we have that*
>
> $$\{p_{f_1}(k_0), p_{f_2}(k_0), \ldots, p_{f_{l_0}}(k_0)\}$$
>
> *is an a.p. of length l_0 where $f_i = g \cup \{(k_0, i)\}$.*

Lemma 6.4 *Let $k \geq 3$. Then $C(k,\Omega)$ is true For any $\Omega \subseteq [k]$.*

Proof Let $k_0 \in [k]$ be such that $[k_0 - 1] \subseteq \Omega$ and $k_0 \notin \Omega$. Let $\Omega' := (\Omega \smallsetminus [k_0 - 1]) \cup \{k_0\}$. It suffices to prove that $C(k,\emptyset)$ is true and $C(k,\Omega)$ implies $C(k,\Omega')$. Notice that $k_0 = 1$ if $1 \notin \Omega$.

Suppose that $\Omega = \emptyset$. For any $N \gg 1$ and $l_0 \in \mathbb{N}$, let $\epsilon = 1/k$. Suppose $S \subseteq [N]$, $\mu_N(S) > 1 - 1/k$, and $c : S \to \Theta$ with $|\Theta| \in \mathbb{N}$. Then S contains k consecutive numbers $p_{\langle\rangle} \subseteq S$. Let $\mathcal{P}_S = \{p_{\langle\rangle}\}$. The collection $\mathcal{P}_S$ of one a.p. satisfies (i) and (iii) of Definition 6.3 vacuously. For each $k_0 \in [k]$ there is only one choice for $g \in [l_0]^{\Omega\smallsetminus\{k_0\}}$, i.e., $g = \langle\rangle = \emptyset$. Let $\theta_g = c(p_{\langle\rangle}(k_0))$. Then (ii) is true trivially. Hence $C(k,\emptyset)$ is true.

Assume that $C(k,\Omega)$ is true. We show that $C(k,\Omega')$ is also true.

Fix $N \gg 1$ and $l_0 \in \mathbb{N}$. Let $\epsilon = \dfrac{1}{l_0 + k(2k+1)^2 l_0^k}$. Suppose $S \subseteq [N]$ with $\mu_N(S) > 1 - \epsilon$ and a coloring $c : S \to \Theta$ with $|\Theta| \in \mathbb{N}$. Let $\gamma := SD(S) \geq \mu_N(S)$. We can find an a.p. P with $|P| \gg 1$ such that $\mu_{|P|}(S \cap P) = \gamma$. Since

Θ is finite, there is a $\theta_0 \in \Theta$ such that $\mu_{|P|}(A) > 0$ where $A = P \cap S \cap c^{-1}(\theta_0)$. Let $\alpha := SD_{S\cap P,\gamma}(A)$. Then $\alpha \geq \mu_{|P|}(A) > 0$. Let $P' \subseteq P$ be an a.p. of difference d with $|P'| = N' \gg 1$ such that $\mu_{N'}(S \cap P') = \gamma$ and $\mu_{N'}(A \cap P') = \alpha$. Let $\varphi : P' \to [N']$ be an affine map, i.e., $\varphi(x) = (x - \min P')/d + 1$ for each $x \in P'$. Let $S' = \varphi(S \cap P')$ and $A' = \varphi(A \cap P')$. Then $SD(S') = \mu_{N'}(S') = \gamma$ and $SD_{S',\gamma}(A') = \mu_{N'}(A') = \alpha$.

For any $l', j, n \in \mathbb{N}$ let $\epsilon' > 0$ be ϵ in the definition of $C(k, \Omega)$ with respect to $l' = l_0$. Let S'' be the following set

$$\left\{ x \in [N'] \mid \delta_n((x + [n]) \cap S') > \gamma - \frac{1}{j} \text{ and } \delta_n((x + [n]) \cap A') > \alpha - \frac{1}{j} \right\}.$$

Notice that $\lim_{n\to\infty} \mu_N(S'') = 1$ for each $j \in \mathbb{N}$ by Lemma 2.3. Hence for all sufficiently large $n \in \mathbb{N}$, we have $\mu_{N'}(S'') > 1 - \epsilon'$. Let $\Theta'_n = \Theta^{[n]}$. Then $|\Theta'_n| \in \mathbb{N}$. Let $c'_n : S'' \to \Theta'_n$ be the coloring defined by

$$c'_n(x)(i) := c(\varphi^{-1}(x + i)) \text{ for each } i \in [n].$$

By $C(k, \Omega)$ we can find a collection $\mathcal{P}_{S''} = \{p_f \subseteq S'' \mid f \in [l']^\Omega\}$ of k–a.p.'s satisfying (i)–(iii) of Definition 6.3 with Θ, c, and l_0 replaced by Θ'_n, c'_n, and l', respectively. By countable saturation we can find an $H_{l',j} \gg 1$ such that $\mathcal{P}_{S''}$ satisfies (i)–(iii) of Definition 6.3 with Θ, c, and l_0 replaced by $\Theta'_{H_{l',j}}$, $c'_{H_{l',j}}$, and l', respectively. Notice that we now have $\mu_{N'}(S'') = 1$ by Lemma 2.3. By countable saturation again, we can find $J_{l'} \gg 1$ such that $\mu_{N'}(S'') = 1$, $H_{J_{l'}} \gg 1$, and $\mathcal{P}_{S''}$ satisfies (i)–(iii) of Definition 6.3 with Θ, c, and l_0 replaced by $\Theta'_{H_{J_{l'}}}$, $c'_{H_{J_{l'}}}$, and l', respectively. Notice that $\mathcal{P}_{S''}$ depends on l'. By countable saturation again we can find $L \gg 1$ such that $\mu_{N'}(S'') = 1$, $H_{J_L} \gg 1$, and $\mathcal{P}_{S''}$ satisfies (i)–(iii) of Definition 6.3 with Θ, c, and l_0 replaced by $\Theta'_{H_{J_L}}$, $c'_{H_{J_L}}$, and L, respectively.

Let $H := H_{J_L}$ and $\mathcal{Q}$ be the collection of all k–a.p.'s in $[H]$. For each $t \in C_H$ where C_H is define in (3) let $\mathcal{Q}^{k_0}_t \subseteq \{q \in \mathcal{Q} \mid q(k_0) = t\}$ be such that $\mu_H(\mathcal{Q}^{k_0}_t) = \mu_H(C_H) = 1/(2k + 1)$. The existence of the set $\mathcal{Q}^{k_0}_t$ is guaranteed by Proposition 6.1. For each $k' \in [k] \smallsetminus [k_0 - 1]$ and $t \in [H]$, let $\mathcal{R}^{k'}_t = \{q \in \mathcal{Q} \mid q(k') = t\}$.

Claim 1 *Assume $k_0 > 1$. For each $i_0 < k_0$, there is a $g_{i_0} \in [L]^{[i_0]}$, $T^{k_0}_{i_0} \subseteq C_H$ with $\mu_{|C_H|}(T^{k_0}_{i_0}) = 1$, $R^{k'}_{i_0} \subseteq [H]$ with $\mu_H(R^{k'}_{i_0}) = 1$ for $k' \geq k_0$ such that if*

$f \in [L]^\Omega$ with $f \restriction (\Omega \cap [i_0]) = g_{i_0}$, $t \in T^{k_0}_{i_0}$, and $t' \in R^{k'}_{i_0}$, there is a

$$\mathcal{Q}^{k_0,i_0}_{t,g} \subseteq \{q \in \mathcal{Q}^{k_0}_t \mid p_f(i) + q(i) \in (p_f(i) + [H]) \cap A' \text{ for all } i \leq i_0\}$$

with $\mu_H(\mathcal{Q}^{k_0,i_0}_{t,g}) = \alpha^{i_0}/(2k+1)$ and

$$\mathcal{R}^{k',i_0}_{t',g} = \{q \in \mathcal{R}^{k'}_{t'} \mid p_f(i) + q(i) \in (p_f(i) + [H]) \cap A' \text{ for all } i \leq i_0\}$$

with $\mu_H(\mathcal{R}^{k',i_0}_{t',g}) \leq \alpha^{i_0}$.

Proof of Claim 1 Notice that for if $f, f' \in [L]^\Omega$ with $f \restriction ([i_0] \cap \Omega) = f' \restriction ([i_0] \cap \Omega) = g$, then $c'(p_f(i)) = c'(p_{f'}(i))$. Hence $p_f(i) + q(i) \in A'$ iff $c(\varphi^{-1}(p_f(i) + q(i))) = \theta_0$ iff $c(\varphi^{-1}(p_{f'}(i) + q(i))) = \theta_0$ iff $p_{f'}(i) + q(i) \in A'$. Hence $\mathcal{Q}^{k_0,i_0}_{t,g}$ and $\mathcal{R}^{k',i_0}_{t',g}$ are independent of the choice of f as long as $f \restriction (\Omega \cap [i_0]) = g$. We now prove the claim by induction on i_0.

Assume $i_0 = 1$. Let $T^{k_0}_1 = C_H$ and $R^{k'}_1 = [H]$ for all $k' > k_0$. Fix an $l \in [L]$ and let $g_1 : \{1\} \to [L]$ be such that $g_1(1) = l$. For any $t \in T^{k_0}_1$ and $t' \in R^{k'}_1$, let $f \in [L]^\Omega$ be such that $f(1) = l$, let

$$\mathcal{Q}^{k_0,1}_{t,g_1} \subseteq \{q \in \mathcal{Q}^{k_0}_t \mid p_f(1) + q(1) \in (p_f(1) + [H]) \cap A'\}$$

be such that $\mu_H(\mathcal{Q}^{k_0,1}_{t,g_1}) = \alpha/(2k+1)$ and

$$\mathcal{R}^{k',1}_{t',g_1} = \{q \in \mathcal{R}^{k'}_{t'} \mid p_f(1) + q(1) \in (p_f(1) + [H]) \cap A'\},$$

with $\mu_H(\mathcal{Q}^{k',1}_{t',g_1}) \leq \alpha$. Notice that the existence of $\mathcal{Q}^{k_0,1}_{t,g_1}$ is guaranteed by Proposition 6.1.

Assume that the claim is true for $i_0 - 1 < k_0 - 1$.

We now prove that the claim is also true for $i_0 < k_0$. For each $t \in T^{k_0}_{i_0-1}$ and $t' \in R^{k'}_{i_0-1} \subseteq [H]$ for $k' \geq k_0$ let

$$E^{k_0}_t := \left\{q(i_0) \mid q \in \mathcal{Q}^{k_0,i_0-1}_{t,g_{i_0-1}}\right\} \text{ and } F^{k'}_{t'} := \left\{q(i_0) \mid q \in \mathcal{R}^{k',i_0-1}_{t',g_{i_0-1}}\right\}.$$

Then $\mu_H(E^{k_0}_t) = \alpha^{i_0-1}/(2k+1)$ and $\mu_H(F^{k'}_{t'}) \leq \alpha^{i_0-1}$. Notice that for any $g \in [L]^{\Omega \setminus \{i_0\}}$ with $g \restriction [i_0 - 1] = g_{i_0-1}$, the list $\{p_{f_1}(i_0), p_{f_2}(i_0), \ldots, p_{f_L}(i_0)\} \subseteq S''$ is an a.p. of length $L \gg 1$, where $f_l = g \cup \{(i_0, l)\}$.

By (iii) of Lemma 3.2 there is an $l \in [L]$, a $T^{k_0}_{i_0} \subseteq T^{k_0}_{i_0-1}$, and a $R^{k'}_{i_0} \subseteq R^{k'}_{i_0-1}$ with $\mu_H(T^{k_0}_{i_0}) = 1/(2k+1)$ and $\mu_H(R^{k'}_{i_0}) = 1$ for all $k' \geq k_0$ such that for each $t \in T^{k_0}_{i_0}$ and $t' \in R^{k'}_{i_0}$ we have

$$\mu_H((p_{f_l}(i_0) + E^{k_0}_t) \cap A') = \alpha \mu_H(E^{k_0}_t) = \alpha^{i_0}/(2k+1) \text{ and}$$

$$\mu_H((p_{f_l}(i_0) + F^{k'}_{t'}) \cap A') = \alpha \mu_H(F^{k'}_{t'}) \leq \alpha^{i_0}.$$

Let $g_{i_0} = g_{i_0-1} \cup \{(i_0, l)\} \in [L]^{[i_0]}$. Let

$$\mathcal{Q}^{k_0,i_0}_{t,g_{i_0}} = \{q \in \mathcal{Q}^{k_0,i_0-1}_{t,g_{i_0-1}} \mid p_{f_l}(i_0) + q(i_0) \in (p_{f_l}(i_0) + E^{k_0}_t) \cap A'\} \text{ and}$$

$$\mathcal{R}^{k',i_0}_{t',g_{i_0}} = \{q \in \mathcal{R}^{k',i_{i_0-1}}_{t',g_{i_0-1}} \mid p_{f_l}(i_0) + q(i_0) \in (p_{f_l}(i_0) + F^{k'}_{t'}) \cap A'\}.$$

Then $\mu_H(\mathcal{Q}^{k_0,i_0}_{t,g_{i_0}}) = \mu_H(E^{i_0}_t) = \alpha^{i_0}/(2k+1)$ and $\mu_H(\mathcal{R}^{k',i_0}_{t',g_{i_0}}) = \mu_H(F^{k'}_{t'}) \leq \alpha^{i_0}$ for every $k' \geq k_0$. Notice that if $f, f' \in [L]^\Omega$ with $f \restriction [i_0] = f' \restriction [i_0]$, then $p_f(i) + q(i) \in A'$ iff $p_{f'}(i) + q(i) \in A'$ for all $i \in [i_0]$. Hence $T^{k_0}_{i_0}$, $R^{k_0}_{i_0}$, $\mathcal{Q}^{k_0,i_0}_{t,g_{i_0}}$, and $\mathcal{R}^{k',i_0}_{t',g_{i_0}}$ for $k' \geq k_0$ are independent of $f \restriction (\Omega \setminus [i_0])$. This completes the proof of the claim.

We continue to prove Lemma 6.3. Keep in mind that k_0 can be 1 below.

By the claim above there exists a function $\overline{g} \in [L]^{[k_0-1]}$, a set $T^{k_0} := T^{k_0}_{k_0-1} \subseteq C_H$ with $\mu_{|C_H|}(T^{k_0}) = 1$, and a set $R^{k'} := R^{k'}_{k_0-1} \subseteq [H]$ with $\mu_H(R^{k'}) = 1$ for every $k' \geq k_0$ such that if $f \in [L]^\Omega$ with $f \restriction [k_0 - 1] = \overline{g}$, $t \in T^{k_0}$, and $t' \in R^{k'}$, there is a

$$\mathcal{Q}^{k_0}_{t,A'} \subseteq \{q \in \mathcal{Q}^{k_0}_t \mid p_f(i) + q(i) \in (p_f(i) + [H]) \cap A' \text{ for all } i < k_0\}$$

with $\mu_H(\mathcal{Q}^{k_0}_{t,A'}) = \alpha^{k_0-1}/(2k+1)$ and

$$\mathcal{R}^{k'}_{t',A'} = \{q \in \mathcal{R}^{k'}_{t'} \mid p_f(i) + q(i) \in (p_f(i) + [H]) \cap A' \text{ for all } i < k_0\}$$

with $\mu_H(\mathcal{R}^{k'}_{t',A'}) \leq \alpha^{k_0-1}$.

Notice that if $k_0 = 1$, simply let $\overline{g} = \emptyset$, $T^1 = C_H$, $R^{k'} = [H]$, $\mathcal{Q}^1_{t,A'} = \mathcal{Q}^1_t$, and $\mathcal{R}^{k'}_{t',A'} = \mathcal{R}^{k'}_t$ for all $k' \geq 1$.

We now trim the collection $\mathcal{P}_{S''}$ to a smaller collection $\mathcal{P}'$ by reducing $L \gg 1$ to $l_0 \in \mathbb{N}$, i.e.,

$$\mathcal{P}' = \{p_f \in \mathcal{P}_{S''} \mid f \restriction [k_0 - 1] = \overline{g} \text{ and } f(k') \in [l_0] \text{ for all } k' \in \Omega \setminus [k_0]\}.$$

Notice that all $p_f \in \mathcal{P}'$ can be indexed by $h \in [l_0]^{\Omega \setminus [k_0]}$, i.e., $p_f \in \mathcal{P}'$ can be rewritten as p_h where $h = f \restriction (\Omega \setminus [k_0])$.

We now want to find enough many $t \in T^{k_0}$ such that $\mathcal{Q}^{k_0}_{t,A'}$ contains some q with $p_h(k') + q(k') \in (p_h(k') + [H]) \cap S'$ for every $k' \geq k_0$. Let $\mathcal{L} := [l_0]^{\Omega \setminus [k_0]}$.

A k–a.p. $q \in \overline{\mathcal{Q}} := \bigcup_{t \in T^{k_0}} \mathcal{Q}^{k_0}_{t,A'}$ is called *good* if

$$p_h(k') + q(k') \in (p_h(k') + [H]) \cap S'$$

for every $p_h \in \mathcal{P}'$ and $k' \geq k_0$. Let $\mathcal{Q}_g$ be the collection of all good k–a.p.'s and $\mathcal{Q}_b := \overline{\mathcal{Q}} \setminus \mathcal{Q}_g$. Let $T_g := \{q(k_0) \in T^{k_0} \mid q \in \mathcal{Q}_g\}$.

We show that $\mu_{|C_H|}(T_g) > 1 - 1/l_0$.

Notice that

$$|\mathcal{Q}_b| \leq \sum_{h \in \mathcal{L}} \sum_{k'=k_0}^{k} \sum_{p_h(k')+t \notin S'} |\mathcal{R}^{k'}_{t,A'}|.$$

$$\begin{aligned} \text{So } \ |\mathcal{Q}_g| &= |\overline{\mathcal{Q}}| - |\mathcal{Q}_b| \\ &\geq \sum_{t \in T^{k_0}} |\mathcal{Q}^{k_0}_{t,A'}| - \sum_{h \in \mathcal{L}} \sum_{k'=k_0}^{k} \sum_{p_h(k')+t \notin S'} |\mathcal{R}^{k'}_{t,A'}|. \end{aligned}$$

Hence we have

$$\begin{aligned} &\mu_{|C_H|}(T_g)\alpha^{k_0-1}/(2k+1) = st\left(\frac{|T_g|}{|C_H|}\frac{1}{|T_g|}\sum_{t \in T_g}\frac{1}{H}|\mathcal{Q}^{k_0}_{t,A'}|\right) \\ &\geq st\left(\frac{1}{H|C_H|}|\mathcal{Q}_g|\right) = st\left(\frac{1}{H|C_H|}(|\overline{\mathcal{Q}}| - |\mathcal{Q}_b|)\right) \\ &\geq st\left(\frac{1}{H|C_H|}\sum_{t \in T^{k_0}}|\mathcal{Q}^{k_0}_{t,A'}| - \frac{2k+1}{H^2}\sum_{h \in \mathcal{L}}\sum_{k'=k_0}^{k}\sum_{p_h(k')+t \notin S'}|\mathcal{R}^{k'}_{t,A'}|\right) \\ &\geq \mu_{|C_H|}(T^{k_0})\alpha^{k_0-1}/(2k+1) - (2k+1)k|\mathcal{L}|\epsilon\alpha^{k_0-1} \\ &= \left(1 - \epsilon k(2k+1)^2 l_0^{k-1}\right)\alpha^{k_0-1}/(2k+1), \end{aligned}$$

which implies $\mu_{|C_H|}(T_g) \geq 1 - \epsilon k(2k+1)^2 l_0^{k-1} > 1 - 1/l_0$. We remind the reader that ϵ is defined in the beginning of the proof and $\mu_{N'}(S') = \gamma > 1 - \epsilon$. Hence T_g contains an a.p. r of length l_0.

For each $l \in [l_0]$ choose a $q_l \in \mathcal{Q}_g$ such that $q_l(k_0) = r(l)$. For each $h \in [l_0]^{\Omega \setminus [k_0]}$ and $l \in [l_0]$ let $h_l = h \cup \{(k_0, l)\}$. Notice that

$$[l_0]^{\Omega'} = \{h_l \mid h \in [l_0]^{\Omega \setminus [k_0]} \text{ and } l \in [l_0]\}.$$

For each $h_l \in [l_0]^{\Omega'}$ define

$$\overline{p}_{h_l} := \varphi^{-1}(p_h \oplus q_l) = \{\varphi^{-1}(p_h(i) + q_l(i)) \mid i \in [k]\}.$$

Let

$$\overline{\mathcal{P}}_S = \left\{\overline{p}_{h_l} \mid h_l \in [l_0]^{\Omega'}\right\}.$$

We now verify that $\overline{\mathcal{P}}_S$ serving as $\mathcal{P}_S$ satisfies (i)–(iii) of Definition 6.3 for $C(k, \Omega')$. Let $c' := c_{H_{J_L}}$ be the coloring from S'' to $\Theta^{[H]}$. Since $p_h(i) + q_l(i)$ is in either S' or $A' \subseteq S'$ for every $i \in [k]$, we have that $\overline{p}_{h_l} \subseteq \varphi^{-1}(S') = S$.

If $i < \min \Omega' = k_0$, then $p_h(i) + q_l(i) \in A'$ because $q_l \in \mathcal{Q}^{k_0}_{t,A'}$. Hence $\overline{p}_{h_l}(i) \in A$, which implies $c(\overline{p}_h(i)) = \theta_0$ for every $h_l \in [l_0]^{\Omega'}$. So (i) is verified.

Let $h_l, h'_{l'} \in [l_0]^{\Omega'}$ and $k_1 \in [k]$ such that $h_l \restriction ([k_1] \cap \Omega') = h'_{l'} \restriction ([k_1] \cap \Omega')$. If $k_1 < k_0$, then $c(\overline{p}_{h_l}(k_1)) = c(\overline{p}_{h'_{l'}}(k_1)) = \theta_0$ by (i). If $k_1 \geq k_0$, then $l = l'$. Hence $\overline{p}_{h_l}(k_1) = \varphi^{-1}(p_h(k_1) + q_l(k_1))$ and $\overline{p}_{h'_l}(k_1) = \varphi^{-1}(p_{h'}(k_1) + q_l(k_1))$. Since $c'(p_h(k_1)) = c'(p_{h'}(k_1))$ by (ii) of Definition 6.3 for $\mathcal{P}_{S''}$ and the fact that $f = h \cup g$ and $f' = h' \cup g$ coincide on $\Omega \cap [k_1]$, we have $c(\overline{p}_{h_l}(k_0)) = c(\overline{p}_{h'_l}(k_0))$. So (ii) is verified.

Given $k_1 \in \Omega'$, for each $h_l \in [l_0]^{\Omega'}$ let $h^- = h \restriction (\Omega \setminus ([k_0] \cup \{k_1\}))$ and $(h^-)_{l'} = h^- \cup \{(k_1, l')\}$. Also let $h_l^- = h_l \restriction (\Omega' \setminus \{k_1\})$ and $(h_l^-)_{l'} = h_l^- \cup \{(k_1, l')\}$. If $k_1 = k_0$, then $(h_l^-)_{l'}(k_0) = r(l')$. Hence

$$\{\overline{p}_{(h_l^-)_{l'}}(k_0) \mid l' \in [l_0]\} = \{\varphi^{-1}(p_h(k_0) + r(l')) \mid l' \in [l_0]\}$$

is an a.p. of length l_0. If $k_1 \in \Omega \setminus [k_0]$, then $\{p_{(h^-)_{l'}}(k_1) \mid l' \in [l_0]\}$ is an a.p. of length l_0 by (iii) of Definition 6.3 for $\mathcal{P}_{S''}$. Hence

$$\{\overline{p}_{(h_l^-)_{l'}}(k_1) \mid l' \in [l_0]\} = \{\varphi^{-1}(p_{(h^-)_{l'}}(k_1) + q_l(k_1)) \mid l' \in [l_0]\}$$

is an a.p. of length l_0. Hence (iii) is verified. This completes the proof of the lemma. □

Theorem 6.5 (E. Szemerédi, 1975) *Let $k \in \mathbb{N}$. If $U \subseteq \mathbb{N}$ and $SD(U) > 0$, then U contains nontrivial k-term arithmetic progressions.*

Proof Let $N \gg 1$ and $A = {}^*U \cap [N]$ be such that $SD(A) = \mu_N(A) = \alpha > 0$. By Lemma 6.4 we can assume that $C(k+1, \{k+1\})$ is true after 2^k steps of the induction. Let $l_0 = 1$. There is an $\epsilon > 0$. Let $n \in \mathbb{N}$ be sufficiently large such that $\mu_N(S) > 1 - \epsilon$ where

$$S = \{x \in [N] \mid \mu_n((x + [n]) \cap A) > \alpha/2\}.$$

Let $\Theta = \{\theta \subseteq [n] \mid \mu_n(\theta) > \alpha/2\}$. Let $c : S \to \Theta$ be such that $c(x) = ((x + [n]) \cap A) - x$. Then there exists a collection of $k+1$–a.p.'s $\mathcal{P}_S = \{p_f \mid f \in \{1\}^{\{k+1\}}\} = \{p\}$ and $\theta_0 \in \Theta$ satisfying (i)–(iii) in Definition 6.3. In particular, we have $c(p(i)) = \theta_0$ for all $i = 1, 2, \ldots, k$ for some $\theta_0 \in \Theta$. Since $\mu_n(\theta_0) > \alpha/2$, we can find $a \in \theta_0$. Notice that $p(i) + a \in p(i) + \theta_0 = (p(i) + [n]) \cap A$. Hence $\{p(1) + a, p(2) + a, \ldots, p(k) + a\}$ is a k–a.p. in A. □

Remark 6.6 *(i) One can also find a k-a.p. in A by assuming $C(k, [k-2])$ together with applying the argument in the proof of the claim one more time.*

(ii) We tried but failed to generalize the approach in §5 to give a proof of Theorem 6.5 with a simple induction instead of a Tower of Hanoi type induction in Lemma 6.4.

References

[1] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression.* Collection of articles in memory of Juriĭ Vladimirovič Linnik, Acta Arith. 27 (1975), 199 – 245.

[2] A. Frieze and K. Kannan, *The regularity lemma and approximation schemes for dense problems*, 37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996), 12–20, IEEE Comput. Soc. Press, Los Alamitos, CA, 1996.

[3] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. 31 (1977), 204 – 256.

[4] T. Gowers, *A new proof of Szemerédi's theorem*, GAFA 11 (2001), 465 – 588.

[5] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 245 – 252.

[6] T. Tao, *Szemerédi's proof of Szemerédi's Theorem*, https://terrytao.files.wordpress.com/2017/09/szemeredi-proof1.pdf