Sets of integers that do not contain long arithmetic progressions

Kevin O'Bryant*

Department of Mathematics
The City University of New York
College of Staten Island and The Graduate Center

kevin@member.ams.org

Submitted: Feb 1, 2011; Accepted: Mar 2, 2010; Published: Mar 11, 2011
Mathematics Subject Classification: 11B25

Abstract

Combining ideas of Rankin, Elkin, Green & Wolf, we give constructive lower bounds for $r_k(N)$, the largest size of a subset of $\{1,2,\dots,N\}$ that does not contain a k-element arithmetic progression: For every $\epsilon>0$, if N is sufficiently large, then

$$r_3(N) \geq N\left(\frac{6\cdot 2^{3/4}\sqrt{5}}{e\,\pi^{3/2}}-\epsilon\right)\exp\Big(-\sqrt{8\log N}+\tfrac{1}{4}\log\log N\Big),$$

$$r_k(N) \geq N\,C_k\,\exp\Big(-n2^{(n-1)/2}\sqrt[n]{\log N}+\tfrac{1}{2n}\log\log N\Big),$$

where $C_k>0$ is an unspecified constant, $\log=\log_2$, $\exp(x)=2^x$, and $n=\lceil \log k\rceil$. These are currently the best lower bounds for all k, and are an improvement over previous lower bounds for all $k\neq 4$.

We denote by $r_k(N)$ the maximum possible size of a subset of $\{1,2,\dots,N\}$ that does not contain k numbers in arithmetic progression. Behrend [1] proved that

$$\frac{r_3(N)}{N} \geq C\exp\Big(-(1+\epsilon)\sqrt{8\log N}\Big),$$

where exp and log are the base-2 exponential and logarithm and each occurrence of C is a new positive constant. Sixty years later, Elkin [2] introduced a new idea to Behrend's work and showed that there are arbitrarily large N satisfying

$$\frac{r_3(N)}{N} \geq C\exp\Big(-\sqrt{8\log N}+\tfrac{1}{4}\log\log N\Big),$$

*This work was supported by a grant from The City University of New York PSC-CUNY Research Award Program.

and shortly afterwards Green & Wolf [6] arrived at the same bound by a different method. For $k \geq 1 + 2^{n-1}$, Rankin [10] proved that for each $\epsilon > 0$, if N is sufficiently large then

$$\frac{r_k(N)}{N} \geq C \, \exp\Big(-n \, 2^{(n-1)/2} \, (1+\epsilon) \, \sqrt[n]{\log N}\Big) \, ,$$

where $n = \lceil \log k \rceil$. For $k = 3$, Rankin's construction is the same as that of Behrend. This was subsequently rediscovered in a simpler, but less precise, form by Łaba & Lacey [8]. Together with the obvious $r_k(N) \leq r_{k+1}(N), r_k(N+M) \leq r_k(N) + r_k(M)$, these were the thickest known constructions. The primary interest in the current work is the following corollary of our main theorem.

Corollary 1. *Fix k, and set $n = \lceil \log k \rceil$. There exists a positive constant C such that for all $N \geq 1$*

$$\frac{r_k(N)}{N} \geq C \, \exp\Big(-n 2^{(n-1)/2} \, \sqrt[n]{\log N} + \tfrac{1}{2n} \log\log N\Big)$$

For every $\epsilon > 0$, if N is sufficiently large then

$$\frac{r_3(N)}{N} \geq \left(\frac{6 \cdot 2^{3/4} \sqrt{5}}{e \, \pi^{3/2}} - \epsilon\right) \exp\Big(-\sqrt{8 \log N} + \tfrac{1}{4} \log\log N\Big) \, .$$

The constant in the r_3 bound is around 1.5, so we have the pleasant-to-write consequence:

$$r_3(N) \geq N \, \exp\Big(-\sqrt{8 \log N} + \tfrac{1}{4} \log\log N\Big) \geq N \, 2^{-\sqrt{8 \log N}}$$

for sufficiently large N.

Szemerédi's Theorem states that $r_k(N) = o(N)$, and the task of getting better upper bounds on $r_k(N)$ has been mathematically fruitful. The currently-best upper bounds on $r_k(N)$ (for sufficiently large N) are due to Sanders [11], Green & Tao [5], and Gowers [4], respectively, and are shown here with our lower bounds:

$$\begin{aligned}
-\sqrt{8 \log N} + \frac{1}{4} \log\log N + 1 &\leq \log\left(\tfrac{r_3(N)}{N}\right) \leq -\log\log N + 5 \log\log\log N + C; \\
-\sqrt{8 \log N} + \frac{1}{4} \log\log N + 1 &\leq \log\left(\tfrac{r_4(N)}{N}\right) \leq -C \sqrt{\log\log N}; \\
-n 2^{(n-1)/2} \sqrt[n]{\log N} + \tfrac{1}{2n} \log\log N + C &\leq \log\left(\tfrac{r_k(N)}{N}\right) \leq -2^{-2^{k+9}} \log\log\log N.
\end{aligned}$$

It is naturally tempting to speculate as to whether the upper or lower bound on $r_k(N)$ is closer to the truth. Certainly, the upper bounds have seen a steady stream of substantive improvements, while the main term of the lower bound has remained unchanged for more than a half century. The reader is directed to a discussion on Gil Kalai's blog [7] for some relevant speculative remarks of Gowers and of Kalai's.

To prove our result we need to induct through sets that do not contain more elaborate types of progressions. A k-term D-progression is a sequence of the form

$$Q(1), Q(2), \ldots, Q(k)$$

where Q is a nonconstant polynomial with degree at most D. For example, 1-progressions are proper arithmetic progressions. The sequences $2, 1, 2, 5, 10$ and $1, 2, 4, 7, 11$ are 5-term 2-progressions arising from the polynomials $(j-2)^2+1$ and $(j^2-j+2)/2$. In particular, a progression of integers may contain the same number in different places, and may arise from a polynomial whose coefficients are not integers. Also, note that the class of k-term D-progressions is invariant under both translation and dilation. Let $r_{k,D}(N)$ denote the maximum possible size of a subset of $[1,N] \cap \mathbb{Z}$ that does not contain any k-term D-progressions.

Theorem 1. *Fix positive integers k, D and set $n = \lceil \log(k/D) \rceil$. There exists a positive constant C such that for every N*

$$\frac{r_{k,D}(N)}{N} \geq C \exp\Big(-n 2^{(n-1)/2} D^{(n-1)/n} \sqrt[n]{\log N} + \tfrac{1}{2n} \log\log N\Big).$$

To explain what is new and interesting in the current work, we begin by summarizing the earlier constructions. Behrend's construction [1], while no longer the numerically best or most general, remains the most elegant. His initial observation is that a sphere cannot contain a 3-term arithmetic progression simply because a line and a sphere cannot intersect more than twice. Let S be a set of points in $\mathbb{Z}^d$ all lying on one sphere and having all coordinates positive and smaller than P, and then let A be the image of S under the map $\varphi : \langle x_1, \ldots, x_d \rangle \mapsto \sum_{i=1}^d x_i (2P)^{i-1}$. Because $0 < x_i < P$, addition of two elements of A will not involve any carrying. This φ is therefore a Freiman 2-isomorphism between S and A; that is, $\overline{x}_1 + \overline{x}_2 = \overline{x}_3 + \overline{x}_4$ if and only if $\varphi(\overline{x}_1) + \varphi(\overline{x}_2) = \varphi(\overline{x}_3) + \varphi(\overline{x}_4)$. Since three integers $a < b < c$ are in arithmetic progression if and only if $a + c = b + b$, this proves that A is free of 3-term arithmetic progressions. The only remaining work is to show that there exists a suitably large S, which Behrend did with the pigeonhole principle, and to optimize P and d in terms of N.

Rankin combined three observations. His first observation was that Behrend's use of the pigeonhole principle could be replaced with a number-theoretic result on the number of representations of a huge number as a sum of a large number of squares. The second is that a degree D polynomial cannot intersect a sphere in more than $2D$ points, and so Behrend's argument actually gives a lower bound on $r_{2D+1,D}$. The third is that one can use a set that does not contain k-term $2D$-progressions to build S as a union of concentric spheres with skillfully chosen radii. The corresponding set A (after mapping S as per Behrend, but with the radix $2P$ replaced by something much larger) will necessarily be free of k-term D-progressions. This provided for an inductive bound. For example, $r_9 = r_{9,1}$ is bounded in terms of $r_{9,2}$, which is bounded in terms of $r_{9,4}$, which is then bounded using Rankin's generalized Behrend argument.

Elkin [2] improved Behrend's 3-term construction in two ways. First, he used the central limit theorem (and the pigeonhole principle) to guarantee the existence of a large S; and second, he considered lattice points in a very thin annulus. Using an annulus instead of a sphere leads to a set S that is substantively larger but, unfortunately, does have 3-term arithmetic progressions. After removing a small number of points to eliminate the progressions, Elkin proceeded along the same line as Behrend, needing to optimize d, P, and also the thickness of the annulus.

Green & Wolf [6] give an argument that spiritually similar to Elkin's, but avoids counting lattice points. In the d-dimensional torus, they take S to be the intersection of a small box and an annulus. Using random elements $\overline{\omega}, \overline{\alpha}$ of the torus, they consider the map $\varphi : n \mapsto n\overline{\omega} + \overline{\alpha}$. Letting $A := \{a \colon \varphi(a) \in S\}$, this map is a Freiman 2-isomorphism between A and $\varphi(A)$. The randomness allowed them to easily count the size of A and the number of progressions in A that need to be removed.

In the current work we recast Rankin's ideas using the lessons of Elkin and Green & Wolf. We avoid Rankin's sum-of-squares number theory lemma by taking random $\overline{\omega}, \overline{\alpha}$ (unfortunately, we still need the pigeonhole principle). We find the right generalization of "an arithmetic progression in a thin annulus has a small difference" to D-progressions, and thereby generalize Elkin's result to improve Rankin's bound on $r_{2D+1,D}$. Finally, by taking concentric annuli, we smooth out Rankin's inductive step. We note also that previous work has sometimes suffered[1] from a cavalier treatment of error terms. For example, Elkin's "arbitrarily large N" and Rankin's "$1+\epsilon$" term can be eliminated with a little care. We have taken the opposite tack here, in places working for coefficients that are not important in the final analysis, but which we consider to be of interest. In particular, the refinement for r_3 stated in Corollary 1 constitutes about 15% (by volume) of this work.

1 Notation

Throughout, log and exp refer to the base-2 logarithm and exponential. Vectors are all given overlines, as in $\overline{x}$, and all have dimension d.

The parameters N and d tend to infinity together, with N much larger than d, and all little-oh notation is with respect to N and d. The parameter d is a dimension, and must be an integer, while N need not be an integer. The other fundamental parameters, the integers k and D, are held constant.

We define the difference operator Δ to be the map taking a finite sequence $(a_i)_{i=1}^k$ to the finite sequence $(a_{v+1} - a_v)_{v=1}^{k-1}$. The formula for repeated differencing is then

$$\Delta^n(a_i) = \left(\sum_{i=0}^{n} \binom{n}{i} (-1)^i a_{i+v} \right)_{v=1}^{k-n}.$$

We note that a nonconstant sequence (a_i) with at least $D+1$ terms is a D-progression if and only if $\Delta^{D+1}(a_i)$ is a sequence of zeros. If $a_i = p(i)$, with p a polynomial with degree D and lead term p_D, then $\Delta^D(a_i) = (D! p_D)$, a constant sequence. Note also that Δ is a linear operator. Finally, we will make repeated use of the fact, provable by induction for $1 \le n \le k$, that

$$|\Delta^n(a_i)| \le 2^{n-1} \left(\max_i a_i - \min_i a_i \right).$$

[1]Some would say benefitted.

A k-term type-(n,a,b) progression is a nonconstant sequence $a_1, a_2, \ldots, a_k$ with $k \geq n$, $a_1 = a$, and n-th differences $\Delta^n(a_i)$ the constant nonzero sequence (b). For example, if p is a degree n polynomial (with lead term $p_n \neq 0$) and $k \geq n$, then $p(1), \ldots, p(k)$ is a type $(n, p(1), n!p_n)$ progression.

The open interval $(a-b, a+b)$ of real numbers is denoted $a \pm b$. The interval $[1,N] \cap \mathbb{Z}$ of natural numbers is denoted $[N]$. For positive integers i, the box $(0 \pm 2^{-i-1})^d$, which has Lebesgue measure 2^{-id}, is denoted $\textsc{Box}_i$. We define $\textsc{Box}_0 = [-1/2, 1/2)^d$, and define $\overline{x} \bmod \overline{1}$ to be the unique element $\overline{y}$ of $\textsc{Box}_0$ with $\overline{x} - \overline{y} \in \mathbb{Z}^d$.

A point $\overline{x} = \langle X_1, \ldots, X_d \rangle$ chosen uniformly from $\textsc{Box}_D$ has components X_i independent and uniformly distributed in $(-2^{-D-1}, 2^{-D-1})$. Therefore, $\|\overline{x}\|_2^2 = \sum_{i=1}^d X_i^2$ is the sum of d iidrvs, and is therefore normally distributed as $d \to \infty$. Further $\|\overline{x}\|_2^2$ has mean $\mu_D := 2^{-2D} d/12$ and variance $\sigma_D^2 := 2^{-4D} d/180$.

For any set $A \subseteq [n]$, positive integer D, and sufficiently small positive real number δ, we define $\textsc{Annuli}(A, n, D, \delta)$ in the following manner:

$$\textsc{Annuli}(A,n,D,\delta) := \left\{ \overline{x} \in \textsc{Box}_D \colon \frac{\|\overline{x}\|_2^2 - \mu_D}{\sigma_D} \in \bigcup_{a \in A} \left(z - \frac{a-1}{n} \pm \delta \right) \right\},$$

where $z \in \mu_D \pm \sigma_D$ is chosen to maximize the volume of $\textsc{Annuli}(A,n,D,\delta)$. Geometrically, $\textsc{Annuli}(A,n,D,\delta)$ is the union of $|A|$ spherical shells, intersected with $\textsc{Box}_D$.

2 Lemmas

The following lemma is best-possible for $k = 2D+1$. Improving the bound for larger k comes down to the following problem: if Q has degree D and all of $|Q(1)|, \ldots, |Q(k)|$ are less than 1, then how big can the leading coefficient of Q be? This lemma plays the role that "a line intersects a sphere in at most two points" played for Behrend, and that "a degree D polynomial curve intersects a sphere in at most $2D$ points" played for Rankin. Green & Wolf handle the $D = 1$ case by a simple geometric argument.

Lemma 1 (Sphere-ish polynomials have small-ish lead coefficients)**.** *Let δ, r be real numbers with $0 \leq \delta \leq r$, and let k, D be integers with $D \geq 1, k \geq 2D+1$. If $\overline{P}(j)$ is a polynomial with degree D, and $r - \delta \leq \|\overline{P}(j)\|_2^2 \leq r + \delta$ for $j \in [k]$, then the lead coefficient of $\overline{P}$ has norm at most $2^D\ (2D)!^{-1/2}\ \sqrt{\delta}$.*

Proof. In this paragraph we summarize the proof; in subsequent paragraphs we provide the details. $Q(j) := \|\overline{P}(j)\|_2^2 - r$ is a degree $2D$ polynomial of j, and each of the $2D+1$ real numbers $Q(1), \ldots, Q(2D+1)$ are close to zero. If they were all exactly zero, then Q would have more zeros than its degree and so would necessarily be identically zero. Just having that many values close to 0, however, is already enough to guarantee that the lead coefficient of Q is small.

Let $\overline{P}(j) = \overline{P}_0 + \overline{P}_1 j + \cdots + \overline{P}_D j^D$. We work with the degree $2D$ polynomial

$$Q(j) := \|\overline{P}(j)\|_2^2 - r = \sum_{n=0}^{2D} q_n j^n,$$

and note in particular that $q_{2D} = \|\overline{P}_D\|_2^2$. As $0 \le \delta \le r$, we conclude that $|Q(j)| \le \delta$.

Set $\overline{q}, \overline{Q}$ to be the column vectors $\langle q_0, q_1, \ldots, q_{2D}\rangle^T, \langle Q(1), \ldots, Q(2D+1)\rangle^T$, respectively. Let M be the $(2D+1) \times (2D+1)$ matrix whose (i,j)-component is i^{j-1}. We have the system of equations

$$M\,\overline{q} = \overline{Q},$$

which is nonsingular because M is a Vandermonde matrix. By Cramer's rule, the cofactor expansion of a determinant along the last column, and the triangle inequality,

$$q_{2D} = \frac{\det(M')}{\det(M)} = \frac{1}{\det(M)} \sum_{j=1}^{2D+1} Q(j)(-1)^{j+1} M_{j,2D+1} \le \frac{\sum_{j=1}^{2D+1} |M_{j,2D+1}|}{|\det(M)|}\delta.$$

By the formula for the determinant of a Vandermonde matrix (the relevant minors of M are also Vandermonde matrices), we find that

$$\|\overline{P}_D\|_2^2 = q_{2D} \le \frac{\sum_{j=1}^{2D+1} |M_{j,2D+1}|}{|\det(M)|}\,\delta = \frac{2^{2D}}{(2D)!}\,\delta,$$

completing the proof. □

The next lemma plays the role that "integers whose base-$2b+1$ expansions only use the digits $\{0, 1, \ldots, b\}$ can be added without carrying" played for Behrend, and that "a polynomial with degree D can be evaluated at integers whose base-$B(b)$ expansions only use the digits $\{0, 1, \ldots, b\}$ without carrying, provided that $B(b)$ is sufficiently large" played for Rankin. The $D = 1$ case is directly in the work of Green & Wolf.

Lemma 2 (Tight modular progressions are also non-modular progressions)**.** *Suppose that $p(j)$ is a polynomial with degree D, with D-th coefficient p_D, and set $\overline{x}_j := \overline{\omega}\, p(j) + \overline{\alpha} \bmod \overline{1}$. If $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k$ are in Box_D and $k \ge D+2$, then there is a vector polynomial $\overline{P}(j) = \sum_{i=0}^{D} \overline{P}_i j^i$ with $\overline{P}(j) = \overline{x}_j$ for $j \in [k]$, and $D!\overline{P}_D = \overline{\omega}\, D! p_D \bmod \overline{1}$.*

Proof. Since p has degree D, the $(D+1)$-th differences of $p(1), p(2), \ldots, p(k)$ are zero, and therefore the $(D+1)$-th differences of $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k$ are $\overline{0}$ modulo $\overline{1}$, i.e., all of their components are integers. We will show that in fact all of their components are strictly between -1 and 1, and so they must all be 0.

The $(D+1)$-th differences are given by (valid only for $1 \le v \le k - D - 1$)

$$\Delta^{D+1}(\overline{x}_i)(v) = \sum_{i=0}^{D+1} \binom{D+1}{i} (-1)^i \overline{x}_{v+i}.$$

Denote the i-th component of $\overline{x}_j$ by $x_j^{(i)}$. As $\overline{x}_{v+i} \in \mathrm{Box}_D$, each component of $\overline{x}_{v+i}$ is in $\left(-2^{-D-1}, 2^{-D-1}\right)$. Thus, the h-th component of $\Delta^{D+1}(\overline{x}_i)(v)$ satisfies

$$\left|\sum_{i=0}^{D+1} \binom{D+1}{i} (-1)^i \overline{x}_{v+i}^{(h)}\right| \le \sum_{i=0}^{D+1} \binom{D+1}{i} |\overline{x}_{v+i}^{(h)}| < \sum_{i=0}^{D+1} \binom{D+1}{i} 2^{-(D+1)} = 1,$$

and therefore $\Delta^{D+1}(\overline{x}_i) = (0)$.

Now,

$$D!\overline{P}_D = \Delta^D(\overline{P}(i)) = \Delta^D(\overline{x}_i) \equiv \overline{\omega} D! p_D \pmod{\overline{1}}.$$

As $\overline{P}(i) \in \textsc{Box}_D$ for $1 \le i \le k$, the above binomial-coefficient triangle-inequality argument tells us that the components of $\Delta^D(\overline{P}(i))$ are between $-1/2$ and $1/2$, and so $D!\overline{P}_D = \big(\overline{\omega} D! p_D \bmod \overline{1}\big)$. □

Behrend and Rankin needed to find spheres that contain many lattice points, which is a fundamentally number theoretic issue that they handled with the pigeonhole principle and the circle method, respectively. Here, as in Green & Wolf, we don't need lattice points on spheres but to put a dynamical system in an annulus frequently; this is merely a measure theoretic/geometric issue.

Lemma 3 (Annuli has large volume)**.** *If d is sufficiently large, $A \subseteq [n]$, and $2\delta \le 1/n$, then the volume of* $\textsc{Annuli}(A, n, D, \delta)$ *is at least* $\frac{2}{5} 2^{-dD}|A|\delta$. *Provided that $\delta \log d \to 0$, the volume of* $\textsc{Annuli}(\{1\}, 1, D, \delta)$ *is at least* $(\sqrt{2/\pi} - o(1))\, 2^{-dD}\,\delta$.

Proof. A uniformly chosen element $\overline{x} = \langle X_1, \ldots, X_d \rangle$ of $\textsc{Box}_D$ has the X_i independent and each uniformly distributed in $(-2^{-D-1}, 2^{-D-1})$. Thus $\|\overline{x}\|_2^2$ is the sum of d iidrvs and has mean $\mu_D := 2^{-2D}d/12$ and variance $\sigma_D^2 := 2^{-4D}d/180$. By the central limit theorem (CLT), the random variable

$$\frac{\|\overline{x}\|_2^2 - \mu_D}{\sigma_D}$$

has a normal distribution, as $d \to \infty$, with mean 0 and variance 1. We would like to argue that

$$\begin{aligned}\mathbf{vol}\,\textsc{Annuli}(\{1\}, 1, D, \delta) \ge 2^{-dD}\left(\int_{-\delta}^{\delta} \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx\right) &\ge 2^{-dD}\left(2\delta\frac{e^{-\delta^2/2}}{\sqrt{2\pi}}\right)\\ &= 2^{-dD}\delta\left(\sqrt{2/\pi} - o(1)\right),\end{aligned}$$

but we cannot apply the CLT to an interval that is shrinking as rapidly as $\pm\delta$. We get around this by applying the CLT to an interval that shrinks very slowly, and then using an analytic form of the pigeonhole principle to guarantee an appropriately short subinterval with the needed density.

We could accomplish this using only the classical CLT, but it is expeditious to use the quantitative CLT known as the Berry-Esseen theorem [3, Section XVI.5], which is applicable since

$$\rho_D := \mathbb{E}\left[|X_i^2 - 2^{-2D}/12|^3\right] = 2^{-6D}(3 + 2\sqrt{3})/11340 < \infty.$$

Let I be an interval whose endpoints depend on d. The Berry-Esseen theorem implies that

$$\mathbb{P}\left[\frac{\|\overline{x}\|_2^2 - \mu}{\sigma_D} \in I\right] \ge \frac{1}{\sqrt{2\pi}} \int_I \exp(-x^2/2)\, dx - 2\,\frac{\rho_D}{(\sigma_D/\sqrt{d})^3\sqrt{d}}.$$

First we handle the case $A = \{1\}, n = 1$. We have

$$\begin{aligned}
\mathbb{P}\left[\frac{\|\overline{x}\|_2^2 - \mu_D}{\sigma_D} \in \pm\frac{1}{\log d}\right] &\geq \frac{1}{\sqrt{2\pi}} \int_{-1/\log d}^{1/\log d} \exp(-x^2/2)\, dx - 2\, \frac{\rho_D}{(\sigma_D/\sqrt{d})^3\sqrt{d}} \\
&\geq \frac{1}{\sqrt{2\pi}}\, \frac{2}{\log d} \exp\big(-(1/\log d)^2/2\big) - \frac{3}{\sqrt{d}} \\
&\geq \frac{\sqrt{2/\pi}}{\log d}\left(1 - \tfrac{1}{2}(\log d)^{-4} - 3(\log d)d^{-1/2}\right) \\
&\geq \frac{\sqrt{2/\pi}}{\log d}\left(1 - (\log d)^{-4}\right).
\end{aligned}$$

Let f be the density function of $\frac{\|\overline{x}\|_2^2 - \mu_D}{\sigma_D}$, and let χ_I be the indicator function of I. Since the convolution

$$(f\chi_{\pm 1/\log d}) * \chi_{\pm\delta}$$

is supported on $\pm(1/\log d + \delta)$ and has 1-norm

$$\|f\chi_{\pm 1/\log d})\|_1\, \|\chi_{\pm\delta}\|_1 \geq \left(\frac{\sqrt{2/\pi}}{\log d}\left(1 - (\log d)^{-4}\right)\right) 2\delta,$$

there must be some z with

$$\begin{aligned}
\big((f\chi_{\pm 1/\log d}) * \chi_{\pm\delta}\big)(z) &\geq \frac{\left(\frac{\sqrt{2/\pi}}{\log d}\left(1 - (\log d)^{-4}\right)\right) 2\delta}{2/\log d + 2\delta} \\
&= \delta\sqrt{\frac{2}{\pi}}\left(\frac{1 - (\log d)^{-4}}{1 + \delta\log d}\right) \\
&= \left(\sqrt{2/\pi} - o(1)\right)\delta.
\end{aligned}$$

Consequently, $\mathbf{vol}\,\textsc{Annuli}(\{1\}, 1, D, \delta) \geq \left(\sqrt{\frac{2}{\pi}} - o(1)\right) 2^{-dD}\,\delta$.

Similar calisthenics make the following heuristic argument rigorous. Let G be a normal rv with mean 0 and variance 1:

$$\begin{aligned}
\mathbf{vol}(\textsc{Annuli}(A, n, D, \delta)) &\to 2^{-dD}\, \mathbb{P}_{\overline{\omega},\overline{\alpha}}\left[G \in \bigcup_{a\in A}\left(-\frac{a-1}{n} \pm \delta\right)\right] \\
&\geq 2^{-dD}\, \mathbb{P}_{\overline{\omega},\overline{\alpha}}\left[G \in \big(-1, -1 + 2\delta|A|\big)\right] \\
&= 2^{-dD} \frac{1}{\sqrt{2\pi}} \int_{-1}^{-1+2\delta|A|} \exp(-x^2/2)\, dx \\
&\geq 2^{-dD} \frac{1}{\sqrt{2\pi}} \exp(-1/2) 2\delta|A| \\
&> \frac{2}{5}\, 2^{-dD}|A|\delta,
\end{aligned}$$

where we have used $2\delta \le 1/n$ to force the intervals $-(a-1)/n \pm \delta$ to be disjoint, and also to force $-1 + 2\delta|A| < 0$. Since the final inequality is strict, we can replace the limit in the central limit theorem with a "sufficiently large d" hypothesis. $\square$

We comment that the use of the pigeonhole principle and the CLT in the previous lemma could be removed. The distribution of X_i^2 is explicit, and so we could, at least in principle, work out an explicit form for the density of $\|\overline{x}\|^2$ (similar in spirit to the Irwin-Hall distribution). This would also likely allow one to take $z = 0$.

The $k = 3, D = 1$ case of the following lemma is in Green & Wolf.

Lemma 4 is not best possible. However, the factor 2^{D+1} will turn out to be irrelevant in the final analysis.

Lemma 4 (There are not many types of progressions)**.** *Assume $k \ge D$. There are fewer than $2^{D+1}N^2$ types of k-term progressions with degree at most D contained in $[N]$.*

Proof. We suppose that we have a k-term progression $a_1, \dots, a_k$ contained in $[N]$ of type (D', a, b), and find restrictions on D', a and b. First, fix D'. There are clearly at most N possibilities for a. It is straightforward to prove by induction that for $\ell \in \{1, \dots, D'\}$

$$-2^{\ell-1}N < \Delta^{\ell}(a_i)(v) < 2^{\ell-1}N.$$

Since $\Delta^{D'}(a_i)$ must be a nonzero constant sequence of integers, there are fewer than $2^{D'}N$ possibilities for the constant sequence $(b) = \Delta^{D'}(a_i)$. Summing this total over $1 \le D' \le D$ yields the claim. $\square$

3 A base case and an inductive step

Following Łaba & Lacey (it is implicitly in Rankin), we proceed by induction. The statements of the next two propositions are extremely similar to what appears in Łaba & Lacey, but the proofs are considerably messier.

Proposition 1 (Base Case)**.** *If $k > 2D$, then as $N \to \infty$*

$$\frac{r_{k,D}(N)}{N} \ge \left(\frac{\sqrt{90}}{e\pi^{3/2}} \, \frac{2^D}{D^{1/4}} \binom{2D}{D} - o(1) \right) \frac{\sqrt[4]{2\log N}}{2^{\sqrt{8D\log N}}}. \tag{1}$$

Proposition 2 (Inductive Step)**.** *If $k > 2D$, then there exists a positive constant C*

$$\frac{r_{k,D}(N)}{N} \ge C\, 2^{-dD}\, \frac{r_{k,2D}(N_0)}{N_0},$$

where

$$N_0 := \frac{e\pi}{3\sqrt{5}} \left(4^D \binom{2D}{D} \right)^{-1} \frac{N^{2/d}}{d^{1/2}}.$$

Let A_0 be a subset of $[N_0]$ with cardinality $r_{k,2D}(N_0)$ that does not contain any k-term $2D$-progression, assume $2\delta N_0 \leq 2^{-2D}$, and let

$$A := A(\overline{\omega}, \overline{\alpha}) = \{n \in [N] \colon n\,\overline{\omega} + \overline{\alpha} \bmod \overline{1} \in \textsc{Annuli}(A_0, N_0, D, \delta)\},$$

which we will show is typically (with respect to $\overline{\omega}, \overline{\alpha}$ being chosen uniformly from $\textsc{Box}_0$) a set with many elements and few types of D-progressions. After removing one element from A for each type of progression it contains, we will be left with a set that has large size and no k-term D-progressions. Since $\textsc{Box}_0 \times \textsc{Box}_0$ has Lebesgue measure 1, this argument could be easily recast in terms of Lebesgue integrals, but we prefer the probabilistic notation and language.

Define $T := T(\overline{\omega}, \overline{\alpha})$ to be the set

$$\left\{ a \in [N] \colon \begin{array}{c} \exists b \in \mathbb{R}, D' \in [D] \text{ such that } A(\overline{\omega}, \overline{\alpha}) \text{ contains} \\ \text{a } k\text{-term progression of type } (D', a, b) \end{array} \right\},$$

which is contained in $A(\overline{\omega}, \overline{\alpha})$. Observe that $A \setminus T$ is a subset of $[N]$ and contains no k-term D-progressions, and consequently $r_{k,D}(N) \geq |A \setminus T| = |A| - |T|$ for every $\overline{\omega}, \overline{\alpha}$. In particular,

$$r_{k,D}(N) \geq \mathbb{E}_{\overline{\omega},\overline{\alpha}}\left[|A| - |T|\right] = \mathbb{E}_{\overline{\omega},\overline{\alpha}}\left[|A|\right] - \mathbb{E}_{\overline{\omega},\overline{\alpha}}\left[|T|\right]. \tag{2}$$

First, we note that

$$\mathbb{E}_{\overline{\omega},\overline{\alpha}}\left[|A|\right] = \sum_{n=1}^{N} \mathbb{P}_{\overline{\omega},\overline{\alpha}}\left[n \in A\right] = \sum_{n=1}^{N} \mathbb{P}_{\overline{\alpha}}\left[n \in A\right] = N\,\mathbf{vol}(\textsc{Annuli}(A_0, N_0, D, \delta)).$$

Let $E(D', a, b) = 1$ if A contains a k-term progression of type (D', a, b), and otherwise set $E(D', a, b) = 0$. We have

$$|T| \leq \sum_{(D',a,b)} E(D', a, b),$$

where the sum extends over all types (D', a, b) for which $D' \in [D]$ and there is a D'-progression of that type contained in $[N]$; by Lemma 4 there are fewer than $2^{D+1}N^2$ such types.

Suppose that A has a k-term progression of type (D', a, b), with $D' \in [D]$. Let p be a degree D' polynomial with lead term $p_{D'} \neq 0$, $p(1), \ldots, p(k)$ a D'-progression contained in A, and $\Delta^{D'}(p(i)) = (b)$. Then

$$\overline{x}_i := p(i)\,\overline{\omega} + \overline{\alpha} \bmod \overline{1} \in \textsc{Annuli}(A_0, N_0, D, \delta) \subseteq \textsc{Box}_D\,.$$

By Lemma 2, the $\overline{x}_i$ are a D'-progression in $\mathbb{R}^d$, say $\overline{P}(j) = \sum_{i=0}^{D'} \overline{P}_i j^i$ has $\overline{P}(j) = \overline{x}_j$ and $D'!\overline{P}_{D'} = D'!p_{D'}\,\overline{\omega} \bmod \overline{1} = b\,\overline{\omega} \bmod \overline{1}$. By elementary algebra

$$Q(j) := \frac{\|\overline{P}(j)\|_2^2 - \mu_D}{\sigma_D} - z$$

is a degree $2D'$ polynomial in j, and since $\overline{P}(j) = \overline{x}_j \in \textsc{Annuli}(A_0, N_0, D, \delta)$ for $j \in [k]$, we know that

$$Q(j) \in \bigcup_{a \in A_0} \left(-\frac{a-1}{N_0} \pm \delta \right)$$

for all $j \in [k]$, and also $Q(1), \ldots, Q(k)$ is a $2D'$-progression. Define the real numbers $a_j \in A_0$, $\epsilon_j \in \pm\delta$ by

$$Q(j) = -\frac{a_j - 1}{N_0} + \epsilon_j.$$

We need to handle two cases separately: either the sequence (a_i) is constant or it is not. Suppose first that it is not constant. Since $a_i \in A_0$, a set without k-term $2D$-progressions, we know that $\Delta^{2D+1}(a_i) \neq (0)$, and since (a_i) is a sequence of integers, for some v

$$|\Delta^{2D+1}(a_i)(v)| \geq 1.$$

Consider:

$$(0) = \Delta^{2D+1}(Q(i)) = \frac{1}{N_0}\Delta^{2D+1}(a_i) + \Delta^{2D+1}(\epsilon_i),$$

whence

$$|\Delta^{2D+1}(\epsilon_i)(v)| = \frac{1}{N_0}|\Delta^{2D+1}(a_i)(v)| \geq \frac{1}{N_0}.$$

Since $|\epsilon_i| < \delta$, we find that $|\Delta^{2D+1}(\epsilon_i)(v)| < 2^{2D+1}\delta$, and since we assumed that $2\delta N_0 \leq 2^{-2D}$, we arrive at the impossibility

$$\frac{1}{N_0} \leq |\Delta^{2D+1}(\epsilon_i)(v)| < 2^{2D+1}\delta \leq 2^{2D} \cdot \frac{2^{-2D}}{N_0} = \frac{1}{N_0}.$$

Now assume that (a_i) is a constant sequence, say $a := a_i$, so that

$$Q(j) \in -\frac{a-1}{N_0} \pm \delta$$

for all $j \in [k]$. This translates to

$$\|\overline{P}(j)\|_2^2 \in \mu_D - (z - \frac{a-1}{N_0})\sigma_D \pm \delta\sigma_D.$$

Using Lemma 1, the lead coefficient $\overline{P}_{D'}$ of $\overline{P}(j)$ satisfies

$$\begin{aligned} \|D'!\overline{P}_{D'}\|_2 \leq D'!\, 2^{D'}(2D')!^{-1/2}\sqrt{\delta\sigma_D} \leq D!\, 2^D(2D)!^{-1/2}\sqrt{\delta\sigma_D} \\ = \left(\frac{4^D\, D!^2}{(2D)!}\right)^{1/2}\sqrt{\sigma_D\delta} = \sqrt{F\sigma_D\delta}, \end{aligned}$$

where $F := 4^D/\binom{2D}{D}$. We have deduced that $E(D', a, b) = 1$ only if

$$a\,\overline{\omega} + \overline{\alpha} \bmod 1 \in \textsc{Annuli}(A_0, N_0, D, \delta) \quad \text{and} \quad \|b\,\overline{\omega} \bmod 1\|_2 \leq \sqrt{F\sigma_D\delta}.$$

Since $\overline{\alpha}$ is chosen uniformly from $\textsc{Box}_0$, we notice that

$$\mathbb{P}_{\overline{\alpha}}\left[a\,\overline{\omega}+\overline{\alpha} \bmod 1 \in \textsc{Annuli}(A_0,N_0,D,\delta)\right] = \mathbf{vol}\,\textsc{Annuli}(A_0,N_0,D,\delta),$$

independent of $\overline{\omega}$. Also, we notice that the event $\{\|b\,\overline{\omega} \bmod 1\|_2 \le \sqrt{F\sigma_D\delta}\}$ is independent of $\overline{\alpha}$, and that since b is an integer, $\overline{\omega} \bmod \overline{1}$ and $b\,\overline{\omega} \bmod \overline{1}$ are identically distributed. Therefore, the event $\{\|b\,\overline{\omega} \bmod 1\|_2 \le \sqrt{F\sigma_D\delta}\}$ has probability at most[2]

$$\mathbf{vol}\,\textsc{Ball}(\sqrt{F\sigma_D\delta}) = \frac{2\pi^{d/2}(\sqrt{F\sigma_D\delta})^d}{\Gamma(d/2)d},$$

where $\textsc{Ball}(x)$ is the d-dimensional ball in $\mathbb{R}^d$ with radius x. It follows that

$$\mathbb{P}_{\overline{\omega},\overline{\alpha}}\left[E(D',a,b)=1\right] \le \mathbf{vol}\,\textsc{Annuli}(A_0,N_0,D,\delta)\cdot \mathbf{vol}\,\textsc{Ball}(\sqrt{F\sigma_D\delta}),$$

and so

$$\mathbb{E}_{\overline{\omega},\overline{\alpha}}\left[|T|\right] \le 2^{D+1}N^2\,\mathbf{vol}\,\textsc{Annuli}(A_0,N_0,D,\delta)\cdot \mathbf{vol}\,\textsc{Ball}(\sqrt{F\sigma_D\delta}).$$

Equation (2) now gives us

$$\frac{r_{k,D}(N)}{N} \ge \mathbf{vol}(\textsc{Annuli}(A_0,N_0,D,\delta))\left(1-2^{D+1}N\,\mathbf{vol}\,\textsc{Ball}(\sqrt{F\sigma_D\delta})\right).$$

Setting

$$\delta := \frac{1}{\pi F}\left(\frac{d}{(d+2)2^{D+1}}\right)^{2/d}\frac{\Gamma(d/2)^{2/d}}{N^{2/d}\sigma_D} \sim \frac{3\sqrt{5}}{e\pi}\binom{2D}{D}\frac{d^{1/2}}{N^{2/d}},$$

we observe that

$$1-2^{D+1}N\frac{2\pi^{d/2}(F\delta^{1/2}d^{1/4})^d}{\Gamma(d/2)d} = \frac{d}{d+2}.$$

3.1 Finish proof of Proposition 1

We set

$$d := \left\lfloor \sqrt{\frac{2\log N}{D}} \right\rfloor,$$

[2]In fact, since we will shortly choose δ so that $F\sigma_D\delta \to 0$, this upper bound cannot be improved.

so that $\delta \log d \to 0$, and

$$\begin{aligned}
\frac{r_{k,D}(N)}{N} &\geq \frac{d}{d+2}\ \mathbf{vol}\,\textsc{Annuli}(\{1\},1,D,\delta) \\
&\geq \frac{d}{d+2}\left(\sqrt{\frac{2}{\pi}} - o(1)\right) 2^{-dD}\,\delta \\
&\geq \frac{d}{d+2}\left(\sqrt{\frac{2}{\pi}} - o(1)\right) 2^{-dD}\,\frac{1}{\pi F}\left(\frac{d}{(d+2)2^{D+1}}\right)^{2/d}\frac{\Gamma(d/2)^{2/d}}{N^{2/d}\sigma_D} \\
&= \left(\frac{\sqrt{2}}{\pi^{3/2}F} - o(1)\right) 2^{-dD}\,\frac{\Gamma(d/2)^{2/d}}{N^{2/d}\sigma_D} \\
&= \left(\frac{\sqrt{2}}{\pi^{3/2}F} - o(1)\right) 2^{-dD}\,\frac{(1+o(1))d/2e}{N^{2/d}2^{-2D}\sqrt{d/180}} \\
&\geq \left(\frac{2^{2D}\sqrt{360}}{2e\pi^{3/2}F} - o(1)\right) 2^{-dD}\,\frac{\sqrt{d}}{N^{2/d}} \\
&= \left(\frac{\sqrt{90}}{e\pi^{3/2}}\binom{2D}{D} - o(1)\right)\sqrt{d}\,\exp\big(-(dD + \frac{2}{d}\log N)\big).
\end{aligned}$$

Define the error term $\epsilon(N)$ by

$$dD + \frac{2}{d}\log N = \sqrt{8D\log N} + \epsilon(N),$$

and observe that for any integer ℓ, we have $\epsilon(x)$ monotone increasing on $[2^{\ell^2 D/2}, 2^{(\ell+1)^2 D/2})$, while N being in that interval gives $d = \ell$. By algebra $\epsilon(2^{\ell^2 D/2}) = 0$, and also

$$\lim_{N \to \exp((d+1)^2 D/2)} \epsilon(N) = \frac{D}{d}.$$

It follows that $\epsilon(N) \leq D/(\sqrt{2(\log N)/D} - 1)$.

From this, we see that

$$\begin{aligned}
\exp\big(-(dD + \frac{2}{d}\log N)\big) &\geq \exp(-\sqrt{8D\log N})\exp\left(\frac{D}{\sqrt{2(\log N)/D} - 1}\right) \\
&= (1+o(1))\exp(-\sqrt{8D\log N}),
\end{aligned}$$

which completes the proof of Proposition 1.

3.2 Finish proof of Proposition 2

We set

$$N_0 := \frac{e\pi}{3\sqrt{5}}\left(4^D\binom{2D}{D}\right)^{-1}\frac{N^{2/d}}{d^{1/2}},$$

which accomplishes $\frac{1}{4}2^{-2D} \leq 2\delta N_0 \leq 2^{-2D}$. With this δ, N_0, and Lemma 3 we have

$$\frac{r_{k,D}(N)}{N} \geq \frac{d}{d+2}\,\mathbf{vol}\,\textsc{Annuli}(A_0, N_0, D, \delta) \geq \frac{d}{d+2}\,\frac{2}{5}\,2^{-dD}\,|A_0|\,\delta$$
$$\geq \frac{1}{2}\,\frac{2}{5}\,2^{-dD}\,\frac{|A_0|}{N_0}\,\delta N_0 = C\,2^{-dD}\,\frac{r_{k,2D}(N_0)}{N_0}.$$

4 Proof of Theorem 1

We proceed by induction, with the base case of $n = 2$ following immediately from Proposition 1. We now assume that Theorem 1 holds for n, assume that $k > 2^n D$, and show that

$$\frac{r_{k,D}(N)}{N} \geq C\frac{(\log N)^{1/(2n+2)}}{\exp\left((n+1)2^{n/2}D^{n/(n+1)}\sqrt[n+1]{\log N}\right)}.$$

By Proposition 2, we have

$$\frac{r_{k,D}(N)}{N} \geq C\frac{1}{2^{dD}}\,\frac{r_{k,2D}(N_0)}{N_0},$$

with $N_0 = CN^{2/d}d^{-1/2}$. Since $k > 2^n D = 2^{n-1}(2D)$, the inductive hypothesis gives us

$$\begin{aligned}\frac{r_{k,D}(N)}{N} &\geq C\frac{1}{2^{dD}}\frac{(\log N_0)^{1/(2n)}}{\exp\left(n2^{(n-1)/2}(2D)^{(n-1)/n}\sqrt[n]{\log N_0}\right)}\\ &= C\frac{(\log N_0)^{1/(2n)}}{\exp\left(dD + n2^{(n-1)/2}(2D)^{(n-1)/n}\sqrt[n]{\log C - \frac{1}{2}\log d + \frac{2}{d}\log N}\right)}\\ &\geq C\frac{(\log N_0)^{1/(2n)}}{\exp\left(dD + n2^{(n-1)/2}(2D)^{(n-1)/n}\sqrt[n]{\frac{2}{d}\log N}\right)},\end{aligned}$$

with the final inequality coming from d being sufficiently large.

Setting

$$d := \left\lfloor 2^{n/2}\left(\frac{\log N}{D}\right)^{1/(n+1)}\right\rfloor$$

we arrive at the error term and bound for it:

$$dD + n2^{(n-1)/2}(2D)^{(n-1)/n}\sqrt[n]{\frac{2}{d}\log N} = (n+1)2^{n/2}D^{n/(n+1)}(\log N)^{1/(n+1)} + \epsilon(N)$$

where

$$\epsilon(N) \leq (1+o(1))\frac{(n+1)D^{(n+2)/(n+1)}}{n\,2^{n/2+1}\,(\log N)^{1/(n+1)}} \leq \frac{C}{(\log N)^{1/(n+1)}}.$$

Thus,

$$\exp\biggl(-\left(dD + n2^{(n-1)/2}(2D)^{(n-1)/n}\sqrt[n]{\frac{2}{d}\log N}\right)\biggr) \geq$$
$$(1+o(1))\exp\left(-(n+1)2^{n/2}D^{n/(n+1)}(\log N)^{1/(n+1)}\right).$$

5 Further Thoughts

The approach here works *mutatis mutandis* for constructing a subset of an arbitrary set $\mathcal{N}$ of N integers. The number of progressions in $\mathcal{N}$ becomes a critical parameter, and the inductive step is somewhat more technical. The specific changes are detailed in [9].

Further, the methods here can serve as a basic outline for constructing thick subsets of a large arbitrary set that does not contain nontrivial solutions to a linear system of equations. This problem has seen recent progress due to Shapira [12], but a universal thick construction remains elusive.

References

[1] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **32** (1946), 331–332. MR 0018694 (8,317d)

[2] Michael Elkin, *An improved construction of progression-free sets* (January 28, 2008), 20 pp., available at `arXiv:0801.4310`. Version 1.

[3] William Feller, *An introduction to probability theory and its applications, volume II.*

[4] Timothy Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588, DOI 10.1007/s00039-001-0332-9. MR **1844079 (2002k:**11014)

[5] Ben Green and Terence Tao, *New bounds for Szemerédi's theorem. II. A new bound for $r_4(N)$*, Analytic number theory, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204. MR **2508645 (2010b:**11016)

[6] Ben Green and Julia Wolf, *A note on Elkin's improvement of Behrend's constructions* (October 5, 2008), 4 pp., available at `arXiv:0810.0732`. Version 1.

[7] Gil Kalai and W. T. Gowers, *Pushing Behrend Around*, Combinatorics and more, Weblog of Gil Kalai, (2009). gilkalai.wordpress.com/2008/07/10/pushing-behrend-around.

[8] Izabella Łaba and Michael T. Lacey, *On sets of integers not containing long arithmetic progressions* (August 22, 2001), 8 pp., available at `arXiv:math.CO/0108.155`. Version 1.

[9] Kevin O'Bryant, *Thick subsets that do not contain arithmetic progressions* (June 21, 2010), available at `arxiv.org/abs/0912.1494`.

[10] R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh Sect. A **65** (1960/1961), 332–344 (1960/61). MR 0142526 (26 #95)

[11] Tom Sanders, *On Roth's theorem on progressions* (October 30, 2010), 13 pages pp., available at `arxiv:1011.0104`.

[12] Asaf Shapira, *A proof of Green's conjecture regarding the removal properties of sets of linear equations* (August 21, 2008), available at `arxiv.org/abs/0807.4901`.