*Mathematika* 61 (2015) 49–62 © University College London
doi:10.1112/S0025579314000175

# ON IMPROVING ROTH'S THEOREM IN THE PRIMES

ERIC NASLUND

*Abstract.* Let $A \subset \{1, \dots, N\}$ be a set of prime numbers containing no non-trivial arithmetic progressions. Suppose that $A$ has relative density $\alpha = |A|/\pi(N)$, where $\pi(N)$ denotes the number of primes in the set $\{1, \dots, N\}$. By modifying Helfgott and De Roton's work [Improving Roth's theorem in the primes. *Int. Math. Res. Not. IMRN* **2011** (4) (2011), 767–783], we improve their bound and show that

$$\alpha \ll \frac{(\log\log\log N)^6}{\log\log N}.$$

§1. *Introduction.* In 1936, Erdös and Turán [3] conjectured that if a set $A \subset \mathbb{N} = \{1, 2, 3 \dots\}$ does not contain a non-trivial $k$ term arithmetic progression then

$$\limsup_{x\to\infty} \frac{1}{x} \sum_{n\leqslant x} 1_A(x) = 0,$$

or in other words its natural upper density is zero. In 1953, Roth [10] proved this result for three-term arithmetic progressions. Quantitatively he showed that any progression-free set of integers $A$ satisfies

$$|A \cap \{1, 2, \dots, N\}| \ll \frac{N}{\log\log N}.$$

Roth's theorem has been improved significantly over the last 60 years by Heath-Brown, [7], Szemerédi [12], Bourgain [1, 2] and most recently Sanders [11], who obtained the bound

$$|A \cap \{1, 2, \dots, N\}| \ll \frac{N(\log\log N)^5}{\log N}.$$

Moving to the set of prime numbers, which we will denote by $\mathcal{P}$, we define the relative density of a set $A \subset \mathcal{P}$ up to $N$ to be

$$\alpha(N) = \frac{|A \cap \{1, 2, \dots, N\}|}{|\mathcal{P} \cap \{1, 2, \dots, N\}|}. \tag{1.1}$$

In 1939, Van Der Corput [13] showed that $\mathcal{P}$ contains infinitely many non-trivial three-term arithmetic progressions. Green [4] proved an analogue of

Received 13 February 2013, published online 9 October 2014.
MSC (2010): 11B25, 11N13, 11N36, 11P32, 11B30 (primary).

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

Roth's theorem inside the primes, showing that if $A \subset \mathcal{P}$ contains no non-trivial arithmetic progressions, then

$$\alpha(N) \ll \left(\frac{\log\log\log\log\log N}{\log\log\log\log N}\right)^{1/2},$$

where the notation $f(N) \ll g(N)$ means that there exists an absolute constant $C$ such that $f(N) \leqslant Cg(N)$ for all $N \geqslant 1$. Helfgott and De Roton [8] improved this density bound, removing two logs from the denominator and obtaining

$$\alpha(N) \ll \frac{\log\log\log N}{(\log\log N)^{1/3}}. \tag{1.2}$$

Their result implicitly uses the best quantitative bound on Roth's Theorem in the integers, and when the proof is run through again with Sander's bound, the density recovered is

$$\alpha(N) \ll \frac{(\log\log\log N)^{5/2}}{(\log\log N)^{1/2}}. \tag{1.3}$$

Our main result is the following theorem.

THEOREM 1. *Suppose that $A \subset \mathcal{P} \cap [1, N]$ has relative density $\alpha$ and contains no non-trivial arithmetic progressions. Then*

$$\alpha \ll \frac{(\log\log\log N)^6}{\log\log N}.$$

Our proof parallels that of Helfgott and De Roton, and we look at the convolution of the indicator function of the set of primes and the indicator function of a set $\Sigma$. We gain a factor of two in the exponent by using the $L^{2k}$ norm, where $k$ is a slowly growing function of $N$, rather than the $L^2$ norm. Using this higher norm introduces several combinatorial difficulties which are dealt with in §2 and in the proof of proposition 2. This $L^{2k}$ norm bound gives greater control over the outliers, and allows us to choose a larger subset on which the convolution is uniformly bounded from below. As in Helfgott and De Roton, the bound on Roth's theorem yields a lower bound on the size of the three term progression operator applied to this uniform set. If the set $\Sigma$ is chosen correctly, the three-term progression operator of the convolution cannot be too far from that of the indicator function of the primes, which gives the desired density bound for the primes.

1.1. *Preliminaries and notation.* For two functions $f, g : \mathbb{N} \to \mathbb{R}$, we write $f \ll g$, or $f(x) = O(g(x))$ if there exists a constant $C > 0$ such that $|f(n)| \leqslant Cg(n)$ for all positive integers $n$. Often we will look at when $f \ll g$ for sufficiently large $n$, which means that there exists $N_0, C > 0$ with $|f(n)| \leqslant Cg(n)$ for all $n \geqslant N_0$.

To denote $(1/|S|)\sum_{x\in S} f(x)$, the expectation of $f$ over the set $S$, we write $\mathbb{E}_{x\in S} f(x)$. Given a function $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, where $N$ is a prime, we define the

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

Fourier transform to be

$$\hat{f}(t) = \mathbb{E}_{x\in\mathbb{Z}/N\mathbb{Z}} f(x)\mathrm{e}^{2\pi \mathrm{i}xt/N}.$$

The convolution operation is given by

$$(f * g)(x) = \mathbb{E}_{y\in\mathbb{Z}/N\mathbb{Z}} f(y)g(x-y), \tag{1.4}$$

which is suitably normalized so that

$$\widehat{f * g}(t) = \hat{f}(t)\hat{g}(t). \tag{1.5}$$

The $L^k$ and $\ell^k$ norms are defined to be

$$\|f\|_{L^k(\mathbb{Z}/N\mathbb{Z})} = (\mathbb{E}_{x\in\mathbb{Z}/N\mathbb{Z}}|f(x)|^k)^{1/k},$$

and

$$\|\hat{f}\|_{\ell^k(\mathbb{Z}/N\mathbb{Z})} = \Bigg(\sum_{x\in\mathbb{Z}/N\mathbb{Z}} |\hat{f}(x)|^k\Bigg)^{1/k}.$$

When there is no ambiguity, we will omit the notation $\ell^k(\mathbb{Z}/N\mathbb{Z})$ and $L^k(\mathbb{Z}/N\mathbb{Z})$, and simply write $\|\cdot\|_k$. We will make use of the fact that the inner product $\langle f, g\rangle_{L^2(\mathbb{Z}/N\mathbb{Z})} = \mathbb{E}_{x\in\mathbb{Z}/N\mathbb{Z}} f(x)\overline{g(x)}$, satisfies Plancherel's identity

$$\mathbb{E}_{x\in\mathbb{Z}/N\mathbb{Z}} f(x)\overline{g(x)} = \sum_{t\in\mathbb{Z}/N\mathbb{Z}} \widehat{f}(t)\overline{\hat{g}(t)}, \tag{1.6}$$

from which we obtain $\|f\|_{L^2(G)} = \|\hat{f}\|_{\ell^2(G)}$. Given functions $f, g, h : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, we let $\Lambda(f, g, h)$ denote the three-term arithmetic progression operator defined by

$$\Lambda(f, g, h) = \mathbb{E}_{x,d\in\mathbb{Z}/N\mathbb{Z}} f(x)g(x+d)h(x+2d).$$

If $1_A$ is the indicator function of a set $A \subset \mathbb{Z}/N\mathbb{Z}$, then $\Lambda(1_A, 1_A, 1_A)$ counts the total number of three-term progressions in $A$, including the trivial progressions. For a set $\Sigma \subset \mathbb{Z}/N\mathbb{Z}$, we let $|\Sigma|$ denote the cardinality of $\Sigma$, and $\mu(\Sigma) = |\Sigma|/N$ denote the relative measure.

§2. *Sieving the primes.* Let $A \subset [1, N]$ be a subset of the primes with $|A| = \alpha(N/\log N)$, and suppose that $\alpha \geqslant (\log N)^{-1/4}$. We will remove the small primes using the '$W$ trick', which allows us to effectively apply certain sieve results later on. Let $W = \prod_{p\leqslant z} p$ be the product of the primes less than $z$. Splitting into the different arithmetic progressions modulo $W$, there will be exactly $\phi(W)$ non-trivial residue classes. By the pigeon hole principle there exists an arithmetic progression $AP(b) = \{b + nW : 1 \leqslant n \leqslant N/W\}$ with

$$|AP(b) \cap A| \geqslant \alpha\frac{N}{\log N}\frac{1}{\phi(W)} - \frac{W}{\phi(W)} \tag{2.1}$$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

where the $W/\phi(W)$ on the right-hand side appears since we are not including the primes up to $W$. Let $P$ be the least prime larger than $3N/W$, so that $3(N/W) < P \leqslant 6(N/W)$, and let $A_0 \subset [1, P]$ be the set

$$A_0 = \left\{ n = \frac{m-b}{W} : m \in AP(b) \cap A \right\},$$

noting that an arithmetic progression in $A_0$ can be lifted to a progression in $A$. Using equation (2.1) along with some basic asymptotics for the number of primes, we can find a lower bound for the size of $A_0$. Note that

$$\log W = \sum_{p \leqslant z} \log p = \theta(z) \sim z, \tag{2.2}$$

so that $W \approx \mathrm{e}^z$, and

$$\frac{W}{\phi(W)} = \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right)^{-1} \sim \mathrm{e}^{\gamma} \log z$$

where $\gamma = \lim_{x \to \infty} (\sum_{n \leqslant x} 1/n - \log x) \approx 0.577$ is the Euler–Mascheroni constant. Choosing $z = \frac{1}{4} \log N$, and $N$ sufficiently large, we may assume that

$$\tfrac{4}{5} z \leqslant \log W \leqslant \tfrac{4}{3} z,$$

which means that the modulus has size $N^{1/5} \leqslant W \leqslant N^{1/3}$, and

$$\log z \leqslant \frac{W}{\phi(W)} \leqslant 2 \log z. \tag{2.3}$$

Equation (2.3) along with the inequality $P/6 \leqslant N/W$ implies that and $(P \log z)/6 \leqslant N/\phi(W)$, and so by (2.1) we have that

$$\begin{aligned} |A_0| &\geqslant \alpha P \frac{\log z}{6 \log N} - 2 \log z \\ &\geqslant \alpha P \frac{\log z}{10 \log N} \end{aligned}$$

for $N$ sufficiently large. Each arithmetic progression inside $A_0$ corresponds to an arithmetic progression in $A$, and so if $A_0$ contains a three-term arithmetic progression, then $A$ must as well. With this in mind, we shift our attention to the progressions inside $A_0$. Define $a = (\log N / \log z) 1_{A_0}$ to be the normalized indicator function for the set $A_0$, which is supported on $[0, P/3]$ since we chose $P > 3N/W$. This function satisfies

$$\|a\|_1 = \mathbb{E}_{n \in \mathbb{Z}/P\mathbb{Z}} a(n) \geqslant \frac{\alpha}{10}, \tag{2.4}$$

and

$$\|a\|_2^2 = \mathbb{E}_{n \in \mathbb{Z}/P\mathbb{Z}} a(n)^2 = \frac{\log N}{\log z} \mathbb{E}_{n \in \mathbb{Z}/P\mathbb{Z}} a(n) = \frac{\log N}{\log z} \|a\|_1.$$

In §3 we will examine the key quantity $\Lambda(a, a, a)$ in detail, and show that it cannot be too small when $\alpha$ is large. To do this, we will need a bound on the $L^{2k}$ norm of the convolution of $a$ and an indicator function $1_{\Sigma}$, which is discussed in the following section.

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

2.1. *Bounding the $L^{2k}$ norm of the convolution.* Our goal is to provide bounds on the $L^{2k}$ norm of $a * \sigma$ where $\sigma$ is the indicator function of a set $\Sigma$.

PROPOSITION 2. *Given $z$, $N$, $P$ and $a(n)$ as above, let $k$ be an integer in the range $1 \leqslant k \leqslant \frac{1}{2}\log^{1/3} z$. Suppose that $\sigma = 1_\Sigma/\mu(\Sigma)$ is the normalized indicator function of a set $\Sigma \subset [-P/3, P/3] \subset \mathbb{Z}/P\mathbb{Z}$. Then for $N$ greater than some fixed $N_0$, we have the bound*

$$\|a * \sigma\|_{2k} \ll k + \left(\frac{\log N}{\log z}\right)^{1-1/2k} |\Sigma|^{-1/2k}.$$

We will make use of a theorem of Klimov on the Selberg sieve. Theorem 3 of [9] is as follows.

THEOREM 3 (Klimov). *Let $1 \leqslant i \leqslant k$, $1 \leqslant n \leqslant M$, $v_0 \leqslant v \leqslant \sqrt{M}/(\log^{2k} M)$, for a fixed $v_0$, and define $M_v(q_i, l_i)$ to be the number of integers $n$ for which $p \nmid q_i n + l_i$, for each $p \leqslant v$, and each $1 \leqslant i \leqslant k$. Then if $u_0 = O(\exp(\log^B v))$ for a fixed constant $B > 0$, for $k \geqslant 2$, we have*

$$M_v(q_i, l_i) \leqslant \frac{M}{\log^k v} k! \prod_p \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}\left(1 + O\left(\frac{\log\log u_0}{\log v}\right)\right),$$

*where $u_0 = \max(q_i, u_{i,j})$, $u_{i,j} = |l_i q_j - l_j q_i|$ $(1 \leqslant j \leqslant k)$, and $\rho(p)$ is the number of $n \in \{1, \dots, p\}$ such that*

$$\prod_{i=1}^{k} (q_i n + l_i) \equiv 0 \bmod p.$$

We elect to reference the above theorem, rather than Theorem 5.7 from Halberstam and Richert [6], as Klimov's result allows us to make the dependence on the number variables explicit. As a direct corollary, we have the following result.

LEMMA 4. *Let $k \geqslant 2$, and let $W$ and $b$ satisfy $\log b \leqslant 2\log P$ and $\log W \leqslant 2\log P$, where $P \geqslant P_0$ for some fixed $P_0$. Suppose that we have $k$ pairwise distinct $b_i$, all relatively prime to $W$, such that $b_i \leqslant b$ for $1 \leqslant i \leqslant k$, and that $k \leqslant \log P/(12\log\log P)$. Then*

$$\begin{aligned}&|\{n \leqslant P : b_1 + nW, \dots, b_k + nW \text{ all prime}\}| \\ &\ll P \frac{3^k k!}{\log^k P} \prod_p \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k},\end{aligned}$$

*where the constant does not depend on $k$.*

*Proof.* In Theorem 3, let $q_i = W$, $l_i = b_i$, $M = P$, $v = P^{1/3}$, and so that $u_0 \leqslant 2bW$. To apply Klimov's theorem, we must have $v \leqslant \sqrt{P}/(\log^{2k}(P))$, and $u_0 = O(\exp(\log^B v))$. The first condition follows from the assumption that

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

$k \leqslant \log P/(12 \log\log P)$. The upper bound from $u_0$ is satisfied with $B = 15$ since $\log(2bW) \leqslant 5 \log P \leqslant 15 \log v$ by the hypothesis of the lemma. It follows that

$$\frac{\log\log u_0}{\log v} = O\left(\frac{\log\log N}{\log N}\right),$$

and so

$$M_v(W, b_i) \ll P \frac{3^k k!}{\log^k P} \prod_p \left(1 - \frac{\rho(p)}{p}\right)(1 - 1/p)^{-k},$$

as desired. □

We note that if $k > \log P/(12 \log\log P)$, the term $3^k k!/\log^k P$ will be larger than one, and so the bound is weaker than the trivial upper of $P$. Using this lemma, we prove Proposition 2.

*Proof.* By the definition of the $L^{2k}$ norm and the convolution, we have that

$$\|a * \sigma\|_{2k}^{2k} = \mathbb{E}_x |(a * \sigma)(x)|^{2k} = \mathbb{E}_x |\mathbb{E}_y \sigma(y) a(x - y)|^{2k}.$$

Expanding the sum, $\|a * \sigma\|_{2k}^{2k}$ is bounded above by

$$\mathbb{E}_{y_1, \ldots, y_{2k}} |\sigma(y_1)| \cdots |\sigma(y_{2k})| \mathbb{E}_x a(x - y_1) \cdots a(x - y_{2k}).$$

Since $a$ is supported on $[0, P/3]$, and $\sigma$ is supported on $[-P/3, P/3]$, there can be no wrap around inside $\mathbb{Z}/P\mathbb{Z}$, and we have the upper bound

$$\mathbb{E}_x a(x - y_1) \cdots a(x - y_{2k}) \leqslant 1/P \left(\frac{\log N}{\log z}\right)^{2k} |\mathcal{A}(y_1, y_2, \ldots, y_{2k})| \qquad (2.5)$$

where

$$\mathcal{A}(y_1, y_2, \ldots, y_{2k}) = \{n \leqslant P : b + (n - y_1)W, \ldots, b + (n - y_{2k})W \text{ all prime}\}.$$

The size of this set of primes is bounded above by the number of $n \leqslant P$ such that $b + (n - y_1)W$ is prime, and so by the Brun–Titchmarsh inequality

$$|\mathcal{A}(y_1, y_2, \ldots, y_{2k})| \leqslant \frac{2PW}{\phi(W) \log(P/W)},$$

which we combine with the inequalities $P/W \geqslant N^{1/3}$ and $W/\phi(W) \leqslant 2 \log z$ to obtain

$$|\mathcal{A}(y_1, y_2, \ldots, y_{2k})| \leqslant \frac{12P \log z}{\log N}. \qquad (2.6)$$

However, we will need our sieving lemma in the case where each of the $y_i$ is distinct. Let

$$I_l = \{(y_1, \ldots, y_{2k}) : y_i \leqslant P/2, \text{ with } l \text{ or less distinct coordinates } y_i\},$$
$$J_l = \{(y_1, \ldots, y_{2k}) : y_i \leqslant P/2, \text{ with exactly } l \text{ distinct coordinates } y_i\},$$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

so that we may bound our quantity from above by a sum from $r = 1$ to $2k$, and over $(y_1, \dots, y_{2k}) \in I_r$ or $J_r$. We split this into two cases. When $r < 2k$, we will bound above by the sum over $I_r$, and when $r = 2k$, we will use a sum over elements in the set $J_{2k}$. For $r < 2k$, we are looking at

$$\frac{1}{P^{2k+1}} \sum_{r=1}^{2k-1} \sum_{(y_1,\dots,y_{2k})\in I_r} |\sigma(y_1)| \cdots |\sigma(y_{2k})| \sum_x a(x-y_1)\cdots a(x-y_{2k}) \quad (2.7)$$

which by (2.5) and (2.6) is

$$\ll \left(\frac{\log N}{\log z}\right)^{2k-1} \sum_{r=1}^{2k-1} \frac{1}{P^{2k}} \sum_{(y_1,\dots,y_{2k})\in I_r} |\sigma(y_1)| \cdots |\sigma(y_{2k})|.$$

Each term in the sum over $r$ on the right-hand side may be bounded above by $1/|\Sigma|^{2k-r}$ since

$$\begin{aligned}\frac{1}{P^{2k}} \sum_{(y_1,\dots,y_{2k})\in I_r} |\sigma(y_1)| \cdots |\sigma(y_{2k})| &= \frac{1}{|\Sigma|^{2k-r}} (\mathbb{E}_{y_1,\dots,y_r} \sigma(y_1)\cdots\sigma(y_r)) \\ &= \frac{1}{|\Sigma|^{2k-r}}.\end{aligned}$$

If $|\Sigma| = 1$, the lemma is trivial, and if $|\Sigma| \geqslant 2$ we have

$$\sum_{r=1}^{2k-1} \frac{1}{|\Sigma|^{2k-r}} \leqslant \frac{1}{|\Sigma|-1} \leqslant \frac{2}{|\Sigma|},$$

which implies that the quantity in (2.7) is

$$\ll \left(\frac{\log N}{\log z}\right)^{2k-1} \frac{2}{|\Sigma|}.$$

To bound the cardinality of $\mathcal{A}(y_1, \dots, y_{2k})$ for $(y_1, \dots, y_{2k}) \in J_{2k}$, we apply Lemma 4. The conditions of the lemma are satisfied as $b$, $W \leqslant N^{1/3}$ and $k \leqslant \log P/(12 \log\log P)$ since $k \leqslant \frac{1}{2}\log^{1/3} z$. Thus, for any $2k$-tuple $(y_1, \dots, y_{2k}) \in J_{2k}$,

$$|\mathcal{A}(y_1, y_2, \dots, y_{2k})| \ll \frac{3^{2k}(2k)!P}{(\log P)^{2k}} \prod_p \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-2k}. \quad (2.8)$$

On the right-hand side of equation (2.8) $\rho$ is defined so that $\rho(p) = 0$ if $p \leqslant z$, $\rho(p) = 2k$ if $p > z$ and $p \nmid (y_i - y_j)$ for all $i \neq j$. The product over all $p > z$ where $\rho(p) = 2k$ is extremely well behaved, and we have that

$$\prod_{p>z} \left(1 - \frac{2k}{p}\right)\left(1 - \frac{1}{p}\right)^{-2k}$$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

is bounded from above and below by absolute constants. To see why, note that

$$\log \prod_{p>z}\left(1-\frac{2k}{p}\right)\left(1-\frac{1}{p}\right)^{-2k} = 2k\sum_{p>z}\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{1}{p}\right)^{j} - \sum_{p>z}\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{2k}{p}\right)^{j}$$
$$= \sum_{p>z}\sum_{j=2}^{\infty}\frac{1}{j}\left(2k\left(\frac{1}{p}\right)^{j} - \left(\frac{2k}{p}\right)^{j}\right),$$

and as $k \leqslant \frac{1}{2}\log z$, we have the lower bound

$$\sum_{p>z}\sum_{j=2}^{\infty}\frac{1}{j}\left(2k\left(\frac{1}{p}\right)^{j} - \left(\frac{2k}{p}\right)^{j}\right) \geqslant -\sum_{j=2}^{\infty}\frac{1}{j}(\log z)^{j}\sum_{p>z}\frac{1}{p^{j}}$$
$$\geqslant -\sum_{j=2}^{\infty}\frac{1}{j}(\log z)^{j}\sum_{n>z}\frac{1}{n^{j}}$$
$$\gg -\frac{\log^{2} z}{z},$$

and similarly we have the upper bound

$$\sum_{p>z}\sum_{j=2}^{\infty}\frac{1}{j}\left(2k\left(\frac{1}{p}\right)^{j} - \left(\frac{2k}{p}\right)^{j}\right) \leqslant 0$$

as $2k(1/p)^{j} \leqslant (2k/p)^{j}$ for all $j$. From these bounds, it follows that for a constant independent of $k$,

$$\prod_{p}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-2k} \ll (\log z)^{2k}\prod_{\substack{p>z\\ p\mid y_i-y_j}}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-2k}.$$

Since $|y_i - y_j| \leqslant P$, and any integer $y \leqslant P$ can have at most $\log_z P$ prime factors greater than $z$, we see that there are at most $(2k)^2(\log P/\log z)$ primes greater than $z$ which could divide some difference $y_i - y_j$. For each $p$ that divides $y_i - y_j$ for some $i, j$, the worst case is when $\rho(p) = 1$, and we may assume that this is the case to obtain an upper bound. As $(1-1/p)^{-1} \leqslant z/(z-1)$ for $p > z$, we have

$$\prod_{\substack{p>z\\ p\mid y_i-y_j}}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-2k} \leqslant \left(\frac{z}{z-1}\right)^{(4k^2)(2k-1)\log P/\log z}.$$

Recall that $\log P \leqslant \log N \leqslant 4z$ since $z = \frac{1}{4}\log N$. The exponent is bounded above by

$$(4k^2)(2k-1)\frac{\log P}{\log z} \leqslant 8k^3\frac{\log P}{\log z}$$
$$\leqslant \frac{32k^3}{\log z}z$$
$$\leqslant 4z,$$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

where the final inequality follows from the assumption that $k \leqslant \frac{1}{2}(\log z)^{1/3}$. Since

$$\left(\frac{z}{z-1}\right)^{4z} \leqslant 5\mathrm{e}^4$$

for $z \geqslant 2$, we obtain the inequality

$$\prod_{\substack{p>z \\ p \mid y_i - y_j}} \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-2k} \leqslant 5\mathrm{e}^4,$$

and so equation (2.8) becomes

$$|\mathcal{A}(y_1, y_2, \ldots, y_{2k})| \leqslant C3^{2k}(2k)!P\frac{(\log z)^{2k}}{(\log P)^{2k}},$$

for an absolute constant $C$. Thus, for any $(y_1, \ldots, y_{2k}) \in J_{2k}$,

$$\mathbb{E}_x a(x - y_1)\cdots a(x - y_{2k}) \leqslant C3^{2k}(2k)!.$$

Since

$$\frac{1}{P^{2k}} \sum_{(y_1,\ldots,y_{2k})\in J_{2k}} |\sigma(y_1)|\cdots|\sigma(y_{2k})| \leqslant \mathbb{E}_{y_1,\ldots,y_{2k}}|\sigma(y_1)|\cdots|\sigma(y_{2k})| = 1,$$

the sum over all the $2k$-tuples in $J_{2k}$ is $\leqslant 1$. Combining the work done so far, we have proven that

$$\|a * \sigma\|_{2k}^{2k} \ll 3^{2k}(2k)! + \frac{2k}{|\Sigma|}\left(\frac{\log N}{\log z}\right)^{2k-1}.$$

The stated result then follows from the fact that $(n+m)^{1/2k} \leqslant n^{1/2k} + m^{1/2k}$ for $n, m, k \geqslant 1$, and since

$$(3^{2k}(2k)!)^{1/2k} \leqslant 4k, \quad \text{and} \quad (3k)^{1/2k} \leqslant \mathrm{e}^{3/2e} \leqslant 2. \qquad \square$$

§3. *Main theorem.* Let $a$, $N$, $W$, $P$ and $z$ be defined as in §2. Following [8], we define

$$R = \mathrm{Spec}_\delta(a) \cup \{1\} = \{x \in \mathbb{Z}/P\mathbb{Z} : |\widehat{a}(x)| \geqslant \delta\} \cup \{1\},$$

and

$$B = B(R, \epsilon) = \{n \in \mathbb{Z}/P\mathbb{Z} : \forall x \in R, \|nx/P\| \leqslant \epsilon\},$$

where $\|x\|$ denotes the distance from $x$ to the nearest integer. The set $B$ is called a *Bohr set* with radius $\epsilon$ and frequency set $R$. Set $\sigma = (1/\mu(B))1_B$ to be the normalized indicator function of the Bohr set $B$. By including the element 1 in the set $R$, it follows that $\sigma$ will be supported on $[-P/4, P/4]$ inside $\mathbb{Z}/P\mathbb{Z}$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

when $\epsilon < \frac{1}{4}$. Let $h = a * \sigma$ be our prime indicator smoothed out by the Bohr set $B$. Note that

$$\|\sigma\|_1 = \frac{1}{\mu(B)}\mathbb{E}_{x\in\mathbb{Z}/P\mathbb{Z}}1_B = 1,$$

and

$$\begin{aligned}\|h\|_1 &= \mathbb{E}_{x\in\mathbb{Z}/P\mathbb{Z}}\mathbb{E}_{y\in\mathbb{Z}/P\mathbb{Z}}\frac{1_B(y)}{\mu(B)}a(x-y)\\ &= \|\sigma\|_1\|a\|_1,\end{aligned}$$

so that by (2.4), which gave the bound $\|a\|_1 \geqslant \alpha/10$, we have

$$\|h_1\| \geqslant \frac{\alpha}{10}. \tag{3.1}$$

Our goal is to show that there is little difference between the three-term arithmetic progression operator applied to $a$ and $h$, and then prove that $\Lambda(h, h, h)$ is large. Let

$$\Delta = |\Lambda(a, a, a) - \Lambda(h, h, h)|,$$

where $\Lambda$ is the three term arithmetic progression operator. In Helfgott and De Roton's paper [8], equation (2.6) on page 7 states the following result.

LEMMA 5. *For the above definition of* $\Delta, \epsilon, \delta$ *we have*

$$\Delta \ll \epsilon + \delta^{3/5}.$$

The proof of this lemma makes use of Green and Tao's results on the restriction theory of the Selberg sieve [5]. Applying Proposition 2 with $\sigma = a$, we find that

$$\|a * a\|_2 \ll 1,$$

and so $\|\hat{a}\|_4 \ll 1$ since $\|a*a\|_2^2 = \|\hat{a}\|_4^4$ by (1.6) and (1.5). This yields the bound $|R| \leqslant C_4\delta^{-4}$, on the size of the dimension of the Bohr set $B$, for an absolute constant $C_4$, as

$$|R|\delta^4 \leqslant \sum_t |\hat{a}(t)|^4 \leqslant C_4.$$

A well-known pigeon hole argument tells us that $|B(R, \epsilon)| \geqslant N\epsilon^{|R|}$, and so

$$\log|B| \geqslant \log N - |R||\log\epsilon| \geqslant \log N - C_4\delta^{-4}|\log\epsilon|.$$

We note that an equation nearly identical to the above appears on page 9 of [8]. We chose to deduce it again since the bound on $\|\hat{a}\|_4$ was obtained in a different way. From now on, we will assume that $\epsilon, \delta$ satisfy

$$C_4\delta^{-4}|\log\epsilon| \leqslant \tfrac{1}{2}\log N, \tag{3.2}$$

so that

$$|B| \geqslant N^{1/2}.$$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

For $k \leqslant \frac{1}{2}(\log z)^{1/3}$, Proposition 2 allows us to bound the $\ell^{2k}$ norm of $h$. Using the inequality

$$\tfrac{1}{2}\log N \geqslant (\log\log N)^{4/3},$$

which holds for all $N \geqslant 2$, along with the fact that

$$k \leqslant \tfrac{1}{2}(\log z)^{1/3} \leqslant \tfrac{1}{2}(\log\log N)^{1/3},$$

we see that

$$N^{1/2} \geqslant (\log N)^{2k} \geqslant \left(\frac{\log N}{\log z}\right)^{2k},$$

and consequently

$$|B| \geqslant N^{1/2} \geqslant \left(\frac{\log N}{\log z}\right)^{2k}.$$

Proposition 2 then implies that

$$\|h\|_{2k} = \|a * \sigma\|_{2k} \ll k. \tag{3.3}$$

Using this bound $L^{2k}$ norm of $h = a * \sigma$, along with Sanders bound on Roth's theorem, we are able to show that $\Lambda(h, h, h)$ must be large.

PROPOSITION 6. *There exists positive constants $c_1$, $N_0 > 0$ such that if $N \geqslant N_0$, for any $1 \leqslant k \leqslant \frac{1}{2}(\log z)^{1/3}$ we have*

$$\Lambda(h, h, h) \gg \exp\left(-c_1\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right)$$

*where $q_{2k} = (1 - 1/2k)^{-1}$.*

To prove this proposition, we will make use of the following lemma which allows us to find a large subset where the function $h$ is bounded below uniformly.

LEMMA 7. *Let $q, p > 1$ be such that $1/q + 1/p = 1$, and let $f : \mathbb{Z}/P\mathbb{Z} \to \mathbb{R}^+$ be a function with $\|f\|_1 \geqslant \alpha$, and $\|f\|_p \leqslant C$ for some $C$. Then there exists a subset $L \subset \mathbb{Z}/P\mathbb{Z}$ such for all $n \in L$ $f(n) \geqslant \alpha/2$, and*

$$\left(\frac{\alpha}{2C}\right)^q \leqslant \mu(L).$$

*Proof.* Define $L = \{n : f(n) \geqslant \alpha/2\}$, so that $L$ is the largest possible set satisfying the first condition. Then, if $1_L(n)$ is the indicator function for $L$ we have that

$$\alpha \leqslant \mathbb{E}_{n\in\mathbb{Z}/P\mathbb{Z}} f(n) \leqslant \frac{\alpha}{2} + \mathbb{E}_{n\in\mathbb{Z}/P\mathbb{Z}} f(n)1_L(n).$$

Applying Hölder's inequality yields

$$\mathbb{E}_{n\in\mathbb{Z}/P\mathbb{Z}} f(n)1_L(n) \leqslant \|1_L\|_q \|f\|_p \leqslant C(\mu(L))^{1/q}$$

by our assumption that $\|f\|_p \leqslant C$. The lemma then follows from the resulting inequality $\alpha/2 \leqslant C(\mu(L))^{1/q}$. □

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

Sanders improvement to Roth's theorem [11] states that if $\Xi \subset [1, N]$ with density $\xi$, then

$$\Lambda(1_\Xi, 1_\Xi, 1_\Xi) \gg \exp\left(-c\xi^{-1}\left(\log\frac{1}{\xi}\right)^5\right). \tag{3.4}$$

Using this result along with (3.3) and Lemma 7, we are ready to finish the proof of Proposition 6.

*Proof.* By equation (3.3), it follows that $\|h\|_{2k} \ll k$. Applying Lemma 7 to the function $h$, we obtain a subset $L \subset \mathbb{Z}/P\mathbb{Z}$ with

$$\mu(L) \gg \left(\frac{\alpha}{k}\right)^{q_{2k}},$$

and $h(n) \geqslant \alpha/20$ for all $n \in L$, where $1/q_{2k} = 1 - 1/2k$. Restricting $h$ to this subset $L$, we obtain the lower bound

$$\Lambda(h, h, h) \geqslant \frac{\alpha^3}{20^3}\Lambda(1_L, 1_L, 1_L). \tag{3.5}$$

Applying the bound in (3.4) to our set $L$, we have that

$$\Lambda(1_L, 1_L, 1_L) \gg \exp\left(-c\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right)$$

for some constant $c$, since the density of $L$ is $\gg (\alpha/k)^{q_{2k}}$. By equation (3.5) it follows that there is a constant $c_1$ such that

$$\Lambda(h, h, h) \gg \exp\left(-c_1\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right),$$

as desired. □

Lemma 5 tells us that

$$|\Lambda(h, h, h) - \Lambda(a, a, a)| \ll \epsilon + \delta^{3/5},$$

and Proposition 6 implies that $\Lambda(h, h, h)$ must be very large. Recalling equation (3.2), the requirement that $\epsilon, \delta$ satisfy $C_4\delta^{-4}|\log\epsilon| \leqslant \frac{1}{2}\log N$, we are ready to put everything together and give a precise lower bound for the size of $\Lambda(a, a, a)$. We now prove Theorem 1.

*Proof.* Suppose that $A \subset \mathcal{P}$ contains no non-trivial arithmetic progressions. Then $A_0$ contains only the trivial three-term arithmetic progressions, and we have that the three-term arithmetic progression operator is bounded above by

$$\Lambda(a, a, a) \ll \frac{1}{P}\left(\frac{\log N}{\log z}\right)^2. \tag{3.6}$$

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

By equation (3.6), Lemma 5 and Proposition 6 , we must have

$$\frac{1}{P}\left(\frac{\log N}{\log z}\right)^2 + \epsilon + \delta^{3/5} \gg \exp\left(-c_1\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right). \tag{3.7}$$

In the above, for $N$ sufficiently large, the term $1/P(\log N/\log z)^2$ will be negligible compared with the right-hand side as we assumed that $\alpha > (\log N)^{-1/4}$. Choosing $\epsilon$ and $\delta$ small enough will lead us to a contradiction. In particular, there exists a fixed postie constant $\eta > 0$, independent of $N$ and $k$, such that choosing

$$\epsilon = \delta^{3/5} = \eta \exp\left(-c_1\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right),$$

makes inequality (3.7) impossible. These values of $\epsilon, \delta$ satisfy the necessary constraint $C_4|\log\epsilon|\delta^{-4} \leqslant \frac{1}{2}\log N$ as long as

$$\exp\left(c_2\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right) \ll \frac{1}{2}\log N \tag{3.8}$$

for some new constant $c_2$, where the $\ll$ has consumed $\eta$. When

$$\alpha \geqslant k(2c_2)^{1/q_{2k}}\frac{(\log\log\log N)^{5/q_{2k}}}{(\log\log N)^{1/q_{2k}}},$$

we have that

$$\exp\left(c_2\left(\frac{\alpha}{k}\right)^{-q_{2k}}\left(\log\frac{1}{\alpha}\right)^5\right) \leqslant \sqrt{\log N},$$

and so inequality (3.8) holds for sufficiently large $N$. Letting $2k = [\log\log\log N]$, which satisfies the necessary bound $k \leqslant \frac{1}{2}(\log z)^{1/3}$, we have that

$$\left(\frac{\log\log N}{\log\log\log N}\right)^{1/2k} \sim e,$$

and so

$$k(2c_2)^{1/q_{2k}}\frac{(\log\log\log N)^{5/q_{2k}}}{(\log\log N)^{1/q_{2k}}} \geqslant C\frac{(\log\log\log N)^6}{\log\log N},$$

for some constant $C$. This means that we will have a contradiction when $N$ is sufficiently large and the density $\alpha$ satisfies

$$\alpha \geqslant C\frac{(\log\log\log N)^6}{\log\log N}$$

for some absolute constant $C$, which proves the desired result. □

*Remark* 8. The proof suggests that we might need the condition $N \geqslant N_0$ for some fixed constant $N_0$ in the main theorem. Note, however, that this is not necessary given how the result is phrased, as the constant in the $\ll$ will be so large that it accounts for this.

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press

*Acknowledgements.* I am very grateful to Julia Wolf for her support, encouragement and generous help. I am especially thankful for her patience during our conversations on the Selberg sieve. I would like to thank Greg Martin for his helpful comments, and Daniel Fiorilli for leading me to the paper of Klimov, as well as the anonymous referee for his many useful suggestions.

## *References*

**1.** J. Bourgain, On triples in arithmetic progression. *Geom. Funct. Anal.* **9**(5) (1999), 968–984.
**2.** J. Bourgain, Roth's theorem on progressions revisited. *J. Anal. Math.* **104** (2008), 155–192.
**3.** P. Erdös and P. Turán, On some sequences of integers. *J. Lond. Math. Soc.* **S1–11**(4) (1936), 261–264.
**4.** B. Green, Roth's theorem in the primes. *Ann. of Math.* (2) **161**(3) (2005), 1609–1636.
**5.** B. Green and T. Tao, Restriction theory of the Selberg sieve, with applications. *J. Théor. Nombres Bordeaux* **18**(1) (2006), 147–182.
**6.** H. Halberstam and H.-E. Richert, *Sieve Methods* (*London Mathematical Society Monographs* **4**), Academic Press (London, 1974).
**7.** D. R. Heath-Brown, Integer sets containing no arithmetic progressions. *J. Lond. Math. Soc.* (2) **35**(3) (1987), 385–394.
**8.** H. A. Helfgott and A. de Roton, Improving Roth's theorem in the primes. *Int. Math. Res. Not. IMRN* **2011**(4) (2011), 767–783.
**9.** N. I. Klimov, Combination of elementary and analytic methods in the theory of numbers. *Uspehi Mat. Nauk (N.S.)* **13**(3(81)) (1958), 145–164.
**10.** K. F. Roth, On certain sets of integers. *J. Lond. Math. Soc.* **28** (1953), 104–109.
**11.** T. Sanders, On Roth's theorem on progressions. *Ann. of Math.* (2) **174**(1) (2011), 619–636.
**12.** E. Szemerédi, Integer sets containing no arithmetic progressions. *Acta Math. Hungar.* **56**(1–2) (1990), 155–158.
**13.** J. G. van der Corput, Über Summen von Primzahlen und Primzahlquadraten. *Math. Ann.* **116**(1) (1939), 1–50.

Eric Naslund,
Princeton University Mathematics Department,
Fine Hall Room 304,
Princeton, NJ 08544-1044,
USA
E-mail: naslund@math.princeton.edu

https://doi.org/10.1112/S0025579314000175 Published online by Cambridge University Press