A SHORT PROOF OF THE MULTIDIMENSIONAL SZEMERÉDI THEOREM IN THE PRIMES

JACOB FOX AND YUFEI ZHAO

Abstract. Tao conjectured that every dense subset of \( P_d \), the \( d \)-tuples of primes, contains constellations of any given shape. This was very recently proved by Cook, Magyar, and Titichetrakun and independently by Tao and Ziegler. Here we give a simple proof using the Green-Tao theorem on linear equations in primes and the Furstenberg-Katznelson multidimensional Szemerédi theorem.

Let \( P_N \) denote the set of primes at most \( N \), and let \( [N] := \{1, 2, \ldots, N\} \). Tao [13] conjectured the following result as a natural extension of the Green-Tao theorem [8] on arithmetic progressions in the primes and the Furstenberg-Katznelson [7] multidimensional generalization of Szemerédi’s theorem. Special cases of this conjecture were proven in [5] and [12]. The conjecture was very recently resolved by Cook, Magyar, and Titichetrakun [6] and independently by Tao and Ziegler [14].

Theorem 1. Let \( d \) be a positive integer, \( v_1, \ldots, v_k \in \mathbb{Z}^d \), and \( \delta > 0 \). Then, if \( N \) is sufficiently large, every subset \( A \) of \( P_N^d \) of cardinality \( |A| \geq \delta |P_N|^d \) contains a set of the form \( a + tv_1, \ldots, a + tv_k \), where \( a \in \mathbb{Z}^d \) and \( t \) is a positive integer.

In this note we give a short alternative proof of the theorem, using the landmark result of Green and Tao [9] (which is conditional on results later proved in [10] and with Ziegler in [11]) on the asymptotics for the number of primes satisfying certain systems of linear equations, as well as the following multidimensional generalization of Szemerédi’s theorem established by Furstenberg and Katznelson [7].

Theorem 2 (Multidimensional Szemerédi theorem [7]). Let \( d \) be a positive integer, \( v_1, \ldots, v_k \in \mathbb{Z}^d \), and \( \delta > 0 \). If \( N \) is sufficiently large, then every subset \( A \) of \( [N]^d \) of cardinality \( |A| \geq \delta N^d \) contains a set of the form \( a + tv_1, \ldots, a + tv_k \), where \( a \in \mathbb{Z}^d \) and \( t \) is a positive integer.

To prove Theorem 1, we begin by fixing \( d, v_1, \ldots, v_k, \delta \). Using Theorem 2, we can fix a large integer \( m > 2d/\delta \) so that any subset of \( [m]^d \) with at least \( \delta m^d/2 \) elements contains a set of the form \( a + tv_1, \ldots, a + tv_k \), where \( a \in \mathbb{Z}^d \) and \( t \) is a positive integer.

We next discuss a sketch of the proof idea. The Green-Tao theorem [8] (also see [3, 4] for some recent simplifications) states that there are arbitrarily long arithmetic progressions in the primes. It follows that for \( N \) large, \( P_N^d \) contains homothetic copies of \( [m]^d \). We use a Varnavides-type argument [15] and consider a random homothetic copy of the grid \([m]^d\) inside \( P_N^d \). In expectation, the set \( A \) should occupy at least a \( \delta/2 \) fraction of the random homothetic copy of \([m]^d\). This follows from a linearity of expectation argument. Indeed, the Green-Tao-Ziegler result [9, 10, 11] and a second moment argument imply that most points of \( P_N^d \) appear in about the expected number of such copies of the grid \([m]^d\). Once we find a homothetic copy of \([m]^d\) containing at least \( \delta m^d/2 \) elements of \( A \), we obtain by Theorem 2 a subset of \( A \) of the form \( a + tv_1, \ldots, a + tv_k \), as desired.

To make the above idea actually work, we first apply the \( W \)-trick as described below. This is done to avoid certain biases in the primes. We also only consider homothetic copies of \([m]^d\) with

The first author was supported by a Packard Fellowship, NSF CAREER award DMS-1352121, by an Alfred P. Sloan Fellowship, and by an MIT NEC Corporation Fund Award, and the second author was supported by an internship at Microsoft Research New England and a Microsoft Research PhD Fellowship.
common difference $r \leq N/m^2$ in order to guarantee that almost all elements of $P^d_N$ are in about the same number of such homothetic copies of $[m]^d$.

**Remarks.** This argument also produces a relative multidimensional Szemerédi theorem, where the complexity of the linear forms condition on the majorizing measure depends on $d, v_1, \ldots, v_k$ and $\delta$. It seems plausible that the dependence on $\delta$ is unnecessary; this was shown for the one-dimensional case in [3]. Our arguments share some features with those of Tao and Ziegler [14], who also use the results in [9, 10, 11]. However, the proof in [14] first establishes a relativized version of the Furstenberg correspondence principle and then proceeds in the ergodic theoretic setting, whereas we go directly to the multidimensional Szemerédi theorem. Cook, Magyar, and Titichetrakun [6] take a different approach and develop a relative hypergraph removal lemma from scratch, and they also require a linear forms condition whose complexity depend on $\delta$.

Conditional on a certain polynomial extension of the Green-Tao-Ziegler result (c.f. the Bateman-Horn conjecture [1]), one can also combine this sampling argument with the polynomial extension of Szemerédi’s theorem by Bergelson and Leibman [2] to obtain a polynomial extension of Theorem 1.

The hypothesis that $|A| \geq \delta |P^d_N|$ implies that

$$
\sum_{n_1, \ldots, n_d \in [N]} 1_A(n_1, \ldots, n_d)\Lambda'(n_1) \cdots \Lambda'(n_d) \geq (\delta - o(1)) N^d,
$$

where $1_A$ is the indicator function of $A$, and $o(1)$ denotes some quantity that goes to zero as $N \to \infty$, and $\Lambda'(p) = \log p$ for prime $p$ and $\Lambda'(n) = 0$ for nonprime $n$.

Next we apply the $W$-trick [9, §5]. Fix some slowly growing function $w = w(N)$; the choice $w := \log \log \log N$ will do. Define $W := \prod_{p \leq w} p$ to be the product of all primes at most $w$. For each $b \in [W]$ with $\gcd(b, W) = 1$, define

$$
\Lambda'_{b,W}(n) := \frac{\phi(W)}{W} \Lambda(W_n + b)
$$

where $\phi(W) = \#\{b \in [W] : \gcd(b, W) = 1\}$ is the Euler totient function. Also define

$$
1_{A b_1, \ldots, b_d,W}(n_1, \ldots, n_d) := 1_A(W n_1 + b_1, \ldots, W n_d + b_d).
$$

By (1) and the pigeonhole principle, we can choose $b_1, \ldots, b_d \in [W]$ all coprime to $W$ so that

$$
\sum_{1 \leq n_1, \ldots, n_d \leq N/W} 1_{A b_1, \ldots, b_d,W}(n_1, \ldots, n_d)\Lambda'_1,1_{b_1,W}(n)\Lambda'_2,2_{b_2,W}(n) \cdots \Lambda'_d, d_{b_d,W}(n) \geq (\delta - o(1)) \left( \frac{N}{W} \right)^d.
$$

We shall write

$$
\tilde{N} := \lceil N/W \rceil, \quad R := \lceil \tilde{N}/m^2 \rceil, \quad \tilde{A} := \prod_{b \in [W]} 1_{A b_1, \ldots, b_d,W} \quad \text{and} \quad \tilde{\Lambda}_j := \Lambda'_{b_j,W}.
$$

(all depending on $N$). So (2) reads

$$
\sum_{n_1, \ldots, n_d \in [\tilde{N}]} \tilde{A}(n_1, \ldots, n_d)\tilde{\Lambda}_1(n_1)\tilde{\Lambda}_2(n_2) \cdots \tilde{\Lambda}_d(n_d) \geq (\delta - o(1)) \tilde{N}^d
$$

The Green-Tao result [9] (along with [10, 11]) says that $\Lambda'_{b_j,W}$ acts pseudorandomly with average value about 1 in terms of counts of linear forms. The statement below is an easy corollary of [9, Thm. 5.1].

**Theorem 3** (Pseudorandomness of the $W$-tricked primes). Fix a linear map $\Psi = (\psi_1, \ldots, \psi_k) : \mathbb{Z}^d \to \mathbb{Z}^l$ (in particular $\Psi(0) = 0$) where no two $\psi_i, \psi_j$ are linearly dependent. Let $K \subseteq [-\tilde{N}, \tilde{N}]^d$ be any convex body. Then, for any $b_1, \ldots, b_l \in [W]$ all coprime to $W$, we have

$$
\sum_{n \in K \cap \mathbb{Z}^d} \prod_{j \in [l]} \Lambda'_{b_j,W}(\psi_j(n)) = \# \{n \in K \cap \mathbb{Z}^d : \psi_j(n) > 0 \ \forall j\} + o(\tilde{N}^d).
$$
Recall that $R$ by (3) and Theorem 3 we have
\[ \sum_{n_1, \ldots, n_d \in \tilde{N}} \left( \sum_{i_1, \ldots, i_d \in [m]} \bar{A}(n_1 + i_1 r, \ldots, n_d + i_d r) \right) \prod_{j \in [d]} \prod_{i \in [m]} \bar{A}_j(n_j + ir) \]
\[ \geq (\delta m^d - dm^{d-1} - o(1)) R \tilde{N}^d. \]  

Proof of Theorem 1 (assuming Lemma 4). By Theorem 3 we have
\[ \sum_{n_1, \ldots, n_d \in \tilde{N}} \prod_{r \in [R]} \prod_{j \in [d]} \prod_{i \in [m]} \bar{A}_j(n_j + ir) = (1 + o(1)) R \tilde{N}^d, \]
So by (4), for sufficiently large $N$, there exists some choice of $n_1, \ldots, n_d \in \tilde{N}$ and $r \in [R]$ so that
\[ \sum_{i_1, \ldots, i_d \in [m]} \bar{A}(n_1 + i_1 r, \ldots, n_d + i_d r) \geq \frac{1}{2} \delta m^d. \]
This means that a certain dilation of the grid $[m]^d$ contains at least $\delta m^d/2$ elements of $A$, from which it follows by the choice of $m$ that it must contain a set of the form $a + tv_1, \ldots, a + tv_k$. \qed

Lemma 4 follows from the next lemma by summing over all choices of $i_1, \ldots, i_d \in [m]$.

Lemma 5. Suppose $\bar{A}$ satisfies (3). Fix $i_1, \ldots, i_d \in [m]$. Then we have
\[ \sum_{n_1, \ldots, n_d \in \tilde{N}} \bar{A}(n_1 + i_1 r, \ldots, n_d + i_d r) \prod_{j \in [d]} \prod_{i \in [m]} \bar{A}_j(n_j + ir) \geq \left( \delta - \frac{d}{m} - o(1) \right) R \tilde{N}^d. \]  

Proof. By a change of variables $n'_j = n_j + i_j r$ for each $j$, we write the LHS of (5) as
\[ \sum_{r \in [R]} \left( \sum_{n'_1, \ldots, n'_d \in [m]} \sum_{n'_j - i_j r \in [\tilde{N}]} \bar{A}(n'_1, \ldots, n'_d) \prod_{j \in [d]} \prod_{i \in [m]} \bar{A}_j(n'_j + (i - i_j) r). \]  
Recall that $R = \lceil \tilde{N}/m^2 \rceil$. Note that (6) is at least
\[ \sum_{r \in [R]} \tilde{N}/m< n'_1, \ldots, n'_d \leq \tilde{N} \bar{A}(n'_1, \ldots, n'_d) \prod_{j \in [d]} \prod_{i \in [m]} \bar{A}_j(n'_j + (i - i_j) r). \]
By (3) and Theorem 3 we have
\[ \sum_{\tilde{N}/m< n'_1, \ldots, n'_d \leq \tilde{N}} \bar{A}(n_1, \ldots, n_d) \bar{A}_1(n_1) \bar{A}_2(n_2) \cdots \bar{A}_d(n_d) \geq \left( \delta - \frac{d}{m} - o(1) \right) \tilde{N}^d \]  
-the difference between the left-hand side sums of (3) and (8) consists of terms with $(n_1, \ldots, n_d)$ in some box of the form $[\tilde{N}]^{j-1} \times \tilde{N}/m \times [\tilde{N}]^{d-j}$, which can be upper bounded by using $\bar{A} \leq 1$, applying Theorem 3, and then taking the union bound over $j \in [d]$). It remains to show that
\[ (7) - R \cdot (\text{LHS of } (8)) = o(\tilde{N}^{d+1}). \]
We have

\[
(7) - R \cdot \text{(LHS of (8))}
\]

\[
= \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \tilde{A}(n_1', \ldots, n_d') \left( \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j' + (i - j)r) - \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j') \right)
\]

\[
= \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \tilde{A}(n_1', \ldots, n_d') \left( \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j') \right) \left( \sum_{r \in [R]} \left( \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j' + (i - j)r) - 1 \right) \right)
\]

By the Cauchy-Schwarz inequality and 0 \leq \tilde{A} \leq 1, the above expression can be bounded in absolute value by \( \sqrt{ST} \), where

\[
S = \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j'),
\]

\[
T = \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \left( \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j') \right) \left( \sum_{r \in [R]} \left( \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j' + (i - j)r) - 1 \right) \right)^2
\]

\[
= T_1 - 2T_2 + T_3,
\]

and

\[
T_1 = \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j') \prod_{i \in [m] \setminus \{i_j\}} \tilde{A}_j(n_j' + (i - j)r) \tilde{A}_j(n_j' + (i - j)r'),
\]

\[
T_2 = \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \prod_{\tilde{j} \in [d]} \tilde{A}_j(n_j') \prod_{i \in [m] \setminus \{i_j\}} \tilde{A}_j(n_j' + (i - j)r),
\]

\[
T_3 = \sum_{\tilde{N}/m < n_1', \ldots, n_d' \leq \tilde{N}} \prod_{r, r' \in [R]} \tilde{A}_j(n_j').
\]

By Theorem 3 we have \( S = O(\tilde{N}^d) \), and \( T_1, T_2, T_3 \) pairwise differ by \( o(\tilde{N}^{d+2}) \), so that \( T = o(\tilde{N}^{d+2}) \). Thus \( \sqrt{ST} = o(\tilde{N}^{d+1}) \), as desired. \( \blacksquare \)

**Acknowledgment.** We thank David Conlon, Ben Green, and Terence Tao, and the anonymous referee for helpful comments.

**References**

[1] P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. 16 (1962), 363–367.

[2] V. Bergegbson and A. Leibman, *Polynomial extensions of van der Waerden’s and Szemerédi’s theorems*, J. Amer. Math. Soc. 9 (1996), no. 3, 725–753.

[3] D. Conlon, J. Fox, and Y. Zhao, *A relative Szemerédi theorem*, arXiv:1305.5440.

[4] D. Conlon, J. Fox, and Y. Zhao, *The Green–Tao theorem: an exposition*, arXiv:1403.2957.

[5] B. Cook, A. Magyar, *Constellations in \( \mathbb{P}^d \)*, Int. Math. Res. Not. IMRN 2012 (2012), no. 12, 2794–2818.

[6] B. Cook, A. Magyar, and T. Titichetrakun, *A multidimensional Szemerédi theorem in the primes*, arXiv:1306.3025.

[7] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. 34 (1978), 275–291.
[8] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, Ann. of Math. 167 (2008), no. 2, 481–547.

[9] ———, Linear equations in primes, Ann. of Math. 171 (2010), no. 3, 1753–1850.

[10] ———, The Möbius function is strongly orthogonal to nilsequences, Ann. of Math. 175 (2012), no. 2, 541–566.

[11] B. Green, T. Tao, and T. Ziegler, An inverse theorem for the Gowers $U^{k+1}[N]$-norm, Ann. of Math. 176 (2012), no. 2, 1231–1372.

[12] A. Magyar, and T. Titichetrakun, Corners in dense subsets of $\mathbb{F}^d$, arXiv:1306.3026.

[13] T. Tao, The Gaussian primes contain arbitrarily shaped constellations, J. Anal. Math. 99 (2006), 109–176.

[14] T. Tao and T. Ziegler, A multi-dimensional Szemerédi theorem for the primes via a correspondence principle, arXiv:1306.2886.

[15] P. Varnavides, On certain sets of positive density, J. London Math. Soc. 34 (1959), 358–360.

Department of Mathematics, MIT, Cambridge, MA 02139-4307
E-mail address: fox@math.mit.edu

Department of Mathematics, MIT, Cambridge, MA 02139-4307
E-mail address: yufeiz@math.mit.edu