Mathematical Proceedings of the Cambridge Philosophical Society

VOL. 177 NOVEMBER 2024 PART 3

Math. Proc. Camb. Phil. Soc. (2024), **177**, 371–413 © The Author(s), 2024. Published by Cambridge University Press on behalf of Cambridge Philosophical Society.
doi: 10.1017/S0305004124000264

First published online 27 November 2024

Improved bounds for five-term arithmetic progressions

BY JAMES LENG[†]
Department of Mathematics, UCLA, Los Angeles, CA 90095, *USA.*
e-mail: James.leng@math.ucla.edu

ASHWIN SAH AND MEHTAAB SAWHNEY[‡]
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, *USA.*
e-mails: asah@mit.edu, msawhney@mit.edu

(*Received* 10 *April* 2024*; accepted* 22 *May* 2024)

Abstract

Let $r_5(N)$ be the largest cardinality of a set in $\{1, \dots, N\}$ which does not contain 5 elements in arithmetic progression. Then there exists a constant $c \in (0, 1)$ such that

$$r_5(N) \ll \frac{N}{\exp((\log\log N)^c)}.$$

Our work is a consequence of recent improved bounds on the U^4-inverse theorem of J. Leng and the fact that 3-step nilsequences may be approximated by locally cubic functions on shifted Bohr sets. This, combined with the density increment strategy of Heath–Brown and Szemerédi, codified by Green and Tao, gives the desired result.

2020 Mathematics Subject Classification: 05D10 (Primary)

[†]Leng supported by an NSF Graduate research fellowship grant no. DGE-2034835.

[‡]Sah and Sawhney were supported by NSF Graduate research fellowship program DGE-2141064.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

1. *Introduction*

Let $[N] = \{1, \dots, N\}$ and $r_k(N)$ denote the size of the largest $S \subseteq \{1, \dots, N\}$ such that S has no k-term arithmetic progressions. The first nontrivial upper bound on $r_3(N)$ came from work of Roth [28] which proved

$$r_3(N) \ll N(\log\log N)^{-1}.$$

A long series of works improved the bounds in the case $k = 3$. This includes works of Heath–Brown [16], Szemerédi [35], Bourgain [5, 6], Sanders [30, 31], Bloom [2], and Bloom and Sisask [3]. Recently, in breakthrough work, Kelley and Meka [18] proved that

$$r_3(N) \ll N \exp(-\Omega((\log N)^{1/12}));$$

modulo the constant $1/12$ this matches the lower bound construction of Behrend [1]. The constant $1/12$ was refined to $1/9$ in work of Bloom and Sisask [4].

For higher k, establishing that $r_k(N) = o(N)$ was a long standing conjecture of Erdös and Turán. In seminal works, Szemerédi [33] first established the estimate $r_4(N) = o(N)$ and then in later work Szemerédi [34] established his eponymous theorem that

$$r_k(N) = o(N).$$

Due to the uses of van der Waerden theorem and the regularity lemma (which was introduced in this work), Szemerédi's estimate was exceedingly weak. In seminal work Gowers [7, 8] introduced higher order Fourier analysis and proved the first "reasonable" bounds for Szemerédi's theorem:

$$r_k(N) \ll N(\log\log N)^{-c_k}.$$

The only previous improvement to this result for $k \geq 4$ was work on the case $k = 4$ of Green and Tao [13, 14] which ultimately established that

$$r_4(N) \ll N(\log N)^{-c}.$$

Our main result is a "quasi-logarithmic" bound in the case $k = 5$.

THEOREM 1·1. *There is* $c \in (0, 1)$ *such that*

$$r_5(N) \ll \frac{N}{\exp((\log\log N)^c)}.$$

Remark. Throughout this paper, we will abusively write log for $\max(\log(\cdot), e^e)$. This is to avoid trivial issues regarding inputs which are small.

Our work relies crucially on a recent improved inverse theorem for the Gowers U^4-norm due to the first author [23, theorem 7]. We first formally define the Gowers U^k-norm.

Definition 1·2. Given $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ and $k \geq 1$, we define

$$\|f\|_{U^k}^{2^k} = \mathbb{E}_{x,h_1,\dots,h_k \in \mathbb{Z}/N\mathbb{Z}} \Delta_{h_1,\dots,h_k} f(x)$$

where $\Delta_h f(x) = f(x)\overline{f(x+h)}$ is the multiplicative discrete derivative (extended to vectors h in the natural way). This is known to be well-defined and a norm (seminorm if $k = 1$).

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

THEOREM 1·3. *There exists an absolute constant $C \geq 1$ such that the following holds. Let N be prime, let $\delta > 0$, and suppose that $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is a 1-bounded function with*

$$\|f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \geq \delta.$$

There exists a degree 3 nilsequence $F(g(n)\Gamma)$ such that it has dimension bounded by $(\log(1/\delta))^C$, complexity bounded by C, such that F is 1-Lipschitz, and such that

$$\left|\mathbb{E}_{n\in[N]} f(n)\overline{F(g(n)\Gamma)}\right| \geq 1/\exp((\log(1/\delta))^C).$$

A key manoeuvre in this paper is our deduction of a variant of the U^4-inverse theorem which lends itself to the analysis of multiple nilsequences simultaneously and may be of independent interest. Although it is known that having a large U^4-norm does not necessarily imply direct correlation with a cubic phase function due to the existence of bracket polynomials, one can hope to achieve correlation on a dense, structured host set (for us, a Bohr set).

PROPOSITION 1·4. *There exists an absolute constant $C \geq 1$ such that the following holds. Let N be prime, let $\delta > 0$, and suppose that $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is a 1-bounded function with*

$$\|f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \geq \delta.$$

There exist $S \subseteq (1/N)\mathbb{Z}$ with $|S| \ll (\log(1/\delta))^C$, and $y \in \mathbb{Z}/N\mathbb{Z}$ such that the following holds: there is a locally cubic phase function $\phi : y + B(S, 1/100) \to \mathbb{R}/\mathbb{Z}$ such that

$$|\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}} \mathbb{1}_{t\in y+B(S,1/100)} f(t)e(-\phi(t))| \gg 1/\exp((\log(1/\delta))^C).$$

We refer the reader to Definitions 2·1 and 2·2 for precise definitions of Bohr set and locally cubic phase function. We remark that an analogous result to Proposition 2·3 for higher U^k-norms is false; this can be seen by examining the function $e(\{an\{bn\}\}\{cn\}dn)$. We discuss this issue in more detail in Section 1·1.

1·1. *Proof outline*

The starting point of our work involves combining the density increment strategy of Heath-Brown [**16**] and Szemerédi [**35**], which was reformulated in a robust manner by Green and Tao [**13**] when proving that $r_4(N) \ll N\exp(-\Omega(\sqrt{\log\log N}))$, with the improved U^4-inverse theorem Theorem 1·3 of the first author. The crucial difference between the density increment strategy of Roth [**28**] versus Heath-Brown [**16**] and Szemerédi [**35**] is that one finds correlations with "many functions" to deduce a density increment.

If we apply Theorem 1·3 directly with the density increment strategy as codified by Green and Tao [**13**], we would at the very least need that, given a polynomial sequence $g(n)$ with $g(0) = \mathrm{id}_G$ on a nilmanifold G/Γ of degree 3 with complexity M and dimension d, we have

$$\min_{1\leq n\leq N} d_{G/\Gamma}(\mathrm{id}_G, g(n)) \ll M^{O(d^{O(1)})}N^{-1/d^{O(1)}}.$$

When $g(n)$ is a polynomial sequence of degree 3 on the torus such results can be derived from work of Schmidt [**32**] on small fractional parts of polynomials. While directly deriving such a result for nonabelian nilmanifolds does not appear implausible, at present the distribution theory of nilmanifolds only has such "polynomial in d" dependencies when dealing with test functions having vertical frequency, due to work of the first author [**22**].

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

The crucial step in our work therefore is solving such a Schmidt-type problem for nilmanifolds via representing 3-step nilmanifolds as sums of exponentials of locally cubic functions on Bohr sets (see Proposition 2·3). The analogue of such a result for 2-step nilmanifolds (without bounds) appears in work of Green and Tao [**10**, proposition 2·3]. That such a result is true is a miracle of small numbers which is most easily seen via bracket polynomials. (In work of Green and Tao [**13**] regarding $r_4(N)$, one operates with a version of the U^3-inverse theorem proven directly for locally quadratic functions on Bohr sets.)

At an informal level, the U^4-inverse theorem may be rephrased as follows: if $\|f\|_{U^4}$ is large for 1-bounded f, then f correlates with a bracket exponential phase $e(H(n))$ where $H(n)$ is (essentially) a sum of terms of the form

$$an^3, an^2\{bn\}, an\{bn^2\}, an\{bn\}\{cn\}, an\{\{bn\}\{cn\}\}, \{an\}\{bn\}\{cn\}, an\{bn\{cn\}\} \bmod 1$$

plus terms which are obviously of "lower degree". By the work of Green and Tao, various lower order terms may be viewed as quadratic functions on Bohr sets. Now note that

$$\{x\}y + y\{x\} = xy - \lfloor x\rfloor \lfloor y\rfloor + \{x\}\{y\}.$$

Therefore

$$\{x\}y + y\{x\} = xy + \{x\}\{y\} \bmod 1.$$

Furthermore note that $e(\{x\}\{y\})$ is, after appropriate smoothing, a Lipschitz function on $(\mathbb{R}/\mathbb{Z})^2$ and therefore is well-approximated by a weighted sum of exponentials of the form $e(kx + \ell y)$ for $k, \ell \in \mathbb{Z}$. Given this (and noting the analogous fact for $e(\{x\}\{y\}\{z\})$) we may rewrite our basis of degree 3 functions as

$$an^3, an^2\{bn\}, an\{bn\}\{cn\} \bmod 1;$$

the most crucial of these manipulations is

$$an\{bn\{cn\}\} = abn^2\{cn\} - \{an\}bn\{cn\} \bmod 1.$$

The miracle is that we do not have any nested $\{\cdot\}$ and all of the brackets surround linear functions. Therefore, upon restricting the fractional parts to lie in certain narrow ranges away from the discontinuities in the fraction part (i.e., Bohr set-type conditions) the functions $an^2\{bn\}, an\{bn\}\{cn\}$ are seen to be "locally cubic". That is, discrete fourth-order derivatives vanish given that all points in the corresponding 4-dimensional hypercube lie in an appropriate Bohr set. The existence of such a miracle can be seen by examining carefully all "fractional part" expressions required in work of Green–Tao–Ziegler [**12**, appendix E] on the U^4-inverse theorem. For the U^5-inverse theorem and higher, we have the function $e(\{an\{bn\}\}\{cn\}dn)$ and fractional part identities do not allow one to "remove iterated floor functions".

To actually prove the desired representation of a step 3 nilsequence, we partition the nilmanifold via a partition of unity. Operating within a partition of unity, we may manipulate the nilmanifold (as in [**10**, proposition 2·3]) via explicitly operating in Mal'cev coordinates of the first kind. This allows us to manipulate the floor and fractional expressions as suggested by the bracket polynomial formalism in the previous paragraphs. Identifying various fundamental domains with the torus and applying Fourier expansion appropriately eventually gives the desired decomposition.

We remark for technical reasons it turns out to be useful to manipulate our nilsequence to be N-periodic and living on a nilmanifold for which the Lie bracket structure constants

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

are integers which are sufficiently divisible. The first task is accomplished via appropriately quantifying a construction of Manners [**24**] (see Proposition C·2) and the second is accomplished via a "clearing denominators" trick on the Mal'cev basis (see Lemma C·1). One can see that such a manipulation might be useful by noting that if the structure constants of G/Γ are sufficiently divisible then $\psi_{\exp}(\Gamma) = \mathbb{Z}^d$ where $\psi_{\exp}$ is the map to Mal'cev coordinates of the first kind (see Lemma 2·5). In general $\psi_{\exp}(\Gamma)$ is not even a lattice in $\mathbb{Q}^d$.

Finally, given a correlation with a locally cubic function on a Bohr set, we are now in position to run the density increment strategy of Heath-Brown [**16**] and Szemerédi [**35**] as codified by Green and Tao [**13**]. Our proof is very close to that of Green and Tao [**13**], although there are slight simplifications afforded in our situation since the dimension of our Bohr sets are "quasi-logarithmic". In particular, it is possible to operate by considering single atoms in the Bohr partition and avoid the use of regular Bohr sets. Our situation at this point is closely analogous to having access to the U^3-inverse theorem of Sanders [**29**] and aiming to prove bounds of the form $r_4(N) \ll N \exp(-(\log\log N)^c)$.

Notation. We use standard asymptotic notation. Given functions $f = f(n)$ and $g = g(n)$, we write $f = O(g), f \ll g, g = \Omega(f)$, or $g \gg f$ to mean that there is a constant C such that $|f(n)| \leq Cg(n)$ for sufficiently large n. We write $f \asymp g$ or $f = \Theta(g)$ to mean that $f \ll g$ and $g \ll f$, and write $f = o(g)$ or $g = \omega(f)$ to mean $f(n)/g(n) \to 0$ as $n \to \infty$. Subscripts on asymptotic notation indicate dependence of the bounds on those parameters. We will use the notation $[x] = \{1, 2 \ldots, \lfloor x \rfloor\}$.

We use the notations of Appendix A with regards to nilmanifolds. Write $\Delta_h f(x) = f(x)\overline{f(x+h)}$ for the multiplicative discete derivative and $\partial_h f(x) = f(x) - f(x+h)$ for the additive discrete derivative (for functions over appropriate domains and codomains). The Gowers U^s-norm on a finite abelian group G is then defined via

$$\|f\|_{U^s}^{2^s} := \mathbb{E}_{x,h_1,\ldots,h_s \in G} \Delta_{h_1,\ldots,h_s} f(x)$$

for $f\colon G \to \mathbb{C}$, which is known to be well-defined and a norm for $s \geq 2$ and a seminorm for $s = 1$.

1·2. *Organisation of paper*

In Section 2 we prove the main technical result regarding approximating nilsequences as local cubics on Bohr sets. In Section 3 we deduce the main result via a density increment argument. In Appendix A we collect various conventions and definitions regarding nilmanifolds. In Appendix B we essentially reproduce [**13**, appendix A] and note that it verbatim extends to higher degree polynomials. In Appendix C, we manipulate Theorem 1·3 to prove Theorem C·3 which gives correlation with a periodic nilsequence where the underlying nilmanifold has appropriately divisible Lie bracket structure constants. Finally, in Appendix D we collect a number of deferred technical lemmas.

2. *Converting nilsequences to local cubics on a Bohr set*

The key idea is to work with a presentation of our nilsequence coming from Theorem 1·3, and manipulate it into an approximate form composed of locally cubic functions on Bohr sets.

We will first define Bohr sets formally.

Definition 2·1. Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$ and $\rho \in (0, 1)$, we define the (centered) Bohr set

$$B(S, \rho) := \{x \in \mathbb{Z}/N\mathbb{Z} \colon \|\xi x/N\|_{\mathbb{R}/\mathbb{Z}} < \rho \text{ for all } \xi \in S\}.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Given $\alpha = (\alpha_\xi)_{\xi \in S} \in (\mathbb{R}/\mathbb{Z})^S$, we define the uncentered Bohr set

$$B_\alpha(S, \rho) := \{x \in \mathbb{Z}/N\mathbb{Z} \colon \|\xi x/N - a_\xi\|_{\mathbb{R}/\mathbb{Z}} < \rho \text{ for all } \xi \in S\}.$$

The parameter $|S|$ is the rank and ρ is the radius.

We next define the notion of local polynomial structure; we will care specifically about the case of degree $s = 3$.

Definition 2·2. Given $S \subseteq G$ and $f \colon S \to H$, we say that f is locally degree $s \geq 0$ if for all $x, h_1, \dots, h_{s+1}$ such that $x + \sum_{j=1}^{s+1} \epsilon_j h_j \in S$ for all $\epsilon = (\epsilon_1, \dots, \epsilon_{s+1}) \in \{0, 1\}^{s+1}$, we have

$$\partial_{h_1, \dots, h_{s+1}} f(x) = 0.$$

Remark. We will primarily be concerned with $S \subseteq \mathbb{Z}/N\mathbb{Z}$ and $H = \mathbb{R}/\mathbb{Z}$.

PROPOSITION 2·3. *Suppose we are given a degree 3 nilmanifold G/Γ of dimension d, complexity M, and all Lie bracket structure constants divisible by 12. Furthermore suppose that $F \colon G/\Gamma \to \mathbb{C}$ is smooth with Lipschitz norm bounded by 1 and let $\eta \in (0, 1/2)$.*

Then there exist data such that

$$\sup_{n \in \mathbb{Z}/N\mathbb{Z}} \left| F(g(n)\Gamma) - \sum_{i \in I} w_i \mathbb{1}_{n \in y_i + B(S, 1/100)} e(\phi_i(n)) \right| \leq \eta$$

such that:

(i) $S \subseteq (1/N)\mathbb{Z}$ *with size at most* $d + 1$;

(ii) I *is an index set of size at most* $O(M\eta^{-1})^{O(d^{O(1)})}$ *and* $|w_i| \leq O(M\eta^{-1})^{O(d^{O(1)})}$ *for all* $i \in I$;

(iii) $y_i \in \mathbb{Z}/N\mathbb{Z}$ *for all* $i \in I$;

(iv) $\phi_i \colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ *is locally cubic on* $y_i + B(S, 1/100)$.

Then, Proposition 1·4 directly follows from a slightly modified version of Theorem 1·3 (namely Theorem C·3) and Proposition 2·3. Theorem C·3 allows one to essentially assume that the underlying nilsequence is N-periodic and various structure constants are sufficiently divisible.

Proof of Proposition 1·4. By Theorem C·3, there exists a nilmanifold G/Γ with function F and a polynomial sequence $g(n)$ such that

$$\left| \mathbb{E}_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) \overline{F(g(n)\Gamma)} \right| \geq 1/\exp((\log(1/\delta))^C) =: 2\eta$$

with $g(0) = \mathrm{id}_G$ and G/Γ having dimension d, complexity at most M and all structure constants divisible by 12 with

$$d \leq (\log(1/\delta))^{O(1)}$$
$$M \leq \exp((\log(1/\delta))^{O(1)}).$$

We now approximate $F(g(n)\Gamma)$ by a function $H(n)$ such that $\sup_{n \in \mathbb{Z}/N\mathbb{Z}} |F(g(n)\Gamma) - H(n)| \leq \eta$ using Proposition 2·3. As f is 1-bounded, we immediately have

$$\left| \mathbb{E}_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) \overline{H(n)} \right| \geq \left| \mathbb{E}_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) \overline{F(g(n)\Gamma)} \right| - \mathbb{E}_{n \in \mathbb{Z}/N\mathbb{Z}} |H(n) - F(g(n)\Gamma)| \geq \eta.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

We have some additional guarantees on the structure of H (in particular, some associated data S, I, w_i, y_i, ϕ_i). Applying Pigeonhole on the index set I and using that the w_i are sufficiently bounded, there exist $y \in \mathbb{Z}/N\mathbb{Z}$, S a set of at most size $d+1$, $\rho = 1/100$, and $\phi\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ locally cubic on $y + B(S, \rho)$ such that

$$\left|\mathbb{E}_{n\in\mathbb{Z}/N\mathbb{Z}} f(n)\mathbb{1}_{n\in y+B(S,\rho)} e(-\phi(n))\right| \geq 1/\exp((\log(1/\delta))^{O(1)}).$$

In order to prove Proposition 2·3, we will need a number of quantitative and structural lemmas about nilsequences. These arguments are deferred to Appendix D. We first require the following quantitative partition of unity for nilmanifolds.

LEMMA 2·4. *Fix $\varepsilon \in (0, 1/2)$ and a nilmanifold G/Γ of degree s, dimension d, and complexity M. There exists an index set I and a collection of nonnegative smooth functions $\tau_j\colon G/\Gamma \to \mathbb{R}$ for $j \in I$ such that*:

(i) $L := (M/\varepsilon)^{O_s(d^{O_s(1)})}$;

(ii) for all $g \in G$, we have $\sum_{j\in I} \tau_j(g\Gamma) = 1$;

(iii) $|I| \leq L$;

(vi) for each $j \in I$, there exists $\beta \in [-L, L]^d$ so that for any $g\Gamma \in \operatorname{supp}(\tau_j)$ there exist $g' \in g\Gamma$ such that $\log_G(g') \in \prod_{i=1}^d [\beta_i - \varepsilon, \beta_i + \varepsilon]$;

(v) τ_j are L-Lipschitz on G/Γ.

We also require the algebraic fact that if the Lie bracket structure constants of G/Γ are sufficiently divisible then $\psi_{\exp}(\Gamma) = \mathbb{Z}^d$.

LEMMA 2·5. *There exists an integer $C_s \geq 1$ such that the following holds. Fix a nilmanifold G/Γ of degree s and a Mal'cev basis $\mathcal{X}$ such that all Lie bracket structure constants of G/Γ are divisible by C_s. Then $\psi_{\exp}(\Gamma) = \mathbb{Z}^d$.*

Remark. We may take $C_3 = 12$.

Next we need the conversion between the nilmanifold distance and distance in coordinates of the first kind (on $\mathbb{R}^d$).

LEMMA 2·6. *Fix a nilmanifold G/Γ of degree s, dimension d, and complexity M. If $\|\psi_{\exp}(x) - \psi_{\exp}(y)\|_\infty \leq \varepsilon$ and $\|\psi_{\exp}(x)\|_\infty \leq L$ then*

$$d_{G/\Gamma}(x\Gamma, y\Gamma) \leq \varepsilon(LM)^{O_s(d^{O_s(1)})}.$$

We also require the following basic estimate regarding Fourier expansion of Lipschitz functions on the torus; this follows immediately by quantifying the proof in [**36**, proposition 1·1·13] (see e.g. [**26**, lemma A·8]).

LEMMA 2·7. *Fix $0 < \varepsilon < 1/2$, and let $F\colon (\mathbb{R}/\mathbb{Z})^d \to \mathbb{C}$ with $\|F\|_{\mathrm{Lip}} \leq L$ with metric $d(x, y) = \max_{1\leq i\leq d} \|x_i - y_i\|_{\mathbb{R}/\mathbb{Z}}$ for $x, y \in (\mathbb{R}/\mathbb{Z})^d$. There exists an absolute constant $C > 0$ such that there exist c_ξ with $\sum_\xi |c_\xi| \leq (3CLd\varepsilon^{-1})^{5d}$ and*

$$\sup_{x\in(\mathbb{R}/\mathbb{Z})^d} \left|F(x) - \sum_{|\xi|\leq(CLd\varepsilon^{-1})^2} c_\xi e(\xi\cdot x)\right| \leq \varepsilon.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

We finally require the following elementary lemma regarding how local degree acts with respect to multiplication; this is once again deferred to Appendix D. (Recall we are defining local degree with respect to ∂ as opposed to Δ, and the following cannot be altered to only require that the functions reduced mod 1 have local degree.)

LEMMA 2·8. *If $S \subseteq G$ and $f_j \colon S \to \mathbb{C}$ are locally degree d_j functions, then $g = \prod_{j=1}^{k} f_k$ is a locally degree $\sum_{j=1}^{k} d_j$ function.*

We are ready to proceed.

Proof of Proposition 2·3. We break the proof into a series of steps.

Step 1. Polynomial sequence and group multiplication in coordinates. Note that as we are dealing with nilpotent groups of step at most 3 we have

$$\log(e^X e^Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] - [Y, [X, Y]])$$

by the Baker–Campbell–Hausdorff formula. Let $d_j = \dim G_1 - \dim G_{j+1}$ and let our Mal'cev basis be $(X_k)_{k\in[d]}$ such that $(X_k)_{k>d_j}$ spans $\log G_{j+1}$ for $j \in \{0, 1, 2\}$. Note that by definition Γ is the set of elements of the form $\prod_{i=1}^{d} \exp(t_i X_i)$ for $t_i \in \mathbb{Z}$. Recall that we have

$$[X_i, X_j] = \sum_{k=1}^{d} c_{ijk} X_k$$

for c_{ijk} integers divisible by 12 and of absolute value at most M due to the complexity assumption. As $[\log G_i, \log G_j] \subseteq \log G_{i+j}$, we have that $c_{ijk} = 0$ for $k \leq d_{a+b+1}$ if $i > d_a$ and $j > d_b$.

As $g(n)$ is a polynomial sequence with $g(0) = \mathrm{id}_G$, by Taylor expansion we have

$$g(n) = g_1^n g_2^{\binom{n}{2}} g_3^{\binom{n}{3}}$$

with $g_i \in G_i$. Therefore

$$\begin{aligned}\log(g(n)) &= \log\left(\exp(n\log(g_1))\exp\left(\binom{n}{2}\log(g_2)\right)\exp\left(\binom{n}{3}\log(g_3)\right)\right)\\ &= n\log(g_1) + \binom{n}{2}\log(g_2) + \frac{n}{2}\binom{n}{2}[\log(g_1), \log(g_2)] + \binom{n}{3}\log(g_3)\end{aligned}$$

with $[\log(g_1), \log(g_2)] \in \log(G_3)$. Therefore we have

$$g(n) = \exp\left(\sum_{j=1}^{d} p_j(n) X_j\right) \tag{2·1}$$

with $p_j(0) = 0$ and p_j being at most degree 1 if $1 \leq j \leq d_1$, at most degree 2 if $d_1 + 1 \leq j \leq d_2$, and at most degree 3 if $d_2 + 1 \leq j \leq d_3 = d$. We next realise multiplication of two elements corresponding to Mal'cev coordinates of the first kind as

$$\begin{aligned}&(x_1, \ldots, x_d) * (y_1, \ldots, y_d)\\ &= (x_1 + y_1, \ldots, x_{d_1} + y_{d_1}, x_{d_1+1} + y_{d_1+1} + \phi_{d_1+1}(x_{\leq d_1}, y_{\leq d_1}), \ldots, x_{d_2} + y_{d_2}\\ &+ \phi_{d_2}(x_{\leq d_1}, y_{\leq d_1}), x_{d_2+1} + y_{d_2+1} + \phi_{d_2+1}(x_{\leq d_2}, y_{\leq d_2}) + \varphi_{d_2+1}(x_{\leq d_1}, y_{\leq d_1}), \ldots, x_d + y_d\\ &+ \phi_d(x_{\leq d}, y_{\leq d}) + \varphi_d(x_{\leq d_1}, y_{\leq d_1}))\end{aligned}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

where ϕ_j are bilinear forms with *integral* coefficients and φ_j are cubic forms with *integral* coefficients. This is an immediate consequence of the Baker–Campbell–Hausdorff formula and using the assumption that the Lie bracket structure constants are all divisible by 12. Furthermore these coefficients are of height at most $O(Md)^{O(1)}$ and since $[G_2, G_2] \subseteq G_4 = \mathrm{Id}_G$, the bilinear form ϕ_j for $j \in [d_1+1, d_2]$ has all coordinates 0 on the box $[d_1+1, d_2]^2$.

Step 2. Explicit representation of nilsequence with coordinates. We can represent the coordinates of $g(n)\Gamma$ in a fundamental domain with respect to Mal'cev coordinates of the first kind via iterated floor and fractional parts. These coordinates are not smooth at the boundary and thus we decompose F via a partition of unity (and using different fundamental domains for each part). This will allow us to manipulate such coordinate functions without needing to worry very precisely about the minor discontinuities.

Let $L = O(M)^{O(d^{O(1)})}$. We write

$$F = \sum_{i \in I} F_i$$

with $F_i = F\tau_i$ where τ_i is as in Lemma 2·4 applied with parameter $\varepsilon = 10^{-3}$. This will allow us to represent $F\tau_i$ on the fundamental domain (with respect to Mal'cev coordinates of the first kind) $\prod_{j=1}^{d} [\beta_j^{(i)} - 1/2, \beta_j^{(i)} + 1/2)$, where $\beta^{(i)} \in [-L, L]^d$ and

$$\mathrm{supp}\,(F\tau_i) \subseteq \mathrm{supp}\,(\tau_i) \subseteq \prod_{j=1}^{d} [\beta_j^{(i)} - 10^{-3}, \beta_j^{(i)} + 10^{-3}).$$

We now define nonstandard (shifted) floor and fractional part functions for each $i \in I$ and $j \in [d]$ so that

$$\begin{aligned} \{t\}_{i,j} &\equiv t \bmod 1 \\ t &= \{t\}_{i,j} + \lfloor t \rfloor_{i,j} \\ \{t\}_{i,j} &\in [\beta_j^{(i)} - 1/2, \beta_j^{(i)} + 1/2). \end{aligned}$$

For nearly the entire remainder of the proof, we will focus on massaging the representation of $F_i(x)$ into a more convenient form. Given $x \in G$ such that $\psi_{\exp}(x) = (x_1, \dots, x_d)$, consider γ with Mal'cev coordinates of the first kind:

$$\begin{aligned} &(-\lfloor x_j \rfloor_{i,j})_{j \in [d_1]}, \quad (-\lfloor x_j + \phi_j(x_{\le d_1}, -\lfloor x_{\le d_1} \rfloor_i) \rfloor_{i,j})_{d_1+1 \le j \le d_2}, \\ &(-\lfloor x_j + \varphi_j(x_{\le d_1}, -\lfloor x_{\le d_1} \rfloor_i) + \phi_j(x_{\le d_2}, x^*_{\le d_2}) \rfloor_{i,j})_{d_2+1 \le j \le d_3}, \end{aligned} \tag{2·2}$$

where $x^*_{\le d_2}$ has coordinates equal to those on the first line of (2·2) (so it implicitly depends on i) and $\lfloor x_{\le d_1} \rfloor_i = (\lfloor x_1 \rfloor_{i,1}, \dots, \lfloor x_1 \rfloor_{i,d_1})$. Furthermore let $\{x_{\le d_1}\}_i = x - \lfloor x_{\le d_1} \rfloor_i$. Note that $\gamma \in \Gamma$ by Lemma 2·5.

Thus $x\gamma$ has coordinates

$$\begin{aligned} &(\{x_j\}_{i,j})_{j \in [d_1]}, \quad (\{x_j + \phi_j(x_{\le d_1}, -\lfloor x_{\le d_1} \rfloor_i)\}_{i,j})_{d_1+1 \le j \le d_2}, \\ &(\{x_j + \varphi_j(x_{\le d_1}, -\lfloor x_{\le d_1} \rfloor_i) + \phi_j(x_{\le d_2}, x^*_{\le d_2})\}_{i,j})_{d_2+1 \le j \le d} \end{aligned} \tag{2·3}$$

which is in the specified fundamental domain $\prod_{j=1}^{d} [\beta_j^{(i)} - 1/2, \beta_j^{(i)} + 1/2)$ for F_i. Recall also that ϕ_j for $j \in [d_2+1, d]$ has certain 0 coefficients.

Let $x^*_{[d_1+1, d_2]}$ denote the vector with d_2 coordinates, the first d_1 of which are zero and the last $d_2 - d_1$ of which are $(-\lfloor x_j + \phi_j(x_{\le d_1}, -\lfloor x_{\le d_1} \rfloor_i) \rfloor_{i,j})_{d_1+1 \le j \le d_2}$. Let $y_j = x_j +$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

$\phi_j(x_{\leq d_1}, -\lfloor x_{\leq d_1}\rfloor_i)$, $y^*_{[d_1+1,d_2]}$ analogously be a vector of these coordinates and d_1 many 0s, and let $\{y_{[d_1+1,d_2]}\}^*_i = (0,\dots,0,\{y_{d_1+1}\}_{i,d_1+1},\dots,\{y_{d_2}\}_{i,d_2})$.

Note that we have the identity

$$-y\lfloor z\rfloor_2 = \lfloor y\rfloor_1 z - yz - \lfloor y\rfloor_1\lfloor z\rfloor_2 + \{y\}_1\{z\}_2,$$

where the subscripts 1 and 2 indicate potentially different shift types. Therefore

$$\begin{aligned}\phi_j(x_{\leq d_2}, x^*_{\leq d_2}) &= \phi_j(x_{\leq d_2}, -\lfloor x_{\leq d_1}\rfloor_i) + \phi_j(x_{\leq d_2}, x^*_{[d_1+1,d_2]})\\ &= \phi_j(x_{\leq d_2}, -\lfloor x_{\leq d_1}\rfloor_i) + \phi_j(x_{\leq d_1}, x^*_{[d_1+1,d_2]})\\ &= \phi_j(x_{\leq d_2}, -\lfloor x_{\leq d_1}\rfloor_i) + \phi_j(\lfloor x_{\leq d_1}\rfloor_i, y^*_{[d_1+1,d_2]}) - \phi_j(x_{\leq d_1}, y^*_{[d_1+1,d_2]})\\ &\qquad + \phi_j(\lfloor x_{\leq d_1}\rfloor_i, x^*_{[d_1+1,d_2]}) + \phi_j(\{x_{\leq d_1}\}_i, \{y_{[d_1+1,d_2]}\}^*_i).\end{aligned}$$

The equality $\phi_j(x_{\leq d_2}, x^*_{[d_1+1,d_2]}) = \phi_j(x_{\leq d_1}, x^*_{[d_1+1,d_2]})$ follows as ϕ_j has no nonzero coordinates on the box $[d_1+1, d_2]^2$.

This implies that

$$\begin{aligned}&(\{x_j + \varphi_j(x_{\leq d_1}, -\lfloor x_{\leq d_1}\rfloor_i) + \phi_j(x_{\leq d_2}, x^*_{\leq d_2})\}_{i,j})_{d_2+1\leq j\leq d}\\ &= (\{x_j + \varphi_j(x_{\leq d_1}, -\lfloor x_{\leq d_1}\rfloor_i) + \phi_j(x_{\leq d_2}, -\lfloor x_{\leq d_1}\rfloor_i) + \phi_j(\lfloor x_{\leq d_1}\rfloor_i, y^*_{[d_1+1,d_2]})\\ &\quad - \phi_j(x_{\leq d_1}, y^*_{[d_1+1,d_2]}) + \phi_j(\{x_{\leq d_1}\}_i, \{y_{[d_1+1,d_2]}\}^*_i)\}_{i,j})_{d_2+1\leq j\leq d};\end{aligned}$$

we are able to drop $\phi_j(\lfloor x_{\leq d_1}\rfloor_i, x^*_{[d_1+1,d_2]})$ as ϕ_j has integral coefficients and we are taking fractional parts. Using the above general identity and this expression, we thus have coordinates of the form

$$\begin{aligned}&(\{x_j\}_{i,j})_{j\in[d_1]},\quad (\{x_j - \phi_j(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i)\}_{i,j})_{d_1+1\leq j\leq d_2},\\ &(\{x_j + \varphi^*_j(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i) + \phi^*_j(x_{\leq d_2}, \lfloor x_{\leq d_1}\rfloor_i)\\ &\quad + \sigma_j(x_{\leq d_2}, x_{\leq d_1}) + \phi_j(\{x_{\leq d_1}\}_i, \{y_{[d_1+1,d_2]}\}^*)\}_{i,j})_{d_2+1\leq j\leq d_3}.\end{aligned} \tag{2·4}$$

Here φ^*_j are degree at most 3 polynomials and ϕ^*_j, σ_j are bilinear forms. Additionally, all coefficients are integral with heights bounded by $(O(Md))^{O(1)}$.

Let us briefly take stock of what has been accomplished in the last manipulation. Note that there are no longer any "iterated" floor expressions and the only terms being "floored" are $x_{\leq d_1}$. The Bohr sets we will ultimately take therefore are determined by these coordinates only.

We now "lift" F_i from a function on G/Γ to $\widetilde{F}_i$ on $(\mathbb{R}/\mathbb{Z})^d$. This can be done via identifying $(\mathbb{R}/\mathbb{Z})^d$ with the fundamental domain $\prod_{j=1}^d [\beta^{(i)}_j - 1/2, \beta^{(i)}_j + 1/2)$ with respect to $\psi_{\exp}$. Now, $\widetilde{F}_i$ is seen to be $O(LM/\varepsilon)^{O(d^{O(1)})}$-Lipschitz on $(\mathbb{R}/\mathbb{Z})^d$ by Lemma 2·6. (We are also using that the support of $\widetilde{F}_i$ is close to the center of the torus, so the torus metric and ℓ^∞ on $\mathbb{R}^d$ are the same where it matters.)

Therefore it is sufficient to consider $\widetilde{F}_i$ with coordinates given by

$$\begin{aligned}&(x_j)_{j\in[d_1]},\quad (x_j - \phi_j(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i))_{d_1+1\leq j\leq d_2},\\ &(x_j + \varphi^*_j(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i) + \phi^*_j(x_{\leq d_2}, \lfloor x_{\leq d_1}\rfloor_i) + \sigma_j(x_{\leq d_2}, x_{\leq d_1})\\ &\quad + \phi_j(\{x_{\leq d_1}\}_i, \{y_{[d_1+1,d_2]}\}^*_i))_{d_2+1\leq j\leq d_3}\end{aligned} \tag{2·5}$$

which live on the torus.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

We now "smooth" $\{x_{\leq d_1}\}_i$ and $\{y_{[d_1+1,d_2]}\}_i^*$. We replace $\{\cdot\}_{i,j}$ with a 1-periodic function $\{\cdot\}_{i,j,\mathrm{sm}}$ which agrees with $\{x\}_{i,j}$ if $\{x\}_{i,j} \in [\beta_j^{(i)} - 1/4, \beta_j^{(i)} + 1/4)$ and is $O(1)$-Lipschitz when viewed as a function on $\mathbb{R}/\mathbb{Z}$. Given this we write $\{x_{\leq d_1}\}_{\mathrm{sm}}$ and $\{y_{[d_1+1,d_2]}\}_{\mathrm{sm}}^*$ for the associated vectors where smooth fractional parts have been used.

We claim that it suffices to consider $\widetilde{F}_i$ with coordinates given by

$$\begin{aligned}&(x_j)_{j\in[d_1]}, \quad (x_j - \phi_j(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i))_{d_1+1\leq j\leq d_2},\\ &(x_j + \varphi_j^*(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i) + \phi_j^*(x_{\leq d_2}, \lfloor x_{\leq d_1}\rfloor_i) + \sigma_j(x_{\leq d_2}, x_{\leq d_1})\\ &\quad + \phi_j(\{x_{\leq d_1}\}_{\mathrm{sm}}, \{y_{[d_1+1,d_2]}\}_{\mathrm{sm}}^*))_{d_2+1\leq j\leq d_3}.\end{aligned} \tag{2·6}$$

Note that if $\{x_{\leq d_1}\}_{\mathrm{sm}} \neq \{x_{\leq d_1}\}_i$ or $\{y_{[d_1+1,d_2]}\}_{\mathrm{sm}} \neq \{y_{[d_1+1,d_2]}\}_i^*$ we immediately see that one of the coordinates in the first two lines has been forced outside the support of $\widetilde{F}_i$ and thus the value is already by construction zero in both coordinates. Otherwise the coordinates in (2·5) and (2·6) match and the representation has been unchanged.

Step 3. Fourier expansion of nilsequence. Now, $\widetilde{F}_i\colon (\mathbb{R}/\mathbb{Z})^d \to \mathbb{C}$ is a Lipschitz function. Therefore by Lemma 2·7 with parameter τ,

$$\sup_{z\in(\mathbb{R}/\mathbb{Z})^d} \left|\widetilde{F}_i(z) - \sum_{|\xi|\leq O(M\tau^{-1})^{O(d^{O(1)})}} c_\xi e(\xi\cdot z)\right| \leq \tau$$

with $|c_\xi| \leq O(M\tau^{-1})^{O(d^{O(1)})}$.

Using this Fourier representation, we have

$$\begin{aligned}\sup_{x\in G}\Bigg|F_i(x\Gamma) - \sum_{|\xi|\leq O(M\tau^{-1})^{O(d^{O(1)})}} c_\xi e\Bigg(&\sum_{j=1}^d T_{\xi,j}x_j + \widetilde{\varphi}_\xi(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i) + \widetilde{\phi}_\xi(x_{\leq d_2}, \lfloor x_{\leq d_1}\rfloor_i)\\ &+ \widetilde{\sigma}_\xi(x_{\leq d_1}, x_{\leq d_2}) + \widetilde{\phi}'_\xi(\{x_{\leq d_1}\}_{\mathrm{sm}}, \{y_{[d_1+1,d_2]}\}_{\mathrm{sm}}^*)\Bigg)\Bigg| \leq \tau\end{aligned}$$

such that $\widetilde{\varphi}_\xi$ is at most a degree 3 polynomial, and $\widetilde{\phi}_\xi, \widetilde{\sigma}_\xi, \widetilde{\phi}'_\xi$ are bilinear forms. Furthermore all coefficients involved are integers bounded by $(M\tau^{-1})^{O(d^{O(1)})}$.

The final smoothing we perform before we specialise to the polynomial sequence under consideration is to remove the $\{\cdot\}_{\mathrm{sm}}\{\cdot\}_{\mathrm{sm}}$ terms. Note that the function

$$(x, y) \to e(T\{x\}_{i,j,\mathrm{sm}}\{y\}_{i,j,\mathrm{sm}})$$

is $O(TL^{O(1)})$-Lipschitz. We apply Lemma 2·7 with $\varepsilon = (M\tau^{-1})^{-O(d^{O(1)})}$ sufficiently small and replace each of the possible $d_1\cdot(d_2-d_1)$ possible combinations for the "smoothed parts" simultaneously in each term. We find that

$$\begin{aligned}&\sup_{x\in G}\Bigg|F_i(x\Gamma)\\ &-\sum_{k\in I} c_k e\Bigg(\sum_{j=1}^d T_{k,j}x_j + \widetilde{\varphi}_k(x_{\leq d_1}, \lfloor x_{\leq d_1}\rfloor_i) + \widetilde{\phi}_k(x_{\leq d_2}, \lfloor x_{\leq d_1}\rfloor_i) + \widetilde{\sigma}_k(x_{\leq d_2}, x_{\leq d_1})\Bigg)\Bigg| \leq 2\tau\end{aligned} \tag{2·7}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

with $|I| \leq (M\tau^{-1})^{O(d^{O(1)})}$, $|c_k| \leq (M\tau^{-1})^{O(d^{O(1)})}$, $\widetilde{\varphi_k}$ is an at most degree 3 polynomial, $\widetilde{\phi_k}$ and $\widetilde{\sigma_k}$ are bilinear forms, and all coefficients of these forms bounded by $(M\tau^{-1})^{O(d^{O(1)})}$. Explicitly, we used Lemma 2·7 on $(\mathbb{R}/\mathbb{Z})^2$ with parameter ε many times and multiplied the representations together.

We now specialise to the case of interest. Note that our primary concern is with the case where $x = g(n)$ and therefore $x_j = p_j(n)$ where $\deg p_j \leq i$ for $j \leq d_i$ if $i \in \{1, 2, 3\}$. Using the representation (2·7) and collecting terms we have

$$\sup_{n\in\mathbb{Z}} \left| F_i(g(n)\Gamma) - \sum_{k\in I} c_k e(H_k(n)) \right| \leq 2\tau,$$

where $H_k \colon \mathbb{Z} \to \mathbb{R}$ is a sum of terms of the form:

(i) $\alpha n^3 + \beta n^2 + \gamma n$;
(ii) $(\alpha n^2 + \beta n + \gamma)\lfloor p_j(n)\rfloor_{i,j}$ for $1 \leq j \leq d_1$;
(iii) $(\alpha n + \beta)\lfloor p_j(n)\rfloor_{i,j}\lfloor p_k(n)\rfloor_{i,k}$ for $1 \leq j \leq k \leq d_1$;
(vi) $\gamma \lfloor p_j(n)\rfloor_{i,j}\lfloor p_k(n)\rfloor_{i,k}\lfloor p_\ell(n)\rfloor_{i,\ell}$ for $1 \leq j \leq k \leq \ell \leq d_1$;

note that one can always collect terms so that there are at most $d^{O(1)}$ terms in each expression.

Step 4. N-periodicity and introducing Bohr sets. Let $\rho = 1/100$, which will be the radius of our Bohr sets. Now, note that the expressions $e(H_k(n))$ are not "obviously" N-periodic. We will artificially make them so via a partition of unity argument. There exist smooth $\chi_1, \ldots, \chi_{10^3}$ such that $\chi_j \colon (\mathbb{R}/\mathbb{Z}) \to \mathbb{R}$ are nonnegative with $\operatorname{supp}(\chi_j) \subseteq [j/10^3, (j+2)/10^3) \bmod 1$ and $\sum_{j=1}^{10^3} \chi_j = 1$. We have

$$F_i(g(n)\Gamma) = \sum_{j\in[10^3]} \chi_j(n/N) F_i(g(n)\Gamma),$$

$$\sup_{\substack{n\in\mathbb{Z}\\ j\in[10^3]}} \left| F_i(g(n)\Gamma)\chi_j(n/N) - \chi_j(n/N) \sum_{k\in I} c_k e(H_k(n)) \right| \leq O(\tau).$$

Now fix $j \in [10^3]$; such a term only contributes when $n/N \in [j/10^3, (j+2)/10^3) \bmod 1$. Next note that each $p_\ell(n) = \alpha_\ell n$ for $\ell \leq d_1$ and because the initial nilsequence $g(n)\Gamma$ is N-periodic, by [**23**, lemma A·12] we have that $\alpha_\ell \in (1/N)\mathbb{Z}$.

Note that in order for $\chi_j(n/N)F_i(g(n)\Gamma)$ to be nonzero, we need that:

(i) $n/N \equiv [j/10^3, (j+2)/10^3) \bmod 1$;
(ii) $\{p_\ell(n)\}_{i,\ell} \in [\beta_\ell^{(i)} - 10^{-3}, \beta_\ell^{(i)} + 10^{-3})$ for all $\ell \in [d_1]$.

These conditions are clearly N-periodic and in fact all such n lie in a shifted Bohr set with frequency set $S = \{1/N\} \cup \{\alpha_\ell\}_{1\leq\ell\leq d_1} \subseteq (1/N)\mathbb{Z}$. If the corresponding shifted Bohr set is empty we know that $\chi_j(n/N)F_i(g(n)\Gamma) = 0$ and we approximate $\chi_j(n/N)F_i(g(n)\Gamma)$ by 0. Else there is $y_j \in \mathbb{Z}/N\mathbb{Z}$ such that $y_j/N \equiv [j/10^3, (j+2)/10^3) \bmod 1$ and $\{\alpha_\ell y_j\}_{i,\ell} \in [\beta_\ell^{(i)} - 10^{-3}, \beta_\ell^{(i)} + 10^{-3})$ for $\ell \in [d_1]$. Letting J_i be the set of j for which this shifted Bohr set is

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

nonempty, we have

$$\sup_{\substack{n\in\mathbb{Z}\\ j\in J_i}}\left|F_i(g(n)\Gamma)\chi_j(n/N)-\mathbb{1}_{n\in y_j+B(S,\rho)}\chi_j(n/N)\sum_{k\in I}c_k e(H_k(n))\right|\le O(\tau).$$

We now apply Fourier expansion on χ_j using Lemma 2·7 on the torus $\mathbb{R}/\mathbb{Z}$ and appropriately chosen error parameter $\tau'=(M\tau^{-1})^{-O(d^{O(1)})}$. We find

$$\sup_{\substack{n\in\mathbb{Z}\\ j\in J_i}}\left|F_i(g(n)\Gamma)\chi_j(n/N)-\mathbb{1}_{n\in y_j+B(S,\rho)}\chi_j(n/N)\sum_{k\in I,|\xi|\le(\tau')^{-O(1)}}c_k d_\xi e(H_k(n)+\xi n/N)\right|\le O(\tau) \tag{2·8}$$

with $|d_\xi|\le(\tau')^{-O(1)}$.

Given $n\in y_j+B(S,\rho)$, we will now replace $H_k(n)$ by $H_{k,j}(n)$ so that the latter is N-periodic (and it will be locally cubic on $y_j+B(S,\rho)$). We fix an interval $I_j\in[-N,N)$ of integers of length at most $\lceil N/50\rceil$ such that for each $n\in y_j+B(S,\rho)$ there is a unique integer $n'\in I_j$ such that $n'\equiv n \bmod N$. This exists since $1/N\in S$. We then define $H_{k,j}(n)\equiv H_k(n') \mod 1$ for such $n\in y_j+B(S,\rho)$ and $H_{k,j}(n)=0$ otherwise; note that $H_{k,j}(n)$ is taking values in $\mathbb{R}/\mathbb{Z}$ which is sufficient as we are plugging these values into the exponential function.

It is easy to see that (2·8) holds with $H_k(n)+\xi\cdot n/N$ replaced by $H_{k,j}(n)$ since everything in the inequality except for potentially $H_k(n)$ is N-periodic (recall $S\subseteq(1/N)\mathbb{Z}$), and there is a supremum over $n\in\mathbb{Z}$ on the outside. So, we have

$$\sup_{\substack{n\in\mathbb{Z}\\ j\in J_i}}\left|F_i(g(n)\Gamma)\chi_j(n/N)-\mathbb{1}_{n\in y_j+B(S,\rho)}\sum_{k\in I,|\xi|\le(\tau')^{-O(1)}}c_k d_\xi e(H_{k,j}(n)+\xi n/N)\right|\le O(\tau),$$

where $H_{k,j}(n)$ is N-periodic by construction. Taking $\tau=\eta M^{-O(d^{O(1)})}$ and summing over j and i then finishes the proof modulo showing that $H_{k,j}(n)$ is a locally cubic function on $y_j+B(S,\rho)$.

Step 5. Local cubicity on shifted Bohr sets. Fix $i,j\in J_i$, and $k\in I$. We wish to show local cubicity of $H_{k,j}$ on $y_j+B(S,\rho)$. Suppose that $n,h_1,h_2,h_3\in\mathbb{Z}/N\mathbb{Z}$ are such that $n+\sum_{\ell=1}^3\epsilon_\ell h_\ell\in y_j+B(S,\rho)$ for all $(\epsilon_1,\epsilon_2,\epsilon_3)\in\{0,1\}^3$. We consider the representatives of $n+\sum_{\ell=1}^3\epsilon_\ell h_\ell$ in I_j; we see that there exist $n',{h_1}',{h_2}',{h_3}'\in\mathbb{Z}$ such that $n'+\sum_{\ell=1}^3\epsilon_\ell {h_\ell}'\equiv n+\sum_{\ell=1}^3\epsilon_\ell h_\ell \bmod N$ and $n'+\sum_{\ell=1}^3\epsilon_\ell {h_\ell}'\in I_j$ for all $\epsilon\in\{0,1\}^3$. Indeed, one can look at the representatives of each $n+\sum_{\ell=1}^3\epsilon_\ell h_\ell$ in I_j, and use the fact that if $t_1+t_2\equiv t_3+t_4 \bmod N$ with $t_i\in I_j$ then $t_1+t_2=t_3+t_4$ (since I_j has length at most $\lceil N/50\rceil$).

Therefore it suffices to prove that $H_k\colon\mathbb{Z}\to\mathbb{R}$ is locally cubic on $y_j+B(S,\rho)$. Recalling the classification of $H_k(n)$ as a sum of terms of various kinds, it suffices to verify this for each individual function type which is combined to give $H_k(n)$.

Note that pure polynomials are always the correct degree (on all of $\mathbb{Z}$). By Lemma 2·8 it suffices to verify that $\lfloor\alpha_\ell n\rfloor_{i,\ell}$ is locally linear on $y_j+B(S,\rho)$ for $1\le\ell\le d_1$.

Given $n,n+h_1,n+h_2,n+h_1+h_2\in y_j+B(S,\rho)$ we have that

$$\begin{aligned}&\lfloor\alpha_\ell n\rfloor_{i,\ell}-\lfloor\alpha_\ell(n+h_1)\rfloor_{i,\ell}-\lfloor\alpha_\ell(n+h_2)\rfloor_{i,\ell}+\lfloor\alpha_\ell(n+h_1+h_2)\rfloor_{i,\ell}\in\mathbb{Z},\\ &\left|\lfloor\alpha_\ell n\rfloor_{i,\ell}-\lfloor\alpha_\ell(n+h_1)\rfloor_{i,\ell}-\lfloor\alpha_\ell(n+h_2)\rfloor_{i,\ell}+\lfloor\alpha_\ell(n+h_1+h_2)\rfloor_{i,\ell}\right|\\ &=\left|\{\alpha_\ell n\}_{i,\ell}-\{\alpha_\ell(n+h_1)\}_{i,\ell}-\{\alpha_\ell(n+h_2)\}_{i,\ell}+\{\alpha_\ell(n+h_1+h_2)\}_{i,\ell}\right|\le 1/2,\end{aligned}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

the inequality using that $\rho = 1/100$ and every $\alpha_\ell(n + \epsilon_1 h_1 + \epsilon_2 h_2)$ must be close in $\mathbb{R}/\mathbb{Z}$ to $\alpha_\ell y_j$. Additionally, we recall $\{\alpha_\ell y_j\} \in [\beta_\ell^{(i)} - 10^{-3}, \beta_\ell^{(i)} + 10^{-3})$ which implies that there is no discontinuity in $\{\cdot\}_{i,\ell}$ in the area of consideration.

Therefore

$$\lfloor \alpha_\ell n \rfloor_{i,\ell} - \lfloor \alpha_\ell (n + h_1) \rfloor_{i,\ell} - \lfloor \alpha_\ell (n + h_2) \rfloor_{i,\ell} + \lfloor \alpha_\ell (n + h_1 + h_2) \rfloor_{i,\ell} = 0$$

for $n, n + h_1, n + h_2, n + h_1 + h_2 \in y_j + B(S, \rho)$ as desired. This (finally) completes the proof.

3. *Proof of Theorem 1·1 given Proposition 1·4*

In this section, we convert Proposition 1·4 into a density increment using the strategy of Heath-Brown [16] and Szemerédi [35]. Our treatment is closely modeled on that of Green and Tao [13]; the crucial idea is that one may group together a large number of phases before passing to a subprogression.

Throughout this section we will consider functions $f\colon [N'] \to [-1, 1]$ corresponding to sets and shifted indicators. By Bertrand's postulate, we may find a prime N such that $1024N' \le N \le 2048N'$. We may thus embed $[N']$ inside $\mathbb{Z}/(N\mathbb{Z})$ and lift f to $\mathbb{Z}/N\mathbb{Z}$ (mapping inputs not congruent to an element of $[N']$ to 0). We define the quintilinear operator

$$\Lambda(f_1, \ldots, f_5) = \mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} \prod_{k=1}^{5} f_k(x + (k-1)y) \text{ and } \Lambda(f) = \Lambda(f, f, f, f, f).$$

We now state the key claim for this section, which we will prove in Section 3.3

PROPOSITION 3·1. *Fix a constant $c > 0$. There exist $c' > 0$ and $C > 0$ such that the following holds. Consider a function $f\colon [N'] \to [0, 1]$ such that $\mathbb{E}_{n \in [N']} f(n) = \delta > 0$ and a prime N such that $1024N' \le N \le 2048N'$. Let $M(\delta) = \exp((\log(1/\delta))^C)$. At least one of the following possibilities always holds*:

(i) $N' \le \exp(M(\delta))$;

(ii) $\left|\Lambda(f) - \Lambda(\delta \mathbb{1}_{[N']})\right| \le c\delta^5$;

(iii) there exists an arithmetic progression P of length at least $N^{1/M(\delta)}$ such that

$$\mathbb{E}_{n \in P} f(n) \ge (1 + c')\delta.$$

With this, we can prove the main result.

Proof of Theorem 1·1 given Proposition 3·1. Suppose that $A \subseteq [N]$ has size δN and has no 5-term arithmetic progressions. We now perform the density increment strategy using $A_0 = A$, $N_0' = N$, and $\delta_0 = \delta$.

For each N_i' choose a prime N_i between $1024N_i' \le N_i \le 2048N_i'$. If $N_i' \le M(\delta_i)$, we immediately terminate. Else note that

$$\left|\Lambda(\mathbb{1}_{A_i}) - \Lambda(\delta_i \mathbb{1}_{[N_i']})\right| \ge \left|\Lambda(\delta_i \mathbb{1}_{[N_i']})\right| - |A_i| N_i^{-2} \gg \delta_i^5$$

as A_i is free of 5-term arithmetic progressions. Therefore the third case in the trichotomy must hold and there exists a long progression P_i on which the density of A_i increases by a

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

factor of at least $(1+c')$. Let A_{i+1} be $A_i \cap P_i$ rescaled so that P_i starts at 0 and has common difference 1, let ${N_{i+1}}' = |P_i|$, and let $\delta_{i+1} = |A_i \cap P_i|/|P_i| \geq (1+c')\delta_i$.

Since the density of a set cannot exceed 1, we must terminate in at most $O(\log(1/\delta))$ iterations. If we terminate at i, we must have

$$N^{(1/M(\delta))^{O(\log(1/\delta))}} \leq {N_i}' \leq \exp(M(\delta_i)) \leq \exp(M(\delta)).$$

Rearranging this gives exactly that

$$\log N \leq M(\delta)^{O(\log(1/\delta))}$$

or, as desired,

$$\delta \ll \frac{1}{\exp((\log\log N)^{\Omega(1)})}.$$

3·1. *Inverse theorem relative to linear and cubic factors*

We introduce a framework for studying functions satisfying a correlation as given by Proposition 1·4, considering a σ-algebra (or *factor*) which incorporates the information of the approximate value of our linear and cubic functions on $[N]$. Our treatment closely follows that of Green and Tao [13] (and uses elements from Peluse and Prendiville [25]).

Definition 3·2. We define a *factor* $\mathcal{B}$ of $\mathbb{Z}/N\mathbb{Z}$ to be a partition $\mathbb{Z}/N\mathbb{Z} = \bigsqcup_{B\in\mathcal{B}} B$. We define $\mathcal{B}(x)$ for $x \in \mathbb{Z}/N\mathbb{Z}$ to be the part of $\mathcal{B}$ that contains x.

We say $\mathcal{B}'$ refines $\mathcal{B}$ if every part of $\mathcal{B}$ can be written as a disjoint union of parts of $\mathcal{B}'$. We define a *join* of a sequence of factors to be the partition (discarding empty parts)

$$\mathcal{B}_1 \vee \cdots \vee \mathcal{B}_d := \{B_1 \cap \cdots \cap B_d \colon B_i \in \mathcal{B}_i\}.$$

Define a function $\phi\colon S \to \mathbb{R}/\mathbb{Z}$ for $S \subseteq \mathbb{Z}/N\mathbb{Z}$ to be *irrational* if ϕ takes on irrational values. Define the factor $\mathcal{B}_{\phi,K}$ with respect to ϕ of resolution K as the partition

$$((\mathbb{Z}/N\mathbb{Z}) \setminus S) \sqcup \bigsqcup_{0\leq j\leq K-1} \{x \in \mathbb{Z}/N\mathbb{Z} \colon \|\phi(x) - j/K\|_{\mathbb{R}/\mathbb{Z}} \leq 1/(2K)\}.$$

(Since ϕ is irrational this is indeed a disjoint partition.) We further say that ϕ respects a factor $\mathcal{B}$ if $\mathcal{B}$ refines $S \sqcup ((\mathbb{Z}/N\mathbb{Z}) \setminus S)$.

We define a factor of complexity (d_1, d_2) and resolution K via the data of irrational linear functions $\phi_1, \ldots, \phi_{d_1}$ (defined on $\mathbb{Z}/N\mathbb{Z}$) and irrational locally cubic functions ${\phi_1}', \ldots, {\phi_{d_2}}'$ (defined on subsets of $\mathbb{Z}/N\mathbb{Z}$) which respect $\bigvee_{1\leq j\leq d_1} \mathcal{B}_{\phi_j,K}$. The associated factor is $\bigvee_{1\leq j\leq d_1} \mathcal{B}_{\phi_j,K} \vee \bigvee_{1\leq j\leq d_1} \mathcal{B}_{{\phi_j}',K}$.

Finally, given a factor $\mathcal{B}$ we define $\Pi_{\mathcal{B}} f$ by

$$\Pi_{\mathcal{B}} f(x) = \mathbb{E}_{y\in\mathcal{B}(x)} f(y).$$

Remark. We ensure all functions ϕ we consider are irrational in order to avoid issues regarding the hitting exactly the boundary of the factor. Note that if a locally cubic function ϕ respects a partition $\mathcal{B}$ then we see that the support set is in the σ-algebra generated by $\mathcal{B}$.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

In particular, we can treat ϕ as either "undefined" on an atom $\mathcal{B}$ or as locally cubic on the associated atom.

We now restate Proposition 1·4 in the language of factors.

LEMMA 3·3. *There exists $C > 0$ such that the following holds. Fix $\eta > 0$ and let f be a function $f\colon \mathbb{Z}/N\mathbb{Z} \to [0,1]$ such that $\|f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \geq \eta$.*

Let $M(\eta) = \exp((\log(1/\eta))^C)$ and fix an integer $K \geq \exp(M(\eta))$. Suppose that $N \geq 8K$. There exist $d \leq M(\eta)$ and a factor $\mathcal{B}$ of complexity (d,1) and resolution K such that

$$\|\Pi_{\mathcal{B}} f\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \geq \exp(-M(\eta)).$$

Proof. By Proposition 1·4, there exist $\rho = 1/100$, $S \subseteq (1/N)\mathbb{Z}$ with $|S| \ll (\log(1/\eta))^C$ (say $|S| = d$), and $y \in \mathbb{Z}/N\mathbb{Z}$ such that $B = B(S,\rho)$ is a Bohr set and $\phi\colon y + B \to \mathbb{R}/\mathbb{Z}$ is a locally cubic phase function such that

$$|\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}} \mathbb{1}_{t\in y+B} f(t) e(-\phi(t))| \geq \exp(-(\log(1/\eta))^C/2).$$

Let C' be a sufficiently large constant and assume $K \geq \exp((\log(1/\eta))^{C'})$. Take $\phi_i\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ to be $\phi_i(x) = a_i x - a_i y + \alpha$ where α is an irrational which is smaller than $(2K)^{-1}$ and $a_i \in S$.

Note that $y + B$ is exactly the set

$$T_1 = \left\{x \in \mathbb{Z}/N\mathbb{Z}\colon \sup_{a_i\in S} \|a_i x - a_i y\|_{\mathbb{R}/\mathbb{Z}} \leq \rho\right\}.$$

Note that the set

$$T_2 = \left\{x \in \mathbb{Z}/N\mathbb{Z}\colon \sup_{a_i\in S} \|a_i x - a_i y + \alpha\|_{\mathbb{R}/\mathbb{Z}} \leq (2K)^{-1}\cdot(2\lfloor K\rho\rfloor - 1)\right\}$$

is measurable with respect to the factor $\bigvee_{1\leq j\leq d} \mathcal{B}_{\phi_j,K}$. Also, $T_2 \subseteq T_1$ since $|\alpha| \leq 1/(2K)$ and $(2K)^{-1}\cdot(2\lfloor K\rho\rfloor - 1) + (2K)^{-1} \leq \rho$. Furthermore we have

$$T_1 \setminus T_2 \subseteq \left\{x \in \mathbb{Z}/N\mathbb{Z}\colon \sup_{a_i\in S} \|a_i x - a_i y\|_{\mathbb{R}/\mathbb{Z}} \in [\rho - 2/K, \rho)\right\}$$

and thus we have

$$|T_2 \setminus T_1| \leq |S|\cdot(5/K)N = 5dN/K$$

as long as $8K \leq N$. Here we have used that $S \subseteq \mathbb{Z}/N\mathbb{Z}$ where N is prime.

So, since C' is large with respect to C and f is 1-bounded we have

$$|\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}} \mathbb{1}_{t\in T_2} f(t) e(-\phi(t))| \gg \exp(-(\log(1/\eta))^C/2).$$

The locally cubic function we will consider is $\phi^*\colon T_2 \to \mathbb{R}/\mathbb{Z}$ given by $\phi^*(x) = \phi(x) + \alpha$. This is well-defined as $T_2 \subseteq T_1 = y + B$. Let $\mathcal{B} = \bigvee_{1\leq j\leq d} \mathcal{B}_{\phi_j,K} \vee \mathcal{B}_{\phi^*,K}$ and note that

$$\sup_{x\in\mathbb{Z}/N\mathbb{Z}} \left|e(-\phi(x)) - \Pi_{\mathcal{B}}[e(-\phi^*(x))]\right| \leq 2\pi/K$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

since $z \mapsto e(z)$ is a 2π-Lipschitz function on $\mathbb{R}/\mathbb{Z}$. Therefore

$$|\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}} f(t)\Pi_{\mathcal{B}}[e(-\phi^*(x))]| \geq \exp(-(\log(1/\eta))^C).$$

$\Pi_{\mathcal{B}}$ is self-adjoint with respect to the standard inner product (see e.g. [**25**, lemma 4·3(ii)]) so

$$\begin{aligned}\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}}|\Pi_{\mathcal{B}}[f(t)]| &\geq |\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}}\Pi_{\mathcal{B}}[f(t)]e(-\phi^*(t))| = |\mathbb{E}_{t\in\mathbb{Z}/N\mathbb{Z}} f(t)\Pi_{\mathcal{B}}[e(-\phi^*(t))]| \\ &\geq \exp(-(\log(1/\eta))^C).\end{aligned}$$

This gives the desired result.

We now prove an associated Koopman–von Neumann theorem; our proof is closely modeled after [**13**, theorem 5·6].

LEMMA 3·4. *There exists $C > 0$ such that the following holds. Fix $\eta > 0$, let f be a function $f\colon \mathbb{Z}/N\mathbb{Z} \to [0,1]$, and let $\mathcal{B}^*$ denote an initial factor.*

Let $M(\eta) = \exp(\log(1/\eta)^C)$ and let $N \geq 10M(\eta)$. There exist $d_1, d_2 \leq M(\eta)$ such that the following holds. There exists an integer $K \leq M(\eta)$ and a factor of $\mathcal{B}'$ of complexity (d_1, d_2) and resolution K such that if $\mathcal{B} = \mathcal{B}^* \vee \mathcal{B}'$ then

$$\|f - \Pi_{\mathcal{B}} f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \leq \eta.$$

Proof. We create the factor $\mathcal{B}'$ as a join given by applying Lemma 3·3 iteratively. Set $\mathcal{B}_0 = \mathcal{B}^*$.

(i) If $\|f - \Pi_{\mathcal{B}_i} f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \leq \eta$, we terminate. Set $i^* = i$ and output $\mathcal{B}' = \bigvee_{0\leq i'<i^*} \mathcal{B}_{i'}{}'$.

(ii) If $\|f - \Pi_{\mathcal{B}_i} f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \geq \eta$, then set $K = \lceil \exp(M(\eta))\rceil$. By Lemma 3·3 there exists $\mathcal{B}_i{}'$ such that

$$\|\Pi_{\mathcal{B}_i{}'}(f - \Pi_{\mathcal{B}_i} f)\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \geq \exp(-(\log(1/\eta))^{C'}).$$

Here $\mathcal{B}_i{}' = \bigvee_{1\leq j\leq d_i} \mathcal{B}_{\phi_j^{(i)},K} \vee \mathcal{B}_{\phi_i{}',K}$ where $\phi_j^{(i)}$ (for $j \leq d_i \leq \exp((\log(1/\eta))^{O(1)})$) are irrational linear functions on $\mathbb{Z}/N\mathbb{Z}$ and $\phi_i{}'$ is a locally cubic function respecting the factor $\bigvee_{1\leq j\leq d} \mathcal{B}_{\phi_j,K}$. We set $\mathcal{B}_{i+1} = \mathcal{B}_i \vee \mathcal{B}_i{}'$.

Note that $\mathcal{B} = \mathcal{B}^* \vee \mathcal{B}' = \mathcal{B}_{i^*}$. Observe that

$$\begin{aligned}\|\Pi_{\mathcal{B}_i{}'}(f - \Pi_{\mathcal{B}_i} f)\|_{L^1(\mathbb{Z}/N\mathbb{Z})} &\leq \|\Pi_{\mathcal{B}_i{}'}(f - \Pi_{\mathcal{B}_i} f)\|_{L^2(\mathbb{Z}/N\mathbb{Z})} = \|\Pi_{\mathcal{B}_i{}'}\Pi_{\mathcal{B}_{i+1}}(f - \Pi_{\mathcal{B}_i} f)\|_{L^2(\mathbb{Z}/N\mathbb{Z})} \\ &\leq \|\Pi_{\mathcal{B}_{i+1}}(f - \Pi_{\mathcal{B}_i} f)\|_{L^2(\mathbb{Z}/N\mathbb{Z})} = \|\Pi_{\mathcal{B}_{i+1}} f - \Pi_{\mathcal{B}_i} f\|_{L^2(\mathbb{Z}/N\mathbb{Z})} \\ &= (\|\Pi_{\mathcal{B}_{i+1}} f\|^2_{L^2(\mathbb{Z}/N\mathbb{Z})} - \|\Pi_{\mathcal{B}_i} f\|^2_{L^2(\mathbb{Z}/N\mathbb{Z})})^{1/2}.\end{aligned}$$

The final equality is the Pythagorean theorem with respect to projections (this follows from e.g. [**25**, lemma 4·3(iv)]).

Thus $\|\Pi_{\mathcal{B}_{i+1}} f\|^2_{L^2(\mathbb{Z}/N\mathbb{Z})} - \|\Pi_{\mathcal{B}_i} f\|^2_{L^2(\mathbb{Z}/N\mathbb{Z})} \geq \exp(-(\log(1/\eta))^{O(1)})$ and hence the iteration lasts only $\exp((\log(1/\eta))^{O(1)})$ steps. Therefore the final $\mathcal{B}'$ is generated by at most $\exp((\log(1/\eta))^{O(1)})$ linear functions and a similar number of locally cubic functions. This completes the proof.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

3·2. *Density increment onto cubic Bohr set*

We next need that Λ is controlled by the U^4-norm and the L^1-norm. The proof is by now standard and hence is omitted (see [13, lemma 3·2] and [9, theorem 3·2]).

LEMMA 3·5. *Let N be prime and $f_i\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$. Then we have*:

$$|\Lambda(f_1,f_2,f_3,f_4,f_5)| \leq \min_{1\leq i\leq 5}\|f_i\|_{L^1(\mathbb{Z}/N\mathbb{Z})}\prod_{j\neq i}\|f_j\|_{L^\infty(\mathbb{Z}/N\mathbb{Z})};$$

$$|\Lambda(f_1,f_2,f_3,f_4,f_5)| \leq \min_{1\leq i\leq 5}\|f_i\|_{U^4(\mathbb{Z}/N\mathbb{Z})}\prod_{j\neq i}\|f_j\|_{L^\infty(\mathbb{Z}/N\mathbb{Z})}.$$

Given this we may now prove that there exists a density increment onto a piece of the partition given by Lemma 3·4. The proof is identical to [13, lemma 5·8].

PROPOSITION 3·6. *Fix $c>0$. There exist $C>0$ and $c'>0$ such that the following holds. Given a function $f\colon [N'] \to [0,1]$ such that $\mathbb{E}_{n\in[N']}f(n)=\delta>0$, extend it by 0s to a function on $\mathbb{Z}/N\mathbb{Z}$ where N is prime with $1024N' \leq N \leq 2048N'$. Suppose that*

$$|\Lambda(f)-\Lambda(\delta\mathbb{1}_{[N']})| \geq c\delta^5.$$

Let $\mathcal{B}^* = [N'] \sqcup ((\mathbb{Z}/N\mathbb{Z})\setminus[N'])$ be an initial factor. Let $M(\delta)=\exp((\log(1/\delta))^C)$ and suppose that $N \geq M(\delta)$. Then there exist $d_1,d_2,K\leq M(\delta)$, a factor $\mathcal{B}$ of complexity (d_1,d_2) and resolution K, and an atom Ω^* of $\mathcal{B}\vee\mathcal{B}^*$ such that:

(i) $\mathbb{P}_{n\in\mathbb{Z}/N\mathbb{Z}}[n\in\Omega^*] \geq \exp(-\exp((\log(1/\delta))^C))$;

(ii) $\mathbb{E}[f(n)|n\in\Omega^*] \geq (1+c')\delta$.

Proof. We may clearly assume δ is sufficiently small. By Lemma 3·4, there exists a factor $\mathcal{B}$ of complexity (d_1,d_2) and resolution K with $d_1,d_2,K\leq\exp((\log(1/\delta))^{O(1)})$ such that

$$\|f-\Pi_{\mathcal{B}\vee\mathcal{B}^*}f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \leq c\delta^5/10.$$

By telescoping and the second part of Lemma 3·5 we have

$$\left|\Lambda(f)-\Lambda(\Pi_{\mathcal{B}\vee\mathcal{B}^*}f)\right| \leq 5\|f-\Pi_{\mathcal{B}\vee\mathcal{B}^*}f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \leq c\delta^5/2.$$

Therefore we have

$$|\Lambda(\Pi_{\mathcal{B}\vee\mathcal{B}^*}f)-\Lambda(\delta\mathbb{1}_{[N']})| \geq c\delta^5/2.$$

Let $\mathcal{B}'=\mathcal{B}\vee\mathcal{B}^*$ and $g=\min(\Pi_{\mathcal{B}'}f,(1+c')\delta)$ and assume that $c'\leq\max(c,1)/10^5$. Let $\Omega'=\{x\in\mathbb{Z}/N\mathbb{Z}\colon g(x)\neq\Pi_{\mathcal{B}'}f(x)\}$ and notice that

$$|\Lambda(\Pi_{\mathcal{B}'}f)-\Lambda(g)| \leq 5\|\Pi_{\mathcal{B}'}f-g\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \leq 5\mathbb{P}[n\in\Omega']$$

$$|\Lambda(\delta\mathbb{1}_{[N']})-\Lambda(g)| \leq 5(1+c')^4\delta^4\|\delta\mathbb{1}_{[N']}-g\|_{L^1(\mathbb{Z}/N\mathbb{Z})}$$

$$\|g-\delta\mathbb{1}_{[N']}\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \leq \mathbb{P}[n\in\Omega'] + \|\delta\mathbb{1}_{[N']}-\Pi_{\mathcal{B}'}f\|_{L^1(\mathbb{Z}/N\mathbb{Z})}.$$

The first and second inequality follow from the first part of Lemma 3·5 while the final inequality follows from the triangle inequality.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Suppose that the proposition is false. Note that there are at most $2(2K)^{d_1+d_2}$ atoms of $\mathcal{B}'$ and define an atom to be small if it has measure at most $(2K)^{-d_1-d_2} \cdot (c\delta^5)/40$. We therefore have that on every atom which is not small,

$$\Pi_{\mathcal{B}'} f \leq (1+c')\delta$$

(else we take Ω^* to be that atom) and therefore $\mathbb{P}[n \in \Omega'] \leq c\delta^5/20$. Thus by the first inequality above and the triangle inequality we find

$$|\Lambda(\delta \mathbb{1}_{[N']}) - \Lambda(g)| \geq |\Lambda(\delta \mathbb{1}_{[N']}) - \Lambda(\Pi_{\mathcal{B}'} f)| - 5\mathbb{P}[n \in \Omega^*] \geq c\delta^5/4.$$

By the second inequality, this implies that

$$\|\delta \mathbb{1}_{[N']} - g\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \geq c\delta/100.$$

By the third inequality we have that

$$\|\delta \mathbb{1}_{[N']} - \Pi_{\mathcal{B}'} f\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \geq c\delta/200.$$

For a function h let $h_+ = \max(h, 0)$. For mean zero functions, $\|h\|_{L^1(\mathbb{Z}/N\mathbb{Z})} = 2\|h_+\|_{L^1(\mathbb{Z}/N\mathbb{Z})}$. This allows us to derive the contradiction since

$$\begin{aligned}\|\delta \mathbb{1}_{[N']} - \Pi_{\mathcal{B}'} f\|_{L^1(\mathbb{Z}/N\mathbb{Z})} &= 2\|(\Pi_{\mathcal{B}'} f - \delta \mathbb{1}_{[N']})_+\|_{L^1(\mathbb{Z}/N\mathbb{Z})} \\ &\leq 2c'\delta + 2((2K)^{-d_1-d_2} \cdot (c\delta^5)/40)(2(2K)^{d_1+d_2}) \leq 4c'\delta < c\delta/200.\end{aligned}$$

3·3. *Density increment onto subprogression*

We now finish the argument by partitioning factors in our cubic Bohr partition into long progressions. This is a technical modification of an argument of Green and Tao [**13**] for multiple quadratics (the case of a single polynomial is handled in Gowers [**9**]); as input they must provide, in the case of quadratics, an explicit implicit constant for a result of Schmidt [**32**] which provides recurrence for multiple polynomials mod 1 simultaneously. We require the analogous result for arbitrary degrees.

PROPOSITION 3·7. *Fix an integer $k \geq 1$. There exist $c = c(k) > 0$ such that the following holds. Let $\alpha_1, \ldots, \alpha_d$ be real numbers. Then*

$$\min_{1 \leq n \leq N} \max_{1 \leq i \leq d} \|\alpha_i n^k\|_{\mathbb{R}/\mathbb{Z}} \ll_k dN^{-c/d^2}.$$

We defer the proof of this input to Appendix B; it is an essentially verbatim copy of [**13**, appendix A].

Proof of Proposition 3·1. Step 1. Setup. Let $M = \exp((\log(1/\delta))^C)$ where $C > 0$ will be chosen later. Let us assume $N' > \exp(M(\delta))$ and $|\Lambda(f) - \Lambda(\delta \mathbb{1}_{[N']})| > c\delta^5$. Our aim is to find a progression of length at least $N^{1/M(\delta)}$ where our density is incremented.

Let $\mathcal{B}^* = [N'] \sqcup ((\mathbb{Z}/N\mathbb{Z}) \setminus [N'])$ be an initial factor. Let $M'(\delta) = \exp((\log(1/\delta))^{C'})$ where C' is chosen appropriately so as to apply Proposition 3·6. By assumption we have $N \geq M'(\delta)$. So there exist $d_1, d_2, K \leq M'(\delta)$, a factor $\mathcal{B}$ of complexity (d_1, d_2) and resolution K, and an atom Ω^* of $\mathcal{B} \vee \mathcal{B}^*$ such that:

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

(i) $\mathbb{P}_{n\in\mathbb{Z}/N\mathbb{Z}}[n\in\Omega^*]\geq \exp(-\exp((\log(1/\delta))^{C'}))$;

(ii) $\mathbb{E}[f(n)|n\in\Omega^*]\geq (1+2c')\delta$.

Now our aim is to partition Ω^* into few arithmetic progressions, namely at most $X=7^dN^{1-\Omega(1/d^7)}$ progressions where $d=M'(\delta)$. The number of elements of Ω^* in progressions of length at most $c'\delta|\Omega^*|/X$ is at most $c'\delta|\Omega^*|$. Letting the union of short progressions be Ω', we have that

$$\mathbb{E}[f(n)|n\in\Omega^*\setminus\Omega']\geq(1+c')\delta.$$

Therefore there is a progression of length longer than $c'\delta|\Omega^*|/X$ with density at least $(1+c')\delta$. We have

$$c'\delta|\Omega^*|/X\geq N^{1/M(\delta)}$$

due to our assumption $N'>\exp(M(\delta))$, assuming $C>0$ is chosen sufficiently large. (This is where the dependence in the first item of Proposition 3·1 comes from.)

We may write

$$\Omega^*=[N']\cap\bigcap_{i\in I_1}\{x\colon \|\phi_i(x)-j_i/K\|_{\mathbb{R}/\mathbb{Z}}\leq 1/(2K)\}\cap\bigcap_{i\in I_2}\{x\colon \|\varphi_i(x)-j_i/K\|_{\mathbb{R}/\mathbb{Z}}\leq 1/(2K)\},$$

where ϕ_i are linear functions on $\mathbb{Z}/N\mathbb{Z}$ and $\varphi_i\colon T\to\mathbb{R}$ are locally cubic functions on

$$T:=[N']\cap\bigcap_{i\in I_1}\{x\colon \|\phi_i(x)-j_i/K\|_{\mathbb{R}/\mathbb{Z}}\leq 1/(2K)\}.$$

We additionally have $|I_1|,|I_2|,K\leq M'(\delta)=:d$. (The fact that $\Omega^*\subseteq[N']$ is evident from $\mathbb{E}[f(n)|n\in\Omega^*]>0$, recalling that f is extended by 0s in Proposition 3·6.)

Step 2. Partitioning T into progressions where the cubic phases are near-constant. Our goal will be to partition T into subprogressions on which each φ_i for $i\in I_2$ is roughly constant mod 1. In fact, we will find a collection of progressions, say of the form $\{a+bn\colon n\in[L]\}$, so that for $n\in[L]$ we have $\varphi_i(a+bn)=\kappa+P_i^{(a,b)}(n) \mod 1$, where $P_i^{(a,b)}$ is a real polynomial of degree at most 3 with coefficients of size $O(L^{-4})$.

We first partition

$$T=\bigsqcup_{j\in J_1}T_j,$$

where T_j are progressions and $|J_1|\ll 2^{|I_1|}N^{1-1/(|I_1|+1)}$. To do this, we may use [**13**, proposition 6·3] (which is a simple application of Kronecker approximation, or Proposition 3·7 for $k=1$).

Since T_j is a progression, we may write $T_j=\{a_j+d_j,\dots,a_j+M_jd_j\}$, where $M_j=|T_j|$. Since each φ_i is locally cubic on this progression, we may write

$$\varphi_i(a_j+d_jn)=\alpha_i^{(j)}n^3+\beta_i^{(j)}n^2+\gamma_i^{(j)}n+\delta_i^{(j)} \bmod 1$$

for $i\in I_2$ and $n\in[M_j]$ where $j\in J_1$. Next, we perform three intermediate partitions of T_j in order to iteratively "reduce the degree" of the function φ_i until we see that it is roughly constant on our subprogressions.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

First, choose $1 \le r_j \le M_j^{1/3}$ so that

$$\max_{i \in I_2} \|\alpha_i^{(j)} r_j^3\|_{\mathbb{R}/\mathbb{Z}} \ll dM_j^{-c(3)/(3d^2)}.$$

Then we can partition $[M_j]$ into at most $M_j^{1-c(3)/(120d^2)}$ progressions of difference r_j and length roughly $M_j^{c(3)/(120d^2)}$. This corresponds to a partition

$$T_j = \bigsqcup_{k_3 \in D_3} T_{j,k_3}$$

with $|D_3| \le M_j^{1-c(3)/(120d^2)}$ and $|T_{j,k_3}| \sim M_j^{c(3)/(120d^2)}$. Furthermore, we have

$$T_{j,k_3} = \{a_j + (b_{j,k_3} + r_j t)d_j \colon t \in [|T_{j,k_3}|]\}$$

for some integer b_{j,k_3}. We have

$$\begin{aligned}\varphi_i(a_j + (b_{j,k_3} + r_j n)d_j) &= \alpha_i^{(j)}(b_{j,k_3} + r_j n)^3 + \beta_i^{(j)}(b_{j,k_3} + r_j n)^2 + \gamma_i^{(j)}(b_{j,k_3} + r_j n) + \delta_i^{(j)} \\ &= \alpha_i^{(j,k_3)} n^2 + \beta_i^{(j,k_3)} n + \gamma_i^{(j,k_3)} + \{\alpha_i^{(j)} r_j^3\} n^3 \bmod 1.\end{aligned}$$

By construction, $\{\alpha_i^{(j)} r_j^3\}$ is close to 0 mod 1 (namely, it is $\ll dM_j^{-c(3)/(3d^2)} \ll |T_{j,k_3}|^{-30}$).

We can thus consider the quadratic function obtained by removing this part, and repeat the same process on each T_{j,k_3}. We obtain iterative decompositions

$$T_{j,k_3} = \bigsqcup_{k_2 \in D_2^{(k_3)}} T_{j,k_3,k_2}, \quad T_{j,k_3,k_2} = \bigsqcup_{k_1 \in D_1^{(k_3,k_2)}} T_{j,k_3,k_2,k_1},$$

where:

$$\begin{aligned}|D_2^{(k_3)}| \le |T_{j,k_3}|^{1-c(2)/(120d^2)}, \quad |D_1^{(k_3,k_2)}| \le |T_{j,k_3,k_2}|^{1-c(1)/(120d^2)}; \\ |T_{j,k_3,k_2}| \sim |T_{j,k_3}|^{c(2)/(120d^2)}, \quad |T_{j,k_3,k_2,k_1}| \sim |T_{j,k_3,k_2}|^{c(1)/(120d^2)}.\end{aligned}$$

Additionally, we will find that if $T_{j,k_3,k_2,k_1} = \{a + bn \colon n \in [|T_{j,k_3,k_2,k_1}|]\}$ then

$$\varphi_i(a + bn) = \alpha_i^{(j,k_3,k_2,k_1)} n^3 + \beta_i^{(j,k_3,k_2,k_1)} n^2 + \gamma_i^{(j,k_3,k_2,k_1)} n + \kappa_i^{(j,k_3,k_2,k_1)} \bmod 1,$$

where

$$|\alpha_i^{(j,k_3,k_2,k_1)}| + |\beta_i^{(j,k_3,k_2,k_1)}| + |\gamma_i^{(j,k_3,k_2,k_1)}| \ll |T_{j,k_3,k_2,k_1}|^{-4}.$$

The number of such T_{j,k_3,k_2,k_1} partitioning T_j is in total at most

$$M_j^{1-c(3)/(12d^2)} (M_j^{c(3)/(12d^2)})^{1-c(2)/(12d^2)} \left((M_j^{c(3)/(12d^2)})^{c(2)/(12d^2)}\right)^{1-c(1)/(12d^2)} \le M_j^{1-\Omega(1/d^6)}.$$

(In the rare cases of any j such that $M_j \le \exp(O(d^6))$ we may simply partition M_j into progressions of length 1 appropriately.)

Finally, choosing appropriate value $p = \Omega(1/d^6)$, the total number of subprogressions is

$$\sum_{j \in J_1} M_j^{1-p} \le |J_1|^p \left(\sum_{j \in J_1} M_j\right)^{1-p} \le |J_1|^p N^{1-p} \le N^{1-\Omega(1/d^7)}.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Step 3. Partitioning Ω^* into few progressions. From the previous part, we have a partition of T into at most $N^{1-\Omega(1/d^7)}$ progressions of various lengths so that for a progression $\{a + bn : n \in [L]\}$ that appears, we have

$$\varphi_i(a + bn) = \kappa + P_i^{(a,b)}(n) \bmod 1,$$

where $P_i^{(a,b)}$ is a real polynomial of degree at most 3 with coefficients of size $O(L^{-4})$. Recalling $\Omega^* \subseteq T$, we can now intersect these progressions with Ω^*. Note that Ω^* merely imposes a condition on each φ_i for $i \in I_2$. Since each $P_i^{(a,b)}(n)$ is at most degree 3, within the range $n \in [L]$ it hits the cutoffs for the condition for φ_i at most $2 \cdot 3 = 6$ times in the real numbers. Since $P_i^{(a,b)}(n)$ has small coefficients (so is small on $n \in [L]$), there is no rounding wraparound and the same holds for its image mod 1 over the domain $n \in [L]$.

The upshot is that each progression is cut into at most 7 pieces by the condition on φ_i defining Ω^* for each $i \in I_2$. This gives at most $7^{|I_2|} \leq 7^d$ total pieces.

Thus, we have a partition of Ω^* into at most $7^d N^{1-\Omega(1/d^7)}$ pieces, which combined with the argument above completes the proof.

Appendix A. Conventions regarding nilsequences and effective equidistribution

We begin this appendix by giving the precise definition of the complexity of a nilmanifold; this definition is exactly as in [37, definition 6·1].

Definition A·1. Let $s \geq 1$ be an integer and let $K > 0$. A *filtered nilmanifold* G/Γ *of degree s and complexity at most K* consists of the following:

(i) a nilpotent, connected, and simply connected Lie group G of dimension m, which can be identified with its Lie algebra $\log G$ via the exponential map $\exp : \log G \to G$;

(ii) a filtration $G_\bullet = (G_i)_{i \geq 0}$ of closed connected subgroups G_i of G with

$$G = G_0 = G_1 \geqslant G_1 \geqslant \cdots \geqslant G_s \geqslant G_{s+1} = \mathrm{Id}_G$$

such that $[G_i, G_j] \subseteq G_{i+j}$ for all $i, j \geq 0$ (equivalently $[\log G_i, \log G_j] \subseteq \log G_{i+j}$);

(iii) a discrete cocompact subgroup Γ of G; and

(iv) a linear basis $\mathcal{X} = \{X_1, \ldots, X_m\}$ of $\log G$, known as a *Mal'cev basis*.

We, furthermore, require that this data obeys the following conditions:

(1) for $1 \leq i, j \leq m$, one has Lie algebra relations

$$[X_i, X_j] = \sum_{i,j<k\leq m} c_{ijk} X_k$$

for rational numbers c_{ijk} of height at most K (we will often refer to these as the *Lie bracket structure constants*);

(2) for each $1 \leq i \leq s$, the Lie algebra $\log G_i$ is spanned by $\{X_j : m - \dim(G_i) < j \leq m\}$; and

(3) the subgroup Γ consists of all elements of the form $\exp(t_1 X_1) \cdots \exp(t_m X_m)$ with $t_i \in \mathbb{Z}$.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

We note that the conditions imply $[G, G_s] = \mathrm{Id}_G$, i.e., G_s is contained in the center of G (commutes with every element).

Next, we will define polynomial sequences in filtered nilpotent groups. This concrete definition is equivalent (by [**11**, lemma 6·7]) to the one given in [**11**].

Definition A·2. We adopt the conventions of Definition A·1. Let G be a filtered nilpotent group of degree s. A function $g\colon \mathbb{Z} \to G$ is a *polynomial sequence* if there exist elements $g_i \in G_i$ for $i = 0, \ldots, s$ such that

$$g(n) = g_0 g_1^{\binom{n}{1}} \cdots g_s^{\binom{n}{s}},$$

where $\binom{n}{i} = 1/i! \prod_{j=0}^{i-1} (n-j)$, for all $n \in \mathbb{Z}$. We say a polynomial sequence $g(n)$ is N-periodic with respect to a lattice Γ if

$$g(n)g(n+N)^{-1} \in \Gamma$$

for all $n \in \mathbb{Z}$.

We will denote the set of polynomial sequences $g\colon \mathbb{Z} \to G$ relative to the filtration $G_\bullet$ of G by poly $(\mathbb{Z}, G_\bullet)$. It turns out that poly $(\mathbb{Z}, G_\bullet)$ is a group under the natural multiplication of sequences–this is due to Lazard [**19**] and Leibman [**20, 21**].

Now we can define Mal'cev coordinates, the explicit metrics on G and G/Γ used in our work, and the precise definition of the Lipschitz norm of functions on G/Γ. These definitions are exactly as in [**11**, appendix A].

Definition A·3. We adopt the conventions of Definition A·1. Given a Mal'cev basis $\mathcal{X}$ and $g \in G$, there exists $(t_1, \ldots, t_m) \in \mathbb{R}^m$ such that

$$g = \exp(t_1 X_1 + t_2 X_2 + \ldots t_m X_m).$$

We define *Mal'cev coordinates of first kind* $\psi_{\exp} = \psi_{\exp,\mathcal{X}}\colon G \to \mathbb{R}^m$ *for* g *relative to* $\mathcal{X}$ by

$$\psi_{\exp}(g) := (t_1, \ldots, t_m).$$

Given $g \in G$ there also exists $(u_1, \ldots, u_m) \in \mathbb{R}^m$ such that

$$g = \exp(u_1 X_1) \cdots \exp(u_m X_m),$$

and we define the *Mal'cev coordinates of second kind* $\psi = \psi_{\mathcal{X}}\colon G \to \mathbb{R}^m$ *for* g *relative to* $\mathcal{X}$ by

$$\psi(g) := (u_1, \ldots, u_m).$$

We then define a metric $d = d_{\mathcal{X}}$ on G by

$$d(x, y) :=$$

$$\inf \left\{ \sum_{i=1}^{n} \min \left(\|\psi(x_i x_{i+1}^{-1})\|, \|\psi(x_{i+1} x_i^{-1})\| \right) \colon n \in \mathbb{N}, x_1, \ldots, x_{n+1} \in G, x_1 = x, x_{n+1} = y \right\},$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

where $\|\cdot\|$ denotes the ℓ^∞-norm on $\mathbb{R}^m$, and define a metric on G/Γ by

$$d(x\Gamma, y\Gamma) = \inf_{\gamma,\gamma' \in \Gamma} d(x\gamma, y\gamma').$$

Furthermore, for any function $F\colon G/\Gamma \to \mathbb{C}$, we define

$$\|F\|_{\mathrm{Lip}} := \|F\|_\infty + \sup_{\substack{x,y \in G/\Gamma \\ x \neq y}} \frac{|F(x) - F(y)|}{d(x,y)}.$$

We recall the notion of a horizontal character and the notion of a function F having a vertical frequency; our definitions are exactly as in [**11**, definitions 1·5, 3·3, 3·4, 3·5].

Definition A·4. Given a filtered nilmanifold G/Γ, the *horizontal torus* is defined to be

$$(G/\Gamma)_{\mathrm{ab}} := G/[G,G]\Gamma.$$

A *horizontal character* is a continuous homomorphism $\eta\colon G \to \mathbb{R}/\mathbb{Z}$ that annihilates Γ; such characters may be equivalently viewed as characters on the horizontal torus. A horizontal character is *nontrivial* if it is not identically zero.

Furthermore, if the nilmanifold G/Γ has degree s, the vertical torus is defined to be

$$G_s/(G_s \cap \Gamma).$$

A *vertical character* is a continuous homomorphism $\xi\colon G_s \to \mathbb{R}/\mathbb{Z}$ that annihilates $\Gamma \cap G_s$. Setting $m_s = \dim G_s$, one may use the last m_s coordinates of the Mal'cev coordinate map to identify G_s and $G_s/(G_s \cap \Gamma)$ with $\mathbb{R}^{m_s}$ and $\mathbb{R}^{m_s}/\mathbb{Z}^{m_s}$, respectively. Thus, we may identify any vertical character ξ with a unique $k \in \mathbb{Z}^{m_s}$ such that $\xi(x) = k \cdot x$ under this identification $G_s/(\Gamma \cap G_s) \cong \mathbb{R}^{m_s}/\mathbb{Z}^{m_s}$. We refer to k as the *frequency* of the character ξ, we write $|\xi| := \|k\|_\infty$ to denote the magnitude of the frequency ξ, and say that a function $F\colon G/\Gamma \to \mathbb{C}$ *has a vertical frequency* ξ if

$$F(g_s \cdot x) = e(\xi(g_s))F(x)$$

for all $g_s \in G_s$ and $x \in G/\Gamma$.

Appendix B. Explicit dimension dependence for Schmidt's polynomial recurrence

We now prove the explicit version of Schmidt's polynomial recurrence (Proposition 3·7). The proof we provide is essentially verbatim from [**13**, appendix A]. For those familiar with the proof presented there, the use of theta functions is unrelated to the fact that one is finding small fractional parts of quadratic polynomials. Roughly speaking, the theta function is only used as a smooth approximation of the neighbourhood of points in a lattice which has nice Fourier properties.

Definition B·1. Suppose that Λ is a lattice of full rank in $\mathbb{R}^d$. The dual lattice, denoted Λ^*, is $\Lambda^* = \{\xi \in \mathbb{R}^d : \xi \cdot m \in \mathbb{Z} \text{ for all } m \in \Lambda\}$. For any $t > 0$ and $x \in \mathbb{R}^d$, we define

$$\Theta_\Lambda(t,x) := \sum_{m \in \Lambda} e^{-\pi t \|x-m\|_2^2}.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Finally define

$$A_\Lambda := \Theta_{\Lambda^*}(1,\vec{0}) = \sum_{\xi\in\Lambda^*} e^{-\pi\|\xi\|_2^2} = \det(\Lambda)\sum_{m\in\Lambda} e^{-\pi\|m\|_2^2}.$$

We next need a version of Weyl's inequality; we take the statement from [11, proposition 4·3].

PROPOSITION B·2. *There exists $C = C(k) > 0$ such that the following holds. Suppose that $g\colon \mathbb{Z}\to\mathbb{R}$ is a polynomial of degree k and $\delta\in(0,1/2)$. If $N\geq\delta^{-C}$ and*

$$\left|\mathbb{E}_{n\in[N]} e(g(n))\right| \geq \delta$$

then there exists a positive integer ℓ such that $\ell\leq\delta^{-C}$ and

$$\|\ell g\|_{C^\infty[N]} \leq \delta^{-C}.$$

Remark. For a polynomial $P(n) = a_0 + \cdots + a_k n^k$, we have $\|P\|_{C^\infty[N]} = \max_{0<i\leq k} N^i\|a_i\|_{\mathbb{R}/\mathbb{Z}}$.

A crucial object of study for $\alpha\in\mathbb{R}^d$ will be

$$F_{\Lambda,\alpha}(N) := \det(\Lambda)\mathbb{E}_{|n|\leq N}\Theta_\Lambda(1,n^k\alpha) = \sum_{\xi\in\Lambda^*} e^{-\pi\|\xi\|_2^2}\mathbb{E}_{|n|\leq N} e(n^k\xi\cdot\alpha);$$

the final equality is a consequence of the Poisson summation formula (cf. [13, (A·3), (A·5)]).

The following appears as [13, lemma A·5 (i),(ii),(iii)] except with trivial modification for the differing degree.

LEMMA B·3. *Let Λ be a lattice of full rank in $\mathbb{R}^d$, let $\alpha\in\mathbb{R}^d$, and $N>0$. We have the following properties*:

(i) (contraction on N) For any $c\in(10/N,1)$, we have $F_{\Lambda,\alpha}(N)\gg cF_{\Lambda,\alpha}(cN)$;

(ii) (dilation of α) For any integer $q\geq 1$, we have $F_{\Lambda,\alpha}(N)\gg\frac{1}{q}F_{\Lambda,q^k\alpha}(N/q)$;

(iii) (stability) Suppose that $\|\widetilde{\alpha}-\alpha\|_2\leq\varepsilon N^{-k}$ with $\varepsilon\in(0,1)$. Then $F_{\Lambda,\alpha}(N)\gg F_{(1+\varepsilon)\Lambda,(1+\varepsilon)\widetilde{\alpha}}(N)$.

Proof. We prove each of the items in turn. For the first item, note that $\Theta_\Lambda > 0$ so

$$F_{\Lambda,\alpha}(N) = \det(\Lambda)\frac{\sum_{|n|\leq N}\Theta_\Lambda(1,n^k\alpha)}{2\lfloor N\rfloor+1} \geq \det(\Lambda)\frac{\sum_{|n|\leq cN}\Theta_\Lambda(1,n^k\alpha)}{2\lfloor N\rfloor+1} \geq \frac{cF_{\Lambda,\alpha}(cN)}{2}.$$

For the second item, if $q>N$ note that

$$F_{\Lambda,\alpha}(N) = \det(\Lambda)\mathbb{E}_{|n|\leq N}\Theta_\Lambda(1,n^k\alpha) \geq \det(\Lambda)\frac{\Theta(1,\vec{0})}{2N+1} \geq \frac{F_{\Lambda,q^k\alpha}(N/q)}{3q}.$$

For $q\leq N$, we have

$$F_{\Lambda,\alpha}(N) = \det(\Lambda)\frac{\sum_{|n|\leq N}\Theta_\Lambda(1,n^k\alpha)}{2\lfloor N\rfloor+1} \geq \det(\Lambda)\frac{\sum_{|n|\leq N/q}\Theta_\Lambda(1,n^k\cdot q^k\alpha)}{2\lfloor N\rfloor+1} \geq \frac{F_{\Lambda,q^k\alpha}(N/q)}{4q}.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

We now handle the final item; let $X := \|n^k\alpha - m\|_2$, $\widetilde{X} = \|n^k\widetilde{\alpha} - m\|_2$, and note that $|X - \widetilde{X}| \le \|n^k(\alpha - \widetilde{\alpha})\|_2 \le \varepsilon$. We have that $4\pi + \pi(1+\varepsilon)^2\widetilde{X}^2 \ge \pi X^2$ for all $X \ge 0$ and $\widetilde{X} \in [X \pm \varepsilon]$; the inequality is trivial for $X \le 2$ and for $X \ge 2$ we have that $(1+\varepsilon)(X-\varepsilon) - X = \varepsilon(X - (1+\varepsilon)) \ge 0$. Therefore

$$e^{-\pi X^2} \ge e^{-4\pi} \cdot e^{-\pi(1+\varepsilon)^2\widetilde{X}^2} = e^{-4\pi} \cdot e^{-\pi\|n^k(1+\varepsilon)\alpha - (1+\varepsilon)m\|_2^2}.$$

Summing over $m \in \Lambda$ and $|n| \le N$, the desired result follows.

The key point (which is identical to [13, lemma A·6] modulo citing Weyl's inequality for higher degree polynomials) is noting that if $F_{\Lambda,\alpha}(N)$ is small then there exists a vector in Λ^* which is nearly orthogonal to α.

LEMMA B·4. *(Schmidt alternative). There exists $C = C(k) > 0$ such that the following holds. Suppose that $\alpha \in \mathbb{R}^d$ and $\Lambda \subseteq \mathbb{R}^d$ is a full rank lattice. Let N be a positive integer. Then one of the following always holds*:

(i) $F_{\Lambda,\alpha}(N) \ge 1/2$;

(ii) There exist positive integer $q \ll dA_\Lambda^C$ and primitive $\xi \in \Lambda^* \setminus \{0\}$ such that

$$\|\xi\|_2 \ll \sqrt{d} + \sqrt{\log A_\Lambda} \text{ and } \|q\xi \cdot \alpha\|_{\mathbb{R}/\mathbb{Z}} \ll A_\Lambda^C N^{-k}.$$

($\xi \in \Lambda^*$ is primitive if $\xi/n \notin \Lambda^*$ for all $n \ge 2$.)

Proof. We claim that it suffices to prove that if $F_{\Lambda,\alpha}(N) \le 1/2$, then there exists $q \ll A_\Lambda^C$ and a vector

$$\|\xi\|_2 \ll \sqrt{d} + \sqrt{\log A_\Lambda} \text{ and } \|q\xi \cdot \alpha\|_{\mathbb{R}/\mathbb{Z}} \ll A_\Lambda^C N^{-k}.$$

Note that the shortest vector in Λ^* is easily seen to be $\gg A_\Lambda^{-1}$ by considering the contribution to A_Λ by scalar multiples of the shortest vector. Therefore ξ is a multiple of $\widetilde{\xi}$ which is primitive, we have $\|\xi\|_2/\|\widetilde{\xi}\|_2 \ll (\sqrt{d} + \sqrt{\log A_\Lambda})A_\Lambda$, and therefore q can be replaced by $q' = q\|\xi\|_2/\|\widetilde{\xi}\|_2 \ll (\sqrt{d} + \sqrt{\log A_\Lambda})A_\Lambda \cdot A_\Lambda^C \ll d \cdot A_\Lambda^{C+1}$. This gives the desired result.

Let $M = 4(\sqrt{d} + \sqrt{\log A_\Lambda})$ and note that

$$\begin{aligned}
1/2 \le |F_{\Lambda,\alpha}(N) - 1| &= \left| \sum_{\xi \in \Lambda^* \setminus \{0\}} e^{-\pi\|\xi\|_2^2} \mathbb{E}_{|n| \le N} e(n^k\xi \cdot \alpha) \right| \\
&\le \sum_{\substack{\xi \in \Lambda^* \\ \|\xi\|_2 \ge M}} e^{-\pi\|\xi\|_2^2} + \left(\sup_{\substack{\xi \in \Lambda^* \\ \|\xi\|_2 \le M}} \left| \mathbb{E}_{|n| \le N} e(n^k\xi \cdot \alpha) \right| \right) \cdot \sum_{\xi \in \Lambda^*} e^{-\pi\|\xi\|_2^2} \\
&\le e^{-\pi M^2/2} \sum_{\xi \in \Lambda^*} e^{-\pi\|\xi\|_2^2/2} + A_\Lambda \sup_{\substack{\xi \in \Lambda^* \\ \|\xi\|_2 \le M}} \left| \mathbb{E}_{|n| \le N} e(n^k\xi \cdot \alpha) \right| \\
&= e^{-\pi M^2/2} 2^{d/2} \det(\Lambda) \sum_{m \in \Lambda} e^{-2\pi\|m\|_2^2} + A_\Lambda \sup_{\substack{\xi \in \Lambda^* \\ \|\xi\|_2 \le M}} \left| \mathbb{E}_{|n| \le N} e(n^k\xi \cdot \alpha) \right| \\
&\le 1/4 + A_\Lambda \sup_{\substack{\xi \in \Lambda^* \\ \|\xi\|_2 \le M}} \left| \mathbb{E}_{|n| \le N} e(n^k\xi \cdot \alpha) \right|,
\end{aligned}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

where the equality follows from Poisson summation. We therefore have that

$$\left|\mathbb{E}_{|n|\leq N} e(n^k\xi\cdot\alpha)\right| \geq 1/(4A_\Lambda)$$

and the result follows from Proposition B·2 as desired.

The following appears as [**13**, lemma A·7].

LEMMA B·5. *Let $\Lambda'\subseteq\mathbb{R}^{d-1}$ and $\Lambda\subseteq\mathbb{R}^d$ be full-rank lattices with $\Lambda'\subseteq\Lambda$, embedding $\mathbb{R}^{d-1}$ in $\mathbb{R}^d$ by putting 0 in the final coordinate. Suppose that $\alpha\in\mathbb{R}^d$ and $\alpha'\in\mathbb{R}^{d-1}$ satisfy $\alpha-\alpha'\in\Lambda$. Then*

$$F_{\Lambda,\alpha}(N)\geq\frac{\det(\Lambda)}{\det(\Lambda')}F_{\Lambda',\alpha'}(N).$$

We now combine Lemma B·4 and Lemma B·5 exactly as in [**13**, proposition A·8].

LEMMA B·6. *There exists $C=C(k)>0$ such that the following holds. Suppose $\alpha\in\mathbb{R}^d$ and $\Lambda\subset\mathbb{R}^d$ is full rank. If $N>(dA_\Lambda)^C$ and $F_{\Lambda,\alpha}(N)\leq 1/2$, then there exist $\Lambda'\subseteq\mathbb{R}^{d-1}$, $\alpha'\in\mathbb{R}^{d-1}$, and $N'\gg N(dA_\Lambda)^{-C}$ such that*

$$\begin{aligned}A_{\Lambda'}&\ll(\sqrt{d}+\sqrt{\log A_\Lambda})A_\Lambda,\\ F_{\Lambda,\alpha}(N)&\geq(dA_\Lambda)^{-C}F_{\Lambda',\alpha'}(N^*).\end{aligned}$$

Proof. By Lemma B·4 we have that there exist primitive $\xi\in\Lambda^*\setminus\{0\}$ and $q\ll dA_\Lambda^C$ such that

$$\|\xi\|_2\ll\sqrt{d}+\sqrt{\log A_\Lambda}\text{ and }\|q\xi\cdot\alpha\|_{\mathbb{R}/\mathbb{Z}}\ll A_\Lambda^C N^{-k}.$$

By applying a rotation to both α and Λ, we may assume that $\xi=\xi_d e_d$. Note that $|\xi_d|\gg A_\Lambda^{-1}$ and $\|\xi\cdot q^k\alpha\|_{\mathbb{R}/\mathbb{Z}}\ll dA_\Lambda^{O(1)}N^{-k}$. Thus there exists $\beta\in\mathbb{R}^d$ such that $\beta\cdot\xi\in\mathbb{Z}$ and

$$\|\beta-q^k\alpha\|_2\leq|\xi_d|^{-1}\|\xi\cdot q^k\alpha\|_{\mathbb{R}/\mathbb{Z}}\ll dA_\Lambda^{O(1)}N^{-k}.$$

We take $N_*\gg(dA_\Lambda)^{-O(1)}N$ such that $\|\beta-q^k\alpha\|_2\leq N_*^{-k}/d$. We have that

$$\begin{aligned}F_{\Lambda,\alpha}(N)&\gg(dA_\Lambda)^{-O(1)}F_{\Lambda,\alpha}(N^*)\gg(dA_\Lambda)^{-O(1)}F_{\Lambda,q^k\alpha}(N^*/q)\\&\gg(dA_\Lambda)^{-O(1)}F_{(1+1/d)\Lambda,(1+1/d)\beta}(N^*/q),\end{aligned}$$

where we have applied the first, second, and then third item of Lemma B·3. Let $\pi:(x_1,\dots,x_d)\to(x_1,\dots,x_{d-1})$ and take $\alpha'=(1+1/d)\pi(\beta)$, $\Lambda'=\pi((1+1/d)\Lambda)$, and $N'=N^*/q$. (That Λ' is a lattice uses that ξ is primitive.) Finally by Lemma B·5,

$$F_{(1+1/d)\Lambda,(1+1/d)\beta}(N^*/q)\geq\frac{\det((1+1/d)\Lambda)}{\det(\Lambda')}\cdot F_{\Lambda',\alpha'}(N')=((1+1/d)\xi_d)F_{\Lambda',\alpha'}(N')$$

and using the lower bound on $|\xi_d|$ we have the desired lower bound on $F_{\Lambda,\alpha}(N)$. For the upper bound of $A_{\Lambda'}$, by positivity of the exponential function we have

$$\frac{A_{\pi(\Lambda)}}{\det(\pi(\Lambda))}\leq\frac{A_\Lambda}{\det(\Lambda)}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

and therefore

$$A_{\Lambda'} \leq (1+1/d)^d A_{\pi(\Lambda)} \leq eA_\Lambda \frac{\det(\pi(\Lambda))}{\det(\Lambda)} = eA_\Lambda \cdot \|\xi\|_2$$

as desired.

Iterating Lemma B·6 exactly as is done in [**13**, proposition A·9] (see arXiv version for a corrected statement) yields the following.

PROPOSITION B·7. *There exists $C = C(k) > 0$ such that the following holds. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice of full rank with $\det(\Lambda) \geq 1$. Then for any integer N we have*

$$F_{\Lambda,\alpha}(N) \gg d^{-Cd^2} A_\Lambda^{-Cd}.$$

We now prove Proposition 3·7; our deduction once again is identical to [**13**, proposition A·2].

Proof of Proposition 3·7. Let R be a quantity to be chosen later; $R \gg d^3$ throughout. Let $\Lambda := R\mathbb{Z}^d$ and $A_\Lambda = R^d(\sum_{m\in R\mathbb{Z}} e^{-\pi m^2})^d \leq 2^{O(d)}R^d$. This along with Proposition B·7 implies that

$$F_{\Lambda,\alpha}(N) \gg R^{-O(d^2)}$$

or

$$\mathbb{E}_{|n|\leq N} \sum_{m\in R\mathbb{Z}^d} e^{-\pi\|n^k\alpha - m\|_2^2} \gg R^{-O(d^2)}.$$

For $N \gg (2R)^{\Omega(d^2)}$, there is $n \in \{1,\dots N\}$ such that

$$\sum_{m\in R\mathbb{Z}^d} e^{-\pi\|n^k\alpha - m\|_2^2} \gg R^{-O(d^2)}.$$

If $\|n^k\alpha - m\|_2 \geq \sqrt{R}$ for all $m \in R\mathbb{Z}^d$, we have

$$\begin{aligned}\sum_{m\in R\mathbb{Z}^d} e^{-\pi\|n^k\alpha - m\|_2^2} &\leq e^{-\pi R/2} \sum_{m\in R\mathbb{Z}^d} e^{-\pi\|n^k\alpha - m\|_2^2/2} \\ &= e^{-\pi R/2}\frac{2^{d/2}}{\det(\Lambda)} \sum_{\xi\in\Lambda^*} e^{-2\pi\|\xi\|_2^2}\exp(\xi\cdot n^k\alpha) \\ &\leq e^{-\pi R/2}\frac{2^{d/2}A_\Lambda}{\det(\Lambda)} \ll e^{-\pi R/2}2^{O(d)}.\end{aligned}$$

This is a contradiction; thus if $N \geq (2R)^{\Omega(d^2)}$ and $R \gg d^3$ then there exists $1 \leq n \leq N$ such that

$$\|n^k\alpha_j\|_{\mathbb{R}/\mathbb{Z}} \leq 1/\sqrt{R}$$

for all $j = 1,\dots,d$. The result follows upon taking $R = N^{c/d^2}$ for an appropriately small constant c (if N is small enough that $R \gg d^3$ fails to hold, the result is trivial).

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Appendix C. Constructing an periodic nilsequence with integral structure constants

For technical reasons, it will be advantageous to construct nilsequences where the underlying nilmanifold has structure constants which are integral (and in fact divisible by a sufficient fixed integer). This follows via a straightforward lifting procedure where one replaces the underlying Mal'cev basis $\{e_1, \dots, e_{\dim(G)}\}$ for Γ with $\{e_1^L, \dots, e_{\dim(G)}^L\}$ for a sufficiently large constant L. This is primarily to avoid needing various fractional part identities in the proof of Proposition 2·3.

LEMMA C·1. *Fix an integer $K \geq 1$. Suppose we are given a degree s nilmanifold G/Γ with dimension d and complexity M and $F\colon G/\Gamma \to \mathbb{C}$ with Lipschitz norm bounded by L (with respect to the metric $d_{G/\Gamma}$).*

There exist $\widetilde{\Gamma}, \widetilde{F}$ such that $\widetilde{F}\colon G/\widetilde{\Gamma} \to \mathbb{C}$ such that for any $g \in G$ we have

$$F(g\Gamma) = \widetilde{F}(g\widetilde{\Gamma}).$$

Furthermore we have that $\widetilde{\Gamma}$ has a Mal'cev basis $\widetilde{X}$ with all Lie bracket structure constants (the values c_{ijk} in Definition A·1) being integral and divisible by K, that $G/\widetilde{\Gamma}$ (with $\widetilde{X}$) has complexity bounded by $O_s(K) \cdot \exp(O(M))$, and that F has Lipschitz norm bounded by $L \cdot (K \cdot \exp(O(M)))^{O_s(d^{O_s(1)})}$.

Proof. Let $\mathcal{X} = \{X_1, \dots, X_d\}$ denote the Mal'cev basis for the Lie algebra $\mathfrak{g}$ implicit in the complexity bound given for G/Γ. In order for $\mathcal{X}$ to be a Mal'cev basis we have:

(i) for each $j = 0, \dots, d$, the subspace $\mathfrak{h}_j := \operatorname{span}(X_{j+1}, \dots, X_d)$ is a Lie algebra ideal in $\mathfrak{g}$, and hence $H_j := \exp(\mathfrak{h}_j)$ is a normal Lie subgroup of G;

(ii) if $d_i = \dim(G_i)$ then $G_i = H_{d-d_i}$;

(iii) each $g \in G$ may be written uniquely as $\exp(t_1 X_1) \cdots \exp(t_m X_m)$ for $t_i \in \mathbb{R}$;

(iv) Γ are exactly the elements which can be written in the form $\exp(t_1 X_1) \cdots \exp(t_m X_m)$ with $t_i \in \mathbb{Z}$ (and is a discrete cocompact subgroup).

We also require that the Lie algebra relations are appropriately filtered, corresponding to the containments $[G_i, G_j] \subseteq G_{i+j}$.

We take $\widetilde{\mathcal{X}} = \{R \cdot X_1, \dots, R \cdot X_d\} =: \{\widetilde{X}_1, \dots, \widetilde{X}_d\}$ for a large positive integer R to be chosen later. We take $\widetilde{\Gamma} = \langle \exp(RX_i) \rangle$ and claim that $\widetilde{\mathcal{X}}$ is a valid Mal'cev basis for $G, \widetilde{\Gamma}$. All conditions for verifying a Mal'cev basis are trivial for us except for the fourth bullet point. The key point is verifying that every element of the subgroup generated by elements $\exp(RX_i)$ can be presented in the desired form.

We take $R = C_s \cdot \operatorname{lcm}(1, \dots, M) \cdot K$ for some appropriate positive integer C_s. Note that

$$[\widetilde{X}_i, \widetilde{X}_j] = \sum_{i,j<k\leq m} c_{ijk} \cdot R^2 X_k = \sum_{i,j<k\leq m} (Rc_{ijk})\widetilde{X}_k.$$

The crucial point here is that $Rc_{ijk} \in C_s\mathbb{Z}$ by construction. Let $e_i = \exp(X_i)$ and suppose we have a word in $\widetilde{\Gamma}$ given by

$$w = e_{i_1}^{\pm R} e_{i_2}^{\pm R} \cdots e_{i_t}^{\pm R},$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

where each $\pm$ denotes an appropriate sign. Note $e_i^R = \exp(\widetilde{X}_i)$. We first use the Baker–Campbell–Hausdorff formula to write

$$w = \exp\left(\sum_{j=1}^{d} w_j \widetilde{X}_j\right).$$

Note that $w_j \in \mathbb{Z}$ since the Lie bracket structure constants Rc_{ijk} are in $C_s\mathbb{Z}$, and Baker–Campbell–Hausdorff only goes to finite height due to the bounded step of the nilpotent Lie group. Now we iteratively pull out a single Mal'cev basis element "one at a time". We have

$$\exp(-w_1\widetilde{X}_1)w = \exp(-w_1\widetilde{X}_1)\exp\left(\sum_{j=1}^{d} w_j\widetilde{X}_j\right) = \exp\left(\sum_{j=2}^{d} w_j'\widetilde{X}_j\right)$$

for some $w_j' \in \mathbb{Z}$ using Baker–Campbell–Hausdorff again. Note that we have eliminated use of $\widetilde{X}_1$ in the right side (which uses the observation that $c_{ijk} = 0$ if $k \le \max(i,j)$). We iterate this, obtaining

$$\left(\prod_{j=d}^{1} \exp(-w_j^*\widetilde{X}_j)\right)w = \exp(0) = \mathrm{id}_G$$

for appropriate $w_j^* \in \mathbb{Z}$ (note that the product has decreasing indices left-to-right). Solving for w gives the result. This completes the proof that $\widetilde{\mathcal{X}}$ is a valid Mal'cev basis.

Now we define $\widetilde{F}$ and verify various claims regarding bounds. Note that $G/\widetilde{\Gamma}$ clearly has complexity bounded by $O(RM)$ which is as desired. Since $\widetilde{\Gamma} \subseteq \Gamma$, we may lift F from G/Γ to $\widetilde{F}$ on $G/\widetilde{\Gamma}$ in the obvious manner, composing with a quotient. It suffices to verify that this does not ruin the Lipschitz norm. Let $M' = \exp(O(M)) \cdot O_s(K)$; we have that the diameter of $G/\widetilde{\Gamma}$ is bounded by $(M')^{O_s(d^{O_s(1)})}$ (cf. [11, lemma A·16] with explicit dimension dependence [22, lemma B·8]). Since $\|\widetilde{F}\|_\infty = \|F\|_\infty \le \|F\|_{\mathrm{Lip}} \le L$, it suffices to consider points which are within $d_{G/\widetilde{\Gamma}}$ distance ε when verifying the Lipschitz bound for $\widetilde{F}$ as long as we are willing to lose a factor of ε^{-1}.

Note that $d_{G,\widetilde{\mathcal{X}}}(x,y) = (1/R)d_{G,\mathcal{X}}(x,y)$ since Mal'cev coordinates of the second kind in $\mathcal{X}, \widetilde{\mathcal{X}}$ differ by a factor of R. Now consider $x, y \in G/\widetilde{\Gamma}$ such that $d_{G/\widetilde{\Gamma}}(x,y) \le (M')^{-O_s(d^{O_s(1)})}$. There exist representatives of $\widetilde{x}$ and $\widetilde{y}$ for x and y with respect to $\widetilde{\Gamma}$ such that

$$d_{G/\widetilde{\Gamma}}(x,y) = d_{G,\widetilde{\mathcal{X}}}(\widetilde{x},\widetilde{y}) \text{ and } d_{G,\widetilde{\mathcal{X}}}(\widetilde{x},\mathrm{Id}_G) + d_{G,\widetilde{X}}(\widetilde{y},\mathrm{Id}_G) \le (M')^{O_s(d^{O_s(1)})}.$$

Additionally, we claim that since x, y are sufficiently close in $G/\widetilde{\Gamma}$, we find that

$$d_{G/\Gamma}(\widetilde{x}\Gamma,\widetilde{y}\Gamma) = d_{G,\mathcal{X}}(\widetilde{x},\widetilde{y}).$$

The argument is as follows. If this is not true, then there is $\gamma \in \Gamma \setminus \{0\}$ with $d_{G,\mathcal{X}}(\widetilde{x},\widetilde{y}\gamma) < d_{G,\mathcal{X}}(\widetilde{x},\widetilde{y})$. By approximate left-invariance of the metric on G ([11, lemma A·5] with explicit dimension dependence [22, lemma B·4]) we see

$$d_{G,\mathcal{X}}(\widetilde{y}^{-1}\widetilde{x},\gamma) \le (M')^{O_s(d^{O_s(1)})}Rd_{G/\widetilde{\Gamma}}(x,y) \le (M')^{O_s(d^{O_s(1)})}\varepsilon.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Additionally,

$$d_{G,\mathcal{X}}(\widetilde{y}^{-1}\widetilde{x}, \mathrm{id}_G) \leq (M')^{O_s(d^{O_s(1)})}\varepsilon.$$

The triangle inequality yields

$$d_{G,\mathcal{X}}(\gamma, \mathrm{id}_G) \leq (M')^{O_s(d^{O_s(1)})}\varepsilon.$$

Now, the metric is comparable to the L^∞ distance in second kind Mal'cev coordinates up to a factor of $(M')^{O_s(d^{O_s(1)})}$ ([**11**, lemma A·4] with explicit dimension dependence [**22**, lemma B·3]) as long as ε is sufficiently small here, so we find

$$1 \leq (M')^{O_s(d^{O_s(1)})}\varepsilon.$$

This as a contradiction as long as we take sufficiently small ε (which will be admissible in the bounds we need). Finally,

$$\begin{aligned}|\widetilde{F}(x\widetilde{\Gamma}) - \widetilde{F}(y\widetilde{\Gamma})| &= |F(\widetilde{x}\Gamma) - F(\widetilde{y}\Gamma)| \leq Ld_{G/\Gamma}(\widetilde{x}, \widetilde{y}) = Ld_{G,\mathcal{X}}(\widetilde{x}, \widetilde{y}) = LRd_{G,\widetilde{\mathcal{X}}}(\widetilde{x}, \widetilde{y})\\ &= LRd_{G/\widetilde{\Gamma}}(x, y)\end{aligned}$$

and we are done.

Having constructed a nilsequence with integral Lie bracket structure constants, it will also prove useful to construct a nilsequence which is additionally periodic. For this purpose, we quantify a construction of Manners [**24**] which demonstrates that one may lift a nilsequence along a subset of the support. We first give a quantified version of [**24**, theorem 1·5].

PROPOSITION C·2. *There is an integer $C_s \geq 1$ so that the following holds. Fix an integer $K \geq 1$. Suppose we have a degree s nilmanifold G/Γ with complexity M and $F\colon G/\Gamma \to \mathbb{C}$ with Lipschitz norm L. Furthermore suppose we have a polynomial sequence $g\colon \mathbb{Z} \to G$ and a smooth function $\phi\colon \mathbb{R}/\mathbb{Z} \to [0, 1]$ with $\mathrm{supp}\,(\phi) \in [(3K)^{-1}, (2K)^{-1}]$.*

There exists a degree s nilmanifold $\widetilde{G}/\widetilde{\Gamma}$ with a polynomial sequence $\tilde{g}\colon \mathbb{Z} \to \widetilde{G}$ such that

$$\widetilde{F}(\widetilde{g}(x)) = \phi(x/N)F(g(x \bmod N)),$$

where $x \bmod N$ is interpreted to lie in $\{0, \dots, N-1\}$. Furthermore we may assume:

- *(i) $\widetilde{G}/\widetilde{\Gamma}$ has complexity bounded by $O_K(1)^{O_s(1)} + O_s(M)$ and dimension at most $O_s(d)$;*
- *(ii) $\widetilde{F}$ has Lipschitz norm bounded by $L \cdot O_{\phi,K}(M^{O_s(d^{O_s(1)})})$;*
- *(iii) $\widetilde{g}(n)^{-1}\widetilde{g}(n+N) \in \widetilde{\Gamma}$ for all $n \in \mathbb{Z}$;*
- *(iv) if all Lie bracket structure constants of G/Γ are integers divisible by K then the Lie bracket structure constants of $\widetilde{G}/\widetilde{\Gamma}$ are contained in $K(C_s)^{-1}\mathbb{Z}$.*

Given this we deduce a variant of the U^4-inverse theorem from Theorem 1·3. This argument is essentially a combination of Lemma C·1, Proposition C·2, and the argument of [**24**, theorem 1·4].

THEOREM C·3. *Suppose we have an integer $K \geq 1$. There exists a constant $C = C_K > 0$ such that the following holds. Let N be prime, let $\delta > 0$, and suppose that $f\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is*

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

a 1-bounded function with

$$\|f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \geq \delta.$$

There exists a degree 3 nilsequence $F(g(n)\Gamma)$ such that it has dimension bounded by $(\log(1/\delta))^C$, complexity bounded by $O_K(1)$, such that F is 1-Lipschitz, and such that

$$\left|\mathbb{E}_{n\in[N]} f(n)\overline{F(g(n)\Gamma)}\right| \geq 1/\exp((\log(1/\delta))^C).$$

Furthermore, all Lie bracket structure constants of G (the c_{ijk} in Definition A·1) are integers divisible by K, $g(n)g(n+N)^{-1} \in \Gamma$ for all $n \in \mathbb{Z}$, and $g(0) = \mathrm{id}_G$.

Remark. In all our applications, we will take $K = 12$.

We deduce Theorem C·3 from Proposition C·2; this is essentially as in the proof of [**24**, theorem 1·4].

Proof of Theorem C·3. By adjusting constants appropriately, we may assume that N is larger than an absolute constant (and for small N apply the U^2-inverse theorem noting all Lie bracket structure constants for an abelian nilmanifold are 0).

We take a partition of unity of $\mathbb{R}/\mathbb{Z}$ denoted $\phi_1, \dots, \phi_{20K} \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ such that $\mathrm{supp}\,(\phi_j) \subseteq [j/(20K), j/(20K) + 1/(10K)] \bmod 1$. We abusively denote $\phi_i \colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ via $\phi_i(n) := \phi_i(n/N)$. Recalling that $\|\cdot\|_{U^4(\mathbb{Z}/N\mathbb{Z})}$ is a norm, we have

$$\|f\|_{U^4(\mathbb{Z}/N\mathbb{Z})} \leq \sum_{i=1}^{20K} \|\phi_i f\|_{U^4(\mathbb{Z}/N\mathbb{Z})}.$$

Thus there exists an index k such that $\|\phi_k f\|_{U^k(\mathbb{Z}/N\mathbb{Z})} \geq \delta/(20K)$. Applying a translation, we may assume that $\mathrm{supp}\,(\phi_k f)$ is contained in $[\lceil N/(3K)\rceil, \lfloor N/(2K)\rfloor] \bmod N$.

Applying Theorem 1·3, we have that there exists degree 3 nilsequence $F_1(g_1(n)\Gamma_1)$ on G/Γ_1 with complexity $O(1)$, dimension $(\log(1/\delta))^{O(1)}$, and Lipschitz constant 1 such that

$$|\mathbb{E}_{n\in[N]} \phi_k(n/N) f(n)\overline{F_1(g_1(n)\Gamma)}| \geq 1/\exp((\log(1/\delta))^{O(1)}).$$

By Lemma C·1, we may construct Γ_2, F_2 with Lie bracket structure constants for G/Γ_2 being integers divisible by $K \cdot C_s$ (where $s = 3$) such that

$$|\mathbb{E}_{n\in[N]} \phi_k(n/N) f(n)\overline{F_2(g(n)\Gamma_2)}| \geq 1/\exp((\log(1/\delta))^{O(1)}).$$

Furthermore G/Γ_2 has complexity bounded by $O(K)$ and F_2 has Lipschitz norm bounded by $(2K)^{O(d^{O(1)})}$. Next, we may write

$$\mathbb{E}_{n\in[N]} \phi_k(n/N) f(n)\overline{F_2(g_2(n)\Gamma_2)} = \mathbb{E}_{n\in[N]} f(n)\overline{\phi_k(n/N)F_2(g_2(n)\Gamma_2)}$$

and apply Proposition C·2 to $\phi(x/N)F_2(g_2(x)\Gamma_2)$. In particular, we obtain F_3, G_3, Γ_3, and $g_3(n)$ such that

$$\mathbb{E}_{n\in[N]} f(n)\overline{\phi_k(n/N)F_2(g_2(n)\Gamma_2)} = \mathbb{E}_{n\in\mathbb{Z}/N\mathbb{Z}} f(n)\overline{F_3(g_3(n)\Gamma)}$$

with F_3 having Lipschitz norm $(O_K(1))^{O(d^{O(1)})}$, G_3/Γ_3 having complexity $O_K(1)$, $g_3(n)$ being N-periodic, and all Lie structure constants divisible by K.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Note that we do not necessarily have $g_3(0) \neq \mathrm{id}_{G_3}$. This may be repaired by defining $g(n) = \{g_3(0)\}^{-1} g_3(n) g_3(0)^{-1} \{g_3(0)\}$ where $\{g_3(0)\}^{-1} g_3(0) = [g_3(0)] \in \Gamma_3$ and $\{g_3(0)\}$ has Mal'cev coordinates of the second kind bounded by $1/2$. Taking

$$F(x\Gamma_3) := F_3(\{g_3(0)\}x\Gamma_3)$$

we have

$$\mathbb{E}_{n\in\mathbb{Z}/N\mathbb{Z}} f(n)\overline{F_3(g_3(n)\Gamma_3)} = \mathbb{E}_{n\in\mathbb{Z}/N\mathbb{Z}} f(n)\overline{F(g(n)\Gamma_3)}.$$

As left-multiplication by bounded elements preserves the metric up to an admissible multiplicative factor ([**22**, lemma B·4]), we find that F has Lipschitz norm $(O_K(1))^{O(d^{O(1)})}$. Thus, rescaling F to have Lipschitz norm 1 will keep the correlation sufficiently large. Furthermore as left-multiplying a fixed element and right-multiplying a fixed element in Γ_3 does not change whether our polynomial sequence is N-periodic on G_3/Γ_3 this completes the proof.

We now prove Proposition C·2. We use the construction of Manners [**24**, theorem 1·5]; we modify the construction slightly to match the definition of filtration used in this paper.

Proof of Proposition C·2. *Step 1. Setup and constructing the proposed Mal'cev basis.* We consider the given degree s nilpotent Lie group G with filtration

$$G = G_0 = G_1 \geqslant G_2 \geqslant \cdots \geqslant G_s \geqslant \mathrm{Id}_G\,.$$

Let $\mathfrak{g}_j = \log_G(G_j)$.

For $i \geq 0$, let G^{+i} denote the prefiltration $G_i \geqslant G_{i+1} \geqslant G_{i+2} \geqslant \cdots \geqslant G_s \geqslant \mathrm{Id}_G$. Let $H_i = \mathrm{poly}(\mathbb{Z}, G^{+i})$, where we define polynomial sequences with respect to prefiltrations as in [**15**, definition B·1] with the prefiltration $\mathbb{Z} \geqslant \mathbb{Z} \geqslant \{0\} \geqslant \cdots$ on the domain $\mathbb{Z}$ (see also [**11**, p. 28], which includes the formal definition of prefiltrations).

By the Filtered Lazard–Leibman theorem of Green, Tao and Ziegler [**15**, proposition B·6], H_i are not only groups but also can be given the prefiltration

$$H_0 \geqslant H_1 \geqslant H_2 \geqslant \cdots \geqslant H_s \geqslant \mathrm{Id}_{H_0}\,.$$

This implies that one has the genuine filtration

$$H_1 = H_1 \geqslant H_2 \geqslant \cdots \geqslant H_s \geqslant \mathrm{Id}_{H_0}\,.$$

We now define the semidirect product $H_i \rtimes \mathbb{R}$ by defining a group operation $*$ via

$$(x \mapsto g(x), t) * (x \mapsto g'(x), t') = (x \mapsto g(x+t')g'(x), t+t').$$

Note that this is slightly different than that given in the work of Manners [**24**], since we take right-quotients by the cocompact subgroup, and in fact this is rather the opposite group of the semidirect product of the opposite groups (but we suppress such notational dependence). The semidirect product in question is embedding $t \in \mathbb{R}$ as the "shift by t" automorphism. Additionally, we abusively identify $g \in \mathrm{poly}(\mathbb{Z}, G^{+i})$ with a function $\mathbb{R} \to G^{+i}$ defined by using Taylor expansion and then allowing the argument to vary over reals instead of integers using the Lie group structure; this extension is unique due to a generalisation of the identity theorem.

One can easily prove that

$$H_0 \rtimes \mathbb{R} \geqslant H_1 \rtimes \mathbb{R} \geqslant H_2 \rtimes \{0\} \geqslant \cdots \geqslant H_s \rtimes \{0\} \geqslant \mathrm{Id}_{H_0 \rtimes \mathbb{R}}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

is a prefiltration; this follows immediately from the proof given in [17, proposition 14]. (The only nontrivial aspect involves the $\rtimes\mathbb{R}$ component, and it is not too hard to show that it suffices to check relevant properties at the level of the generators (id_{H_0}, t) since we have the above prefiltration on H_0.)

Let $X_1, \dots, X_d$ denote the Mal'cev basis on the Lie algebra associated to G (which is adapted to the s-step filtration given at the beginning). We construct the desired Mal'cev basis for $H_0 \rtimes \mathbb{R}$ as follows:

(i) Let $Z = (0, K)$;

(ii) let $Y_{i,k} = (n \mapsto \binom{n}{k} X_i, 0)$; if $X_i \in \mathfrak{g}_j \setminus \mathfrak{g}_{j+1}$ define the *level* of $Y_{i,k}$ to be $j - k$. We restrict attention to $Y_{i,k}$ of nonnegative level and level at most s. (Note that there are no contributions with $j = 0$.)

(iii) we order the Mal'cev basis left-to-right by increasing level; within level we order with increasing k, and then in increasing order of i. Z is inserted as the first element of level 1.

Remark. By Taylor expansion (see [15, lemma B·9]) we see that this is a Lie basis of the necessary group, filtered with respect to the desired prefiltration. Note that *a priori* the Lie algebra of poly $(\mathbb{Z}, G^{+0})$ is some abstract space, but we can utilise the embedding $H_0 = \mathrm{poly}\,(\mathbb{Z}, G^{+0}) \hookrightarrow G^{\mathbb{Z}}$ and the corresponding representation of Lie algebras to yield the above form of writing elements of the tangent space.

Our goal is now to verify that the discrete cocompact subgroup $\widetilde{\Gamma} = \mathrm{poly}\,(\mathbb{Z}, \Gamma) \rtimes (K\mathbb{Z})$ is the push-forward of $\mathbb{Z}^{\dim(\widetilde{G})}$ for Mal'cev coordinates of the second kind. This algorithmic proof will also immediately verify the remaining property of the Mal'cev basis (regarding rationality of the Lie bracket structure constants). After this, we will give the necessary nilsequence lifting construction and verify the necessary properties.

Step 2. Verifying Mal'cev properties modulo the semidirect product. We first prove that the $Y_{i,k}$ graded as above form a Mal'cev basis for

$$H_0 \geqslant H_1 \geqslant H_2 \geqslant \cdots \geqslant H_s \geqslant \mathrm{Id}_{H_0}$$

with discrete cocompact subgroup $\Gamma' = \mathrm{poly}\,(\mathbb{Z}, \Gamma)$. We proceed by induction upwards by step. Note that handling the final step is trivial as the group H_s is abelian.

Assume the inductive hypothesis that $H_{j+1} \cap \Gamma'$ is the image of the integer lattice in Mal'cev coordinates of the second kind (which necessarily uses only those basis elements of level at least $j + 1$). Using Taylor expansion (see [15, lemma B·9]), we may write an element $H_j \cap \Gamma'$ as

$$n \mapsto \prod_{i=0}^{s-j} g_{i+j}^{\binom{n}{i}}, \tag{C·1}$$

where $g_i \in G_{i+j} \cap \Gamma$. ($g_i \in \Gamma$ easily follows from our polynomial lying in $\Gamma' = \mathrm{poly}\,(\mathbb{Z}, \Gamma)$ sequence.)

Now, for our given value of j, we prove by an induction on ℓ that the set of polynomials of the form

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

$$n \mapsto \prod_{i=0}^{s-j} g_{i+j}^{\binom{n}{i}}$$

with $g_{i+j} \in G_{i+j+1} \cap \Gamma$ for $0 \le i \le \ell$ and $g_i \in G_{i+j} \cap \Gamma$ for $\ell + 1 \le i \le s - j$ is generated by the level $j+1$ elements and the largest $s - j - \ell$ "types" of level j elements of the Mal'cev basis for H_0. (Here "types" means in the sense of the parameter k we used to define the ordering on the $Y_{i,k}$ above.) The case $\ell = s - j$ is exactly the above inductive assumption regarding $H_{j+1} \cap \Gamma'$, and the case $\ell = 0$ is exactly what we need to complete the (outer) induction.

We now suppose that this is known for some $s - j \ge \ell \ge 1$, and wish to prove the result for $\ell - 1$. We may write

$$\left(n \mapsto \prod_{i=0}^{s-j} g_{i+j}^{\binom{n}{i}}\right) = \left(n \mapsto \prod_{i=0}^{\ell-1} g_{i+j}^{\binom{n}{i}}\right)\left(n \mapsto g_{\ell+j}^{\binom{n}{\ell}}\right)\left(n \mapsto \prod_{i=\ell+1}^{s-j} g_{i+j}^{\binom{n}{i}}\right) = ABC,$$

where A, B, C are defined in the obvious way. Note that $A \in H_{j+1} \cap \Gamma'$. We may write $g_{\ell+j} = \prod_r \exp(t_r X_r)$ for $t_r \in \mathbb{Z}$ and $X_r \in \mathfrak{g}_{\ell+j}$. Note that

$$D^{-1}B = \left(\prod_r \left(n \mapsto \exp(t_r X_r)^{\binom{n}{\ell}}\right)\right)^{-1}\left(n \mapsto g_{\ell+j}^{\binom{n}{\ell}}\right) \in H_j \cap \Gamma',$$

where D is defined in the obvious way. The Taylor expansion of the polynomial $(D^{-1}B)C$ (in a similar form to (C·1)) trivially has its coordinates of index $0, \dots, \ell - 1$ being the identity (plugging in the values $n = 0, 1, \dots, \ell - 1$). The ℓth Taylor coordinate is in $G_{\ell+j+1} \cap \Gamma'$, considering the value $n = \ell$. Thus

$$ABC = D(D^{-1}AD)(D^{-1}BC),$$

where $D^{-1}BC$ has the form

$$n \mapsto \prod_{i=\ell}^{s-j} {g_{i+j}}'^{\binom{n}{i}}$$

for ${g_{i+j}}' \in G_{i+j} \cap \Gamma$ and additionally ${g_{\ell+j}}' \in G_{\ell+j+1} \cap \Gamma$.

Note that $D^{-1}BC$ is of the correct form of Taylor series to apply the inductive hypothesis (since ${g_{\ell+j}}' \in G_{\ell+j+1}$), yielding

$$D^{-1}BC = \prod_{\mu \in M} \exp(t_\mu X_\mu) \cdot F = EF,$$

where M ranges over the $s - j - \ell$ "types" of level j elements of the Mal'cev basis of H_0, from $\ell + 1$ to $s - j$ inclusive, and is arranged in the correct order within the product E, and $F \in H_{j+1} \cap \Gamma'$. Finally, we may write

$$ABC = DE(E^{-1}(D^{-1}AD)E \cdot F),$$

and note that H_{j+1} is normal in H_0 due to the original prefiltration we established, so we find $E^{-1}(D^{-1}AD)E \in H_{j+1} \cap \Gamma'$. Thus we have

$$ABC = DEF',$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

where D, E form the necessary ordered product for the Mal'cev second kind representation, and $F' \in H_{j+1} \cap \Gamma'$. Finally, we recall the induction hypothesis for $H_{j+1} \cap \Gamma'$ to finish.

Step 3. Semidirect product and Lie bracket structure constants. We now handle the generator Z. Note that given an element $(p, t) \in \text{poly}(\mathbb{Z}, \Gamma) \rtimes (K\mathbb{Z})$, we may write

$$(p, t) = (\,\mathrm{id}_G, t) * (p, 0)$$

and therefore if Z was the highest element of the ordering (which would correspond to a situation where the filtration contains $H_1 \rtimes \{0\}$ instead of $H_1 \rtimes \mathbb{R}$) we would easily be done. Instead, though, it is inserted right before the Mal'cev basis components for H_1. Note however that

$$(\,\mathrm{id}_G, -K) * (p^{-1}, 0) * (\,\mathrm{id}_G, K) * (p, 0) = (p^{-1}, -K) * (p, K) = (x \mapsto p^{-1}(x + K)p(x), 0)$$

and recall that derivatives of polynomials in $H_j \cap \text{poly}(\mathbb{Z}, \Gamma)$ are in $H_{j+1} \cap \text{poly}(\mathbb{Z}, \Gamma)$ due to the definitions of the shifted filtrations on G. Therefore we may commute $(\,\mathrm{id}_G, t)$ across the product of the level 0 terms in the representation above, and we will introduce only terms of level 1 and higher (and they will lie in $\widetilde{\Gamma}$). Then using the Mal'cev representation for $H_1 \cap \Gamma'$ finishes. This completes our discussion that the generators listed form a Mal'cev basis.

We now compute the Lie bracket structure constants associated to the Mal'cev basis. We will compute all structure constants via the identity

$$[V, W] = \frac{d}{ds}\frac{d}{dt} \exp(sV) \exp(tW) \exp(-sV) \exp(-tW)\Big|_{u,v=0}$$

which holds for any Lie group and the associated Lie bracket.

We first compute the constants associated with Z. Note that

$$\begin{aligned}
[Y_{i,k}, Z] &= \frac{d}{dt_0}\frac{d}{dt_1}\left(n \mapsto \exp(t_0 X_i)^{\binom{n}{k}}, 0\right)(\,\mathrm{id}_G, Kt_1)\left(n \mapsto \exp(t_0 X_i)^{-\binom{n}{k}}, 0\right)(\,\mathrm{id}_G, -Kt_1)\Big|_{t_0,t_1=0} \\
&= \frac{d}{dt_0}\frac{d}{dt_1}\left(n \mapsto \exp(t_0 X_i)^{\binom{n}{k}-\binom{n-Kt_1}{k}}, 0\right)\Big|_{t_0,t_1=0} \\
&= \log_{H_0 \rtimes \mathbb{R}}\left(n \mapsto \exp(X_i)^{-\frac{d}{dt_1}\binom{n-Kt_1}{k}\big|_{t_0=0}}, 0\right) = \left(n \mapsto -\frac{d}{dt_1}\binom{n - Kt_1}{k} X_i\Big|_{t_1=0}, 0\right),
\end{aligned}$$

where the last line follows from considering the relationship between the differential structure on H_0 and on G (which allows us to "implicitly differentiate in the obvious way"). The coefficients of $(d/dt_1)\binom{n-Kt_1}{k}\big|_{t_1=0}$ are all integer multiples of $K^t/k!$ for $1 \le t \le k-1$. We may easily express this in terms of various $Y_{i,k'}$ for $k' < k$, and the structure constants clearly have the desired form, lying in $K(C_s)^{-1}\mathbb{Z}$ for an appropriate integer C_s.

We next compute the structure constants associated with two polynomials. Notice that

$$\begin{aligned}
&[Y_{i,k}, Y_{j,\ell}] \\
&= \frac{d}{dt_1}\frac{d}{dt_2}\left(n \mapsto \exp(t_1 X_i)^{\binom{n}{k}}, 0\right)\left(n \mapsto \exp(t_2 X_j)^{\binom{n}{\ell}}, 0\right)\left(n \mapsto \exp(t_1 X_i)^{-\binom{n}{k}}, 0\right) \\
&\quad \times \left(n \mapsto \exp(t_2 X_j)^{-\binom{n}{\ell}}, 0\right)\Big|_{t_1,t_2=0}
\end{aligned}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

$$= \frac{d}{dt_1}\frac{d}{dt_2}\left(n \mapsto \exp\left(t_1 t_2 \binom{n}{k}\binom{n}{\ell}[X_i, X_j]\right), 0\right)\Big|_{t_1,t_2=0}$$
$$= \left(n \mapsto \binom{n}{k}\binom{n}{\ell}[X_i, X_j], 0\right).$$

The second equality follows from using Baker–Campbell–Hausdorff to multiply the polynomial sequences (as pointwise functions, say) and collapse them, and then noting that we may discard terms with higher powers than t_1^1 and t_2^1. The only relevant remaining term is the $t_1 t_2$ which has the claimed form.

Note that as $\binom{n}{k}\binom{n}{\ell}$ is an integer-valued polynomial, by Polya's classification of integer polynomials [27], it follows that this may be written as an integral combination of binomial coefficients with coefficients bounded by $O_s(1)$. We obtain a representation in terms of various $Y_{i',k'}$ where $k' \leq k + \ell$ and $X_{i'}$ shows up in the expansion of $[X_i, X_j]$. We are done with this part of the proof, which includes verifying the well-definedness of various lifted filtrations, Mal'cev bases, and also the rationality properties of the Lie bracket structure constants.

Step 4. Lifting the nilsequence and checking quantitative dependences. We now actually define the desired function. We are given a polynomial sequence $g(x)$ and we define $q(x) = g(xN/K)$. We define

$$\widetilde{g}(x) = (q, 0) * (\mathrm{id}_G, Kx/N).$$

Note that $\widetilde{g}(x)$ is N-periodic as $\widetilde{g}(x+N) = \widetilde{g}(x) * (\mathrm{id}_G, K)$. We now define $\widetilde{F}$ on $\mathrm{poly}(\mathbb{R}, G)/\mathrm{poly}(\mathbb{Z}, \Gamma) \times [0, K)$ via

$$\widetilde{F}(x \cdot \mathrm{poly}(\mathbb{Z}, \Gamma), t) = \phi(t/K) \cdot F(x(0)\Gamma).$$

This extends uniquely to $H_0 \rtimes \mathbb{R}$ if we enforce right-$\widetilde{\Gamma}$-invariance, which yields

$$\widetilde{F}(x \cdot \mathrm{poly}(\mathbb{Z}, \Gamma), t) = \phi(\{t/K\}) \cdot F(x(-K\lfloor t/K \rfloor)\Gamma).$$

That $\widetilde{F}$ is appropriately Lipschitz and that we can then descend this construction to a proper filtration will be checked last.

The key point is the following computation:

$$\begin{aligned}\widetilde{F}(\widetilde{g}(x)) &= \widetilde{F}((q, 0) * (\mathrm{id}_G, Kx/N)) = \widetilde{F}((q(\cdot + Kx/N), Kx/N)) \\ &= \phi(\{x/N\})F(q(Kx/N - K\lfloor x/N \rfloor)) = \phi(\{x/N\})F(g(x - N\lfloor x/N \rfloor)) \\ &= \phi(\{x/N\})F(g(x \bmod N)).\end{aligned}$$

We now check that the function $|\widetilde{F}|$ is appropriately Lipschitz with respect to the Mal'cev basis specified. Since $\widetilde{F}$ is bounded by $\|F\|_\infty \|\phi\|_\infty$, note that for $x, y \in (H_0 \rtimes \mathbb{R})/\widetilde{\Gamma}$ if $d_{(H_0 \rtimes \mathbb{R})/\widetilde{\Gamma}}(x, y) \geq \varepsilon$ we have

$$\frac{\widetilde{F}(x) - \widetilde{F}(y)}{d_{(H_0 \rtimes \mathbb{R})/\widetilde{\Gamma}}(x, y)} \leq 2\varepsilon^{-1}\|F\|_\infty \|\phi\|_\infty.$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

Fix $\varepsilon \leq O_K(M)^{-O_s(d)^{O_s(1)}}$ to be chosen later. The diameter of $(H_0 \rtimes \mathbb{R})/(\operatorname{poly}(\mathbb{Z}, \Gamma) \rtimes K\mathbb{Z})$ is bounded by $O_K(M)^{O_s(d)^{O_s(1)}}$ by [22, lemma B·8][1], so there exist $\widetilde{x}, \widetilde{y}$ such that

$$d_{(H_0 \rtimes \mathbb{R})/\widetilde{\Gamma}}(x, y) = d_{(H_0 \rtimes \mathbb{R})}(\widetilde{x}, \widetilde{y}) \text{ and } d_{G,\widetilde{\mathcal{X}}}(\widetilde{x}, \mathrm{Id}_G) + d_{G,\widetilde{\mathcal{X}}}(\widetilde{y}, \mathrm{Id}_G) \leq O_K(M)^{O_s(d^{O_s(1)})}.$$

Furthermore we may chose the fundamental representatives $\widetilde{x}$ and $\widetilde{y}$ such that the second coordinate is in $[0,K)$ (by multiplying by an appropriate element of $\mathrm{Id}_G \rtimes (K\mathbb{Z})$ on the right and noting that the metric is right-invariant). Note that if the second coordinate of the representative $\widetilde{x}$ is outside the range $[1/4, 2/3]$, the assumption on the support of ϕ guarantees that $\widetilde{F}(\widetilde{x}\Gamma) = \widetilde{F}(\widetilde{y}\Gamma) = 0$ and therefore verifying the Lipschitz constant in this case is trivial. (This uses that ε is sufficiently small.)

Else we denote $\widetilde{x} = (x^*, x')$ with $x^* \in H_0$ and $x' \in [1/4, 2/3]$ and $\widetilde{y}$ analogously. We now have

$$\begin{aligned}\left|\widetilde{F}(\widetilde{x}\widetilde{\Gamma}) - \widetilde{F}(\widetilde{y}\widetilde{\Gamma})\right| &= \left|\phi(x'/K)F(x^*(0)\Gamma) - \phi(y'/K)F(y^*(0)\Gamma)\right| \\ &\leq \|\phi\|_\infty \left|F(x^*(0)\Gamma) - F(y^*(0)\Gamma)\right| + \left|\phi(x'/K) - \phi(y'/K)\right| \|F\|_\infty.\end{aligned}$$

Let ψ denote the second Mal'cev coordinates with respect to the basis specified for $(H_0 \rtimes \mathbb{R})/\widetilde{\Gamma}$. Note that x',y' are controlled only by Z in the Mal'cev basis and thus we can bound the second term by

$$\left|\phi(x^*/K) - \phi(y^*/K)\right| \lesssim_K \|\psi'\|_\infty \cdot \left|\psi(\widetilde{x}) - \psi(\widetilde{y})\right|.$$

To bound the first term, note that $d_{G/\Gamma}(x^*(0)\Gamma, y^*(0)\Gamma) \leq d_G(x^*(0), y^*(0))$.

Let $\widetilde{x} = \exp((x'/K)Z) \prod_{k,\ell} \exp(x_{k,\ell} Y_{k,\ell})$ and analogously for $\widetilde{y}$ where the product over $Y_{k,\ell}$ is taken in the ordering specified for the Mal'cev basis. We have

$$x^*(0) = \prod_{k,\ell} \exp\left(\binom{0}{\ell} x_{k,\ell} X_k\right) \text{ and } y^*(0) = \prod_{k,\ell} \exp\left(\binom{0}{\ell} y_{k,\ell} X_k\right).$$

Note that the only terms which contribute to the above product are $Y_{k,0} = (n \mapsto \binom{n}{0} X_k) = (n \mapsto X_k)$. If $X_k \in \mathfrak{g}_j \setminus \mathfrak{g}_{j+1}$, this has level precisely j and thus the product (removing terms which are the identity) are in the correct order for the Mal'cev basis on G.

Therefore we have

$$\begin{aligned}d_G(x^*(0), y^*(0)) &\leq O_K(M)^{O_s(d^{O_s(1)})} \max_k |x_{k,0} - y_{k,0}| \\ &\leq O_K(M)^{O_s(d^{O_s(1)})} \left(\max_{k,\ell} |x_{k,\ell} - y_{k,\ell}| + |x'/K - y'/K|\right) \\ &\leq O_K(M)^{O_s(d^{O_s(1)})} d_{H_0 \rtimes \mathbb{R}}(\widetilde{x}, \widetilde{y}).\end{aligned}$$

The first line follows from [22, lemma B·3] applied on the Mal'cev basis $\mathcal{X}$ of G and the second inequality is trivial. The final inequality follows from [22, lemma B·3] and that $\max_{k,\ell} |x_{k,0} - y_{k,0}| + |x'/K - y'/K|$ corresponds (up to constants) to the L^∞ distance in the Mal'cev coordinates when Z is placed first in the order and one may return to the original

[1] We have proven that the constructed basis is a filtered basis for the Lie algebra. Separating by level and then by "type", we see that the filtered basis has nesting property of degree $\leq 2s^2$. [22, lemma B·8] is proven under only the assumption that the basis for the Lie algebra exhibits this nesting property.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

distance at the cost of $O_K(M)^{O_s(d)^{O_s(1)}}$ by [22, lemma B·3]. (Shifting Z to the first position is clearly a 1-rational change of basis.)

Step 5. Reducing to a proper filtration. We now (finally) perform the last reduction to place our polynomial sequence on a proper filtration. By inspection we have

$$\widetilde{g}(x) = (h_0^*)(h_1^*)^{\binom{x}{1}}$$

with $h_i^* \in H_i \rtimes \mathbb{R}$ for $i \leq 1$. Let $h_0^* = \{h_0^*\}[h_0^*]$ such that $[h_0^*] \in \widetilde{\Gamma}$ and $\|\psi(\{h_0^*\})\| \leq 1$.

For $\widetilde{p} \in H_1 \rtimes \mathbb{R}$ define $F^*(\widetilde{p}\widetilde{\Gamma}) = \widetilde{F}(\{h_0^*\}\widetilde{p}\widetilde{\Gamma})$, and note that

$$\begin{aligned}\widetilde{F}(\widetilde{g}(x)\widetilde{\Gamma}) &= \widetilde{F}(\{h_0^*\}[h_0^*](h_1^*)^{\binom{x}{1}}\widetilde{\Gamma}) = \widetilde{F}(\{h_0^*\}([h_0^*]h_1^*[h_0^*]^{-1})^{\binom{x}{1}}[h_0^*]\widetilde{\Gamma})\\ &= F^*(([h_0^*]h_1^*[h_0^*]^{-1})^{\binom{x}{1}}\widetilde{\Gamma}).\end{aligned}$$

Since left-multiplication by bounded elements is suitably Lipschitz and since $H_1 \rtimes \mathbb{R}$ is normal in $H_0 \rtimes \mathbb{R}$, we have that $([h_0^*]h_1^*[h_0^*]^{-1})^{\binom{x}{1}}$ is a polynomial sequence with respect to the filtration $H_1 \rtimes \mathbb{R} = H_1 \rtimes \mathbb{R} \geqslant H_2 \rtimes \{0\} \geqslant \cdots \geqslant H_s \rtimes \{0\} \geqslant \mathrm{Id}_{H_1 \rtimes \mathbb{R}}$.

Note that pre- or post-multiplication by a fixed constant does not change N-periodicity. Furthermore note that Z and all level 1 and higher elements in the order given by removing all level 0 elements gives a valid Mal'cev basis with respect to the proper filtration

$$H_1 \rtimes \mathbb{R} = H_1 \rtimes \mathbb{R} \geqslant H_2 \rtimes \{0\} \geqslant \cdots \geqslant H_s \rtimes \{0\} \geqslant \mathrm{Id}_{H_1 \rtimes \mathbb{R}}\,.$$

Thus the polynomial sequence $([h_0]h_1^*[h_0]^{-1})^{\binom{x}{1}}$ with respect to the function $\widetilde{F}(\,\cdot\,)$ on nilmanifold $(H_1 \rtimes \mathbb{R})/(\widetilde{\Gamma} \cap (H_1 \rtimes \mathbb{R}))$ with this filtration and Mal'cev basis provides the desired construction.

Appendix D. Deferred lemmas

We first prove the necessary partition of unity result on a nilmanifold.

Proof of Lemma 2·4. The key trick is to consider a "smoothed sum" of fundamental domains and perform a partition of unity. Let δ be a constant to be specified later and let $f(x) \geq 0$ be a smooth bump function on $\mathbb{R}$ with $\mathrm{supp}\,(f) \in [-1/4, 1/4]$, $\int_{\mathbb{R}} f(x)dx = 1$, and $\|f\|_\infty \leq O(1)$. Let H_j be smooth functions indexed by $j \in S$ of size $O(1/\delta)$ which are nonnegative with $\mathrm{supp}\,(H_j) \in [j\delta, (j+2)\delta]$, $\|H_j\|_\infty \leq O(1)$, and $\sum_{j\in S} H_j(x)$ equal to 1 for $|x| \leq 3/2$ and 0 for $|x| \geq 2$.

For each $g \in G$ and $\beta \in \mathbb{R}$, there exists a unique $\gamma_i \in \Gamma$ such that $\psi(g\gamma_i) \in [-\beta, 1-\beta)^d$ (see e.g. [22, lemma B·6]). Define for $\beta \in \mathbb{R}$ the function $T_\beta : G \to \Gamma$ such that $\psi(gT_\beta(g)) \in [-\beta, 1-\beta)^d$; note that this function suffers from discontinuities at the boundaries of $[-\beta, 1-\beta)^d$.

Note that for all $g \in G$ we have

$$\begin{aligned}1 &= \int_{\beta\in\mathbb{R}} f(\beta)d\beta = \int_{\beta\in\mathbb{R}} \prod_{i=1}^{d} \Big(\sum_{j\in S} H_j(\psi_i(gT_\beta(g)))\Big) f(\beta)d\beta\\ &= \sum_{j_1,\dots,j_d\in S} \int_{\beta\in\mathbb{R}} \prod_{k=1}^{d} H_{j_k}(\psi_k(gT_\beta(g)))f(\beta)d\beta.\end{aligned}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

The functions in our collection will be indexed by $j_1, \dots, j_d$ and are defined by

$$\tau_{j_1,\dots,j_d}(g\Gamma) = \int_{\beta\in\mathbb{R}} \prod_{k=1}^{d} H_{j_k}(\psi_k(gT_\beta(g)))f(\beta)d\beta.$$

This is a well-defined function since $gT_\beta(g) = g\gamma T_\beta(g\gamma)$ for all $\gamma \in \Gamma$. Furthermore note that $\|\tau_{j_1,\dots,j_d}\|_\infty \le \exp(O(d))$. We will use these as our partition of unity.

Let $L = O(M)^{O_s(d^{O_s(1)})}$ with implicit constants chosen sufficiently large. To check the Lipschitz property, it suffices to consider $x\Gamma, y\Gamma$ such that $d_{G/\Gamma}(x\Gamma, y\Gamma) \le L^{-1}$. We may find $\widetilde{x}, \widetilde{y} \in G$ such that

$$d_{G/\Gamma}(x\Gamma, y\Gamma) = d_G(\widetilde{x}, \widetilde{y}),\ \psi(\widetilde{x}), \psi(\widetilde{y}) \in [-1, 2)^d,\ \text{and } d_G(\widetilde{x}, \mathrm{id}_G) + d_G(\widetilde{y}, \mathrm{id}_G) \le L.$$

If $d_G(\gamma, \mathrm{id}_G) \ge L^3$ then note that for $g \in G$ with $d_G(g, \mathrm{id}_G) \le L$ we have

$$\begin{aligned} d_G(g\gamma, \mathrm{id}_G) &\ge d_G(g\gamma, g) - d_G(g, \mathrm{id}_G) \ge L^{-1}d_G(\gamma, \mathrm{id}_G) - d_G(g, \mathrm{id}_G) \\ &\ge L^{-1}(L^3) - L \ge L^2/2 > 3, \end{aligned}$$

using quantitative approximate left-invariance of the metric ([22, lemma B·4]). Therefore we have $d_G(T_\beta(\widetilde{x}), \mathrm{id}_G) + d_G(T_\beta(\widetilde{y}), \mathrm{id}_G) \le 2L^3$ for $|\beta| \le 2$. Let $\widetilde{\Gamma} \subseteq \Gamma$ be the set of all $\widetilde{\gamma} \in \Gamma$ such that $d_G(\widetilde{\gamma}, \mathrm{id}_G) \le 2L^3$ and note that $|\widetilde{\Gamma}| \le L^{O_s(d^{O_s(1)})}$.

We now show the desired properties of a partition of unity. Let β be drawn from the probability distribution $f(\beta)$. Let $\eta := d_G(\widetilde{x}, \widetilde{y})$ and note that for $\gamma \in \widetilde{\Gamma}$ we have $\|\psi(\widetilde{x}\gamma) - \psi(\widetilde{y}\gamma)\| \le \eta \cdot L^{O_s(d^{O_s(1)})} =: \eta'$. Therefore

$$\begin{aligned} \mathbb{P}_\beta[T_\beta(\widetilde{x}) \ne T_\beta(\widetilde{y})] &\le \sum_{\gamma\in\widetilde{\Gamma}} \mathbb{P}_\beta[\psi(\widetilde{x}\gamma) \in [-\beta, 1-\beta)^d \cap \psi(\widetilde{y}\widetilde{x}^{-1}(\widetilde{x}\gamma)) \notin [-\beta, 1-\beta)^d] \\ &\le \sum_{\gamma\in\widetilde{\Gamma}} \sum_{i=1}^{d} \mathbb{P}_\beta[\psi_i(\widetilde{x}\gamma) \in [-\beta-\eta', -\beta] \cup [1-\beta-\eta', 1-\beta)] \le \eta\delta^{-1}L^{O_s(d^{O_s(1)})}. \end{aligned}$$

Thus

$$\begin{aligned} \left|\tau_{j_1,\dots,j_d}(\widetilde{x}\Gamma) - \tau_{j_1,\dots,j_d}(\widetilde{y}\Gamma)\right| &\le \int_{\beta\in\mathbb{R}} \left| \prod_{k=1}^{d} H_{j_k}(\psi(\widetilde{x}T_\beta(\widetilde{x}))) - \prod_{k=1}^{d} H_{j_k}(\psi(\widetilde{y}T_\beta(\widetilde{x})))\right| f(\beta)d\beta \\ &\quad + \delta^{-O(d)} \cdot \mathbb{P}_\beta[T_\beta(\widetilde{x}) \ne T_\beta(\widetilde{y})] \\ &\le \delta^{-O(d)}\eta L^{O_s(d^{O_s(1)})}, \end{aligned}$$

which demonstrates the necessary Lipschitz bound.

Finally we check the supports; note that in order for $\tau_{j_1,\dots,j_d}(g\Gamma)$ to be nonzero, there exists $\gamma \in \Gamma$ such that $\psi(g\gamma) \in \prod_{k=1}^{d} [j_k\delta, (j_k+2)\delta]$. Let $\widetilde{g}$ be such that $\psi(\widetilde{g}) = ((j_k+1)\delta)_{1\le k\le d} \in [-2, 2]^d$. We have $\|\psi(g\gamma) - \psi(\widetilde{g})\| \le \delta$. Thus, taking the fundamental domain for G/Γ which is $\prod_{k=1}^{d} [(j_k+1)\delta - 1/2, (j_k+1)\delta + 1/2)$ (with respect to ψ) we have that the support is contained within a δ-ball of the center. Note as $\psi_{\exp} \circ \psi^{-1}$ is a polynomial with

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

coefficients bounded by $O(M)^{O_s(d^{O_s(1)})}$ and degree $O_s(1)$ ([22, lemma B·1]) we have the desired result taking $\delta = \varepsilon M^{-O_s(d^{O_s(1)})}$ (and appropriately modifying the definition of L).

We next need that sufficiently divisible structure constants prove that $\psi_{\exp}^{-1}(\Gamma)$.

Proof of Lemma 2·5. First, consider any element of Γ which can be written $\gamma = \prod_{i=1}^{d} \exp(t_i X_i)$ for $t_i \in \mathbb{Z}$. We inductively prove that $\prod_{i=j}^{d} \exp(t_i X_i)$ is in $\psi_{\exp}^{-1}(\mathbb{Z}^d)$. Suppose we have shown that $\prod_{i=j+1}^{d} \exp(t_i X_i) = \exp(\sum_{i=j+1}^{d} u_i X_i)$ with $u_i \in \mathbb{Z}$, for some $1 \leq j \leq d-1$ (the base case $j = d-1$ is obvious). Then

$$\prod_{i=j}^{d} \exp(t_i X_i) = \exp(t_j X_j) \exp\left(\sum_{i=j+1}^{d} u_i X_i\right) = \exp\left(t_j X_j + \sum_{i=j+1}^{d} u_i X_i + \cdots\right),$$

where the remainder is a finite list of commutators of $t_j X_j$ and $\sum_{i=j+1}^{d} u_i X_i$ coming from the Baker–Campbell–Hausdorff formula. As the coefficients are rationals with denominators bounded by C_s, it follows immediately from the condition on divisibility of Lie bracket structure constants that the inductive step holds. We deduce $\psi_{\exp}(\Gamma) \subseteq \mathbb{Z}^d$.

For the reverse inclusion, we also use induction. We show for all $1 \leq j \leq d-1$ that for any $u_i \in \mathbb{Z}$, $\exp(\sum_{i=j}^{d} u_i X_i)$ can be written in the form $\prod_{i=j}^{d} \exp(t_i X_i)$ for $t_i \in \mathbb{Z}$. Suppose we have the result for $j+1$, so that we have $\exp(\sum_{i=j+1}^{d} u_i X_i) = \prod_{i=j+1}^{d} \exp(t_i X_i)$. Then

$$\exp(-u_j X_j) \exp\left(\sum_{i=j}^{d} u_i X_i\right) = \exp\left(\sum_{i=j+1}^{d} u_i' X_i\right)$$

for $u_i' \in \mathbb{Z}$ using the Baker–Campbell–Hausdorff formula, the divisibility assumption, and the filtered nature of the Mal'cev basis. (The term on the far left is tailored to cancel the X_j part.) Again, the inductive step immediately follows. This demonstrates $\psi_{\exp}^{-1}(\mathbb{Z}^d) \subseteq \Gamma$, and we are done.

We next prove a comparison estimate between the distance in Mal'cev coordinates of the first kind and the associated torus metric.

Proof of Lemma 2·6. Note that

$$d_{G/\Gamma}(x\Gamma, y\Gamma) \leq d_G(x, y).$$

As $\psi \circ \psi_{\exp}^{-1}$ is a polynomial of degree $O_s(1)$ with coefficients bounded by $O(M)^{O_s(d^{O_s(1)})}$ ([22, lemma B·1]) we have that

$$\|\psi(x)\|_\infty \leq (LM)^{O_s(d^{O_s(1)})} \text{ and } \|\psi(x) - \psi(y)\|_\infty \leq \varepsilon (LM)^{O_s(d^{O_s(1)})}.$$

The result then follows from [22, lemma B·3].

We now give the short proof of Lemma 2·8.

Proof of Lemma 2·8. It is evidently enough to show it for two functions f_1, f_2. Additionally, recall the product rule

$$\partial_h(fg) = (\partial_h f) \cdot (T_h g) + f \cdot (\partial_h g), \tag{D·1}$$

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

where $T_h g(x) = g(x+h)$ defines the translation operator. Note that all translation operators commute with all other operations such as discrete derivatives and products, and other translations. Also, this is valid at a point x so long as $x, x+h$ are in the domains of f, g.

Let $s = d_1 + d_2$. Suppose we are given $x, h_1, \dots, h_{s+1}$ with $x + \{0, h_1\} + \cdots + \{0, h_{s+1}\} \subseteq S$. We have, by iterating (D·1),

$$\partial_{h_1,\dots,h_{s+1}}(f_1 f_2)(x) = \sum_{T_1 \sqcup T_2 = [s+1]} (\partial_{(h_j)_{j\in T_1}} f_1)(x) \cdot T_{\sum_{j\in T_1} h_j}(\partial_{(h_j)_{j\in T_2}} f_2)(x). \tag{D·2}$$

This is seen to be valid since every element of the cube $x + \{0, h_1\} + \cdots + \{0, h_{s+1}\}$ is in the domain S. Now, for every disjoint partition $T_1 \sqcup T_2 = [s+1]$ we have either $|T_1| \geq d_1 + 1$ or $|T_2| \geq d_2 + 1$. By the given condition of f_j being locally degree $d_j + 1$, we see one of the two terms in (D·2) is always 0. The result follows.

Acknowledgements. We thank Zach Hunter for various minor corrections.

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.
[2] T. F. BLOOM. A quantitative improvement for Roth's theorem on arithmetic progressions. *J. Lond. Math. Soc.* (2) **93** (2016), 643–663.
[3] T. F. BLOOM and O. SISASK. Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions. ArXiv:2007.03528.
[4] T. F. BLOOM and O. SISASK. An improvement to the Kelley-Meka bounds on three-term arithmetic progressions. ArXiv:2309.02353.
[5] J. BOURGAIN. On triples in arithmetic progression, *Geom. Funct. Anal.* **9** (1999), 968–984.
[6] JEAN BOURGAIN. Roth's theorem on progressions revisited. *J. Anal. Math.* **104** (2008), 155–192.
[7] W. T. GOWERS. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* **8** (1998), 529–551.
[8] W. T. GOWERS. Arithmetic progressions in sparse sets. Current developments in mathematics, 2000 (Int. Press, Somerville, MA, 2001), pp. 149–196.
[9] W. T. GOWERS. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), 465–588.
[10] B. GREEN and T. TAO. Quadratic uniformity of the Möbius function. *Ann. Inst. Fourier (Grenoble)* **58** (2008), 1863–1935.
[11] B. GREEN and T. TAO. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.* (2) **175** (2012), 465–540.
[12] B. GREEN, T. TAO, and T. ZIEGLER. An inverse theorem for the Gowers U^4-norm, *Glasg. Math. J.* **53** (2011), 1–50.
[13] B. GREEN and T. TAO. *New bounds for Szemerédi's theorem. II. A new bound for* $r_4(N)$. Analytic number theory (Cambridge University Press, Cambridge, 2009), pp. 180–204.
[14] B. GREEN and T. TAO. New bounds for Szemerédi's theorem, III: a polylogarithmic bound for $r_4(N)$. *Mathematika* **63** (2017), 944–1040.
[15] BEN GREEN, T. TAO and T. ZIEGLER. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math.* (2) **176** (2012), 1231–1372.
[16] D. R. HEATH-BROWN. Integer sets containing no arithmetic progressions. *J. London Math. Soc.* (2) **35** (1987), 385–394.
[17] B. HOST and B. KRA. *Nilpotent structures in ergodic theory*, Math. Surv. Monogr. **236** (American Mathematical Society, Providence, RI, 2018).
[18] Z. KELLEY and R. MEKA. Strong bounds for 3-progressions, ArXiv:2302.05537.
[19] M. LAZARD. Sur les groupes nilpotents et les anneaux de Lie. *Ann. Sci. École Norm. Sup.* (3) **71** (1954), 101–190.
[20] A. LEIBMAN. Polynomial sequences in groups. *J. Algebra* **201** (1998), 189–206.
[21] A. LEIBMAN. Polynomial mappings of groups. *Israel J. Math.* **129** (2002), 29–60.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press

[22] J. LENG. Efficient equidistribution of nilsequences. ArXiv:2312.10772.
[23] J. LENG. Efficient equidistribution of periodic nilsequences and applications. ArXiv:2306.13820.
[24] F. MANNERS. Periodic nilsequences and inverse theorems on cyclic groups. ArXiv:1404.7742.
[25] S. PELUSE and S. PRENDIVILLE. A polylogarithmic bound in the nonlinear Roth theorem. *Int. Math. Res. Not. IMRN* (2022), 5658–5684.
[26] S. PELUSE, A. SAH and M. SAWHNEY. Effective bounds for Roth's theorem with shifted square common difference. ArXiv:2309.08359.
[27] GEORG PÓLYA. Über ganzwertige ganze funktionen. *Rendiconti del Circolo Matematico di Palermo* (1884-1940) **40** (1915), 1–16.
[28] K. F. ROTH. On certain sets of integers. II. *J. London Math. Soc.* **29** (1954), 20–26.
[29] T. SANDERS. On the Bogolyubov-Ruzsa lemma. *Anal. PDE* **5** (2012), 627–655.
[30] TOM SANDERS. On Roth's theorem on progressions. *Ann. of Math.* (2) **174** (2011), 619–636.
[31] TOM SANDERS. On certain other sets of integers. *J. Anal. Math.* **116** (2012), 53–82.
[32] WOLFGANG M. SCHMIDT. *Small fractional parts of polynomials*, Regional Conference Series in Mathematics, no. 32 (American Mathematical Society, Providence, RI, 1977).
[33] E. SZEMERÉDI. *On sets of integers containing no four elements in arithmetic progression*. Number Theory (Colloq., János Bolyai Math. Soc., Debrecen, 1968), Colloq. Math. Soc. János Bolyai, vol. 2 (North-Holland, Amsterdam-London, 1970), pp. 197–204.
[34] E. SZEMERÉDI. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27** (1975), 199–245.
[35] E. SZEMERÉDI. Integer sets containing no arithmetic progressions. *Acta Math. Hungar.* **56** (1990), 155–158.
[36] T. TAO. *Higher order Fourier analysis*, Grad. Stud. Math. **142** (American Mathematical Society, Providence, RI, 2012).
[37] T. TAO and J. TERÄVÄINEN. Quantitative bounds for Gowers uniformity of the Möbius and von Mangoldt functions. ArXiv:2107.02158.

https://doi.org/10.1017/S0305004124000264 Published online by Cambridge University Press