COMPOSITIO MATHEMATICA

A non-flag arithmetic regularity lemma and counting lemma

Daniel Altman

Compositio Math. **161** (2025), 635–679.

doi: 10.1112/S0010437X24007693

FOUNDATION
COMPOSITIO
MATHEMATICA



LONDON
MATHEMATICAL
SOCIETY
EST. 1865



https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Compositio Math. 161 (2025) 635–679
doi:10.1112/S0010437X24007693

A non-flag arithmetic regularity lemma and counting lemma

Daniel Altman

Abstract

Green and Tao's arithmetic regularity lemma and counting lemma together apply to systems of linear forms which satisfy a particular algebraic criterion known as the 'flag condition'. We give an arithmetic regularity lemma and counting lemma which apply to all systems of linear forms.

Contents

1 Introduction **636**
1.1 Context and summary . . . 636
1.2 Results . . . 636
1.3 On the flag condition . . . 639
1.4 Outline of the paper . . . 639
2 Examples, discussion of methods and qualitative results **640**
2.1 Non-flag examples . . . 640
2.2 Moving to the Lie algebra . . . 642
2.3 Notions of irrationality . . . 643
2.4 The counting lemma . . . 646
2.5 Factorisation . . . 647
3 Description of $\mathfrak{g}^{\Psi}(S_\bullet)$ and comparison with $\log G^{\Psi}$ **650**
4 The counting lemma **652**
5 The factorisation theorem for linear irrationality **657**
5.1 Additive decomposition for linear irrationality . . . 657
5.2 Multiplicative factorisation for strong irrationality . . . 660
6 The strongly irrational arithmetic regularity lemma **663**
7 Applications **666**
7.1 Recovering the flag counting lemma . . . 666
7.2 Another proof of the Gowers–Wolf conjecture . . . 667
Appendix A. Details on polynomial sequences in Lie groups and Lie algebras **668**
A.1 Smoothness norms . . . 668
A.2 Equidistribution in the Lie algebra . . . 669

Received 6 January 2023, accepted in final form 24 September 2024, published online 2 July 2025.
2020 Mathematics Subject Classification 11B30 (primary).
Keywords: regularity lemma; nilsequence; counting lemma.
© The Author(s), 2025. The publishing rights in this article are licensed to Foundation Compositio Mathematica under an exclusive licence.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

A.3 Rational and small polynomial sequences 673
A.4 Factorising for filtration irrationality 673
A.5 On the Lipschitz property 675
Appendix B. Some quantitative integer linear algebra **677**
B.1 Hermite normal form 677
Acknowledgements **678**
References **678**

1. Introduction

1.1 Context and summary

The use of 'regularity lemmas' in combinatorics dates back to Szeméredi's proof of his eponymous theorem on arithmetic progressions in dense sets of integers [Sze75]. Szemerédi's regularity lemma and developments thereof have since become powerful tools in combinatorics which have been brought to bear on a far wider range of problems than the study of arithmetic progressions in sets of integers. Where Szemerédi's regularity lemma is a structural result about graphs, Green introduced an arithmetic analogue of this brand of regularity results in [Gre05], and used it there to solve some problems in additive combinatorics. From the perspective of applications to additive patterns in the integers, the state of the art is Green and Tao's arithmetic regularity lemma [GT10].

Green and Tao's arithmetic regularity lemma is phrased as a decomposition result for bounded functions which consequently may be written as the sum of a nilsequence, a Gowers uniform function, and an error. Green and Tao [GT10] demonstrated a number of applications for which the contributions of the uniform and error components are negligible and so the problem is essentially reduced to an analysis of nilsequences. Complementing the arithmetic regularity lemma in [GT10] is an arithmetic counting lemma which gives a general formula (via equidistribution on nilmanifolds) for computing averages on linear patterns in the nilsequence component.

It has come to light (see [GT20, Tao20]) that this counting lemma [GT10, Theorem 1.11] applies only to systems of linear forms which satisfy the so-called *flag condition* (see § 1.3 for definitions), whence the arithmetic regularity and counting lemma strategy may be exercised only on this particular set of linear patterns. Further to this, there are examples (see § 2.1) which demonstrate that the output of the arithmetic regularity lemma [GT10, Theorem 1.2] is too coarse to determine the distribution of non-flag linear patterns in any prospective counting lemma.

The goal of this paper is to give an arithmetic regularity lemma and complementary counting lemma which are able to handle all linear patterns. The main results are a stronger arithmetic regularity lemma, Theorem 1.4, which feeds into a new counting lemma, Theorem 1.2, which itself is able to handle all sets of linear forms. We remark that our description of the nilmanifold on which we obtain equidistribution in the counting lemma is less explicit than what is worked out for the Leibman group in [GT10, § 3] (though this is partially remedied in § 3 here). Nonetheless, the broadstrokes 'arithmetic regularity and counting' strategy for solving problems in additive combinatorics can now be applied to all linear patterns of finite complexity.

We also give a couple of applications; these are discussed in the next subsection.

1.2 Results

Throughout this paper, any Lie group G that we discuss will be assumed to be nilpotent, connected and simply connected without further mention. Furthermore, Γ will be used to denote a

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

discrete cocompact subgroup of G. Given a flag system of linear forms Ψ, a filtered Lie group $G_\bullet$ and a polynomial sequence $g : \mathbb{Z} \to G$ which is suitably 'irrational' with respect to this filtration, Green and Tao's counting lemma determines (quantitatively) the distribution of the tuple $g^\Psi(\boldsymbol{n}) := ((g(\psi_1(\boldsymbol{n})), \dots, g(\psi_t(\boldsymbol{n})))\Gamma^t)_{\boldsymbol{n}\in\mathbb{Z}^D}$ in G^t/Γ^t from the data $\{G_\bullet, \Psi\}$. Unfortunately, these data are insufficient to determine the distribution of g^Ψ in the general, non-flag setting in the sense that there exist a non-flag system of linear forms Ψ and distinct polynomial sequences g, $\tilde{g}$, both of which are irrational in the same $G_\bullet$, such that g^Ψ distributes differently to $\tilde{g}^\Psi$; see Examples 2.2 and 2.3. Instead of dealing with a filtered Lie group $G_\bullet$, we consider a sequence of subspaces $S_\bullet$ of the Lie algebra $\mathfrak{g}$ of G. One may define a notion of irrationality, which we will call 'linear irrationality', with respect to a sequence of subspaces such that for all linearly irrational polynomial sequences g with respect to $S_\bullet$, one may determine the distribution of g^Ψ from the data $\{S_\bullet, \Psi\}$. It will prove convenient to describe the nilmanifold on which g^Ψ equidistributes as a Lie subalgebra of $\mathfrak{g}^t$. We denote this Lie subalgebra by $\mathfrak{g}^\Psi(S_\bullet)$, define it in terms of Ψ and $S_\bullet$ in Definition 2.14 and then compute it more explicitly in § 3.

Working in the Lie algebra, we denote group multiplication by $*$.[1]

DEFINITION 1.1. Define the operation $* : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by $x * y := \log(\exp(x)\exp(y))$.

We denote group multiplication in this way partially to avoid the temptation to accidentally distribute over addition. By the Baker–Campbell–Hausdorff formula,[2] we have an explicit, finite expansion

$$x * y = x + y + \tfrac{1}{2}[x, y] + \tfrac{1}{12}[x, [x, y]] + \cdots .$$

We now state our counting lemma. Although there are a number of requisite definitions which have not yet been provided (rational basis, Definition 4.1; rational subspace, Definition 2.12; complexity of a rational subspace, Definition 4.6; linear irrationality, Definitions 4.7 and 4.8; quantitative equidistribution, Definitions A.9 and 4.14; Hermite basis, Definition B.2), we hope that the reader may obtain a rough appreciation of the content of the result. It is proven as Theorem 4.16 in § 4.

THEOREM 1.2 (Counting lemma). *Let $N, A, M \geqslant 2$ and $d, s, t \geqslant 1$ be integers. Let $\Psi = (\psi_1, \dots, \psi_t)$ be a collection of linear forms each mapping $\mathbb{Z}^D$ to $\mathbb{Z}$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra with rational basis $\mathcal{X}$, dimension d and rational structure constants of height at most M. Let $S_\bullet$ be a sequence of rational subspaces of complexity M in $\mathfrak{g}$. Suppose that $p(n)$ is an (A, N)-linearly irrational polynomial sequence of degree s in $S_\bullet$. Then there are constants $0 < c_1, c_2 = O_{d,s,D,t,\Psi}(1)$ such that the polynomial sequence $p^\Psi(\boldsymbol{x}) := (p(\psi_1(\boldsymbol{x})), \dots, p(\psi_t(\boldsymbol{x})))_{\boldsymbol{x}\in[N]^D}$ is $O(M^{c_1}/A^{c_2})$-equidistributed in $\mathfrak{g}^\Psi(S_\bullet)$ with respect to $\mathcal{X}^\Psi$, the Hermite basis for $\mathfrak{g}^\Psi(S_\bullet)$ in $(\mathfrak{g}^t, \mathcal{X}^t)$.*

The next task is to obtain a factorisation result for the notion of linear irrationality. The following theorem says that any polynomial sequence may be written as the product of an 'error' sequence, a sequence which is linearly irrational with respect to some sequence of subspaces $S_\bullet$, and a rational sequence. Furthermore, the factorisation process is such that if we begin with a polynomial sequence which is irrational in the Green–Tao sense (we henceforth call this 'filtration irrational'), then the linearly irrational factorisation is still (though slightly less) Green–Tao (i.e. filtration) irrational. There are a number of additional definitions (filtration,

[1] We will not need much Lie theory in this paper, but the reader may wish to recall Lie's third theorem (see, for example, [Kir08, Theorem 3.42 and Corollary 3.43]), that every real Lie algebra is the Lie algebra of a connected, simply connected Lie group (unique up to isomorphism).

[2] See [CG90, pp. 11–13 and Theorem 1.2.1].

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Definition 2.4; polynomial sequence adapted to a filtration, Definition 2.5; quantitative rationality, Definition 5.2; quantitative smallness, Definition 5.3; quantitative filtration irrationality, Definitions 5.6 and 5.7) required for a rigorous interpretation.

Theorem 1.3 (Factorisation of polynomial sequences). *Let $N, A, M \geq 2$ and $d, s, t \geq 1$ be integers such that $AM > d$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra with filtration $\mathfrak{g}_\bullet$, rational basis $\mathcal{X}$, dimension d and rational structure constants of height at most M. Let p be a polynomial sequence adapted to $\mathfrak{g}_\bullet$ and sequence of subspaces $S_\bullet \leq \mathfrak{g}_\bullet$. Finally, insist that $S_\bullet, \mathfrak{g}_\bullet$ are of complexity $A^{O(1)}$ in $\mathfrak{g}$, where here and henceforth $O(1)$ terms may depend on d, s. Then we may write $p = e * p' * r$ where e, p', r are polynomial sequences in $\mathfrak{g}$ adapted to $\mathfrak{g}_\bullet$, p' is (A, N)-linearly irrational in a sequence of subspaces $S'_\bullet$ of complexity $A^{O(1)}$, r is $(AM)^{O(1)}$-rational, and e is $((AM)^{O(1)}, N)$-small.*

This theorem is proven in § 5 as Theorem 5.10 (where we actually prove a bit more about the factorisation; see Theorem 5.10). Our final goal is to provide an arithmetic regularity lemma which produces a virtual nilsequence[3] which is both filtration irrational and linearly irrational; we call such a nilsequence 'strongly irrational' (cf. Definition 5.11). Though filtration irrationality is not strictly necessary to determine distribution, it is useful in applications and, furthermore, means that our arithmetic regularity lemma is more compatible with, and explicitly strengthens, what is done in [GT10].

Theorem 1.4 (Strongly irrational arithmetic regularity lemma). *Let $f : [N] \to [0, 1]$, let $s \geq 1$, let $\varepsilon > 0$, and let $\mathcal{F} : \mathbb{R}^+ \to \mathbb{R}^+$ be a growth function. Then there exists a quantity $M = O_{s,\varepsilon,\mathcal{F}}(1)$ and a decomposition*

$$f = f_{\mathrm{nil}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

of f into functions $f_{\mathrm{nil}}, f_{\mathrm{unf}} : [N] \to [-1, 1]$ such that:

(i) *(f_{nil} structured) f_{nil} is a degree $\leq s$, $(\mathcal{F}(M), N)$-strongly irrational virtual polynomial nilsequence of complexity $\leq M$;*
(ii) *(f_{sml} small) $\|f_{\mathrm{sml}}\|_{L^2[N]} \leq \varepsilon$;*
(iii) *(f_{unf} very uniform) $\|f_{\mathrm{nil}}\|_{U^{s+1}[N]} \leq 1/\mathcal{F}(M)$;*
(iv) *(nonnegativity) f_{nil} and f_{sml} take values in $[0, 1]$.*

We also provide a couple of applications. The first is the modest task of recovering (with polynomial bounds) the flag counting lemma for strongly irrational nilsequences.

Theorem 1.5 (Equidistribution in the Leibman group for flag systems). *Let $N, A, M \geq 2$. Let $\Psi = (\psi_1, \dots, \psi_t)$ be a flag system of linear forms, each mapping $\mathbb{Z}^D$ to $\mathbb{Z}$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra of dimension d. Let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$ which passes through a degree s filtration $\mathfrak{g}_\bullet$ and with respect to which $\mathfrak{g}$ has rational structure constants of height at most M. Let $S_\bullet \leq \mathfrak{g}_\bullet$ be a sequence of subspaces of complexity at most M. If p is (A, N)-strongly irrational in $(\mathfrak{g}_\bullet, S_\bullet)$, then there are constants $0 < c_1, c_2 = O_{d,s,D,t,\Psi}(1)$ such that p^Ψ is $O(M^{c_1}/A^{c_2})$-equidistributed on $[N]^D$ in $\log G^\Psi$, where G^Ψ is the Leibman group for $G_\bullet, \Psi$.*

This is proven in § 7.1, where we show that under appropriate assumptions, $\mathfrak{g}^\Psi(S_\bullet) = \log G^\Psi$.

[3]The terminology 'virtual nilsequence' is due to Green–Tao and, roughly speaking, refers to a function f for which one may partition $[N]$ into a bounded number of subprogressions of equal step such that f is a nilsequence on each such subprogression. We refer the reader to [GT10, Definition 1.9] for the full definition.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Next, in § 7.2 we give another resolution of a conjecture of Gowers and Wolf [GW10, Conjecture 2.5] which was recovered recently in the generality of all linear forms in [Alt22] by employing a somewhat specialised linear algebraic trick. The point of § 7.2 is to provide a less *ad hoc* argument. Indeed, Green and Tao's proof in the flag setting [GT10, Theorem 1.13] may be followed after doing some modest algebraic computations (Proposition 7.5) and employing the strongly irrational arithmetic regularity lemma, Theorem 1.4, and counting lemma, Theorem 1.2, in place of the original irrational arithmetic regularity lemma [GT10, Theorem 1.2] and counting lemma [GT10, Theorem 1.11], respectively.

THEOREM 1.6 (Gowers–Wolf conjecture). *Let $N \geq 2$. Let $\Psi = (\psi_1, \dots, \psi_t)$ be a collection of linear forms each mapping $\mathbb{Z}^D$ to $\mathbb{Z}$, and let $s \geq 1$ be an integer such that the polynomials $\psi_1^{s+1}, \dots, \psi_t^{s+1}$ are linearly independent. For $i = 1, \dots t$, let $f_i : [-N, N] \to \mathbb{C}$ be functions bounded in magnitude by 1 (and defined to be zero outside of $[-N, N]$). For all $\varepsilon > 0$ there exists $\delta > 0$ such that if $\min_i \|f_i\|_{U^{s+1}[-N,N]} \leq \delta$, then*

$$\left| \mathbb{E}_{\boldsymbol{x} \in [-N,N]^D} \prod_{i=1}^{t} f_i(\psi_i(\boldsymbol{x})) \right| \leq \varepsilon.$$

1.3 On the flag condition

Let $\Psi = (\psi_1, \dots, \psi_t)$ where each ψ_i maps $\mathbb{Z}^D$ to $\mathbb{Z}$. For $j \in \mathbb{N}^+$, let Ψ^j denote the vector of functions $(\psi_1^j, \dots, \psi_t^j)$, where multiplication in each coordinate is conducted pointwise. One may write $\Psi^j(\boldsymbol{x}) = \sum_{m \in \mathcal{M}} v_{m,j} m(\boldsymbol{x})$, where $v_{m,j} \in \mathbb{Z}^t$, and $m \in \mathcal{M}$ is a 'monomial map' which maps $\boldsymbol{x} = (x_1, \dots, x_D)$ to some monomial in the variables $\{x_1, \dots, x_D\}$. To elucidate the notation, we compute with the specific example

$$\Psi(x, y) := (\psi_1(x, y), \psi_2(x, y), \psi_3(x, y), \psi_4(x, y)) := (y, 2x + 2y, x + 3y, x),$$

that

$$\Psi^1(x, y) = (0, 2, 1, 1)x + (1, 2, 3, 0)y,$$

and

$$\Psi^2(x, y) = (0, 4, 1, 1)x^2 + (1, 4, 9, 0)y^2 + (0, 8, 6, 0)xy.$$

For $i \in \mathbb{N}^+$ we define the vector spaces $V_i = \operatorname{span}_{\mathbb{R}}\{v_{m,i}\}_{m \in \mathcal{M}}$ and let $V_\Psi = V_1$. In fact, it is not difficult to see that $V_i = V_\Psi^i$ where V_Ψ^i is the smallest vector space containing i-fold products of elements of V_Ψ, where multiplication is conducted coordinatewise.

A system of linear forms Ψ satisfies the *flag condition* if $V_\Psi^i \leq V_\Psi^j$ whenever $i \leq j$. One may check that the system of linear forms in the example above does not satisfy the flag condition (i.e. *is not flag*) because V_Ψ is not contained in V_Ψ^2.

Throughout this article we use the notation V_Ψ to denote the vector subspace of $\mathbb{R}^t$ spanned by the image of a given sequence of linear forms Ψ.

1.4 Outline of the paper

In § 2 we begin by providing examples which motivate the transition from Lie group to Lie algebra. Then we work out the qualitative theory in the Lie algebra setting, giving a qualitative counting lemma, Theorem 2.16, and qualitative factorisation theorem, Proposition 2.28. We encourage a reader who wishes to understand the ideas behind Theorems 1.2 and 1.3 to at least glean from this section the main algebraic strategy before progressing to the quantitative sections that follow. The theory is substantially cleaner here and we hope that this section convinces the reader that this is the right setup in general. This section may also be of independent interest.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

The qualitative counting lemma in § 2 describes a particular Lie algebra $\mathfrak{g}^{\Psi}(S_\bullet)$ on which we obtain equidistribution. This is the same Lie algebra that appears in the statement of Theorem 1.2. In § 3 we give more explicit and tenable algebraic descriptions of this Lie algebra.

In § 4, we prove our quantitative counting lemma, Theorem 1.2. In § 5, we prove our quantitative factorisation of polynomial sequences for linear irrationality, Theorem 1.3. In § 6, we prove our quantitative 'strongly irrational' arithmetic regularity lemma, Theorem 1.4. Finally, in § 7 we give applications as discussed previously.

2. Examples, discussion of methods and qualitative results

In this section we begin with some examples which demonstrate key phenomena in the qualitative setting. This motivates our methods and setup and, in particular, how they deviate from those in [GT10]. Then we work out the infinitary distribution of polynomial sequences on linear patterns. This includes most of the key ideas and phenomena that will appear in the quantitative setting (which is what we will ultimately use for our applications in § 7), but is not muddied by quantitative bounds or the technical baggage introduced in their pursuit. This section will also be interspersed with examples. We hope that this section is edifying in preparation for the more arduous task of reading §§ 4 and 5.

Recall that a filtration on a nilpotent, connected, simply connected Lie group G is a decreasing sequence of simply connected Lie subgroups $G = G_0 = G_1 \geq G_2 \geq \cdots$ with the property that $[G_i, G_j] \leq G_{i+j}$ for all i, j. A polynomial sequence $g(n) := \prod_i g_i^{\binom{n}{i}}$ on G is said to be adapted to a filtered Lie group $G_\bullet$ if $g_i \in G_i$ for all i. As a general disclaimer, in this section we may use some standard terminology from this area without explicit introduction. We refer the reader to [GT10] or [GT12] for the development of the basic theory used here (on such an expedition, readers really need only acquaint themselves with basic qualitative definitions to make this rather informal section legible; we provide our own definitions when doing things rigorously later on).

2.1 Non-flag examples

In the flag case, Green and Tao's counting lemma computes averages of the form $\mathbb{E}_{\boldsymbol{x}} \prod_{i=1}^{t} \phi(\psi_i(\boldsymbol{x}))$ where $\phi = f \circ g$ is a suitably 'irrational' nilsequence in some filtered nilmanifold $(G/\Gamma, G_\bullet)$. Roughly speaking, if $G_\bullet$ is the 'correct' filtered nilmanifold for the polynomial sequence g (i.e. if g cannot be written as a polynomial sequence on cosets of some smaller filtered subnilmanifold), then the distribution of $g^{\Psi}(\boldsymbol{x}) := (g(\psi_1(\boldsymbol{x})), \ldots, g(\psi_t(\boldsymbol{x})))$ in G^t/Γ^t may be determined from the data $\{G_\bullet, \Psi\}$.

Definition 2.1 (The Leibman group). Let $G_\bullet$ be a filtered group and let Ψ be a family of linear forms. Then define $G^{\Psi} := \langle g_i^{v_i}, g_i \in G_i, v_i \in V_{\Psi}^i \rangle$, where V_Ψ is as defined in § 1.3.

The flag counting lemma says that if g is irrational in $G_\bullet$, then g^{Ψ} equidistributes in $G^{\Psi}/G^{\Psi} \cap \Gamma^t$, and so the average $\mathbb{E}_{\boldsymbol{x}} \prod_{i=1}^{t} \phi(\psi_i(\boldsymbol{x}))$ is asymptotically equal to the integral $\int_{G^{\Psi}/G^{\Psi}\cap\Gamma^t} f^{\otimes t}$.

When moving to non-flag systems of forms, not only does g^{Ψ} not necessarily equidistribute on the Leibman group, but the notion of irrationality originally considered by Green and Tao does not capture enough information about the polynomial sequence g to determine its distribution. The following example appears in [GT20] and [Tao20].

Example 2.2. Let $\Psi : \mathbb{Z}^2 \to \mathbb{Z}^4$ be our non-flag system of linear forms from above:

$$\Psi(x, y) = (\psi_1(x, y), \psi_2(x, y), \psi_3(x, y), \psi_4(x, y)) = (y, 2x + 2y, x + 3y, x).$$

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Then, in $\mathbb{R}^4$,

$$V_\Psi = \operatorname{span}\{(0,2,1,1),(1,2,3,0)\} = \ker\{(3,0,-1,1),(2,-1,0,2)\},$$
$$V_\Psi^2 = \operatorname{span}\{(0,4,1,1),(1,4,9,0),(0,4,3,0)\} = \ker\{(24,3,-4,-8)\},$$

and $V_\Psi^3 = \mathbb{R}^4$.

Let $\alpha, \beta \in \mathbb{R}$ be such that $\{1, \alpha, \beta\}$ satisfy no nontrivial linear relations over $\mathbb{Q}$. Let $f : \mathbb{R}^2/\mathbb{Z}^2 \to \mathbb{C}$ be smooth and let $g(n) = (\alpha n, \beta n^2)$ determine a polynomial sequence on $\mathbb{R}^2$. Then (suppressing in our notation the quotient $\mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2$) we have that $\phi(n) := f(\alpha n, \beta n^2)$ is an irrational nilsequence in the filtered Lie group $\mathbb{R}^2$ with filtration $G_1 = \mathbb{R}^2$, $G_2 = 0 \times \mathbb{R}$, essentially because the linear part $(\alpha, 0)n$ of $g(n)$ equidistributes in $G_1/G_2\Gamma \cong (\mathbb{R}/\mathbb{Z}) \times 0$, and the quadratic part $(0, \beta)n^2$ equidistributes in $G_2/G_2 \cap \Gamma \cong 0 \times \mathbb{R}/\mathbb{Z}$. (We refer the reader to [GT10, Definition A.6] for a formal definition of the irrationality of a polynomial sequence.)

We would like our counting lemma to determine the average

$$\mathbb{E}_{x,y}(f \circ g)(\psi_1(x,y))(f \circ g)(\psi_2(x,y))(f \circ g)(\psi_3(x,y))(f \circ g)(\psi_4(x,y)). \tag{1}$$

To do so we find where $g^\Psi(x,y) := (g(\psi_1(x,y)), \dots, g(\psi_4(x,y)))$ distributes within in $(\mathbb{R}^2/\mathbb{Z}^2)^4$ whereupon the average (1) may be computed by the relatively easy task of integrating over a subtorus.

Here the Leibman group is

$$\{(a,b) : a \in V_\Psi, b \in V_\Psi + V_\Psi^2\},$$

suitably interpreted as a subgroup of $(\mathbb{R}^2)^4$. On the other hand,

$$g^\Psi(x,y) = ((0,2,1,1)\alpha x + (1,2,3,0)\alpha y, (0,4,1,1)\beta x^2 + (1,4,9,0)\beta y^2 + 2(0,4,3,0)\beta xy)$$

is constrained to the following subgroup of G^Ψ:

$$\{(a,b) : a \in V_\Psi, b \in V_\Psi^2\}.$$

Since $V_\Psi^2 \subsetneq V_\Psi + V_\Psi^2$ we have that g^Ψ takes values in a strict subgroup of the Leibman group and so will certainly not equidistribute there. (Note, on the other hand, that if the flag condition was satisfied by the system of linear forms, i.e. $V_\Psi \leq V_\Psi^2$ so $V_\Psi^2 = V_\Psi + V_\Psi^2$, then the subgroup is equal to the Leibman group and indeed Green and Tao's counting lemma holds.)

Now define $\tilde{g}(n) = (\alpha n, \gamma n + \beta n^2)$, where $\{1, \alpha, \beta, \gamma\}$ is a $\mathbb{Q}$-linearly independent set. Then $\tilde{g}$ is irrational in the same filtration as g and in this instance $\tilde{g}^\Psi$ is not constrained to the same subgroup of G^Ψ as g^Ψ was. In fact, $\tilde{g}^\Psi$ equidistributes in G^Ψ. The upshot is that in the non-flag setting, the irrationality of a polynomial sequence in $G_\bullet$ is not enough to determine the distribution of g^Ψ.

In light of this example, one might conjecture in general that the subgroup of G^t on which g^Ψ equidistributes may be expressed in terms of the sequence of subgroups on which the distinct degree parts of g equidistribute. In our example with $g(n) = (\alpha, 0)n + (0, \beta)n^2$ and $\tilde{g}(n) = (\alpha, \gamma)n + (0, \beta)n^2$, we have $(\alpha, 0)n$ equidistributing in $\mathbb{R} \times 0$ whereas $(\alpha, \gamma)n$ equidistributes in $\mathbb{R}^2$ and indeed this precisely accounts for how the distribution of g^Ψ differs from that of $\tilde{g}^\Psi$. It turns out that when the nilmanifold in question is abelian, one can indeed compute the distribution of g^Ψ in G^t/Γ^t if the distribution of the coefficients of g are known.[4] However,

[4]The justification for this claim should not yet be obvious. The reader may wish to return to this statement after seeing the definition of linear irrationality Definition 2.13 and our qualitative counting lemma, Theorem 2.16. Then the proof of this claim follows from the fact that in the abelian case, the distribution of the coefficients of a polynomial sequence determines the sequence of rational subspaces in which it is linearly irrational, which by Theorem 2.16 determines the distribution of g^Ψ.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

this is not the case in the nonabelian setting, as the following example of linear sequences on the Heisenberg group demonstrates.

Example 2.3. Let G be the Heisenberg group with filtration determined by $G_2 = [G, G]$. Let Ψ, ψ_i, V be as above and let Γ be the obvious lattice on G. Let $v_1 = (1, 2, 3, 0)$ and $v_2 = (0, 2, 1, 1)$ so $V_\Psi = \mathrm{span}(v_1, v_2)$. Let

$$g_1 = \begin{pmatrix} 1 & \alpha & \gamma \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{g}_1 = \begin{pmatrix} 1 & \alpha & \frac{1}{2}\alpha\beta \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{pmatrix},$$

where $1, \alpha, \beta, \alpha\beta, \gamma$ satisfy no nontrivial linear relations over $\mathbb{Q}$. Then both $g(n) = g_1^n$ and $\tilde{g}(n) = \tilde{g}_1^n$ equidistribute in G/Γ; however, as we now show, g^Ψ distributes differently to $\tilde{g}^\Psi$ in G^4/Γ^4. (Furthermore, both polynomial sequences are irrational by Definition A.6 of [GT10].)

Both g^Ψ and $\tilde{g}^\Psi$ take values in the Leibman group G^Ψ. Define η on $G^\Psi \leq G^4$ by

$$\eta(h_1, h_2, h_3, h_4) = w \cdot (h_{1,2}, h_{2,2}, h_{3,2}, h_{4,2}),$$

where

$$h_i = \begin{pmatrix} 1 & h_{i,0} & h_{i,2} \\ 0 & 1 & h_{i,1} \\ 0 & 0 & 1 \end{pmatrix} \in G$$

for each i and $w = (24, 3, -4, -8) \in V_\Psi^{2\perp}$. An easy but mundane computation (using crucially that $w \in V_\Psi^{2\perp}$) reveals that, in fact, η is a homomorphism from G^Ψ to $\mathbb{R}$. We compute (with slight abuse of notation where we use a single matrix with vector-valued entries to denote a 4-tuple of matrices)

$$g^\Psi(x, y) = g_1^{\Psi(x,y)} = g_1^{xv_1 + yv_2} = \begin{pmatrix} 1 & \alpha(xv_1 + yv_2) & \gamma(xv_1 + yv_2) + \binom{xv_1 + yv_2}{2}\alpha\beta \\ 0 & 1 & \beta(xv_1 + yv_2) \\ 0 & 0 & 1 \end{pmatrix}$$

and so

$$\begin{aligned} \eta(g^\Psi(x, y)) = {} & w \cdot v_1 \left(\gamma - \tfrac{1}{2}\alpha\beta\right) x + w \cdot v_2 \left(\gamma - \tfrac{1}{2}\alpha\beta\right) y \\ & + w \cdot v_1^2 \left(\tfrac{1}{2}\alpha\beta\right) x^2 + 2w \cdot v_1 v_2 \left(\tfrac{1}{2}\alpha\beta\right) xy + w \cdot v_2^2 \left(\tfrac{1}{2}\alpha\beta\right) y^2. \end{aligned}$$

Since $w \in V_\Psi^{2\perp}$, we have that $\eta(g^\Psi(x, y)) = w \cdot v_1(\gamma - \frac{1}{2}\alpha\beta)x + w \cdot v_2(\gamma - \frac{1}{2}\alpha\beta)y$, where $w \cdot v_1, w \cdot v_2 \neq 0$. Of course, completing the same computation with $\tilde{g}$ yields the same result with $\frac{1}{2}\alpha\beta$ in place of γ and so, in particular, $\eta(\tilde{g}^\Psi(x, y)) = 0$. That is, $\tilde{g}^\Psi$ is constrained to the subgroup of G^Ψ given by $\ker \eta$ whereas g^Ψ is not. (As an application of our qualitative counting lemma, Theorem 2.16, we show in Example 2.19 that g^Ψ in fact equidistributes on $G^\Psi/G^\Psi \cap \Gamma^4$ and $\tilde{g}^\Psi$ equidistributes on $\ker \eta / \ker \eta \cap \Gamma^4$.)

2.2 Moving to the Lie algebra

Let us now interpret the previous example in the Lie algebra $\mathfrak{g}$ of the Heisenberg group G. Recall that $\log : G \to \mathfrak{g}$ is a bijection given by

$$\log \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & a & c - \frac{1}{2}ab \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}.$$

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

For convenience and brevity we use the coordinates $(a, b, c - \frac{1}{2}ab)$ to denote the latter matrix in the previous sentence. The Lie bracket in these coordinates is

$$[(a, b, c), (x, y, z)] = (0, 0, ay - bx). \tag{2}$$

The algebra $\mathfrak{g}$ inherits a filtration $\mathfrak{g}_0 = \mathfrak{g}_1 = \mathfrak{g}$ and $\mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}] = \{(0, 0, \mathbb{R})\}$. In these coordinates, we have that $p(n) := \log g(n) = (\alpha, \beta, \gamma - \frac{1}{2}\alpha\beta)n$ and $\tilde{p}(n) := \log \tilde{g}(n) = (\alpha, \beta, 0)n$.

Let us digress slightly and consider a general polynomial sequence p on some Lie algebra $\mathfrak{g}$ with linear forms Ψ whereupon we are interested in determining where p^Ψ equidistributes within $\mathfrak{g}^t$.

Definition 2.4 (Filtration in the Lie algebra). A *filtration* on $\mathfrak{g}$ of *degree* at most s is a sequence $(\mathfrak{g}_i)_{i=1}^s$ of Lie subalgebras of $\mathfrak{g}$ such that $\mathfrak{g}_0 = \mathfrak{g}_1 \geq \mathfrak{g}_2 \geq \cdots \geq \mathfrak{g}_s$ and such that for all i, j, $[\mathfrak{g}_i, \mathfrak{g}_j] \leq \mathfrak{g}_{i+j}$ where we define $\mathfrak{g}_{s+1} = \mathfrak{g}_{s+2} = \cdots = 0$.

Definition 2.5 (Polynomial sequence adapted to a filtration). A polynomial sequence defined by $p(n) := \sum_{i=1}^s a_i n^i$ in $\mathfrak{g}$ is *adapted to* $\mathfrak{g}_\bullet$ if $a_i \in \mathfrak{g}_i$ for all i.

It will prove convenient to deal with $\mathfrak{g} \otimes \mathbb{R}^t$ rather than $\mathfrak{g}^t$, where the map from the former to the latter on elementary tensors is given by $\varphi : a \otimes v \mapsto (v_1 a, v_2 a, \ldots, v_t a)$. This gives a vector space isomorphism and we may obtain a Lie algebra isomorphism by endowing $\mathfrak{g} \otimes \mathbb{R}^t$ with a Lie bracket as follows:

$$\begin{aligned}
[a \otimes v, b \otimes w] &:= \varphi^{-1}([(v_1 a, \ldots, v_t a), (w_1 b, \ldots, w_t b)]) \\
&= \varphi^{-1}(([v_1 a, w_1 b], \ldots, [v_t a, w_t b])) \\
&= \varphi^{-1}((v_1 w_1 [a, b], \ldots, v_t w_t [a, b])) \\
&= [a, b] \otimes vw,
\end{aligned} \tag{3}$$

where multiplication in the second entry is conducted coordinatewise. With this setup, and henceforth suppressing our φ notation and implicitly identifying $\mathfrak{g}^t$ with $\mathfrak{g} \otimes \mathbb{R}^t$, if $p(n) = \sum a_i n^i$, then $p^\Psi(\boldsymbol{x}) := (p(\psi_1(\boldsymbol{x})), \ldots, p(\psi_t(\boldsymbol{x}))) = \sum a_i \otimes \Psi^i(\boldsymbol{x}) \in \mathfrak{g} \otimes \mathbb{R}^t$.

Returning to our example, $\tilde{p}^\Psi$ is contained in the subspace $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi$ at all inputs. We claim that $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2$ is the smallest Lie algebra containing the subspace $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi$. Indeed, we first observe that $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2$ is closed under the Lie bracket and so is certainly a Lie algebra which contains $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi$. On the other hand, for $a, b, x, y \in \mathbb{R}$ and $v, w \in V_\Psi$,

$$[(a, b, 0) \otimes v, (x, y, 0) \otimes w] := [(a, b, 0), (x, y, 0)] \otimes vw = (0, 0, ay - bx) \otimes vw,$$

and elements of this form span $\mathfrak{g}_2 \otimes V_\Psi^2$.

We show later as an application of our qualitative counting lemma, Theorem 2.16, in Example 2.19 that in fact $\tilde{p}^\Psi$ equidistributes in $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2$ (modulo an appropriate lattice). Importantly, $(\mathbb{R}, \mathbb{R}, 0) \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2$ is a proper subalgebra of $\mathfrak{g} \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2 = \log G^\Psi$ and indeed p^Ψ equidistributes in $\log G^\Psi$ (again, modulo an appropriate lattice).

2.3 Notions of irrationality

In this subsection we restrict our attention to polynomial sequences p with $p(0) = 0 \in \mathfrak{g}$. In particular, notions of irrationality are only defined for polynomial sequences satisfying this property.

We begin by introducing some definitions, in particular that of a *rational structure* on $\mathfrak{g}$. A *rational structure* on $\mathfrak{g}$ is a rational Lie subalgebra $\mathfrak{g}_\mathbb{Q}$ of $\mathfrak{g}$ such that the vector space isomorphism $\mathfrak{g} \cong \mathfrak{g}_\mathbb{Q} \otimes \mathbb{R}$ holds. Then, let $\mathfrak{g}_\mathbb{Q}^* \subset \mathfrak{g}^*$ be the linear maps on $\mathfrak{g}$ which map $\mathfrak{g}_\mathbb{Q}$ to $\mathbb{Q}$. In the coordinates

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

in our Heisenberg example above, we have implicitly used the rational structure $\mathbb{Q}^3 \leq \mathbb{R}^3$. In what follows, let $\mathrm{Hom}_{\mathbb{Q}}(\mathfrak{g}, \mathbb{R})$ denote the Lie algebra homomorphisms from $\mathfrak{g}$ to $\mathbb{R}$ which map $\mathfrak{g}_{\mathbb{Q}}$ to $\mathbb{Q}$; we refer to elements of $\mathrm{Hom}_{\mathbb{Q}}(\mathfrak{g}, \mathbb{R})$ as *horizontal characters* on $\mathfrak{g}$.[5] Finally, many statements in this document are made with respect to a rational structure $\mathfrak{g}_{\mathbb{Q}}$ on $\mathfrak{g}$; at times the underlying rational structure may not be referenced explicitly.

Definition 2.6 (Qualitative additive irrationality). A polynomial sequence p in $\mathfrak{g}$ is *additively irrational* in $\mathfrak{g}$ with respect to $\mathfrak{g}_{\mathbb{Q}}$ if for all nontrivial $\eta \in \mathrm{Hom}_{\mathbb{Q}}(\mathfrak{g}, \mathbb{R})$ we have that $(\eta \circ p)(\boldsymbol{x}) \notin \mathbb{Q}$ for some $\boldsymbol{x}$.

In the following we invoke Lie's third theorem to glean the simply connected Lie group G from the Lie algebra $\mathfrak{g}$.

Definition 2.7 (Qualitative equidistribution with respect to a rational structure). A polynomial sequence p in $\mathfrak{g}$ *equidistributes* with respect to a rational structure $\mathfrak{g}_{\mathbb{Q}}$ if $\exp \circ p$ equidistributes in G/Γ (as per Definition A.4), where Γ is *any* lattice in G such that $\mathrm{span}_{\mathbb{Q}} \log \Gamma = \mathfrak{g}_{\mathbb{Q}}$.

It is not at all *a priori* clear that the above is well defined (in particular, that such a lattice exists and, furthermore, that the property of equidistribution is independent of the choice of such lattice). This is addressed in Appendix A.2.

The following theorem of Leibman is our main tool for determining qualitative equidistribution. It says that the above two definitions are equivalent.

Theorem 2.8 (Leibman's criterion for qualitative equidistribution in the Lie algebra). *Let D be a positive integer, let $\mathfrak{g}$ be a real, finite-dimensional, nilpotent Lie algebra and let $\mathfrak{g}_{\mathbb{Q}}$ be a rational structure on $\mathfrak{g}$. Let $(p(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D}$ be a polynomial sequence on $\mathfrak{g}$. Then $(p(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D}$ equidistributes in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$ if and only if p is additively irrational in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$.*

We deduce the above statement from Leibman's original theorem [Lei05] in Theorem A.8.

For appropriately (i.e. *linearly*) irrational p, we show in the proof of Theorem 2.16 that p^{Ψ} is additively irrational in an explicit subalgebra of $\mathfrak{g}^t$, whereupon we have determined its distribution by Theorem 2.8. Before defining linear irrationality, we transfer Green and Tao's notion of irrationality [GT10, Definition A.6] to the Lie algebra. For the purposes of our current discussion, we give the following qualitative translation in which one removes the complexity requirement on 'i-horizontal characters' [GT10, Definition A.5] and takes the limit as $N \to \infty$.

Definition 2.9 (Qualitative Green–Tao irrationality). An element $g \in G_i$ is i-irrational if for all nontrivial continuous homomorphisms $\eta : G_i \to \mathbb{R}$ which vanish on G_{i+1} and $[G_j, G_{i-j}]$ for all j, and which map the lattice Γ to $\mathbb{Q}$, we have $\eta(g) \notin \mathbb{Q}$.

(A direct translation works with functions that map Γ to $\mathbb{Z}$, but the above is equivalent and more natural in the qualitative setting.) By basic Lie theory one may interpret this in the Lie algebra with filtration $\mathfrak{g}_\bullet := (\log G_i)_{i=1}^s$.

[5] These are qualitative, Lie algebra analogues of horizontal characters on the Lie group. Here and throughout, by a 'horizontal character' on the Lie group, we mean a continuous homomorphism $G \to \mathbb{R}$ which maps Γ to $\mathbb{Z}$. We note that this definition coincides with the terminology 'i-horizontal character' from [GT10], but differs slightly from the usage of the terminology 'horizontal character' in [GT12]: in our definition and in [GT10], the relevant maps take values in $\mathbb{R}$, whereas in [GT12, Definition 1.5], the relevant map is defined to take values in $\mathbb{R}/\mathbb{Z}$. Whether we are referring to a horizontal character on the Lie group or on the Lie algebra should be clear from context.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

DEFINITION 2.10. Let $\mathfrak{g}$ have filtration $\mathfrak{g}_\bullet = (\mathfrak{g}_i)_{i=0}^s$. For $i = 1, \dots, s$, define the subalgebra $\mathfrak{h}_i$ to be the smallest Lie algebra containing $\mathfrak{g}_{i+1}$ and $[\mathfrak{g}_j, \mathfrak{g}_{i-j}]$ for all j.

The following definition is what one obtains when one lifts qualitative Green–Tao irrationality to the Lie algebra. We call it 'filtration irrationality'.

DEFINITION 2.11 (Qualitative filtration irrationality). We say $a \in \mathfrak{g}_i$ is *i-filtration irrational* if for all $l \in \mathfrak{g}_{\mathbb{Q}}^*$ which are nontrivial on $\mathfrak{g}_i$ and vanish on $\mathfrak{h}_i$, we have $l(a) \notin \mathbb{Q}$. A polynomial sequence $p(n) = \sum_{i=1}^s a_i n^i$ is *filtration irrational* in $\mathfrak{g}_\bullet$ if a_i is i-filtration irrational for $i = 1, \dots, s$.

It is standard [CG90, Theorem 5.1.8] that for any lattice $\Gamma \leq G$, $\operatorname{span}_{\mathbb{Q}} \log \Gamma$ is a rational structure in $\mathfrak{g}$. With the rational structure $\mathfrak{g}_{\mathbb{Q}}$ on $\mathfrak{g}$ chosen in this way, one may check that every l as in Definition 2.11 induces a function η on G_i with the properties in Definition 2.9 (and, conversely, every η induces an $l \in \mathfrak{g}_{\mathbb{Q}}^*$ which vanishes on $\mathfrak{h}_i$).

In the examples in the previous subsection, we were somewhat vague about the notion of irrationality. The reader may wish to return to Examples 2.2 and 2.3 to verify that the polynomial sequences there are filtration irrational and, thus, ultimately that filtration irrationality is insufficient to determine the distribution of p^Ψ in the non-flag setting.

DEFINITION 2.12 (Rational subspace). A subspace S of $\mathfrak{g}$ is *rational* with respect to a rational structure $\mathfrak{g}_{\mathbb{Q}}$ if it may be written as the kernel of a set of elements of $\mathfrak{g}_{\mathbb{Q}}^*$.

DEFINITION 2.13 (Qualitative linear irrationality). Let S be a rational subspace of $\mathfrak{g}$. An element $a \in S$ is *linearly irrational* in S if for all $l \in \mathfrak{g}_{\mathbb{Q}}^*$ which are nontrivial on S, we have $l(a) \notin \mathbb{Q}$. A polynomial sequence $p(n) := \sum_{i=1}^s a_i n^i$ is *linearly irrational* in a sequence of subspaces $S_\bullet = (S_i)_{i=1}^s$ if a_i is linearly irrational in S_i for $i = 1, \dots, s$.

The notion of linear irrationality does not refer to the Lie algebra structure on $\mathfrak{g}$ and so may be defined analogously for vector spaces with fixed rational structures. We note that linear irrationality is robust under rational (vector space) automorphisms.

In our Heisenberg example above, recalling that α, β, γ and 1 are linearly independent over $\mathbb{Q}$, we have that $\tilde{p}(n) = (\alpha, \beta, 0)n$ is linearly irrational in $S_\bullet = (S_1 := \mathbb{R} \times \mathbb{R} \times 0, S_2 := 0)$, but is not linearly irrational in $(\mathbb{R}^3, 0)$, or in $(\mathbb{R} \times \mathbb{R} \times 0, 0 \times 0 \times \mathbb{R})$. Define

$$p'(n) = (\alpha, \beta, 2\alpha + \beta)n + (0, 0, \gamma)n^2, \quad p''(n) = (\alpha, \beta, 2\alpha + \beta + 1/3)n + (0, 0, \gamma)n^2.$$

Then p' is linearly irrational in $(\ker(2, 1, -1), 0 \times 0 \times \mathbb{R})$. On the other hand, p'' is not linearly irrational in any sequence of subspaces of $\mathfrak{g}$ because there is $l \in \mathfrak{g}_{\mathbb{Q}}^*$ with $0 \neq l(a_1) \in \mathbb{Q}$.

In general, any polynomial sequence which has a coefficient a_i with $0 \neq l(a_i) \in \mathbb{Q}$ for some $l \in \mathfrak{g}_{\mathbb{Q}}^*$ is not linearly irrational in any sequence of subspaces. If, on the other hand, all coefficients $\{a_i\}_i$ of p satisfy the property that $l(a_i) \in \mathbb{Q}$ implies $l(a_i) = 0$, then p is linearly irrational in the sequence of rational subspaces $S_\bullet = (S_i)_{i=1}^s$ defined by $S_i = (a_i^\perp \cap \mathfrak{g}_{\mathbb{Q}}^*)^\perp$.

Linear irrationality is a stronger notion than filtration irrationality in the sense that if p is linearly irrational with respect to some $S_\bullet$, then there exists some $\mathfrak{g}'_\bullet$ (potentially on a subalgebra $\mathfrak{g}' \leq \mathfrak{g}$) with respect to which p is filtration irrational (see Lemma 3.4). The converse is not true, as is demonstrated by the example of p'' above which is filtration irrational with respect to the filtration given by $\mathfrak{g}_1 = \mathfrak{g}$ and $\mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}] = (0, 0, \mathbb{R})$; indeed, for nontrivial $l \in \mathfrak{g}_{\mathbb{Q}}^*$ which vanishes on $\mathfrak{h}_1 = \mathfrak{g}_2$, we have that $l((\alpha, \beta, 2\alpha + \beta + 1/3))$ is a nonzero rational linear combination of α and β, which cannot lie in $\mathbb{Q}$ since α and β have been chosen to be rationally independent (and $(0, 0, \gamma)$ is trivially filtration irrational in $\mathfrak{g}_2$). Finally, also note that if p is adapted to a rational

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

filtration $\mathfrak{g}_\bullet$ and is linearly irrational with respect to $S_\bullet$ then we automatically have that $S_i \leq \mathfrak{g}_i$ for all i and so $[S_i, S_j] \subset \mathfrak{g}_{i+j}$ for all i, j.

Unfortunately the notions of linear and filtration irrationality are not quite as simple in the quantitative setting. For example, a polynomial sequence may be quantitatively filtration/linearly irrational in two distinct sequences of subspaces; see Example 4.2. The relationship between the two notions of irrationality is also not so clear cut in the quantitative setting (though they are related; see Lemma 7.1). For this reason, we explicitly keep track of the sequence of subspaces with respect to which we are asserting quantitative linear irrationality and independently the filtered Lie algebra with respect to which we are asserting quantitative filtration irrationality.

2.4 The counting lemma

We can now prove a qualitative counting lemma for polynomial sequences on linear patterns. Again we restrict our attention to polynomial sequences with $p(0) = 0 \in \mathfrak{g}$; in the general case, one simply obtains equidistribution on a coset of the algebra studied in this section. Taking for granted our algebraic preliminaries, Theorem 2.16 is rather straightforward.

Definition 2.14. Given a system of t linear forms Ψ and a sequence of rational subspaces $S_\bullet = (S_i)_{i=1}^s$ of $\mathfrak{g}_\bullet$, define the Lie algebra $g^\Psi(S_\bullet)$ as the smallest Lie subalgebra of $\mathfrak{g}^t$ which contains $S_i \otimes V_\Psi^i$ for all i (where recall V_Ψ is defined in § 1.3).

The qualitative counting lemma may be stated informally as follows: if a polynomial sequence p is linearly irrational in $S_\bullet$, then p^Ψ equidistributes in $\mathfrak{g}^\Psi(S_\bullet)$. We offer a more explicit description of $g^\Psi(S_\bullet)$ in § 3.

Lemma 2.15. *Let $p : \mathbb{R}^D \to \mathbb{R}$ be a polynomial such that $p(\mathbb{Z}^D) \subset \mathbb{Q}$. Then the coefficients of p (with respect to the monomial basis) lie in $\mathbb{Q}$.*

Proof. For univariate polynomials ($D = 1$), the coefficients may be recovered by evaluating the polynomial at sufficiently distinct points and basic linear algebra over $\mathbb{Q}$. The multivariate case follows from an easy induction on the number of variables. □

We are ready to prove the qualitative counting lemma.

Theorem 2.16 (Qualitative counting lemma). *Suppose that p is a linearly irrational polynomial sequence in $S_\bullet$. Then the polynomial sequence $p^\Psi(\boldsymbol{x})$ equidistributes on $\mathfrak{g}^\Psi(S_\bullet)$ with respect to the rational structure $\mathfrak{g}^\Psi(S_\bullet) \cap \mathfrak{g}_\mathbb{Q}^t$.*

Proof. We use Theorem 2.8 and so show that p^Ψ is additively irrational in $\mathfrak{g}^\Psi(S_\bullet)$. Let $p(n) := \sum_{i=1}^s a_i n^i$. Let $\eta \in \mathrm{Hom}_\mathbb{Q}(\mathfrak{g}^\Psi(S_\bullet), \mathbb{R})$.[6] In $\mathfrak{g}^\Psi(S_\bullet) \leq \mathfrak{g} \otimes \mathbb{R}^t$, we have that

$$\eta(g^\Psi(\boldsymbol{x})) = \sum_{i=1}^s \eta(a_i \otimes \Psi(\boldsymbol{x})^i) = \sum_{m \in \mathcal{M}} \sum_{i=1}^s \eta(a_i \otimes v_{m,i}) m(\boldsymbol{x}), \tag{4}$$

where the $\{v_{m,i}\}_{m \in \mathcal{M}}$ span V_Ψ^i as in § 1.3. By Lemma 2.15, if $\eta(g^\Psi(\boldsymbol{x})) \in \mathbb{Q}$ for all $\boldsymbol{x} \in \mathbb{Z}^D$, then each $\eta(a_i \otimes v_{m,i}) \in \mathbb{Q}$. But each $\eta(\cdot \otimes v_{m,i})$ lies in $\mathfrak{g}_\mathbb{Q}^*$ and so by the linear irrationality of a_i in S_i we have that, in fact, $\eta(a_i \otimes v_{m,i}) = 0$. The kernel of a rational homomorphism is a rational subspace and so $\eta(S_i \otimes v_{m,i}) = 0$, since S_i is necessarily the smallest rational subspace containing a_i. Next, the $\{v_{m,i}\}_{m \in \mathcal{M}}$ span V_Ψ^i, so we have $\eta(S_i \otimes V_\Psi^i) = 0$ for all i. Furthermore, since η is a Lie algebra homomorphism to an abelian Lie algebra, η vanishes on the smallest

[6] Recall that we use the notation $\mathrm{Hom}_\mathbb{Q}$ to refer to homomorphisms of Lie algebras which map $\mathfrak{g}_\mathbb{Q}$ to $\mathbb{Q}$.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Lie algebra containing the subspaces $S_i \otimes V_\Psi^i$ and so η is trivial on $\mathfrak{g}^\Psi(S_\bullet)$. This completes the proof. □

Remark 2.17. In the flag setting, if η is a horizontal character on G^t (or a subgroup thereof), the 'pullback' $\eta(\cdot^{v_{m,i}})$ is an i-horizontal character on G_i (see [GT10, Definition A.5]). However, in the non-flag setting, $\eta(\cdot^{v_{m,i}})$ may fail to possess obviously discernible structure as a map on G_i (cf. Example 2.3 and the horizontal character η from there). The key technical point in the above is that horizontal characters on G^t, when lifted to maps on $\mathfrak{g}^t$ (or subalgebras thereof), pull back to linear maps on $\mathfrak{g}$. In the notation of the proof of Theorem 2.16, $\eta(\cdot \otimes v_{m,i})$ is a rational linear map on $\mathfrak{g}$ whence linear irrationality is sufficient to determine the distribution of p^Ψ.

Note that Theorem 2.16 yields equidistribution statements for polynomial sequences which are not filtration irrational.

Example 2.18. In the Heisenberg group with usual coordinates, let $p(n) = (\alpha, \alpha, 0)n$. Then p is linearly irrational in $S_\bullet := (S_1 := \{(x, x, 0) : x \in \mathbb{R}\}, S_2 := 0)$. Furthermore, we have from (2) that $[S_1, S_1] = 0$ (as is expected since this example is really coming from an abelian group). Theorem 2.16 says that p^Ψ equidistributes on $S_1 \otimes V_\Psi$.

We are also able to prove the equidistribution claims we made in Examples 2.2 and 2.3, and in § 2.2.

Example 2.19 (Equidistribution for Examples 2.2 and 2.3). In the abelian setting we may identify the Lie group with the Lie algebra. Let $p(n) = (\alpha, 0)n + (0, \beta)n^2$ on $\mathbb{R}^2$ be the first polynomial sequence from Example 2.2. Theorem 2.16 yields that p^Ψ equidistributes on $(\mathbb{R} \times 0) \otimes V_\Psi + (0 \times \mathbb{R}) \otimes V_\Psi^2$. On the other hand, the second polynomial sequence from Example 2.2 $\tilde{p}(n) = (\alpha, \gamma)n + (0, \beta)n^2$ equidistributes on $\mathbb{R}^2 \otimes V_\Psi + (0 \times \mathbb{R}) \otimes V_\Psi^2$.

Moving to Example 2.3, we saw in § 2.2 that the relevant polynomial sequences may be transferred to the Lie algebra as $p(n) = (\alpha, \beta, \gamma - \frac{1}{2}\alpha\beta)n$ and $\tilde{p}(n) = (\alpha, \beta, 0)n$. Theorem 2.16 says that p^Ψ equidistributes in the smallest Lie algebra containing $\mathbb{R}^3 \otimes V$, which is $\mathbb{R}^3 \otimes V_\Psi + (0 \times 0 \times \mathbb{R}) \otimes V_\Psi^2 = \mathfrak{g} \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2$. We also obtain that $\tilde{p}^\Psi$ equidistributes in $(\mathbb{R} \times \mathbb{R} \times 0) \otimes V_\Psi + \mathfrak{g}_2 \otimes V_\Psi^2$. One may furthermore check that this subalgebra maps to the subgroup $\ker \eta$ from Example 2.3 under the exponential map.

In each of the examples above, computing explicitly the Lie algebra $\mathfrak{g}^\Psi(S_\bullet)$ is not a difficult task. Of course, this may not be the case in general. We address this in § 3.

2.5 Factorisation

We have demonstrated how to obtain an equidistribution result for polynomial sequences which satisfy a special property: linear irrationality. Ultimately, one wants to understand the equidistribution of arbitrary polynomial sequences (we again focus on p with $p(0) = 0$). The purpose of this section is to show (Proposition 2.28) that one may write an arbitrary p as a product $p = p' * r$ where p' is linearly irrational and r is rational (and, as before, $*$ is group multiplication conducted in the Lie algebra). Then, roughly speaking, one may partition $\mathbb{Z}^D$ into appropriately spaced subprogressions where we understand the equidistribution of p on each one.

Part of the power of Green and Tao's counting lemma is that it requires a surprisingly small amount of data from a polynomial sequence g to determine the distribution of g^Ψ in the flag case. In the non-flag case, we require more data to determine distribution and so the task of

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

reducing a general nilsequence to something that our counting lemma can handle (i.e. factorising for linear irrationality) is inevitably a little more involved. This subsection may be viewed as a warm up for the quantitative factorisation theory in § 5.

Linear irrationality is a linear algebraic notion and indeed one may, by natural linear algebraic arguments, write an arbitrary $a \in \mathfrak{g}$ as the sum of a linearly irrational element and a rational element. This is the content of the upcoming lemma. We do not provide a proof immediately because Lemma 2.20 is a special case of the more general Lemma 2.22, which is proven shortly afterwards.

Lemma 2.20. *Let $a \in \mathfrak{g}$. Then there exists a_p which is linearly irrational in some rational subspace $S \leq \mathfrak{g}$ and $a_r \in \mathfrak{g}_{\mathbb{Q}}$ such that $a = a_p + a_r$.*

One may apply this to each coefficient of an arbitrary polynomial sequence p to obtain an additive decomposition $p = p' + r$, where p' is linearly irrational and r is rational. This additive decomposition does not, however, immediately determine a factorisation with respect to group multiplication, which is ultimately what is required.

To later obtain our multiplicative factorisation, we need a stronger additive decomposition, Lemma 2.22, of which Lemma 2.20 is a special case. First we need the following definition of linear irrationality relative to a subspace.

Definition 2.21 (Relative linear irrationality). Let U, T be rational subspaces of $\mathfrak{g}$. An element $a \in U$ is *linearly irrational in U* mod T if for all $l \in \mathfrak{g}^*_{\mathbb{Q}}$ which are nontrivial on U and which vanish on $U \cap T$ we have $l(a) \notin \mathbb{Q}$.

In what follows we use $T_{\mathbb{Q}}$ to denote the space $T \cap \mathfrak{g}_{\mathbb{Q}}$, and similarly for other subspaces. If T is rational with respect to $\mathfrak{g}_{\mathbb{Q}}$, then $T_{\mathbb{Q}}$ is a rational vector space of full dimension (and indeed topologically dense) inside T.

Lemma 2.22 (Qualitative additive decomposition for linear irrationality). *Let U, T be rational subspaces of $\mathfrak{g}$ and let $a \in U$ be linearly irrational in U mod T. Then there exist $a_p \in U$ and $a_r \in (U \cap T)_{\mathbb{Q}}$ such that a_p is linearly irrational in some rational subspace $U' \leq U$ and such that $a = a_p + a_r$.*

Proof. We identify $\mathfrak{g}$ with $U \oplus U^{\perp}$. Let $L \leq \mathfrak{g}^*_{\mathbb{Q}}$ be the subspace of elements l which are trivial on $U^{\perp}$ and such that $l(a) \in \mathbb{Q}$. We may assume that L has positive dimension, since otherwise a is already linearly irrational in U. Choose a basis $\mathcal{B}$ for L.

Next, we may assume that $U \cap T$ has positive dimension, otherwise again a is already linearly irrational in U by its linear irrationality in $U \cap T$. We claim that the elements of $\mathcal{B}$ remain linearly independent when restricted to maps on $U \cap T$. Indeed, if not, choose a nontrivial linear combination of the elements of $\mathcal{B}$ which is the trivial map on $U \cap T$ to obtain a contradiction with the linear irrationality of a in U mod T. Furthermore, since U, T are rational subspaces, so too is $U \cap T$ and so $\mathcal{B}$ is linearly independent when restricted to a set of maps on the $\mathbb{Q}$-vector space $(U \cap T)_{\mathbb{Q}}$. Thus, the map from $(U \cap T)_{\mathbb{Q}}$ to $\mathbb{Q}^{|\mathcal{B}|}$ given by $x \mapsto (l(x))_{l \in \mathcal{B}}$ is full rank and, in particular, surjective. Therefore, there exists $a_r \in (U \cap T)_{\mathbb{Q}}$ such that $l(a) = l(a_r)$ for all $l \in \mathcal{B}$ and, thus, for all $l \in L$.

Note that since $a_r \in U_{\mathbb{Q}}$, the subspace of $\mathfrak{g}^*_{\mathbb{Q}}$ of elements which are trivial on $U^{\perp}$ and such that $l(a - a_r) \in \mathbb{Q}$ is also equal to L, and for all of these elements $l(a - a_r) = 0$. Ultimately, letting $l \in \mathfrak{g}^*_{\mathbb{Q}}$ and writing $l = l' + l''$ where l' is trivial on $U^{\perp}_{\mathbb{Q}}$ and l'' is trivial on $U_{\mathbb{Q}}$, we see that if $l(a - a_r) \in \mathbb{Q}$ then $l'(a - a_r) \in \mathbb{Q}$, so that $l' \in L$, and so $l(a - a_r) = 0$. Thus $a - a_r$ is linearly irrational in $\ker\{l \in \mathfrak{g}^*_{\mathbb{Q}} : l(a - a_r) \in \mathbb{Q}\} \leq U$. □

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Note that the above lemma does not depend at all on the Lie bracket structure and so holds in vector spaces.

DEFINITION 2.23. Let $\mathfrak{g}_\bullet$ be a filtration of degree s on $\mathfrak{g}$ and let $U_\bullet, T_\bullet \leq \mathfrak{g}_\bullet$ be two sequences of rational subspaces with $U_i, T_i \leq \mathfrak{g}_i$ for all i. A polynomial sequence $p := \sum_{i=1}^s a_i n^i$ in $\mathfrak{g}$ is *linearly irrational in* $U_\bullet$ mod $T_\bullet$ if a_i is linearly irrational in U_i mod T_i for all i.

DEFINITION 2.24. A polynomial sequence $p(n) := \sum_{1=1}^s a_i n^s$ adapted to $\mathfrak{g}_\bullet$ is *rational* if $a_i \in (\mathfrak{g}_i)_\mathbb{Q}$ for all i.

In what follows, let $(p)_i$ denote the ith coefficient of a polynomial sequence p. Applying Lemma 2.22 to each coefficient of a polynomial sequence we obtain the following.

COROLLARY 2.25. *Let $U_\bullet, T_\bullet$ be sequences of rational subspaces of $\mathfrak{g}_\bullet$ and let p be a polynomial sequence which is linearly irrational in $U_\bullet$ mod $T_\bullet$. Then there are polynomials p' and r adapted to $U_\bullet$ and $(U_\bullet \cap T_\bullet)_\mathbb{Q} := ((U_i \cap T_i)_\mathbb{Q})_{i=1}^s$, respectively, such that p' is linearly irrational in some sequence of rational subspaces $U'_\bullet \leq U_\bullet$ and $p = p' + r$.*

Now we pursue a factorisation with respect to the $*$ operation. We need the fact that the set of polynomial maps adapted to $\mathfrak{g}_\bullet$ form a group. This result (in the Lie group, and in greater generality) is originally due to Leibman [Lei02]. See also [GT12, Proposition 6.2] for a related statement and different proof. The result that we need in the Lie algebra is the following more obvious lemma, which follows easily from the Baker–Campbell–Hausdorff formula (we omit the proof).

LEMMA 2.26. *The set of polynomials adapted to $\mathfrak{g}_\bullet$ forms a group under the $*$ operation.*

LEMMA 2.27. *The product of two rational polynomial sequences is rational.*

Proof. This follows immediately from the Baker–Campbell–Hausdorff formula $x * y = x + y + \frac{1}{2}[x, y] + \cdots$ and the fact the rational subspaces $(\mathfrak{g}_i)_\mathbb{Q}$ are rational subalgebras which obey the same filtration as $\mathfrak{g}_\bullet$. □

PROPOSITION 2.28 (Qualitative multiplicative factorisation for linear irrationality). *Let p be a polynomial sequence in $\mathfrak{g}$ adapted to $\mathfrak{g}_\bullet$. We may write $p = p' * r$ where p' is linearly irrational with respect to some sequence of rational subspaces $S'_\bullet \leq \mathfrak{g}_\bullet$ and r is a rational polynomial sequence, both adapted to $\mathfrak{g}_\bullet$.*

Proof. Invoke Corollary 2.25 with $U_\bullet = \mathfrak{g}_\bullet$ and $T_\bullet = (\mathfrak{g}, \ldots, \mathfrak{g})$ to obtain p_1 which is linearly irrational in some sequence $S_\bullet^{(1)}$ and r_1 adapted to $(S_\bullet)_\mathbb{Q}$ such that $p = p_1 + r_1$. By the Baker–Campbell–Hausdorff formula and the fact that $\mathfrak{g}/\mathfrak{g}_2$ is abelian we have that

$$p = p_1 * r_1 \mod \mathfrak{g}_2.$$

Set $T_\bullet^{(2)} := (\mathfrak{g}_2, \mathfrak{g}_2, \ldots, \mathfrak{g}_2)$. Then $p * r_1^{-1}$ is a polynomial sequence adapted to $\mathfrak{g}_\bullet$ by Lemma 2.26, and is linearly irrational in $(S_\bullet^{(1)} + T_\bullet^{(2)})$ mod $T_\bullet^{(2)}$.[7] Next apply Corollary 2.25 with $U_\bullet = S_\bullet^{(1)} + T_\bullet^{(2)}$ and $T_\bullet = T_\bullet^{(2)}$ to $p * r_1^{-1}$ to find p_2 which is linearly irrational in some $S_\bullet^{(2)}$ and r_2 which is adapted to $(S_\bullet^{(1)} \cap T_\bullet^{(2)})_\mathbb{Q}$ such that $p * r_1^{-1} = p_2 + r_2$. By the Baker–Campbell–Hausdorff formula and the fact that $[\mathfrak{g}, \mathfrak{g}_2] \subset \mathfrak{g}_3$ we then have $p * r_1^{-1} = p_2 * r_2$ mod $\mathfrak{g}_3$ and so

$$p = p_2 * r_2' \mod \mathfrak{g}_3,$$

[7] Here we recall that all elements of the filtration $\mathfrak{g}_\bullet$ are rational in $\mathfrak{g}$, so in particular $T_\bullet$ is a sequence of rational subspaces.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

where $r_2' := r_2 * r_1$ is rational by Lemma 2.27. Continue in this way to eventually find $p' := p_s$, $S'_\bullet := S^{(s)}_\bullet$ and $r := r'_s$. □

3. Description of $\mathfrak{g}^\Psi(S_\bullet)$ and comparison with $\log G^\Psi$

Our counting lemma, Theorem 1.2, states that if p is a polynomial sequence which is quantitatively linearly irrational in $S_\bullet$, then p^Ψ quantitatively equidistributes in a Lie subalgebra $\mathfrak{g}^\Psi(S_\bullet)$ of $\mathfrak{g}^t$. We have proven a qualitative version of this result, Theorem 2.16, which gives qualitative equidistribution on the same Lie algebra $\mathfrak{g}^\Psi(S_\bullet)$. So far, however, we have only described $\mathfrak{g}^\Psi(S_\bullet)$ as the smallest Lie subalgebra of $\mathfrak{g}^t$ which contains $S_i \otimes V^i_\Psi$ for all i. The goal of this section is to give descriptions of $\mathfrak{g}^\Psi(S_\bullet)$ which are more explicit and amenable to applications.

Let $S_\bullet := (S_i)_{i=1}^s$ be a sequence of rational subspaces of $\mathfrak{g}$ with respect to a fixed rational structure $\mathfrak{g}_\mathbb{Q}$. Let $W_1 = S_1$ and iteratively define $W_i = \operatorname{span}\{S_j, [W_j, W_{i-j}] \text{ for } j = 1, \dots, i-1\}$ for $i = 2, \dots, s$. That is, W_i comprises all degree i commutators in the S_j, where the degree of a commutator is the sum of the subscripts it contains. Note that if $\mathfrak{g}_\bullet$ is a filtration such that $S_i \le \mathfrak{g}_i$ for all i, then $W_i \le \mathfrak{g}_i$ for all i.

Proposition 3.1 (General description of $\mathfrak{g}^\Psi(S_\bullet)$). *Let $S_\bullet$ be a sequence of rational subspaces of $\mathfrak{g}_\bullet$. Let Ψ be a system of linear forms. Then $\mathfrak{g}^\Psi(S_\bullet) = \sum_{i=1}^s W_i \otimes V^i_\Psi$.*

Proof. Recall from (3) the Lie bracket operation on $\mathfrak{g} \otimes \mathbb{R}^t$. It follows immediately that for all i, j we have $[W_i \otimes V^i_\Psi, W_j \otimes V^j_\Psi] \subset W_{i+j} \otimes V^{i+j}_\Psi$ (where $W_{s+1} = W_{s+2} = \cdots = 0$) and so $\sum_{i=1}^s W_i \otimes V^i_\Psi$ is indeed a Lie subalgebra of $\mathfrak{g}^t$. It is also immediate from definitions that $\sum_{i=1}^s W_i \otimes V^i_\Psi$ is generated by $\{S_i \otimes V^i_\Psi\}_{i=1}^s$ under commutators and linear spans. □

Definition 3.2. Recall that we use G^Ψ to denote the Leibman group (Definition 2.1); we call $\log G^\Psi = \sum_{i=1}^s \mathfrak{g}_i \otimes V^i_\Psi$ the *Leibman algebra*.

We now pursue Proposition 3.7, a description of $\log G^\Psi$ which is more readily comparable to that of $\mathfrak{g}^\Psi(S_\bullet)$ in the above proposition. In so doing, we make the assumption that $S_i + \mathfrak{h}_i = \mathfrak{g}_i$ for all i (recall Definition 2.10). The goal of the upcoming lemmas is to motivate why this is reasonable.

For a sequence of subspaces $S_\bullet \le \mathfrak{g}_\bullet$, we say that a filtration $\mathfrak{g}'_\bullet \le \mathfrak{g}_\bullet$ is *minimal* for $S_\bullet$ if $S_\bullet \le \mathfrak{g}'_\bullet$ and there is no proper subfiltration $\mathfrak{g}''_\bullet < \mathfrak{g}'_\bullet$ with $S_i \le \mathfrak{g}''_i$ for all i.

Lemma 3.3. *For a sequence of subspaces $S_\bullet \le \mathfrak{g}_\bullet$, if $\mathfrak{g}_\bullet$ is minimal for $S_\bullet$ then $S_i + \mathfrak{h}_i = \mathfrak{g}_i$ for all i, where the $\mathfrak{h}_i$ are the subalgebras from Definition 2.10.*[8]

Proof. If not, for the smallest i such that $S_i + \mathfrak{h}_i \ne \mathfrak{g}_i$, set $\mathfrak{g}'_i = S_i + \mathfrak{h}_i$, and otherwise set $\mathfrak{g}'_j = \mathfrak{g}_j$ for all $j \ne i$. We first observe that $S_i + \mathfrak{h}_i$ is closed under the Lie bracket operation, since $S_i \le \mathfrak{g}_i$ and $[\mathfrak{g}, \mathfrak{g}_i] \le \mathfrak{g}_{i+1} \le \mathfrak{h}_i$; thus, $S_i + \mathfrak{h}_i$ is indeed a Lie subalgebra of $\mathfrak{g}$. Furthermore, since $\mathfrak{g}'_i \le \mathfrak{g}_i$ and $\mathfrak{g}'_j = \mathfrak{g}_j$ for all other j, to check that $\mathfrak{g}'_\bullet$ is a filtration, we just need to check that $\mathfrak{g}_{i+1} \le \mathfrak{g}'_i$ and $[\mathfrak{g}_k, \mathfrak{g}_{i-k}] \le \mathfrak{g}'_i$ for all $k = 0, \dots, i$. But both of these properties follow from the fact that $\mathfrak{h}_i \le \mathfrak{g}'_i$. Thus, we have constructed a proper subfiltration $\mathfrak{g}'_\bullet < \mathfrak{g}_\bullet$ with $S_\bullet \le \mathfrak{g}'_\bullet$, contradicting the minimality of $\mathfrak{g}_\bullet$. □

Lemma 3.4. *Let p be linearly irrational with respect to a sequence of subspaces $S_\bullet \le \mathfrak{g}_\bullet$. Then p is filtration irrational with respect to the filtration $\mathfrak{g}'_\bullet \le \mathfrak{g}_\bullet$ which is minimal for $S_\bullet$.*

[8] The converse is also true though we do not need such a statement.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Proof. Write $p(n)=\sum_{i=1}^{s}a_in^i$. Suppose for a contradiction that p is not filtration irrational in the minimal filtration $\mathfrak{g}'_\bullet$ for $S_\bullet$. By expanding the definition of filtration irrationality (Definition 2.11), let j be such that a_j is not filtration irrational in $\mathfrak{g}'_j$ and let $l\in\mathfrak{g}'^*_{\mathbb{Q}}$ be nontrivial on $\mathfrak{g}'_j$, vanish on $\mathfrak{h}'_j$ and have $l(a_j)\in\mathbb{Q}$. Since a_j is linearly irrational in S_j, the map l must be trivial on S_j. Thus, we must have, in fact, $l(a_j)=0$, and so $a_j\in\ker l$. Let $\mathfrak{g}''_j=\ker l\cap\mathfrak{g}'_j<\mathfrak{g}'_j$. Note that $S_j\le\mathfrak{g}''_j$. We are now able to contradict the minimality of $\mathfrak{g}'_\bullet$ in much the same way as we did in the proof of Lemma 3.3.

Since $\ker l$ contains $\mathfrak{h}'_j$, so too does $\mathfrak{g}''_j$ and so we see that $\mathfrak{g}''_j$ is indeed closed under the Lie bracket, that $\mathfrak{g}'_{j+1}\le\mathfrak{g}''_j$ and, furthermore, that $[\mathfrak{g}'_k,\mathfrak{g}'_{j-k}]\le\mathfrak{g}''_j$ for all $k=0,\dots,j$. Thus, defining $\mathfrak{g}''_i:=\mathfrak{g}'_i$ for all $i\neq j$, we obtain a strictly smaller filtration $\mathfrak{g}''_\bullet$ which contains $S_\bullet$, contradicting the minimality of $\mathfrak{g}'_\bullet$. □

LEMMA 3.5. *Let $p(n):=\sum_{i=1}^{s}a_in^i$ be filtration irrational in $\mathfrak{g}_\bullet$. Then for any sequence of rational subspaces $T_\bullet$ with $a_i\in T_i\le\mathfrak{g}_i$, we have $T_i+\mathfrak{h}_i=\mathfrak{g}_i$ for all i.*

In particular, if p is linearly irrational in $S_\bullet$, then $S_i+\mathfrak{h}_i=\mathfrak{g}_i$ for all i.

Proof. That $T_i+\mathfrak{h}_i\le\mathfrak{g}_i$ is immediate. By the definition of filtration irrationality (Definition 2.11), for all nontrivial $l\in\mathfrak{g}^*_{i\mathbb{Q}}$ which vanish on $\mathfrak{h}_i$ we have $l(a_i)\notin\mathbb{Q}$. In particular, $a_i^\perp\cap\mathfrak{h}_i^\perp\cap\mathfrak{g}^*_{i\mathbb{Q}}$ contains only the trivial (i.e. zero) linear map (where here orthogonal complements are taken in $\mathfrak{g}_i$ and $\mathfrak{g}_i^*$). Therefore,

$$\mathfrak{g}_i=(a_i^\perp\cap\mathfrak{g}^*_{i\mathbb{Q}}\cap\mathfrak{h}_i^\perp)^\perp=(a_i^\perp\cap\mathfrak{g}^*_{i\mathbb{Q}})^\perp+\mathfrak{h}_i. \tag{5}$$

However, $(a_i^\perp\cap\mathfrak{g}^*_{i\mathbb{Q}})^\perp$ is the smallest rational subspace of $\mathfrak{g}_i$ containing a_i, so $T_i\ge(a_i^\perp\cap\mathfrak{g}^*_{i\mathbb{Q}})^\perp$. Combining this observation with (5), we deduce that $T_i+\mathfrak{h}_i\ge\mathfrak{g}_i$. □

The following is the main ingredient in the proof of Proposition 3.7. Recall that we let $W_1=S_1$ and iteratively define $W_i=\operatorname{span}\{S_j,[W_j,W_{i-j}]\text{ for }j=1,\dots,i-1\}$ for $i=2,\dots,s$.

LEMMA 3.6. *Suppose that the sequence of subspaces $(S_i)_{i=1}^s\le\mathfrak{g}_\bullet$ satisfies $S_i+\mathfrak{h}_i=\mathfrak{g}_i$ for all i. Then the sequence of subspaces $(W_i)_{i=1}^s$ satisfies*

$$W_i+\mathfrak{g}_{i+1}=\mathfrak{g}_i,$$

for all i.

Proof. The proof proceeds by induction and the base case is easy from definitions. Expanding the definition of W_i, we show that

$$S_i+\sum_{j=1}^{i-1}[W_j,W_{i-j}]+\mathfrak{g}_{i+1}=\mathfrak{g}_i. \tag{6}$$

Now for $j=1,\dots,i-1$ we have $W_j\le\mathfrak{g}_j$ and so by the filtration $\mathfrak{g}_\bullet$ we get

$$[W_j+\mathfrak{g}_{j+1},W_{i-j}+\mathfrak{g}_{i+1-j}]+\mathfrak{g}_{i+1}=[W_j,W_{i-j}]+\mathfrak{g}_{i+1}.$$

By induction, therefore, (6) is equivalent to

$$S_i+\sum_{j=1}^{i-1}[\mathfrak{g}_j,\mathfrak{g}_{i-j}]+\mathfrak{g}_{i+1}=\mathfrak{g}_i.$$

However, the left-hand side of the above is just $S_i+\mathfrak{h}_i$. □

Recall the definition of the Leibman algebra, Definition 3.2.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Proposition 3.7 (Comparison of Leibman algebra and $\mathfrak{g}^{\Psi}(S_\bullet)$). *Suppose $S_\bullet$ is a sequence of rational subspaces of $\mathfrak{g}_\bullet$ such that $S_i + \mathfrak{h}_i = \mathfrak{g}_i$. Then*

$$\sum_{i=1}^{s} \mathfrak{g}_i \otimes V_\Psi^i = \sum_{i=1}^{s} W_i \otimes \left(\sum_{j=1}^{i} V_\Psi^j \right) = \mathfrak{g}^{\Psi}(S_\bullet) + \sum_{i=1}^{s} W_i \otimes \left(\sum_{j=1}^{i-1} V_\Psi^j \right).$$

Proof. From Lemma 3.6 we have $W_i + \mathfrak{g}_{i+1} = \mathfrak{g}_i$ and so $\mathfrak{g}_i = \sum_{j=i}^{s} W_j$. The result follows by switching the order of summation. □

Corollary 3.8. *If Ψ satisfies the flag property and $\mathfrak{g}_\bullet$ is minimal for $S_\bullet$, then $\mathfrak{g}^{\Psi}(S_\bullet)$ is equal to the Leibman algebra.*

Proof. If Ψ satisfies the flag property, then for all i we have $\sum_{j=1}^{i} V_\Psi^j = V_\Psi^i$; invoke Propositions 3.1 and 3.7. □

4. The counting lemma

Our first goal in the quantitative setting is to obtain a counting lemma for polynomial sequences which are (quantitatively) linearly irrational. Throughout this section and the next, we assume that the dimension d of $\mathfrak{g}$, the degree s of $\mathfrak{g}_\bullet$, the number of linear forms t and the number of variables D in these linear forms are all of bounded size; in particular, we always allow constants to depend on these parameters without necessarily indicating this in our notation.

We need to fix a basis for $\mathfrak{g}$ in order to quantify various notions. (Indeed, in the absence of a choice of basis for $\mathfrak{g}$ we do not even have a canonical metric or (Lebesgue) measure on $\mathfrak{g}$.)

Recall that the *structure constants* of a basis $\mathcal{X} := \{X_1, \dots, X_{\dim \mathfrak{g}}\}$ of a real Lie algebra $\mathfrak{g}$ are the real numbers c_{ijk} satisfying

$$[X_i, X_j] = \sum_{k=1}^{\dim \mathfrak{g}} c_{ijk} X_k.$$

Definition 4.1 (Rational basis). A basis $\mathcal{X}$ for $\mathfrak{g}$ is *rational* if its structure constants lie in $\mathbb{Q}$.

The $\mathbb{Q}$-span of a rational basis is a rational structure on $\mathfrak{g}$. In what follows, we fix a rational basis $\mathcal{X}$ for $\mathfrak{g}$ and the corresponding rational structure $\mathfrak{g}_{\mathbb{Q}} := \operatorname{span}_{\mathbb{Q}}(\mathcal{X})$. If statements whose meaning implicitly depends upon a basis are made without reference to one, it may be assumed that $\mathcal{X}$ is the chosen basis. Furthermore, we denote by $\langle \mathcal{X} \rangle$ the abelian group generated by $\mathcal{X}$, that is, $\operatorname{span}_{\mathbb{Z}} \mathcal{X}$. We also need to conduct quantitative arithmetic on subspaces of $\mathfrak{g}$ and so for essentially the same reasons, we need to identify $\mathbb{Z}$-bases for sublattices obtained by intersecting $\langle \mathcal{X} \rangle$ with a rational subspace S. This may be done canonically by invoking the Hermite normal form. For details and definitions see Appendix B, but for the purposes of a more casual reading of this section, one may interpret *Hermite basis* to mean 'a canonically chosen basis $\mathcal{X}' = \mathcal{X}'_{S,\mathcal{X}}$ for S such that $\mathcal{X}'$ is also a $\mathbb{Z}$-basis for the lattice $S \cap \langle \mathcal{X} \rangle$'.

Note that this construction does not use the Lie algebra structure on $\mathfrak{g}$ and we may speak of Hermite bases of a vector subspace of a vector space with given basis. We caution that if $T \leq S \leq \mathfrak{g}$ are rational subspaces, then $\mathcal{X}'_{T,\mathcal{X}}$ (the Hermite basis for T in $\mathfrak{g}$ with respect to $\mathcal{X}$) need not equal $\mathcal{X}''_{T,\mathcal{X}'_{S,\mathcal{X}}}$ (the Hermite basis for T in S with respect to $\mathcal{X}'_{S,\mathcal{X}}$, which itself is the Hermite basis for S in $\mathfrak{g}$ with respect to $\mathcal{X}$). We also note that in the case that $S = \mathfrak{g}$, we have $\mathcal{X}' = \mathcal{X}$.

In the qualitative setting in § 2, if a polynomial sequence was linearly irrational with respect to any sequence of rational subspaces, then it was linearly irrational with respect to the sequence

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

of minimal rational subspaces of $\mathfrak{g}$ which contained (respectively) the sequence of polynomial coefficients.

We need to be a little more careful in the quantitative setting. Even in the single-variable case (that is, for the trivial sequence of linear forms $\Psi := (\psi_1(n))$ where $\psi_1(n) = n$) and for A arbitrarily large, (A, N)-irrationality (in the original Green–Tao sense [GT10, Definition A.6]) does not preclude a polynomial sequence from equidistributing on more than one nilmanifold of 'complexity' $A^{O(1)}$.

Example 4.2. Let A be a large integer and let $N > 100A^3$. Let $p(n) := (1/(10A-1), 1/10A)n$ define a polynomial sequence in the abelian Lie algebra $\mathfrak{g} := \mathbb{R}^2$ with basis $\mathcal{X} := \{(1,0),(0,1)\}$ and lattice $\langle \mathcal{X} \rangle = \mathbb{Z}^2$.

First, let us verify that in the limit as $A \to \infty$, the quantitative distribution obtained from the Green–Tao counting lemma agrees with our qualitative counting lemma, Theorem 2.16. Here, $a_1 = (1/(10A-1), 1/10A)$ and the smallest rational subspace containing a_1 is $S_1 := \{(x,y) \in \mathbb{R}^2 : (10A-1)x = 10Ay\}$. Furthermore, S_1 is a subalgebra of $\mathbb{R}^2$ and has 'complexity' $A^{O(1)}$ because it is defined by linear relations of height $A^{O(1)}$.[9] We claim that p is $o_{A\to\infty}(1)$ equidistributed in S_1. Clearly S_1 is isomorphic to $\mathbb{R}$, and $\varphi : (x,y) \mapsto x/(10A)$ is the isomorphism which maps $\mathcal{X}' = \{(10A, 10A-1)\}$ to $1 \in \mathbb{R}$. Thus, we may analyse $p'(n) = n/(10A(10A-1))$ in $\mathbb{R}/\mathbb{Z}$.[10] Set $a_1' := \varphi(a_1) = 1/10A(10A-1)$. For any nonzero $m \in \mathbb{Z}$ with $|m| \le A$ we have $\|ma_1'\|_{\mathbb{R}/\mathbb{Z}} \ge 1/10A(10A-1) > A/N$. Thus, by [GT10, Definition A.6], p' is (A,N)-irrational in $\mathbb{R}/\mathbb{Z}$, so p is (A,N)-irrational in S_1 with respect to $\mathcal{X}'$ and so invoking the single-variable version of Green and Tao's counting lemma [GT10, Lemma 3.7] we have that p is $o_{A\to\infty}(1)$ equidistributed in S_1, as is predicted by Theorem 2.16.

On the other hand, given any nonzero linear map $l : \mathfrak{g} \to \mathbb{R}$ which maps $\mathcal{X}$ to $\mathbb{Z}$ and has height at most A, we have by the coprimality of $10A$ and $10A-1$ that $\|l(a_1)\|_{\mathbb{R}/\mathbb{Z}} \ge 1/((10A)(10A-1)) > A/N$, and so p is (A,N)-irrational in $\mathfrak{g}$. Therefore, by [GT10, Lemma 3.7], p is $o_{A\to\infty}(1)$ equidistributed in all of $\mathbb{R}^2/\mathbb{Z}^2$.

Remark 4.3. Green and Tao's quantitative factorisation of polynomial sequences [GT10, Lemma 2.9] states that given a polynomial sequence g on a nilmanifold, *at least one* of the following is true: (1) g is (A,N)-irrational in the nilmanifold; or (2) one may find a smooth sequence β and a rational sequence γ such that $g = \beta * g' * \gamma$, where g' takes values in a complexity $O_A(1)$ subnilmanifold of strictly smaller total dimension. Example 4.2 exhibits a situation where these two eventualities are not mutually exclusive (note that the smooth and rational parts in our example are trivial).

DEFINITION 4.4 (Integer linear map). We call an element of $\mathfrak{g}^*$ *integer* (with respect to $\mathcal{X}$) if it maps $\mathcal{X}$ to $\mathbb{Z}$. We denote by $\mathfrak{g}^*_{\mathcal{X}}$ the set of integer linear maps with respect to $\mathcal{X}$ and we denote by $\mathrm{Hom}_{\mathcal{X}}(\mathfrak{g}, \mathbb{R})$ those elements of $\mathfrak{g}^*_{\mathcal{X}}$ which vanish on $[\mathfrak{g}, \mathfrak{g}]$ (i.e. are Lie algebra homomorphisms).

In the following, we refer to the Hermite basis $\mathcal{X}'_{S,\mathcal{X}}$ whose formal definition is given in Definition B.2.

DEFINITION 4.5 (Complexity of an integer linear map). The *complexity* of an integer linear map $l : \mathfrak{g} \to \mathbb{R}$ on a rational subspace $S \le \mathfrak{g}$ with respect to $\mathcal{X}$ is the ℓ_∞ norm of l with respect to the dual basis of $\mathcal{X}'_{S,\mathcal{X}}$.

[9] The complexity of a subspace is defined rigorously soon. See Definition 4.6.

[10] We transfer this example to $\mathbb{R}$ with lattice $\mathbb{Z}$ where it is more clear what the definition of the complexity of a horizontal character should be. One may check that transferring in this way is equivalent to our upcoming formal Definition 4.5.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Definition 4.6 (Complexity of a subspace). Given a Lie algebra $\mathfrak{g}$ with rational basis $\mathcal{X}$ and a proper rational subspace $S \leq \mathfrak{g}$, we say that S has *complexity* at most C in $\mathfrak{g}$ (with respect to $\mathcal{X}$) if S may be written as the kernel of a set of integer linear maps, each of which have complexity at most C with respect to $\mathcal{X}$. We say that $\mathfrak{g}$ has complexity 1 in itself.

Definition 4.7 (Quantitative linear irrationality). For a Lie algebra $\mathfrak{g}$ and a rational subspace $S \leq \mathfrak{g}$ we say that $a \in S$ is (A, ε)-*linearly irrational in* S with respect to $\mathcal{X}$ if, for all integer linear maps $l : \mathfrak{g} \to \mathbb{R}$ which are nontrivial on S and have complexity at most A on S, we have $\|l(a)\|_{\mathbb{R}/\mathbb{Z}} \geq \varepsilon$.

Definition 4.8 (Quantitative linear irrationality for polynomial sequences). For a Lie algebra $\mathfrak{g}$ and a sequence of rational subspaces $S_\bullet := (S_i)_{i=1}^s$, $S_i \leq \mathfrak{g}$, a polynomial sequence $(p(\boldsymbol{n}))_{\boldsymbol{n} \in [N]^D}$ in $\mathfrak{g}$ is (A, N)-*linearly irrational in* $S_\bullet$ with respect to $\mathcal{X}$ if for all i, all coefficients of all degree i terms of p are $(A, A/N^i)$-linearly irrational in S_i.

We record some easy consequences of these definitions for more convenient use later on. The proofs are just linear algebra and are omitted. We refer the reader to Appendix B for some further details. Recall that throughout this section we are assuming that $\dim \mathfrak{g} = O(1)$.

Lemma 4.9 (Complexity of subspaces under operations). *Let $C > 1$. Let S, T be rational subspaces of $\mathfrak{g}$ of complexity at most C with respect to $\mathcal{X}$ and let U be a rational subspace of a distinct dimension $O(1)$ Lie algebra $\mathfrak{h}$ of complexity at most C with respect to a rational basis $\mathcal{Y}$. Then:*

(i) *$S + T$ has complexity at most $C^{O(1)}$;*
(ii) *$[S, T]$ has complexity at most $M^{O(1)}C^{O(1)}$, where M is an upper bound on the height of the largest (rational) structure constant of $\mathfrak{g}$ with respect to $\mathcal{X}$;*
(iii) *$S \otimes U$ has complexity at most $C^{O(1)}$ in $\mathfrak{g} \otimes \mathfrak{h}$ with respect to $\mathcal{X} \otimes \mathcal{Y}$.*

Lemma 4.10 (Complexity of maps under lifting and restriction). *Let $A, B, C > 1$ be integers and let S, T be rational subspaces of $\mathfrak{g}$ of complexity at most C with respect to $\mathcal{X}$ such that $S \leq T$. Then:*

(i) *the elements of $\mathcal{X}'_{S,\mathcal{X}}$ may be written as integer linear combinations of the elements of $\mathcal{X}$ with integers of size at most $C^{O(1)}$;*
(ii) *if l is an integer linear map on T of complexity at most A, then $l|_S$ is an integer linear map of complexity at most $C^{O(1)}A$ on S;*
(iii) *if l is an integer linear map on S of complexity at most B, then there is an integer linear map $\tilde{l}$ on T of complexity at most $C^{O(1)}B$ such that $\tilde{l}|_S = Kl$, where $K = C^{O(1)}$.*

Example 4.2 demonstrated that there are elements which are (A, ε)-linearly irrational with respect to more than one subspace. In this example, the nontrivial subspace was of complexity $A^{O(1)}$. The following lemma gives a sense in which this is the only way in which such a phenomenon can occur.

Lemma 4.11. *Let $A > 1$. Let $a \in \mathfrak{g}$ and let S, T be subspaces of $\mathfrak{g}$ of complexity at most A which contain a. There is a constant $k > 1$ (which may depend on $\dim \mathfrak{g}$) such that for every $\varepsilon > 0$ if a is (A^k, ε)-linearly irrational in T, then $T \leq S$.*

Proof. We show that $S^\perp \leq T^\perp$. Let $l : \mathfrak{g} \to \mathbb{R}$ have complexity at most A and lie in $S^\perp$. Since $a \in S$ we have $l(a) = 0$. Furthermore, l has complexity at most $A^{O(1)}$ with respect to $\mathcal{X}'_{\mathcal{X},T}$ by

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Lemma 4.10 and so for $k = O(1)$ sufficiently large, this contradicts the (A^k, ε)-linear irrationality of a in T, unless $l \in T^\perp$. Letting $\mathfrak{g}^*_{\mathcal{X}}(A)$ denote the set of integer linear maps of complexity at most A with respect to $\mathcal{X}$, we have so far that $S^\perp \cap \mathfrak{g}^*_{\mathcal{X}}(A) \subset T^\perp$ and taking the vector space closure on both sides we have $\mathrm{span}_{\mathbb{R}}\{S^\perp \cap \mathfrak{g}^*_{\mathcal{X}}(A)\} \subset T^\perp$. However, by the definition of the complexity of S, we have that $S^\perp$ may be written as the span of complexity $\leq A$ maps and so $S^\perp = \mathrm{span}_{\mathbb{R}}\{S^\perp \cap \mathfrak{g}^*_{\mathcal{X}}(A)\}$. This completes the proof. □

In particular, there is a polynomial threshold above which linear irrationality may be obtained on at most one rational subspace of bounded complexity. In general, however, the sequence of subspaces on which we are asserting linear irrationality is not determined by the polynomial sequence and so our counting lemma takes this as additional input.

Similarly to the use of Leibman's criterion for equidistribution in the qualitative setting, we introduce here Green and Tao's criterion for quantitative equidistribution. First a quantitative version of additive irrationality; see § A.1 for a definition of the $\|\cdot\|_{C^\infty[\boldsymbol{N}]}$-norm. In the following, we use $\mathrm{Hom}_{\mathcal{X}}(\mathfrak{g}, \mathbb{R})$ to denote the Lie algebra homomorphisms from $\mathfrak{g}$ to $\mathbb{R}$ which map $\mathcal{X}$ to $\mathbb{Z}$.

Definition 4.12 (Quantitative additive irrationality). Let $A > 1$, let D be a positive integer, let $N_1, \dots, N_D \geq 1$ be integers and let $[\boldsymbol{N}]$ denote the set $[N_1] \times \cdots \times [N_D]$. A polynomial sequence p on $\mathfrak{g}$ is $(A, \boldsymbol{N})$-additively irrational with respect to $\mathcal{X}$ if for all nontrivial $\eta \in \mathrm{Hom}_{\mathcal{X}}(\mathfrak{g}, \mathbb{R})$ of complexity at most A, we have $\|\eta \circ p\|_{C^\infty[\boldsymbol{N}]} > A$.

Remark 4.13. Though we have not done so here, one may develop the various notions of irrationality in a transparently analogous way to highlight that these notions essentially only differ by the subalgebras of $\mathfrak{g}$ on which the set of maps under consideration are required to vanish (i.e. $[\mathfrak{g}, \mathfrak{g}]$ for additive irrationality, $\mathfrak{h}_i$ for filtration irrationality and none for linear irrationality). One may define quantitative additive irrationality for elements of (rather than just polynomial sequences in) $\mathfrak{g}$, and define quantitative filtration/linear irrationality for multivariate polynomial sequences; these definitions are not required in this document and so are not provided.

Next we define quantitative equidistribution in $(\mathfrak{g}, \mathcal{X})$, and defer to Appendix A.2 for more details.

Definition 4.14 (Quantitative equidistribution in the Lie algebra). Let $\mathcal{X}$ be a rational basis in $\mathfrak{g}$. Then a polynomial sequence $(p(\boldsymbol{n}))_{\boldsymbol{n}\in[\boldsymbol{N}]}$ is δ-equidistributed on $[\boldsymbol{N}]$ in $(\mathfrak{g}, \mathcal{X})$ if $(\exp(p(\boldsymbol{n})))_{\boldsymbol{n}\in[\boldsymbol{N}]}$ is δ-equidistributed on $[\boldsymbol{N}]$ in G/Γ, where Γ is the smallest multiplicative lattice containing $\exp(\langle\mathcal{X}\rangle) = \exp(\mathrm{span}_{\mathbb{Z}}\,\mathcal{X})$.

Green and Tao show [GT15, Theorem p.3] that quantitative additive irrationality implies quantitative equidistribution. Their result concerns equidistribution in a nilmanifold; we defer to Appendix A for a deduction of the following Lie algebra version.

Theorem 4.15 (Green and Tao's criterion for quantitative equidistribution in the Lie algebra). *Let $0 < \delta < 1/2$, let $N_1, \dots, N_D \geq 1$ and let $[\boldsymbol{N}]$ denote the set $[N_1] \times \cdots \times [N_D]$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra with dimension d, step s and basis $\mathcal{X}$ which is $1/\delta$-rational. Let $(p(\boldsymbol{n}))_{\boldsymbol{n}\in\mathbb{Z}^D}$ be a polynomial sequence in $\mathfrak{g}$. There are positive constants $c_{d,s,D}, c'_{d,s,D}$ such that at least one of the following is true:*

(i) *there is some $N_i < \delta^{-c'_{d,s,D}}$;*
(ii) *$(p(\boldsymbol{n}))_{\boldsymbol{n}\in[\boldsymbol{N}]}$ is δ-equidistributed in $(\mathfrak{g}, \mathcal{X})$;*
(iii) *p is not $(\delta^{-c_{d,s,D}}, \boldsymbol{N})$-additively irrational in $(\mathfrak{g}, \mathcal{X})$.*

Furthermore, when $N_1 = \cdots = N_D$, the first option above may be removed.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

We are ready to prove the main theorem of this section. We only prove a result for averages over a cube $[N]^D$ (i.e. when $N_1 = \cdots = N_D = N$), but it should be clear how to obtain a result for the more general case in which not all sides have equal length.

Theorem 4.16 (Quantitative counting lemma, Theorem 1.2). *Let $N, A, M \geq 2$. Let $\Psi = (\psi_1, \ldots, \psi_t)$ be a collection of linear forms each mapping $\mathbb{Z}^D$ to $\mathbb{Z}$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra with rational basis $\mathcal{X}$, dimension d and rational structure constants of height at most M. Let $S_\bullet$ be a sequence of rational subspaces of complexity M in $\mathfrak{g}$. Suppose that $p(n)$ is an (A, N)-linearly irrational polynomial sequence of degree s in $S_\bullet$. Then there are constants $0 < c_1, c_2 = O_{d,s,D,t,\Psi}(1)$ such that the polynomial sequence $p^\Psi(\boldsymbol{x}) := (p(\psi_1(\boldsymbol{x})), \ldots, p(\psi_t(\boldsymbol{x})))_{\boldsymbol{x} \in [N]^D}$ is $O(M^{c_1}/A^{c_2})$-equidistributed on $\mathfrak{g}^\Psi(S_\bullet)$ with respect to $\mathcal{X}^\Psi$, the Hermite basis for $\mathfrak{g}^\Psi(S_\bullet)$ in $(\mathfrak{g}^t, \mathcal{X}^t)$.*

Proof. We let constants depend on s, d, D, t, Ψ. Suppose that p is (A, N)-linearly irrational in $(S_i)_{i=1}^s$. Write $p(n) = \sum_{i=1}^s a_i n^i$, where a_i is $(A, A/N^i)$-linearly irrational in S_i (see Definition 4.8).

Since M bounds the complexity of $S_\bullet$ and the height of the structure constants of $\mathfrak{g}$, the description of $\mathfrak{g}^\Psi(S_\bullet)$ in Proposition 3.1 and repeated application of Lemma 4.9 yields that $\mathfrak{g}^\Psi(S_\bullet)$ is of complexity at most $M^{O(1)}$ in $\mathfrak{g} \otimes \mathbb{R}^t$ with respect to $\mathcal{X} \otimes \mathcal{Z}$, where $\mathcal{Z}$ is the standard basis for $\mathbb{Z}^t$. Denote by $\mathcal{X}^\Psi$ the Hermite basis for $\mathfrak{g}^\Psi(S_\bullet)$ and $\mathcal{X} \otimes \mathcal{Z}$ and note that $\mathcal{X}^\Psi$ has rational structure constants of height at most $M^{O(1)}$. Let $\delta > 0$ be a parameter which we will choose later and suppose that p^Ψ is not δ-equidistributed on $[N]^D$ in $\mathfrak{g}^\Psi(S_\bullet)$ with respect to $\mathcal{X}^\Psi$.

Invoking Theorem 4.15 with respect to the basis $\mathcal{X}^\Psi$, we obtain a nontrivial $\eta \in \mathrm{Hom}_{\mathcal{X}^\Psi}(\mathfrak{g}^\Psi(S_\bullet), \mathbb{R})$ which has complexity at most $M^{O(1)}\delta^{-O(1)}$ and such that $\|\eta \circ p^\Psi\|_{C^\infty[N]^D} < M^{O(1)}\delta^{-O(1)}$. We have then, recalling the definition of the $C^\infty_{\mathcal{M}}$ norms from Definition A.2 and invoking Lemma A.3, that $\|s!\eta \circ p^\Psi\|_{C^\infty_{\mathcal{M}}[N]^D} < M^{O(1)}\delta^{-O(1)}$. Now we compute

$$s!\eta(p^\Psi(\boldsymbol{x})) = \sum_{i=1}^s s!\eta(a_i \otimes \Psi(\boldsymbol{x})^i) = \sum_{m \in \mathcal{M}} \sum_{i=1}^s s!\eta(a_i \otimes v_{m,i}) m(\boldsymbol{x}), \tag{7}$$

where, as in § 1.3, $v_{m,i} \in \mathbb{Z}^t$ is the coefficient of the monomial $m(\boldsymbol{x})$ in the polynomial $\Psi^i(\boldsymbol{x})$. Thus, by the definition of the $C^\infty_{\mathcal{M}}[N]^D$ norm Definition A.2, we have for all m, i that

$$\|s!\eta(a_i \otimes v_{m,i})\|_{\mathbb{R}/\mathbb{Z}} < M^{O(1)}\delta^{-O(1)}/N^i. \tag{8}$$

By Lemma 4.10, there is $\tilde{\eta} \in (\mathfrak{g} \otimes \mathbb{R}^t)^*$ of complexity $M^{O(1)}\delta^{-O(1)}$ with respect to $\mathcal{X} \otimes \mathcal{Z}$ such that $\tilde{\eta}|_{\mathfrak{g}^\Psi(S_\bullet)} = M^{O(1)}s!\eta$. It follows then that $\tilde{\eta}(\cdot \otimes v_{m,i})$ is an integer linear map on $\mathfrak{g}$ of complexity $M^{O(1)}\delta^{-O(1)}$ with respect to $\mathcal{X}$ and so invoking Lemma 4.10 once again, we have that $\tilde{\eta}(\cdot \otimes v_{m,i})|_{S_i}$ has complexity $M^{O(1)}\delta^{-O(1)}$ with respect to $\mathcal{X}_i'$, the Hermite basis for S_i.

Now note that (8) implies that $\|\tilde{\eta}(a_i \otimes v_{m,i})\|_{\mathbb{R}/\mathbb{Z}} < M^{O(1)}\delta^{-O(1)}/N^i$ for all i, m. Therefore, if $A \geq M^{O(1)}\delta^{-O(1)}$ (the complexity of the map $\tilde{\eta}(\cdot \otimes v_{m,i})|_{S_i}$) and $A \geq M^{O(1)}\delta^{-O(1)}$ so that $\|\tilde{\eta}(a_i \otimes v_{m,i})\|_{\mathbb{R}/\mathbb{Z}} < A/N^i$, then the $(A, A/N^i)$-linear irrationality of a_i in S_i implies that $\tilde{\eta}(\cdot \otimes v_{m,i})$ is trivial on S_i, i.e. $\tilde{\eta}(S_i \otimes v_{m,i}) = 0$. Since the $\{v_{m,i}\}_{m \in \mathcal{M}}$ span V^i_Ψ, we have $\tilde{\eta}(S_i \otimes V^i_\Psi) = 0$ and so from definitions $\eta(S_i \otimes V^i_\Psi) = 0$. Since η is a Lie algebra homomorphism to an abelian Lie algebra, η vanishes on the smallest Lie subalgebra of $\mathfrak{g}^\Psi(S_\bullet)$ containing the subspaces $S_i \otimes V^i_\Psi$, which is $\mathfrak{g}^\Psi(S_\bullet)$ itself. That is, η is trivial on $\mathfrak{g}^\Psi(S_\bullet)$, a contradiction. Thus, p^Ψ is indeed (upon solving for δ) $(M^{O(1)}/A^{O(1)})$-equidistributed in $\mathfrak{g}^\Psi(S_\bullet)$. □

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

5. The factorisation theorem for linear irrationality

As in the qualitative setting, a counting lemma is of little use without a method to reduce a general polynomial sequence to something it can handle. Our main result of this section is Theorem 5.10, a factorisation theorem for linear irrationality in much the same spirit as [GT10, Lemma 2.10] and [GT12, Theorem 1.19], which may be viewed as factorisation theorems for the notions of filtration and additive irrationality respectively. Throughout this section we assume d, s, t and D are of bounded size and let constants depend on them. As we did in the qualitative setting, we begin with an additive decomposition.

5.1 Additive decomposition for linear irrationality

In this subsection we prove an additive decomposition of vector space elements into linearly irrational, rational and small parts (Proposition 5.4), and as a corollary obtain a similar decomposition for polynomial sequences (Corollary 5.5).

DEFINITION 5.1 (Quantitative relative linear irrationality). Let $A>1$ and $\varepsilon>0$, let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$, and let U,T be rational subspaces. An element $a\in U$ is (A,ε)-*linearly irrational in* U mod T *with respect to* $\mathcal{X}$ if, for all integer linear maps l which have complexity at most A in U with respect to $\mathcal{X}'_{\mathcal{X},U}$, which are nontrivial on U and which vanish on $U\cap T$, we have that $\|l(a)\|_{\mathbb{R}/\mathbb{Z}}\geq\varepsilon$.

DEFINITION 5.2 (Quantitative rationality). Let $a\in\mathfrak{g}$, $A\in\mathbb{N}^+$. We say that a is A-rational with respect to $\mathcal{X}$ if $a\in\mathfrak{g}_{\mathbb{Q}}$ and, when written with respect to the basis $\mathcal{X}$, the vector for a has entries in $(1/A)\mathbb{Z}$. A polynomial sequence p in $\mathfrak{g}$ is A-rational if all of its coefficients are A-rational.

DEFINITION 5.3 (Quantitative smallness). Let $\varepsilon>0$. An element $a\in\mathfrak{g}$ is ε-small with respect to $\mathcal{X}$ if $\|a\|_\infty\leq\varepsilon$, where $\|\cdot\|_\infty$ is the ℓ_∞ norm of a with respect to the basis $\mathcal{X}$. A polynomial sequence $(p(\boldsymbol{n}))_{\boldsymbol{n}\in[N]^D}$ in $\mathfrak{g}$ is (A,N)-*small* with respect to $\mathcal{X}$ if, for all monomials $m\in\mathcal{M}$, we have that c_m is $A/N^{\deg m}$-small, where c_m is the coefficient of $m(\boldsymbol{n})$ in p.

The following decomposition of vector space elements into linearly irrational, rational and error terms allows us to perturb an element which is linearly irrational in a quotient space by elements of the quotient to obtain an element which is linearly irrational in the original space. Note that this result includes the case where the quotient space is trivial (that is, $T=U$ in the upcoming notation), whereupon we obtain a decomposition of an arbitrary element in the vector space.

PROPOSITION 5.4 (Quantitative decomposition of vector space elements). *Let $A>1$ be an integer, let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$, let U,T be rational subspaces of $\mathfrak{g}$, each of complexity $A^{O(1)}$. Then there is a constant $c>1$ (which may depend on $\dim U,\dim T$) such that the following is true. Let $0<\varepsilon<1/2$ and $a\in U$, and suppose that a is $(A^c,A^c\varepsilon)$-linearly irrational in U mod T. There are elements $a_s\in U\cap T$, $a_r\in(U\cap T)_{\mathbb{Q}}$ such that a_r is $A^{O(1)}$-rational, a_s is $A^{O(1)}\varepsilon$-small and $a-a_r-a_s$ is (A,ε)-linearly irrational in a subspace $U'\leq U$ of complexity $A^{O(1)}$ with respect to $\mathcal{X}$.*

Proof. In the first instance, observe that if a is (A,ε)-linearly irrational in U, then we are done and so supposing it is not, we have an integer linear map l of complexity at most A with respect to the Hermite basis $\mathcal{X}'_{U,\mathcal{X}}=\mathcal{X}'$ which is nontrivial on U such that $\|l(a)\|_{\mathbb{R}/\mathbb{Z}}<\varepsilon$. Note also that l must be nontrivial as a map on $U\cap T$ due to the linear irrationality of a in U mod T. The basic idea of the proof is to iterate this observation until we are able to

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

identify a_r, a_s as in the statement of the theorem. We now explain the induction by which this iteration is carried out, and why this process must eventually terminate, producing a_r, a_s as desired.

Let L_1 be the $1 \times \dim U$ matrix whose entries correspond to l with respect to $\mathcal{X}'$. This is the $m = 1$ (i.e. base) case of the following induction. By induction on m, we may suppose that we have an $m \times \dim U$ integer matrix $L = L_m$ whose rows are linearly independent, integer linear maps $U \to \mathbb{R}$ with respect to the basis $\mathcal{X}' := \mathcal{X}'_{U,\mathcal{X}}$ and which possess the following properties. First, the projection of the rows are linearly independent as maps in $(U \cap T)^*$. Furthermore, for each row l, we have that l has complexity $A^{O(1)}$ with respect to $\mathcal{X}'$ and that $\|l \cdot a\|_{\mathbb{R}/\mathbb{Z}} \leq A^{O(1)}\varepsilon$. We show in the following that either:

(i) there are elements $a_s \in U \cap T$, $a_r \in (U \cap T)_{\mathbb{Q}}$ such that with respect to $\mathcal{X}'$, a_r is $A^{O(1)}$-rational, $\|a_s\|_\infty = A^{O(1)}\varepsilon$ and $a - a_r - a_s$ is (A, ε)-linearly irrational in $\ker L \leq U$ (i.e. the proof is complete upon observing that with respect to $\mathcal{X}$, a_s is still $A^{O(1)}\varepsilon$-small and a_r is still $A^{O(1)}$-rational); or

(ii) we may find some L_{m+1} which fits the description above with $m + 1$ in place of m.

Note that in situation (i), we set $U' = \ker L$. In situation (ii), the constants implicit in the $O(1)$ notation may change when we pass from m to $m + 1$. However, this can occur only finitely many times since $\dim(U \cap T) = O(1)$ and the rows of each L_m are linearly independent in $(U \cap T)^*$. In particular, we must, at some point, find ourselves in situation (i), completing the proof. It remains to prove the above claim.

Let $L = L_m$ be as above. For $i = 1, \ldots, m$, we have by supposition that $l_i \cdot a = n_i + e_i$, for some integer n_i and $0 < |e_i| < \varepsilon$. We may write this simultaneously for $i = 1, \ldots, k$ as $La = n + e$ where n, e are vectors consisting of the n_i, e_i, respectively.

We claim that there is a solution to $Lx = n$ with $x \in U \cap T$ and such that each entry of x with respect to $\mathcal{X}'$ lies in $(1/A^{O(1)})\mathbb{Z}$; we let this be a_r. Let M be the change of basis matrix from $\mathcal{X}'' = \mathcal{X}''_{\mathcal{X},U\cap T}$ to $\mathcal{X}'$, so LM is the linear map corresponding to L where input vectors are given with respect to $\mathcal{X}''$. Note that M is an integer matrix since $\langle \mathcal{X}' \rangle = U \cap \langle \mathcal{X} \rangle \supset \mathcal{X}''$ and, furthermore, complexity bounds ensure that M has entries of size at most $A^{O(1)}$. Recall that by the induction statement L has rank m as a family of linear maps restricted to $U \cap T$. Thus, LM has rank m. Let L' be the invertible square submatrix of LM obtained by choosing m linearly independent columns. Then it is a standard linear algebra fact that the simultaneous equations $L'x' = n$ are solvable in $(1/|\det L'|)\mathbb{Z}$. Observing that $|\det L'| = A^{O(1)}$ and then padding the solution x' with $\dim(U \cap T) - m$ zero entries to form an appropriate vector $\tilde{x}$ yields a solution $LM\tilde{x} = n$ in $((1/A^{O(1)})\mathbb{Z})^{\dim U \cap T}$. Setting $x = M\tilde{x}$ proves the claim made at the beginning of the paragraph.

Next, we claim that there is a solution to $Lx = e$ with $x \in U \cap T$ and $\|x\|_\infty = A^{O(1)}\varepsilon$. This may be proven, for example, by choosing L' as above, setting $x' = L'^{-1}e$, upper bounding $\|L'^{-1}\|_\infty$ directly using Cramer's rule and then augmenting x' with zero entries to form a vector $\tilde{x}$ with $LM\tilde{x} = e$. Let $a_s = M\tilde{x} \in U \cap T \subset U$ so that $La_s = e$ and $\|a_s\|_\infty = A^{O(1)}\varepsilon$ (where here the ℓ_∞ norm is of course with respect to the basis $\mathcal{X}'$).

Now, $a_p := a - a_r - a_s$ is contained in $\ker L$. If a_p is (A, ε)-linearly irrational in $\ker L$, then we are in situation (i) and we are done, so suppose not. Then there is some nontrivial integer linear map l of complexity at most A in $\ker L$ such that $\|l(a_p)\|_{\mathbb{R}/\mathbb{Z}} < \varepsilon$. In addition, it follows from definitions that $\ker L$ is of complexity $A^{O(1)}$ and so by Lemma 4.10 we may find an integer linear map $\tilde{l} \in U^*$ of complexity $A^{O(1)}$ such that $\tilde{l}|_{\ker L} = A^{O(1)}l$. Furthermore, since l is nontrivial on

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

$\ker L$, we have that $\tilde{l}$ does not lie in the row span of L. We claim that, in fact, (the projection of) the rows of L and $\tilde{l}$ are linearly independent in $(U \cap T)^*$. Suppose not for a contradiction with the $(A^c, A^c\varepsilon)$-linear irrationality of a in $U \bmod T$. Then there is some linear combination $b\tilde{l} + \sum_{i=1}^m b_i l_i = 0 \in (U \cap T)^*$, where $b \neq 0$ because $\{l_1, \ldots, l_m\}$ are linearly independent in $(U \cap T)^*$. In fact, since the entries of each $l_i, \tilde{l}$ are integers of size $A^{O(1)}$, one may check, e.g., by standard Gaussian elimination that we can insist that the constants b_i, b are themselves integers of size $A^{O(1)}$ (with a different constant $O(1)$ which may depend on the dimension of the space). For these b, b_i, let $l' = b\tilde{l} + \sum_{i=1}^m b_i l_i \in (U \cap T)^\perp \subset U^*$. Then $l'(a) = l'(a_p)$ (since $a_r, a_s \in U \cap T$), $l'(a_p) = b\tilde{l}(a_p)$ (since $a_p \in \ker L$), and $b\tilde{l}(a_p) = bA^{O(1)} l(a_p)$ (by the construction of $\tilde{l}$ and since $a_p \in \ker L$). Thus,

$$\|l'(a)\|_{\mathbb{R}/\mathbb{Z}} = \|bA^{O(1)} l(a_p)\|_{\mathbb{R}/\mathbb{Z}} \leq bA^{O(1)} \|l(a_p)\|_{\mathbb{R}/\mathbb{Z}} < A^{O(1)}\varepsilon.$$

However, l' is of complexity $A^{O(1)}$ with respect to $\mathcal{X}'$. For c chosen suitably large, this contradicts the $(A^c, A^c\varepsilon)$-linear irrationality of a in $U \bmod T$. Thus, we have that (the projection of) the rows of L and $\tilde{l}$ are linearly independent in $(U \cap T)^*$.

Finally, let C be the smallest positive integer such that $C\tilde{l}(a_r) \in \mathbb{Z}$; note that $C = A^{O(1)}$. Then we compute

$$\begin{aligned}\|C\tilde{l}(a)\|_{\mathbb{R}/\mathbb{Z}} &\leq \|C\tilde{l}(a_p)\|_{\mathbb{R}/\mathbb{Z}} + \|C\tilde{l}(a_r)\|_{\mathbb{R}/\mathbb{Z}} + \|C\tilde{l}(a_s)\|_{\mathbb{R}/\mathbb{Z}} \\ &\leq CA^{O(1)}\varepsilon + 0 + CA^{O(1)}\varepsilon \leq A^{O(1)}\varepsilon.\end{aligned}$$

One may then form L_{m+1} by adding the row $C\tilde{l}$ to the bottom of $L = L_m$. This puts us in situation (ii) and completes the proof. □

COROLLARY 5.5 (Quantitative additive decomposition of polynomial sequences). *Let $A, N > 1$ be integers, let $U_\bullet = (U_i)_{i=1}^s$ be a sequence of rational subspaces of $\mathfrak{g}$ each of complexity $A^{O(1)}$ and let T be a rational subspace of $\mathfrak{g}$ which is also of complexity $A^{O(1)}$. There is a constant $c > 1$ (which may depend on $\dim U$, $\dim T$) such that the following is true. Let p be a polynomial sequence defined by $p(n) = \sum_{i=1}^s a_i n^i$ such that $a_i \in U_i$ for all i and which is (A^c, N)-linearly irrational in $U_\bullet \bmod T$.*[11] *Then there are polynomial sequences e, r in $U_\bullet \cap T$ of degree s and a sequence of subspaces $(U_i')_{i=1}^s \leq U_\bullet$ with each U_i' of complexity $A^{O(1)}$ such that e is $(A^{O(1)}, N)$-small, r is $A^{O(1)}$-rational, and $p - e - r$ is (A, N)-linearly irrational in $U_\bullet' := (U_i')_{i=1}^s$.*

Proof. Let $p(n) = \sum_{i=1}^s a_i n^i$. If p is (A, N)-linearly irrational in $U_\bullet$, then we are done so suppose it is not and conduct the following argument for each index i such that a_i fails to be $(A, A/N^i)$-linearly irrational in U_i. By definition, we have that each such a_i is $(A^c, A^c/N^i)$-linearly irrational in $U_i \bmod T$. Trivially then, a_i is $(A^{c-1}, A^c/N^i)$-linearly irrational in $U_i \bmod T$. Now invoke Proposition 5.4 for a_i and let c_i be a constant (obtained from the statement of Proposition 5.4) so that we may find $a_{e,i} \in U_i \cap T$ which is $(A^{O(1)}/N^i)$-small and $a_{r,i} \in U \cap T$ which is $A^{O(1)}$-rational such that $a_i' := a_i - a_{r,i} - a_{e,i}$ is $(A^{(c-1)/c_i}, A/N^i)$-linearly irrational in some subspace $U_i' \leq U_i$ of complexity $A^{O(1)}$. If $(c-1)/c_i \geq 1$, then a_i' is $(A, A/N^i)$-linearly irrational in U_i'; insist that c is chosen so that this is the case. Let $r(n) = \sum_{i=1}^s a_{r,i} n^i$ and $e(n) = \sum_{i=1}^s a_{e,i} n^i$. The result follows. □

[11] The operations are conducted pointwise on the sequence, i.e. this may be translated to mean that each coefficient a_i is $(A^c, A^c/N^i)$-linearly irrational in $U_i \bmod T$.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

5.2 Multiplicative factorisation for strong irrationality

The goal of this subsection is to prove Theorem 5.10. This may be viewed as analogous to Green and Tao's factorisation theorems for polynomial sequences (see [GT10, GT12]) for the notion of linear irrationality. We also want to control the following notion of irrationality which comes from [GT10] (see, in particular, [GT10, Appendix A]).

In the following, we need to recall the definition of the subalgebras $\mathfrak{h}_i$: Definition 2.10.

Definition 5.6 (Filtration irrationality). Let $A > 1, \varepsilon > 0$. We say $a \in \mathfrak{g}_i$ is (A, ε)-filtration irrational in $\mathfrak{g}_i$ with respect to a rational basis $\mathcal{X}$ for $\mathfrak{g}$ if $\|l(a)\|_{\mathbb{R}/\mathbb{Z}} \geq \varepsilon$ for all $l \in \mathfrak{g}_{\mathbb{Q}}^*$ which are nontrivial on $\mathfrak{g}_i$, have complexity at most A on $\mathfrak{g}_i$ and which vanish on $\mathfrak{h}_i$.

One may equivalently define $a \in \mathfrak{g}_i$ to be (A, ε)-filtration irrational in $\mathfrak{g}_i$ if and only if it is (A, ε)-linearly irrational in $\mathfrak{g}_i$ mod $\mathfrak{h}_i$, as per Definition 5.1.

Definition 5.7 (Filtration irrationality for polynomial sequences). Let $A, N > 1$ and let $\mathfrak{g}$ have rational basis $\mathcal{X}$. Let p be a polynomial sequence adapted to $\mathfrak{g}_\bullet$ which is defined by $p(n) := \sum_{i=1}^{s} a_i n^i$. We say that p is (A, N)-*filtration irrational* if a_i is $(A, A/N^i)$-filtration irrational in $\mathfrak{g}_i$ for every i. Furthermore, we say that p is (A, N)-*filtration irrational in* $\mathfrak{g}$ mod $\mathfrak{g}_j$ if a_i is $(A, A/N^i)$-filtration irrational in $\mathfrak{g}_i$ for $i = 1, \ldots, j-1$.

Before proving Theorem 5.10, we need some lemmas. Henceforth, for a univariate polynomial sequence q, let $(q)_j$ denote the coefficient of the term of degree j in q. The following equation is an easy consequence of definitions.

Lemma 5.8. *For polynomials q, q' adapted to $\mathfrak{g}_\bullet$ we have*

$$(q * q')_j = (q)_j + (q')_j \mod \mathfrak{h}_j,$$

for all j.

The upcoming lemma bounds the extent to which the filtration irrationality of an element $a \in \mathfrak{g}$ can be diminished under perturbation by a rational element and a small element. This will ultimately allow us to bound the extent to which factorising a polynomial sequence for linear irrationality diminishes the filtration irrationality of that sequence (see, in particular, the second dot point of Theorem 5.10).

Lemma 5.9. *Let $A > R > 1$ be integers and let $\delta, \varepsilon > 0$ be such that $\delta - (\dim \mathfrak{g}_i) A\varepsilon > 0$. Suppose that $a, r, s \in \mathfrak{g}_i$ such that a is (A, δ)-filtration irrational in $\mathfrak{g}_i$, r is R-rational in $\mathfrak{g}_i$ and s is ε-small. Then $a - r - s$ is $(A/R, (\delta - (\dim \mathfrak{g}_i) A\varepsilon)/R)$-filtration irrational in $\mathfrak{g}_i$.*

Proof. Let $\beta = (\delta - (\dim \mathfrak{g}_i) A\varepsilon)/R$ and $b = a - r - s$. We prove the contrapositive: suppose b is not $(A/R, \beta)$-filtration irrational. We may thus choose an integer linear map l, nontrivial on $\mathfrak{g}_i$, which vanishes on $\mathfrak{h}_i$ and with complexity A/R in $\mathfrak{g}_i$ such that $\|l(b)\|_{\mathbb{R}/\mathbb{Z}} < \beta$. Then Rl has complexity A on $\mathfrak{g}_i$ and

$$\|Rl(a)\|_{\mathbb{R}/\mathbb{Z}} \leq \|Rl(b)\|_{\mathbb{R}/\mathbb{Z}} + \|Rl(r)\|_{\mathbb{R}/\mathbb{Z}} + \|Rl(s)\|_{\mathbb{R}/\mathbb{Z}} < R\beta + 0 + (\dim \mathfrak{g}_i) A\varepsilon = \delta,$$

so a is not (A, δ)-filtration irrational. □

The main goal of the upcoming theorem is to provide a factorisation theorem for the notion of linear irrationality. Recall that we have already proven Proposition 2.28, a qualitative version of the upcoming theorem, which may be of interest to a reader wishing to extricate the important parts of the upcoming proof. In the quantitative setting, we also want our factorisation to have

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

the following feature: if we start with a sequence that is sufficiently filtration irrational, then we can factorise to a sequence which is both filtration irrational and linearly irrational. This property is also proven in the following theorem.

THEOREM 5.10 (Quantitative factorisation of polynomial sequences). *Let $N, A, M \geq 2$ and $d, s, t \geq 1$ be integers such that $AM > d$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra with filtration $\mathfrak{g}_\bullet$ of degree s, rational basis $\mathcal{X}$, dimension d and rational structure constants of height at most M. Let p be a polynomial sequence adapted to $\mathfrak{g}_\bullet$ and sequence of subspaces $S_\bullet \leq \mathfrak{g}_\bullet$. Finally, insist that $S_\bullet, \mathfrak{g}_\bullet$ are of complexity $A^{O(1)}$ in $\mathfrak{g}$, where here and henceforth $O(1)$ terms may depend on d, s. Then we may write $p = e * p' * r$ where e, p', r are polynomial sequences in $\mathfrak{g}$ adapted to $\mathfrak{g}_\bullet$, p' is (A, N)-linearly irrational in a sequence of subspaces $S'_\bullet$ of complexity $A^{O(1)}$, r is $(AM)^{O(1)}$-rational and e is $((AM)^{O(1)}, N)$-small. Furthermore:*

- *writing $p(n) = \sum_{i=1}^{s} a_i n^i$, if j is the smallest integer such that a_j is not (A, N)-linearly irrational in S_j, then we may take $S'_k = S_k$ for all $k = 1, \dots, j-1$ and have S'_j strictly contained in S_j;*
- *there is a constant $c > 1$ such that if $B > AM$ and p is (B^c, N)-filtration irrational, then p' is (B, N)-filtration irrational.*

Proof. The factorisation process proceeds inductively. We demonstrate the first couple of steps, from which the general process should be clear.

If p is (A, N)-linearly irrational in $S_\bullet$, then we are done trivially. Suppose not, let $b \geq 1$ be a constant to be chosen later, and write $p(n) = \sum_{i=1}^{s} a_i n^i$ and let j be the smallest integer such that a_j is not (A^b, N)-linearly irrational in S_j (since $b \geq 1$, such a j certainly exists). We begin by applying Corollary 5.5 to the degree $\geq j$ terms of p with $U_\bullet = S_\bullet$ and $T_\bullet = (\mathfrak{g}, \dots, \mathfrak{g})$ (note that p is arbitrarily linearly irrational in $U_\bullet$ mod T) to obtain $p_1 := p - \tilde{e}_1 - \tilde{r}_1$, where $\tilde{e}_1, \tilde{r}_1$ are polynomial sequences adapted to $\mathfrak{g}_\bullet$ with all terms of degree at least j, $\tilde{e}_1$ is $((AM)^{O(1)}, N)$-small, $\tilde{r}_1$ is $A^{O(1)}$-rational and p_1 is (A^b, N)-linearly irrational in some sequence of subspaces $S^{(1)}_\bullet$ with $S^{(1)}_k = S_k$ for $k < j$ and $S^{(1)}_j < S_j$. Then set

$$\tilde{p}_2 := \tilde{e}_1^{-1} * p * \tilde{r}_1^{-1} = p - \tilde{e}_1 - \tilde{r}_1 - \tfrac{1}{2}([\tilde{e}_1, p] + [\tilde{e}_1, r_1] + [p, \tilde{r}_1]) - \cdots .$$

Note that $\tilde{p}_2 = p_1 \mod \mathfrak{g}_2$ and, furthermore, that $(\tilde{p}_2)_k = (p_1)_k$ for all $k \leq j$ (since every term coming from a Lie bracket will have degree at least $j+1$). The former fact implies that $\tilde{p}_2$ is (A^b, N)-linearly irrational in $(S^{(1)}_\bullet + \mathfrak{g}_2) \bmod \mathfrak{g}_2$.

Next we apply Corollary 5.5 to the degree $> j$ terms of $\tilde{p}_2$ with $U_\bullet = S^{(1)}_\bullet + \mathfrak{g}_2$ and $T = (\mathfrak{g}_2, \dots, \mathfrak{g}_2)$. We obtain $p_2 := \tilde{p}_2 - \tilde{e}_2 - \tilde{r}_2$, where $\tilde{e}_2, \tilde{r}_2$ are polynomial sequences in $\mathfrak{g}_2$, adapted to $\mathfrak{g}_\bullet$ and with all terms of degree $> j$, $\tilde{e}_2$ is $(A^{O(1)}, N)$-small, $\tilde{r}_2$ is $A^{O(1)}$-rational and p_2 is $(A^{b/c_2}, N)$-linearly irrational (where c_2 is the constant from the statement of Corollary 5.5) in a sequence of subspaces $S^{(2)}_\bullet$ with $S^{(2)}_k = S^{(1)}_k = S_k$ for $k < j$ and $S^{(2)}_j = S^{(1)}_j < S_j$.

If $s = 2$ (i.e. $\mathfrak{g}_3 = 0$), we conclude

$$\begin{aligned} p &= \tilde{e}_1 * \tilde{p}_2 * \tilde{r}_1 \\ &= \tilde{e}_1 * (\tilde{e}_2 + p_2 + \tilde{r}_2) * \tilde{r}_1 \\ &= \tilde{e}_1 * \tilde{e}_2 * p_2 * \tilde{r}_2 * \tilde{r}_1, \end{aligned}$$

where in the final line we have used the fact that $\tilde{e}_2, \tilde{r}_2 \in \mathfrak{g}_2$, the Baker–Campbell–Hausdorff formula and that $[\mathfrak{g}, \mathfrak{g}_2] \subset \mathfrak{g}_3$. Furthermore, Lemma A.15 yields that $e_2 := \tilde{e}_1 * \tilde{e}_2$ is $((AM)^{O(1)}, N)$-small and $r_2 := \tilde{r}_2 * \tilde{r}_1$ is $(AM)^{O(1)}$-rational.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Otherwise, set

$$\tilde{p}_3 := \tilde{e}_2^{-1} * \tilde{p}_2 * \tilde{r}_2^{-1} = \tilde{p}_2 - \tilde{e}_2 - \tilde{r}_2 - \tfrac{1}{2}([\tilde{e}_2, \tilde{p}_2] + [\tilde{e}_2, \tilde{r}_2] + [\tilde{p}_2, \tilde{r}_2]) - \cdots .$$

Note that $\tilde{p}_3 = p_2 \mod \mathfrak{g}_3$ and, furthermore, that $(\tilde{p}_3)_k = (\tilde{p}_2)_k = (p_1)_k$ for all $k \leq j$. Apply Corollary 5.5 to the degree $> j$ terms of $\tilde{p}_3$ to obtain a linearly irrational p_3 and proceed as before.

If we begin with $b \geq \prod_{i=2}^{s} c_i$, then the process terminates with the polynomial sequence p_s which is (A, N)-linearly irrational in $S_\bullet^{(s)}$, a sequence of complexity $A^{O(1)}$ subspaces of $\mathfrak{g}$ with $S_k^{(s)} = S_k$ for all $k < j$ and $S_j^{(s)} = S_j^{(1)} < S_j$.

Now we proceed to the statement about filtration irrationality. Let c be a large constant whose value we will decide upon later and let $b_1 := c$. Suppose that p_i is (B^{b_i}, N)-filtration irrational in $\mathfrak{g} \bmod \mathfrak{g}_{i+1}$ and recall our (implicit) definitions

$$p_{i+1} := \tilde{p}_{i+1} - \tilde{e}_{i+1} - \tilde{r}_{i+1} := \tilde{e}_i^{-1} * (\tilde{e}_i + p_i + \tilde{r}_i) * \tilde{r}_i^{-1} - \tilde{e}_{i+1} - \tilde{r}_{i+1}, \tag{9}$$

where $\tilde{e}_i, \tilde{r}_i \in \mathfrak{g}_i$ and $\tilde{e}_{i+1}, \tilde{r}_{i+1} \in \mathfrak{g}_{i+1}$, and

$$e_i := \prod_{j=1}^{i} \tilde{e}_j, \quad r_i := \prod_{j=1}^{i} \tilde{r}_j. \tag{10}$$

It follows from (9) that $p_{i+1} = p_i$ modulo $\mathfrak{g}_{i+1}$. For $j = 1, \ldots, i$ we have that $\mathfrak{h}_j \geq \mathfrak{g}_{i+1}$ and so the jth coefficient of p_{i+1} inherits $(B^{b_i}, B^{b_i}/N^j)$-filtration irrationality from that of p_i. To check the filtration irrationality of p_{i+1} in $\mathfrak{g} \bmod \mathfrak{g}_{i+2}$, it remains to check the filtration irrationality of the $(i+1)$th coefficient of p_{i+1}. To this end, we analyse

$$\begin{aligned}(p_{i+1})_{i+1} &= (e_i^{-1} * p * r_i^{-1} - \tilde{e}_{i+1} - \tilde{r}_{i+1})_{i+1} \\ &= (e_i^{-1} * p * r_i^{-1})_{i+1} - (\tilde{e}_{i+1})_{i+1} - (\tilde{r}_{i+1})_{i+1} \\ &= (p)_{i+1} - (e_i)_{i+1} - (r_i)_{i+1} - (\tilde{e}_{i+1})_{i+1} - (\tilde{r}_{i+1})_{i+1} \mod \mathfrak{h}_{i+1},\end{aligned}$$

where we have used Lemma 5.8 in the second line. It follows from definitions and Lemma A.15 that both $(e_i)_{i+1}$ and $(\tilde{e}_{i+1})_{i+1}$ are $(AM)^{O(1)}/N^{i+1}$-small so that $(e_i)_{i+1} + (\tilde{e}_{i+1})_{i+1}$ is $(AM)^{O(1)}/N^{i+1}$-small by the triangle inequality. In addition, both $(r_i)_{i+1}$ and $(\tilde{r}_{i+1})_{i+1}$ are $(AM)^{O(1)}$-rational so $(r_i)_{i+1} + (\tilde{r}_{i+1})_{i+1}$ is $(AM)^{O(1)}$-rational. We have by supposition that $(p)_{i+1}$ is $(B^c, B^c/N^{i+1})$-filtration irrational, and so is trivially $(B^{c-k}, B^c/N^{i+1})$-filtration irrational for any constant $k > 0$. Invoking Lemma 5.9 (and to do so recalling that $B > AM > \dim \mathfrak{g}$), we have that $(p_{i+1})_{i+1}$ is $(B^{c-k}/(AM)^{O(1)}, (B^c/N^{i+1} - (\dim \mathfrak{g})B^{c-k}(AM)^{O(1)}/N^{i+1})/(AM)^{O(1)})$-filtration irrational in $\mathfrak{g}_{i+1}$, which in turn is also $(B^{c-k-O(1)}, (B^{c-O(1)} - B^{c-k+O(1)})/N^{i+1})$-filtration irrational since $B > AM > \dim \mathfrak{g}$. Setting $k = O(1)$ large enough depending on the other $O(1)$ terms, we may arrange for this coefficient to be $(B^{c-O(1)}, B^{c-O(1)}/N^{i+1})$-irrational. Therefore, p_{i+1} is $(B^{b_{i+1}}, N)$-filtration irrational where $b_{i+1} := \min(b_i, c - O(1))$ and $c = O(1)$ has been chosen sufficiently large. By induction, if $c = O(1)$ is chosen sufficiently large, p_s will indeed be (B, N)-filtration irrational. This completes the proof. □

As mentioned previously, the above theorem allows us to factorise for linear irrationality whilst maintaining some level of filtration irrationality. It will be convenient in future to refer to polynomial sequences which are both linearly and filtration irrational.

Definition 5.11 (Strong irrationality). If a polynomial sequence p is both (A, N)-filtration irrational in $\mathfrak{g}_\bullet$ and (A, N)-linearly irrational in some sequence of rational subspaces $S_\bullet$, then we say that it is (A, N)-*strongly irrational* in $(\mathfrak{g}_\bullet, S_\bullet)$.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

6. The strongly irrational arithmetic regularity lemma

In this section we prove our strongly irrational arithmetic regularity lemma, Theorem 1.4. The main ingredients are Green and Tao's non-irrational regularity lemma and the factorisation theory we developed in the previous section.

THEOREM 6.1 (Non-irrational arithmetic regularity lemma [GT10, Proposition 2.7]). *Let $N > 1$, let $f : [N] \to [0,1]$, let $s \geq 1$, let $\varepsilon > 0$ and let $\mathcal{F} : \mathbb{R}^+ \to \mathbb{R}^+$ be a growth function. Then there exists a quantity $M = O_{s,\varepsilon,\mathcal{F}}(1)$ and a decomposition*

$$f = f_{\mathrm{nil}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

of f into functions $f_{\mathrm{nil}}, f_{\mathrm{unf}} : [N] \to [-1,1]$ such that:

(i) *(f_{nil} structured) f_{nil} is a degree $\leq s$ polynomial nilsequence of complexity $\leq M$;*
(ii) *(f_{sml} small) $\|f_{\mathrm{sml}}\|_{L^2[N]} \leq \varepsilon$;*
(iii) *(f_{unf} very uniform) $\|f_{\mathrm{unf}}\|_{U^{s+1}[N]} \leq 1/\mathcal{F}(M)$;*
(iv) *(nonnegativity) f_{nil} and $f_{\mathrm{nil}} + f_{\mathrm{sml}}$ take values in $[0,1]$.*

In fact, Theorem 1.4 produces a (virtual) nilsequence which is both linearly irrational and filtration irrational. Although the latter is not strictly necessary to determine distribution and indeed is omitted in § 2, it is useful in applications (where otherwise one might have to replicate a similar factorisation on a case-by-case basis). Throughout this section we assume that d, s, D, t are of size $O(1)$.

The first step is to bootstrap our factorisation theory from the previous section to obtain a factorisation where p' is (much) more strongly irrational (by an arbitrary growth function, say) than e is small and r is rational. To do so, we first factorise for filtration irrationality using [GT10, Lemma 2.10] and then repeatedly factorise for linear irrationality using Proposition 5.10 and a reduction argument on the lexicographical ordering of the sequence $\dim(S_\bullet) := (\dim S_1, \dots, \dim S_s)$.

The following is what is needed from [GT10].

PROPOSITION 6.2 (Factorising for filtration irrationality [GT10, Lemma 2.10]). *Let $N, M_0 > 1$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra of dimension at most M_0 with filtration $\mathfrak{g}_\bullet$ of degree at most M_0. Let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$ with structure constants of height at most M_0. Let p be a polynomial sequence adapted to $\mathfrak{g}_\bullet$ with $p(0) = 0$. For any growth function $\mathcal{F} : \mathbb{R}^+ \to \mathbb{R}^+$, there exists $M_0 \leq M \leq O_{M_0,\mathcal{F}}(1)$ and a factorisation $p = e * p' * r$ where e is $(O_M(1), N)$-small, where r is $O_M(1)$-rational and where p' is $(\mathcal{F}(M), N)$-filtration irrational in a subfiltration $\mathfrak{g}'_\bullet \leq \mathfrak{g}_\bullet$ of complexity $O_M(1)$.[12] This filtration irrationality is measured with respect to $\mathcal{X}'$, a Mal'cev basis[13] for $\mathfrak{g}'_\bullet$ which may be written as a linear combination of $\mathcal{X}$ with rational coefficients of height $O_M(1)$.*

Proof. The result [GT10, Lemma 2.10] is phrased in slightly different language and there is some checking to be done to show that their statement implies ours. We do this in Appendix A.4. □

The following proposition is the 'strongly irrational' analogue of the previous proposition.

PROPOSITION 6.3 (Factorising for strong irrationality). *Let $N, M_0 > 1$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra of dimension at most M_0 with filtration $\mathfrak{g}_\bullet$ of degree at most M_0. Let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$ with structure constants of height at most M_0. Let p be a polynomial*

[12] That is, for all i, $\mathfrak{g}'_i$ is a subspace of $\mathfrak{g}_i$ of complexity $O_M(1)$.
[13] See [GT12, Definition 2.1].

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

sequence adapted to $\mathfrak{g}_\bullet$ *with* $p(0)=0$. *For any growth function* $\mathcal{F}:\mathbb{R}^+\to\mathbb{R}^+$, *there exists* $M_0\leq M\leq O_{M_0,\mathcal{F}}(1)$ *and a factorisation* $p=e*\tilde{p}*r$, *where* e *is* $(O_M(1),N)$*-small, where* r *is* $O_M(1)$*-rational and where* $\tilde{p}$ *is* $(\mathcal{F}(M),N)$*-strongly irrational in some* $(\mathfrak{g}'_\bullet,S_\bullet)$ *where* $\mathfrak{g}'_\bullet$ *is a subfiltration of* $\mathfrak{g}_\bullet$ *of complexity* $O_M(1)$ *in* $\mathfrak{g}$ *and* $S_\bullet$ *is a sequence of subspaces in* $\mathfrak{g}'_\bullet$ *of complexity* $O_M(1)$. *The filtration irrationality is measured with respect to* $\mathcal{X}'$, *a Mal'cev basis for* $\mathfrak{g}'_\bullet$ *where each element may be written as a linear combination of* $\mathcal{X}$ *with rational coefficients of height at most* $O_M(1)$.

Proof. Let $\mathcal{F}'$ be a growth function depending on $\mathcal{F}$ and M_0 to be determined later. For the meantime, insist that $\mathcal{F}'\geq(\mathcal{F}\circ\mathcal{F})^c$ at all inputs, where c is the constant from the statement of Theorem 5.10. Invoke Proposition 6.2 to factorise $p=e*p'*r$ whereupon there is $M'_0=O_{M_0,\mathcal{F}'}(1)=O_{M_0,\mathcal{F}}(1)$ such that p' is $(\mathcal{F}'(M'_0),N)$-filtration irrational in some $(\mathfrak{g}'_\bullet,\mathcal{X}')$ of complexity $O_{M'_0}(1)$, e is $O_{M'_0}(1)$-small and r is $O_{M'_0}(1)$-rational. Note also that $\mathfrak{g}'$ has rational structure constants of height $O_{M'_0}(1)$ with respect to $\mathcal{X}'$ since the elements of $\mathcal{X}'$ are rational linear combinations of the elements of $\mathcal{X}$ with coefficients of height $O_{M'_0}(1)$.

If p' is $(\mathcal{F}(M'_0),N)$-linearly irrational in $S_\bullet=\mathfrak{g}'_\bullet$, then we are done upon setting $M:=M'_0$. Otherwise, set $M_1:=\mathcal{F}(M'_0)=O_{M_0,\mathcal{F}}(1)$ and invoke Theorem 5.10 to obtain a factorisation $p'=e_1*p_1*r_1$ where e_1 is $(O_{M_1}(1),N)$-small, p_1 is (M_1,N)-linearly irrational in some sequence of subspaces $S^{(1)}_\bullet\leq\mathfrak{g}'_\bullet$ of complexity $O_{M_1}(1)$, and r_1 is $O_{M_1}(1)$-rational. Note that p_1 is (M_1,N)-linearly irrational in $S^{(1)}_\bullet$ but is not (M_1,N)-linearly irrational in $\mathfrak{g}'_\bullet$, and so we may conclude from the first bullet point of Theorem 5.10 that $\dim(S_\bullet)$ (that is, the vector of dimensions) is lexicographically strictly less than $\dim(\mathfrak{g}'_\bullet)$. Recall also that $\mathcal{F}'(M'_0)\geq\mathcal{F}(\mathcal{F}(M'_0))^c=\mathcal{F}(M_1)^c$ and so the $(\mathcal{F}'(M'_0),N)$-filtration irrationality of p' allows us to conclude as per the second bullet point of Theorem 5.10 that p_1 is $(\mathcal{F}(M_1),N)$-filtration irrational in $\mathfrak{g}'_\bullet$.

If, in fact, p_1 is $(\mathcal{F}(M_1),N)$-linearly irrational in $S^{(1)}_\bullet$, then observe that e_1 and r_1 are $(O_{M_0,\mathcal{F}}(1),N)$-small and $O_{M_0,\mathcal{F}}(1)$-rational, respectively, with respect to $\mathcal{X}$ and so we are done with $M=M_1$. Otherwise, set $M_2:=\mathcal{F}(M_1)$, insist that $\mathcal{F}'\geq(\mathcal{F}\circ\mathcal{F}\circ\mathcal{F})^c$ at all inputs, and proceed as before to obtain some p_2 which is (M_2,N)-linearly irrational on a sequence of subspaces $S^{(2)}_\bullet$ with $\dim(S^{(2)}_\bullet)$ lexicographically strictly less than $\dim(S^{(1)}_\bullet)$. Now we have $p=e_1*e_2*p_2*r_2*r_1$, where one may use Lemma A.15 to see that e_1*e_2 and r_2*r_1 are $(O_{M_0,\mathcal{F}}(1),N)$-small and $O_{M_0,\mathcal{F}}(1)$-rational, respectively, with respect to $\mathcal{X}$. If, in fact, p_2 is $(\mathcal{F}(M_2),N)$-linearly irrational, then we are done; if not, proceed as above. Since $\mathfrak{g}'$ has dimension at most M_0 and filtration of degree at most M_0, the vector $\dim(S^{(i)}_\bullet)$ can lexicographically decrease at most $O_{M_0}(1)$ times. Thus, the process must terminate at some p_i which is $(\mathcal{F}(M_i),N)$-strongly irrational in some $(\mathfrak{g}',S^{(i)}_\bullet)$ of complexity $O_{M_i}(1)$, with e_i which is $(O_{M_i}(1),N)$-small, and r_i which is $O_{M_i}(1)$-rational. This proves the claim. □

For convenience, in this section we deal only with (virtual) nilsequences for which the rational basis 'passes through' the filtration $\mathfrak{g}_\bullet$. Since the non-irrational arithmetic regularity lemma, Theorem 6.1, produces nilsequences for which this restriction holds, we do not practically lose anything by making this simplifying assumption.

Definition 6.4. We say that a basis $\mathcal{X}:=\{X_1,\dots,X_n\}$ for a vector space S *passes through* a sequences of rational subspaces $S_1\geq S_2\geq\dots$ if $\{X_{n-\dim S_i+1},\dots,X_n\}$ is a basis for S_i for all i. We may also say that each S_i passes through $\mathcal{X}$.

Definition 6.5 (Virtual nilsequence). A *virtual nilsequence* of complexity M at scale N is a function $\phi:[N]\to\mathbb{C}$ of the form $\phi(n)=f(p(n)*\Lambda,n\pmod q,n/N)$ where $q\leq M$ is a positive

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

integer, p is a polynomial sequence with respect a filtered nilpotent Lie algebra $\mathfrak{g}_\bullet$ of dimension $d \leq M$ and degree $s \leq M$, Λ is an additive lattice in $\mathfrak{g}_\bullet$ with $\mathbb{Z}$-basis $\mathcal{X}$ which passes through $\mathfrak{g}_\bullet$ and for which $\mathfrak{g}$ has rational structure constants of height at most M and f is Lipschitz with Lipschitz constant at most M.[14]

Let $S_\bullet$ be a sequence of subspaces of $\mathfrak{g}_\bullet$ of complexity at most M. We, furthermore, say that ϕ is (A, N)-filtration (respectively, linearly/strongly) irrational in $\mathfrak{g}_\bullet$ (respectively, $(\mathfrak{g}_\bullet, S_\bullet)$) if p is (A, N)-filtration (respectively, linearly/strongly) irrational in $\mathfrak{g}_\bullet$ (respectively, $(\mathfrak{g}_\bullet, S_\bullet)$) with respect to $\mathcal{X}$.

Remark 6.6. In the above, Λ may not be a multiplicative subgroup (i.e. with respect to the $*$ operation), in which case this definition is equivalent to asking that f is automorphic with respect to the multiplicative closure of Λ in $\mathfrak{g}$. We see in Appendix A.2 that this multiplicative closure is indeed a multiplicative lattice which is not too much larger than Λ.

Our conclusion to the argument follows a similar path to that of [GT10, Theorem 1.2] and may essentially be viewed as a Lie algebra analogue of what is done there. We note/caution that in the following proposition, the notion of complexity of a nilsequence used is that of Green and Tao, whereas the notion of complexity of a virtual nilsequence is that of our Definition 6.5.

PROPOSITION 6.7. *Let $s > 0$, $N, M_0 > 1$, let $\mathcal{F} : \mathbb{R}^+ \to \mathbb{R}^+$ be a growth function, and let $\phi : \mathbb{Z} \to [0, 1]$ be a degree $\leq s$ nilsequence of complexity $\leq M_0$. Then there exists $M = O_{s,M_0,\mathcal{F}}(1)$ such that f (when restricted to $[N]$) is also a $(\mathcal{F}(M), N)$-strongly irrational degree $\leq s$ virtual nilsequence of complexity $\leq M$.*

Proof. As usual, we let all constants depend on s. Write $\phi(n) = f(p(n)) = f(p(n) * \log \Gamma)$ where p is a polynomial sequence on some filtered nilpotent Lie algebra $\mathfrak{g}_\bullet$ and where f is automorphic with respect to some lattice $\Gamma \leq G$. Let $\mathcal{X}_{\text{Mal}}$ be the Mal'cev basis which exhibits ϕ as having complexity $\leq M_0$. By replacing p with $p * [p(0)]^{-1}$,[15] we may assume that the coordinates of $p(0)$ are of size $M_0^{O(1)}$ (this will be needed in the proof of the Lipschitz property in Appendix A.5). By [GT12, Lemma A.8], there is an integer $q_0 = M_0^{O(1)}$ such that $\mathcal{X} := q_0 \mathcal{X}_{\text{Mal}}$ yields $\langle \mathcal{X} \rangle \subset \log \Gamma$. Since $\mathcal{X}_{\text{Mal}}$ has structure constants of height at most M_0, we have that $\mathcal{X}$ has structure constants of height $M_0^{O(1)}$. Write $p = p(0) * (p(0)^{-1} * p)$, invoke Proposition 6.3 on the bracketed polynomial (with M_0 replaced by $M_0^{O(1)}$ and for an unspecified growth function $\mathcal{F}'$) and inherit the notation from the statement of that proposition so $\phi(n) = f(p(0) * e(n) * p'(n) * r(n) * \log \Gamma)$. From Lemma A.16 we have that $r(n) * \log \Gamma$ is q-periodic, where $q = O_M(1)$ and we recall that $M = O_{M_0,\mathcal{F}'}(1)$ from Proposition 6.3. Now define $\tilde{f} : \mathfrak{g} \times \mathbb{Z}/q\mathbb{Z} \times [0, 1] \to [0, 1]$ by

$$\tilde{f}(x, s, y) := f(p(0) * e(Ny) * x * r(s) * \log \Gamma),$$

whenever $y \in (1/N)\mathbb{Z}$ and by Lipschitz extension to $[0, 1]$. Note that the q-periodicity of the orbit $r(n) * \log \Gamma$ yields that $\tilde{f}$ is indeed a well-defined function on its second argument.

[14] As in [GT10], the underlying choice of metric does not matter too much. It may be taken to be the direct sum of the metrics on each of the factors, where the metric on the first factor can be adopted from [GT12, Definition 2.2], the metric on $\mathbb{Z}/q\mathbb{Z}$ is obtained by embedding this group into $\mathbb{R}/\mathbb{Z}$ and the metric on $[0, 1]$ is just the usual metric there. Details are worked out in Appendix A.5.

[15] Here we recall the notation from [GT12, Lemma A.14]: given $g \in G$, we may write $g = \{g\}[g]$ where $\{g\}$ has Mal'cev coordinates in $[0, 1)$, and $[g]$ lies in Γ. One may interpret this notation analogously in the Lie algebra.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Next we claim that there exists a positive integer $l = O_M(1)$ such that $\tilde{f}(x * (l\langle\mathcal{X}\rangle), s, y) = \tilde{f}(x, s, y)$, so that $\tilde{f}$ is automorphic with respect to the additive lattice $l\langle\mathcal{X}\rangle$ in its first argument. Since $\langle\mathcal{X}\rangle \subset \log\Gamma$, it suffices to find l such that for all $n \in \mathbb{Z}$ we have $(l\langle\mathcal{X}\rangle) * r(n) \subset r(n) * \langle\mathcal{X}\rangle$. However, the existence of such an l of size $O_M(1)$ follows easily from the fact that $r(n)$ is $O_M(1)$-rational with respect to $\mathcal{X}$, the Baker–Campbell–Hausdorff formula and our bound on the heights of the rational structure constants of $\mathfrak{g}$ with respect to the basis $\mathcal{X}$. Let $\Lambda = l\langle\mathcal{X}\rangle = \langle l\mathcal{X}\rangle$. Since p' is $(\mathcal{F}'(M), N)$-strongly irrational in $(\mathfrak{g}_\bullet, S_\bullet)$ with respect to $\mathcal{X}$, it is $(\gg_M \mathcal{F}'(M), N)$-strongly irrational in $(\mathfrak{g}_\bullet, S_\bullet)$ with respect to $l\mathcal{X}$.

We claim that $\tilde{f}$ is $O_M(1)$-Lipschitz; the details of this claim are worked out in Appendix A.5. The result then follows after replacing M with a suitable quantity $O_M(1)$ and letting $\mathcal{F}$ be sufficiently rapidly growing depending on $\mathcal{F}'$. □

We are finally ready to recover the main result of this section.

Theorem 6.8 (Strongly irrational arithmetic regularity lemma, Theorem 1.4). *Let $f : [N] \to [0,1]$, let $s \geq 1$, let $\varepsilon > 0$ and let $\mathcal{F} : \mathbb{R}^+ \to \mathbb{R}^+$ be a growth function. Then there exists a quantity $M = O_{s,\varepsilon,\mathcal{F}}(1)$ and a decomposition*

$$f = f_{\mathrm{nil}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

of f into functions $f_{\mathrm{nil}}, f_{\mathrm{unf}} : [N] \to [-1,1]$ such that:

(i) *(f_{nil} structured) f_{nil} is a degree $\leq s$, $(\mathcal{F}(M), N)$-strongly irrational virtual nilsequence of complexity $\leq M$;*
(ii) *(f_{sml} small) $\|f_{\mathrm{sml}}\|_{L^2[N]} \leq \varepsilon$;*
(iii) *(f_{unf} very uniform) $\|f_{\mathrm{unf}}\|_{U^{s+1}[N]} \leq 1/\mathcal{F}(M)$;*
(iv) *(nonnegativity) f_{nil} and $f_{\mathrm{nil}} + f_{\mathrm{sml}}$ take values in $[0,1]$.*

Proof. This is an easy corollary of Theorem 6.1 and Proposition 6.7. □

Remark 6.9. Technically a strongly irrational virtual nilsequence consists of data $(p, \mathfrak{g}, \mathcal{X}, \mathfrak{g}_\bullet, S_\bullet)$. However, recalling Lemma 4.11, if $\mathcal{F}$ is sufficiently rapidly growing, we may drop the data $S_\bullet$ since it may be uniquely recovered from the polynomial sequence p. Furthermore, by Lemmas 7.1 and 3.6, one may also recover the filtration $\mathfrak{g}_\bullet$ from p. Thus, for sufficiently strongly irrational virtual nilsequences, the data $(p, \mathfrak{g}, \mathcal{X})$ suffices.

7. Applications

7.1 Recovering the flag counting lemma

This subsection is not an application as much as it is a sanity check (Corollary 7.2) that in the flag case, Green and Tao's arithmetic regularity lemma and counting lemma may be recovered by the results in this article. Its purpose is also to prove Lemma 7.1 which is useful in other applications.

Recall Definition 2.10, that of the subalgebras $\mathfrak{h}_i$ associated to the filtered Lie algebra $\mathfrak{g}_\bullet$. The following is a quantitative version of Lemma 3.5.

Lemma 7.1. *Suppose that p is (A, N)-filtration irrational in $\mathfrak{g}_\bullet$ where the rational basis $\mathcal{X}$ passes through $\mathfrak{g}_\bullet$. Then, for each $i = 1, \dots, s$, for any subspace $T_i \leq \mathfrak{g}_i$ which contains a_i and has complexity at most A, we have $T_i + \mathfrak{h}_i = \mathfrak{g}_i$.*

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Proof. From Definition 2.10 and the filtration on $\mathfrak{g}_\bullet$ we have $\mathfrak{h}_i \leq \mathfrak{g}_i$ and so $T_i + \mathfrak{h}_i \leq \mathfrak{g}_i$. The slightly less-trivial containment to show is $T_i + \mathfrak{h}_i \supseteq \mathfrak{g}_i$. Let $Q = Q(A)$ be the $\mathbb{R}$-span of all complexity $\leq A$ linear maps on $\mathfrak{g}_i$.[16] Since a_i is $(A, A/N^i)$-filtration irrational, we have, in particular, that $l(a_i) \neq 0$ for all nontrivial $l \in Q$ which vanish on $\mathfrak{h}_i$. Thus, in $\mathfrak{g}_i^*$ we have $a_i^\perp \cap Q \cap \mathfrak{h}_i^\perp = \{0\}$ so $(a_i^\perp \cap Q)^\perp + \mathfrak{h}_i = \mathfrak{g}_i$. However, $\mathfrak{g}^*_{i\mathcal{X}}(A) \subset Q$ where $\mathfrak{g}^*_{i\mathcal{X}}(A)$ is the set of integer linear maps on $\mathfrak{g}_i$ of complexity at most A, and so $(a_i^\perp \cap Q)^\perp \subset (a_i^\perp \cap \mathfrak{g}^*_{i\mathcal{X}}(A))^\perp$. However, $(a_i^\perp \cap \mathfrak{g}^*_{i\mathcal{X}}(A))^\perp$ is, by definition, the smallest rational subspace of $\mathfrak{g}_i$ which contains a_i and has complexity at most A. Therefore, $T_i \supset (a_i^\perp \cap \mathfrak{g}^*_{i\mathcal{X}}(A))^\perp$ and, ultimately, $T_i + \mathfrak{h}_i \supset \mathfrak{g}_i$. □

COROLLARY 7.2 (Theorem 1.5). *Let $N, A, M \geq 2$. Let $\Psi = (\psi_1, \dots, \psi_t)$ be a flag system of linear forms, each mapping $\mathbb{Z}^D$ to $\mathbb{Z}$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra of dimension d. Let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$ which passes through a degree s filtration $\mathfrak{g}_\bullet$ and with respect to which $\mathfrak{g}$ has rational structure constants of height at most M. Let $S_\bullet \leq \mathfrak{g}_\bullet$ be a sequence of subspaces of complexity at most M. If p is (A, N)-strongly irrational in $(\mathfrak{g}_\bullet, S_\bullet)$, then there are constants $0 < c_1, c_2 = O_{d,s,D,t,\Psi}(1)$ such that p^Ψ is $O(M^{c_1}/A^{c_2})$-equidistributed on $[N]^D$ in $\log G^\Psi$, where G^Ψ is the Leibman group for $G_\bullet, \Psi$.*

Proof. This follows from our counting lemma (Theorem 1.2), Lemma 7.1 and Proposition 3.7. □

7.2 Another proof of the Gowers–Wolf conjecture

In their original 'arithmetic regularity lemma and counting lemma' paper, Green and Tao proved the Gowers–Wolf conjecture [GW10, Conjecture 2.5] for flag systems of linear forms.

THEOREM 7.3 [GT10, Theorem 1.13]. *Let $N \geq 2$. Let $\Psi = (\psi_1, \dots, \psi_t)$ be a collection of linear forms each mapping $\mathbb{Z}^D$ to $\mathbb{Z}$ which satisfy the flag condition, and let $s \geq 1$ be an integer such that the polynomials $\psi_1^{s+1}, \dots, \psi_t^{s+1}$ are linearly independent. For $i = 1, \dots t$, let $f_i : [-N, N] \to \mathbb{C}$ be functions bounded in magnitude by 1 (and defined to be zero outside of $[-N, N]$). For all $\varepsilon > 0$ there exists $\delta > 0$ such that if $\min_i \|f_i\|_{U^{s+1}[-N,N]} \leq \delta$, then*

$$\left| \mathbb{E}_{\boldsymbol{x} \in [-N,N]^D} \prod_{i=1}^{t} f_i(\psi_i(\boldsymbol{x})) \right| \leq \varepsilon.$$

The restriction to flag patterns in the above theorem is a consequence of the use of the arithmetic regularity lemma and complementary counting lemma from [GT10], which itself is restricted to flag patterns. The author [Alt22] recently resolved the full result by using a linear algebraic trick to reduce to the flag case whereupon the flag arithmetic regularity lemma and counting lemma from [GT10] sufficed.[17]

THEOREM 7.4 [Alt22, Theorem 1.1]. *Theorem 7.3 holds without the restriction that Ψ satisfies the flag condition.*

Our arithmetic regularity lemma, Theorem 1.4, and counting lemma, Theorem 1.2, mean that the original strategy from [GT10] may be applied directly to resolve the Gowers–Wolf conjecture in full generality. We do not reiterate the proof here since it may be fairly cleanly decomposed as a union of what is already done in [GT10, Theorem 1.13] and the upcoming Proposition. We

[16] Since $\mathcal{X}$ passes through $\mathfrak{g}_\bullet$ the complexity of a map on $\mathfrak{g}$ is equal to that of its restriction on any $\mathfrak{g}_i$.
[17] The reader might also be interested to consult [Man21] where, even more recently, Manners resolved the problem over finite fields with polynomial bounds [Man21, Theorem 1.1.5].

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

do note that along the way we will need to obtain a version of Theorem 1.2 which allows for averages over shifted sublattices of $\mathbb{Z}^D$. This is directly analogous to how [GT10, Theorem 3.6] differs from [GT12, Theorem 8.6]. Such a pursuit is routine but technical and tedious; we do not pursue it here to avoid adding further length to this document.

Recall the definition of the sequence of subspaces $(W_i)_{i=1}^s$ given a sequence of subspaces $(S_i)_{i=1}^s$: let $W_1 = S_1$ and iteratively define $W_i = \mathrm{span}\{S_j, [W_j, W_{i-j}]$ for $j = 1, \dots, i-1\}$ for $i = 2, \dots, s$. Recall also the definition of the subalgebras $\mathfrak{h}_i$ given a filtration $\mathfrak{g}_\bullet$: Definition 2.10.

Proposition 7.5. *Suppose $A > M \geq 1$. Suppose that p is (A, N)-strongly irrational in $(\mathfrak{g}_\bullet, S_\bullet)$ where $\mathfrak{g}$ has rational structure constants of height at most M with respect to a basis $\mathcal{X}$ of $\mathfrak{g}$ which passes through $\mathfrak{g}_\bullet$, that the elements of $S_\bullet$ have complexity at most M with respect to $\mathcal{X}$, and that Ψ has the property that $V_\Psi^j = \mathbb{R}^t$. Then there are constants $0 < c_1, c_2 = O_{d,s,D,t,\Psi}(1)$ such that p^Ψ is $O(M^{c_1}/A^{c_2})$-equidistributed on $[N]^D$ in $\mathfrak{g}^\Psi(S_\bullet) \leq \mathfrak{g}^t \cong \mathfrak{g} \otimes \mathbb{R}^t$, where $\mathfrak{g}^\Psi(S_\bullet)$ contains $\mathfrak{g}_j \otimes \mathbb{R}^t$.*

Proof. The counting lemma, Theorem 1.2, and the (A, N)-linear irrationality of p yields that p^Ψ is $O(M^{c_1}/A^{c_2})$-equidistributed on $\mathfrak{g}^\Psi(S_\bullet)$; we need to show that $\mathfrak{g}^\Psi(S_\bullet)$ contains $\mathfrak{g}_j \otimes \mathbb{R}^t$. First, $V_\Psi^j = \mathbb{R}^t$ implies that $V_\Psi^i = \mathbb{R}^t$ for all $i \geq j$ and so by Proposition 3.1, $\mathfrak{g}^\Psi(S_\bullet)$ contains $\sum_{i=j}^s W_j \otimes \mathbb{R}^t$. In addition, the (A, N)-filtration irrationality of p ensures by Lemma 7.1 that $S_i + \mathfrak{h}_i = \mathfrak{g}_i$ for all i, and so $W_i + \mathfrak{g}_{i+1} = \mathfrak{g}_i$ for all i by Lemma 3.6. Thus $\sum_{i=j}^s W_i = \mathfrak{g}_j$ and so $\mathfrak{g}^\Psi(S_\bullet)$ does indeed contain $\mathfrak{g}_j \otimes \mathbb{R}^t$. □

Corollary 7.6. *Theorem 7.3 holds without the restriction that Ψ satisfies the flag condition.*

Proof. Apply the argument from the proof of [GT10, Theorem 1.13] with our strongly irrational arithmetic regularity lemma, Theorem 1.4, in place of [GT10, Theorem 1.2], and use Proposition 7.5 in place of [GT10, Theorem 1.11]. □

Appendix A. Details on polynomial sequences in Lie groups and Lie algebras

In this appendix we may use terminology from [GT10, Appendix A] and [GT12, Appendix A] without further introduction. As we have throughout this article, we assume here that the dimension d and degree s of $\mathfrak{g}_\bullet$ are of size $O(1)$.

A.1 Smoothness norms

Green and Tao introduced the following 'smoothness norms'.

Definition A.1 [GT12, Definition 8.2]. Let $N_1, \dots, N_D \geq 1$ be integers and let $[\boldsymbol{N}]$ denote the set $[N_1] \times \cdots \times [N_D]$. Let $f : \mathbb{Z}^D \to \mathbb{R}/\mathbb{Z}$ be a polynomial map with Taylor expansion $f(\boldsymbol{n}) = \sum_{\boldsymbol{j}} \alpha_{\boldsymbol{j}} \binom{\boldsymbol{n}}{\boldsymbol{j}}$. Then define

$$\|f\|_{C^\infty[\boldsymbol{N}]} := \sup_{\boldsymbol{j} \neq 0} \prod_{i=1}^D N_i^{j_i} \|\alpha_{\boldsymbol{j}}\|_{\mathbb{R}/\mathbb{Z}}.$$

The main theorems in [GT10] and [GT12] are written with respect to these smoothness norms. Note that one may naturally extend this definition to polynomials $f : \mathbb{R}^D \to \mathbb{R}$. It is convenient for us to work with a slightly different definition. Recall that we use $\mathcal{M}$ to denote the set of 'monomial maps' which take $\boldsymbol{x} = (x_1, \dots, x_D)$ to a monomial in the variables $\{x_1, \dots, x_D\}$. The monomial map which takes $\boldsymbol{x}$ to the constant 1 is denoted by 1.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

DEFINITION A.2 (Smoothness norms with respect to monomial basis). Let $f:\mathbb{R}^D\to\mathbb{R}$ have monomial expansion $f(\boldsymbol{x})=\sum_{m\in\mathcal{M}}\beta_m m(\boldsymbol{x})$. Then define

$$\|f\|_{C^\infty_{\mathcal{M}}[N]^D}:=\sup_{1\neq m\in\mathcal{M}}N^{\deg m}\|\beta_m\|_{\mathbb{R}/\mathbb{Z}}.$$

LEMMA A.3. *Let $f:\mathbb{R}^D\to\mathbb{R}$ be a multivariate polynomial of degree s. Then $\|s!f\|_{C^\infty_{\mathcal{M}}[N]^D}\ll_{s,D}\|f\|_{C^\infty[N]^D}$.*

Proof. Let $\varepsilon=\|f\|_{C^\infty[N]^D}$.

By definition, we may write $f=f_n+f_\varepsilon$, where, when developed with respect to the multinomial basis, the coefficients of f_n are integers and the coefficients of the degree d terms of f_ε are of magnitude at most ε/N^d. Writing f_n with respect to the monomial basis yields coefficients in $(1/s!)\mathbb{Z}$ and so $s!f_n$ has integer valued coefficients (with respect to the monomial basis). Writing $s!f_\varepsilon$ with respect to the monomial basis yields coefficients of degree d terms with magnitude at most $O_{s,D}(\varepsilon/N^d)$. □

A.2 Equidistribution in the Lie algebra

The goal of this subsection is to explain what is meant by, and to justify, Theorems 2.8 and 4.15. We deduce them from theorems in [Lei05] and [GT12] respectively. We first deal with the qualitative setting (i.e. deducing Theorem 2.8 from [Lei05]).

When we speak of qualitative equidistribition on a nilmanifold, we mean the following.

DEFINITION A.4 (Qualitative nilmanifold equidistribution). Let G/Γ be a nilmanifold. Recall that this comes with a unique normalised Haar measure, with respect to which we will integrate. A sequence $(g(\boldsymbol{n})\Gamma)_{\boldsymbol{n}\in\mathbb{Z}^D}$ on G/Γ *equidistributes* if for all continuous $F:G/\Gamma\to\mathbb{C}$ we have

$$\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{n}\in[N]^D}F(g(\boldsymbol{n})\Gamma)\to\int_{G/\Gamma}F.$$

Leibman's theorem says that a polynomial sequence equidistributes if and only if it does not lie in (a coset of) the kernel of a nontrivial horizontal character on G.

THEOREM A.5 (Leibman's criterion for qualitative equidistribution [Lei05]). *Let D be a positive integer, let G/Γ be a nilmanifold and let $g:\mathbb{Z}^D\to G$ be a polynomial sequence. Then exactly one of the following is true:*

(i) *$(g(\boldsymbol{n})\Gamma)_{\boldsymbol{n}\in\mathbb{Z}^D}$ equidistributes in G/Γ; or*
(ii) *there exists a nontrivial horizontal character $\eta:G\to\mathbb{R}$ which maps Γ to $\mathbb{Z}$ and such that $\eta\circ g$ is constant.*

Given two lattices Γ_0 and Γ_1 in G, we say that they are *commensurable* if $\Gamma_0\cap\Gamma_1$ is a lattice in G of finite index in both Γ_0 and Γ_1. It is not difficult to check directly from the above definition that if Γ_0 and Γ_1 are commensurable, then a polynomial sequence g equidistributes on G/Γ_0 if and only if it equidistributes on G/Γ_1. Therefore, we may extend the definition of equidistribution on a particular nilmanifold G/Γ to consider simultaneously all lattices commensurable to Γ. This may be neatly understood in the Lie algebra $\mathfrak{g}$ by the following result.

THEOREM A.6 [CG90, Theorem 5.1.12]. *Let Γ_0,Γ_1 be lattices in G. Then Γ_0 and Γ_1 determine the same rational structure on $\mathfrak{g}$ (i.e. $\operatorname{span}_{\mathbb{Q}}(\log\Gamma_0)=\operatorname{span}_{\mathbb{Q}}(\log\Gamma_1)$) if and only if Γ_0 and Γ_1 are commensurable.*

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

We remark also that given any rational structure $\mathfrak{g}_{\mathbb{Q}}$ on $\mathfrak{g}$, there is a lattice Γ in $G = \exp \mathfrak{g}$ (invoke Lie's third theorem) such that $\mathfrak{g}_{\mathbb{Q}} = \operatorname{span}_{\mathbb{Q}}(\log \Gamma)$ (see [CG90, Theorem 5.1.8]). Finally, any polynomial sequence G yields (under the logarithm map) a polynomial sequence in $\mathfrak{g}$ and, conversely. By the previous discussion, we may define the property of equidistribution of a polynomial sequence in $\mathfrak{g}$ with respect to a rational structure $\mathfrak{g}_{\mathbb{Q}}$ as that of the corresponding polynomial sequence in G with respect to any lattice Γ such that $\operatorname{span}_{\mathbb{Q}}(\log \Gamma) = \mathfrak{g}_{\mathbb{Q}}$.

Definition A.7 (Equidistribution with respect to a rational structure). A polynomial sequence p in $\mathfrak{g}$ *equidistributes in* $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$ if $\exp p$ equidistributes in G/Γ, where Γ is any lattice in G such that $\operatorname{span}_{\mathbb{Q}} \log \Gamma = \mathfrak{g}_{\mathbb{Q}}$.

By our previous discussion, this definition is indeed well-defined. We are ready to prove the following. The proof is very straightforward, but we include it to reiterate what is meant by equidistribution in the Lie algebra. Recall that in the following (and, indeed, in all previous discussions of irrationality) we have restricted to polynomial sequences with $p(0) = 0$.

Theorem A.8 (Leibman's criterion for qualitative equidistribution in the Lie algebra, Theorem 2.8). *Let D be a positive integer, let $\mathfrak{g}$ be a real, finite-dimensional, nilpotent Lie algebra and let $\mathfrak{g}_{\mathbb{Q}}$ be a rational structure on $\mathfrak{g}$. Let $(p(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D}$ be a polynomial sequence on $\mathfrak{g}$. Then $(p(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D}$ equidistributes in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$ if and only if p is additively irrational in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$.*

Proof. Suppose p is not additively irrational in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$, so by Definition 2.6, there exists a nontrivial $\eta \in \operatorname{Hom}_{\mathbb{Q}}(\mathfrak{g}, \mathbb{R})$ such that $((\eta \circ p)(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D} \subset \mathbb{Q}$. Let Γ be a lattice in $G = \exp \mathfrak{g}$ such that $\operatorname{span}_{\mathbb{Q}} \log \Gamma = \mathfrak{g}_{\mathbb{Q}}$ (cf. [CG90, Theorem 5.1.8] for its existence). Then $\eta(\log \Gamma) \subset \mathbb{Q}$. In fact, since $\log \Gamma$ is contained in some additive lattice itself contained in $\mathfrak{g}_{\mathbb{Q}}$ (see [CG90, Theorem 5.4.2]), there is a positive integer M such that $\eta(\log \Gamma) \subset (1/M)\mathbb{Z}$. Thus, there exists η', a large integer multiple of η, so that $\eta'(\log \Gamma) \subset \mathbb{Z}$ and $((\eta' \circ p)(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D} \subset \mathbb{Z}$. By standard Lie theory (see [Hal03, Theorem 3.7]) there is a unique Lie group homomorphism $\eta_0 : G \to \mathbb{R}/\mathbb{Z}$ such that $\eta' = d\eta_0$ and so $\eta_0 \circ \exp = \exp \circ \eta'$ (where here the second $\exp : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ is just the quotient map). Writing this out we have

$$\bar{0} = ((\exp \circ \eta' \circ p)(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D} = ((\eta_0 \circ \exp \circ p)(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D},$$

where $\bar{0}$ is the residue class of 0 in $\mathbb{R}/\mathbb{Z}$. Note η_0 is not trivial. Thus, the polynomial sequence $\exp \circ p$ does not equidistribute in G/Γ by Theorem A.5, and so p does not equidistribute in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$ by definition.

In the other direction, suppose that $(p(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D}$ does not equidistribute in $(\mathfrak{g}, \mathfrak{g}_{\mathbb{Q}})$, so $\exp \circ p$ does not equidistribute in G/Γ where Γ is a lattice such that $\operatorname{span}_{\mathbb{Q}}(\log \Gamma) = \mathfrak{g}_{\mathbb{Q}}$. By Theorem A.5, there is a nontrivial horizontal character $\eta_0 : G \to \mathbb{R}/\mathbb{Z}$ which maps Γ to $\mathbb{Z}$ and such that $\eta_0 \circ \exp \circ p$ is constant in $\mathbb{R}/\mathbb{Z}$. Standard Lie theory yields that the differential $d\eta_0$ is a nontrivial Lie algebra homomorphism $\mathfrak{g} \to \mathbb{R}$ and $\eta_0 \circ \exp = \exp \circ d\eta_0$. Set $\eta = d\eta_0$. Thus, $\exp \circ \eta \circ p$ is constant (in fact, zero) in $\mathbb{R}/\mathbb{Z}$, so $\eta \circ p \subset \mathbb{Z}$, and p is not additively irrational. □

Now we proceed to the matter of quantitative equidistribution. In the main text, we used the notation $\langle \mathcal{X} \rangle$ to refer to the additive group generated by a basis $\mathcal{X}$, that is, $\operatorname{span}_{\mathbb{Z}} \mathcal{X}$. In this appendix we also have to deal with $\langle \mathcal{X} \rangle_*$, the $*$-multiplicative group generated by $\mathcal{X}$, and so we generally use $\operatorname{span}_{\mathbb{Z}} \mathcal{X}$ to refer to the additive group in order to make the distinction more notationally clear.

In the following definition, we refer to a Lipschitz function on G/Γ, which obviously presupposes the existence of a metric on this space. We refer the reader to [GT12] for details on this; they are not important for the purposes of this document.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

DEFINITION A.9 (Quantitative nilmanifold equidistribution). Let $0<\delta<1/2$, let G/Γ be a nilmanifold, let $N_1,\dots,N_D\geq 1$ and let $[\boldsymbol{N}]$ denote the set $[N_1]\times\cdots\times[N_D]$. We say that a polynomial sequence $(g(\boldsymbol{n})\Gamma)_{\boldsymbol{n}\in\mathbb{Z}^D}$ is *δ-equidistributed on* $[\boldsymbol{N}]$ in G/Γ if for all Lipschitz functions $F:G/\Gamma\to\mathbb{C}$ we have

$$|\mathbb{E}_{\boldsymbol{n}\in[\boldsymbol{N}]}F(g(\boldsymbol{n})\Gamma)-\int_{G/\Gamma}F|\leq\delta\|F\|_{\mathrm{Lip}}.$$

If $N_1=\cdots=N_D=:N$, then we may also say that $(g(\boldsymbol{n})\Gamma)_{\boldsymbol{n}\in\mathbb{Z}^D}$ is *δ-equidistributed at scale N*.

Green and Tao proved the following; [GT12] may be consulted for terminology/notation which has not otherwise been introduced in this document.

THEOREM A.10 [GT15, Theorem, p. 3]. *Let $0<\delta<1/2$, let $N_1,\dots,N_D\geq 1$ and let $[\boldsymbol{N}]$ denote the set $[N_1]\times\cdots\times[N_D]$. Let G/Γ be a nilmanifold with dimension d, step s and Mal'cev basis $\mathcal{X}$ which is $1/\delta$-rational. Let $(g(\boldsymbol{n}))_{\boldsymbol{n}\in\mathbb{Z}^D}$ be a polynomial sequence in G. Then one of the following holds:*

(i) *there is some $N_i\ll\delta^{-O_{d,s,D}(1)}$;*

(ii) *$(g(\boldsymbol{n})\Gamma)_{\boldsymbol{n}\in[\boldsymbol{N}]}$ is δ-equidistributed on $[\boldsymbol{N}]$ in G/Γ; or*

(iii) *there exists a nontrivial horizontal character η on G/Γ with $0<\|\eta\|\ll\delta^{-O_{d,s,D}(1)}$ such that $\|\eta\circ g\|_{C^\infty[\boldsymbol{N}]}\ll\delta^{-O_{d,s,D}(1)}$.*

Furthermore, when $N_1=\cdots=N_D$, the first option above may be removed.

Our first task in the quantitative setting is to define equidistribution of a polynomial sequence in the Lie algebra with respect to a choice of (rational) basis.

DEFINITION A.11 (Quantitative equidistribution in the Lie algebra, Definition 4.14). Let $\mathcal{X}$ be a rational basis in $\mathfrak{g}$. Then a polynomial sequence $(p(\boldsymbol{n}))_{\boldsymbol{n}\in[\boldsymbol{N}]}$ is δ-equidistributed on $[\boldsymbol{N}]$ in $(\mathfrak{g},\mathcal{X})$ if $(\exp(p(\boldsymbol{n})))_{\boldsymbol{n}\in[\boldsymbol{N}]}$ is δ-equidistributed on $[\boldsymbol{N}]$ in G/Γ, where Γ is the smallest multiplicative lattice containing $\mathrm{span}_{\mathbb{Z}}\mathcal{X}$.

Let us check that this definition is sensible. First, the intersection of two multiplicative lattices that contain $\mathcal{X}$ is clearly a discrete multiplicative group which spans $\mathfrak{g}$; this is enough to conclude [CG90, Corollary 5.4.5] that it is cocompact in G and so indeed itself a multiplicative lattice; thus, we may indeed speak of the smallest lattice containing $\mathrm{span}_{\mathbb{Z}}\mathcal{X}$.

Next, in the context of constructing nilsequences, a function on $\mathfrak{g}$ is of course multiplicatively automorphic with respect to the set $\mathrm{span}_{\mathbb{Z}}\mathcal{X}$ if and only if it is multiplicatively automorphic with respect to $\langle\mathrm{span}_{\mathbb{Z}}\mathcal{X}\rangle_*$.

The following lemma offers some further validation for the above definition.

LEMMA A.12. *Let Λ be an additive lattice in $\mathfrak{g}$. Then the smallest $*$-multiplicative subgroup $\langle\Lambda\rangle_*$ containing Λ is a multiplicative lattice in $\mathfrak{g}$. Furthermore, if $\mathcal{X}$ is a $\mathbb{Z}$-basis for Λ which has rational structure constants of height at most M, then we have*

$$\Lambda\subset\langle\Lambda\rangle_*\subset\frac{1}{M^{O(1)}}\Lambda,$$

where implicit constants may depend on $\dim\mathfrak{g}$.

Proof. First, it is not difficult to construct a basis $\tilde{\mathcal{X}}:=\{\tilde{X}_1,\dots,\tilde{X}_d\}$ for $\mathfrak{g}$ consisting of rational linear combinations of $\mathcal{X}$ with coefficients of height $M^{O(1)}$ which satisfies the nesting property

$$[\mathfrak{g},\tilde{X}_i]\subset\mathrm{span}_{\mathbb{R}}\{\tilde{X}_{i+1},\dots,\tilde{X}_d\};$$

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

see [GT12, Proposition A.9] for details. Thus, we may assume without loss of generality that $\mathcal{X}$ satisfies this nesting property; we use the notation $\mathcal{X} := \{X_1, \dots, X_d\}$. We claim that there are integers $k_1, \dots, k_d$ each of size $M^{O(1)}$ such that upon setting $\mathcal{X}' := \{(1/k_1)X_1, \dots, (1/k_d)X_d\}$ we have that $\operatorname{span}_{\mathbb{Z}} \mathcal{X}'$ is a multiplicative lattice in $\mathfrak{g}$. Then

$$\Lambda \subset \langle \Lambda \rangle_* \subset \operatorname{span}_{\mathbb{Z}} \mathcal{X}' \subset \frac{1}{M^{O(1)}} \Lambda,$$

and so we are done.

It follows from the Baker–Campbell–Hausdorff formula that multiplication in coordinates with respect to $\mathcal{X}$ is a polynomial map $* : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ of degree $O(1)$ and with rational coefficients of height $M^{O(1)}$. Let P_i be polynomials such that $a * b = (P_1(a, b), \dots, P_d(a, b))$, with respect to $\mathcal{X}$. Then by the Baker–Campbell–Hausdorff formula and the nesting property of $\mathcal{X}$, we have that

$$P_i(a, b) = a_i + b_i + \tilde{P}_i(a_1, \dots, a_{i-1}, b_1, \dots, b_{i-1}),$$

where $\tilde{P}_i$ is a polynomial with all terms of degree at least 2.

Let $k_1, \dots, k_d$ be positive integers and form $\mathcal{X}'$ as above. Then we have that

$$P'_i(a, b) = k_i P_i(a_1/k_1, \dots, a_d/k_d, b_1/k_1, \dots, b_d/k_d),$$

where the P'_i are the polynomials which define multiplication in $\mathfrak{g}$ with respect to $\mathcal{X}'$. In particular, defining $\tilde{P}'_i$ analogously, we have

$$\tilde{P}'_i(a, b) = k_i \tilde{P}_i(a_1/k_1, \dots, a_{i-1}/k_{i-1}, b_1/k_1, \dots, b_{i-1}/k_{i-1}).$$

It follows that we may inductively choose k_i of size $M^{O(1)}$ depending on P_i and $(k_1, \dots, k_{i-1})$ so that $\tilde{P}'_i$ has integer coefficients and thus so too does P'_i. Choosing the k_i in this way we have $\operatorname{span}_{\mathbb{Z}} \mathcal{X}' * \operatorname{span}_{\mathbb{Z}} \mathcal{X}' \subset \operatorname{span}_{\mathbb{Z}} \mathcal{X}'$. In addition, $\operatorname{span}_{\mathbb{Z}} \mathcal{X}'$ is clearly closed under inverses, and so is a multiplicative subgroup. Furthermore, $\operatorname{span}_{\mathbb{Z}} \mathcal{X}'$ is clearly discrete and its cocompactness in $\mathfrak{g}$ follows from [CG90, Theorem 5.1.6]. This completes the proof. □

Lemma A.13. *Let Λ be an additive lattice in $\mathfrak{g}$ with $\mathbb{Z}$-basis $\mathcal{X}$, which has rational structure constants of height at most M. Then the multiplicative lattice $\langle \Lambda \rangle_*$ has a Mal'cev basis whose elements are rational linear combinations of $\mathcal{X}$ with coefficients of height $M^{O(1)}$.*

Proof. The existence of a Mal'cev basis for $\langle \Lambda \rangle_*$ follows from Lemma A.12 and [CG90, Theorem 5.1.6]. It is not difficult to use (the proofs of) these results to recover the quantitative statement of this lemma. We omit the details. □

Deducing the Lie algebra version of Green and Tao's criterion for quantitative equidistribution from its Lie group predecessor is similar to what we have already done for Leibman's criterion in the qualitative setting. We only sketch the proof in the following.

Theorem A.14 (Green and Tao's criterion for quantitative equidistribution in the Lie algebra, Theorem 4.15). *Let $0 < \delta < 1/2$, let $N_1, \dots, N_D \geq 1$ and let $[\boldsymbol{N}]$ denote the set $[N_1] \times \cdots \times [N_D]$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra with dimension d, step s and basis $\mathcal{X}$ which is $1/\delta$-rational. Let $(p(\boldsymbol{n}))_{\boldsymbol{n} \in \mathbb{Z}^D}$ be a polynomial sequence in $\mathfrak{g}$. There are positive constants $c_{d,s,D}, c'_{d,s,D}$ such that at least one of the following is true:*

(i) *there is some $N_i < \delta^{-c'_{d,s,D}}$;*
(ii) *$(p(\boldsymbol{n}))_{\boldsymbol{n} \in [\boldsymbol{N}]}$ is δ-equidistributed on $[\boldsymbol{N}]$ in $(\mathfrak{g}, \mathcal{X})$;*
(iii) *p is not $(\delta^{-c_{d,s,D}}, \boldsymbol{N})$-additively irrational in $(\mathfrak{g}, \mathcal{X})$.*

Furthermore, when $N_1 = \cdots = N_D$, the first option above may be removed.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Proof. Let all constants depend on s, d, D. Suppose that parts (i) and (ii) do not hold and denote $G = \exp \mathfrak{g}$ (as before, we are implicitly invoking Lie's third theorem) and $\Gamma = \exp(\langle \mathrm{span}_{\mathbb{Z}} \mathcal{X} \rangle_*)$ so that, by definition, $\exp \circ p$ is not δ-equidistributed on $[\boldsymbol{N}]$ in G/Γ. Let $\mathcal{X}_{\mathrm{Mal}}$ be the Mal'cev basis for Γ which is produced by Lemma A.13; it is $\delta^{-O(1)}$-rational. By Theorem A.10, there is a horizontal character η_0 on G/Γ with $0 < \|\eta_0\| < \delta^{O(1)}$ such that $\|\eta_0 \circ g\|_{C^\infty[\boldsymbol{N}]} < \delta^{-O(1)}$. Then one may argue using basic Lie theory as in the qualitative setting to find from η_0 some nontrivial $\eta \in \mathrm{Hom}_{\mathcal{X}_{\mathrm{Mal}}}(\mathfrak{g}, \mathbb{R})$ with complexity (as in our Definition 4.5) $\delta^{-O(1)}$ such that $\|\eta \circ p\|_{C^\infty[\boldsymbol{N}]} < \delta^{-O(1)}$. But the elements of $\mathcal{X}_{\mathrm{Mal}}$ are rational linear combinations of $\mathcal{X}$ of height $\delta^{-O(1)}$, so there is some $C = \delta^{-O(1)}$ such that $\eta' := C\eta$ is a nontrivial element of[18] $\mathrm{Hom}_{\mathcal{X}}(\mathfrak{g}, \mathbb{R})$ of complexity $\delta^{-O(1)}$. Thus, we have

$$\|\eta' \circ p\|_{C^\infty[\boldsymbol{N}]} \leq C\|\eta \circ p\|_{C^\infty[\boldsymbol{N}]} < \delta^{-O(1)},$$

and so p is not $(\delta^{-O(1)}, \boldsymbol{N})$-additively irrational in $(\mathfrak{g}, \mathcal{X})$. □

A.3 Rational and small polynomial sequences

LEMMA A.15 (Smallness and rationality under products). *Let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$ such that the Lie bracket has rational structure constants of height at most M with respect to $\mathcal{X}$. Let e, e' be polynomial sequences adapted to $\mathfrak{g}_\bullet$ which are (A, N)-small and let r, r' be polynomial sequences adapted to $\mathfrak{g}_\bullet$ which are A-rational. Then $e * e'$ and $r * r'$ are polynomial sequences adapted to $\mathfrak{g}_\bullet$, the former is $((AM)^{O(1)}, N)$-small and the latter is $(AM)^{O(1)}$-rational.*

Proof. That $e * e'$ and $r * r'$ are polynomial sequences adapted to $\mathfrak{g}_\bullet$ is standard. The rest follows from definitions and some easy, but disproportionately tedious, arguments. We omit the details. □

LEMMA A.16. *Suppose that r is a polynomial sequence in $\mathfrak{g}$ which is A-rational with respect to $\mathcal{X}$. Suppose that $\mathfrak{g}$ has rational structure constants of height at most M. Then for any $*$-multiplicative subgroup Γ of $\mathfrak{g}$ which contains $\langle \mathcal{X} \rangle$, the orbit $r(n) * \Gamma$ is $(AM)^{O(1)}$-periodic.*

Proof. We show that there is an integer q of size $(AM)^{O(1)}$ such that $p(n)^{-1} * p(n+q) \in \mathrm{span}_{\mathbb{Z}} \mathcal{X}$ for all integers n. For the meantime let q be undetermined. Note $p(n)^{-1} = -p(n)$ and let p' be the polynomial sequence in 2 variables such that $qp'(n, q) = p(n+q) - p(n)$. By the Baker–Campbell–Hausdorff formula we have

$$p(n)^{-1} * p(n+q) = p(n+q) - p(n) - \tfrac{1}{2}[p(n), p(n+q)] + \cdots .$$

We also have from definitions that $[p(n), p(n+q)] = q[p(n), p'(n, q)]$, and indeed every term in the Baker–Campbell–Hausdorff expansion will be a polynomial in n and q which is divisible by q and which has coefficients in $((1/(AM)^{O(1)})\mathbb{Z})^{\deg \mathfrak{g}}$, so that ultimately the polynomial $p(n)^{-1} * p(n+q)$ has this property. Therefore, a specific value of q may be chosen of size at most $(AM)^{O(1)}$ so that, for all $n \in \mathbb{Z}$, $p(n)^{-1} * p(n+q) \in \mathrm{span}_{\mathbb{Z}} \mathcal{X}$. □

A.4 Factorising for filtration irrationality

In this section, we sketch how to deduce Proposition 6.2 from [GT10, Lemma 2.10]. It is mostly a matter of transferring between related notions of complexity whose details we largely omit. The main additional ingredient is the following.

[18] Recall that $\mathrm{Hom}_{\mathcal{X}}(\mathfrak{g}, \mathbb{R})$ denotes the set of Lie algebra homomorphisms from $\mathfrak{g}$ to $\mathbb{R}$ which map $\mathcal{X}$ to $\mathbb{Z}$.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Lemma A.17. *Let $M > 1$ be a positive integer. Let $\mathfrak{g}$ have rational basis $\mathcal{X}$ which has rational structure constants of height at most M. Then there exists a multiplicative lattice $\Gamma \leq G$ such that $M^{O(1)} \operatorname{span}_{\mathbb{Z}} \mathcal{X} \subseteq \log \Gamma \subseteq (1/M^{O(1)}) \operatorname{span}_{\mathbb{Z}} \mathcal{X}$.*[19]

Proof. First, we argue that there is a strong Mal'cev basis[20] $\mathcal{X}'$ such that

$$M^{O(1)} \operatorname{span}_{\mathbb{Z}} \mathcal{X}' \subseteq \operatorname{span}_{\mathbb{Z}} \mathcal{X} \subseteq \frac{1}{M^{O(1)}} \operatorname{span}_{\mathbb{Z}} \mathcal{X}'. \tag{A1}$$

This is equivalent to finding a strong Mal'cev basis whose elements may be written as rational linear combinations of $\mathcal{X}$ with coefficients of height at most $M^{O(1)}$. Once one makes the observation that the elements of the descending central series have complexity $M^{O(1)}$ in $\mathfrak{g}$ with respect to $\mathcal{X}$ as a consequence on the bound on the height of the rational structure constants, the rest of the claim is established in the proof of [GT12, Proposition A.9].

Next, by working through the details of [CG90, Theorem 5.1.8(b)] (in particular, making note of quantitative bounds), one obtains that there is an integer $K = M^{O(1)}$ such the set $\mathbb{Z}^d \subset \mathbb{R}^d$ in *Mal'cev coordinates*[21] with respect to $K\mathcal{X}'$ is a multiplicative lattice (and so, in fact, $K\mathcal{X}'$ is a Mal'cev basis for this lattice). Let Γ be this lattice. Then [GT12, Lemma A.2] yields that

$$M^{O(1)} \operatorname{span}_{\mathbb{Z}} K\mathcal{X}' \subseteq \log \Gamma \subseteq \frac{1}{M^{O(1)}} \operatorname{span}_{\mathbb{Z}} K\mathcal{X}'.$$

Putting this together with (A1) proves the lemma. □

Next we restate [GT10, Lemma 2.10] from which we derive Proposition 6.2. The statement is copied verbatim (including notation) from [GT10] and is included here for the convenience of the reader.

Lemma A.18 [GT10, Lemma 2.10]. *Let $N > 1$. Let $(G/\Gamma, G_\bullet)$ be a degree $\leq s$ filtered nilmanifold of complexity $\leq M_0$ and let $g \in \operatorname{poly}(\mathbb{Z}, G_\bullet)$. For any growth function $\mathcal{F}'$, we can find a quantity $M_0 \leq M \leq O_{M,\mathcal{F}'}(1)$ and a factorisation $g = \beta g' \gamma$ where:*

- *$\beta \in \operatorname{poly}(\mathbb{Z}, G_\bullet)$ is $(O_M(1), N)$-smooth;*
- *$g' \in \operatorname{poly}(\mathbb{Z}, G_\bullet)$ is $(\mathcal{F}'(M), N)$-irrational in a subnilmanifold $(G'/\Gamma', G'_\bullet)$ of $(G/\Gamma, G_\bullet)$ of complexity $O_M(1)$; and*
- *$\gamma \in \operatorname{poly}(\mathbb{Z}, G_\bullet)$ is $O_M(1)$-periodic.*

We are ready to prove Proposition 6.2.

Proposition A.19 (Proposition 6.2). *Let $N, M_0 > 1$. Let $\mathfrak{g}$ be a real nilpotent Lie algebra of dimension at most M_0 with filtration $\mathfrak{g}_\bullet$ of degree at most M_0. Let $\mathcal{X}$ be a rational basis for $\mathfrak{g}$ with structure constants of height at most M_0. Let p be a polynomial sequence adapted to $\mathfrak{g}_\bullet$ with $p(0) = 0$. For any growth function $\mathcal{F} : \mathbb{R}^+ \to \mathbb{R}^+$, there exists $M_0 \leq M \leq O_{M_0,\mathcal{F}}(1)$ and a factorisation $p = e * p' * r$ where e is $(O_M(1), N)$-small, where r is $O_M(1)$-rational and where p' is $(\mathcal{F}(M), N)$-filtration irrational in a subfiltration $\mathfrak{g}'_\bullet \leq \mathfrak{g}_\bullet$ of complexity $O_M(1)$.*[22] *This filtration irrationality is measured with respect to $\mathcal{X}'$, a Mal'cev basis*[23] *for $\mathfrak{g}'_\bullet$ which may be written as a linear combination of $\mathcal{X}$ with rational coefficients of height $O_M(1)$.*

[19] In the language of [GT12], $\mathcal{X}$ is then a *$M^{O(1)}$-rational weak basis* for G/Γ.

[20] Here our usage of 'strong Mal'cev basis' is as in [CG90]: a basis $\{X_1, \ldots, X_d\}$ for $\mathfrak{g}$ such that $\operatorname{span}_{\mathbb{R}}\{X_i, \ldots, X_d\}$ is an ideal for all i.

[21] Our usage of 'Mal'cev coordinates' is as in [GT12]. These coordinates are also known in the literature as *coordinates of the second kind*.

[22] That is, for all i, $\mathfrak{g}'_i$ is a subspace of $\mathfrak{g}_i$ of complexity $O_M(1)$.

[23] See [GT12, Definition 2.1].

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Proof. We sketch the argument as a detailed proof is both straightforward and tedious. We refer the reader to [GT12, Appendix A] for more details on Mal'cev bases and quantitative aspects thereof. We use terminology from there without further introduction.

Let Γ be the multiplicative lattice produced by Lemma A.17 with input $\mathcal{X}$ and M_0. Then G/Γ will have complexity $M_0^{O(1)}$ as a nilmanifold (cf. [GT10, Definition 1.4]).[24] Invoke Lemma A.18 for the polynomial sequence $g = \exp \circ p$ and with lattice Γ. One may go inside the proof of [GT10, Lemma 2.10] and also use [GT12, Lemma A.2] to show that the $(O_M(1), N)$-smooth (respectively, $O_M(1)$-periodic) sequence produced by [GT10, Lemma 2.10] may instead by taken to be $(O_M(1), N)$-small (respectively, $O_M(1)$-rational) by our definition.[25] Inferring the complexity of $\mathfrak{g}'_\bullet := (\log G'_i)_{i=1}^s$ from that of $G'_\bullet$ is an easy exercise in expanding definitions. Furthermore, the $(\mathcal{F}(M), N)$-filtration irrationality of p' with respect to an appropriate basis in $\mathfrak{g}'$ may be inferred from [GT12, Proposition A.10] and the irrationality of g' in $G'/(G' \cap \Gamma)$. □

A.5 On the Lipschitz property

In this appendix we sketch a proof of the claim made in the proof of Proposition 6.7: that $\tilde{f}$ is $O_M(1)$-Lipschitz. In the rest of this appendix, we follow [GT12] and use ψ to denote the Mal'cev coordinate map. However, we abuse notation and use ψ to be both a map on the Lie group G and the Lie algebra $\mathfrak{g}$ (in both cases to $\mathbb{R}^{\dim \mathfrak{g}}$); usage will be clear based on context.

Recall that we define $\tilde{f} : \mathfrak{g} \times \mathbb{Z}/q\mathbb{Z} \times [0,1] \to [0,1]$ by

$$\tilde{f}(x, s, y) := f(\{p(0)\} * e(Ny) * x * r(s) * \log \Gamma),$$

with notation and assumptions inherited from the proof of Proposition 6.7. Recall, in particular, that in Proposition 6.7 we preprocessed the polynomial sequence $p(n)$ so that the coordinates of $p(0)$ are of size $O_M(1)$. For brevity, we use the notation $p := p(0)$ for the rest of this appendix.

By the triangle inequality, it suffices to show that $\tilde{f}$ is $O_M(1)$-Lipschitz in each component. At this point, the reader may wish to recall the definite of the metric d on G/Γ: see [GT12, Definition 2.2]. To this end we compute (omitting on this occasion the $*$ symbol for multiplication to save space),

$$\begin{aligned} |\tilde{f}(x, s, y) - \tilde{f}(x', s, y)| &= |f(pe(Ny)xr(s)\log\Gamma) - f(pe(Ny)x'r(s)\log\Gamma)| \\ &\le O_M(d_{G/\Gamma}(pe(Ny)xr(s)\log\Gamma, pe(Ny)x'r(s)\log\Gamma)) \\ &= O_M\Big(\inf_{\gamma,\gamma'\in\Gamma} d(pe(Ny)xr(s)\log\gamma, pe(Ny)x'r(s)\log\gamma')\Big), \end{aligned} \tag{A2}$$

where we have used that f is $O_M(1)$-Lipschitz in the second line.

Recall from the proof of Proposition 6.7 the notation $l\langle\mathcal{X}\rangle$. Now define $\log\Gamma'$ to be the multiplicative closure of $l\langle\mathcal{X}\rangle$; we have shown previously (Lemmas A.12 and A.13) that $\log\Gamma'$ is a mutliplicative lattice and admits a Mal'cev basis whose elements are height $M^{O(1)}$ rational linear combinations of the elements of $l\mathcal{X}$. Let $\mathcal{X}'$ be this basis. Recall also from the proof of Proposition 6.7 the fact that

$$l\langle\mathcal{X}\rangle \subset r(n) * \langle\mathcal{X}\rangle * r(n)^{-1} \subset r(n) * \log\Gamma * r(n)^{-1},$$

for all $n \in \mathbb{Z}$. Taking $*$-multiplicative closures, we see that $\log\Gamma' * r(n) \subset r(n) * \log\Gamma$ for all n.

[24] This is not entirely trivial as one needs to construct from $\mathcal{X}$ a Mal'cev basis for Γ to exhibit the complexity of G/Γ. This is done in [GT12, Lemma A.8 and Proposition A.9] together with our Lemma A.17 which yields that $\mathcal{X}$ is a $M^{O(1)}$-rational weak basis for Γ as per [GT12, Definition A.7]. Alternatively, one may go inside the proof of Lemma A.17 to see that the basis $K\mathcal{X}'$ is in fact an appropriate Mal'cev basis for Γ.

[25] With respect to the Mal'cev basis used in [GT10, Lemma 2.10], but then the same is true with respect to $\mathcal{X}$ since the two bases are related by rational linear combinations of height $M_0^{O(1)}$.

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

Continuing from (A2), and using that $\log\Gamma' * r(s) \subset r(s) * \log\Gamma$, we have

$$|\tilde{f}(x,s,y) - \tilde{f}(x',s,y)| \leq O_M\Big(\inf_{\gamma,\gamma'\in\Gamma'} d(pe(Ny)x\log\gamma r(s), pe(Ny)x'\log\gamma' r(s))\Big).$$

Furthermore, by [GT12, Lemma A.17], the above is true when the Mal'cev basis underlying the metric d is changed from $\mathcal{X}_{\mathrm{Mal}}$ to $\mathcal{X}'$. We denote this latter metric by d'. This, together with the right-invariance of d', gives

$$|\tilde{f}(x,s,y) - \tilde{f}(x',s,y)| \leq O_M\Big(\inf_{\gamma,\gamma'\in\Gamma'} d'(pe(Ny)x\log\gamma, pe(Ny)x'\log\gamma')\Big).$$

Next, using [GT12, Lemma A.14], we reduce x, x' to lie in a fundamental domain for G/Γ' such that $\{x\}, \{x'\}$ has Mal'cev coordinates in $[0,1)$. In addition, utilising the right-invariance of d' once more to multiply by $(\log\gamma)^{-1}$, and then reparametrising to take the infimum only over one $\gamma\in\Gamma'$, we obtain

$$|\tilde{f}(x,s,y) - \tilde{f}(x',s,y)| \leq O_M\Big(\inf_{\gamma\in\Gamma'} d'(pe(Ny)\{x\}, pe(Ny)\{x'\}\log\gamma)\Big). \tag{A3}$$

Removing the factors p and $e(Ny)$ from the left of x, x' in each component is a little trickier since the metric d' is only *approximately* left invariant (see [GT12, Lemma A.5]). In particular, we need the fact that the coordinates of p and $e(Ny)$ are bounded for all $y\in[0,1]$ and, furthermore, that the infimum over γ is obtained by some γ of bounded coordinate size.

Lemma A.20. *Let ψ denote the Mal'cev coordinate map: $G\to\mathbb{R}^{\dim\mathfrak{g}}$ with respect to $\mathcal{X}'$. We have that $\|\psi(p)\|_\infty = O_M(1)$ and that $\|\psi(e(Ny))\|_\infty = O_M(1)$ for all $y\in[0,1]$.*

Proof. The first bound follows from the fact that p was constructed to lie in a bounded fundamental domain for G/Γ, and the fact that $\mathcal{X}'$ consists of elements which are height $O_M(1)$ rational linear combinations of the elements of the Mal'cev basis for G/Γ.

Next, we recall that $e(\cdot)$ is $(O_M(1), N)$-small with respect to $\mathcal{X}$, so

$$\|e(Ny)\|_\infty = \|\sum_{i=1}^{s} c_i(Ny)^i\|_\infty \leq \sum_{i=1}^{s}\|c_i\|_\infty(Ny)^i \leq \sum_{i=1}^{s} O_M\Big(\frac{(Ny)^i}{N^i}\Big) = O_M(1),$$

for all $y\in[0,1]$. We may then deduce the corresponding bound for $\|\psi(e(Ny))\|_\infty$ by invoking [GT12, Lemma A.2]. □

Lemma A.21. *Let $Q\geq 1$ and let $y, z\in G$ satisfy $\|\psi(z)\|_\infty, \|\psi(y)\|_\infty \leq Q$. Let K be an integer such that $\mathcal{X}'$ is K-rational. If Q and $\gamma_0\in\Gamma'$ satisfy $d'(y, z\gamma_0)\leq \max(K, 2\inf_{\gamma\in\Gamma'} d'(y, z\gamma)) \leq Q$, then $\|\psi(\gamma_0)\|_\infty \ll Q^{O(1)}$.*

In particular, under these assumptions, there exists $\gamma_1\in\Gamma'$ which satisfies both $\inf_{\gamma\in\Gamma'} d'(y, z\gamma) = d'(y, z\gamma_1)$, and $\|\psi(\gamma_1)\|_\infty \ll Q^{O(1)}$.

Proof. By the right-invariance of d', we have that $d'(y, z\gamma) = d'(\mathrm{id}_G, z\gamma y^{-1})$, and so $d'(\mathrm{id}_G, z\gamma_0 y^{-1}) \leq Q$. Furthermore, we have Q-rationality with respect to $\mathcal{X}'$. Thus, by [GT12, Lemma A.4 (ii)], we have $\|\psi(z\gamma_0 y^{-1})\|_\infty \ll Q^{O(1)}$. Writing $\psi(\gamma_0) = \psi(z^{-1}(z\gamma_0 y^{-1})y)$, we see that $\|\psi(\gamma_0)\|_\infty \ll Q^{O(1)}$ by invoking [GT12, Lemma A.3].

For the 'in particular' statement, we observe that there are only finitely many $\gamma_0\in\Gamma'$ with $\|\psi(\gamma_0)\|_\infty \ll Q^{O(1)}$ (because the elements of $\psi(\gamma_0)$ are integers), and so the infimum comes from a minimum over such elements. Let γ_1 attain this minimum. □

Returning to (A3), we have

$$|\tilde{f}(x,s,y) - \tilde{f}(x',s,y)| \leq O_M\Big(\inf_{\gamma\in\Gamma'} d'(pe(Ny)\{x\}, pe(Ny)\{x'\}\log\gamma)\Big).$$

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

From Lemma A.20 and [GT12, Lemma A.3] we know that both $\|\psi(pe(Ny))\|_\infty$ and $\|\psi(pe(Ny))\|_\infty$ are of size $O_M(1)$. Furthermore, [GT12, Lemma A.16] says that nilmanifolds are bounded and, in particular, $\inf_{\gamma\in\Gamma'} d'(\{x\},\{x'\}\gamma) = O_M(1)$. Let γ_1 be an element obtained from Lemma A.21 with $\|\psi(\gamma_1)\|_\infty = O_M(1)$ and

$$\inf_{\gamma\in\Gamma'} d'(\{x\},\{x'\}\log\gamma) = d'(\{x\},\{x'\}\log\gamma_1).$$

We are now finally in a position to use the approximate left-invariance of d'. In particular, invoking this approximate left-invariance [GT12, Lemma A.5] in the second line,

$$\begin{aligned}\inf_{\gamma\in\Gamma'} d'(pe(Ny)\{x\}, pe(Ny)\{x'\}\log\gamma) &\le d'(pe(Ny)\{x\}, pe(Ny)\{x'\}\log\gamma_1)\\ &\le O_M(d'(\{x\},\{x'\}\log\gamma_1))\\ &= O_M\Big(\inf_{\gamma\in\Gamma'} d'(\{x\},\{x'\}\log\gamma)\Big)\\ &= O_M(d'_{G/\Gamma'}(x,x')).\end{aligned}$$

This completes the proof that $\tilde f$ is $O_M(1)$-Lipschitz in its first coordinate. The $O_M(1)$-Lipschitz property in the second coordinate is trivial: one combines the observation that $q = O_M(1)$, so for $s, s' \in \mathbb{Z}/q\mathbb{Z}$, $d(s,s') \gg_M 1$, with the facts that f is $O_M(1)$-Lipschitz and that the nilmanifold G/Γ is $O_M(1)$-bounded [GT12, Lemma A.16]. Finally, the $O_M(1)$-Lipschitz property in the third coordinate follows from very similar arguments to that of the first coordinate given previously.

Appendix B. Some quantitative integer linear algebra

In this appendix we define what a Hermite basis is and collect some (very standard) results that one might use to prove Lemmas 4.10 and 4.9, which are essentially all of the form 'integer linear algebra engenders at most polynomial growth in the size of the integers with which we are computing'.

B.1 Hermite normal form

Given a vector space $\mathfrak{g}$ with basis $\mathcal{X}$ and a rational subspace S, we want to define a basis $\mathcal{X}'$ for S such that $\mathrm{span}_{\mathbb{Z}}\mathcal{X}' = (\mathrm{span}_{\mathbb{Z}}\mathcal{X})\cap S$. To do so, we use the theory of the Hermite normal form. Everything in this subappendix is completely standard and is included here for the reader's convenience. For a reference on the Hermite normal form, a reader may consult, for example, [Coh93, § 2.4].

Let m, n be positive integers and let M be an $m\times n$ integer valued matrix.

Definition/Theorem B.1. *The Hermite normal form of M is the unique matrix H of the form $H = MU$ where U is a unimodular matrix and H has the following properties:*

(i) *H is lower triangular;*

(ii) *for any column, the highest nonzero entry (the 'pivot') is strictly below that of all columns to its left;*

(iii) *all pivots are positive;*

(iv) *for all rows with a pivot, all elements to the right of the pivot are zero and all elements to the left of the pivot are non-negative and strictly smaller than the pivot.*

For $\mathfrak{g}, \mathcal{X}, S$ as above, the rationality of S in $\mathfrak{g}$ with respect to the rational structure induced by $\mathcal{X}$ yields that $\mathrm{span}_{\mathbb{Z}}\mathcal{X}\cap S$ is a full rank lattice in S. Let $\mathcal{Z}$ be any $\mathbb{Z}$-basis for this lattice and

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

let $M_{\mathcal{Z}}$ be the $\dim \mathfrak{g} \times \dim S$ integer matrix whose columns are the elements of $\mathcal{Z}$ with respect to $\mathcal{X}$. Let H be the Hermite normal form of $M_{\mathcal{Z}}$. The previous definition/theorem yields that, in fact, H is independent of the choice of basis $\mathcal{Z}$, since any other $M_{\mathcal{Z}'} = M_{\mathcal{Z}}U$ for some unimodular U. Thus, to any $\mathcal{X}, S$ we may associate a unique matrix $H_{\mathcal{X},S}$ in Hermite normal form whose columns are a $\mathbb{Z}$-basis for $\operatorname{span}_{\mathbb{Z}} \mathcal{X} \cap S$.

Definition B.2 (Hermite basis). Let $\mathfrak{g}, \mathcal{X}, S, H_{\mathcal{X},S}$ be as before. We define the *Hermite basis* $\mathcal{X}'_{\mathcal{X},S}$ to be the (ordered) set of columns of the Hermite normal form $H_{\mathcal{X},S}$ with respect to $\mathcal{X}$.

The following lemma is a corollary of any sensible algorithm to compute the Hermite normal form.

Lemma B.3. *Let $m, n = O(1)$. Let M be a $m \times n$ matrix with integer entries of size at most C. Then the Hermite normal form of M has entries of size $C^{O(1)}$.*

As a result, given bounds on the complexity of a subspace, one may bound the size of the entries of its Hermite basis.

The following lemma is essentially all of the linear algebra that is needed to prove Lemmas 4.10 and 4.9. We omit a proof.

Lemma B.4. *Let $m, n = O(1)$. Let M be a $m \times n$ full rank matrix with integer entries of size at most C. Then $\ker M \cap \mathbb{Z}^n$ is a lattice which has a $\mathbb{Z}$-basis consisting of integer vectors whose elements have size at most $C^{O(1)}$.*

Conversely, given a $\mathbb{Z}$-basis for a rank-m integer lattice L in $\mathbb{R}^n$ whose vectors have elements of size at most C, there is an $m \times n$ integer matrix M whose entries have size at most $C^{O(1)}$ such that $\ker M \cap \mathbb{Z}^n = L$.

Acknowledgements

The author is grateful to Ben Green for many helpful conversations. The author is also grateful to Terence Tao for helpful comments on an earlier version of this document, and to an anonymous referee for a very careful reading of this paper and very valuable feedback.

Conflicts of interest

None.

Journal information

Compositio Mathematica is owned by the Foundation Compositio Mathematica and published by the London Mathematical Society in partnership with Cambridge University Press. All surplus income from the publication of *Compositio Mathematica* is returned to mathematics and higher education through the charitable activities of the Foundation, the London Mathematical Society and Cambridge University Press.

References

Alt22 D. Altman, *On a conjecture of Gowers and Wolf*, Discrete Anal. (2022), 10.

Coh93 H. Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics, vol. 138 (Springer, Berlin, 1993).

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press

CG90 L. J. Corwin and F. P. Greenleaf, *Representations of nilpotent Lie groups and their applications. Part I: Basic theory and examples*, Cambridge Studies in Advanced Mathematics, vol. 18 (Cambridge University Press, Cambridge, 1990).

GW10 W. T. Gowers and J. Wolf, *The true complexity of a system of linear equations*, Proc. Lond. Math. Soc. (3) **100** (2010), 155–176.

Gre05 B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), 340–376.

GT10 B. Green and T. Tao, An arithmetic regularity lemma, an associated counting lemma, and applications, in *An irregular mind*, Bolyai Society Mathematical Studies, vol. 21 (János Bolyai Mathematical Society, Budapest, 2010), 261–334.

GT12 B. Green and T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), 465–540.

GT15 B. Green and T. Tao, *On the quantitative distribution of polynomial nilsequences – erratum*, Preprint (2015), arxiv:1311.6170.

GT20 B. Green and T. Tao, *An arithmetic regularity lemma, associated counting lemma, and applications*, Preprint (2020), arXiv:1002.2028v3.

Hal03 B. C. Hall, *Lie groups, Lie algebras, and representations: An elementary introduction*, Graduate Texts in Mathematics (Springer, Cham, 2003).

Kir08 A. Kirillov Jr., *An introduction to Lie groups and Lie algebras*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 2008).

Lei02 A. Leibman, *Polynomial mappings of groups*, Israel J. Math. **129** (2002), 29–60.

Lei05 A. Leibman, *Pointwise convergence of ergodic averages for polynomial actions of $\mathbb{Z}^d$ by translations on a nilmanifold*, Ergodic Theory Dynam. Systems **25** (2005), 215–225.

Man21 F. Manners, *True complexity and iterated Cauchy–Schwarz*, Preprint (2021), arXiv:2109.05731.

Sze75 E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

Tao20 T. Tao, A correction to "An arithmetic regularity lemma, an associated counting lemma, and applications" (2020), https://terrytao.wordpress.com/2020/11/26/a-correction-to-an-arithmetic-regularity-lemma-an-associated-counting-lemma-and-applications/.

Daniel Altman daniel.h.altman@gmail.com
Department of Mathematics, Sloan Mathematical Center, Stanford University, 450 Jane Stanford Way Building 380, Stanford, CA 94305, USA

https://doi.org/10.1112/S0010437X24007693 Published online by Cambridge University Press