J. Inst. Math. Jussieu (2018) **17**(1), 107–135
doi:10.1017/S1474748015000389

ON THE SQUARE-FREE REPRESENTATION FUNCTION OF A NORM FORM AND NILSEQUENCES

LILIAN MATTHIESEN

Institut de Mathématiques de Jussieu – Paris Rive Gauche, UMR 7586, Bâtiment Sophie Germain, Case 7012, 75205 Paris Cedex 13, France
(lilian.matthiesen@imj-prg.fr)

(Received 19 October 2014; revised 6 October 2015; accepted 6 October 2015; first published online 11 November 2015)

Abstract We show that the restriction to square-free numbers of the representation function attached to a norm form does not correlate with nilsequences. By combining this result with previous work of Browning and the author, we obtain an application that is used in recent work of Harpaz and Wittenberg on the fibration method for rational points.

Keywords: norm form equations; weak approximation; higher degree uniformity; nilsequences

2010 *Mathematics subject classification*: 14G05; 11D57; 11B30; 11N37

1. Introduction

Let $K_1, \dots, K_r$ be finite extensions of $\mathbb{Q}$ of degree at least 2, and let $n_1, \dots, n_r \geqslant 2$ denote their respective degrees. For each $1 \leqslant i \leqslant r$, let $\{\omega_1^{(i)}, \dots, \omega_{n_i}^{(i)}\}$ be a $\mathbb{Z}$-basis for the ring of integers of K_i, and denote by

$$\mathbf{N}_{K_i}(x_1, \dots, x_{n_i}) = N_{K_i/\mathbb{Q}}(x_1\omega_1 + \cdots + x_{n_i}\omega_{n_i})$$

the corresponding norm form, where $N_{K_i/\mathbb{Q}}$ is the field norm. One of the central results from [1], stated as [1, Theorem 1.3], proves weak approximation for varieties $X \subset \mathbb{A}_{\mathbb{Q}}^{n_1+\cdots+n_r+s}$ defined by the system of equations

$$0 \neq \mathbf{N}_{K_i}(x_1^{(i)}, \dots, x_{n_i}^{(i)}) = f_i(u_1, \dots, u_s), \quad (1 \leqslant i \leqslant r), \tag{1.1}$$

where $s \geqslant 2$ and $f_1, \dots, f_r \in \mathbb{Z}[u_1, \dots, u_s]$ are pairwise non-proportional linear forms. This weak approximation result is deduced from an asymptotic formula for the number of (suitably restricted) integral points on X (see [1, Theorem 5.2]).

Our aim here is to develop refinements of both [1, Theorem 5.2] and the weak approximation result that allow one to deduce the linear case of a conjecture due to Harpaz and Wittenberg [8] (see [8, § 9] for details). Building on this linear case, they establish the following very strong fibration theorem for the existence of rational points.

While working on this paper the author was supported by a postdoctoral fellowship from the Fondation Sciences Mathématiques de Paris.

Theorem [8, Theorem 9.27]. *Let X be a smooth, proper, irreducible variety over $\mathbb{Q}$, and let $f : X \to \mathbb{P}^1_{\mathbb{Q}}$ be a dominant morphism, with rationally connected geometric generic fibre. Suppose further that all non-split fibres lie over rational points of $\mathbb{P}^1_{\mathbb{Q}}$. If $X_c(\mathbb{Q})$ is dense in $X_c(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}(X_c)}$ for every rational point c of a Hilbert subset of $\mathbb{P}^1_{\mathbb{Q}}$, then $X(\mathbb{Q})$ is dense in $X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}(X)}$.*

The results we discuss here analyse instead of (1.1) the following system of equations:

$$0 \neq \mathbf{N}_{K_i}(x_1^{(i)}, \dots, x_{n_i}^{(i)}) = f_i(u_1, \dots, u_s)\mu(f_i(u_1, \dots, u_s))^2, \quad (1 \leqslant i \leqslant r), \tag{1.2}$$

where μ denotes the Möbius function. Note that counting integral solutions to this system is a question concerning the representation of square-free integers by norm forms.

The weak approximation type result relevant to [8] is the following.

Theorem 1.1. *Let $K_1, \dots, K_r$ be finite extensions of $\mathbb{Q}$ of degree at least 2. Let $f_1, \dots, f_r \in \mathbb{Z}[u_1, \dots, u_s]$ be pairwise non-proportional linear forms. Let S be a finite set of primes that contains all primes $p \leqslant C$ for some constant C only depending on $f_1, \dots, f_r$ and $K_1, \dots, K_r$. Let $\mathbf{u} \in \mathbb{Z}^s$ be a vector such that $f_i(\mathbf{u})$ is non-zero and a local integral norm from K_i at all places of S and also at the real place. Then there exists a vector $\mathbf{u}' \in \mathbb{Z}^s$ such that the following hold.*

1. *$\mathbf{u}'$ is arbitrarily close to $\mathbf{u}$ at the places of S.*
2. *$\mathbf{u}'$ belongs to any given open convex cone of $\mathbb{R}^s$ which contains $\mathbf{u}$.*
3. *$f_i(\mathbf{u}')$ is square free outside S in the sense that $v_p(f_i(\mathbf{u}')) \geqslant 2$ only if $p \in S$, and $f_i(\mathbf{u}')$ is the norm of an integral element of K_i for all i.*

Just as in the case of the weak approximation result from [1], this result is a corollary to an asymptotic formula for the number of (suitably restricted) integral solutions to (1.2), which we state as Theorem 1.3 below. The deduction of Theorem 1.1 from Theorem 1.3 will be carried out in § 2. In order to state Theorem 1.3, we proceed by introducing a square-free representation function for any given norm form $\mathbf{N}_K$ associated to a field extension $K/\mathbb{Q}$ of degree $n \geqslant 2$.

Let $\{\omega_1, \dots, \omega_n\}$ denote the basis with respect to which $\mathbf{N}_K$ is defined. As in [1, § 2], we let $\mathfrak{D}_+ \subset \mathbb{R}^n$ denote a fundamental domain for the equivalence relation that identifies two vectors $\mathbf{x}$ and $\mathbf{y}$ if and only if $x_1\omega_1 + \cdots + x_n\omega_n$ and $y_1\omega_1 + \cdots + y_n\omega_n$ are associated by a unit of positive norm. Define the representation function $R : \mathbb{Z} \to \mathbb{Z}_{\geqslant 0}$ by setting

$$R(m) = \mathbf{1}_{m \neq 0} \cdot \#\{\mathbf{x} \in \mathbb{Z}^n \cap \mathfrak{D}_+ : \mathbf{N}_K(\mathbf{x}) = m\}, \tag{1.3}$$

for any $m \in \mathbb{Z}$. This is a special case of the representation functions considered in [1, Definition 5.1]. Here, we will be interested in the following restrictions of R.

Definition 1.2. Let R be the function defined in (1.3), and let S be a finite set of primes. Then we let $R_S^* : \mathbb{Z} \to \mathbb{Z}_{\geqslant 0}$ denote the restriction of R to integers m that are square free outside S. That is, we define

$$R_S^*(m) = \mu^2\Big(\prod_{p \notin S} p^{v_p(m)}\Big) R(m)$$

for $m \in \mathbb{Z}$.

Remark. The representation function R is in general not multiplicative, not even away from S. For this reason, inclusion–exclusion arguments cannot be carried out in a straightforward way, and therefore the main results of this paper do not follow directly from those in [1].

Finally, we define, as in [1, § 4], a local count of representations by setting

$$\varrho(q, A) = \{\mathbf{x} \in (\mathbb{Z}/q\mathbb{Z})^n : \mathbf{N}_K(\mathbf{x}) \equiv A \pmod{q}\}$$

for any $q \in \mathbb{N}$ and $A \in \mathbb{Z}/q\mathbb{Z}$.

With this notation, the following result is the asymptotic result for R_S^* that corresponds to [1, Theorem 5.2].

Theorem 1.3. *Let $K_1, \ldots, K_r$ be finite extensions of $\mathbb{Q}$ of degree at least 2, and let $n_1, \ldots, n_r$ denote their respective degrees. Let $S_1, \ldots, S_r$ be finite sets of primes. For each $i \in \{1, \ldots, r\}$, let R_i be the representation function of a norm form associated to $K_i/\mathbb{Q}$, and let $R_i^* := R_{i\,S_i}^*$ denote its restriction to integers that are square free outside S_i. Let $\mathfrak{K} \subset [-1, 1]^s$ be a convex body. Further, suppose that $f_1, \ldots, f_r \in \mathbb{Z}[u_1, \ldots, u_s]$ are pairwise non-proportional linear forms, and assume that $|f_i(\mathfrak{K})| \leqslant 1$ for $1 \leqslant i \leqslant r$. Given any modulus $q \in \mathbb{N}$ and a vector $\mathbf{a} \in (\mathbb{Z}/q\mathbb{Z})^s$ such that $v_p(f_i(\mathbf{a})) < v_p(q)$ for all $p \in S_i$ and $i \in \{1, \ldots, r\}$, we then have*

$$\sum_{\substack{\mathbf{u}\in\mathbb{Z}^s\cap T\mathfrak{K} \\ \mathbf{u}\equiv\mathbf{a} \pmod{q}}} \prod_{i=1}^{r} R_i^*(f_i(\mathbf{u})) = \beta_\infty \prod_p \beta_p \cdot T^s + o(T^s), \quad (T \to \infty),$$

where

$$\beta_\infty = \sum_{\boldsymbol{\epsilon}\in\{\pm\}^r} \mathrm{vol}(\mathfrak{K} \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \cdots \times \mathbb{R}_{\epsilon_r})) \prod_{i=1}^{r} \kappa_i^{\epsilon_i}$$

with

$$\kappa_i^{\epsilon_i} = \mathrm{vol}\{\mathbf{x} \in \mathfrak{D}_i^+ : 0 < \epsilon_i \mathbf{N}_{K_i}(\mathbf{x}) \leqslant 1\} \quad \textit{and} \quad \mathbf{f} = (f_1, \ldots, f_r) : \mathbb{Z}^s \to \mathbb{Z}^r,$$

and

$$\beta_p = \lim_{m\to\infty} \frac{1}{p^{ms}} \sum_{\substack{\mathbf{u}\in(\mathbb{Z}/p^m\mathbb{Z})^s \\ \mathbf{u}\equiv\mathbf{a} \pmod{p^{v_p(q)}}}} \prod_{i=1}^{r} \left(1 - \mathbf{1}_{p\notin S_i} \frac{\varrho_i(p^2, 0)}{p^{2n_i}}\right) \frac{\varrho_i(p^m, f_i(\mathbf{u}))}{p^{m(n_i-1)}},$$

for each prime p. Furthermore, the product $\prod_p \beta_p$ is absolutely convergent.

Remark. In all of this work, R_S^* could be replaced by the more general representation function that arises from replacing R by a function $R(\mathfrak{X}, M, \mathbf{b})$ as defined in [1, Definition 5.1]. This would allow one to prove a weak approximation result that takes not only the variables $\mathbf{u}$ from (1.2) into account, but also the variables $\mathbf{x}_i$. While working with a general function $R(\mathfrak{X}, M, \mathbf{b})$ requires essentially no additional work, we decided to restrict ourselves here to the special case of (1.3) for reasons of notational simplicity.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

The proof of [1, Theorem 5.2] uses the methods that were introduced in Green and Tao [3], and as such the two main steps in the proof are the construction of a family of pseudo-random majorants for a W-tricked version of the function R and the proof that this new function is orthogonal to nilsequences. The pseudo-random majorants constructed in [1] for the R-function are in fact pseudo-random majorants for the W-tricked version of R^*_S as well. Thus, the main step that is missing is to check orthogonality with nilsequences.

We prove the convergence of the product of local factors in § 2. As all remaining parts work exactly as in [1], our main focus here is to establish a non-correlation estimate (Theorem 1.4 below) for R^*_S that corresponds to [1, Proposition 6.3] with R replaced by its square-free version. Both § 3 and § 4 contain technical lemmas needed in the proof of Theorem 1.4 in § 5.

The statement of Theorem 1.4 requires a W-trick. In contrast to the case of R handled in [1], there is a lot of flexibility in the choice of the W-trick here, as the exceptionally large values of R, which were problematic before, occur at integers that are not square free. Given any integer $N > 0$, let $w(N) = \log\log N$, and set

$$W(N) = \prod_{p \leqslant w(N)} p^{\alpha(p)}, \tag{1.4}$$

where $\alpha(p) \in \mathbb{N}$ is such that

$$p^{\alpha(p)-1} < \log N \leqslant p^{\alpha(p)}.$$

Observe that $\alpha(p) > 1$. Taking N sufficiently large allows us to assume that any given finite set S of primes is contained in the set of primes less than $w(N)$, and moreover that $q | W(N)$ for any given integer q. Given a representation function R^*_S and any integer $T > 0$, we define the following set of 'unexceptional' residues:

$$\mathscr{A}(R^*_S, N) = \left\{ A \bmod W(N) : \begin{array}{ll} 0 \leqslant v_p(A) \leqslant 1 & \text{if } p < w(N),\ p \notin S \\ 0 \leqslant v_p(A) < v_p(W(N))/3 & \text{if } p \in S \\ \varrho(W(N), A) > 0 & \end{array} \right\}. \tag{1.5}$$

With the exception of integers that are divisible to a large order by some prime p from S, the support of R^*_S is contained in the set of numbers whose residues modulo $W(N)$ belong to $\mathscr{A}(R^*_S, N)$. Since S is finite, we can avoid the exceptional set by fixing the S-part $\prod_{p \in S} m^{v_p(m)}$ of integers m under consideration and taking N sufficiently large so that $v_p(m) < v_p(W(N))/3$.

With the W-trick in place, we are now ready to reveal the main result of this paper, which states that the W-tricked version of R^*_S is orthogonal to nilsequences.

Theorem 1.4. *Let G/Γ be a nilmanifold of dimension $m_G \geqslant 1$, let $G_\bullet$ be a filtration of G of degree $\ell \geqslant 1$, and let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ be a polynomial sequence. Suppose that G/Γ has a Q-rational Mal'cev basis $\mathscr{X}$ for some $Q \geqslant 2$, defining a metric $d_{\mathscr{X}}$ on G/Γ. Suppose that $F : G/\Gamma \to [-1, 1]$ is a Lipschitz function. Let N and $T = T(N)$ be positive integer parameters that satisfy $N^{1-\varepsilon} \ll_\varepsilon T \leqslant N$ for all $\varepsilon > 0$. Then, for $\epsilon \in \{\pm\}$, $W = W(N)$, and $A \in \mathbb{Z}$ with $A \pmod W \in \mathscr{A}(R^*_S, N)$ and $0 \leqslant \epsilon A < W$, we have the estimate*

$$\left|\frac{W}{T}\sum_{0<\epsilon m\leqslant T/W}\left(R_S^*(Wm+A)-\kappa^{\epsilon}\frac{\varrho(W,A)}{W^{n-1}}\prod_{p>w(N)}\left(1-\frac{\varrho(p^2,0)}{p^{2n}}\right)\right)F(g(|m|)\Gamma)\right|$$
$$\ll_{m_G,\ell,E} Q^{O_{m_G,\ell,E}(1)}\frac{1+\|F\|_{\mathrm{Lip}}}{(\log\log\log N)^E}\frac{\varrho(W,A)}{W^{n-1}}\prod_{p>w(N)}\left(1-\frac{\varrho(p^2,0)}{p^{2n}}\right),$$

for any $E>0$ and provided that N is sufficiently large.

General remarks. We assume familiarity with [1] throughout this paper. The ideas and proofs that we present in §§ 2 and 5 are very closely related to the material from [1]. The main new observation is the fact that these ideas can be made to work in the case of the square-free representation function by means of the new technical lemmas we prove in §§ 3 and 4.

2. Local factors and the deduction of Theorem 1.1

The aim of this section is to deduce Theorem 1.1 from Theorem 1.3. This deduction partially relies on the following proposition, which asymptotically evaluates the local factors from Theorem 1.3. Note that this proposition implies the final part of Theorem 1.3, namely that the product of local factors is absolutely convergent.

Proposition 2.1. *Let $L=\max_{1\leqslant i\leqslant r}\{\|f_i\|,s,r,n_i,|D_{K_i}|\}$, where $\|f_i\|$ denotes the maximum of the absolute values of the coefficients of f_i. Then the local factors from Theorem 1.3 satisfy both the following.*

(1) *$\beta_p=1+O_L(p^{-2})$ whenever $p\nmid q$.*

(2) *$\beta_p=O_{L,q}(1)$ at all primes p.*

In particular, there is $L'=O_L(1)$ such that $\beta_p>0$ provided that $p\nmid q$ and $p\geqslant L'$.

Proof. For every prime p, let

$$\beta_p'=\lim_{m\to\infty}\frac{1}{p^{ms}}\sum_{\substack{\mathbf{u}\in(\mathbb{Z}/p^m\mathbb{Z})^s\\ \mathbf{u}\equiv\mathbf{a}\ (\mathrm{mod}\ p^{v_p(q)})}}\prod_{i=1}^{r}\frac{\varrho_i(p^m,f_i(\mathbf{u}))}{p^{m(n_i-1)}}.$$

Then

$$\beta_p=\beta_p'\prod_{i=1}^{r}\left(1-\mathbf{1}_{p\notin S_i}\frac{\varrho_i(p^2,0)}{p^{2n_i}}\right),$$

and β_p' is the local factor that appears in [1, Theorem 5.2], but with $\varrho_i(p^m,f_i(\mathbf{u});p^{v_p(q)})$ replaced by $\varrho_i(p^m,f_i(\mathbf{u}))$. The proof of [1, Proposition 5.5] implies that $\beta_p'=O_{L,q}(1)$ and that $\beta_p'=1+O_L(p^{-2})$ whenever $p\nmid q$. Indeed, the second part of [1, Proposition 5.5] rests on the bound $\varrho_i(p^m,f_i(\mathbf{u});p^{v_p(q)})\leqslant\varrho_i(p^m,f_i(\mathbf{u}))$, and therefore includes a proof of assertion (1); assertion (2) follows from a direct application of [1, Proposition 5.5] with $M=q$, since $\varrho_i(p^m,f_i(\mathbf{u});p^{v_p(q)})=\varrho_i(p^m,f_i(\mathbf{u}))$ when $p\nmid q$. Thus, it remains to

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

show that

$$1 - \frac{\varrho_i(p^2, 0)}{p^{2n_i}} = 1 + O(p^{-2})$$

for all primes p and $i \in \{1, \ldots, r\}$. By [1, Lemma 4.5], there is $C > 0$ such that

$$\frac{\varrho_i(p^2, 0)}{p^{2(n_i-1)}} \leqslant C2^{n_i};$$

that is,

$$\frac{\varrho_i(p^2, 0)}{p^{2n_i}} \leqslant C\frac{2^{n_i}}{p^2}, \tag{2.1}$$

for each $1 \leqslant i \leqslant r$. □

We are now in a position to prove Theorem 1.1. The proof below is similar to that given in [1, § 5.3], but is significantly easier, since we only consider weak approximation in the variables $\mathbf{u}$ and not in the variables $\mathbf{x}_i$.

Proof of Theorem 1.1 assuming Theorem 1.3. First of all, we may assume that S contains all primes $p \leqslant L'$, where L' is given by Proposition 2.1 above. For $1 \leqslant i \leqslant r$, let R_i be the representation function of some norm form $\mathbf{N}_{K_i}$, and let R_i^* be its restriction to integers m that are square free outside S. Then it suffices to show that, given any $\varepsilon > 0$, there exists a vector $\mathbf{u}' \in \mathbb{Z}^s$ such that the following hold.

(1) $|\mathbf{u}' - \mathbf{u}|_p < \varepsilon$ for all $p \in S$.

(2) $|t\mathbf{u}' - \mathbf{u}| < \varepsilon$ for some $t > 0$.

(3) $\prod_{i=1}^r R_i^*(f_i(\mathbf{u}')) > 0$.

For every $\varepsilon > 0$ there is an integer Q composed of primes from S such that condition (1) is implied by the congruence

$$\mathbf{u}' \equiv \mathbf{u} \pmod{Q},$$

and such that $v_p(Q) > v_p(f_i(\mathbf{u}))$ for every $p \in S$. Further, let

$$\mathfrak{K}(\mathbf{u}; \varepsilon) = \{\mathbf{v} \in \mathbb{R}^s : |\mathbf{u} - \mathbf{v}| < \varepsilon\},$$

and note that, whenever $T > 0$ and $\mathbf{u}' \in T\mathfrak{K}(\mathbf{u}; \varepsilon)$, the second condition is satisfied. Thus it is enough to show that

$$N(T) = \sum_{\substack{\mathbf{u}' \in \mathbb{Z}^s \cap T\mathfrak{K}(\mathbf{u}; \varepsilon') \\ \mathbf{u}' \equiv \mathbf{u} \pmod{Q}}} \prod_{i=1}^r R_i^*(f_i(\mathbf{u}')) > 0$$

for some value of $T > 0$. Theorem 1.3 implies that

$$N(T) \geqslant T^s \operatorname{vol}(\mathfrak{K}(\mathbf{u}, \varepsilon) \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \cdots \times \mathbb{R}_{\epsilon_r}))\kappa_1^{\epsilon_1} \ldots \kappa_r^{\epsilon_r} \prod_p \beta_p + o(T^s),$$

where $\epsilon_i = \operatorname{sign} f_i(\mathbf{u})$. Hence, the result follows by taking T sufficiently large, provided that we can show that the product of local factors on the right-hand side is positive.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

To see this, first note that every factor $\kappa_i^{\epsilon_i} = \mathrm{vol}\{\mathbf{x} \in \mathfrak{D}_i^+ : 0 < \epsilon_i \mathbf{N}_{K_i}(\mathbf{x}) \leqslant 1\}$ is positive, since $f_i(\mathbf{u}) \neq 0$ is a local norm from K_i at the real place. Indeed, there is a vector $\mathbf{x}_i \in \mathbb{R}^{n_i} \cap \mathfrak{D}_i^+$ such that $f_i(\mathbf{u}) = \mathbf{N}_{K_i}(\mathbf{x}_i)$, and, since $\epsilon_i = \mathrm{sign}\, f_i(\mathbf{u})$, we have

$$t\mathbf{x}_i \in \{\mathbf{x} \in \mathfrak{D}_i^+ : 0 < \epsilon_i \mathbf{N}_{K_i}(\mathbf{x}) < 1\}$$

for every sufficiently small $t > 0$. By continuity of $\mathbf{N}_{K_i}$, the above set is open and has positive volume, since it is non-empty.

Since $f_i(\mathbf{u}) \neq 0$ for all $i \in \{1, \ldots, r\}$, the open set

$$\mathfrak{K}(\mathbf{u}, \varepsilon) \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \cdots \times \mathbb{R}_{\epsilon_r})$$

contains $\mathbf{u}$, which again implies that $\mathrm{vol}(\mathfrak{K}(\mathbf{u}, \varepsilon) \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \cdots \times \mathbb{R}_{\epsilon_r})) > 0$.

By Proposition 2.1, we have $\prod_{p \notin S} \beta_p > 0$, so it remains to check that $\beta_p > 0$ whenever $p \in S$. We proceed as in [1, § 5.3].

Let p be any element from S, and recall that for $1 \leqslant i \leqslant r$ there is an integral element $k_i \in K_i$ such that $f_i(\mathbf{u}) = N_{K_i \otimes_{\mathbb{Q}} \mathbb{Q}_p / \mathbb{Q}_p}(k_i)$. This implies that for every $m > 0$ there is a vector $\mathbf{x}_i \in \mathbb{Z}^{n_i}$ such that

$$f_i(\mathbf{u}) \equiv \mathbf{N}_{K_i}(\mathbf{x}_i) \pmod{p^m},$$

and hence

$$\prod_{i=1}^{r} \varrho_i(p^m, f_i(\mathbf{u})) \geqslant 1.$$

Choosing

$$m = 2\left(1 + v_p(Q) + \sum_{i=1}^{r} v_p(f_i(\mathbf{u})) + \sum_{i=1}^{r} v_p(n_i)\right),$$

we apply [1, Lemma 3.4] with $A = f_i(\mathbf{u})$, $G = \mathbf{N}_{K_i}$, and $\ell = 0$ to deduce that

$$\prod_{i=1}^{r} \frac{\varrho_i(p^{m'}, f_i(\tilde{\mathbf{u}}))}{p^{m'(n_i-1)}} = \prod_{i=1}^{r} \frac{\varrho_i(p^m, f_i(\mathbf{u}))}{p^{m(n_i-1)}} \geqslant \prod_{i=1}^{r} \frac{1}{p^{m(n_i-1)}}$$

whenever $m' > m$ and $\tilde{\mathbf{u}} \in (\mathbb{Z}/p^{m'}\mathbb{Z})^s$ is such that $\tilde{\mathbf{u}} \equiv \mathbf{u} \pmod{p^m}$. For any given $m' > m$, there are $p^{(m'-m)s}$ admissible choices for $\tilde{\mathbf{u}}$. Note that $m > v_p(Q)$. Thus,

$$\beta_p \geqslant \lim_{m' \to \infty} \frac{1}{p^{m's}} \sum_{\substack{\tilde{\mathbf{u}} \in (\mathbb{Z}/p^{m'}\mathbb{Z})^s \\ \tilde{\mathbf{u}} \equiv \mathbf{u} \pmod{p^{m'}}}} \prod_{i=1}^{r} \frac{\varrho_i(p^{m'}, f_i(\tilde{\mathbf{u}}))}{p^{m'(n_i-1)}} \geqslant \frac{1}{p^{m(n_1+\cdots+n_r+s-r)}} > 0,$$

which completes the proof. □

3. R_S^* in arithmetic progressions

This section contains two lemmas about the mean value of R_S^* in arithmetic progressions. These will be required in the proof of Theorem 1.4 in § 5.

Lemma 3.1. *Let S be a finite set of primes, and let R_S^* be the corresponding restriction of the representation function. Let N and q be positive integers such that $p | q$ for every prime $p < w(N)$ and such that $v_p(q) \neq 1$ for all $p \notin S$. Suppose further that $A \in \{1, \ldots, q\}$ is*

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

an integer such that $0 \leqslant v_p(A) \leqslant 1$ *whenever* $p|q$ *and* $p \notin S$. *Let* $\epsilon \in \{\pm\}$, *and let* κ^ϵ *be the constant that appears in* [1, Lemma 6.1]. *Then, provided that* N *is sufficiently large,*

$$\sum_{\substack{0<\epsilon m\leqslant x\\ m\equiv A \pmod q}} R_S^*(m) = \kappa^\epsilon x \frac{\varrho(q,A)}{q^n} \prod_{p\nmid q}\left(1-\frac{\varrho(p^2,0)}{p^{2n}}\right) + O(qx^{1-\frac{1}{20n}}). \tag{3.1}$$

Proof. We shall deduce this result from [1, Lemma 6.1], which states that

$$\sum_{\substack{0<\epsilon m\leqslant x\\ m\equiv A' \pmod{q'}}} R(m) = \frac{\varrho(q',A')}{q'^n}\kappa^\epsilon x + O(q' x^{1-1/n}), \tag{3.2}$$

for any positive integer q', any $A' \in \mathbb{Z}$, and $\epsilon \in \{\pm\}$. Since

$$R_S^*(m) = \sum_{\substack{d^2|m\\ \gcd(d,q)=1}} \mu(d) R(m)$$

for $m \equiv A \pmod q$, we have

$$\sum_{\substack{0<\epsilon m\leqslant x\\ m\equiv A \pmod q}} R_S^*(m) = \sum_{\substack{d\leqslant x^{1/2}\\ \gcd(d,q)=1}} \mu(d) \sum_{\substack{0<\epsilon m\leqslant x\\ m\equiv A \pmod q\\ m\equiv 0 \pmod{d^2}}} R(m). \tag{3.3}$$

If d is sufficiently small, then the inner sum may be evaluated by means of (3.2). Invoking the Chinese remainder theorem, it follows that

$$\begin{aligned}\sum_{\substack{d<x^{1/10n}\\ \gcd(d,q)=1}} \mu(d) &\sum_{\substack{0<\epsilon m\leqslant x\\ m\equiv A \pmod q\\ m\equiv 0 \pmod{d^2}}} R(m)\\ &= \sum_{\substack{d\leqslant x^{1/10n}\\ \gcd(d,q)=1}} \left(\mu(d)\frac{\varrho(d^2,0)}{d^{2n}}\frac{\varrho(q,A)}{q^n}\kappa^\epsilon x + O(d^2 q x^{1-\frac{1}{n}})\right)\\ &= \sum_{\substack{d\leqslant x^{1/10n}\\ \gcd(d,q)=1}} \mu(d)\frac{\varrho(d^2,0)}{d^{2n}}\frac{\varrho(q,A)}{q^n}\kappa^\epsilon x + O(q x^{1-\frac{1}{n}+\frac{3}{10n}}).\end{aligned} \tag{3.4}$$

We aim to extend the summation in d to all positive integers that are co-prime to q. By multiplicativity of ϱ we deduce from (2.1) that

$$\mu^2(d)\frac{\varrho(d^2,0)}{d^{2n}} \leqslant \frac{(C2^n)^{\omega(d)}}{d^2} \ll_{C,n,\varepsilon} d^{-2+\varepsilon} \tag{3.5}$$

for any $\varepsilon > 0$. Since

$$\sum_{d>x^{1/10n}} d^{-2+\varepsilon} \ll x^{\frac{1}{10n}(-2+\varepsilon+1)} \ll x^{-\frac{1}{20n}}$$

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

for ε sufficiently small, (3.4) is seen to equal

$$\begin{aligned}
&\kappa^{\epsilon} x \frac{\varrho(q,A)}{q^n}\Bigg(\sum_{\substack{d\geqslant 1\\ \gcd(d,q)=1}} \mu(d)\frac{\varrho(d^2,0)}{d^{2n}} + O(x^{-\frac{1}{20n}})\Bigg) + O(qx^{1-\frac{1}{n}+\frac{3}{10n}})\\
&\quad = \kappa^{\epsilon} x \frac{\varrho(q,A)}{q^n}\prod_{p\nmid q}\bigg(1-\frac{\varrho(p^2,0)}{p^{2n}}\bigg) + O(x^{1-\frac{1}{20n}}) + O(qx^{1-\frac{1}{n}+\frac{3}{10n}}), \qquad (3.6)
\end{aligned}$$

where we made use of the trivial bound $\varrho(q,A) \ll q^n$.

In order to bound the tail of the summation in d from (3.3), we consider a fixed square-free integer d with $\gcd(d,q)=1$, and let $d = p_1 \dots p_k$ be its prime factorization. As shown in [1, Lemma 8.1], the representation function R may be uniformly bounded above by a multiplicative function of constant average order. More precisely, we have $R(m) \ll r_K(|m|)$ for all $m \neq 0$, where r_K is the multiplicative function describing the Dirichlet coefficients of the Dedekind zeta function of K; i.e., $\zeta_K(s) = \sum_n r_K(n) n^{-s}$. Invoking two basic properties of r_K, namely that $r_K(m) \ll \tau(m)^n$ and that $\sum_{m\leqslant x} r_K(m) \ll x$ (cf. [1, (2.8) and (2.10)]), we deduce that

$$\begin{aligned}
\sum_{\substack{m\equiv A \ (\mathrm{mod}\ q)\\ m\equiv 0\ (\mathrm{mod}\ d^2)\\ 0<\epsilon m\leqslant x}} R(m) &\ll \sum_{(b_1,\dots,b_k)\in\mathbb{N}_0^k} r_K(d^2 p_1^{b_1}\dots p_k^{b_k}) \sum_{0<m\leqslant x/(d^2 p_1^{b_1}\dots p_k^{b_k})} r_K(m)\\
&\ll \sum_{\substack{(b_1,\dots,b_k)\in\mathbb{N}_0^k\\ p_1^{b_1}\dots p_k^{b_k}\leqslant x}} \tau(d^2 p_1^{b_1}\dots p_k^{b_k})^n \frac{x}{d^2 p_1^{b_1}\dots p_k^{b_k}}\\
&\ll x\prod_{i=1}^k \sum_{b_i\geqslant 2}\frac{\tau(p_i^{b_i})^n}{p_i^{b_i}} \ll x\prod_{i=1}^k\sum_{b_i\geqslant 2}\bigg(\frac{2^n}{p_i}\bigg)^{b_i} \ll x\prod_{i=1}^k \frac{2^{2n+1}}{{p_i}^2} \ll x\frac{\tau(d^2)^{n+1}}{d^2}.
\end{aligned}$$

Thus, for any $C_0 > 2$, we have

$$\begin{aligned}
\Bigg|\sum_{\substack{x^{1/C_0}\leqslant d\leqslant x^{1/2}\\ \gcd(d,q)=1}} \mu(d) \sum_{\substack{m\equiv A\ (\mathrm{mod}\ q)\\ m\equiv 0\ (\mathrm{mod}\ d^2)\\ 0<\epsilon m\leqslant x}} R(m)\Bigg| &\leqslant \sum_{\substack{x^{1/C_0}\leqslant d\leqslant x^{1/2}\\ \gcd(d,q)=1\\ \mu^2(d)=1}} \sum_{\substack{m\equiv A\ (\mathrm{mod}\ q)\\ m\equiv 0\ (\mathrm{mod}\ d^2)\\ 0<\epsilon m\leqslant x}} R(m)\\
&\ll x \sum_{\substack{x^{1/C_0}\leqslant d\leqslant x^{1/2}\\ \gcd(d,q)=1}} \frac{\tau(d^2)^{n+1}}{d^2}\\
&\ll_\varepsilon x \sum_{\substack{x^{1/C_0}\leqslant d\leqslant x^{1/2}\\ \gcd(d,q)=1}} d^{-2+\varepsilon}\\
&\ll_\varepsilon x^{1-\frac{1-\varepsilon}{C_0}}. \qquad (3.7)
\end{aligned}$$

Combining this estimate for $C_0 = 10n$ with (3.4) and (3.6) completes the proof. □

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

Our next aim is to establish the 'major arc estimate' that is required in order to deduce Theorem 1.4 from a non-correlation estimate that only involves sufficiently *equidistributed* polynomial nilsequences. The following lemma corresponds to [1, Lemma 6.2], and it shows that the W-tricked version of R_S^* has constant average values on certain subprogressions. Recall the definition of $\mathscr{A}(R_S^*, N)$ from (1.5).

Lemma 3.2 (Major arc estimate). *Let $\epsilon \in \{\pm\}$, and let $N > 0$ be an integer. Suppose that $A \in \mathscr{A}(R_S^*, N)$, and let q_0 be a $w(N)$-smooth number. Let $x, x' \in \mathbb{Z}_{>0}$ be parameters that satisfy $x \asymp x'$. Then*

$$\frac{W(N)}{x} \sum_{\substack{m \equiv A \pmod{W} \\ 0 < \epsilon m \leqslant x}} R_S^*(m) = \frac{W(N) q_0}{x'} \sum_{\substack{m \equiv A + W q_1 \pmod{W q_0} \\ 0 < \epsilon m \leqslant x'}} R_S^*(m) + O(q_0^2 W(N)^2 x^{-\frac{1}{20n}})$$

for any $q_1 \in \mathbb{Z}$.

Proof. We shall employ the lifting result [1, Lemma 3.4], which in our context states the following. Let $m \geqslant 1$, $A' \neq 0$, and assume that

$$v_p(A') + v_p(n) < \frac{m}{2}.$$

Then we have

$$\frac{\varrho(p^m, A')}{p^{m(n-1)}} = \frac{\varrho(p^{m+1}, A' + kp^m)}{p^{(m+1)(n-1)}}, \tag{3.8}$$

uniformly for $k \in \mathbb{Z}/p\mathbb{Z}$.

Since the definition of $\mathscr{A}(R_S^*, N)$ guarantees that all of the above assumptions are satisfied, we deduce that

$$\frac{\varrho(W, A)}{W^{n-1}} = \frac{\varrho(W q_0, A + W q_1)}{(W q_0)^{n-1}}. \tag{3.9}$$

The lemma now follows from an application of Lemma 3.1 to each of the two sums over R_S^* from the statement, combined with an application of identity (3.9). □

4. Polynomial subsequences of multiparameter nilsequences

In this section we recall some of the background on equidistribution of multiparameter polynomial nilsequences and prove several technical results that analyse to what extent equidistribution properties are preserved when passing to certain subsequences or families of subsequences.

Throughout what follows, $[x]$ denotes the set of integers $\{1, \ldots, \lfloor x \rfloor\}$. We shall be working with the quantitative notion of equidistribution that was introduced by Green and Tao in [4, Definition 8.5].

Definition 4.1 (Quantitative equidistribution). Let G/Γ be a nilmanifold equipped with Haar measure, and let $\delta > 0$. A finite sequence

$$(g(\mathbf{n})\Gamma)_{\mathbf{n} \in [N_1] \times \cdots \times [N_t]}$$

taking values in G/Γ is called δ-*equidistributed* if

$$\left|\frac{1}{|N_1|\dots|N_t|}\sum_{\mathbf{n}\in[N_1]\times\cdots\times[N_t]} F(g(\mathbf{n})\Gamma) - \int_{G/\Gamma} F\right| \leqslant \delta\|F\|_{\mathrm{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \to \mathbb{C}$. It is said to be *totally* δ-*equidistributed* if, moreover,

$$\left|\frac{1}{|P_1|\dots|P_t|}\sum_{\mathbf{n}\in P_1\times\cdots\times P_t} F(g(\mathbf{n})\Gamma) - \int_{G/\Gamma} F\right| \leqslant \delta\|F\|_{\mathrm{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \to \mathbb{C}$, and for all collections of arithmetic progressions $P_i \subset \{1,\dots,N_i\}$ of length $|P_i| \geqslant \delta N_i$ for $1 \leqslant i \leqslant t$.

The most relevant measures in the analysis of quantitative equidistribution of polynomial sequences are the smoothness norms. These, too, were introduced in [4]; see also [6].

Definition 4.2 (Smoothness norms). Let $f : \mathbb{Z}^t \to \mathbb{R}/\mathbb{Z}$ be a polynomial of degree d, and suppose that

$$f(n_1,\dots,n_t) = \sum_{\substack{(i_1,\dots,i_t)\in\mathbb{Z}^t_{\geqslant 0}\\ i_1+\cdots+i_t\leqslant d}} \beta_{i_1,\dots,i_t} n_1^{i_1}\dots n_t^{i_t}.$$

Then

$$\|f\|_{C^\infty_*[N_1]\times\cdots\times[N_t]} := \sup_{(i_1,\dots,i_t)\neq\mathbf{0}} N_1^{i_1}\dots N_t^{i_t}\|\beta_{i_1,\dots,i_t}\|.$$

Finally, recall that a continuous additive homomorphism $\eta : G \to \mathbb{R}/\mathbb{Z}$ is called a horizontal character if it annihilates Γ. The equidistribution properties of multiparameter nilsequences can be analysed through horizontal characters on G/Γ via a theorem of Green and Tao [4, Theorem 8.6], which we state below. See [6] for a proof.

Theorem 4.3 [4, 6]. *Let* $0 < \delta < 1/2$*, and let* $m, t, d, N_1, \dots, N_t \geqslant 1$ *be positive integers. Suppose that* G/Γ *is an* m*-dimensional nilmanifold equipped with a* $\frac{1}{\delta}$*-rational Mal'cev basis* $\mathscr{X}$ *adapted to some filtration* $G_\bullet$ *of degree* ℓ*, and that* $g \in \mathrm{poly}(\mathbb{Z}^t, G_\bullet)$*. Then either* $(g(\mathbf{n})\Gamma)_{\mathbf{n}\in[N_1]\times\cdots\times[N_t]}$ *is* δ*-equidistributed, or else there is some horizontal character* η *with* $0 < |\eta| \ll \delta^{-O_{\ell,m,t}(1)}$ *such that*

$$\|\eta\circ g\|_{C^\infty_*[N_1]\times\cdots\times[N_t]} \ll \delta^{-O_{\ell,m,t}(1)}.$$

In the case of polynomial nilsequences, the quantitative notions of equidistribution and total equidistribution are equivalent, with polynomial dependence in the equidistribution parameter. The following lemma handles the non-trivial direction of this equivalence.

Lemma 4.4. *Suppose that* $\delta : \mathbb{N} \to (0, 1/2)$ *is such that* $\delta(x)^{-T} \ll_T x$ *for* $T > 0$*. Let* $m, \ell, t,$ *and* N *be positive integers, and suppose that* G/Γ *is an* m*-dimensional nilmanifold*

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

equipped with a $\frac{1}{\delta(N)}$-rational Mal'cev basis $\mathscr{X}$ adapted to some filtration $G_\bullet$ of degree ℓ. Then there is a constant $1 \leqslant C \ll_{\ell,m} 1$ such that the following holds.

Let $E > C$, and let $g \in \mathrm{poly}(\mathbb{Z}^t, G_\bullet)$. Suppose that the finite sequence $(g(\mathbf{n})\Gamma)_{\mathbf{n}\in[N]^t}$ is $\delta(N)^E$-equidistributed. Then $(g(\mathbf{n})\Gamma)_{\mathbf{n}\in[N]^t}$ is totally $\delta(N)^{E/C}$-equidistributed, provided that N is sufficiently large.

Proof. This is a rather straightforward generalization of the computation carried out in the proof of [11, Lemma 6.2]. □

We note aside that the factorization theorem for nilsequences [4, Theorem 1.19] will allow us to assume that E is sufficiently large for the condition $E > C$ of the above lemma to be satisfied in all instances when we make use of it.

The multiparameter nilsequences that will be most relevant to the proof of Theorem 1.4 are those that arise as the composition $g \circ P$ of a polynomial $P \in \mathbb{Z}[X_1, \dots, X_t]$ and a one-parameter nilsequence g.

Lemma 4.5. *Suppose that $\delta : \mathbb{N} \to (0, 1/2)$ is a function that satisfies $\delta(x)^{-T} \ll_T x$ for all $T > 0$. Let m, ℓ, t, and N be positive integers, and suppose that G/Γ is an m-dimensional nilmanifold equipped with a $\frac{1}{\delta(N)}$-rational Mal'cev basis $\mathscr{X}$ adapted to some filtration $G_\bullet$ of degree ℓ. Let $P \in \mathbb{Z}[X_1, \dots, X_t]$ be a homogeneous polynomial of degree t, fixed once and for all, and let all implied constants be allowed to depend on the coefficients of P in any way. Then there is a constant $1 \leqslant C \ll_{t,\ell} 1$ such that the following holds.*

Let $E > C$, and let $g \in \mathrm{poly}(\mathbb{Z}^t, G_\bullet)$. Suppose that $(g(n)\Gamma)_{n\leqslant N}$ is totally $\delta(N)^E$-equidistributed. Then

$$(g(P(n_1, \dots, n_t)\Gamma))_{(n_1,\dots,n_t)\in[N^{1/t}]\times\cdots\times[N^{1/t}]}$$

is totally $\delta(N)^{E/C}$-equidistributed whenever N is sufficiently large.

Proof. In order to apply Lemma 4.4 and Theorem 4.3 to the sequence $(g(P(\mathbf{n})))_{\mathbf{n}\in\mathbb{Z}^t}$, we require first of all a filtration $G'_\bullet$ with respect to which $(g(P(\mathbf{n})))_{\mathbf{n}\in\mathbb{Z}^t}$ is a polynomial sequence. It follows from [4, Lemma 6.7] that g has a representation of the form $g(n) = a_1^{P_1(n)} \dots a_k^{P_k(n)}$, where $a_1, \dots, a_k \in G$, and where $P_1, \dots, P_k \in \mathbb{Z}[X]$ are polynomials of degree at most ℓ. Thus, $g(P(\mathbf{n})) = a_1^{P_1\circ P(\mathbf{n})} \dots a_k^{P_k\circ P(\mathbf{n})}$, where each polynomial $P_i \circ P$ has degree at most $\ell' = \ell \cdot \max_{1\leqslant i\leqslant k}(\deg P_i)$. Define a filtration $G'_\bullet$ by setting $G'_j = G_{\lceil j/\ell' \rceil}$ for $0 \leqslant j \leqslant \ell'\ell$. Then it is immediate (cf. the discussion following [4, Lemma 6.7]) that $(a_i^{P_i\circ P(\mathbf{n})})_{\mathbf{n}\in\mathbb{Z}^t}$ belongs to $\mathrm{poly}(G'_\bullet, \mathbb{Z}^t)$ for each $1 \leqslant i \leqslant k$. By Leibman's theorem [9] (see [4] for a different proof), the set $\mathrm{poly}(G'_\bullet, \mathbb{Z}^t)$ forms a group. Thus it follows that $(g(P(\mathbf{n})))_{\mathbf{n}\in\mathbb{Z}^t} \in \mathrm{poly}(\mathbb{Z}^t, G'_\bullet)$.

Finally, observe that the given Mal'cev basis $\mathscr{X}$ is a Mal'cev basis adapted to $G'_\bullet$ as well. Indeed, part (ii) of [4, Definition 2.1] follows immediately from the corresponding statement for $G_\bullet$, since $\{G'_i : 0 \leqslant i \leqslant s\} \subseteq \{G_i : 0 \leqslant i \leqslant s\}$.

We are now in the position to start with the proof of the lemma. Suppose that $B \geqslant 1$ and that

$$(g(P(n_1, \dots, n_t)\Gamma))_{(n_1,\dots,n_t)\in[N^{1/t}]^t}$$

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

fails to be totally $\delta(N)^B$-equidistributed. Then, by Lemma 4.4, there is a constant $1 \leqslant C_1 \ll_{\ell,m} 1$ such that the above sequence also fails to be $\delta(N)^{C_1 B}$-equidistributed. Thus, Theorem 4.3 implies that there is a non-trivial horizontal character $\eta : G/\Gamma \to \mathbb{C}$ of modulus $|\eta| \ll \delta(N)^{-O_{\ell,m,t}(B)}$ such that

$$\|\eta \circ g \circ P\|_{C_*^\infty[N^{1/t}]^t} \ll \delta(N)^{-O_{\ell,m,t}(B)}.$$

Writing

$$P(n)^j = \sum_{\substack{i_1,\ldots,i_t \geqslant 0 \\ i_1+\cdots+i_t = tj}} \gamma^{(j)}_{i_1,\ldots,i_t} n_1^{i_1} \ldots n_t^{i_t} \quad (1 \leqslant j \leqslant \ell),$$

then all coefficients $\gamma^{(j)}_{i_1,\ldots,i_t}$ are bounded. If, further,

$$\eta \circ g(n) = \sum_{i=j}^{\ell} \beta_j n^j,$$

then

$$\|\eta \circ g \circ P\|_{C_*^\infty[N^{1/t}]^t} = \sup_{\substack{1 \leqslant j \leqslant \ell \\ i_1+\cdots+i_t = tj}} N^j \|\beta_j \gamma^{(j)}_{i_1,\ldots,i_t}\| \ll \delta(N)^{-O_{\ell,m,t}(B)}. \tag{4.1}$$

Let γ be the least common multiple of all non-zero coefficients $\gamma^{(j)}_{i_1,\ldots,i_t}$, and observe that $\gamma \ll_{\ell,m,t} 1$. Since $\delta(x)^{-T} \ll_T x$, we have

$$\|\beta_j \gamma^{(j)}_{i_1,\ldots,i_t}\| \ll \delta(N)^{-O_{\ell,t,m}(B)} N^{-j} = o_{\ell,t,m}(1)$$

whenever $i_1 + \cdots + i_t = tj$ and $1 \leqslant j \leqslant \ell$. Hence, given any $A > 0$ and provided that N is sufficiently large with respect to ℓ, m, t and A, then $\|A'\beta_j \gamma^{(j)}_{i_1,\ldots,i_t}\| = A'\|\beta_j \gamma^{(j)}_{i_1,\ldots,i_t}\|$ for any positive real $A' \leqslant A$. In particular, we may assume that $\|\beta_j \gamma\| \ll_{\ell,t,m} \|\beta_j \gamma^{(j)}_{i_1,\ldots,i_t}\|$ whenever $\gamma^{(j)}_{i_1,\ldots,i_t}$ is non-zero. Since for every $j \in \{1, \ldots, \ell\}$ at least one of the coefficients $\gamma^{(j)}_{i_1,\ldots,i_t}$ of P^j is non-zero, the above and (4.1) imply that

$$\|\gamma \eta \circ g\|_{C_*^\infty[N]} = \sup_{1 \leqslant j \leqslant \ell} N^j \|\beta_j \gamma\| \ll \delta^{-O_{\ell,m,t}(B)},$$

provided that N is sufficiently large. Since $\gamma\eta$ is a non-trivial horizontal character of modulus $|\gamma\eta| \ll \delta^{-O_{\ell,m,t}(B)}$, we deduce that (cf. [10, Propositions 14.2 and 14.3]) there is a constant $C_2 \asymp_{\ell,m,t} 1$ such that $(g(n)\Gamma)_{n \leqslant N}$ fails to be totally $\delta(N)^{C_2 B}$-equidistributed. Choosing $C = \max(1, C_2)$, the result follows for every $E > C$ by setting $B = E/C$. Indeed, when $E = CB \geqslant C_2 B$, then the above conclusion that $(g(n)\Gamma)_{n \leqslant N}$ fails to be totally $\delta(N)^{C_2 B}$-equidistributed contradicts the assumption that this sequence is totally $\delta(N)^E$-equidistributed. □

Our next aim is to extend the above lemma in a way that allows us to replace the homogeneous polynomial P by an inhomogeneous polynomial of the form

$$\mathbf{x} \mapsto \frac{P(Wq\mathbf{x} + \mathbf{y}) - A'}{Wq},$$

where $W = W(N)$ is given by (1.4), where $q \in \mathbb{N}$, where $\mathbf{y} \in \mathbb{Z}^t$ is such that $0 \leqslant y_i < Wq$ for $1 \leqslant i \leqslant t$, and where $A' \in \mathbb{Z}$ is such that $P(\mathbf{y}) \equiv A' \pmod{Wq}$ and $|A'| < Wq$.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

Lemma 4.6. *Let N be a positive integer, and suppose that $T = T(N)$ satisfies $N^{1-\varepsilon} \ll_{\varepsilon} T \leqslant N$ for all $\varepsilon > 0$. Let $\delta : \mathbb{N} \to (0, 1/2)$ be a function that satisfies $\delta(x)^{-1} \ll_{\varepsilon} x^{\varepsilon}$ for all $\varepsilon > 0$. Let m, ℓ and t be positive integers, and suppose that G/Γ is an m-dimensional nilmanifold equipped with a $\frac{1}{\delta(N)}$-rational Mal'cev basis $\mathscr{X}$ adapted to some filtration $G_{\bullet}$ of degree ℓ. Let $g \in \mathrm{poly}(\mathbb{Z}, G_{\bullet})$ be any polynomial sequence, and let $P \in \mathbb{Z}[X_1, \ldots, X_t]$ be a fixed homogeneous polynomial of degree t. All implied constants are allowed to depend on the coefficients of P in any way. Suppose that $S : \mathbb{N} \to \mathbb{N}$ satisfies $S(x) \ll_{\varepsilon} x^{\varepsilon}$ for all $\varepsilon > 0$.*

Then there is a constant $1 \leqslant C \ll_{m,\ell,t} 1$ such that the following holds. Let $E > C$, and suppose that for every $w(N)$-smooth integer $\tilde{q} \leqslant S(N)$ the sequence $(g(\tilde{q}n)\Gamma)_{n \leqslant T/\tilde{q}}$ is totally $\delta(N)^E$-equidistributed. Further, let $q > 0$ be a $w(N)$-smooth integer that satisfies the bound $(Wq)^{t\ell^2} \leqslant S(N)$, where $W = W(N)$. Then, provided that N and T are sufficiently large,

$$\left(g\left(\frac{P(Wq\mathbf{x}+\mathbf{y}) - A'}{Wq}\right)\Gamma\right)_{\mathbf{x} \in \left[\left(\frac{T}{(Wq)^{t-1}}\right)^{1/t}\right]^t}$$

is a totally $\delta(N)^{E/C}$-equidistributed sequence for every choice of $\mathbf{y} \in \mathbb{Z}^t$ such that $0 \leqslant y_i < Wq$ for $1 \leqslant i \leqslant t$, and for $A' \in \mathbb{Z}$ such that $P(\mathbf{y}) \equiv A' \pmod{Wq}$ and $|A'| < Wq$.

Proof. Let us write

$$\tilde{g}(\mathbf{x}) = g\left(\frac{P(Wq\mathbf{x}+\mathbf{y}) - A'}{Wq}\right).$$

As in the previous proof, there is a refinement $G'_{\bullet}$ of the filtration $G_{\bullet}$ such that the new filtration is adapted to the basis $\mathscr{X}$, and its degree is of order $O_{\ell,t}(1)$, and such that $(\tilde{g}(\mathbf{x})\Gamma)_{\mathbf{x} \in \mathbb{Z}^t} \in \mathrm{poly}(G'_{\bullet}, \mathbb{Z}^t)$.

Let $B > 1$, and suppose that

$$\left(\tilde{g}(\mathbf{x})\Gamma\right)_{\mathbf{x} \in \left[\left(\frac{T}{(Wq)^{t-1}}\right)^{1/t}\right]^t}$$

fails to be totally $\delta(N)^B$-equidistributed. Then, as in the proof of the previous lemma, Lemma 4.4 and Theorem 4.3 imply that there is a non-trivial horizontal character $\eta : G/\Gamma \to \mathbb{C}$ such that $|\eta| \ll \delta(N)^{-O_{\ell,m,t}(B)}$ and

$$\|\eta \circ \tilde{g}\|_{C^{\infty}_{*}\left[\left(\frac{T}{(Wq)^{t-1}}\right)^{1/t}\right]^t} \ll \delta^{-O_{m,\ell,t}(B)}.$$

Suppose that

$$P(\mathbf{n})^j = \sum_{\substack{i_1,\ldots,i_t \geqslant 0 \\ i_1+\cdots+i_t = tj}} \gamma^{(j)}_{i_1,\ldots,i_t} n_1^{i_1} \ldots n_t^{i_t}, \quad (1 \leqslant j \leqslant \ell),$$

and note that for each j at least one of the coefficients $\gamma^{(j)}_{i_1,\ldots,i_t}$ is non-zero. Furthermore, suppose that

$$\eta \circ g(n) = \sum_{j=0}^{\ell} \beta_j n^j,$$

where $\beta_j \neq 0$ for at least one value $j > 0$.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

We proceed by analysing the coefficients of the polynomial map $\mathbf{x} \mapsto \eta \circ \tilde{g}(\mathbf{x}) = \eta \circ g(\frac{P(Wq\mathbf{x}+\mathbf{y})-A'}{Wq})$. To begin with, observe that

$$\frac{P(Wq\mathbf{x}+\mathbf{y})-A'}{Wq} = (Wq)^{t-1}P(\mathbf{x}) + P'(\mathbf{x}),$$

for some polynomial P' of degree $t-1$ with coefficients of order $O((Wq)^{t-1})$. Inserting this information into the above expression for $\eta \circ g$, we obtain

$$\eta \circ g\left(\frac{P(Wq\mathbf{x}+\mathbf{y})-A'}{Wq}\right) = \sum_{j=0}^{\ell} \beta_j \sum_{\substack{i_1,\ldots,i_t \geqslant 0 \\ i_1+\cdots+i_t=tj}} (Wq)^{(t-1)j} \gamma^{(j)}_{i_1,\ldots,i_t} x_1^{i_1} \ldots x_t^{i_t}$$
$$+ \sum_{j=0}^{\ell} \beta_j \sum_{\substack{i_1,\ldots,i_t \geqslant 0 \\ i_1+\cdots+i_t\leqslant tj-1}} c^{(j)}_{i_1,\ldots,i_t} x_1^{i_1} \ldots x_t^{i_t},$$

where $|c^{(j)}_{i_1,\ldots,i_t}| \ll (Wq)^{(t-1)j}$. If $\eta \circ \tilde{g}$ has the representation

$$\eta \circ \tilde{g}(\mathbf{x}) = \sum_{\substack{i_1,\ldots,i_t \geqslant 0 \\ i_1+\cdots+i_t\leqslant t\ell}} \alpha_{i_1,\ldots,i_t} x_1^{i_1} \ldots x_t^{i_t},$$

then

$$\sup_{i_1+\cdots+i_t\leqslant t\ell} \left(\frac{T}{(Wq)^{t-1}}\right)^{\frac{i_1+\cdots+i_t}{t}} \|\alpha_{i_1,\ldots,i_t}\| \ll \delta^{-O_{\ell,m,t}(B)},$$

or, in other words,

$$\|\alpha_{i_1,\ldots,i_t}\| \ll \delta^{-O_{\ell,m,t}(B)} (Wq)^{(i_1+\cdots+i_t)\frac{t-1}{t}} T^{-(i_1+\cdots+i_t)/t} \tag{4.2}$$

holds uniformly for all admissible tuples $(i_1, \ldots, i_t)$. Since $W(N)q \ll_\varepsilon T^\varepsilon$ and $\delta^{-1}(x) \ll_\varepsilon x^\varepsilon$, we in fact have the following 'graded' bounds in terms of the value $j = (i_1 + \cdots + i_t)/t$:

$$\|\alpha_{i_1,\ldots,i_t}\| \ll_\varepsilon T^{-\frac{i_1+\cdots+i_t}{t}+\varepsilon+o(1)}. \tag{4.3}$$

Let γ be, as before, the least common multiple of non-zero coefficients $\gamma^{(j)}_{i_1,\ldots,i_t}$. We aim to deduce from (4.2) and (4.3) similar bounds with a graded decay depending on j for the coefficients β_j. While it seems difficult to achieve this directly for the quantities $\|\beta_j\|$, we will obtain such bounds for certain related quantities $\|\beta_j q_j\|$, $1 \leqslant j \leqslant \ell$, where each q_j is a $w(N)$-smooth integer that is small compared to $S(N)$; it will, in fact, take the form

$$q_j := (Wq)^{(t-1)(\ell+(\ell-1)+\cdots+j)} \gamma^{1+\ell-j}.$$

Note that

$$\alpha_{i_1,\ldots,i_t} = \beta_d (Wq)^{(t-1)\ell} \gamma^{(\ell)}_{i_1,\ldots,i_t}$$

whenever $i_1 + \cdots + i_t = t\ell$. By (4.2), this immediately yields

$$\|\beta_\ell (Wq)^{(t-1)\ell} \gamma\| \ll \delta^{-O_{\ell,m,t}(1)} (Wq)^{(t-1)\ell} T^{-\ell}.$$

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

More generally, we have

$$\alpha_{i_1,\dots,i_t} = \beta_j (Wq)^{(t-1)j} \gamma^{(j)}_{i_1,\dots,i_t} + \sum_{k>j} \beta_k c^{(k)}_{i_1,\dots,i_t}$$

when $i_1 + \cdots + i_t = tj$. Multiplying through by q_{j+1}, this identity allows us to employ a downwards-inductive argument, taking advantage of the graded decay bounds that can be assumed inductively for $\|\beta_k q_k\|$ with $k > j$. Indeed, by applying (4.2) to $\alpha_{i_1,\dots,i_t}$, and the induction hypotheses to $\beta_k q_k$ for $k > j$, we deduce that

$$\|\beta_j q_j\| \ll \delta^{-O_{\ell,m,t}(B)} q_j T^{-j}.$$

It follows that

$$\sup_{1 \leqslant j \leqslant \ell} (T/q_\ell)^j \|q_\ell^j \beta_j\| \ll \delta^{-O_{\ell,m,t}(B)}.$$

If N and T are sufficiently large, then Lemma 4.4 and Theorem 4.3 imply that there is $C_1 \asymp_{\ell,m,t} 1$ such that $g(q_\ell n)_{n \leqslant T/q_\ell}$ fails to be totally $\delta(N)^{C_1 B}$-equidistributed. Setting $C = \max(C_1, 1)$, the result follows for every $E > C$ by choosing $B = E/C$. Indeed, if $E = BC > BC_1$, and if g satisfies all hypotheses from the statement, then, in particular, $g(q_\ell n)_{n \leqslant T/q_\ell}$ is totally $\delta(N)^E$-equidistributed. This is a contradiction, and it shows that

$$\big(\tilde{g}(\mathbf{x})\Gamma\big)_{\mathbf{x} \in \left[\left(\frac{T}{(Wq)^{t-1}}\right)^{1/t}\right]^t}$$

is in fact totally $\delta(N)^{E/C}$-equidistributed. □

The next lemma is in spirit closely related to the previous one. It shows that the assumptions that the previous lemma makes on the polynomial sequence g imply that these assumptions, with a slightly different constant E, are also met by any sequence of the form $g \circ L$ for certain linear polynomials L. This result will allow us to replace g by $g \circ L$ in the conclusion of Lemma 4.6, and thus to easily deal with a necessary restriction to subprogressions in § 5. We note aside that this result generalizes to higher-degree polynomials.

Lemma 4.7. *Let N and T be positive integers, and suppose that $T = T(N)$ satisfies $N^{1-\varepsilon} \ll_\varepsilon T \leqslant N$ for all $\varepsilon > 0$. Let $\delta : \mathbb{N} \to (0, 1/2)$ be a function that satisfies $\delta(x)^{-1} \ll_\varepsilon x^\varepsilon$ for all $\varepsilon > 0$. Let m_G and ℓ be positive integers, and suppose that G/Γ is an m_G-dimensional nilmanifold equipped with a $\frac{1}{\delta(N)}$-rational Mal'cev basis $\mathscr{X}$ adapted to some filtration $G_\bullet$ of degree ℓ. Let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ be any polynomial sequence, and let $S : \mathbb{N} \to \mathbb{N}$ be a function such that $S(x) \ll_\varepsilon x^\varepsilon$ for all $\varepsilon > 0$.*

Then there is a constant $1 \leqslant C \ll_{m_G,\ell} 1$ such that the following holds. Let $E > C$, and suppose that for every $w(N)$-smooth integer $\tilde{q} \leqslant S(N)$ the sequence $(g(\tilde{q}n)\Gamma)_{n \leqslant N/\tilde{q}}$ is totally $\delta(N)^E$-equidistributed. Let $L(m) = am + b$ be a linear polynomial with $0 \leqslant b < a$ and a $w(N)$-smooth leading constant a, and let q be a $w(N)$-smooth integer such that $qa \leqslant S(N)^{1/\ell^{\ell+1}}$. Then the finite sequence

$$(g(aqm + b))\Gamma)_{m \leqslant T/(aq)}$$

is totally $\delta(N)^{E/C}$-equidistributed.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

Proof. Assuming that all conditions of [10, Proposition 15.4] are satisfied, this result, applied with $P = L$, will guarantee the existence of a $w(N)$-smooth integer $\tilde{q}$ and a constant $C = O_{m_G,\ell}(1)$ such that for every $0 \leqslant r < \tilde{q}$ the sequence

$$(g(aq(\tilde{q}m + r) + b))\Gamma)_{m \leqslant T/(aq\tilde{q})}$$

is totally $\delta(N)^{E/C}$-equidistributed. This, however, implies that the sequence

$$(g(aqm + b))\Gamma)_{m \leqslant T/(aq)}$$

itself is totally $\delta(N)^{E/C}$-equidistributed, which will prove the lemma.

It remains to check that all conditions are satisfied. The integer $\tilde{q}$ produced by [10, Proposition 15.4] comes from an application of [10, Proposition 15.2]. The proof of the latter proposition reveals that we can take $\tilde{q} = (aq)^{C'}$ for some positive integer $C' = O_\ell(1)$. It is moreover possible to read of an explicit upper bound of the form $C' \leqslant \ell^{\ell+1}$; cf. the lines before [10, equation (15.4)] where t is introduced, and note that C' corresponds to the quantity td. In order for the proof of [10, Proposition 15.4] including its application of [10, Proposition 15.2] to work in the setting of the current lemma, it suffices to know that for every $w(N)$-smooth integer $q' \leqslant (aq)^{C'}$ the sequence $(g(q'm)\Gamma)_{m \leqslant T/q'}$ is totally $\delta(N)^E$-equidistributed. This, however, is guaranteed by the assumption that $aq \leqslant S(N)^{1/\ell^{\ell+1}}$. □

The following lemma will be used to carry out an inclusion–exclusion argument that allows us to reduce estimates involving R_S^* to estimates involving R.

Lemma 4.8. *Let N and T be positive integers, and suppose that $T = T(N)$ satisfies $N^{1-\varepsilon} \ll_\varepsilon T \leqslant N$ for all $\varepsilon > 0$. Suppose that $\delta : \mathbb{N} \to (0, 1/2)$ satisfies*

$$\delta(x)^{-1} \asymp (\log w(x))^C$$

for some positive constant C. Let m, ℓ, and t be positive integers, and suppose that G/Γ is an m-dimensional nilmanifold equipped with a $\frac{1}{\delta(N)}$-rational Mal'cev basis $\mathscr{X}$ that is adapted to some filtration $G_\bullet$ of degree ℓ. Further, let $g \in \mathrm{poly}(\mathbb{Z}^t, G_\bullet)$ be a polynomial sequence. Given any integer $d \geqslant 1$, let $\mathbf{x}_d \in \{0, \ldots, d^2\}^t$ be a fixed vector.

Then there are constants $C_0 > 2t$ and $E_0 > 1$, both of order $O_{m,\ell,t}(1)$, such that, provided that $N \gg_{m,\ell,t} 1$ is sufficiently large, the following holds for every $E > E_0$.

Suppose that $(g(\mathbf{n})\Gamma)_{\mathbf{n} \in [T^{1/t}]^t}$ is totally $\delta(N)^E$-equidistributed. Then, for every integer K such that $1 < K < T^{1/C_0}$, all but $o(\delta(N)\frac{K}{\log w(N)})$ of the sequences

$$(g(d^2\mathbf{n} + \mathbf{x}_d)\Gamma)_{\mathbf{n} \in [T^{1/t}d^{-2}]^t}$$

for $d \in \{n \in [K, 2K) : \gcd(n, W(N)) = 1\}$ are totally $\delta(N)^{E/E_0}$-equidistributed.

Proof. Let $W = W(N)$, and recall that $\prod_{p<w(N)}(1 - p^{-1}) \asymp \frac{1}{\log w(N)}$. Let us write

$$g_d(\mathbf{x}) = g(d^2\mathbf{x} + \mathbf{x}_d).$$

Let $B > 1$, and suppose $B = E/E_0$, with E_0 to be defined at the end of the proof. Suppose further that there is some K, $1 < K < N^{1/C_0}$, such that $\gg \delta(N)\frac{K}{\log w(N)}$ of the integers

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

$d \in [K, 2K)$ with $\gcd(d, W(N)) = 1$ are exceptional. In each of these cases, Lemma 4.4 and Theorem 4.3 imply that there is a non-trivial horizontal character $\eta_d : G/\Gamma \to \mathbb{C}$ of modulus $|\eta_d| \ll \delta(N)^{-O_{m,\ell,t}(B)}$ such that

$$\|\eta_d \circ g_d\|_{C^\infty_*[T^{1/t}d^{-2}]^t} \ll \delta(N)^{-O_{m,\ell,t}(B)}. \quad (4.4)$$

By the pigeonhole principle, we find some η such that $\eta_d = \eta$ for at least

$$\gg \delta(N)^{O_{m,\ell,t}(B)} \frac{K}{\log w(N)}$$

of the exceptional values of d. Suppose that $\eta \circ g$ has the representation

$$\eta \circ g(\mathbf{x}) = \sum_{\substack{i_1,\dots,i_t \geqslant 0 \\ i_1+\dots+i_t \leqslant \ell}} \beta_{i_1,\dots,i_t} x_1^{i_1} \dots x_t^{i_t}.$$

Writing

$$\eta \circ g_d(\mathbf{x}) = \eta \circ g(d^2\mathbf{x} + \mathbf{x}_d) = \sum_{\substack{i_1,\dots,i_t \geqslant 0 \\ i_1+\dots+i_t \leqslant \ell}} \alpha^{(d)}_{i_1,\dots,i_t} x_1^{i_1} \dots x_t^{i_t},$$

the bound (4.4) translates to

$$\sup_{(i_1,\dots,i_t)\neq \mathbf{0}} \left(\frac{T^{1/t}}{d^2}\right)^{i_1+\dots+i_t} \|\alpha^{(d)}_{i_1,\dots,i_t}\| \ll \delta^{-O_{m,\ell,t}(B)}. \quad (4.5)$$

Note that every coefficient $\alpha^{(d)}_{i_1,\dots,i_t}$ can be expressed in terms of coefficients $\beta_{j_1,\dots,j_t}$ as follows:

$$\alpha^{(d)}_{i_1,\dots,i_t} = d^{2(i_1+\dots+i_t)} \beta_{i_1,\dots,i_t} + \sum_{j_1+\dots+j_t > i_1+\dots+i_t} c_{\mathbf{i},\mathbf{j}} \beta_{j_1,\dots,j_t},$$

with $\mathbf{i} = (i_1, \dots, i_t)$, $\mathbf{j} = (j_1, \dots, j_t)$ and integral coefficients $c_{\mathbf{i},\mathbf{j}}$ of order $O(d^{2(i_1+\dots+i_t)})$. As in the proof of Lemma 4.6, these identities allow us to deduce downward-inductively information on the coefficients $\beta_{j_1,\dots,j_t}$ from (4.5).

If $i_1 + \dots + i_t = \ell$, then we immediately have

$$\|\beta_{i_1,\dots,i_t} d^{2\ell}\| \ll \delta(N)^{-O_{m,\ell,t}(B)} d^{2\ell} T^{-\ell/t}.$$

In general, we obtain

$$\|\beta_{i_1,\dots,i_t} d^{2(\ell+(\ell-1)+\dots+(i_1+\dots+i_t))}\| \ll \delta(N)^{-O_{m,\ell,t}(B)} d^{2(\ell+(\ell-1)+\dots+(i_1+\dots+i_t))} T^{-(i_1+\dots+i_t)/t}.$$

Thus,

$$\|\beta_{i_1,\dots,i_t} d^k\| \ll \delta(N)^{-O_{m,\ell,t}(B)} d^k T^{-(i_1+\dots+i_t)/t},$$

for $k = \ell^2$. The above bound holds for $\gg \delta(N)^{O_{m,\ell,t}(B)} \frac{K}{\log w(N)} \gg \delta(N)^{O_{m,\ell,t}(B)} K$ values of $d \in [K, 2K)$. Employing the Waring-type result given in [5, Lemma 3.3], we deduce that there are at least $\gg_\ell \delta(N)^{O_{\ell,t,m}(B)} K^k$ positive integers $n \leqslant 10^k K^k$ such that

$$\|\beta_{i_1,\dots,i_t} n\| \ll \delta(N)^{-O_{m,\ell,t}(B)} K^k T^{-(i_1+\dots+i_t)/t}.$$

Setting $C_0 = 4tk$, so that $K^k \ll T^{1/2t}$, we deduce from the strong recurrence lemma recorded in [5, Lemma 3.4] that for each $\beta_{i_1,\dots,i_t}$ there is a non-zero integer

$$q_{i_1,\dots,i_t} \ll \delta(N)^{-O_{m,\ell,t}(B)}$$

such that

$$\|q_{i_1,\dots,i_t}\beta_{i_1,\dots,i_t}\| \ll \delta(N)^{-O_{m,\ell,t}(B)}T^{-(i_1+\dots+i_t)/t}.$$

Let q be the least common multiple of all the $q_{i_1,\dots,i_t}$. Then $q\eta$ is a non-trivial horizontal character of modulus $|q\eta| \ll \delta(N)^{-O_{m,\ell,t}(B)}$ with the property that

$$\|q\eta \circ g\|_{C_*^\infty[T^{1/t}]^t} \ll \delta(N)^{-O_{m,\ell,t}(B)}.$$

Recall that $B = E/E_0$. Choosing E_0 sufficiently large with respect to m, ℓ, and t, we deduce from Theorem 4.3 and Lemma 4.4 that $(g(\mathbf{n})\Gamma)_{\mathbf{n}\in[T^{1/t}]^t}$ fails to be totally $\delta(N)^E$-equidistributed, which is a contradiction. □

The final lemma of this section states the following. Given a multiparameter sequence, then we obtain a natural collection of one-parameter sequences by fixing all but one of the parameters. The lemma shows that if the multiparameter sequence is equidistributed then so are almost all of the one-parameter sequences from this collection.

Lemma 4.9. *Let m, t, ℓ, N, and T be positive integers. Suppose that $N^{1-\varepsilon} \ll_\varepsilon T \leqslant N$ for all $\varepsilon > 0$, and let $\delta : \mathbb{N} \to (0, 1/2)$ be such that $\delta(x)^{-1} \ll_\varepsilon x^\varepsilon$ for all $\varepsilon > 0$. Let G/Γ be an m-dimensional nilmanifold together with a $\frac{1}{\delta(N)}$-rational Mal'cev basis adapted to some filtration $G_\bullet$ of degree ℓ. Suppose that $g \in \mathrm{poly}(\mathbb{Z}^t, G_\bullet)$. Any fixed choice of integers $a_1, \dots, a_{t-1}$ gives rise to an element $g_{a_1,\dots,a_{t-1}}$ of $\mathrm{poly}(\mathbb{Z}, G_\bullet)$ by setting $g_{a_1,\dots,a_{t-1}}(n) = g(a_1, \dots, a_{t-1}, n)$. Then there is a constant $1 \leqslant C \ll_{m,t,\ell} 1$ such that the following holds for all $E > C$, provided that T is sufficiently large.*

If $(g(\mathbf{n})\Gamma)_{n\in[T^{1/t}]^t}$ is $\delta(N)^E$-equidistributed, then

$$(g_{a_1,\dots,a_{t-1}}(n)\Gamma)_{n\leqslant T^{1/t}}$$

is totally $\delta^{E/C}(N)$-equidistributed for all but $o(\delta(N)^{O_{m,t,\ell}(E/C)}T^{\frac{t-1}{t}})$ choices of

$$1 \leqslant a_1, \dots, a_{t-1} \leqslant T^{1/t}.$$

Proof. Let $B > 1$ denote the ratio $B = E/C$, with C to be determined at the end of the proof. Suppose there are $\gg \delta(N)^{O_{m,t,\ell}(B)}T^{(t-1)/t}$ tuples $(a_1, \dots, a_{t-1}) \in [1, T^{1/t}]^{t-1}$ for which $(g_{a_1,\dots,a_{t-1}}(n)\Gamma)_{n\leqslant T^{1/t}}$ fails to be totally $\delta(N)^B$-equidistributed.

Applying Lemma 4.4 and Theorem 4.3, we find non-trivial horizontal characters $\eta_{a_1,\dots,a_{t-1}}$ of modulus $\ll \delta(N)^{-O_{m,t,\ell}(B)}$ such that

$$\|\eta_{a_1,\dots,a_{t-1}} \circ g_{a_1,\dots,a_{t-1}}\|_{C_*^\infty[T^{1/t}]} \ll \delta(N)^{-O_{m,\ell,t}(B)}. \quad (4.6)$$

By the pigeonhole principle, there is some character η such that $\eta = \eta_{a_1,\dots,a_{t-1}}$ for at least $\gg \delta(N)^{O_{m,\ell,t}(B)}T^{(t-1)/t}$ of the exceptional tuples $(a_1, \dots, a_{t-1})$. We continue to only consider this subset of exceptional $(t-1)$-tuples.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

Suppose that

$$\eta \circ g(\mathbf{n}) = \sum_{i_1,\dots,i_t} \gamma_{i_1,\dots,i_t} n_1^{i_1} \dots n_t^{i_t}.$$

Then

$$\eta \circ g_{a_1,\dots,a_{t-1}}(n) = \sum_{j=1}^{\ell} n^j \sum_{\substack{i_1+\dots+i_t \\ \leqslant \ell t - j}} a_1^{i_1} \dots a_{t-1}^{i_{t-1}} \gamma_{i_1,\dots,i_{t-1},j},$$

and, for any of the exceptional tuples from above, (4.6) translates to

$$\sup_{1\leqslant j\leqslant \ell} T^{j/t} \left\| \sum_{\substack{i_1+\dots+i_t \\ \leqslant \ell t - j}} a_1^{i_1} \dots a_{t-1}^{i_{t-1}} \gamma_{i_1,\dots,i_{t-1},j} \right\| \ll \delta(N)^{-O_{m,t,\ell}(B)}. \tag{4.7}$$

Each of the coefficients of $\eta \circ g_{a_1,\dots,a_{t-1}}(n)$ can be regarded as a polynomial in $t-1$ variables that is evaluated at the point $(a_1, \dots, a_{t-1})$. These polynomials take the form

$$P_j(n_1, \dots, n_{t-1}) = \sum_{\substack{i_1+\dots+i_t \\ \leqslant \ell t - j}} n_1^{i_1} \dots n_{t-1}^{i_{t-1}} \gamma_{i_1,\dots,i_{t-1},j}.$$

The bounds (4.7) show that to each of these polynomials [6, Proposition 2.2] applies with $\varepsilon = \delta(N)^{-O_{m,t,\ell}(B)} T^{-j/t}$. Thus there is $Q_j \ll \delta(N)^{-O_{m,t,\ell}(B)}$ such that

$$\|Q_j P_j\|_{C^\infty_*[T^{1/t}]^{t-1}} = \sup_{1\leqslant i_1+\dots+i_{t-1}\leqslant \ell t - j} T^{(i_1+\dots+i_{t-1})/t} \|Q_j \gamma_{i_1,\dots,i_{t-1},j}\| \ll \delta(N)^{-O_{m,t,\ell}(B)} T^{-j/t},$$

and hence

$$\sup_{1\leqslant i_1+\dots+i_{t-1}\leqslant \ell t - j} T^{(i_1+\dots+i_{t-1}+j)/t} \|Q_j \gamma_{i_1,\dots,i_{t-1},j}\| \ll \delta(N)^{-O_{m,t,\ell}(B)}.$$

Since $\delta(x)^{-1} \ll_\varepsilon x^\varepsilon$ and hence $\|Q_j \gamma_{i_1,\dots,i_{t-1},j}\| = o(1)$, we can introduce a factor $Q_1 \dots Q_\ell / Q_j$ into the latter expression and deduce that

$$\sup_{1\leqslant i_1+\dots+i_t\leqslant \ell t} T^{(i_1+\dots+i_t)/t} \|Q_1 \dots Q_\ell \gamma_{i_1,\dots,i_t}\| \ll \delta(N)^{-O_{m,t,\ell}(B)},$$

provided that T and N are sufficiently large. For $B = E/C$ with $C > 1$ sufficiently large depending only on m, t, and ℓ, the latter bound implies in view of Theorem 4.3 that $(g(\mathbf{n})\Gamma)_{\mathbf{n}\in[T^{1/t}]^t}$ fails to be $\delta(N)^E$-equidistributed. This contradicts our assumptions, and completes the proof. □

5. Proof of Theorem 1.4

The general strategy for proving results like Theorem 1.4 is to deduce them from the special case in which the nilsequence involved is equidistributed. This strategy was introduced in [5, § 2]. The transition from the equidistributed statement to the general one is achieved through an application of the factorization theorem for nilsequences [4, Theorem 1.19], or a consequence thereof. For technical reasons we require a

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

factorization result of the form given in [12, Theorem 3], which is a slight generalization of [10, Proposition 16.4] and arises by iterative application of [4, Theorem 1.19]. Combining this factorization theorem with the major arc estimate given in Lemma 3.2, we will deduce Theorem 1.4 from the following adaptation of [1, Proposition 6.4] to the square-free version of R.

Proposition 5.1. *Let N and $T = T(N)$ be positive integers such that $N^{1-\varepsilon} \ll_\varepsilon T \leqslant N$ for all $\varepsilon > 0$. Let $\epsilon \in \{\pm\}$, and let $W = W(N)$. Let S be a finite set of primes, all bounded by $w(N)$, and let A be an integer such that $A \pmod W \in \mathscr{A}(R^*_S, N)$ and $0 \leqslant \epsilon A < W$. Suppose that $\delta : \mathbb{Z} \to (0, 1/2)$ satisfies $\delta^{-1}(x) \asymp (\log w(x))^C$ for some positive constant C. Assume that $(G/\Gamma, d_{\mathscr{X}})$ is an m_G-dimensional nilmanifold with a filtration $G_\bullet$ of degree ℓ adapted to a $\frac{1}{\delta(N)}$-rational Mal'cev basis. Let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$. Let $S : \mathbb{N} \to \mathbb{N}$ be such that $S(x) \ll_\varepsilon x^\varepsilon$ for all $\varepsilon > 0$. Finally, let $E > 0$, and suppose that for every $w(N)$-smooth number $\tilde{q} \leqslant S(N)$ the finite sequence $(g(\tilde{q}m)\Gamma)_{0<m\leqslant T/\tilde{q}}$ is totally $\delta(N)^E$-equidistributed in G/Γ.*

Then the following holds. There exists a constant $E_0 > 1$, only depending on m_G, ℓ and $n := [K : \mathbb{Q}]$, such that, for every Lipschitz function $F : G/\Gamma \to [-1, 1]$ of mean value $\int_{G/\Gamma} F = 0$, for every $w(N)$-smooth number $q > 0$ such that $(Wq)^{n\ell^2} \leqslant S(N)$, and for every $0 \leqslant b < q$, we have

$$\left|\frac{1}{T}\sum_{0<\epsilon m\leqslant T} R^*_S(W(qm+b)+A)F(g(|m|)\Gamma)\right|$$
$$\ll_{m_G,\ell,n} \delta(N)^{E/E_0}(1+\|F\|_{\mathrm{Lip}})\frac{\varrho(W,A)}{W^{n-1}}\prod_{p>w(N)}\left(1-\frac{\varrho(p^2,0)}{p^2}\right)$$

as $N \to \infty$.

Proof of Theorem 1.4 assuming Proposition 5.1. We follow the strategy from [5, § 2]. To prove Theorem 1.4 we may restrict attention to the case where $Q \leqslant \log\log\log N$, as the statement is trivially true otherwise. This allows us to apply [12, Theorem 3] with the following parameters (distinguished by a tilde from already existing ones):

$$\tilde{N} = N, \quad \tilde{T} = T/W(N), \quad \tilde{k} = w(N) = \log\log N, \quad \tilde{Q}_0 = \log\tilde{k}, \quad \tilde{R} = W(N),$$

and, finally, $\tilde{B} = E$ and $\tilde{E} > 2n\ell^2\ell^{\ell+1}$. Observe that $W(N) \ll (\log N)^{C_1 \log\log N}$. Hence, $\tilde{R}$ satisfies the required condition that $\tilde{R}(N)^t \ll_t N$ for all $t > 0$ as $N \to \infty$. By [12, Theorem 3], we therefore obtain an integer Q' such that

$$\log\log\log N \leqslant Q' \ll (\log\log\log N)^{O_{m_G,\ell,E}(1)},$$

and a partition $\mathscr{P}$ of the set $\{1, \ldots, \tilde{T}\}$ into $\ll W^{O_{\ell,m_G,n,E}(1)}$ disjoint subprogressions, each of some $w(N)$-common difference $q(P) \ll W^{O_{\ell,m_G,n,E}(1)}$ and length $\frac{T}{Wq(P)} + O(1)$. Along each progression $P = \{a < n < b : n \equiv r \pmod q\}$ from $\mathscr{P}$, the polynomial sequence g has a factorization

$$g(qm + r) = \varepsilon_P(m) g'_P(m) \gamma_P(m),$$

with the following properties.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

(1) $\varepsilon_P : \mathbb{Z} \to G$ is $(Q', \frac{T}{Wq})$-smooth[1].

(2) $g'_P : \mathbb{Z} \to G'$ takes values in G', and for each $w(N)$-smooth number $\tilde{q} < (qq_{\gamma_P} W)^{\tilde{E}}$ the finite sequence $(g'_P(\tilde{q}m)\Gamma')_{m \leqslant T/(Wq\tilde{q})}$ is totally Q'^{-E}-equidistributed in $G'\Gamma/\Gamma$.

(3) $\gamma_P : \mathbb{Z} \to G$ is a product $\gamma_P(m) = \gamma_1(m) \dots \gamma_t(m)$ of length at most m_G of Q-rational sequences γ_i. It gives rise to a periodic sequence $(\gamma_P(m)\Gamma)_{m\in\mathbb{Z}}$ with $w(N)$-smooth period $q_{\gamma_P} \leqslant Q'$.

Any progression $P \in \mathscr{P}$ as above can be split into $q_{\gamma_P} \leqslant Q'$ subprogressions on which γ_P is constant. Next, we split each of these subprogressions into pieces of diameter between $Q'^{-c^*E}T/W$ and $2Q'^{-c^*E}T/W$ for a small parameter $c^* \in (0, 1)$ which will be determined later. Let $\mathscr{P}^*$ denote the collection of the resulting bounded diameter pieces of all progressions $P \in \mathscr{P}$, and note that each $P' \in \mathscr{P}^*$ is of the form

$$P' = \left\{ qq_\gamma m + r' : \frac{\delta_1 T}{Wqq_\gamma} < m \leqslant \frac{\delta_2 T}{Wqq_\gamma} \right\}, \tag{5.1}$$

for a $w(T)$-smooth period q as before, for some $0 \leqslant r' < qq_\gamma$, and for $\delta_1, \delta_2 \in [0, 1]$ such that $Q'^{-c^*E} \leqslant \delta_2 - \delta_1 \leqslant 2Q'^{-c^*E}$ and either $\delta_1 = 0$ or $\delta_1 > Q'^{-c^*E}$. For each progression $P' \in \mathscr{P}^*$, let $s_{P'}$ denote its smallest element. If $F : G/\Gamma \to \mathbb{C}$ is a Lipschitz function, then the right invariance of the metric $d_{\mathscr{X}}$, defined in [4, Definition 2.2], implies for any m, m' that belong to the same element of $\mathscr{P}^*$ that

$$\begin{aligned} &|F(\varepsilon(m)g'(m)\gamma(m)\Gamma) - F(\varepsilon(m')g'(m)\gamma(m)\Gamma)| \\ &\quad\leqslant \|F\|_{\mathrm{Lip}}\, d_{\mathscr{X}}(\varepsilon(m)g'(m)\gamma(m), \varepsilon(m')g'(m)\gamma(m)) \\ &\quad= \|F\|_{\mathrm{Lip}}\, d_{\mathscr{X}}(\varepsilon(m), \varepsilon(m')) \\ &\quad\leqslant \|F\|_{\mathrm{Lip}}\, |m - m'|Q'W/T \\ &\quad\ll \|F\|_{\mathrm{Lip}}\, Q'^{1-c^*E} \end{aligned}$$

if ε, γ, and g' satisfy the respective conditions (1)–(3) above. This estimate allows one to fix for any $P' \in \mathscr{P}^*$ the contribution of ε. To see this, suppose that P' arises from the progression $P \in \mathscr{P}$, and define

$$\mu_{R^*} := \kappa^{\epsilon} \frac{\varrho(W, A)}{W^{n-1}} \prod_{p > w(N)} \left(1 - \frac{\varrho(p^2, 0)}{p^{2n}}\right).$$

By combining the previous estimate with Lemmas 3.1 and 3.2, we deduce that

$$\begin{aligned} &\sum_{m\in P'} (R^*_S(Wm + A) - \mu_{R^*})F(g(m)\Gamma) \\ &\quad= \sum_{m\in P'} (R^*_S(Wm + A) - \mu_{R^*})F(\varepsilon_P(s_{P'})g'_P(m)\gamma_P(m)\Gamma) \\ &\qquad+ O\big(\|F\|_{\mathrm{Lip}} Q'^{1-c^*E}\big(\mu_{R^*}\#P' + W^{O_{l,m_G,n,E}(1)} T^{1-\frac{1}{20n}}\big)\big). \end{aligned}$$

In view of Theorem 1.4, the error term above is still acceptable when summed over all $P' \in \mathscr{P}*$, which allows us to exclude it from further observation.

[1]The notion of smoothness was defined in [4, Definition 1.18]. A sequence $(\varepsilon(n))_{n\in\mathbb{Z}}$ is said to be (M, N)-*smooth* if both $d_{\mathscr{X}}(\varepsilon(n), \mathrm{id}_G) \leqslant \mathrm{N}$ and $d_{\mathscr{X}}(\varepsilon(n), \varepsilon(n-1)) \leqslant M/N$ hold for all $1 \leqslant n \leqslant N$.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

The remaining sequence $m \mapsto F(\varepsilon(s_{P'})g'(m)\gamma(s_{P'})\Gamma)$ can be reinterpreted as an equidistributed nilsequence, as is shown in [5, *Claim* in § 2]. Indeed, setting

$$\begin{aligned} H_{P'} &:= \gamma_P(s_{P'})^{-1} G' \gamma_P(s_{P'}), \\ (H_{P'})_\bullet &:= \gamma_P(s_{P'})^{-1} G'_\bullet \gamma_P(s_{P'}), \\ \Lambda_{P'} &:= \Gamma \cap H_{P'}, \\ g_{P'}(m) &:= \gamma_P(s_{P'})^{-1} g'_P(m) \gamma_P(s_{P'}) \end{aligned}$$

and

$$F_{P'} : H_{P'}/\Lambda_{P'} \to [-1, 1], \quad F_{P'}(x\Lambda_{P'}) := F(\varepsilon_P(s_{P'})\gamma_P(s_{P'})x\Gamma),$$

the *Claim* guarantees the existence of a Mal'cev basis for $H_{P'}/\Lambda_{P'}$ adapted to the filtration $(H_{P'})_\bullet$ that is $Q'^{O(1)}$-rational in terms of the basis $\mathscr{X}$. This basis induces a metric on $H_{P'}/\Lambda_{P'}$ with respect to which we have on the one hand that $\|F_{P'}\|_{\mathrm{Lip}} \leqslant Q'^{O(1)}$ and on the other hand that each of the sequences $(g_{P'}(\tilde{q}m))_{m \leqslant T/(Wq(P)\tilde{q})}$ for $w(N)$-smooth $\tilde{q} < (q(P)q_{\gamma_P}W)^{2n\ell^2\ell^{\ell+1}}$ is totally $Q'^{-cE+O(1)}$-equidistributed for some $c > 0$ depending only on m_G and d.

We aim to apply Proposition 5.1 making use of the equidistribution properties of $g_{P'}$. To prepare this application, first note that Lemma 4.7 implies that every sequence

$$g_{P',r}(m) := g_{P'}(q_{\gamma_P}m + r), \quad (0 \leqslant r < q_{\gamma_P}),$$

has the property that for every $w(N)$-smooth integer $\tilde{q} < (q(P)q_{\gamma_P}W)^{2n\ell^2}/q_{\gamma_P}$ the sequence $(g_{P',r}(\tilde{q}m))_{m \leqslant T/(Wq(P)q_{\gamma_P}\tilde{q})}$ is totally $Q'^{-c'E+O(1)}$-equidistributed for some $c' > 0$ depending only on m_G and d. Thus, we may set $S_{P'}(N) := (q(P)q_{\gamma_P}W)^{n\ell^2}$ for any $P' \subset P \in \mathscr{P}$. This quantity will play the role of $S(N)$ from Proposition 5.1. It will, in particular, allow us to take $q = q(P)q_{\gamma_P}$ in the application below.

Next, we need to ensure that the mean value of the Lipschitz function Proposition 5.1 will be applied with vanishes. In this regard, note that Lemma 3.1 implies that

$$\sum_{m \in P'} (R_S^*(Wm + A) - \mu_{R^*}) \ll q(P)q_\gamma W(T/W)^{1-\frac{1}{20n}} \ll_\varepsilon T^{1-\frac{1}{20n}+\varepsilon}.$$

This allows us to replace $F_{P'}$ by $(F_{P'} - \int_{H_{P'}/\Lambda_{P'}} F_{P'})$, since

$$\begin{aligned} &\sum_{m \in P'} (R_S^*(Wm + A) - \mu_{R^*}) F(\varepsilon_P(m_{P'})g'_P(m)\gamma_P(m)\Gamma) \\ &= \sum_{\frac{\delta_1 T}{Wqq_\gamma} < m \leqslant \frac{\delta_2 T}{Wqq_\gamma}} (R_S^*(W(q(P)q_\gamma m + r') + A) - \mu_{R^*}) \left(F_{P'}(g_{P',r}(m)\Lambda_{P'}) - \int_{H_{P'}/\Lambda_{P'}} F_{P'} \right) \\ &\quad + O_\varepsilon\big(T^{1-\frac{1}{20n}+\varepsilon}\big). \end{aligned}$$

Here, we made use of the notation from (5.1). Note that the error term is negligible even when summed over all $P' \in \mathscr{P}^*$, and it can be ignored.

We are now ready to apply Proposition 5.1 for each $P' \in \mathscr{P}^*$ to the sum on the left-hand side above with $\delta(N) = (\log\log\log N)^{-1}$ and with $\delta(N)^E$ replaced by $Q'^{-c'E+O(1)}$. Since Proposition 5.1 cannot be applied to the short progression P' directly, we apply it once with T replaced by $\delta_2 T/(Wqq_\gamma)$ and once with T replaced by $\delta_1 T/(Wqq_\gamma)$. Resulting

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

from this double counting, we pick up a factor of $O(Q'^{c^*E})$ along each progression P from the original decomposition $\mathscr{P}$. In total, we obtain

$$\sum_{P\in\mathscr{P}}\sum_{\substack{P'\in\mathscr{P}^*\\P'\subset P}}\sum_{n\in P'}(R_S^*(Wm+A)-\mu_{R^*})\left(F_{P'}(g'(m)\Lambda_{P'})-\int_{H_{P'}/\Lambda_{P'}}F_{P'}\right)$$
$$\ll_{m_G,\ell,n}\left(\sum_{P\in\mathscr{P}}Q'^{c^*E}\#P\right)Q'^{-cE/E_0+O(1)}(1+\|F\|_{\mathrm{Lip}})\frac{\varrho(W,A)}{W^{n-1}}\prod_{p>w(T)}\left(1-\frac{\varrho(p^2,0)}{p^2}\right)$$
$$\ll_{m_G,\ell,n}\frac{T}{W}Q'^{c^*E-cE/E_0+O(1)}(1+\|F\|_{\mathrm{Lip}})\frac{\varrho(W,A)}{W^{n-1}}\prod_{p>w(T)}\left(1-\frac{\varrho(p^2,0)}{p^2}\right).$$

Choosing $c^*<c/(2E_0)$ and recalling that $Q'\geqslant\log\log\log N$, this implies the result. □

Proof of Proposition 5.1. We observe first of all that (2.1) implies that

$$\prod_{p>w(N)}\left(1-\frac{\varrho(p^2,0)}{p^2}\right)\asymp 1,$$

provided that N is sufficiently large, so that the last factor in the bound can be ignored.

Our aim is to deduce this proposition from the proof of [1, Proposition 6.4] by using the decomposition $R_S^*(m)=\sum_{d^2|m,(d,W)=1}\mu(d)R(m)$, valid for all m with $m \pmod W\in\mathscr{A}(R_S^*,N)$, together with Lemmas 4.6, 4.8 and 4.9. Indeed, the decomposition yields

$$\frac{1}{T}\sum_{0<\epsilon m\leqslant T}R_S^*(W(qm+b)+A)F(g(|m|)\Gamma)$$
$$=\sum_{\substack{d\leqslant N^{1/2}\\\gcd(d,W)=1}}\frac{\mu(d)}{T}\sum_{\substack{0<\epsilon m\leqslant T\\W(qm+b)+A\equiv 0 \pmod{d^2}}}R(W(qm+b)+A)F(g(|m|)\Gamma)$$
$$=\sum_{\substack{d\leqslant N^{1/2}\\\gcd(d,W)=1}}\frac{\mu(d)}{T}\sum_{\substack{0<\epsilon m\leqslant B\\m\equiv A+Wb \pmod{Wq}\\m\equiv 0 \pmod{d^2}}}R(m)F\left(g\left(\frac{m-A-Wb}{\epsilon Wq}\right)\Gamma\right),\tag{5.2}$$

where $B:=Wq(T+b)+A\asymp WqT$. For fixed d, the inner sum will be estimated using the strategy from [1, Proposition 6.4]. Due to the additional restriction $m\equiv 0 \pmod{d^2}$ that appears here, we need to work with subsequences of the nilsequence that mattered in [1, Proposition 6.4]. The results from §4 provide the necessary information on equidistribution properties of the new sequences as d varies. Since all these results require d to be sufficiently small, we begin by restricting the summation in d.

The estimate (3.7) allows us to remove large values of d from consideration. In particular, it shows that the summation can be truncated at $d\leqslant N^{1/C_0}$ for any fixed $C_0>2$ at the expense of an error term of order $O(T^{-1}(WqT)^{1-1/O(C_0)})$. Since $Wq\ll_\varepsilon T^\varepsilon$, this error is $o(1)$. In order to obtain a direct sum over F, we aim to move the appearance of the R-function in (5.2) from the argument of the summation to the summation condition.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

For this purpose, let

$$\mathscr{Y} = \{\mathbf{y} \in (\mathbb{Z}/Wq\mathbb{Z})^n : \mathbf{N}_K(\mathbf{y}) \equiv A + Wb \ (\mathrm{mod}\ Wq)\},$$

so that $\#\mathscr{Y} = \varrho(Wq, A + Wb)$. Given $\mathbf{y} \in \mathscr{Y}$ and d with $\gcd(d, W) = 1$, we further set

$$\mathscr{X}_{\mathbf{y},d} = \{\tilde{\mathbf{x}}_d \in (\mathbb{Z}/d^2\mathbb{Z})^n : \mathbf{N}_K(Wq\tilde{\mathbf{x}}_d + \mathbf{y}) \equiv 0 \ (\mathrm{mod}\ d^2)\},$$

so that $\#\mathscr{X}_{\mathbf{y},d} = \varrho(d^2, 0)$. Recall that

$$R(m) = \mathbf{1}_{m\neq 0} \cdot \ \#\{\mathbf{x} \in \mathbb{Z}^n \cap \mathfrak{D}^+ : \mathbf{N}_K(\mathbf{x}) = m\}.$$

Thus (5.2) becomes

$$\sum_{\substack{d \leqslant N^{1/C_0} \\ \gcd(d,W)=1}} \frac{\mu(d)}{T''} \sum_{\mathbf{y}\in\mathscr{Y}} \sum_{\tilde{\mathbf{x}}_d \in \mathscr{X}_{\mathbf{y},d}} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n \\ Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y} \\ \in B^{1/n}\mathfrak{X}(1)}} F\left(g\left(\frac{\mathbf{N}_K(Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y}) - A - Wb}{\epsilon Wq}\right)\Gamma\right) + o(1), \tag{5.3}$$

where $\mathfrak{X}(1) = \{\mathbf{x} \in \mathfrak{D}^+ : 0 < \epsilon\,\mathbf{N}_K(\mathbf{x}) \leqslant 1\}$ is a compact set.

As in [1], we proceed with the analysis of the inner sum over $\mathbf{x}$ by fixing the first $n-1$ coordinates of $\mathbf{x}$. Since $\mathfrak{X}(1)$ is compact, we have $\mathfrak{X}(1) \subset (-\alpha, \alpha)^n$ for some positive constant $\alpha = O(1)$, which allows us to restrict the ranges of the first $n-1$ coordinates we consider: letting $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ denote the projection onto the coordinate plane $\{x_n = 0\}$, it thus suffices to consider the set $\{\mathbf{x} : \pi(\mathbf{x}) = \mathbf{a}\}$, where $\mathbf{a}$ runs over all points in

$$\mathbb{Z}^{n-1} \cap \left(\frac{B^{1/n}(-\alpha, \alpha)^{n-1} - Wq\pi(\tilde{\mathbf{x}}_d) - \mathbf{y}}{Wqd^2}\right).$$

Assuming that T is sufficiently large, the latter set can, in fact, be replaced by

$$\mathscr{Z}_d = \mathbb{Z}^{n-1} \cap \frac{B^{1/n}}{Wqd^2}(-2\alpha, 2\alpha)^{n-1}. \tag{5.4}$$

Returning to (5.3), we consider the argument of g. Since the coefficient of X_n^n in $\mathbf{N}_K(X_1, \ldots, X_n)$ is non-zero, we obtain an integral polynomial of degree n and leading coefficient $\epsilon N_{K/\mathbb{Q}}(\omega_n) d^{2n}(Wq)^{n-1}$ when fixing all but the nth variable in

$$\frac{\mathbf{N}_K(Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y}) - A - Wb}{\epsilon Wq}.$$

If $\pi(\mathbf{x}) = \mathbf{a}$, we let $P_{\mathbf{a};\mathbf{y};d}(x)$ denote this polynomial.

Our next step is to show that most of the sequences $g(P_{\mathbf{a};\mathbf{y};d}(x)\Gamma)_{|x|\ll B^{1/n}/Wqd^2}$ are equidistributed. The assumptions on g and Lemma 4.6 imply that the sequence

$$\left(g\left(\frac{\mathbf{N}_K(Wq\mathbf{x}+\mathbf{y}) - A - Wb}{\epsilon Wq}\right)\Gamma\right)_{\mathbf{x}\in[B^{1/n}/(Wq)]^n}$$

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

is totally $\delta(N)^{E/C'}$-equidistributed for some $1 \leqslant C' \ll_{m_G,\ell,n} 1$, provided that $E > C'$. Applying Lemma 4.8 to this sequence, we thus deduce that there is $1 \leqslant C'' \ll_{m_G,\ell,n} 1$ such that the sequence

$$\left(g\left(\frac{\mathbf{N}_K(Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y})-A-Wb}{\epsilon Wq}\right)\Gamma\right)_{\mathbf{x}\in[B^{1/n}/(Wqd^2)]^n}$$

is totally $\delta(N)^{E/C''}$-equidistributed for all but $o(\delta(N)^{O_{m_G,\ell,n}(E/C'')}K)$ integers d such that $d \in [K, 2K)$ and $\gcd(d, W) = 1$, and for all integers $K \in (1, N^{1/C_0})$, provided that $E > C''$. Finally, for all unexceptional values of d, Lemma 4.9 implies that there is $1 \leqslant C''' \ll_{m_G,\ell,n} 1$ such that, provided that $E > C'''$, the sequence

$$\left(g\left(P_{(a_1,\dots,a_{n-1});\mathbf{y};d}(x)\right)\Gamma\right)_{x\leqslant B^{1/n}/(Wqd^2)}$$

is totally $\delta(N)^{E/C'''}$-equidistributed for all but $o(\delta(N)^{O_{m_G,\ell,n}(E/C''')}(\frac{B^{1/n}}{Wqd^2})^{n-1})$ integer tuples $(a_1,\dots,a_{n-1})$ with $1 \leqslant a_1,\dots,a_{n-1} \leqslant B^{1/n}/(Wqd^2)$.

Before we exploit these equidistribution properties, we aim to bound the contribution of exceptional values of d and $\mathbf{a}$. Using the fact that $\|F\|_\infty \ll 1$, the contribution of exceptional values for d to the main term of (5.3) may trivially be bounded by

$$\begin{aligned}
&\frac{1}{T}\#\mathscr{Y}\sum_{k=0}^{\frac{\log N}{C_0\log 2}}\sum_{\substack{d\sim 2^k\\ \gcd(d,W)=1\\ d \text{ exceptional}}}\varrho(d^2,0)\sum_{\mathbf{a}\in\mathscr{Z}_d}\frac{B^{1/n}}{Wqd^2}\\
&\ll \frac{1}{T}\varrho(Wq, A+Wb)\sum_{k=0}^{\frac{\log N}{C_0\log 2}}\sum_{\substack{d\sim 2^k\\ \gcd(d,W)=1\\ d \text{ exceptional}}}\varrho(d^2,0)\frac{TWq}{(Wqd^2)^n}.
\end{aligned}$$

In view of (3.9), the above is further bounded by

$$\begin{aligned}
&\ll \frac{\varrho(W,A)}{W^{n-1}}\sum_{k=0}^{\frac{\log N}{C_0\log 2}}\sum_{\substack{d\sim 2^k\\ \gcd(d,W)=1\\ d \text{ exceptional}}}\frac{\varrho(d^2,0)}{d^2}\\
&\ll \frac{\varrho(W,A)}{W^{n-1}}\sum_{k=0}^{\frac{\log N}{C_0\log 2}}\delta(N)^{O_{m_G,\ell,n}(E/C'')}(2^{1-\frac{1}{2}})^{-k}\\
&\ll \frac{\varrho(W,A)}{W^{n-1}}\delta(N)^{O_{m_G,\ell,n}(E/C'')},
\end{aligned}$$

where we employed (3.5) with $\varepsilon = \frac{1}{2}$.

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

Recall the definition of $\mathscr{Z}_d$ from (5.4). Then the contribution from exceptional values of $\mathbf{a}$ may be bounded in a similar manner:

$$\begin{aligned}
&\frac{1}{T}\#\mathscr{Y}\sum_{\substack{d\leqslant N^{1/C_0}\\ \gcd(d,W)=1}}\varrho(d^2,0)\sum_{\substack{\mathbf{a}\in\mathscr{Z}_d\\ \mathbf{a}\text{ exceptional}}}\frac{B^{1/n}}{Wqd^2}\\
&\quad\ll\frac{1}{T}\varrho(Wq,A+Wb)\sum_{\substack{d\leqslant N^{1/C_0}\\ \gcd(d,W)=1}}\varrho(d^2,0)\delta(N)^{O_{m_G,\ell,n}(E/C''')}\frac{TqW}{(Wqd^2)^n}\\
&\quad\ll\delta(N)^{O_{m_G,\ell,n}(E/C''')}\frac{\varrho(W,A)}{W^{n-1}}\sum_{k=0}^{\frac{\log N}{C_0\log 2}}\sum_{\substack{d\leqslant N^{1/C_0}\\ \gcd(d,W)=1}}\frac{\varrho(d^2,0)}{d^2}\\
&\quad\ll\delta(N)^{O_{m_G,\ell,n}(E/C''')}\frac{\varrho(W,A)}{W^{n-1}}.
\end{aligned}$$

In the case of unexceptional values of d and $a_1,\dots,a_{n-1}$, we can finally proceed in exactly the same way as in the proof of [1, Proposition 6.4]. First of all, we recall from [1] how the lines $\{\mathbf{x}\in\mathbb{R}:\pi(\mathbf{x})=(a_1,\dots,a_{n-1})\}$ intersect the domain $\mathfrak{X}(1)\subset(-\alpha,\alpha)^n$: let $\mathbf{a}'=(a'_1,\dots,a'_{n-1})$, with $|\mathbf{a}'|<\alpha$, and consider the line $\ell_{\mathbf{a}'}:(-\alpha,\alpha)\to\mathbb{R}^n$ given by $\ell_{\mathbf{a}'}(x)=(\mathbf{a}',x)$. For $\varepsilon\geqslant 0$, let $\partial_\varepsilon\mathfrak{X}(1)\subset\mathbb{R}^n$ denote the set of points at distance at most ε to the boundary of the closure of $\mathfrak{X}(1)$. We note that the set

$$\{x\in(-\alpha,\alpha):\ell_{\mathbf{a}'}(x)\in\mathfrak{X}(1)\setminus\partial_0\mathfrak{X}(1)\}$$

is the union of disjoint open intervals. By removing all intervals of length at most ε, we obtain a collection of at most $2\alpha\varepsilon^{-1}\ll\varepsilon^{-1}$ open intervals $I_1(\mathbf{a}'),\dots,I_{k(\mathbf{a}')}(\mathbf{a}')\in(-\alpha,\alpha)$ such that any $x\in(-\alpha,\alpha)$ satisfies the implication

$$\ell_{\mathbf{a}'}(x)\in\mathfrak{X}(1)\setminus\partial_\varepsilon\mathfrak{X}(1)\implies x\in I_j(\mathbf{a}')\text{ for some }j\in\{1,\dots,k(\mathbf{a}')\}.$$

We will choose a suitable value of ε at the end of the proof.

Observe that any interval $(z_0,z_1)\subset(-\alpha,\alpha)$ can be expressed as a difference of intervals in $(-\alpha,\alpha)$ that have length at least $2\alpha/3$. Indeed, z_0 and z_1 partition $(-\alpha,\alpha)$ into three (possibly empty) intervals, at least one of which has length at least $2\alpha/3$. Thus, one of the three representations

$$(z_0,z_1)=(-\alpha,z_1)\setminus(-\alpha,z_0]=(z_0,\alpha)\setminus[z_1,\alpha)$$

has the required property. For each $\mathbf{a}'$ and $j\in\{1,\dots,k(\mathbf{a}')\}$, we let $I_j(\mathbf{a}')=J_j^{(1)}(\mathbf{a}')\setminus J_j^{(2)}(\mathbf{a}')$ be such a decomposition, where $J_j^{(2)}(\mathbf{a}')$ is possibly empty.

Given any $\mathbf{a}\in\mathbb{Z}^{n-1}$, we let $\mathbf{a}'$ denote from now on the specific vector

$$\mathbf{a}'=B^{-1/n}(Wqd^2\mathbf{a}+\pi(Wq\tilde{\mathbf{x}}_d+\mathbf{y})).$$

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

With this notation, the main term of (5.3) equals

$$\sum_{\substack{d\leqslant N^{1/C_0}\\ \gcd(d,W)=1}} \frac{\mu(d)}{T} \sum_{\mathbf{y}\in\mathscr{Y}} \sum_{\tilde{\mathbf{x}}_d\in\mathscr{X}_{\mathbf{y},d}} \sum_{\substack{\mathbf{a}\in\mathbb{Z}^{n-1}\\ |\mathbf{a}'|<\alpha}}$$
$$\times \sum_{j=1}^{k(\mathbf{a})} \sum_{x\in\mathbb{Z}} \{\mathbf{1}_{B^{-1/n}(Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y})\in J_j^{(1)}(\mathbf{a}')} - \mathbf{1}_{B^{-1/n}(Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y})\in J_j^{(2)}(\mathbf{a}')}\} F(g(P_{\mathbf{a},\mathbf{y}}(x))\Gamma)$$
$$+ O\Bigg(\sum_{\substack{d\leqslant N^{1/C_0}\\ \gcd(d,W)=1}} \frac{\mu(d)}{T''} \sum_{\mathbf{y}\in\mathscr{Y}} \sum_{\tilde{\mathbf{x}}_d\in\mathscr{X}_{\mathbf{y},d}} \#\{\mathbf{x}\in\mathbb{Z}^n : B^{-1/n}(Wq(d^2\mathbf{x}+\tilde{\mathbf{x}}_d)+\mathbf{y})\in\partial_\varepsilon\mathfrak{X}(1)\}\Bigg). \tag{5.5}$$

Here, the error term accounts for all points in the $B^{1/n}\varepsilon$-neighbourhood of the boundary of $B^{1/n}\mathfrak{X}(1)$ that were excluded through the choice of intervals $I_j(\mathbf{a})$. Observe that we made use of the fact that $\|F\|_\infty \leqslant 1$. Since $\mathfrak{X}(1)$ is $(n-1)$-Lipschitz parameterizable, we have $\mathrm{vol}(\partial_\varepsilon\mathfrak{X}(1)) \asymp \varepsilon$. Together with applications of (3.9) and (3.5), this shows that the error term is bounded by

$$\frac{1}{T}\sum_{d:\gcd(d,W)=1} \#\mathscr{Y}\varrho(d^2,0)\frac{\varepsilon B}{(Wqd^2)^n} \ll \varepsilon\frac{B}{TWq}\frac{\varrho(Wq,A+Wb)}{(Wq)^{n-1}}\sum_d \frac{\varrho(d^2,0)}{d^{2n}}$$
$$\ll \varepsilon\frac{\varrho(W,A)}{W^{n-1}}.$$

Note that the set

$$\left\{x\in\mathbb{Z} : \frac{Wq(d^2x+\tilde{x}_{d,n})+y_n}{B^{1/n}} \in J_j^{(1)}(\mathbf{a}')\right\}$$

is a discrete interval of length

$$\#\left\{x\in\mathbb{Z} : \frac{Wq(d^2x+\tilde{x}_{d,n})+y_n}{B^{1/n}} \in J_j^{(1)}(\mathbf{a}')\right\} \asymp \frac{B^{1/n}}{Wqd^2}.$$

Thus, the total $\delta(N)^{E/C'''}$-equidistribution property of $(g(P_{\mathbf{a};\mathbf{y};d}(x))\Gamma)_{x\leqslant B^{1/n}/(Wqd^2)}$ implies that

$$\Bigg| \sum_{\substack{x\in\mathbb{Z}:\\ B^{-1/n}(Wq(d^2x+\tilde{x}_{d,n})+y_n)\\ \in J_j^{(1)}(\mathbf{a}')}} F(g(P_{\mathbf{a};\mathbf{y};d}(x))\Gamma)\Bigg| \ll \delta(N)^{E/C'''}\frac{B^{1/n}}{Wqd^2}\|F\|_{\mathrm{Lip}}.$$

https://doi.org/10.1017/S1474748015000389 Published online by Cambridge University Press

The same holds for $J_j^{(1)}(\mathbf{a}')$ replaced by any non-empty $J_j^{(2)}(\mathbf{a}')$. Hence (5.5) is bounded by

$$\begin{aligned}
&\ll T^{-1}\sum_{d\leqslant N^{1/C_0}} \#\mathscr{Y}\frac{\varrho(d^2,0)}{d^2}\left(\frac{B^{1/n}}{Wq}\right)^{n-1}\varepsilon^{-1}B^{1/n}(Wq)^{-1}\delta(N)^{E/C'''}\|F\|_{\mathrm{Lip}}\\
&\quad+(\varepsilon+\delta(N)^{O_{m_G,\ell,n}(E/C'')}+\delta(N)^{O_{m_G,\ell,n}(E/C''')})\frac{\varrho(W,A)}{W^{n-1}}\\
&\ll T^{-1}\#\mathscr{Y}\frac{B}{(Wq)^n}\varepsilon^{-1}\delta(N)^{E/C'''}\|F\|_{\mathrm{Lip}}\\
&\quad+(\varepsilon+\delta(N)^{O_{m_G,\ell,n}(E/C'')}+\delta(N)^{O_{m_G,\ell,n}(E/C''')})\frac{\varrho(W,A)}{W^{n-1}}\\
&\ll \frac{\varrho(W,A)}{W^{n-1}}\Big(\varepsilon^{-1}\delta(N)^{E/C'''}\|F\|_{\mathrm{Lip}}+\varepsilon+\delta(N)^{O_{m_G,\ell,n}(E/C'')}+\delta(N)^{O_{m_G,\ell,n}(E/C''')}\Big),
\end{aligned}$$

where we applied the bound (3.5), the fact that $\#\mathscr{Y}=\varrho(Wq,A+Wb)$, and the lifting property (3.9). Choosing $\varepsilon=\delta(N)^{E/2C'''}$ and $E_0=C'''$ completes the proof. □

Acknowledgements. This paper owes its existence to questions from Yonatan Harpaz and Olivier Wittenberg. I am grateful to both of them and to Tim Browning for their comments on an earlier version of this paper, to Alexei Skorobogatov for initiating the discussions, and to the referee for valuable suggestions and comments that helped improve the paper.

References

1. T. D. Browning and L. Matthiesen, Norm forms for arbitrary number fields as products of linear polynomials, arXiv:1307.7641 (submitted).
2. B. J. Green and T. C. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. of Math. (2)* **167** (2008), 481–547.
3. B. J. Green and T. C. Tao, Linear equations in primes, *Ann. of Math. (2)* **171** (2010), 1753–1850.
4. B. J. Green and T. C. Tao, The quantitative behaviour of polynomial orbits on nilmanifolds, *Ann. of Math. (2)* **175** (2012), 465–540.
5. B. J. Green and T. C. Tao, The Möbius function is strongly orthogonal to nilsequences, *Ann. of Math. (2)* **175** (2012), 541–566.
6. B. J. Green and T. C. Tao, On the quantitative behaviour of polynomial orbits on nilmanifolds – Erratum, arXiv:1311.6170.
7. B. J. Green, T. C. Tao and T. Ziegler, An inverse theorem for the Gowers $U_{s+1}[N]$-norm, *Ann. of Math. (2)* **176** (2012), 1231–1372.
8. Y. Harpaz and O. Wittenberg, On the fibration method for zero-cycles and rational points, *Ann. of Math. (2)*, to appear.
9. A. Leibman, Polynomial mappings of groups (corrected version), *Israel J. Math.* **129** (2002), 29–60.
10. L. Matthiesen, Linear correlations amongst numbers represented by positive definite binary quadratic forms, *Acta Arith.* **154** (2012), 235–306.
11. L. Matthiesen, Generalized Fourier coefficients of multiplicative functions, arXiv:1405.1018.
12. L. Matthiesen, A consequence of the factorisation theorem for polynomial nilsequences – Corrigendum to Acta Arith. 154 (2012), 235–306, arXiv:1509.06030 (submitted).