

On the representation of friable integers by linear forms

by

ARMAND LACHAND (Vandœuvre-lès-Nancy)

1. Introduction and statement of the result. Given a real number $y > 1$, an integer n is said to be *y-friable* if its greatest prime factor, denoted by $P^+(n)$, satisfies $P^+(n) \leq y$, with the conventions $P^+(\pm 1) = 1$ and $P^+(0) = 0$. On the other hand, an integer n is called *y-sifted* if its smallest prime factor, denoted by $P^-(n)$, satisfies $P^-(n) > y$ with the conventions $P^-(\pm 1) = +\infty$ and $P^-(0) = 0$.

Due to the duality between sifted integers and friable integers, such integers occur in several places in number theory, and their distribution has been intensively studied (see [17] and [10] for survey articles related to integers without large prime factors). A theorem of Hildebrand [16], concerning the number $\Psi(N, y)$ of *y-friable* integers smaller than N , asserts that, for any $\varepsilon > 0$ and uniformly in the domain

$$(1.1) \quad N \geq 3 \quad \text{and} \quad 1 \leq u \leq \frac{\log N}{(\log \log N)^{5/3+\varepsilon}},$$

we have the asymptotic formula

$$(1.2) \quad \Psi(N, N^{1/u}) = N\rho(u) \left(1 + O\left(\frac{u \log(u+1)}{\log N} \right) \right)$$

where ρ is the Dickman function, the unique solution to the delay differential equation

$$\begin{cases} \rho(u) = 1 & \text{if } 0 \leq u \leq 1, \\ u\rho'(u) + \rho(u-1) = 0 & \text{if } u > 1. \end{cases}$$

2010 *Mathematics Subject Classification*: Primary 11N25; Secondary 11N37.

Key words and phrases: friable integers, linear forms, Gowers norms.

Received 30 April 2015; revised 17 September 2017.

Published online 6 November 2017.

Given $F \in \mathbb{Z}[X_1, \dots, X_d]$ and $\mathcal{K} \subset \mathbb{R}^d$, the cardinality

$$\Psi_F(\mathcal{K}, y) := \#\{\mathbf{n} \in \mathcal{K} \cap \mathbb{Z}^d : P^+(F(\mathbf{n})) \leq y\}$$

is of interest. In particular, the factorization algorithm Number Field Sieve (NFS) ⁽¹⁾ rests on the assumption that the cardinality $\Psi_F(\mathcal{K}, y)$ is sufficiently large for some small y , for $F \in \mathbb{Z}[X_1, X_2]$ and $\mathcal{K} \subset \mathbb{R}^2$ a sufficiently regular compact set.

Let $F = F_1^{j_1} \cdots F_t^{j_t}$ be the decomposition of F with F_1, \dots, F_t the distinct irreducible factors of F and k_1, \dots, k_t their respective degrees with $k_1 \geq \dots \geq k_t \geq 1$. If we assume the events “ $F_i(\mathbf{n})$ is y -friable” to be independent, then (1.2) leads to the following conjecture:

$$(1.3) \quad \Psi_F([0, N]^d, N^{1/u}) \underset{N \rightarrow \infty}{\sim} N^d \rho(k_1 u) \cdots \rho(k_t u)$$

for any fixed $u > 0$.

When $d = 2$, the present author [18, 19] proved the validity of (1.3) for an irreducible cubic form F or for $F = F_1 F_2$ where F_1 is a linear form and F_2 is an irreducible quadratic form. For general binary forms F , such a formula seems beyond reach, but there exist some partial results for estimating $\Psi_F([0, N]^2, N^{1/u})$ when u is sufficiently small. Balog, Blomer, Dartyge and Tenenbaum [1] proved the existence of a constant $\alpha_F > 1/d_1$ such that, for any $\varepsilon > 0$ and uniformly for $N \geq 2$, we have

$$(1.4) \quad \Psi_F([0, N]^2, N^{1/\alpha_F + \varepsilon}) \gg_\varepsilon N^2.$$

Let $d \geq 2$ and $t \geq 1$ be integers. In this paper, we focus on binary forms $F = F_1 \cdots F_t$ where F_1, \dots, F_t are affine-linear forms in $\mathbb{Z}[X_1, \dots, X_d]$. The cases $d = 2$ and $t \in \{1, 2\}$ can be deduced from results of [7] related to the distribution of friable integers in arithmetic progressions. The case $d = 2$ and $t = 3$ was essentially considered in a succession of articles of various authors [3, 21, 5, 4, 6, 15]. In [15, Corollary 1], Harper used the Hardy–Littlewood circle method to show the existence of $c > 0$ such that, uniformly for $N \geq 2$ and $y \geq (\log N)^c$, we have

$$\Psi_{X_1 X_2 (X_1 + X_2)}(\mathcal{K}(N), y) \underset{N \rightarrow \infty}{\sim} \mathfrak{S}_0(\alpha, y) \mathfrak{S}_1(\alpha) \frac{\Psi(N, y)^3}{N}$$

where $\mathcal{K}(N) = \{1 \leq n_1, n_2 \leq N : n_1 + n_2 \leq N\}$, $\alpha := \alpha(N, y)$ denotes the unique real solution of the equation

$$\sum_{p \leq y} \frac{\log p}{p^\alpha - 1} = \log N,$$

⁽¹⁾ The interested reader may find a description of this algorithm in [2, Chapter 6].

and

$$\mathfrak{S}_0(\alpha, y) := \prod_{p \leq y} \left(1 + \frac{(p - p^\alpha)^3}{p(p-1)^2(p^{3\alpha-1} - 1)} \right) \prod_{p > y} \left(1 - \frac{1}{(p-1)^2} \right),$$

$$\mathfrak{S}_1(\alpha) := \int_0^1 \int_0^{1-t_1} \alpha^3 (t_1 t_2 (t_1 + t_2))^{\alpha-1} dt_2 dt_1.$$

The celebrated work of Green, Tao and Ziegler [11–14] provides a scheme—the so-called nilpotent Hardy–Littlewood method—to get asymptotic estimates of the average value

$$(1.5) \quad M_{F_1 \dots F_t}(\mathcal{K}; h) := \sum_{\mathbf{n} \in \mathcal{K} \cap \mathbb{Z}^d} h(F_1(\mathbf{n})) \cdots h(F_t(\mathbf{n}))$$

for any system of affine-linear forms $F_1, \dots, F_t \in \mathbb{Z}[X_1, \dots, X_d]$ such that no two forms are affinely related and for any arithmetic function h with a quasi-random behavior. In recent years, this approach has been applied successfully for several functions including the von Mangoldt function Λ (this gives a partial resolution of the generalized Hardy–Littlewood conjecture [11]), the Liouville function λ or the Möbius function μ [11], the divisor function τ [23], the function r_G which counts the number of representations of a binary quadratic form G [24, 25] or, very recently, any multiplicative function that takes values in the unit disk [8].

In this work, we study how the nilpotent Hardy–Littlewood method may be applied to get an asymptotic formula for (1.5) when $h = 1_{S(N^{1/u})}$ is the indicator function of the $N^{1/u}$ -friable integers for bounded $u \geq 1$. This question is not covered by Frantzikinakis and Host’s work [8] since h depends on N in the present case. The main result is the following theorem.

THEOREM 1.1. *Let N, L, d, t and u_0 be positive integers. Suppose that $F = (F_1, \dots, F_t) : \mathbb{Z}^d \rightarrow \mathbb{Z}^t$ is a system of affine-linear forms such that any two forms F_i and F_j are affinely independent over \mathbb{Q} and the non-constant coefficients of the F_i are bounded by L . Then, for any convex body $\mathcal{K} \subset [-N, N]^d$ with $F(\mathcal{K}) \subset [0, N]^t$ and for any $u_1, \dots, u_t \in [0, u_0]$, we have*

$$\sum_{\mathbf{n} \in \mathcal{K} \cap \mathbb{Z}^d} 1_{S(N^{1/u_1})}(F_1(\mathbf{n})) \cdots 1_{S(N^{1/u_t})}(F_t(\mathbf{n})) = \text{Vol}(\mathcal{K}) \prod_{i=1}^t \rho(u_i) + o(N^d)$$

where the implicit constant depends only on t, d, L, u_0 , and $S(y)$ denotes the set of y -friable integers.

Compared to the result of Balog et al. [1], we get two major improvements in the case of linear forms:

- Theorem 1.1 gives an asymptotic equivalent which is consistent with the conjectural formula (1.3), whereas (1.4) only gives a lower bound;

- when $t \geq 4$, formula (1.4) is valid with $\alpha_F = 1 + \frac{2}{t-2}$, while Theorem 1.1 shows that we can choose any positive real number for α_F .

Outline and perspectives. In its original form, the nilpotent Hardy–Littlewood method is concerned with arithmetic functions h which are equidistributed in residue classes to small moduli and supported on a set of integers with positive asymptotic density. For such functions, the problem is reduced to showing that h is suitably Gowers uniform to allow one to deduce asymptotics for $M_{F_1 \dots F_t}(\mathcal{K}; h)$ (see the description of the method in Section 2).

In many applications, the function h may not satisfy the previous two conditions. The method developed in [11, 23, 24] to overcome these difficulties consists of two steps:

- the decomposition of h into a sum of functions which are equidistributed in residue classes to small moduli (W -trick [11, Section 5]),
- the construction of a pseudorandom measure ν dominating h with a view to applying a transference principle [11, Section 10].

For bounded $u \geq 1$, the set of $N^{1/u}$ -friable integers has positive density $\rho(u)$ and is well-behaved in arithmetic progressions of small common difference (see the work of Fouvry and Tenenbaum [7]). In particular, the problem may be directly handled by using the nilpotent Hardy–Littlewood method and showing that h has small Gowers uniformity norm. This may be viewed as an application of the impressive results of Matthesen [22] related to the orthogonality between multiplicative functions and nilsequences. In Section 3 of the present paper, we develop a more direct and simple approach to studying linear correlations of friable integers.

It would be interesting to prove (1.3) for unbounded parameters u . In this case, the sequence of friable integers is too sparse to enable us to directly apply Green–Tao–Ziegler’s work. A major step to get this generalization would be to construct a pseudorandom majorant for $1_{S(N^{1/u})}$.

2. A brief description of the nilpotent Hardy–Littlewood method.

In this section, we recall two important arguments of the nilpotent Hardy–Littlewood method. The generalized von Neumann theorem—due to Gowers [9] and Green–Tao [11]—reduces the estimation of $M_{F_1 \dots F_t}(\mathcal{K}; h)$ defined in (1.5) to the study of the Gowers uniformity norm $\|h\|_{U^{t-1}[N]}$ (see [11, Appendix B] for the definition of Gowers norms).

THEOREM A ([11, Proposition 7.1]). *Let $t, d, L \geq 1$ be integers. Suppose that $h_1, \dots, h_t : [0, N] \rightarrow \mathbb{R}$ are functions bounded by 1 and that $F = (F_1, \dots, F_t) : \mathbb{Z}^d \rightarrow \mathbb{Z}^t$ is a system of affine-linear forms whose non-constant coefficients are bounded by L and any two forms F_i and F_j are*

affinely independent over \mathbb{Q} . Let $\mathcal{K} \subset [-N, N]^d$ be a convex body such that $F(\mathcal{K}) \subset [0, N]^t$. Suppose also that

$$\min_{1 \leq i \leq t} \|h_i\|_{U^{t-1}[N]} \leq \delta$$

for some $\delta > 0$. Then

$$\sum_{\mathbf{n} \in \mathcal{K}} \prod_{i=1}^t h_i(F_i(\mathbf{n})) = o_\delta(N^d) + \kappa(\delta)N^d$$

where $\kappa(\delta) \rightarrow 0$ as $\delta \rightarrow 0$.

Proof. Let $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ be the canonical basis of \mathbb{R}^d and fix $\mathbf{n} = n_1\mathbf{e}_1 + \dots + n_d\mathbf{e}_d \in \mathcal{K}$. Then

$$|F_i(\mathbf{0})| \leq \sum_{i=1}^d |n_i|L + |F_i(\mathbf{n})| \leq (dL + 1)N$$

because $F(\mathbf{n}) \in [0, N]^t$ and $\mathcal{K} \in [0, N]^d$. With the definition of the norm $\|\cdot\|_N$ of [11, (1.1)], we therefore have $\|F\|_N \ll_{d,t} L$, and [11, Proposition 7.1] can be used to get the result. ■

The inverse theorem for the Gowers norms, proved by Green, Tao and Ziegler [14], exhibits the link between linear correlations and polynomial nilsequences. The reader may refer to [12] for definitions and properties of filtered nilmanifolds and polynomial nilsequences.

THEOREM B ([14, Theorem 1.3]). *Let $s \geq 0$ be an integer and let $\delta \in]0, 1]$. Then there exists a finite collection $\mathcal{M}_{s,\delta}$ of s -step nilmanifolds G/Γ , each equipped with some smooth Riemannian metric $d_{G/\Gamma}$, as well as positive constants $C(s, \delta)$ and $c(s, \delta)$ with the following property: Whenever $N \geq 1$ and $h : [0, N] \cap \mathbb{Z} \rightarrow [-1, 1]$ is a function such that*

$$\|h\|_{U^{s+1}[N]} \geq \delta,$$

there exists a filtered nilmanifold $G/\Gamma \in \mathcal{M}_{s,\delta}$, a function $F : G/\Gamma \rightarrow \mathbb{C}$ bounded in magnitude by 1 and with Lipschitz constant at most $C(s, \delta)$ with respect to the metric $d_{G/\Gamma}$, and a polynomial nilsequence $g : \mathbb{Z} \rightarrow G$ such that

$$\left| \sum_{0 \leq n \leq N} h(n)F(g(n)\Gamma) \right| \geq c(s, \delta)N.$$

We now describe the application of the Green–Tao method to the functions $1_{S(N^{1/u_i})}$. For any parameter of friability N^{1/u_i} , we consider the balanced function

$$h_i : \mathbb{N} \rightarrow [-1, 1], \quad n \mapsto 1_{S(N^{1/u_i})}(n) - \rho(u_i).$$

By writing $1_{S(N^{1/u_i})}(n) = h_i(n) + \rho(u_i)$ and using the bound $\rho(u_i) \leq 1$, it follows that

$$\begin{aligned} & \left| \Psi_{F_1 \dots F_t}(\mathcal{K}, N^{1/u}) - \text{Vol}(\mathcal{K}) \prod_{i=1}^t \rho(u_i) \right| \\ & \leq \sum_{\substack{I \subset \{1, \dots, t\} \\ I \neq \emptyset}} \left| \sum_{\mathbf{n} \in \mathcal{K} \cap \mathbb{Z}^d} \prod_{i \in I} h_i(F_i(\mathbf{n})) \right| + O_d(N^{d-1}). \end{aligned}$$

In view of the inverse theorem, the problem is reduced to proving that, for any $i \in \{1, \dots, t\}$, the function h_i does not correlate with nilsequences, namely the upper bound

$$\sum_{n \leq N} h_i(n) F(g(n)\Gamma) = o(N)$$

holds for any $(t - 2)$ -step nilsequences $F(g(n)\Gamma)$.

3. Non-correlation with nilsequences. Let $s \geq 0, u_0 \geq 1$ be integers and let $(G/\Gamma, G_{\mathbb{N}})$ be a filtered nilmanifold of degree s . In this section, we show that for any 1-bounded Lipschitz function $F : G/\Gamma \rightarrow \mathbb{C}$, any polynomial nilsequence $g : \mathbb{Z} \rightarrow G$ adapted to $G_{\mathbb{N}}, 1 \leq u \leq u_0$ and $N \geq 1$, we have

$$(3.1) \quad \sum_{n \leq N} h(n) F(g(n)\Gamma) = o(N(1 + \|F\|_{\text{Lip}, d_{G/\Gamma}}))$$

where $h(n) := 1_{S(N^{1/u})}(n) - \rho(u)$ and the implicit term $o(\cdot)$ only depends on G/Γ and u_0 . In view of Theorems A and B, this will imply Theorem 1.1.

Mathiesen [22] developed a method to bound the correlations of a multiplicative function with polynomial nilsequences, under some density and growth conditions and some control of the second moment. This approach mixes the Montgomery–Vaughan method [26], the factorization theorem for polynomial sequences from Green–Tao [12] and the fact that the W -tricked von Mangoldt function is orthogonal to nilsequences [11]. The main result [22, Theorem 5.1] may be applied directly to the multiplicative function $1_{S(N^{1/u})}(n)$ to get (3.1), once we have checked that the assumptions required are satisfied. In the case of the indicator of friable integers and for any $E \geq 1$, the various hypotheses which defined the set $\mathcal{F}_1(E)$ of [22] can be essentially deduced from the estimate

$$\sum_{\substack{n \leq N \\ n \equiv a \pmod q}} 1_{S(N^{1/u})}(n) \underset{N \rightarrow \infty}{\sim} \frac{N}{q} \rho(u),$$

which holds uniformly for $1 \leq a, q \leq (\log N)^E$ (see [7]).

In the rest of this paper, we give a direct and simple method to establish (3.1), with a different focus than in [22]. The starting point is the Möbius inversion formula in the form

$$1_{S(N^{1/u})}(n) = \sum_{P^-(k) > N^{1/u}} \mu(k) 1_{k|n}.$$

We approximate the indicator $1_{k|n}$ by its mean value $1/k$ for $k \leq N^{1-\tau}$ where the parameter $\tau = o(1) \in]1/\log N, 1[$ will be chosen later. One can write

$$\sum_{1 \leq n \leq N} (1_{S(N^{1/u})}(n) - \rho(u)) F(g(n)\Gamma) = \Sigma_1(F, g) + \Sigma_2(F, g)$$

where

$$\begin{aligned} \Sigma_1(F, g) &:= \sum_{1 \leq n \leq N} h_\tau(n) F(g(n)\Gamma) \text{ with } h_\tau(n) = \sum_{\substack{k \leq N^{1-\tau} \\ P^-(k) > N^{1/u}}} \mu(k) \left(1_{k|n} - \frac{1}{k} \right), \\ \Sigma_2(F, g) &:= \sum_{1 \leq n \leq N} \left(\sum_{\substack{k > N^{1-\tau} \\ P^-(k) > N^{1/u}}} \mu(k) 1_{k|n} + \sum_{\substack{k \leq N^{1-\tau} \\ P^-(k) > N^{1/u}}} \frac{\mu(k)}{k} - \rho(u) \right) F(g(n)\Gamma). \end{aligned}$$

In the definition of the function h_τ , the summation is restricted to the divisors $k \leq N^{1-\tau}$ since the contribution from the interval $N^{1-\tau} < k \leq N$ is negligible (see (3.5) below).

First, we focus on $\Sigma_2(F, g)$. In view of the following series of estimations, valid whenever $\tau u < 1$,

$$\begin{aligned} (3.2) \quad & \sum_{\substack{N^{1-\tau} < k \leq N \\ P^-(k) > N^{1/u}}} \frac{\mu^2(k)}{k} \\ & \ll \sum_{j \geq 1} \sum_{N^{1/u} < p_2 < \dots < p_j \leq N} \frac{1}{p_2 \cdots p_j} \sum_{\max(N^{1/u}, \frac{N^{1-\tau}}{p_2 \cdots p_j}) \leq p_1 \leq \frac{N}{p_2 \cdots p_j}} \frac{1}{p_1} \\ & \ll \tau u \sum_{j \geq 1} \frac{1}{(j-1)!} \left(\sum_{N^{1/u} < p \leq N} \frac{1}{p} \right)^{j-1} \\ & \ll \tau u \sum_{j \geq 1} \frac{1}{(j-1)!} (\log(u) + O(1))^{j-1} \ll \tau u^2, \end{aligned}$$

we have the upper bound

$$(3.3) \quad \sum_{1 \leq n \leq N} \sum_{\substack{k > N^{1-\tau} \\ P^-(k) > N^{1/u}}} \mu^2(k) 1_{k|n} \ll \tau u^2 N.$$

On the other hand, one can handle the sum over $k \leq N^{1-\tau}$ in $\Sigma_2(F, g)$ by using [20, (1.5)], which states that

$$(3.4) \quad \sum_{\substack{k \leq N \\ P^-(k) > N^{1/u}}} \frac{\mu(k)}{k} = \rho(u) \left(1 + O\left(\frac{u \log(u+1)}{\log N} \right) \right)$$

for any $\varepsilon > 0$, uniformly for $x \geq 2$ and $1 \leq u \leq (\log x)^{3/8-\varepsilon}$. Finally, (3.3) and (3.4) yield

$$(3.5) \quad \Sigma_2(F, g) \ll uN \left(\tau u + \frac{\rho(u) \log(u+1)}{\log N} \right).$$

In view of the foregoing, it remains to obtain an upper bound for $\Sigma_1(F, g)$. This is the subject of the following proposition.

PROPOSITION 3.1. *Let $m, s \geq 1$ be integers and let $A > 0$ be a real number. There exists a constant $c(m, s, A) > 0$ with the following property. Whenever $Q, N \geq 2$ are integers, $\tau \in]0, 1/2[$ and $u \geq 1$ are such that $\min(N^\tau, N^{1/u}) \geq (\log N)^{c(m, s, A)}$, $(G/\Gamma, G_{\mathbb{N}})$ is a filtered nilmanifold of degree s and dimension m , \mathcal{X} is a Q -rational Mal'cev basis ⁽²⁾ of $(G/\Gamma, G_{\mathbb{N}})$, $g : \mathbb{Z} \rightarrow G/\Gamma$ is a polynomial nilsequence adapted to $G_{\mathbb{N}}$ and $F : G/\Gamma \rightarrow [-1, 1]$ is a Lipschitz function, we have*

$$(3.6) \quad \sum_{1 \leq n \leq N} h_\tau(n) F(g(n)\Gamma) \leq NQ^{c(m, s, A)} (1 + \|F\|_{\text{Lip}, \mathcal{X}}) 2^u (\log N)^{-A}.$$

Recall that the smooth Riemannian metric $d_{G/\Gamma}$ of Theorem B is equivalent to the metric $d_{\mathcal{X}}$ (see [12, footnote 4 and Definition 2.2]). With the choice $\tau = (\log \log N)^{1+\varepsilon} / \log N$, it follows from (3.5) and (3.6) that

$$\sum_{n \leq N} h(n) F(g(n)\Gamma) = o(N\rho(u)(1 + \|F\|_{\text{Lip}, d_{G/\Gamma}}))$$

for any $\varepsilon > 0$ and uniformly for $1 \leq u \leq (\log \log N)^{1-\varepsilon}$. This implies (3.1) since $1 \leq u \leq u_0$ is contained in this region for any u_0 , which does not depend on N .

The rest of the article is devoted to the proof of Proposition 3.1. The argument follows essentially the proofs of [13, Theorem 1.1] and [23, Theorem 9.1] and we only outline the major differences. A key point is to reduce the problem to establishing (3.6) in the case of a totally equidistributed polynomial nilsequence g , i.e. such that $|P|^{-1} \sum_{n \in P} F(g(n)\Gamma)$ tends to $\int_{G/\Gamma} F$ where P is a subprogression such that $|P| \rightarrow \infty$.

After this reduction, it will be possible to use the following analogue of [13, Proposition 2.1] and [23, Proposition 9.2].

⁽²⁾ The notion of Q -rational Mal'cev basis is introduced in [12, Definitions 2.1 and 2.4] as a specific basis of the Lie algebra \mathfrak{g} of G .

PROPOSITION 3.2. *Let m, s be positive integers. There exist constants $c_0(m, s), c_1(m, s) > 0$ with the following property. Whenever $Q, N \geq 2$ and $\delta \in]0, 1/2[$ are such that $\delta^{-c_0(m,s)} \leq N^\tau$, $P \subset \{1, \dots, N\}$ is an arithmetic progression of size at least N/Q , $(G/\Gamma, G_{\mathbb{N}})$ is a filtered nilmanifold of degree s and dimension m , \mathcal{X} is a Q -rational Mal'cev basis of $(G/\Gamma, G_{\mathbb{N}})$, $g : \mathbb{Z} \rightarrow G/\Gamma$ is a polynomial and δ -totally equidistributed nilsequence ⁽³⁾ adapted to $G_{\mathbb{N}}$ and $F : G/\Gamma \rightarrow [-1, 1]$ is a Lipschitz function such that $\int_{G/\Gamma} F = 0$, we have*

$$\left| \sum_{n \leq N} h_\tau(n) 1_P(n) F(g(n)\Gamma) \right| \ll \delta^{c_1(m,s)} \|F\|_{\text{Lip}, \mathcal{X}} QN (2^u + \log N).$$

Proof that Proposition 3.2 implies Proposition 3.1. Following some ideas of [13], we can assume that $\|F\|_{\text{Lip}, \mathcal{X}} = 1$ and $Q \leq \log N$. Let $B > 0$ be a parameter to be specified at the end of the proof. By [12, Theorem 1.19], there exists an integer M satisfying $\log N \leq M \leq (\log N)^{c(m,s,B)}$ such that we can decompose $g = \varepsilon g' \gamma$ where

- $\varepsilon \in \text{poly}(\mathbb{Z}, G_{\mathbb{N}})$ is (M, N) -smooth [12, Definition 1.18],
- $g' \in \text{poly}(\mathbb{Z}, G_{\mathbb{N}})$ takes values in a rational subgroup $G' \subseteq G$ with Mal'cev basis \mathcal{X}' , and $(g'(n))_{n \leq N}$ is M^{-B} -totally equidistributed in $G'/(G' \cap \Gamma)$ for the metric $d_{\mathcal{X}'}$ [12, Definition 1.10],
- $\gamma \in \text{poly}(\mathbb{Z}, G_{\mathbb{N}})$ is periodic of period $q \leq M$, and $\gamma(n)$ is M -rational for any $n \in \mathbb{Z}$ [12, Definition 1.17].

Next, we reproduce the arguments of Green and Tao based on partitioning and pigeonholing and we use the periodicity and smoothness of γ and ε . In this way, the problem is reduced to showing that

$$(3.7) \quad \left| \sum_{1 \leq n \leq N} h_\tau(n) 1_P(n) F'(g'(n)\Gamma') \right| \ll 2^u N / (M^2 (\log N)^{2A})$$

where P is a subprogression such that

$$|P| \geq \frac{N}{2M^2 (\log N)^A},$$

$(G'/\Gamma', G'_{\mathbb{N}})$ is an m -dimensional nilmanifold of degree s with $M^{C_1(m,s)}$ -rational Mal'cev basis \mathcal{X}' , $F' : G'/\Gamma' \rightarrow [-1, 1]$ is a Lipschitz function such that $\|F'\|_{\text{Lip}, \mathcal{X}'} \leq M^{C_1(m,s)}$ and $g' \in \text{poly}(\mathbb{Z}, G'_{\mathbb{N}})$ is $M^{-C_2(m,s)B + C_1(m,s)}$ -totally equidistributed, for some constants $C_1(m, s), C_2(m, s) > 0$.

⁽³⁾ A sequence $(g(n)\Gamma)_{n \leq N}$ is δ -totally equidistributed if

$$\left| \frac{1}{|P|} \sum_{n \in P} F(g(n)\Gamma) \right| \leq \delta \|F\|$$

for all Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$ with $\int_{G/\Gamma} F = 0$ and all arithmetic progressions $P \subset \{1, \dots, N\}$ of size at least δN .

If we suppose that $\int_{G'/\Gamma'} F' = 0$, then we can apply Proposition 3.2 to the sequence g' , with $M^{C_1(m,s)}$ (resp. $M^{-C_2(m,s)B+C_1(m,s)}$) as parameter of rationality (resp. total equidistribution). For B , $C_1(m,s)$ and $c(m,s,A)$ sufficiently large, the hypotheses on P and δ are satisfied and we get (3.7).

We can reduce to this last case by writing $F' = (F' - \int_{G'/\Gamma'} F') + \int_{G'/\Gamma'} F'$. Indeed, observe that $\int_{G'/\Gamma'} F'$ is bounded by 1, and since the common difference q of P satisfies $q < N^{1/u}$, we get some multiplicative independence when $P^-(k) > N^{1/u}$:

$$\left| \sum_{1 \leq n \leq N} \left(1_{k|n} - \frac{1}{k} \right) 1_P(n) \right| \leq 1.$$

We deduce the major arc estimate

$$\begin{aligned} \left| \sum_{1 \leq n \leq N} h_\tau(n) 1_P(n) \int_{G'/\Gamma'} F' \right| &\leq \sum_{\substack{k \leq N^{1-\tau} \\ P^-(k) > N^{1/u}}} \left| \sum_{1 \leq n \leq N} \left(1_{k|n} - \frac{1}{k} \right) 1_P(n) \right| \\ &\leq |\{k \leq N^{1-\tau} : P^-(k) > N^{1/u}\}| \ll u \frac{N^{1-\tau}}{\log N}, \end{aligned}$$

which implies (3.7) under the condition $N^\tau \geq (\log N)^{c(m,s,A)}$. ■

Proof of Proposition 3.2. We essentially follow [23, proof of Proposition 9.2] and we suppose that $\|F\|_{\text{Lip}, \mathcal{X}} = 1$ and $Q \leq \delta^{-c_1(m,s)}$. For $\mathcal{T} \in]0, 1/2[$ and $j \geq 1$, we define $S_j(\mathcal{T})$ as the set of all integers k satisfying

$$\left| \sum_{2^j/k < n \leq 2^{j+1}/k} 1_P(kn) F(g(kn)\Gamma) \right| > \mathcal{T} \frac{2^j}{k}.$$

From the estimation

$$\sum_{\substack{k \leq N \\ P^-(k) > N^{1/u}}} \frac{\mu^2(k)}{k} \ll \sum_{j \geq 1} \frac{1}{j!} \left(\sum_{N^{1/u} < p \leq N} \frac{1}{p} \right)^j \ll u$$

and the trivial bound $|\{k|n : P^-(k) > N^{1/u}\}| \leq 2^u$ valid whenever $n \leq N$, we can see that $h_\tau(n) \ll 2^u$. It follows that

$$\left| \sum_{n \leq N^{1-\tau/2}} h_\tau(n) 1_P(n) F(g(n)\Gamma) \right| \ll N^{1-\tau/2} 2^u,$$

and therefore we concentrate on the integers $n > N^{1-\tau/2}$.

Since the nilsequence $(g(n)\Gamma)_{n \leq N}$ is δ -totally equidistributed, the contribution from the part $\sum_k \mu(k)/k$ of h_τ can be handled by observing that

$$\sum_{\substack{k \leq N^{1-\tau} \\ P^-(k) > N^{1/u}}} \frac{\mu^2(k)}{k} \left| \sum_{N^{1-\tau/2} \leq n \leq N} 1_P(n) F(g(n)\Gamma) \right| \ll u \delta N.$$

For the remaining terms $\sum_k \mu(k) 1_{k|n}$ of h_τ , we follow [23, proof of Proposition 9.2]. We make a dyadic splitting over the variables k and n and we drop off the condition $P^-(k) > N^{1/u}$:

$$\begin{aligned} & \sum_{k \leq N^{1-\tau}} \left| \sum_{N^{1-\tau/2}/k < n \leq N/k} 1_P(kn) F(g(kn)\Gamma) \right| \\ & \ll \sum_{2^i \leq N^{1-\tau}} \sum_{N^{1-\tau/2}/2 \leq 2^j \leq N} 2^j \left(\sum_{2^i \leq k < 2^{i+1}} \frac{\mathcal{T}}{k} + \sum_{\substack{2^i \leq k < 2^{i+1} \\ k \in S_j(\mathcal{T})}} \frac{1}{k} \right) \\ & \ll \sum_{N^{1-\tau/2}/2 \leq 2^j \leq N} 2^j \left(\mathcal{T} \log N + \sum_{2^i \leq N^{1-\tau}} \frac{1}{2^i} \#(S_j(\mathcal{T}) \cap [2^i, 2^{i+1}]) \right). \end{aligned}$$

Set $\mathcal{T} := \delta^{c_1(m,s)} \leq Q^{-1}$ for a constant $c_1(m,s) > 0$ sufficiently small. In the previous sum, the contribution of the range $N^{1-\tau/2}/2 \leq 2^j \leq \mathcal{T}N$ is negligible and may be bounded by the trivial inequality.

The rest of the proof consists in showing that if $K \leq N^{1-\tau}$, then

$$(3.8) \quad \#(S_j(\mathcal{T}) \cap [K, 2K]) \leq \mathcal{T}K$$

whenever $\mathcal{T}N \leq 2^j \leq N$.

The estimate (3.8) is the analogue of [23, Lemma 9.3] under the constraint $K \leq N^{1-\tau}$ rather than $K \leq N^{1/2}$ and in the special case $\overline{W} = 1$ and $b = 0$. To see this, we follow the discussion of the Type I case of [13, Part 3] and we suppose for contradiction that (3.8) does not hold for some $K \leq N^{1-\tau}$ and $\mathcal{T}N \leq 2^j \leq N$. By reproducing those authors' arguments, we get the existence of a non-trivial horizontal character ψ with $0 < |\psi| \leq \mathcal{T}^{-c_2(m,s)}$ such that, for any $r \geq 1$ and for at least $\mathcal{T}^{c_2(m,s)}K$ values of k , we have

$$\|\partial^r(\psi \circ g_k)(0)\|_{\mathbb{R}/\mathbb{Z}} \leq \mathcal{T}^{-c_2(m,s)}(K/2^j)^r$$

where $g_k(n) = g(kn)$, which is the analogue of [13, (3.7)].

By [13, Lemmas 3.2 and 3.3]—consequences of Waring's theorem—it follows that there exists an integer $q \ll_s 1$ and at least $\mathcal{T}^{c_3(m,s)}K^r$ integers $l \leq 10^s K^r$ such that

$$\|ql\beta_r\|_{\mathbb{R}/\mathbb{Z}} \leq \mathcal{T}^{-c_3(m,s)}(K/2^j)^r,$$

where the β_r 's are defined by

$$(3.9) \quad \psi \circ g(n) = \beta_s n^s + \cdots + \beta_0.$$

To deduce some diophantine information about the β_r 's, we invoke [12, Lemma 3.2] as in [13] after checking that the hypotheses are satisfied. It suffices to see that $r \geq 1$ and $\mathcal{T}^{2c_3(m,s)}/10^s \gg N^{-\tau} \geq (K/2^j)^r$ if the constant $c_1(m,s)$ is chosen sufficiently small. It follows that there exists

$q' \leq \mathcal{T}^{-c_4(m,s)}$ such that

$$\|q' \beta_r\|_{\mathbb{R}/\mathbb{Z}} \leq \mathcal{T}^{-c_4(m,s)} 2^{-rj}$$

for any integer $r \geq 1$. By the definition (3.9), there exists $c_5(m, s) > 0$ sufficiently large such that $q' \leq \mathcal{T}^{-c_5(m,s)}$ and

$$(3.10) \quad \|q'(\psi \circ g)(n)\|_{\mathbb{R}/\mathbb{Z}} \leq 1/10$$

for any $n \leq \mathcal{T}^{c_5(m,s)} 2^j$.

Let $\eta : \mathbb{R}/\mathbb{Z} \rightarrow [-1, 1]$ be a Lipschitz function of $O(1)$ norm, mean value zero, and equal to 1 on $[-1/10, 1/10]$, such that

$$\int_{G/\Gamma} \eta \circ (q'\psi) = 0 \quad \text{and} \quad \|\eta \circ (q'\psi)\|_{\text{Lip}, \mathcal{X}} \leq \mathcal{T}^{-c_5(m,s)}.$$

It follows from (3.10) that

$$\left| \sum_{n \leq \mathcal{T}^{c_5(m,s)} 2^j} \eta(q'\psi(g(n)\Gamma)) \right| \geq \mathcal{T}^{c_5(m,s)} 2^j > \delta \|\eta \circ (q'\psi)\|_{\text{Lip}, \mathcal{X}} \mathcal{T}^{c_5(m,s)} 2^j$$

whenever $c_1(m, s)$ is sufficiently small. This contradicts the hypothesis that $(g(n))_{n \leq N}$ is δ -totally equidistributed, the set of integers less than $\mathcal{T}^{c_5(m,s)} 2^j$ being an arithmetic progression of size at least δN whenever $c_1(m, s)$ is sufficiently small since $2^j \geq \mathcal{T}N$. ■

Acknowledgments. The author would like to thank Trevor Wooley for his suggestion to study this method, Régis de la Bretèche, François Hennecart and Anne de Roton for their interest in this work, and his Ph.D. advisor Cécile Dartyge for her continuous support. The major part of this work was completed while the author was a Ph.D. student at Université de Lorraine. He put the finishing touch while he was a postdoctoral fellow at Aix-Marseille Université.

References

- [1] A. Balog, V. Blomer, C. Dartyge, and G. Tenenbaum, *Friable values of binary forms*, Comment. Math. Helv. 87 (2012), 639–667.
- [2] R. Crandall and C. Pomerance, *Prime Numbers. A Computational Perspective*, 2nd ed., Springer, New York, 2005.
- [3] R. de la Bretèche, *Sommes sans grand facteur premier*, Acta Arith. 88 (1999), 1–14.
- [4] R. de la Bretèche et A. Granville, *Densité des friables*, Bull. Soc. Math. France 142 (2014), 303–348.
- [5] S. Drappeau, *Sur les solutions friables de l'équation $a+b=c$* , Math. Proc. Cambridge Philos. Soc. 154 (2013), 439–463.
- [6] S. Drappeau, *Sommes friables d'exponentielles et applications*, Canad. J. Math. 67 (2015), 597–638.
- [7] É. Fouvry et G. Tenenbaum, *Entiers sans grand facteur premier en progressions arithmétiques*, Proc. London Math. Soc. (3) 63 (1991), 449–494.

- [8] N. Frantzikinakis and B. Host, *Asymptotics for multilinear averages of multiplicative functions*, Math. Proc. Cambridge Philos. Soc. 161 (2016), 87–101.
- [9] W. T. Gowers, *A new proof of Szemerédi’s theorem*, Geom. Funct. Anal. 11 (2001), 465–588.
- [10] A. Granville, *Smooth numbers: computational number theory and beyond*, in: Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography, Math. Sci. Res. Inst. Publ. 44, Cambridge Univ. Press, Cambridge, 2008, 267–323.
- [11] B. Green and T. Tao, *Linear equations in primes*, Ann. of Math. (2) 171 (2010), 1753–1850.
- [12] B. Green and T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) 175 (2012), 465–540.
- [13] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2) 175 (2012), 541–566.
- [14] B. Green, T. Tao, and T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$ -norm*, Ann. of Math. (2) 176 (2012), 1231–1372.
- [15] A. J. Harper, *Minor arcs, mean values, and restriction theory for exponential sums over smooth numbers*, Compos. Math. 152 (2016), 1121–1158.
- [16] A. Hildebrand, *On the number of positive integers $\leq x$ and free of prime factors $> y$* , J. Number Theory 22 (1986), 289–307.
- [17] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, J. Théor. Nombres Bordeaux 5 (1993), 411–484.
- [18] A. Lachand, *Fonctions arithmétiques et formes binaires irréductibles de degré 3*, Ann. Inst. Fourier (Grenoble), to appear.
- [19] A. Lachand, *Valeurs friables d’une forme quadratique et d’une forme linéaire*, Quart. J. Math. 66 (2015), 225–244.
- [20] A. Lachand et G. Tenenbaum, *Notes sur les valeurs moyennes criblées de certaines fonctions arithmétiques*, Quart. J. Math. 66 (2015), 245–250.
- [21] J. C. Lagarias and K. Soundararajan, *Counting smooth solutions to the equation $A + B = C$* , Proc. London Math. Soc. (3) 104 (2012), 770–798.
- [22] L. Matthiesen, *Generalized Fourier coefficients of multiplicative functions*, arXiv: 1405.1018 (2014).
- [23] L. Matthiesen, *Correlations of the divisor function*, Proc. London Math. Soc. (3) 104 (2012), 827–858.
- [24] L. Matthiesen, *Linear correlations amongst numbers represented by positive definite binary quadratic forms*, Acta Arith. 154 (2012), 235–306.
- [25] L. Matthiesen, *Correlations of representation functions of binary quadratic forms*, Acta Arith. 158 (2013), 245–252.
- [26] H. L. Montgomery and R. C. Vaughan, *Exponential sums with multiplicative coefficients*, Invent. Math. 43 (1977), 69–82.

Armand Lachand
Institut Élie Cartan
Université de Lorraine
B.P. 70239
54506 Vandœuvre-lès-Nancy Cedex, France
E-mail: armand.lachand@gmail.com

