PERTURBATIONS OF WEYL SUMS

TREVOR D. WOOLEY

Abstract. Write $f_k(\boldsymbol{\alpha};X)=\sum_{x\leqslant X}e(\alpha_1x+\ldots+\alpha_kx^k)$ $(k\geqslant 3)$. We show that there is a set $\mathfrak{B}\subseteq[0,1)^{k-2}$ of full measure with the property that whenever $(\alpha_2,\ldots,\alpha_{k-1})\in\mathfrak{B}$ and X is sufficiently large, then

$$\sup_{(\alpha_1,\alpha_k)\in[0,1)^2}|f_k(\boldsymbol{\alpha};X)|\leqslant X^{1/2+\delta_k},$$

where

$$\delta_k=\min\left\{\frac{13}{30},\frac{4}{2k-1}\right\}.$$

For $k\geqslant 5$, this improves on work of Flaminio and Forni, in which a Diophantine condition is imposed on α_k, and the exponent of X is $1-2/(3k(k-1))$.

1. Introduction

Consider the exponential sum $f_k(\boldsymbol{\alpha};X)$, defined for $k\geqslant 2$ and $\boldsymbol{\alpha}\in\mathbb{R}^k$ by

$$f_k(\boldsymbol{\alpha};X)=\sum_{1\leqslant x\leqslant X}e(\alpha_1x+\ldots+\alpha_kx^k),\tag{1.1}$$

where, as usual, we write $e(z)=e^{2\pi iz}$. It was shown by H. Weyl [8] that when α_k is irrational, then $\limsup X^{-1}|f_k(\boldsymbol{\alpha};X)|=0$ as $X\to\infty$. Indeed, when α_k satisfies an appropriate Diophantine condition, as is the case for algebraic irrational numbers such as $\sqrt{2}$, then for each $\varepsilon>0$, provided only that X is sufficiently large in terms of k and ε, one has the upper bound

$$|f_k(\boldsymbol{\alpha};X)|\leqslant X^{1-2^{1-k}+\varepsilon}.\tag{1.2}$$

Although such conclusions can be improved by employing the latest developments surrounding Vinogradov's mean value theorem (see, for example [9, Theorem 1.5]), the improved exponents remain very close to 1. Motivated by recent work of Flaminio and Forni [5] concerning equidistribution for higher step nilflows, in this paper we address two basic questions. First, we explore the extent to which the estimate (1.2) can be improved if one is prepared to exclude the perturbing coefficient tuple $(\alpha_1,\ldots,\alpha_{k-1})$ from a set of measure zero. Second, we examine how sensitive such estimates may be to the Diophantine conditions imposed on the lead coefficient α_k.

2010 *Mathematics Subject Classification.* 11L15, 11L07, 11P55.
Key words and phrases. Exponential sums, Hardy-Littlewood method.

Before proceeding further, we introduce some notation associated with Vinogradov's mean value theorem. With $f_k(\boldsymbol{\alpha};X)$ defined via (1.1), the *Main Conjecture* asserts that for all positive numbers s, one has

$$\int_{[0,1)^k} |f_k(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll X^\varepsilon(X^s+X^{2s-\frac{1}{2}k(k+1)}). \tag{1.3}$$

Here and throughout, the implicit constant in Vinogradov's notation may depend on k, s and the arbitrary positive number ε. We denote by $\mathrm{MC}_k(u)$ the assertion that the Main Conjecture (1.3) holds for $1\leqslant s\leqslant u$. We will be interested in the size of the exponential sum $f_k(\boldsymbol{\alpha};X)$ when the coefficients α_i are fixed for certain suffices $i=i_l$ $(1\leqslant l\leqslant t)$ with $1\leqslant i_1<i_2<\ldots<i_t\leqslant k$. The complementary set of suffices

$$\{1,2,\ldots,k\}\setminus\{i_1,i_2,\ldots,i_t\}=\{\iota_1,\iota_2,\ldots\iota_{k-t}\},$$

with $1\leqslant \iota_1<\iota_2<\ldots<\iota_{k-t}\leqslant k$, then corresponds to a $(k-t)$-tuple $(\alpha_{\iota_1},\ldots,\alpha_{\iota_{k-t}})$ that we permit to come from a set $\mathfrak{B}(\boldsymbol{\iota})\subseteq[0,1)^{k-t}$ that is central to our investigations. In order to facilitate concision, throughout this paper we write $\boldsymbol{\alpha}^*$ for $(\alpha_{i_1},\ldots,\alpha_{i_t})$ and $\boldsymbol{\alpha}^\dagger$ for $(\alpha_{\iota_1},\ldots,\alpha_{\iota_{k-t}})$.

Theorem 1.1. *Suppose that $k\geqslant 3$ and $1\leqslant u\leqslant \frac{1}{2}k(k+1)$, and assume $\mathrm{MC}_k(u)$. Let t be a positive integer with $1\leqslant t\leqslant k$, and let $\mathbf{i}$ be a t-tuple of suffices satisfying $1\leqslant i_1<i_2<\ldots<i_t\leqslant k$. Then there exists a set $\mathfrak{B}(\boldsymbol{\iota})\subseteq[0,1)^{k-t}$ of full measure such that, whenever $(\alpha_{\iota_1},\ldots,\alpha_{\iota_{k-t}})\in\mathfrak{B}(\boldsymbol{\iota})$, then for all real numbers X sufficiently large in terms of ε, k and $\boldsymbol{\alpha}^\dagger$, one has*

$$\sup_{\boldsymbol{\alpha}^*\in[0,1)^t}|f_k(\boldsymbol{\alpha};X)|\leqslant X^{1/2+\delta(\mathbf{i})+\varepsilon},$$

where

$$\delta(\mathbf{i})=\frac{t+1+2(i_1+\ldots+i_t)}{4u+2t+2}. \tag{1.4}$$

We extract two corollaries from Theorem 1.1 at the end of §2.

Corollary 1.2. *Suppose that $k\geqslant 3$. Then there exists a set $\mathfrak{B}\subseteq[0,1)^{k-2}$ of full measure such that, whenever $(\alpha_2,\alpha_3,\ldots,\alpha_{k-1})\in\mathfrak{B}$, then for all real numbers X sufficiently large in terms of k and $\alpha_2,\ldots,\alpha_{k-1}$, one has*

$$\sup_{(\alpha_1,\alpha_k)\in[0,1)^2}|f_k(\boldsymbol{\alpha};X)|\leqslant X^{1/2+\delta_k}, \tag{1.5}$$

where

$$\delta_k=\min\left\{\frac{13}{30},\frac{4}{2k-1}\right\}.$$

Moreover, when k is sufficiently large, the same conclusion holds with $\delta_k=1/k+o(1)$.

When k is large, the conclusion of Corollary 1.2 obtains very nearly square-root cancellation for the exponential sum $f_k(\boldsymbol{\alpha};X)$, greatly improving the estimate (1.2). In addition to this emphatic response to the first question posed in our opening paragraph, we note that no condition whatsoever has been imposed on the lead coefficient α_k. Of course, the restriction of the $(k-2)$-tuple

$(\alpha_2, \ldots, \alpha_{k-1})$ to the universal set $\mathfrak{B}$ of measure 1 implicitly imposes some sort of Diophantine condition on these lower order coefficients. Nonetheless, it is clear that there is in general little sensitivity to the lead coefficient.

Flaminio and Forni [5, Corollary 1.2] have derived a conclusion similar to that of Corollary 1.2 in which α_k is subject to a certain Diophantine condition, and the conclusion (1.5) holds with $\frac{1}{2} + \delta_k = 1 - 1/\left(\frac{3}{2}k(k-1)\right)$. Subject to a similar Diophantine condition on α_k, the latest progress on Vinogradov's mean value theorem permits the proof of a similar estimate with $2(k-1)(k-2)$ in place of $\frac{3}{2}k(k-1)$, though without any restriction on $(\alpha_2, \ldots, \alpha_{k-1})$ (simply substitute the conclusion of [13, Theorem 1.2] into the argument of the proof of [11, Theorem 11.1]). Thus, when k is large, the conclusion of Flaminio and Forni obtains barely non-trivial cancellation subject to a Diophantine condition, whereas Corollary 1.2 delivers nearly square-root cancellation.

We have aligned Corollary 1.2 so as to facilitate comparison with the work of Flaminio and Forni [5, Corollary 1.2]. When k is large, the conclusion of Theorem 1.1 offers estimates for $f_k(\boldsymbol{\alpha}; X)$ exhibiting close to square-root cancellation even when the number of fixed coefficients α_i is large. We illustrate such ideas with a further corollary.

Corollary 1.3. *Suppose that k is large, and that i_l $(1 \leqslant l \leqslant t)$ are integers with $1 \leqslant i_1 < i_2 < \ldots < i_t \leqslant k$. Suppose also that $2(i_1 + \ldots + i_t) + t + 1 < k^2/(\log k)$. Then there exists a set $\mathfrak{B}(\boldsymbol{\iota}) \subseteq [0,1)^{k-t}$ of full measure such that, whenever $\boldsymbol{\alpha}^\dagger \in \mathfrak{B}(\boldsymbol{\iota})$, then for all X sufficiently large in terms of k and $\boldsymbol{\alpha}^\dagger$, one has*

$$\sup_{\boldsymbol{\alpha}^* \in [0,1)^t} |f_k(\boldsymbol{\alpha}; X)| \leqslant X^{1/2 + 1/\log k}.$$

The conclusion of Corollary 1.3 shows that, in a suitable sense, almost a positive proportion of the coefficients of $f_k(\boldsymbol{\alpha}; X)$ can be fixed, and yet one nonetheless achieves nearly square-root cancellation on a universal set of full measure for the remaining coefficients.

Our methods extend naturally to deliver equidistribution results for polynomials modulo 1. In this context, when $0 \leqslant a < b \leqslant 1$, we write $Z_{a,b}(\boldsymbol{\alpha}; N)$ for the number of integers n with $1 \leqslant n \leqslant N$ for which

$$a \leqslant \alpha_1 n + \alpha_2 n^2 + \ldots + \alpha_k n^k \leqslant b \pmod 1.$$

Theorem 1.4. *Suppose that $k \geqslant 3$ and $1 \leqslant u \leqslant \frac{1}{2}k(k+1)$, and assume $\mathrm{MC}_k(u)$. Let t be a positive integer with $1 \leqslant t \leqslant k$, and let $\mathbf{i}$ be a t-tuple of suffices satisfying $1 \leqslant i_1 < i_2 < \ldots < i_t \leqslant k$. Then there exists a set $\mathfrak{B}(\boldsymbol{\iota}) \subseteq [0,1)^{k-t}$ of full measure such that, whenever $(\alpha_{\iota_1}, \ldots, \alpha_{\iota_{k-t}}) \in \mathfrak{B}(\boldsymbol{\iota})$, then for all real numbers N sufficiently large in terms of ε, k and $\boldsymbol{\alpha}^\dagger$, one has*

$$|Z_{a,b}(\boldsymbol{\alpha}; N) - (b-a)N| \leqslant N^{1/2 + \nu(\mathbf{i}) + \varepsilon} \quad (0 \leqslant a < b \leqslant 1),$$

where

$$\nu(\mathbf{i}) = \frac{t + 2 + 2(i_1 + \ldots + i_t)}{4u + 2t + 4}. \tag{1.6}$$

We derive the following corollary to Theorem 1.4 at the end of §3. Define

$$\nu_k = \min\left\{\frac{11}{30}, \frac{2}{k}\right\}.$$

Corollary 1.5. *Suppose that $k \geqslant 3$. Then there exists a set $\mathfrak{B} \subseteq [0,1)^{k-1}$ of full measure such that, whenever $(\alpha_1, \alpha_2, \ldots, \alpha_{k-1}) \in \mathfrak{B}$, then for all real numbers N sufficiently large in terms of k and $\alpha_1, \ldots, \alpha_{k-1}$, one has*

$$|Z_{a,b}(\boldsymbol{\alpha}; N) - (b-a)N| \leqslant N^{1/2+\nu_k} \quad (0 \leqslant a < b \leqslant 1).$$

Write $\|\theta\| = \min\{|\theta - m| : m \in \mathbb{Z}\}$. Then by putting $a = 0$ and $b = 2N^{-1/2+\nu_k}$, we obtain as a special case of Corollary 1.5 the following conclusion.

Corollary 1.6. *Suppose that $k \geqslant 3$. Then there exists a set $\mathfrak{B}^* \subseteq [0,1)^{k-1}$ of full measure such that, whenever $(\alpha_1, \ldots, \alpha_{k-1}) \in \mathfrak{B}^*$, then for all real numbers N sufficiently large in terms of k and $\alpha_1, \ldots, \alpha_{k-1}$, one has*

$$\min_{1 \leqslant n \leqslant N} \|\alpha_1 n + \ldots + \alpha_k n^k\| \ll N^{-1/2+\nu_k}. \tag{1.7}$$

There are results available in the literature analogous to (1.7) in which $(\alpha_1, \ldots, \alpha_{k-1})$ is a fixed real $(k-1)$-tuple. Thus one finds that the conclusions of [2, Theorem 5.2] and [11, Theorem 11.3] (as enhanced by utilising [13, Theorem 1.2]) yield an estimate of the shape (1.7) with the exponent $\frac{1}{2} + \nu_k$ replaced by any real number exceeding $1 - 1/\min\{4(k-1)(k-2), 2^{k-1}\}$. These uniform results are considerably weaker than those available via Corollary 1.6.

In contrast to the ergodic methods employed by Flaminio and Forni [5], in this paper we utilise recent progress on Vinogradov's mean value theorem. Of critical importance to us are mean value estimates of the shape

$$\int_{[0,1)^k} |f_k(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll X^{s+\delta}, \tag{1.8}$$

with δ small and s large. Prior to the author's introduction of "efficient congruencing" methods in 2012 (see [9]), available estimates were far too weak to deliver conclusions of the type described in Corollary 1.2. However, the estimate (1.8) is established in [10, Corollary 1.3] with $\delta = 1+\varepsilon$ for $1 \leqslant s \leqslant \frac{1}{4}k^2+k$, and this would suffice for our purposes in the present paper. Recent work of Ford [6, Theorem 1.1] joint with the author establishes (1.8) for any $\delta > 0$ in the same range of s, and even more recently the author [12, Theorem 1.3] has extended the permissible range of s to $1 \leqslant s \leqslant \frac{1}{2}k(k+1) - \frac{1}{3}k + o(k)$, encompassing nearly the whole of the critical interval.

Let X and T be large, and consider a fixed t-tuple $(\alpha_{i_1}, \ldots, \alpha_{i_t}) \in [0,1)^t$. The estimate (1.8) permits one to estimate the measure of the set $\mathfrak{B}_T(X)$ of $(k-t)$-tuples $(\alpha_{\iota_1}, \ldots \alpha_{\iota_{k-t}}) \in [0,1)^{k-t}$ for which $|f_k(\boldsymbol{\alpha}; X)| > T$. Suppose that T is chosen as a function of X for which $\sum_{X=1}^{\infty} \mathrm{mes}(\mathfrak{B}_T(X)) < \infty$, and define $\mathfrak{B}^* \subseteq [0,1)^{k-t}$ to be the set of $(k-t)$-tuples $(\alpha_{\iota_1}, \ldots, \alpha_{\iota_{k-t}})$ for which $\limsup T^{-1}|f_k(\boldsymbol{\alpha}; X)| \geqslant 1$ as $X \to \infty$. Then it follows from the Borel-Cantelli theorem that the set $\mathfrak{B}^*$ has measure 0. One may remove the dependence of these estimates on the fixed t-tuple of coefficients $(\alpha_{i_1}, \ldots, \alpha_{i_t})$ by a suitable application of the mean value theorem, showing that the size of $|f_k(\boldsymbol{\alpha};X)|$ changes little as α_j varies over an interval having length of order X^{-j}. Moreover, we are able to sharpen our estimates by observing that $|f_k(\boldsymbol{\alpha};X)|$ also changes little as X varies over an interval of length small compared to T.

We remark that Pustyl′nikov has work spanning a number of papers (see, for example [7]) which derives conclusions related to those of this paper. Pustyl′nikov makes use of the estimate (1.8) in the classical case $s=k$. In this special case, one may apply Newton's formulae relating symmetric polynomials with the roots of polynomials to derive the formula

$$\int_{[0,1)^k} |f_k(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \sim s!X^s.$$

The point of view taken in [7] is that by taking k sufficiently large, one may gain some control of the value distribution of Weyl sums $f_k(\boldsymbol{\alpha};X)$. The relative strength of the conclusions made available in the present paper rests on the far more powerful mean value estimates stemming from our recent work on Vinogradov's mean value theorem.

Our basic parameter is X, a sufficiently large positive number. In this paper, implicit constants in Vinogradov's notation $\ll$ and $\gg$ may depend on k, u and ε. Whenever ε appears in a statement, either implicitly or explicitly, we assert that the statement holds for each $\varepsilon>0$. We use vector notation in the natural way. When $\mathfrak{A}\subset\mathbb{R}$ is Lebesgue measurable, we write $\mu(\mathfrak{A})$ for its measure. Finally, we write $[\theta]$ for $\max\{n\in\mathbb{Z}:n\leqslant\theta\}$.

The author is grateful to Professors Flaminio and Forni for discussions concerning the problems addressed in this paper, and in particular for providing the author with an early version of their paper [5]. These discussions benefitted from the excellent working conditions and support provided by the Isaac Newton Institute in Cambridge during the program "Interactions between Dynamics of Group Actions and Number Theory" in June 2014. We also thank the referee for useful comments and suggestions.

2. Large values of Weyl sums

Our goal in this section is the proof of Theorem 1.1 and its corollaries. We begin our analysis of $f_k(\boldsymbol{\alpha};X)$ by showing that the magnitude of this Weyl sum changes little when its argument is modified by a small quantity.

Lemma 2.1. *Let $T>0$ and $\boldsymbol{\alpha}\in\mathbb{R}^k$, and suppose that $|f_k(\boldsymbol{\alpha};X)|>T$. Then whenever $\boldsymbol{\beta}\in\mathbb{R}^k$ satisfies*

$$|\beta_j-\alpha_j|\leqslant(4\pi k)^{-1}TX^{-j-1}\quad(1\leqslant j\leqslant k),$$

one has $|f_k(\boldsymbol{\beta};X)|>\frac{1}{2}T$.

Proof. Under the hypotheses of the statement of the lemma, an application of the multidimensional mean value theorem (see [1, Theorem 6-17]) shows that there exists a point $\boldsymbol{\gamma}$ on the line segment connecting $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that

$$\begin{aligned} f_k(\boldsymbol{\beta};X)-f_k(\boldsymbol{\alpha};X)&=\sum_{j=1}^k(\beta_j-\alpha_j)\frac{\partial}{\partial\gamma_j}f_k(\boldsymbol{\gamma};X)\\ &=2\pi i\sum_{j=1}^k(\beta_j-\alpha_j)\sum_{1\leqslant x\leqslant X}x^je(\gamma_1x+\ldots+\gamma_kx^k).\end{aligned}$$

Thus, by making a trivial estimate for the exponential sum defined by the inner summation here, we deduce that

$$\begin{aligned}|f_k(\boldsymbol{\beta};X)|&\geqslant|f_k(\boldsymbol{\alpha};X)|-2\pi\sum_{j=1}^k|\beta_j-\alpha_j|X^{j+1}\\ &>T-(2k)^{-1}\sum_{j=1}^kT=\tfrac12T.\end{aligned}$$

This completes the proof of the lemma. $\square$

We suppose now that i_l $(1\leqslant l\leqslant t)$ are suffices with $1\leqslant i_1<\ldots<i_t\leqslant k$, and we recall the notation introduced in the preamble to the statement of Theorem 1.1 above. It is convenient to write

$$\sigma(\mathbf{i})=i_1+i_2+\ldots+i_t.$$

Our initial objective is to obtain an estimate for the set

$$\mathfrak{B}_T(X)=\{\boldsymbol{\alpha}^\dagger\in[0,1)^{k-t}:|f_k(\boldsymbol{\alpha};X)|>T\text{ for some }\boldsymbol{\alpha}^*\in[0,1)^t\}.\tag{2.1}$$

Lemma 2.2. *Suppose that* $1\leqslant u\leqslant\frac12k(k+1)$, *and assume the hypothesis* $\mathrm{MC}_k(u)$. *Then whenever* T *is a real number with* $0<T\leqslant X$, *one has*

$$\mu(\mathfrak{B}_T(X))\ll X^{u+t+\sigma(\mathbf{i})+\varepsilon}T^{-2u-t}.$$

Proof. For $1\leqslant l\leqslant t$, put

$$\delta_l=(4\pi k)^{-1}TX^{-i_l-1}\quad\text{and}\quad M_l=[\delta_l^{-1}].$$

When $0\leqslant m_l\leqslant M_l$ $(1\leqslant l\leqslant t)$, we define the hypercuboids

$$\mathcal{I}(\mathbf{m})=[m_1\delta_1,(m_1+1)\delta_1]\times\ldots\times[m_t\delta_t,(m_t+1)\delta_t]$$

and

$$\mathcal{M}=[0,M_1]\times\ldots\times[0,M_t].$$

Finally, for each $\mathbf{m}\in\mathcal{M}$, we put

$$\mathfrak{B}_T(\mathbf{m};X)=\{\boldsymbol{\alpha}^\dagger\in[0,1)^{k-t}:|f_k(\boldsymbol{\alpha};X)|>T\text{ for some }\boldsymbol{\alpha}^*\in\mathcal{I}(\mathbf{m})\}.$$

Since $[0,1)^t$ is contained in the union of the sets $\mathcal{I}(\mathbf{m})$ for $\mathbf{m}\in\mathcal{M}$, we see that

$$\mathfrak{B}_T(X)=\bigcup_{\mathbf{m}\in\mathcal{M}}\mathfrak{B}_T(\mathbf{m};X).\tag{2.2}$$

Observe next that when $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ both lie in $\mathcal{I}(\mathbf{m})$ for some $\mathbf{m}\in\mathcal{M}$, then

$$|\alpha_{i_l}-\beta_{i_l}|\leqslant\delta_l=(4\pi k)^{-1}TX^{-i_l-1}\quad(1\leqslant l\leqslant t).$$

Thus we deduce from Lemma 2.1 that whenever $\boldsymbol{\alpha}^\dagger \in \mathfrak{B}_T(\mathbf{m};X)$ for some $\mathbf{m} \in \mathcal{M}$, then $|f_k(\boldsymbol{\alpha};X)| > \frac{1}{2}T$ for all $\boldsymbol{\alpha}^* \in \mathcal{I}(\mathbf{m})$. It follows that

$$\left(\tfrac{1}{2}T\right)^{2u} \mu(\mathfrak{B}_T(\mathbf{m};X))\mu(\mathcal{I}(\mathbf{m})) < \int_{\mathcal{I}(\mathbf{m})} \int_{\mathfrak{B}_T(\mathbf{m};X)} |f_k(\boldsymbol{\alpha};X)|^{2u} \, \mathrm{d}\boldsymbol{\alpha}^\dagger \, \mathrm{d}\boldsymbol{\alpha}^*.$$

But $\mu(\mathcal{I}(\mathbf{m})) = \delta_1 \cdots \delta_t \gg (T/X)^t X^{-\sigma(\mathbf{i})}$, and thus

$$T^{2u+t} X^{-t-\sigma(\mathbf{i})} \mu(\mathfrak{B}_T(\mathbf{m};X)) \ll \int_{\mathcal{I}(\mathbf{m})} \int_{[0,1)^{k-t}} |f_k(\boldsymbol{\alpha};X)|^{2u} \, \mathrm{d}\boldsymbol{\alpha}^\dagger \, \mathrm{d}\boldsymbol{\alpha}^*.$$

Consequently, on recalling (2.2), one arrives at the upper bound

$$\begin{aligned} \mu(\mathfrak{B}_T(X)) &\leqslant \sum_{\mathbf{m}\in\mathcal{M}} \mu(\mathfrak{B}_T(\mathbf{m};X)) \\ &\ll T^{-2u-t} X^{t+\sigma(\mathbf{i})} \sum_{\mathbf{m}\in\mathcal{M}} \int_{\mathcal{I}(\mathbf{m})} \int_{[0,1)^{k-t}} |f_k(\boldsymbol{\alpha};X)|^{2u} \, \mathrm{d}\boldsymbol{\alpha}^\dagger \, \mathrm{d}\boldsymbol{\alpha}^*. \end{aligned}$$

Since the union of the sets $\mathcal{I}(\mathbf{m})$ with $\mathbf{m} \in \mathcal{M}$ is contained in $[0,2)^t$, we reach the point at which we may utilise $\mathrm{MC}_k(u)$, obtaining the estimate

$$\begin{aligned} \mu(\mathfrak{B}_T(X)) &\ll T^{-2u-t} X^{t+\sigma(\mathbf{i})} \int_{[0,2)^k} |f_k(\boldsymbol{\alpha};X)|^{2u} \, \mathrm{d}\boldsymbol{\alpha} \\ &\ll T^{-2u-t} X^{t+\sigma(\mathbf{i})} \cdot 2^k X^{u+\varepsilon}. \end{aligned}$$

The conclusion of the lemma is now immediate. □

We next make a choice for T. Let τ be a positive number, and put

$$T(X) = X^{1/2+\delta(\mathbf{i})+\tau},$$

where $\delta(\mathbf{i})$ is defined as in (1.4). Here we note that

$$\tfrac{1}{2} + \delta(\mathbf{i}) = \frac{(2u+t+1) + (t+1+2\sigma(\mathbf{i}))}{4u+2t+2} = \frac{u+t+1+\sigma(\mathbf{i})}{2u+t+1}. \tag{2.3}$$

Let $(X_n)_{n=1}^\infty$ be any strictly increasing sequence of natural numbers with the property that for large enough values of n, one has

$$T(X_n) \leqslant X_{n+1} - X_n \leqslant 2T(X_n), \tag{2.4}$$

and in the interests of concision, write $T_n = T(X_n)$. For each non-negative integer j, let n_j denote the least integer with the property that

$$X_{n_j} \geqslant 2^j. \tag{2.5}$$

Since the sequence $(X_n)_{n=1}^\infty$ is strictly increasing, it is trivial that $n_j \leqslant 2^j$ for each j. Moreover, when n is large, it follows from (2.4) that whenever $m(n) \geqslant X_n/T_n$, then $X_{n+m(n)} - X_n \geqslant (X_n/T_n)T_n$, whence $X_{n+m(n)} \geqslant 2X_n$. Thus, there is a natural number j_0 with the property that, when $j \geqslant j_0$, then

$$n_{j+1} - n_j \leqslant 2X_{n_j}/T_{n_j}. \tag{2.6}$$

Lemma 2.3. *Suppose that $1 \leqslant u \leqslant \frac{1}{2}k(k+1)$ and assume* $\mathrm{MC}_k(u)$*. Then for any sequence $(X_n)_{n=1}^\infty$ satisfying* (2.4)*, one has*

$$\mu\left(\bigcup_{n=1}^\infty \mathfrak{B}_{T_n}(X_n)\right) < \infty.$$

Proof. On noting the relation (2.3), we find from Lemma 2.2 that

$$\sum_{n=1}^\infty \mu(\mathfrak{B}_{T_n}(X_n)) \ll \sum_{n=1}^\infty X_n^{u+t+\sigma(\mathbf{i})+\varepsilon} T_n^{-2u-t} \leqslant \sum_{n=1}^\infty (T_n/X_n) X_n^{\varepsilon-2\tau}.$$

But when $n_j \leqslant n < n_{j+1}$, one has $X_{n_j} \leqslant X_n < X_{n_{j+1}}$, and thus we infer from (2.5) and (2.6) that

$$\begin{aligned}\mu\left(\bigcup_{n=1}^\infty \mathfrak{B}_{T_n}(X_n)\right) &\ll \sum_{j=0}^\infty \sum_{n_j\leqslant n<n_{j+1}} (T_n/X_n)X_n^{\varepsilon-2\tau}\\ &\ll 1+\sum_{j\geqslant j_0}(n_{j+1}-n_j)(T_{n_j}/X_{n_j})X_{n_j}^{\varepsilon-2\tau}\\ &\ll 1+\sum_{j=0}^\infty (2^j)^{\varepsilon-2\tau}<\infty.\end{aligned}$$

This completes the proof of the lemma. $\square$

We are now equipped to complete the proof of Theorem 1.1. Denote by $A_n(\boldsymbol{\alpha}^\dagger)$ the condition that $|f_k(\boldsymbol{\alpha};X_n)| > T_n$ for some $\boldsymbol{\alpha}^* \in [0,1)^t$. Then the definition (2.1) of $\mathfrak{B}_T(X)$ implies that

$$\mathfrak{B}_{T_n}(X_n) = \{\boldsymbol{\alpha}^\dagger \in [0,1)^{k-t} : A_n(\boldsymbol{\alpha}^\dagger)\}.$$

Put

$$\mathfrak{B}^* = \{\boldsymbol{\alpha}^\dagger \in [0,1)^{k-t} : A_n(\boldsymbol{\alpha}^\dagger) \text{ holds for infinitely many } n\in\mathbb{N}\}.$$

Then it follows from Lemma 2.3 via the Borel-Cantelli lemma that $\mu(\mathfrak{B}^*) = 0$. Consequently, there is a set $\mathfrak{B}_0 = [0,1)^{k-t}\setminus\mathfrak{B}^*$ of full measure having the property that, whenever $\boldsymbol{\alpha}^\dagger \in \mathfrak{B}_0$, then $A_n(\boldsymbol{\alpha}^\dagger)$ holds for at most finitely many $n\in\mathbb{N}$. The latter assertion implies that $|f_k(\boldsymbol{\alpha};X_n)| \leqslant T_n$ for all $\boldsymbol{\alpha}^*\in[0,1)^t$, with the exception of at most finitely many $n\in\mathbb{N}$.

Suppose that $X>0$, and put $X^*=[X]$, so that $f_k(\boldsymbol{\alpha};X)=f_k(\boldsymbol{\alpha};X^*)$. In view of the condition (2.4), when X is sufficiently large there exists $n\in\mathbb{N}$ for which $X_n \leqslant X^* \leqslant X_n+2T_n$. But then, on making a trivial estimate for the exponential function, we have

$$|f_k(\boldsymbol{\alpha};X)-f_k(\boldsymbol{\alpha};X_n)| \leqslant X-X_n \leqslant 2T_n.$$

Whenever $|f_k(\boldsymbol{\alpha};X_n)|\leqslant T_n$, therefore, one finds that

$$|f_k(\boldsymbol{\alpha};X)| \leqslant 3T_n = 3X_n^{1/2+\delta(\mathbf{i})+\tau} \leqslant 3X^{1/2+\delta(\mathbf{i})+\tau}.$$

Then we may conclude that whenever $\boldsymbol{\alpha}^\dagger\in\mathfrak{B}_0$, then for all positive numbers X, one has $|f_k(\boldsymbol{\alpha};X)| \leqslant 3X^{1/2+\delta(\mathbf{i})+\tau}$ for all $\boldsymbol{\alpha}^*\in[0,1)^t$, with the exception of at most those numbers X lying in a bounded interval $(0, X_0]$. Since $\tau > 0$ may be taken arbitrarily small, the conclusion of Theorem 1.1 follows.

The corollaries to Theorem 1.1 are easily confirmed. On the one hand, when $k \geqslant 4$, we find from [6, Theorem 1.1] that $\mathrm{MC}_k(u)$ holds for $u = \left[\frac{1}{4}(k+1)^2\right]$. On the other hand, from [12, Theorem 1.3], one obtains $\mathrm{MC}_k(u)$ when k is large and $u = [\frac{1}{2}k(k+1) - \frac{1}{3}k - 8k^{2/3}]$. In addition, [13, Theorem 1.1] furnishes $\mathrm{MC}_3(6)$. In order to establish Corollary 1.2, we apply Theorem 1.1 with $\mathbf{i} = (1, k)$. In such circumstances, we have $t = 2$ and

$$\delta(\mathbf{i}) = \frac{3 + 2(k+1)}{4u + 6}.$$

Thus, when $k = 3$ one may take $\delta(\mathbf{i}) = 11/30$, and when $k \geqslant 4$ one may take

$$\delta(\mathbf{i}) = \frac{2k+5}{(k^2+2k)+6} \leqslant \min\left\{\frac{13}{30}, \frac{2}{k-1/2}\right\}.$$

In addition, for large k we may instead take

$$\delta(\mathbf{i}) = \frac{2k+5}{2k(k+1) - \frac{4}{3}k + o(k)} = \frac{1}{k - \frac{13}{6} + o(1)} = \frac{1}{k} + o(1).$$

In all situations, we conclude from Theorem 1.1 that there exists a set $\mathfrak{B} \subseteq [0,1)^{k-2}$ of full measure such that, when $(\alpha_2, \alpha_3, \ldots, \alpha_{k-1}) \in \mathfrak{B}$, then for all real numbers X sufficiently large in terms of ε, k and $\alpha_2, \ldots, \alpha_{k-1}$, one has

$$\sup_{(\alpha_1, \alpha_k) \in [0,1)^2} |f_k(\boldsymbol{\alpha}; X)| \leqslant X^{1/2 + \delta(\mathbf{i}) + \varepsilon}.$$

This confirms both of the conclusions of Corollary 1.2.

We turn next to Corollary 1.3. Taking $\mathbf{i} = (i_1, \ldots, i_t)$ and $u = \left[\frac{1}{4}(k+1)^2\right]$, we find that the conclusion of Theorem 1.1 holds with

$$\delta(\mathbf{i}) = \frac{t + 1 + 2(i_1 + \ldots + i_t)}{4[\frac{1}{4}(k+1)^2] + 2t + 2} < \frac{k^2/\log k}{k^2 + 2k + 2t + 2} < \frac{1}{\log k}.$$

Consequently, there exists a set $\mathfrak{B} \subseteq [0,1)^{k-t}$ of full measure such that, when $\boldsymbol{\alpha}^\dagger \in \mathfrak{B}$, then for all real numbers X sufficiently large in terms of k and $\boldsymbol{\alpha}^\dagger$, one has $\sup_{\boldsymbol{\alpha}^* \in [0,1)^t} |f_k(\boldsymbol{\alpha}; X)| \leqslant X^{1/2 + 1/\log k}$. This confirms Corollary 1.3.

3. Equidistribution of polynomials modulo one

We investigate the equidistribution of polynomial sequences by applying the Erdős-Turán inequality (see [3, 4]). This entails estimating the exponential sum $f_k(h\boldsymbol{\alpha}; X)$ for $1 \leqslant h \leqslant H$, with H as large as is feasible. Suppose once more that i_l $(1 \leqslant l \leqslant t)$ are suffices with $1 \leqslant i_1 < i_2 < \ldots < i_t \leqslant k$, with the conventions in the preamble to the statement of Theorem 1.1. When $h \in \mathbb{N}$, we define a set generalising that defined in (2.1) by putting

$$\mathfrak{B}_T^{(h)}(X) = \{\boldsymbol{\alpha}^\dagger \in [0,1)^{k-t} : |f_k(h\boldsymbol{\alpha}; X)| > T \text{ for some } \boldsymbol{\alpha}^* \in [0,1)^t\}.$$

Thus we have

$$\mathfrak{B}_T^{(h)}(X) = \{\boldsymbol{\alpha}^\dagger \in [0,1)^{k-t} : h\boldsymbol{\alpha}^\dagger \in \mathfrak{B}_T(X) \pmod 1\}. \tag{3.1}$$

When $\lambda,\mu\in\mathbb{R}$ and $\mathcal{A}\subseteq\mathbb{R}$, denote by $\lambda(\mathcal{A}+\mu)$ the set $\{\lambda(\theta+\mu):\theta\in\mathcal{A}\}$. Then it follows from (3.1) that

$$\mathfrak{B}_T^{(h)}(X)=\bigcup_{m=0}^{h-1}h^{-1}(\mathfrak{B}_T(X)+m),$$

and hence $\mu(\mathfrak{B}_T^{(h)}(X))=\mu(\mathfrak{B}_T(X))$ for $h\in\mathbb{N}$.

We next introduce the set $\mathfrak{C}_T(X,H)$ consisting of those points $\boldsymbol{\alpha}^\dagger\in[0,1)^{k-t}$ for which one has $|f_k(h\boldsymbol{\alpha};X)|>T$ for some $\boldsymbol{\alpha}^*\in[0,1)^t$ and $h\in\mathbb{N}$ with $1\leqslant h\leqslant H$. Then we have

$$\mathfrak{C}_T(X,H)=\bigcup_{1\leqslant h\leqslant H}\mathfrak{B}_T^{(h)}(X),$$

so that

$$\mu(\mathfrak{C}_T(X,H))\leqslant\sum_{1\leqslant h\leqslant H}\mu(\mathfrak{B}_T^{(h)}(X))\leqslant H\mu(\mathfrak{B}_T(X)).$$

We therefore deduce from Lemma 2.2 that when $1\leqslant u\leqslant\frac{1}{2}k(k+1)$ and $\mathrm{MC}_k(u)$ holds, then one has

$$\mu(\mathfrak{C}_T(X,H))\ll HX^{u+t+\sigma(\mathbf{i})+\varepsilon}T^{-2u-t}. \tag{3.2}$$

We now make a choice for T and H. Let τ be a positive number, and put

$$H(X)=X^{1/2-\nu(\mathbf{i})-2\tau}\quad\text{and}\quad T(X)=X^{1/2+\nu(\mathbf{i})+\tau},$$

where $\nu(\mathbf{i})$ is defined as in (1.6). Note that

$$\tfrac{1}{2}+\nu(\mathbf{i})=\frac{(2u+t+2)+(t+2+2\sigma(\mathbf{i}))}{4u+2t+4}=\frac{u+t+2+\sigma(\mathbf{i})}{2u+t+2}. \tag{3.3}$$

We again consider a sequence of natural numbers $(X_n)_{n=1}^\infty$ satisfying the condition (2.4), and then write $T_n=T(X_n)$ and $H_n=H(X_n)$.

Lemma 3.1. *Suppose that* $1\leqslant u\leqslant\frac{1}{2}k(k+1)$ *and assume* $\mathrm{MC}_k(u)$. *Then for any sequence* $(X_n)_{n=1}^\infty$ *satisfying* (2.4), *one has*

$$\mu\left(\bigcup_{n=1}^\infty\mathfrak{C}_{T_n}(X_n,H_n)\right)<\infty.$$

Proof. We recall the infrastructure associated with the sequence $(X_n)_{n=1}^\infty$ introduced in the preamble to the statement of Lemma 2.3. In view of the relation (3.3), it follows from (3.2) that

$$\sum_{n=1}^\infty\mu(\mathfrak{C}_{T_n}(X_n,H_n))\ll\sum_{n=1}^\infty H_nX_n^{u+t+\sigma(\mathbf{i})+\varepsilon}T_n^{-2u-t}\leqslant\sum_{n=1}^\infty(H_nT_n^2/X_n^2)X_n^{\varepsilon-2\tau}.$$

Note again that when $n_j \leqslant n < n_{j+1}$, one has $2^j \leqslant X_n \leqslant 2^{j+1}$. Thus, since $T_n H_n \leqslant X_n$, we may infer from (2.5) and (2.6) on this occasion that

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} \mathfrak{C}_{T_n}(X_n, H_n)\right) &\ll \sum_{j=0}^{\infty} \sum_{n_j \leqslant n < n_{j+1}} (T_n/X_n) X_n^{\varepsilon - 2\tau} \\ &\ll 1 + \sum_{j \geqslant j_0} (n_{j+1} - n_j)(T_{n_j}/X_{n_j}) X_{n_j}^{\varepsilon - 2\tau} \\ &\ll 1 + \sum_{j=0}^{\infty} (2^j)^{\varepsilon - 2\tau} < \infty.\end{aligned}$$

This completes the proof of the lemma. $\square$

Denote by $B_n(\boldsymbol{\alpha}^\dagger)$ the condition that $|f_k(h\boldsymbol{\alpha}; X_n)| > T_n$ for some $\boldsymbol{\alpha}^* \in [0,1)^t$ and $h \in \mathbb{N}$ with $1 \leqslant h \leqslant H_n$. Then the definition of $\mathfrak{C}_T(X, H)$ implies that

$$\mathfrak{C}_{T_n}(X_n, H_n) = \{\boldsymbol{\alpha}^\dagger \in [0,1)^{k-t} : B_n(\boldsymbol{\alpha}^\dagger)\}.$$

Put

$$\mathfrak{C}^* = \{\boldsymbol{\alpha}^\dagger \in [0,1)^{k-t} : B_n(\boldsymbol{\alpha}^\dagger) \text{ holds for infinitely many } n \in \mathbb{N}\}.$$

Then it follows from Lemma 3.1 via the Borel-Cantelli lemma that $\mu(\mathfrak{C}^*) = 0$. Consequently, there is a set $\mathfrak{C}_0 = [0,1)^{k-t} \setminus \mathfrak{C}^*$ of full measure having the property that, whenever $\boldsymbol{\alpha}^\dagger \in \mathfrak{C}_0$, then $B_n(\boldsymbol{\alpha}^\dagger)$ holds for at most finitely many $n \in \mathbb{N}$. The latter implies that $|f_k(h\boldsymbol{\alpha}; X_n)| \leqslant T_n$ for all $\boldsymbol{\alpha}^* \in [0,1)^t$ and all $h \in \mathbb{N}$ with $1 \leqslant h \leqslant H_n$, with the exception of at most finitely many $n \in \mathbb{N}$.

As in the corresponding treatment of §2, the condition (2.4) ensures that when $\boldsymbol{\alpha}^\dagger \in \mathfrak{C}_0$, then for all $X > 0$ and $h \in \mathbb{N}$ with $h \leqslant X^{1/2 - \nu(\mathbf{i}) - 2\tau}$, one has

$$\sup_{\boldsymbol{\alpha}^* \in [0,1)^t} |f_k(h\boldsymbol{\alpha}; X)| \leqslant 3X^{1/2 + \nu(\mathbf{i}) + \tau}, \tag{3.4}$$

except perhaps for certain numbers X lying in a bounded interval $[0, X_0)$.

The estimate (3.4) provides our basic input for an application of the Erdős-Turán inequality, as decribed in [2, Theorem 2.1]. Suppose that $0 \leqslant a < b \leqslant 1$. Also, write $x_n = \alpha_k n^k + \ldots + \alpha_1 n$ and put $H = X^{1/2 - \nu(\mathbf{i}) - 2\tau}$. Then

$$\begin{aligned}\Bigg| \sum_{\substack{1 \leqslant n \leqslant X \\ x_n \in [a,b] \pmod 1}} 1 - X(b-a) \Bigg| &\leqslant \frac{X}{H+1} + 3 \sum_{1 \leqslant h \leqslant H} h^{-1} \Bigg| \sum_{1 \leqslant n \leqslant X} e(hx_n) \Bigg| \\ &= \frac{X}{H+1} + 3 \sum_{1 \leqslant h \leqslant H} h^{-1} |f_k(h\boldsymbol{\alpha}; X)|.\end{aligned}$$

Consequently, when $\boldsymbol{\alpha}^\dagger \in \mathfrak{C}_0$, one finds from (3.4) that

$$|Z_{a,b}(\boldsymbol{\alpha}; X) - X(b-a)| \leqslant XH^{-1} + 9 \sum_{1 \leqslant h \leqslant H} h^{-1} X^{1/2 + \nu(\mathbf{i}) + \tau} \ll X^{1/2 + \nu(\mathbf{i}) + 2\tau}.$$

Since $\tau > 0$ may be taken arbitarily small, Theorem 1.4 now follows.

The proof of Corollary 1.5 follows on taking $\mathbf{i} = (k)$ and
$$u = \max\left\{6, \left[\tfrac{1}{4}(k+1)^2\right]\right\},$$
so that the conclusion of Theorem 1.4 holds with $\nu(\mathbf{i}) = 3/10$ when $k = 3$, and otherwise, when $k \geqslant 4$, with
$$\nu(\mathbf{i}) = \frac{3+2k}{4u+6} \leqslant \frac{2k+3}{k^2+2k+6} \leqslant \min\left\{\frac{11}{30}, \frac{2}{k}\right\}.$$
Then we conclude that there exists a set $\mathfrak{B}^* \subseteq [0,1)^{k-1}$ of full measure with the property that, whenever $(\alpha_1, \ldots, \alpha_{k-1}) \in \mathfrak{B}^*$, then for all $N \in \mathbb{N}$ sufficiently large in terms of k and $\alpha_1, \ldots, \alpha_{k-1}$, one has
$$|Z_{a,b}(\boldsymbol{\alpha}; N) - (b-a)N| \leqslant N^{1/2+\nu_k} \quad (0 \leqslant a < b \leqslant 1).$$
This completes the proof of Corollary 1.5.

References

[1] T. M. Apostol, *Mathematical analysis: a modern approach to advanced calculus*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1957.

[2] R. C. Baker, *Diophantine inequalities*, London Mathematical Society Monographs, New Series, **1**, Oxford University Press, New York, 1986.

[3] P. Erdős and P. Turán, *On a problem in the theory of uniform distribution. I*, Nederl. Akad. Wetensch. Proc. **51** (1948), 1146–1154.

[4] P. Erdős and P. Turán, *On a problem in the theory of uniform distribution. II*, Nederl. Akad. Wetensch. Proc. **51** (1948), 1262–1269.

[5] L. Flaminio and G. Forni, *On effective equidistribution for higher step nilflows*, preprint available as arXiv:1407.3640.

[6] K. Ford and T. D. Wooley, *On Vinogradov's mean value theorem: strongly diagonal behaviour via efficient congruencing*, Acta Math. **213** (2014), no. 2, 199–236.

[7] L. D. Pustyl′nikov, *New estimates of Weyl sums and the remainder term in the law of distribution of the fractional part of a polynomial*, Ergodic Theory Dynam. Systems **11** (1991), no. 3, 515–534.

[8] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. **77** (1916), 313–352.

[9] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing*, Annals of Math. (2) **175** (2012), no. 3, 1575–1627.

[10] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing, II*, Duke Math. J. **162** (2013), no. 4, 673–730.

[11] T. D. Wooley, *Multigrade efficient congruencing and Vinogradov's mean value theorem*, Proc. London Math. Soc. (in press); preprint available as arXiv:1310.8447.

[12] T. D. Wooley, *Approximating the main conjecture in Vinogradov's mean value theorem*, submitted; preprint available as arXiv:1401.2932.

[13] T. D. Wooley, *The cubic case of the main conjecture in Vinogradov's mean value theorem*, submitted; preprint available as arXiv:1401.3150.

School of Mathematics, University of Bristol, University Walk, Clifton, Bristol BS8 1TW, United Kingdom

E-mail address: `matdw@bristol.ac.uk`