Functiones et Approximatio
XXXIX.2 (2008), 237–260

COUNTING DIOPHANTINE APPROXIMATIONS

Jörg Brüdern

Dedicated to Professor Władysław Narkiewicz

Abstract: A recent development of the Davenport-Heilbronn method for diophantine inequalities is reexamined, and then applied to a class of problems in diophantine approximation. Among other things, an asymptotic formula is obtained for the number of solutions of the simultaneous inequalities $|n_j - \lambda_j n_0| < \varepsilon$ with square-free $n_j \in [1, N]$, whenever the positive real numbers $\lambda_1, \dots, \lambda_r$ and 1 are linearly independent over the rationals.
Keywords: Davenport-Heilbronn method, diophantine approximation, square-free numbers.

1. An illustrative example

The most classical problem in diophantine approximation concerns rational numbers near a given set of positive real numbers, say $\lambda_1, \dots, \lambda_r$. One intends to solve the inequalities

$$|n_j - \lambda_j n_0| < \varepsilon \quad (1 \leq j \leq r) \tag{1.1}$$

in natural numbers n_i. A prominent example is Dirichlet's theorem: If $\varepsilon = N^{-1/r}$, then (1.1) has a solution with $1 \leq n_i \leq N$. Many variants of Dirichlet's theorem have been considered, with the variables n_j restricted to special sets, and ε a function of the "denominator" n_0, tending to zero as n_0 increases. The focus in these researches has mostly been on the existence of solutions, with ε as small as possible. We mention a recent result of Dietmann [9]: For $\theta < \frac{2}{3}$ and subject to a mild condition on $\lambda_1, \dots, \lambda_r$, there are infinitely many square-free numbers n_0 such that the inequalities $|n_j - \lambda_j n_0| < n_0^{-\theta/r}$ have a solution in integers $n_1, \dots, n_r$. When $r = 1$, Heath-Brown [14] preceded this with the stronger conclusion that for irrational $\lambda > 0$ and any $\theta < \frac{2}{3}$ there are infinitely many pairs n_0, n_1 of square-free numbers with $|n_1 - \lambda n_0| < n_0^{-\theta}$. Heath-Brown's work is among the few examples where not only the denominator n_0 but also the numerators n_j with $1 \leq j \leq r$ in (1.1) fall into specified sets.

Mathematics Subject Classification: 11J13, 11D75

In the present communication we address the problem of counting solutions of (1.1), with the variables $n_j \in [1, N]$ restricted to a fairly general class of sets. When specialized to the set of square-free numbers, our result reads as follows.

Theorem 1. *Let $\lambda_1, \ldots, \lambda_r$ denote positive real numbers such that $1, \lambda_1, \ldots, \lambda_r$ are linearly independent over $\mathbb{Q}$. Let $Z_\varepsilon(N)$ denote the number of solutions of* (1.1) *in square-free numbers n_j with $1 \le n_j \le N$. Then there exists a positive real number C and a function $T(N)$ that increases to infinity as N tends to infinity, such that*

$$Z_\varepsilon(N) = CN\varepsilon^r + O(NT(N)^{-1})$$

holds uniformly in $0 < \varepsilon \le 1$.

Note that $T(N)$ depends on $\lambda_1, \ldots, \lambda_r$. This is necessarily so, see §8.

Some of the tools in this paper are easier to describe in the simplest case $r = 1$. Even then Theorem 1 seems to be new but may well be part of the folklore. Traditionally one would search for small values of $n_1 - \lambda n_0$, with n_0 in some special set $\mathcal{N}_0$, with the aid of exponential sums. Methods of Weyl and Vinogradov are standard nowadays in this family of problems, but are set up in a way that deny access to the numerators n_1. Thus we are forced to use more subtle methods originally invented by Davenport and Heilbronn [8]. This amounts to a treatment of the inequality $|n_1 - \lambda n_0| < \varepsilon$, with n_j restricted to a special set $\mathcal{N}_j \subset [1, N]$, very much like a binary additive problem. One writes $\Upsilon(\alpha) = \max(0, 1 - |\alpha|/\varepsilon)$ as a Fourier transform of a function $K(\alpha)$ (see (5.2)), and takes the formula

$$\sum_{\substack{n_0 \in \mathcal{N}_0 \\ n_1 \in \mathcal{N}_1}} \Upsilon(n_1 - \lambda n_0) = \int_{-\infty}^{\infty} F_1(\alpha) F_0(-\lambda\alpha) K(\alpha)\, d\alpha \tag{1.2}$$

with

$$F_j(\alpha) = \sum_{n_j \in \mathcal{N}_j} e(\alpha n)$$

as the starting point. A suitable neighbourhood of the origin contributes a term of size $\varepsilon N^{-1} \# \mathcal{N}_0 \# \mathcal{N}_1$ to (1.2) whereas large α make a negligible contribution due to the decay of $K(\alpha)$. Only very recently Bentkus and Götze [2] developed a method that shows

$$F_1(\alpha) F_0(-\lambda\alpha) = o(\# \mathcal{N}_0 \# \mathcal{N}_1) \tag{1.3}$$

uniformly for a rather large range of α, provided λ is irrational and one has some Weyl type bounds for $F_j(\alpha)$ at hand. Freeman [11] reworked their ideas in a context closer to ours, and we refer to the underlying idea as the Bentkus-Götze-Freeman device. No existing version of it seems to suit our needs. Therefore, in §3, we again review the basic principles and demonstrate in Theorem 3 that the essence of the Bentkus-Götze-Freeman device can be freed from any reference to exponential sums. In the new form, the device is ready for immediate adaption to any concrete situation of the shape (1.3), and multi-dimensional variants thereof.

As is familiar to practioneers in the additive theory of numbers, inequalities like (1.3) alone are insufficient to extract an asymptotic formula from the integral (1.2). In fact one needs to break square root cancellation for an individual exponential sum $F_j(\alpha)$, at least in mean square over suitable minor arcs. In [5] we classified the sets $\mathcal{N}_j$ where this is feasible. It turns out that multiplicative sequences like the square-free numbers featured in Theorem 1, form only a subclass of those sequences where mean square estimates of the required strength are indeed available. We summarize the consequences of [5] in §4, and use this and our version of the Bentkus-Götze-Freeman device to derive an asymptotic formula for a multidimensional analogue of (1.2) in §§5 and 6. The main theorem is formulated in §7, and contains Theorem 1 as a special case. In a sense, §§5–7 may be viewed as a variation on the work in [5], for diophantine approximations rather than equations. However, the convolution integral treatment in §6 is a major digression from the iterative approach in [5] that should also prove useful elsewhere.

The notation used in this paper is mostly standard and otherwise introduced when appropriate. Frequently vector notation is used, and printed in bold. The dimension is often, but not always r. For $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r) \in \mathbb{R}^r$ we write $\|\boldsymbol{\alpha}\| = \max |\alpha_j|$, and $\boldsymbol{\alpha\beta} = \alpha_1\beta_1 + \ldots + \alpha_r\beta_r$ is the usual scalar product.

2. Normalized matrices

The Fourier transform method of Davenport and Heilbronn has been used extensively to study linear systems

$$|\lambda_{i1}x_1 + \lambda_{i2}x_2 + \ldots + \lambda_{is}x_s - \mu_i| < \varepsilon \quad (1 \le i \le r), \tag{2.1}$$

where ε is a (small) positive real number, $\Lambda = (\lambda_{ij})_{\substack{1\le i\le r\\ 1\le j\le s}}$ is a real matrix of rank r, the μ_i are real numbers, and the variables x_j vary over some specific set of integers (typically, the k-th powers, the primes, or as in our Theorem 1, the square-free numbers). If for some i all λ_{ij} are integers, then for small ε the corresponding inequality reduces to an equation. It may also be that equations are "hidden" in the system (2.1), in that an inequality has integral coefficients only after suitable linear combinations of the forms in (2.1) have replaced the original ones. Freeman [12], section 2, has discussed this theme in detail, and we briefly describe his ideas here. Let $z(\Lambda)$ denote the number of rows of Λ that have integer entries only, and let

$$R = R(\Lambda) = \max_{\mathrm{A}\in GL_r(\mathbb{R})} z(\mathrm{A}\Lambda) .$$

We call R the *integral rank* of Λ; this is the number of "hidden equations" in (2.1), and we proceed to make them "visible". There is then a non-singular matrix A such that $\mathrm{A}\Lambda$ has the first R rows integral. This partially motivates the following definition.

A real matrix $\Lambda = (\lambda_{ij})_{\substack{1\le i\le r\\ 1\le j\le s}}$ of rank r and integral rank R is called *normalized* if the first R rows are integral, and the submatrices $(\lambda_{ij})_{1\le i,j\le r}$ and $(\lambda_{ij})_{1\le i,j\le R}$ are non-singular.

Note that a matrix of rank r has an $r\times r$-submatrix that is non-singular. Hence, a renumbering of the columns of Λ suffices to make $(\lambda_{ij})_{1\le i,j\le r}$ non-singular. This corresponds to a renumbering of variables in (2.1), and does not significantly alter the original problem. Now consider the first R rows of the first columns, $(\lambda_{ij})_{\substack{1\le i\le R\\ 1\le j\le r}}$. This matrix must contain an $R\times R$-submatrix that is non-singular, so another renumbering of the first r columns guarantees that $(\lambda_{ij})_{1\le i,j\le r}$ is non-singular. Thus, we have proved the following lemma.

Lemma 2. *Let* $\Lambda = (\lambda_{ij})_{\substack{1\le i\le r\\ 1\le j\le s}}$ *be a real matrix of rank* r *and integral rank* R*. Then, after a suitable reordering of the columns of* Λ*, there is a non-singular* $r\times r$*-matrix* A *such that* $\mathrm{A}\Lambda$ *is normalized.*

Let Λ be a normalized matrix with rows $\mathbf{c}_1,\ldots,\mathbf{c}_r$, and suppose that α_1, …, $\alpha_r\in\mathbb{R}$ are chosen such that

$$\alpha_1\mathbf{c}_1+\ldots+\alpha_r\mathbf{c}_r\in\mathbb{Q}^s\ .$$

Then

$$\alpha_{R+1}=\ldots=\alpha_r=0\ ,$$

for otherwise the integral rank would be at least $R+1$. We shall use this fact below, in the proof of Theorem 3.

We end this section by considering briefly the special case where $s=r+1$. It is that case that is relevant later. Following [1], a real $r\times(r+1)$-matrix Λ is called *highly non-singular* if any $r\times r$-submatrix is non-singular. For such a matrix there exists $\mathrm{A}\in GL_r(\mathbb{R})$ such that

$$\mathrm{A}\Lambda=\begin{pmatrix} -\lambda_1 & 1 & & 0\\ \vdots & & \ddots & \\ -\lambda_r & 0 & & 1\end{pmatrix} \tag{2.2}$$

with non-zero real numbers $\lambda_1,\ldots,\lambda_r$. Note $R(\Lambda)=R(\mathrm{A}\Lambda)$, and that we may still apply row operations to the right hand side of (2.2) to determine the integral rank of Λ. In particular, it is immediate that $R=0$ *if and only if the numbers* $1,\lambda_1,\lambda_2,\ldots,\lambda_r$ *are linearly independent over* $\mathbb{Q}$. This will be needed in §6.

The matrix Λ is *positive* if the system of equations $\Lambda\mathbf{x}=0$ has a solution with all $x_i>0$. A matrix is positive if and only if all λ_i in (2.2) are positive.

3. A technical estimate

In this section we set the scene for our version of the Bentkus–Götze–Freeman device. Some notation is required before we can formulate the result. Let N denote the main parameter, and suppose that

$$1 \le Q \le \tfrac{1}{2}\sqrt{N}. \tag{3.1}$$

Then, the intervals

$$\mathfrak{M}_{q,a} = \{\alpha \in \mathbb{R} : \ |q\alpha - a| \le Q/N\}$$

with $1 \le q \le Q$, $a \in \mathbb{Z}$ and $(a;q) = 1$ are pairwise disjoint, and we denote their union by $\mathfrak{M}(Q)$. Choose a function $H : [1,\infty) \to (0,\infty)$ that decreases monotonically to zero as its argument tends to ∞. Define a function $\Xi : \mathbb{R} \to [0,\infty)$ by

$$\Xi(\alpha) = \begin{cases} 0 & \text{for } \alpha \notin \mathfrak{M}(Q), \\ H(q + N|q\alpha - a|) & \text{for } \alpha \in \mathfrak{M}_{q,a} \cap \mathfrak{M}(Q). \end{cases}$$

We call Ξ the weight on $\mathfrak{M}(Q)$ associated with H.

Now let $\Lambda = (\lambda_{ij})_{\substack{1\le i\le r\\ 1\le j\le s}}$ be a matrix with real entries, of rank r and of integral rank $R \le r$. Suppose that Λ is normalized in the sense of section 2. Fix a function $U(N)$ with $U(N) \to \infty$ as $N \to \inf$, and define

$$\mathcal{M} = \{\boldsymbol{\alpha} \in \mathbb{R}^r : \ \alpha_i \in \mathfrak{M}(U) \ (1 \le i \le R), \ |\alpha_i| \le U/N \ (R < i \le r)\} .$$

The *minor arcs* in our set-up are then defined by

$$\mathfrak{m}(T) = \{\boldsymbol{\alpha} \in \mathbb{R}^r : U(N)/N \le \|\boldsymbol{\alpha}\| \le T\} \backslash \mathcal{M} .$$

We are now in a position to announce the main result of this section.

Theorem 3. *Let Q be a function of N satisfying* (3.1). *Let Ξ be a weight on $\mathfrak{M}(Q)$ associated with a decreasing function H, as above. Let $U(N)$ be a function with $U(N) \to \infty$ as $N \to \infty$. Let Λ be a normalized $r \times s$-matrix with columns $\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_s$. Then, there exists a function $T(N) \to \infty$ with $N \to \infty$ and $T(N) \le U(N)$, and such that*

$$\sup_{\boldsymbol{\alpha} \in \mathfrak{m}(T(N))} \Xi(\boldsymbol{\alpha}\boldsymbol{\lambda}_1)\Xi(\boldsymbol{\alpha}\boldsymbol{\lambda}_2)\dots\Xi(\boldsymbol{\alpha}\boldsymbol{\lambda}_s) \ll T(N)^{-1}. \tag{3.2}$$

This is the promised version of the Bentkus–Götze–Freeman device, freed from explicit references to exponential sums. It may be useful to compare it with the earlier versions by Freeman [10, 11] and Wooley [18] or Parsell [15]. These authors study Weyl sums for k-th powers, or for primes, defined by

$$f(\alpha) = \sum_{x \le N^{1/k}} e(\alpha x^k), \quad g(\alpha) = \sum_{p \le N} (\log p) e(\alpha p) .$$

It is then shown that for irrational $\lambda \in \mathbb{R}$ and some suitable functions $U(N)$ and $T(N)$ as above, one has

$$\sup_{U(N)/N<|\alpha|\le T(N)} |f(\alpha)f(\lambda\alpha)| = o(N^{2/k}), \tag{3.3}$$

$$\sup_{U(N)/N<|\alpha|\le T(N)} |g(\alpha)g(\lambda\alpha)| = o(N^2). \tag{3.4}$$

These bounds are also available from Theorem 3, combined with some version of Weyl's inequality. In fact, one readily finds that for $\alpha \in \mathfrak{M}(N^{1/k})$ one has

$$f(\alpha) \ll N^{1/k}(q + N|q\alpha - a|)^{-1/k}$$

(see Vaughan [17], Theorem 4.1 and Lemma 4.6, for example), and when $\alpha \notin \mathfrak{M}(N^{1/k})$ Weyl's inequality ([17], Lemma 2.4) gives $f(\alpha) \ll N^{\frac{1}{k}-\delta}$ for some $\delta > 0$; the actual value of δ is of no relevance here. We take $H(t) = t^{-1/k}$ to define the weight Ξ on $\mathfrak{M}(N^{1/k})$. The two bounds for $f(\alpha)$ now combine to

$$N^{-1/k} f(\alpha) \ll \Xi(\alpha) + N^{-\delta} ,$$

and (3.3) follows from Theorem 3 with $r = 1$, $s = 2$ and $\Lambda = (1, \lambda)$. A very similar argument can be used to establish (3.4). Multi-dimensional versions of (3.3) for smooth Weyl sums occur in work of Freeman. His important Lemma 5.5 of [12] can also be deduced from our Theorem 3 in the way indicated above. It is hoped that Theorem 3 serves as an easy-to-use reference in any future application of the Bentkus–Götze–Freeman device. In delicate uses of the circle method, ordinary and smooth Weyl sums often occur simultaneously, and in such situations, Theorem 3 is still easily applicable.

One might ask for a further generalization of Theorem 3, with different weights $\Xi_1, \dots, \Xi_s$ on the linear forms $\boldsymbol{\alpha\lambda}_1, \dots, \boldsymbol{\alpha\lambda}_s$. It will be clear from the proof below that such a generalization is indeed possible. However, as we shall see later, Theorem 3 is sufficiently flexible as it stands.

In this paper Theorem 3 is used only with $R = 0$. In this case the sets $\mathcal{M}$ and $\mathfrak{m}(T)$ take the simple shape $\|\boldsymbol{\alpha}\| \le U$ and $U < \|\boldsymbol{\alpha}\| \le T$, respectively, and the proof considerably simplifies. However, in most applications the full strength of Theorem 3 is required. For an example and further comments on the matter we refer to our recent work with Wooley [7].

Proof of Theorem 3. As a first step, we show that for any fixed $T \ge 1$, one has

$$\lim_{N\to\infty} \sup_{\boldsymbol{\alpha}\in\mathfrak{m}(T)} \Xi(\boldsymbol{\alpha\lambda}_1) \dots \Xi(\boldsymbol{\alpha\lambda}_s) = 0. \tag{3.5}$$

We prove (3.5) by contradiction. If (3.5) is false, then there exist a real number ε with $0 < \varepsilon < 1$, an increasing sequence of real numbers N_ν with $N_\nu \to \infty$ as $\nu \to \infty$, and $\boldsymbol{\alpha}_\nu \in \mathfrak{m}(T)$ with

$$\Xi(\boldsymbol{\alpha}_\nu\boldsymbol{\lambda}_1)\Xi(\boldsymbol{\alpha}_\nu\boldsymbol{\lambda}_2) \dots \Xi(\boldsymbol{\alpha}_\nu\boldsymbol{\lambda}_s) > \varepsilon .$$

But $\Xi(\alpha) \le 1$ whence $\Xi(\boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j) > \varepsilon$ for $1 \le j \le s$. In particular, $\boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j \in \mathfrak{M}(Q)$, so that there are unique coprime integers $q_{\nu j}, a_{\nu j}$ with $\boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j \in \mathfrak{M}_{q_{\nu j}, a_{\nu j}}$. Then

$$H(q_{\nu j} + N_\nu |q_{\nu j} \boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j - a_{\nu j}|) > \varepsilon \, .$$

Since H is decreasing, we write $E = H^{-1}(\varepsilon)$ and conclude that

$$1 \le q_{\nu j} \le E, \quad |q_{\nu j} \boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j - a_{\nu j}| \le E N_\nu^{-1}$$

hold for all $\nu \in \mathbb{N}$, all $1 \le j \le s$. From the second inequality, one has

$$|a_{\nu j}| \le E N_\nu^{-1} + rET \max |\lambda_{ij}| \, .$$

In particular, the $2s$-tuples $(q_{\nu 1}, \dots, q_{\nu s}, a_{\nu 1}, \dots, a_{\nu s})$ can take only finitely many values, so at least one of these, say $(q_1, \dots, q_s, a_1, \dots, a_s)$, occurs infinitely often as ν varies. Choosing a subsequence of the N_ν if necessary, we can suppose (after renumbering if necessary) that $q_{\nu j} = q_j$, $a_{\nu j} = a_j$ hold for *all* ν, and we then have

$$1 \le q_j \le E, \quad |q_j \boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j - a_j| \le E N_\nu^{-1}. \tag{3.6}$$

Since $\boldsymbol{\alpha}_\nu \in [-T, T]^r$, and $[-T, T]^r$ is compact, there is a subsequence that converges to a limit $\boldsymbol{\alpha} \in [-T, T]^r$. As before, we may then suppose that the N_ν have been chosen such that $\boldsymbol{\alpha}_\nu$ itself already converges to $\boldsymbol{\alpha}$. With $\nu \to \infty$ in (3.6) it follows that

$$\boldsymbol{\alpha} \boldsymbol{\lambda}_j = \frac{a_j}{q_j} \quad (1 \le j \le s), \tag{3.7}$$

and since Λ is normalized, this is only possible if $\alpha_{R+1} = \ldots = \alpha_r = 0$. By (3.6) and (3.7),

$$|(\boldsymbol{\alpha}_\nu - \boldsymbol{\alpha}) \boldsymbol{\lambda}_j| = \left| \Big(\boldsymbol{\alpha}_\nu \boldsymbol{\lambda}_j - \frac{a_j}{q_j} \Big) + \Big(\frac{a_j}{q_j} - \boldsymbol{\alpha} \boldsymbol{\lambda}_j \Big) \right| \le E N_\nu^{-1} \, .$$

We use this with $1 \le j \le r$. Since Λ is normalized, $(\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_r)$ is non-singular. Hence, there is a constant $C > 0$, depending only on $\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_r$, such that

$$\|\boldsymbol{\alpha}_\nu - \boldsymbol{\alpha}\| \le C E N_\nu^{-1} \, .$$

For $R < j \le r$, this yields

$$|\alpha_{\nu j}| \le C E N_\nu^{-1}. \tag{3.8}$$

When $R = 0$, this contradicts $\boldsymbol{\alpha}_\nu \in \mathfrak{m}(T)$ when ν is large. This confirms (3.5) when $R = 0$.

Now suppose that $0 < R \le r$. For any $\mathbf{x} \in \mathbb{R}^r$ write $\tilde{\mathbf{x}} = (x_1, \dots, x_R)$. Then, by (3.6) and (3.8),

$$\left| \tilde{\boldsymbol{\alpha}}_\nu \tilde{\boldsymbol{\lambda}}_j - \frac{a_j}{q_j} \right| \le E N_\nu^{-1} + \Big| \sum_{i=R+1}^{r} \lambda_{ij} \alpha_{\nu i} \Big| \le r C E N_\nu^{-1} \max |\lambda_{ij}| \, .$$

Since Λ is normalized, the $R\times R$-matrix $\mathrm{A} = (\tilde{\boldsymbol{\lambda}}_1,\dots,\tilde{\boldsymbol{\lambda}}_R)$ is non-singular. The previous display with $1\le j\le R$ then yields

$$\tilde{\boldsymbol{\alpha}}_\nu \mathrm{A} = \Big(\frac{a_1}{q_1},\dots,\frac{a_R}{q_R}\Big) + \tilde{\boldsymbol{\beta}}_\nu \tag{3.9}$$

with

$$\|\tilde{\boldsymbol{\beta}}_\nu\| \le C_1 E N_\nu^{-1}\ ;$$

here $C_1 = rC\max|\lambda_{ij}|$ depends only on Λ. There is exactly one $\tilde{\mathbf{b}}\in\mathbb{R}^R$, one $\tilde{\boldsymbol{\gamma}}_\nu\in\mathbb{R}^R$ with

$$\tilde{\mathbf{b}}\mathrm{A} = \Big(\frac{a_1}{q_1},\dots,\frac{a_R}{q_R}\Big),\quad \tilde{\boldsymbol{\gamma}}_\nu\mathrm{A} = \tilde{\boldsymbol{\beta}}_\nu. \tag{3.10}$$

By (3.9), we then get

$$\tilde{\boldsymbol{\alpha}}_\nu = \tilde{\mathbf{b}} + \tilde{\boldsymbol{\gamma}}. \tag{3.11}$$

By Cramer's rule, applied to the second equation in (3.10), the upper bounds on $\tilde{\boldsymbol{\beta}}_\nu$ imply that $\|\tilde{\boldsymbol{\gamma}}_\nu\|\le C_2EN_\nu^{-1}$ where C_2 again only depends on Λ. Since A is an integral matrix (because Λ is normalized), Cramer's rule also shows that $\tilde{\mathbf{b}} = (d_1/q,\dots,d_r/q)$ with $d_i\in\mathbb{Z}, q\in\mathbb{N}, (d_1;\dots,d_r;q)=1$ and $q|q_1q_2\dots q_r\det A$. In particular, $q\le|\det \mathrm{A}|E^R$. By (3.11), it follows that $\alpha_{\nu i}\in\mathfrak{M}(C_3E+|\det\mathrm{A}|E^R)$ for $1\le i\le R$, and in particular that $\alpha_{\nu i}\in\mathfrak{M}(U(N_\nu))$ when ν is large. But then $\boldsymbol{\alpha}_\nu\in\mathcal{M}$, a contradiction. This establishes (3.5).

This far, we have followed Freeman's argument [12], pp. 2688–2691, with very few details that had to be adjusted to the slightly different context. We now proceed to derive Theorem 3 from (3.5) by an argument that closely follows Wooley [18]. In fact, by (3.5), for any $t\in\mathbb{N}$ there exists a number $N_t\in\mathbb{N}$ such that for all $N\ge N_t$ one has

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}(t)} \Xi(\boldsymbol{\alpha}\boldsymbol{\lambda}_1)\Xi(\boldsymbol{\alpha}\boldsymbol{\lambda}_2)\dots\Xi(\boldsymbol{\alpha}\boldsymbol{\lambda}_s)\le\frac{1}{t}. \tag{3.12}$$

Once N_t is determined, it may be replaced by any larger integer. Hence, we may choose N_t such that $N_t<N_{t+1}$ for all t, and that $U(N_t)>t$. Then $N_t\to\infty$ as $t\to\infty$. We define T by $T(N)=t$ for $N_t\le N<N_{t+1}$. Then T satisfies $T(N)\le U(N)$, is increasing, unbounded, and (3.2) follows from (3.12).

4. Extremal sequences and their exponential sums

This section is devoted to a class of sequences that have been investigated extensively in our recent memoir [5]. We begin by collecting the principal concepts and conclusions of that work, as this is fundamental for our later analysis.

Let $\mathcal{S}\subset\mathbb{N}$ be any subset of the natural numbers, and write

$$\mathcal{S}(N,q,a) = \{s\in\mathcal{S} : s\le N, s\equiv a \bmod q\}\ .$$

Suppose that for any $a \in \mathbb{Z}$, $q \in \mathbb{N}$ there is a real number $s(q,a)$ such that the asymptotic formula

$$\#\mathcal{S}(N,q,a) = s(q,a)N + o(N) \quad (N \to \infty) \tag{4.1}$$

holds; sequences $\mathcal{S}$ with this property are called *distributed*, and $\varrho(\mathcal{S}) = s(1,0)$ is the *density* of $\mathcal{S}$. If a distributed sequence satisfies $\varrho(\mathcal{S}) > 0$ and the relation

$$\sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \Big| \sum_{b=1}^{q} s(q,b) e\Big(\frac{ab}{q}\Big)\Big|^2 = \varrho(\mathcal{S}) \tag{4.2}$$

holds, then $\mathcal{S}$ is an *extremal sequence* in the language of [5]. The series in (4.2) converges for any distributed sequence, and is bounded above by $\varrho(\mathcal{S})$ (this is [5], Lemma 1.1). Moreover, $\mathcal{S}$ is extremal if and only if the function

$$\mathbf{s}(n) = \begin{cases} 1 & s \in \mathcal{S} \\ 0 & s \notin \mathcal{S} \end{cases} \tag{4.3}$$

is the limit of periodic functions in the space of functions $f : \mathbb{N} \to \mathbb{R}$ equipped with the semi-norm

$$\|f\|^2 = \limsup N^{-1} \sum_{n \le N} |f(n)|^2 \ ;$$

for this compare Puchta [16] or this author [4]. Many examples of extremal sequences may be found in [5]: any sequence that has positive density and a multiplicative indicator (4.3) is extremal ([5], Theorem 1.4). In particular, the square-free numbers and more generally the k-free numbers form extremal sequences. Also, sequences that result from a convergent sieve process are extremal. More precisely, for any prime p let $k(p) \in \mathbb{N}_0$, let $\Omega_p \subset (-\frac{1}{2}p^k, \frac{1}{2}p^k]$ with $\#\Omega_p < p^k$, and suppose that $\sum_p p^{-k(p)} \#\Omega_p$ converges. If there exists a constant $C > 0$ such that $\Omega_p \subset [-C, C]$ for all primes p, then the sifted set

$$\{s \in \mathbb{N} : \quad s \notin \Omega_p \bmod p^k\}$$

has density $\prod_p (1 - \frac{\#\Omega_p}{p^{k(p)}})$, and is extremal ([5], Theorem 1.10). In particular, if $h_1 < h_2 < \ldots < h_r$ are natural numbers with $\#\{h_j \bmod p^k\} < p^k$ for all primes p, then the *r-tuples of k-free numbers*

$$\{s \in \mathbb{N} : s + h_j \quad k\text{-free for } 1 \le j \le r\}$$

are extremal. For further examples and other properties of extremal sequences, see Brüdern [4], [5] and Puchta [16].

Let $\mathcal{S}$ denote a distributed sequence, and put

$$S(\alpha) = \sum_{\substack{s \in \mathcal{S} \\ s \le N}} e(\alpha s). \tag{4.4}$$

By Parseval's identity and (4.1), one has

$$\int_0^1 |S(\alpha)|^2\,d\alpha = \varrho(\mathcal{S})N + o(N). \tag{4.5}$$

Recall the definition of $\mathfrak{M}(Q)$ (below (3.1)), let $\mathfrak{N}(Q) = \mathfrak{M}(Q) \cap [0,1]$, and put $\mathfrak{n}(Q) = [0,1]\backslash\mathfrak{N}(Q)$. For *extremal* sequences $\mathcal{S}$, Theorem 1.6 of Brüdern [5] asserts that for any function $Q = Q(N)$ with $Q(N) \to \infty$ as $N \to \infty$ one has

$$\int_{\mathfrak{N}(Q)} |S(\alpha)|^2\,d\alpha = \varrho(\mathcal{S})N + o(N) \tag{4.6}$$

and

$$\int_{\mathfrak{n}(Q)} |S(\alpha)|^2\,d\alpha = o(N). \tag{4.7}$$

The properties (4.6) and (4.7) are equivalent, in view of (4.5), and actually characterize extremal sequences among the distributed ones, but this is not needed here. In the later sections of this paper, we shall make crucial use of (4.7) along with the following upper bound that in contrast to (4.7), is valid for distributed sequences.

Lemma 4. *Let $\mathcal{S}$ denote a distributed sequence. Then there exists an increasing function $Q = Q(N)$ with $Q(N) \to \infty$ as $N \to \infty$, and a decreasing function H with $H(t) \to 0$ as $t \to \infty$ such that the inequality $|S(\alpha)| \ll N\Xi(\alpha)$ holds for all $\alpha \in \mathfrak{M}(Q)$; here S is the exponential sum* (4.4)*, and Ξ is the weight on $\mathfrak{M}(Q)$ associated with H.*

Proof. For any distributed sequence $\mathcal{S}$, there exists a function $T(N)$ such that $T(N)$ and $N/T(N)$ are increasing, and $T(N) \to \infty$ as $N \to \infty$, with the property that

$$|\#\mathcal{S}(N,q,b) - s(q,b)N| \le N/T(N)$$

holds uniformly in $N \ge 1, q \in \mathbb{N}, b \in \mathbb{Z}$ (this is [5], Lemma 2.2). Hence, by (4.4), whenever $(a;q) = 1$,

$$S\Big(\frac{a}{q}\Big) = \sum_{b=1}^{q} \sum_{n\in\mathcal{S}(N,q,b)} e\Big(\frac{ab}{q}\Big) = NG(q,a) + E(N,q) \tag{4.8}$$

where

$$G(q,a) = \sum_{b=1}^{q} s(q,b)e\Big(\frac{ab}{q}\Big)\,,$$

and where $|E(N,q)| \le qN/T(N)$. By partial summation and (4.8),

$$\begin{aligned} S\Big(\frac{a}{q} + \beta\Big) &= e(N\beta)S\Big(\frac{a}{q}\Big) - 2\pi i\beta \int_1^N e(\beta\gamma) \sum_{\substack{n\le\gamma\\ n\in\mathcal{S}}} e\Big(\frac{an}{q}\Big)\,d\gamma \\ &= G(q,a)I(\beta) + e(N\beta)E(N,q) - 2\pi i\beta \int_1^N e(\beta\gamma)E(\gamma,q)\,d\gamma \end{aligned}$$

where

$$I(\beta) = \int_0^N e(\beta\gamma)\, d\gamma. \tag{4.9}$$

Since $N/T(N)$ is increasing, we have $|E(\gamma, q)| \le qN/T(N)$ for $1 \le \gamma \le N$. It follows that

$$\Big|S\Big(\frac{a}{q} + \beta\Big) - G(q,a)I(\beta)\Big| \le \frac{qN}{T(N)}(1 + 2\pi N|\beta|). \tag{4.10}$$

As we have remarked earlier, the sum in (4.2) converges. Hence, there is a function $h(q)$ that decreases to 0 as $q \to \infty$, and such that

$$\sum_{\substack{a=1\\(a,q)=1}}^{q} |G(q,a)|^2 \le h(q)^2$$

holds for all q. In particular, this yields $|G(q,a)| \le h(q)$ for all $(a,q) = 1$. By partial integration, applied to (4.9), we infer from (4.10) the inequality

$$\Big|S\Big(\frac{a}{q} + \beta\Big)\Big| \le \frac{2Nh(q)}{1 + N|\beta|} + \frac{qN}{T(N)}(1 + 2\pi N|\beta|)\ .$$

We may suppose that $h(q) \ge 1/q$ for all $q \in \mathbb{N}$, for otherwise we replace $h(q)$ by $\max(h(q), 1/q)$ without affecting the previous argument. We now extend the definition of h to $[1, \infty)$ by $h(\gamma) = h([\gamma])$. Then h is still a non-increasing function, and

$$\Big|S\Big(\frac{a}{q} + \beta\Big)\Big| \le 2Nh(q)h(1 + N|\beta|) + \frac{qN}{T(N)}(1 + 2\pi N|\beta|)\ .$$

Since h is non-increasing, the inequality $h(u)h(v) \le h(1)h(\sqrt{uv})$ holds for any $u \ge 1$, $v \ge 1$. Hence, $H(t) = h(\sqrt{t})$ is non-increasing, satisfies $H(t) \to 0$ as $t \to \infty$, and one has

$$\Big|S\Big(\frac{a}{q} + \beta\Big)\Big| \le 2h(1)NH(q(1 + N|\beta|)) + \frac{N}{T(N)}q(1 + 2\pi N|\beta|). \tag{4.11}$$

Now choose a function $Q(N)$ with $Q(N) \to \infty$ as $N \to \infty$, and such that $H(2Q(N)) \ge 1/\sqrt{T(N)}$; this is always possible since H and $1/T$ are non-increasing and tend to zero as $N \to \infty$. We may now replace $Q(N)$ with $\min(Q(N), \sqrt{T(N)})$ if necessary to ensure that $Q(N) \le \sqrt{T(N)}$; the inequality $H(2Q(N)) \ge 1/\sqrt{T(N)}$ then still holds. By (4.11), it follows that for $\frac{a}{q} + \beta \in \mathfrak{M}(Q)$, one has

$$\begin{aligned}\Big|S\Big(\frac{a}{q} + \beta\Big)\Big| &\le 2h(1)NH(q(1 + N|\beta)) + 8NQ(N)T(N)^{-1}\\ &\le 2h(1)NH(q(1 + N|\beta)) + 8NT(N)^{-1/2}\\ &\le (2h(1) + 8)H(q(1 + N|\beta|)).\end{aligned}$$

This establishes Lemma 4. ∎

5. An auxiliary asymptotic formula

In this section we launch our attack on simultaneous diophantine approximations. The approach is based on the Fourier transform method of Davenport and Heilbronn. We work with the kernel functions

$$K_\eta(\alpha) = \eta\Big(\frac{\sin \pi\eta\alpha}{\pi\eta\alpha}\Big)^2, \quad \Upsilon_\eta(\alpha) = \max\Big(0, 1 - \frac{|\alpha|}{\eta}\Big) \tag{5.1}$$

where $\eta > 0$ is a real parameter. Note that K_η and Υ_η are Fourier transforms of each other; we only require the classical formula

$$\Upsilon_\eta(\alpha) = \int_{-\infty}^{\infty} K_\eta(\beta) e(-\alpha\beta)\, d\beta. \tag{5.2}$$

Higher-dimensional analogues arise from Fubini's theorem. For any function $f : \mathbb{R} \to \mathbb{C}$ we define $f : \mathbb{R}^r \to \mathbb{C}$ by

$$f(\boldsymbol{\alpha}) = f(\alpha_1)f(\alpha_2)\dots f(\alpha_r); \tag{5.3}$$

confusion should not arise from the ambiguous use of f as the number of variables is usually clear from the context.

Fix the following notation for the rest of this section and the next. Let $\mathcal{S}_0, \mathcal{S}_1, \dots, \mathcal{S}_r$ denote extremal sequences, of densities $\varrho_j = \varrho(\mathcal{S}_j)$. Let $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r) \in \mathbb{R}^r$ and suppose that $1, \lambda_1, \dots, \lambda_r$ are linearly independent over $\mathbb{Q}$. Put

$$\varrho = \varrho_0\varrho_1 \dots \varrho_r, \quad \kappa = \max(1, \lambda_1, \dots, \lambda_r). \tag{5.4}$$

Implicit constants, as well as several constructions below, depend on $\boldsymbol{\lambda}$ as well as the $\mathcal{S}_j$, but not on N or η.

For any $\boldsymbol{\beta} \in \mathbb{R}^r$, we count solutions of the inhomogenous system

$$|x_i - \lambda_i x_0 - \beta_i| < \eta \quad (1 \le i \le r)$$

with a certain weight attached to the variables $x_j \in \mathcal{S}_j$. Let

$$P_\eta(N, \boldsymbol{\beta}) = \sum_{\substack{1 \le x_j \le N \\ x_j \in \mathcal{S}_j \\ j=0,1,\dots,r}} \prod_{i=1}^{r} \Upsilon_\eta(x_i - \lambda_i x_0 - \beta_i). \tag{5.5}$$

Our goal is to establish an asymptotic formula for $P_\eta(N, \boldsymbol{\beta})$. By (5.2) and (5.3), if $S_j(\alpha)$ is the sum (4.4) for $\mathcal{S}_j$,

$$P_\eta(N, \boldsymbol{\beta}) = \int_{\mathbb{R}^r} S_1(\alpha_1) \dots S_r(\alpha_r) S_0(-\boldsymbol{\lambda\alpha}) e(-\boldsymbol{\alpha\beta}) K_\eta(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}. \tag{5.6}$$

This Fourier integral is the traditional point of departure for the Davenport–Heilbronn method. One splits the integral into three parts: a small neighbourhood of the origin (the *major arc*) contributes a prospective main term, and the part where $\|\boldsymbol{\alpha}\|$ is large (the *trivial region*) contributes very little, due to the decay of $K_\eta(\boldsymbol{\alpha})$. In the intermediate region (the *minor arc*), one uses inequalities of Weyl's type and some argument of the type discussed in section 3 for an upper bound. There are two points where we have to digress from this well trodden path. Unlike earlier versions of the method, in our case the dissection will depend not only on the sets $\mathcal{S}_0, \ldots, \mathcal{S}_r$, but also on $\boldsymbol{\lambda}$. This seems to be an intrinsic feature of the Bentkus–Götze–Freeman device, and apparently is unavoidable (compare the discussion in §8). Moreover, in our description of the method, there is no trivial region. Instead, we use ideas from our paper [3] that are based on the Poisson summation formula. It seems that some sort of refinement on the trivial region is actually necessary when the number of variables is too low for a direct use of mean value estimates of Hua's type; this is the case here.

We begin by constructing the major arc. By (4.10), applied to S_j, with $q = 1$ and $a = 0$, there exists a function $V_j(N)$ with $V_j(N) \to \infty$ as $N \to \infty$, and such that whenever $|\alpha| \le V_j(N)/N$ one has

$$|S_j(\alpha) - \varrho_j I(\alpha)| \le N/V_j(N). \tag{5.7}$$

Let $V = \min(V_0, \ldots, V_r)$ and $U^* = V^{1/(r+1)}$. Then $U^*(N) \to \infty$ as $N \to \infty$. For $1 \le U \le U^*(N)$ we define

$$\mathfrak{K} = \{\boldsymbol{\alpha} : \|\boldsymbol{\alpha}\| \le U/N\}, \quad \mathfrak{k} = \mathbb{R}^r/\mathfrak{K} = \{\boldsymbol{\alpha} \in \mathbb{R}^r : \|\boldsymbol{\alpha}\| > U/N\}. \tag{5.8}$$

For $\|\boldsymbol{\alpha}\| \le U^*/N$, one has $|\boldsymbol{\lambda\alpha}| \le r\|\boldsymbol{\lambda}\| U^*(N)/N \le V(N)/N$ for large N, whence by (5.7), for these $\boldsymbol{\alpha}$,

$$S_1(\alpha_1)\ldots S_r(\alpha_r)S_0(-\boldsymbol{\lambda\alpha}) = \varrho I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha}) + O(N^{r+1}/V(N)) \tag{5.9}$$

where ϱ is given by (5.4) and $I(\boldsymbol{\alpha})$ is defined via (4.9) and (5.3). We multiply with the bounded function $K(\boldsymbol{\alpha})e(-\boldsymbol{\alpha\beta})$ and integrate over $\mathfrak{K}$ to infer that

$$\int_{\mathfrak{K}} S_1(\alpha_1)\ldots S_r(\alpha_r)S_0(-\boldsymbol{\lambda\alpha})e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}$$
$$= \varrho \int_{\mathfrak{K}} I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha})e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} + O(NU^rV^{-1}).$$

We now complete the integral on the right hand side to the singular integral

$$P^*_\eta(N,\boldsymbol{\beta}) = \int_{\mathbb{R}^r} I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha})e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}, \tag{5.10}$$

and then proceed to evaluate (5.10) asymptotically. The arguments are largely standard, but some care is required because only very few variables are present. By (5.8) we have

$$\int_{\mathfrak{K}} S_1(\alpha_1)\ldots S_r(\alpha_r)S_0(-\boldsymbol{\lambda\alpha})e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}$$
$$= \varrho P^*_\eta(N,\boldsymbol{\beta}) - \int_{\mathfrak{k}} I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha})e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} + O(NU^{-1}). \tag{5.11}$$

For simplicity, write $\alpha_0 = -\boldsymbol{\lambda\alpha}$ (as a function of $\alpha_1, \dots, \alpha_r$) and put

$$J_l(\boldsymbol{\alpha}) = \prod_{\substack{0 \le j \le r \\ j \ne l}} I(\alpha_j) \,.$$

Then, by Hölder's inequality,

$$\int_{\mathfrak{k}} |I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha})| \, d\boldsymbol{\alpha} \le \prod_{l=0}^{r} \Big(\int_{\mathfrak{k}} |J_l(\boldsymbol{\alpha})|^{\frac{r+1}{r}} \, d\boldsymbol{\alpha} \Big)^{\frac{1}{r+1}}. \tag{5.12}$$

But $J_0(\boldsymbol{\alpha}) = I(\boldsymbol{\alpha})$, and $I(\alpha) \ll N(1+N|\alpha|)^{-1}$, as one finds by partial integration. Consequently,

$$\int_{\mathfrak{k}} |J_0(\boldsymbol{\alpha})|^{\frac{r+1}{r}} \, d\boldsymbol{\alpha} \ll N^{r+1} \int_{\mathfrak{k}} \prod_{j=0}^{r} (1 + N|\alpha_j|)^{-\frac{r+1}{r}} \, d\boldsymbol{\alpha} \ll NU^{-1/r} \,.$$

A cruder argument suffices for $1 \le l \le r$. When $l = r$, substitute α_0 for α_r to infer that

$$\begin{aligned} \int_{\mathfrak{k}} |J_r(\boldsymbol{\alpha})|^{\frac{r+1}{r}} d\boldsymbol{\alpha} &\le \int_{\mathbb{R}^r} |J_r(\boldsymbol{\alpha})|^{\frac{r+1}{r}} d\boldsymbol{\alpha} \\ &= \lambda_r^{-1} \int_{\mathbb{R}^r} |I(\alpha_0)I(\alpha_1) \dots I(\alpha_{r-1})|^{\frac{r+1}{r}} \, d\alpha_0 \dots d\alpha_{r-1} \\ &= \lambda_r^{-1} \Big(\int_{-\infty}^{\infty} |I(\alpha)|^{\frac{r+1}{r}} \, d\alpha \Big)^r \ll N; \end{aligned}$$

here the implicit constant depends on $\boldsymbol{\lambda}$. By symmetry, the same bound holds with J_l in place of J_r, and we infer from (5.12) that

$$\int_{\mathfrak{k}} |I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha})| \, d\boldsymbol{\alpha} \ll NU^{-\frac{1}{r(r+1)}}. \tag{5.13}$$

Since $e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha}) \ll 1$ for $0 < \eta \le 1$, we can use (5.13) to simplify (5.11) to

$$\int_{\mathfrak{K}} S_1(\alpha_1) \dots S_r(\alpha_r) S_0(-\boldsymbol{\lambda\alpha}) e(-\boldsymbol{\alpha\beta}) K_\eta(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} = \varrho P_\eta^*(N, \boldsymbol{\beta}) + O(NU^{-\frac{1}{r(r+1)}}) \tag{5.14}$$

It remains to evaluate $P_\eta^*(N, \boldsymbol{\beta})$, as defined in (5.10). Use the definition of $I(\alpha)$ in (4.9) to write $I(\boldsymbol{\alpha})I(-\boldsymbol{\lambda\alpha})$ as an $r+1$-fold integral. Then, by (5.10) and 5.2,

$$P_\eta^*(N, \boldsymbol{\beta}) = \int_0^N \int_{[0,N]^r} \Upsilon_\eta(\beta_1 - t_1 + \lambda_1 t_0, \dots, \beta_r - t_r + \lambda_r t_0) \, d\mathbf{t} \, dt_0 \,.$$

Note the formal similarity with the definition of $P_\eta(N, \boldsymbol{\beta})$ in (5.5), $P^*(N, \boldsymbol{\beta})$ is a continuous analogue that is easy to compute. We substitute $u_j = t_j - \beta_j - \lambda_j t_0$ in the inner integral and write t for t_0. Then, since $\Upsilon_\eta(-u) = \Upsilon_\eta(u)$,

$$P_\eta^*(N, \boldsymbol{\beta}) = \int_0^N \int_{\mathfrak{B}(t)} \Upsilon_\eta(\mathbf{u}) \, d\mathbf{u} \, dt \tag{5.15}$$

where $\mathfrak{B}(t)$ is the set of $\mathbf{u} \in \mathbb{R}^r$ with $0 \le u_j + \lambda_j t + \beta_j \le N$ for all $j = 1, \dots, r$.

So far, all estimates have been uniform with respect to $\boldsymbol{\beta}$. It is now necessary to assume that $\boldsymbol{\beta}$ is small, whence we suppose that $\|\boldsymbol{\beta}\| \le \sqrt{N}$. Then, for $N^{2/3} \le t \le \kappa^{-1}N - N^{2/3}$, we have $\lambda_j N^{2/3} - N^{1/2} \le \lambda_j t + \beta_j \le N - \lambda_j N^{2/3} + N^{1/2}$, and hence, for large N, the box $\|\mathbf{u}\| \le 1$ is contained in $\mathfrak{B}(t)$. Hence, for these t, and for $0 < \eta \le 1$, one has

$$\int_{\mathfrak{B}(t)} \Upsilon_\eta(\mathbf{u})\, d\mathbf{u} = \Big(\int_{-1}^{1} \Upsilon_\eta(u) du \Big)^r = \eta^r .$$

Similarly, when $t \ge \kappa^{-1}N + N^{2/3}$ and N is large, there is at least one j for which the intersection of the intervals $0 \le u_j + \lambda_j t + \beta_j \le N$ and $|u_j| \le 1$ is empty. Hence, in this case, the inner integral in (5.15) vanishes. In the regions $0 \le t \le N^{2/3}$ and $|t - \kappa^{-1}N| \le N^{2/3}$ it suffices to use the upper bound

$$\int_{\mathfrak{B}(t)} \Upsilon_\eta(\mathbf{u})\, d\mathbf{u} \le \int_{\mathbb{R}^r} \Upsilon_\eta(\mathbf{u})\, d\mathbf{u} = \eta^r$$

that is actually valid for all $t \in \mathbb{R}$. Now split the outer integral into the regions $0 \le t \le N^{2/3}$, $N^{2/3} < t \le \kappa^{-1}N - N^{2/3}$, $\kappa^{-1}N - N^{2/3} < t \le \min(N, \kappa^{-1}N + N^{2/3})$, larger $t \le N$ (if any) do not contribute by the above observations. Then, on collecting together,

$$P^*_\eta(N, \boldsymbol{\beta}) = \kappa^{-1}N\eta^r + O(N^{2/3}\eta^r) .$$

By (5.14) we have proved the first statement in the following lemma.

Lemma 5. *Let $\mathcal{S}_0, \dots, \mathcal{S}_r$ denote extremal sequences, and let $\lambda_1, \dots, \lambda_r$ denote positive real numbers such that $1, \lambda_1, \dots, \lambda_r$ are linearly independent over $\mathbb{Q}$. Then there exists a function $U^*(N)$ with $U^*(N) \to \infty$ as $N \to \infty$, and such that uniformly in $0 < \eta \le 1$, $\|\boldsymbol{\beta}\| \le \sqrt{N}$ and $1 \le U \le U^*(N)$ one has*

$$\int_{\|\boldsymbol{\alpha}\| \le U/N} S_1(\alpha_1) \dots S_r(\alpha_r) S_0(-\boldsymbol{\lambda\alpha}) e(-\boldsymbol{\alpha\beta}) K_\eta(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}$$
$$= \varrho\kappa^{-1}\eta^r N + O(NU^{-\frac{1}{r(r+1)}})$$

and

$$\int_{U/N \le \|\boldsymbol{\alpha}\| \le U^*/N} |S_1(\alpha_1) \dots S_r(\alpha_r) S_0(-\boldsymbol{\lambda\alpha})|\, d\boldsymbol{\alpha} \ll NU^{-\frac{1}{r(r+1)}} .$$

Here κ and ϱ are defined by (5.4).

To establish the final inequality in Lemma 5, apply the triangle inequality to (5.9) and integrate over $U/N \le \|\boldsymbol{\alpha}\| \le U^*/N$. Then, by (5.8),

$$\int_{U/N \le \|\boldsymbol{\alpha}\| \le U^*/N} |S_1(\alpha_1) \dots S_r(\alpha_r) S_0(-\boldsymbol{\lambda\alpha})|\, d\boldsymbol{\alpha}$$
$$\ll \int_{\mathfrak{k}} |I(\boldsymbol{\alpha}) I(-\boldsymbol{\lambda\alpha})|\, d\boldsymbol{\alpha} + NU^{*r}V^{-1} .$$

But $U^* = V^{1/(r+1)}$, and the required bound follows from (5.13).

6. Counting continued

In this section we complete the analysis of the weighted counter $P_\eta(N,\boldsymbol{\beta})$, but before we direct our forces to the treatment of the minor arc portion of the integral (5.6), we pause for two auxiliary identities that will frequently be used later.

Let $S(\alpha)$ denote any exponential sum, as in (4.4). Let ν,μ,μ' denote real numbers with $\nu\neq 0$. Then, provided only that $|\nu|\geq\eta>0$, we deduce from (5.2) the *convolution formula*

$$\int_{-\infty}^{\infty} S(\nu\alpha+\mu)S(-\nu\alpha-\mu')K_\eta(\alpha)\,d\alpha=\sum_{\substack{s,s'\leq N\\ s,s'\in\mathcal{S}}}\Upsilon_\eta(\nu(s-s'))e(\mu s-\mu' s')=S(\mu-\mu'). \tag{6.1}$$

Our second tool concerns functions $f\in L^1(\mathbb{R}/\mathbb{Z})$ that we view as functions of period 1 on the real line. Then, by Lemma 1 of Brüdern and Perelli [6], or Brüdern [3], eqn. (4), for any function $K\in L^1(\mathbb{R})$ that has a compactly supported Fourier transform $\hat{K}$ one has

$$\int_{-\infty}^{\infty} f(\alpha)K(\alpha)e(-\alpha\beta)\,d\alpha=\sum_{n=-\infty}^{\infty}\hat{K}(\beta-n)\int_0^1 f(\alpha)e(-\alpha n)\,d\alpha; \tag{6.2}$$

this is readily proved by Poisson summation ([3], p. 63). In particular, we have

$$\int_{-\infty}^{\infty} f(\alpha)K_\eta(\alpha)\,d\alpha=\int_0^1 f(\alpha)\,d\alpha \quad (0<\eta\leq 1)\,.$$

Two special cases are of interest to us. With $S(\alpha)$ as before and $\mathfrak{m}(Q)$ and $\mathfrak{n}(Q)=\mathfrak{m}(Q)\cap[0,1]$ as in section 3, one has

$$\int_{\mathfrak{m}(Q)}|S(\alpha)|^2K_\eta(\alpha)\,d\alpha=\int_{\mathfrak{n}(Q)}|S(\alpha)|^2\,d\alpha. \tag{6.3}$$

To see this, it suffices to take $f(\alpha)=|S(\alpha)|^2$ when $\alpha\in\mathfrak{m}(Q)$, and $f(\alpha)=0$ otherwise, in the preceding identity. Similarly, one finds

$$\int_{\mathfrak{M}(Q)}K_\eta(\alpha)\,d\alpha=\int_{\mathfrak{N}(Q)}d\alpha\ll Q^2/N. \tag{6.4}$$

We return to the main theme, and use the notation from section 5 throughout. The objective is to estimate the quantity

$$\int_{\mathfrak{k}} S_0(\alpha_0)S_1(\alpha_1)\dots S_r(\alpha_r)e(-\boldsymbol{\alpha\beta})K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}=\mathcal{I}(\mathfrak{k}) \tag{6.5}$$

say; here $\alpha_0=-\boldsymbol{\lambda\alpha}$ as on earlier occasions. Some preparation is required before we can bring Theorem 3 into play. First we apply Lemma 4 to each of the sequences $\mathcal{S}_0,\dots,\mathcal{S}_r$. This supplies functions $H_j(t)$ that decrease to 0 as

$t \to \infty$, and increasing unbounded functions $Q_j(N)$, such that $|S_j(\alpha)| \ll N\Xi_j(\alpha)$ holds for $\alpha \in \mathfrak{M}(Q_j(N))$; here Ξ_j is the weight associated to H_j. Now take $H = \max(H_0, H_1, \dots, H_r)$ and $Q = \min(Q_0, \dots, Q_r)$. Then

$$S_j(\alpha) \ll N\Xi(\alpha) \text{ for } \alpha \in \mathfrak{M}(Q) \tag{6.6}$$

holds for all $0 \le j \le r$. With this choice of $Q(N)$ and $U^*(N)$ in the role of U, apply Theorem 3. By the comment at the end of section 2, the relevant coefficient matrix has integral rank 0, and there exists a function $T(N)$, tending to ∞ with N, with $T(N) \le U^*(N)$ and such that

$$\sup_{U^*(N)/N \le \|\boldsymbol{\alpha}\| \le T(N)} \Xi(\alpha_0)\Xi(\alpha_1)\dots\Xi(\alpha_r) \ll T(N)^{-1}. \tag{6.7}$$

The set $\mathfrak{k}$ is to be dissected into $r+2$ subsets that must be considered separately. We need to define various functions before we can describe this in detail. First, we apply (4.7) in combination with (6.3) to $S = S_j$. This shows that for any function V_j with $V_j(N) \to \infty$ as $N \to \infty$ there exists a function U_j that also increases to infinity, and such that

$$\int_{\mathfrak{m}(V_j)} |S_j(\alpha)|K_\eta(\alpha)\, d\alpha \ll NU_j^{-1}. \tag{6.8}$$

Here we may suppose that $U_j \le V_j$ for otherwise U_j can be replaced by $\min(U_j, V_j)$. Now choose functions $V_0, V_r, V_{r-1}, \dots, V_1$ as follows: let

$$V = V_0 = \min(Q, T^{1/4r}) .$$

This defines a function U_0 with $U_0(N) \to \infty$ as $N \to \infty$, and such that (6.8) holds for $j = 0$. Next, put $V_r = U_0^{1/5r}$ and choose an increasing, unbounded function U_r such that (6.8) holds for $j = r$. Proceed likewise and recursively put

$$V_j = U_{j+1}^{1/5j} \quad (1 \le j \le r-1). \tag{6.9}$$

Note that $U_j \le V_j$ implies the inequalities

$$V_j \le V_{j+1}^{1/5j} \quad (1 \le j \le r-1), \quad V_r^{1/5r} \le V_0 = V .$$

The functions V_j now being defined, we choose $U(N) = \min(U_0(N), V_1(N))$ in the work of section 5. This fixes the definition of $\mathfrak{k}$, and Lemma 5 is applicable. Now write

$$\mathfrak{h}_l = \{\boldsymbol{\alpha} \in \mathfrak{k} : \alpha_l \in \mathfrak{m}(V_l),\ \alpha_j \in \mathfrak{M}(V_j) \text{ for } 1 \le j \le l-1\},$$
$$\mathfrak{h}^* = \{\boldsymbol{\alpha} \in \mathfrak{k} : \alpha_0 \in \mathfrak{m}(V_0),\ \alpha_j \in \mathfrak{M}(V_j) \text{ for } 1 \le j \le r\},$$
$$\mathfrak{H} = \{\boldsymbol{\alpha} \in \mathfrak{k} : \alpha_j \in \mathfrak{M}(V_j) \text{ for } 0 \le j \le r\}.$$

Then $\mathfrak{k}$ is the disjoint union of $\mathfrak{h}_1, \ldots, \mathfrak{h}_r, \mathfrak{h}^*$ and $\mathfrak{H}$. Accordingly, by an appropriate interpretation of (6.5),

$$\mathcal{I}(\mathfrak{k}) = \mathcal{I}(\mathfrak{H}) + \mathcal{I}(\mathfrak{h}^*) + \sum_{l=1}^{r} \mathcal{I}(\mathfrak{h}_l). \tag{6.10}$$

The integral $\mathcal{I}(\mathfrak{H})$ is estimated in three steps. We consider the subsets

$$\mathfrak{H}^* = \{\boldsymbol{\alpha} \in \mathfrak{H} : U^*/N < \|\boldsymbol{\alpha}\| \le T\}, \quad \mathfrak{H}_* = \{\boldsymbol{\alpha} \in \mathfrak{H} : \|\boldsymbol{\alpha}\| > T\}$$

and then first observe that by Lemma 5 we have

$$\mathcal{I}(\mathfrak{H}) \ll |\mathcal{I}(\mathfrak{H}^*)| + |\mathcal{I}(\mathfrak{H}_*)| + NU^{-1/r(r+1)}\,.$$

Note that $\mathfrak{H} \subset \mathfrak{M}^r$ where $\mathfrak{M} = \mathfrak{M}(V)$. By (6.5) and (6.6), then (6.7) followed by (6.4), we find that

$$\begin{aligned}\mathcal{I}(\mathfrak{H}^*) &\ll N^{r+1} \int_{\mathfrak{H}^*} \Xi(\alpha_0)\Xi(\alpha_1)\ldots\Xi(\alpha_r)K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \\ &\ll N^{r+1}T^{-1} \int_{\mathfrak{M}^r} K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll NT^{-1}V^{2r} \ll NT^{-1/2}.\end{aligned}$$

For $\mathfrak{H}_*$ we apply brute force, estimate all $S_j(\alpha_j)$ trivially and then use symmetry in $\boldsymbol{\alpha}$ to confirm the bounds

$$\mathcal{I}(\mathfrak{H}_*) \ll N^{r+1} \int_{\mathfrak{H}_*} K_\eta(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll N^{r+1} \Big(\int_{\mathfrak{M}} K_\eta(\alpha)\,d\alpha \Big)^{r-1} \Big(\int_{\mathfrak{L}(T)} K_\eta(\alpha)\,d\alpha \Big)$$

where $\mathfrak{L}(T) = \mathfrak{M} \cap [T, \infty)$. But straightforward estimates based on (5.1) give

$$\int_{\mathfrak{L}(T)} K_\eta(\alpha)\,d\alpha \ll \sum_{m=[T]}^{\infty} m^{-2} \int_{\mathfrak{N}} d\alpha \ll T^{-1}V^2N^{-1}$$

so that (6.4) confirms the bound $\mathcal{I}(\mathfrak{H}_*) \ll NV^{2r}T^{-1}$. This combines to

$$\mathcal{I}(\mathfrak{H}) \ll NT^{-1/2} + NU^{-1/r(r+1)}. \tag{6.11}$$

We now turn our attention to the set $\mathfrak{h}_l$ with $1 \le l \le r$. Note that $U(N) \le V_j(N)$ for all $j = 0, \ldots, r$. In particular, when $1 \le l \le r$ and $\boldsymbol{\alpha} \in \mathbb{R}^r$ with $\alpha_l \in \mathfrak{m}(V_l)$, then $\|\boldsymbol{\alpha}\| \ge V_l/N$, and hence $\boldsymbol{\alpha} \in \mathfrak{k}$. This shows that $\mathfrak{h}_l$ is a cartesian product. It is convenient to introduce the following notation: for $1 \le l \le r$ and $\boldsymbol{\alpha} \in \mathbb{R}^s$, let $\boldsymbol{\alpha}_l = (\alpha_1, \ldots, \alpha_{l-1}, \alpha_{l+1}, \ldots, \alpha_r)$, and put

$$\mathfrak{F}_l = \{\boldsymbol{\alpha}_l \in \mathbb{R}^{r-1} : \alpha_j \in \mathfrak{M}(V_j)\ (1 \le j \le l-1)\}\,.$$

Then, by (6.5), (5.3) and Fubini's theorem,

$$\mathcal{I}(\mathfrak{h}_l) = \int_{\mathfrak{m}(V_l)} S_l(\alpha_l)e(-\alpha_l\beta_l)F_l(\alpha_l)K_\eta(\alpha_l)\,d\alpha_l \tag{6.12}$$

where

$$F_l(\alpha_l) = \int_{\mathfrak{F}_l} \prod_{\substack{j=0 \\ j\neq l}}^{r} S_j(\alpha_j) e(-\boldsymbol{\alpha}_l \boldsymbol{\beta}_l) K_\eta(\boldsymbol{\alpha}_l)\, d\boldsymbol{\alpha}_l, \tag{6.13}$$

here we continue to use the abbreviation $\alpha_0 = -\boldsymbol{\lambda\alpha}$. By Cauchy's inequality,

$$\mathcal{I}(\mathfrak{h}_l) \le \Big(\int_{\mathfrak{m}(V_l)} |S_l(\alpha)|^2 K_\eta(\alpha)\, d\alpha \Big)^{\frac{1}{2}} \mathcal{J}_l^{\frac{1}{2}} \tag{6.14}$$

where

$$\mathcal{J}_l = \int_{-\infty}^{\infty} |F_l(\alpha)|^2 K_\eta(\alpha)\, d\alpha\ .$$

We expand this last integral. On writing

$$\Phi_l(\boldsymbol{\alpha}_l) = \prod_{\substack{j=1 \\ j\neq l}}^{r} S_j(\alpha_j)$$

and

$$\Psi_l(\boldsymbol{\alpha}_l, \boldsymbol{\alpha}_l') = \int_{-\infty}^{\infty} S_0(-\lambda_l\alpha_l - \boldsymbol{\lambda}_l\boldsymbol{\alpha}_l) S_0(\lambda_l\alpha_l + \boldsymbol{\lambda}_l\boldsymbol{\alpha}_l') K_\eta(\alpha_l)\, d\alpha_l\ ,$$

one finds from (6.13) and Fubini's theorem that

$$\mathcal{J}_l = \int_{\mathfrak{F}_l}\int_{\mathfrak{F}_l} \Phi_l(\boldsymbol{\alpha}_l)\Phi_l(-\boldsymbol{\alpha}_l') e(\boldsymbol{\beta}_l(\boldsymbol{\alpha}_l' - \boldsymbol{\alpha}_l)) \Psi_l(\boldsymbol{\alpha}_l, \boldsymbol{\alpha}_l') K_\eta(\boldsymbol{\alpha}_l) K_\eta(\boldsymbol{\alpha}_l')\, d\boldsymbol{\alpha}_l\, d\boldsymbol{\alpha}_l'\ .$$

We suppose that $0 < \eta \le \min \lambda_j$. Then, by (6.1), $\Psi_l(\boldsymbol{\alpha}_l, \boldsymbol{\alpha}_l') = S_0(\boldsymbol{\lambda}_l(\boldsymbol{\alpha}_l' - \boldsymbol{\alpha}_l))$ and

$$\mathcal{J}_l = \int_{\mathfrak{F}_l}\int_{\mathfrak{F}_l} \Phi_l(\boldsymbol{\alpha}_l)\Phi_l(-\boldsymbol{\alpha}_l') S_0(\boldsymbol{\lambda}_l(\boldsymbol{\alpha}_l' - \boldsymbol{\alpha}_l)) e(\boldsymbol{\beta}_l(\boldsymbol{\alpha}_l' - \boldsymbol{\alpha}_l)) K_\eta(\boldsymbol{\alpha}_l) K_\eta(\boldsymbol{\alpha}_l')\, d\boldsymbol{\alpha}_l\, d\boldsymbol{\alpha}_l'. \tag{6.15}$$

The case $l = 1$ is particularly simple. Here $\mathfrak{F}_1 = \mathbb{R}^{r-1}$. Hence, by (5.3) and a consideration of the underlying diophantine problem, we see that when $l = 1$, the integral in (6.15) does not exceed the number of solutions of the system of inequalities

$$|\lambda_j w - x_j| < \eta, \quad |\lambda_j w - y_j| < \eta \quad (2 \le j \le r) \tag{6.16}$$

in integers w, x_j, y_j, all in the interval $[1, N]$. For $\eta \le 1$, the value of w fixes x_j and y_j within $O(1)$ possibilities. Hence, $\mathcal{J}_1 \ll N$, and by (6.14) and (6.8), it follows that

$$\mathcal{I}(\mathfrak{h}_1) \ll N U_1^{-1/2}. \tag{6.17}$$

The cases $2 \le l \le r$ can be treated along similar lines. We return to (6.15) and carry out all integrations against $\alpha_{l+1}, \dots, \alpha_r, \alpha'_{l+1}, \dots, \alpha'_r$, with the aid of 5.3. On writing $\tilde{\boldsymbol{\alpha}} = (\alpha_1, \dots \alpha_{l-1})$,

$$\mathfrak{G}_l = \mathfrak{M}(V_1) \times \dots \times \mathfrak{M}(V_{l-1})$$

and

$$\Theta_l(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}') = \sum_{w, x_j, y_j} \prod_{j=l+1}^{r} \Upsilon_\eta(x_j - \lambda_j w - \beta_j) e\Big(w \sum_{i=1}^{l-1} \lambda_i(\alpha'_i - \alpha_i)\Big) ,$$

where the summation is over all $w \in \mathcal{S}_0$, $x_j, y_j \in \mathcal{S}$ for $1 \le j \le l-1$ that satisfy (6.16) and lie in the interval $[1, N]$, one finds that 6.15 reduces to

$$\mathcal{J}_l = \int_{\mathfrak{G}_l} \int_{\mathfrak{G}_l} \Theta_l(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}') \prod_{j=1}^{l-1} S_j(\alpha_j) S_j(-\alpha'_j) e(\beta_j(\alpha'_j - \alpha_j)) K_\eta(\tilde{\boldsymbol{\alpha}}) K_\eta(\tilde{\boldsymbol{\alpha}}')\, d\tilde{\boldsymbol{\alpha}} d\tilde{\boldsymbol{\alpha}}' .$$

All remaining exponential sums will now be estimated trivially. As we saw in the case $l = 1$, there are only $O(N)$ tuples w, x_j, y_j in the summation conditions for the sum defining Θ. Hence, $\Theta \ll N$, and therefore, we may conclude from (6.4) that

$$\mathcal{J}_l \ll N^{2l-1} \prod_{j=1}^{l-1} \Big(\int_{\mathfrak{M}(V_j)} K_\eta(\alpha_j)\, d\alpha_j \Big)^2 \ll N(V_1 \dots V_{l-1})^4 .$$

By (6.14) and (6.8), it follows that for $2 \le l \le r$ one has

$$\mathcal{I}(\mathfrak{h}_l) \ll N(V_1 \dots V_{l-1})^2 U_l^{-1/2} \ll N U_l^{-1/10} . \tag{6.18}$$

Finally, we consider $\mathfrak{h}^*$. Here we apply Schwarz's inequality directly to (6.5) to obtain

$$\mathcal{I}(\mathfrak{h}^*) \le (\mathcal{I}_1 \mathcal{I}_2)^{1/2}$$

with

$$\mathcal{I}_1 = \int_{\mathfrak{h}^*} |S_0(-\boldsymbol{\lambda\alpha})|^2 K_\eta(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}, \quad \mathcal{I}_2 = \int_{\mathbb{R}^r} |S_1(\alpha_1) \dots S_r(\alpha_r)|^2 K_\eta(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} .$$

The second integral factorizes into

$$\int_{-\infty}^{\infty} |S_j(\alpha)|^2 K_\eta(\alpha)\, d\alpha = \#\{n \le N : n \in \mathcal{S}_j\} \ll N$$

whence $\mathcal{I}_2 \ll N^r$. For $\mathcal{I}_1$, we substitute $\alpha_0 = -\boldsymbol{\lambda\alpha}$ for α_r as a new variable of integration. Then, with $\boldsymbol{\alpha}_r = (\alpha_1, \dots, \alpha_{r-1})$ as before,

$$\mathcal{I}_1 \le \frac{1}{\lambda_r} \int_{\mathfrak{G}_r} \int_{\mathfrak{m}(V_0)} |S_0(\alpha_0)|^2 K_\eta(\alpha_r) K_\eta(\boldsymbol{\alpha}_r)\, d\alpha_0\, d\boldsymbol{\alpha}_r .$$

Here α_r is the function of α_0 and $\boldsymbol{\alpha}_r$ defined via $\alpha_0 = -\lambda_r\alpha_r - \boldsymbol{\lambda}_r\boldsymbol{\alpha}_r$. We can rewrite the inner integral as

$$\int_{\mathfrak{m}(V_0)} |S_0(\alpha_0)|^2 K_\eta\Big(\frac{\alpha_0 + \boldsymbol{\lambda}_r\boldsymbol{\alpha}_r}{\lambda_r}\Big)\, d\alpha_0 \ .$$

The Fourier transform of $\alpha \mapsto K_\eta(\lambda_r^{-1}(\alpha + \boldsymbol{\lambda}_r\boldsymbol{\alpha}_r))$ is $\beta \mapsto \Upsilon_\eta(\lambda_r\beta)\lambda_r e(\beta\boldsymbol{\lambda}_r\boldsymbol{\alpha}_r)$. Hence, by (6.2), when $\eta \le \lambda_r$,

$$\int_{\mathfrak{m}(V_0)} |S_0(\alpha_0)|^2 K_\eta\Big(\frac{\alpha_0 + \boldsymbol{\lambda}_r\boldsymbol{\alpha}_r}{\lambda_r}\Big)\, d\alpha_0 = \lambda_r \int_{\mathfrak{n}(V_0)} |S_0(\alpha)|^2\, d\alpha \ .$$

When combined with (6.3) and (6.4), it follows that

$$\mathcal{I}_1 \ll NU_0^{-1} \int_{\mathfrak{G}_r} K_\eta(\boldsymbol{\alpha}_r)\, d\boldsymbol{\alpha}_r \ll N^{2-r}U_0^{-1}(V_1 \dots V_{r-1})^2 \ll N^{2-r}U_0^{-1/2} \ .$$

Collecting together, we deduce that $\mathcal{I}(\mathfrak{h}^*) \ll NU_0^{-1/4}$. By (6.10), (6.11), (6.17) and (6.18) we finally confirm the bound $\mathcal{I}(\mathfrak{k}) = o(N)$ as $N \mapsto \infty$ When combined with (5.6) and Lemma 5, we have reached the following conclusion.

Theorem 6. *Let $\mathcal{S}_0, \dots, \mathcal{S}_r$ denote extremal sequences. Let $\lambda_1, \dots, \lambda_r$ be positive real numbers such that $1, \lambda_1, \dots, \lambda_r$ are linearly independent over $\mathbb{Q}$. Define $P_\eta(N, \boldsymbol{\beta})$ by (5.5). There exists a function $U(N)$ with $U(N) \to \infty$ as $N \to \infty$, and such that*

$$P_\eta(N, \boldsymbol{\beta}) = \varrho\kappa^{-1}\eta^r N + O(NU^{-1})$$

holds uniformly in $0 < \eta \le \min \lambda_j$ and $|\boldsymbol{\beta}| \le \sqrt{N}$. Here ϱ and κ are defined by (5.4).

We close this section with a brief comment on the method. The kernel in our Fourier approach factorizes, as in (5.3). It seems very important to use such a kernel. This property has been used ubiquitously, but most notably in (6.12). The second factor in (6.14) reduces, at least in part, to a diophantine system of linear inequalities. This provides a crucial saving over more traditional methods of estimation which would require $2r+1$ variables for r inequalities, at the very least. The price we have to pay for this is rather low; it is the tedious and unsymmetric construction of the functions V_j in (6.9), with consequential asymmetries in the sets $\mathfrak{h}_1, \dots, \mathfrak{h}_l$ later. One can use the techniques of this section also in the context of diophantine equations. For example, one can work along the lines of sections 5 and 6 of this paper to give an alternative, more direct proof of the k-tuple-Theorem 1.7 from our recent memoir [5].

7. The main theorem

The weights can be removed from the counting function $P_\eta(N, \boldsymbol{\beta})$ in Theorem 6, and we may consider more general linear forms than those covered by the methods of sections 5 and 6. With this in view, let $\Lambda = (\lambda_{ij})_{\substack{1\le i\le r\\ 0\le j\le r}}$ denote a real matrix, and fix a choice of extremal sequences $\mathcal{S}_0, \dots, \mathcal{S}_r$. For real numbers $\tau_j > 0$ define $Z_{\boldsymbol{\tau}}(N, \boldsymbol{\mu})$ as the number of solutions of the system

$$\Big|\sum_{j=0}^{r} \lambda_{ij} x_j - \mu_i\Big| \le \tau_i \quad (1 \le i \le r) \tag{7.1}$$

with $x_j \in \mathcal{S}_j$ and $1 \le x_j \le N$ for $0 \le j \le r$.

Theorem 7. *Suppose that the real $r \times (r+1)$-matrix Λ is highly non-singular, positive and of integral rank 0. Then there exists a constant $c > 0$, depending only on Λ, and such that the asymptotic formula*

$$Z_{\boldsymbol{\tau}}(N, \boldsymbol{\mu}) = c\varrho N \tau_1 \tau_2 \dots \tau_r + o(N)$$

holds uniformly for all $|\boldsymbol{\mu}| \le \sqrt{N}$ and for all positive τ_j that are sufficiently small in terms of Λ.

Proof. Choose a matrix $\mathrm{A} \in GL_r(\mathbb{R})$ such that $\mathrm{A}\Lambda$ takes the shape (2.2); then all λ_j are positive, and by the remark at the end of section 2, the numbers $1, \lambda_1, \dots, \lambda_r$ are linearly independent over $\mathbb{Q}$. Let

$$\mathcal{D} = \{\boldsymbol{\alpha} \in \mathbb{R}^r : |\alpha_j - \mu_j| \le \tau_j \quad (1 \le j \le r)\} .$$

Then $\mathrm{A}\mathcal{D}$ is a parallelotope, and the inequalities (7.1) are equivalent with the condition that

$$(\lambda_1 x_0 - x_1, \lambda_2 x_0 - x_2, \dots, \lambda_r x_0 - x_r) \in \mathrm{A}\mathcal{D} .$$

Choose $U(N)$ in accordance with Theorem 6, and let $\eta = U^{-1/4r}$. Suppose that $\tau_j \ge \eta^{1/4}$ holds for all j. Then put

$$\mathcal{D}^+ = \{\boldsymbol{\alpha} : |\alpha_j - \mu_j| \le \tau_j + \sqrt{\eta}\}, \quad \mathcal{D}^- = \{\boldsymbol{\alpha} : |\alpha_j - \mu_j| \le \tau_j - \sqrt{\eta}\}$$

Now consider the set of lattice points $\mathcal{B}^\pm = \eta\mathbb{Z}^r \cap \mathrm{A}\mathcal{D}^\pm$. The volume of $\mathcal{D}^\pm$ is $\tau_1\tau_2 \dots \tau_r(1 + O(\eta^{1/4}))$, and hence the volume $\mathrm{A}\mathcal{D}^\pm$ equals $|\det \mathrm{A}|\tau_1\tau_2 \dots \tau_r(1 + O(\eta^{1/4}))$. By standard lattice point arguments, we find that

$$\#\mathcal{B}^\pm = |\det A|\tau_1 \dots \tau_r \eta^{-r}(1 + O(\eta^{1/4})). \tag{7.2}$$

Now write

$$Z^\pm = \sum_{\boldsymbol{\beta} \in \mathcal{B}^\pm} P_\eta(N, \boldsymbol{\beta}) .$$

By (5.1) and (5.5), one readily confirms that whenever N is large one has

$$Z^- \le Z_{\boldsymbol{\tau}}(N, \boldsymbol{\beta}) \le Z^+ .$$

By Theorem 6 and (7.2), one finds that for some constant c one has

$$Z^{\pm} = \varrho c\tau_1 \dots \tau_r N + O(NU^{-3/4}) + O(N\tau_1\tau_2 \dots \tau_r \eta^{1/4}) .$$

Theorem 7 is now immediate. This also proves Theorem 1 because the square-free numbers form an extremal sequence. ■

It is perhaps of interest that Theorem 7 may be generalized to the case of non-zero integral rank. In such a situation there are hidden diophantine equations in the system (7.1), as we pointed out in §2. The asymptotic formula for $Z_{\boldsymbol{\tau}}(N, \boldsymbol{\mu})$ will reflect this with the appearance of a suitable singular series that, however, is not necessarily positive. When Λ is of integral rank r, then the desired formula is actually contained in [5], chapter 5, and the intermediate cases can be handled by the methods of this paper and the intersection principles of [5], chapter 4.

8. Outroduction

Suppose that $0 < \lambda < \frac{1}{2}$ is an irrational number, and let $Z_\tau(N)$ denote the number of solutions of $|m - \lambda n| < \tau$ with $1 \le n, m \le N$. This is the simplest situation where Theorem 7 is applicable. The work of Hardy and Littlewood [13], section 4.4 is readily modified to show that in the current context the error term in the formula $Z_\tau(N) = 2\tau N + o(N)$ is sharp: for any function $T(N)$ increasing to infinity with N, there exists an irrational $\lambda \in (0, \frac{1}{2})$ such that $|Z_\tau(N) - 2\tau N| > N/T(N)$ holds on a sequences of values for N tending to infinity. In particular, the error term in Theorem 7 crucially depends on λ. One would expect that this is so for any r, and any choice of extremal sequences $\mathcal{S}_0, \dots, \mathcal{S}_r$. However, a more thorough discussion of this matter requires another development of our methods that we intend to communicate on a future occasion.

References

[1] O.D. Atkinson, J. Brüdern, R.J. Cook, *Simultaneous additive congruences to a large prime modulus*, Mathematika **39** (1992), 1–9.

[2] V. Bentkus, F. Götze, *Lattice point problems and the distribution of values of quadratic forms*, Ann. of Math. **2**(150) (1999), 977–1027.

[3] J. Brüdern, *The Davenport-Heilbronn Fourier transform method, and some Diophantine inequalities, Number theory and its applications* (Kyoto, 1997), 59–87, Dev. Math., **2**, Kluwer Acad. Publ., Dordrecht, 1999.

[4] J. Brüdern, An elementary harmonic analysis of arithmetic functions, to appear.

[5] J. Brüdern, *Binary additive problems and the circle method, multiplicative sequences and convergent sieves. Analytic number theory – Essays in honour of Klaus Roth*, (W.Chen, T. Gowers, H. Halberstam, W.M. Schmidt, R.C. Vaughan eds.), University Press, Cambridge, 2008, pp. 91–132.

[6] J. Brüdern, A. Perelli, *Goldbach numbers in sparse sequences*, Ann. Inst. Fourier (Grenoble) **48** (1998), 353–378.
[7] J. Brüdern, T.D. Wooley, Pairs of diagonal cubic forms revisited, To appear.
[8] H. Davenport, H. Heilbronn, *On indefinite quadratic forms in five variables*, J. London Math. Soc. **21** (1946), 185–193.
[9] R. Dietmann, *Simultaneous diophantine approximation by square-free numbers*, Quart. J. Math., **59** (2008), 311–319.
[10] E. Freeman, *Quadratic Diophantine inequalities*, J. Number Theory **89** (2001), 268–307.
[11] E. Freeman, *Asymptotic lower bounds for Diophantine inequalities*, Mathematika **47** (2000), 127–159.
[12] E. Freeman, *Systems of diagonal diophantine inequalities*, Trans. Amer. Math. Soc. **355** (2003), 2675–2713.
[13] G.H. Hardy, J.E. Littlewood, *Some problems of diophantine approximation: the lattice points of a right-angled triangle*, Proc. London Math. Soc. **2**(20) (1921), 15–36.
[14] D.R. Heath-Brown, *Diophantine approximation with square-free numbers*, Math. Z. **187** (1984), 335–344.
[15] S.T. Parsell, *Irrational linear forms in prime variables*, J. Number Theory **97** (2002), 144–156.
[16] J.-C. Puchta, *An additive property of almost periodic sets*, Acta Math. Hungar. **97** (2002), 323–331.
[17] R.C. Vaughan, *The Hardy-Littlewood method. 2nd edn.*, University Press, Cambridge 1997.
[18] T.D. Wooley, *On Diophantine inequalities: Freeman's asymptotic formulae. Proceedings of the Session in Analytic Number Theory and Diophantine Equations*, 32 pp., Bonner Math. Schriften, **360** Univ. Bonn, Bonn 2003.

Address: Institut für Algebra und Zahlentheorie
Universität Stuttgart
Pfaffenwaldring 57
D-70511 Stuttgart

E-mail: bruedern@mathematik.uni-stuttgart.de

Received: 31 December 2007; **revised:** 12 July 2008