

On simultaneous diagonal equations and inequalities

by

J. BRÜDERN (Göttingen) and R. J. COOK (Sheffield)

1. Introduction. Let F_1, \dots, F_R be forms of degree k with real coefficients in N variables. We shall be concerned with simultaneous small values of the system (F_1, \dots, F_R) at integer points. In other words, we seek for solutions to the system of inequalities

$$(1) \quad |F_1(\mathbf{x})| < \tau, \quad \dots, \quad |F_R(\mathbf{x})| < \tau$$

where $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N$. In this generality, an affirmative answer has been found by Schmidt [36] provided that k is odd and N is very large compared with k and R . In the case of integer coefficients the system (1) reduces to a system of diophantine equations,

$$(2) \quad F_1(\mathbf{x}) = F_2(\mathbf{x}) = \dots = F_R(\mathbf{x}) = 0,$$

and this specific case of Schmidt's result is due to Birch [8].

In the present paper we restrict our attention to diagonal forms,

$$(3) \quad F_i = \sum_{j=1}^N \lambda_{ij} x_j^k \quad (1 \leq i \leq R).$$

Then it is possible to find solutions of (1) and (2) in a smaller number of variables. In the case of equations, Davenport and Lewis [29] showed that for k odd the condition

$$(4) \quad N \geq [9R^2 k \log 3Rk]$$

assures the existence of a non-trivial solution of (2). When $k \geq 4$ is even they needed

$$(5) \quad N \geq [48R^2 k^3 \log 3Rk^2]$$

variables and an additional "rank condition" on the matrix of coefficients

$$(6) \quad \Lambda = (\lambda_{ij})$$

to be specified later. Here and throughout it is assumed that no column of Λ vanishes so that all variables occur explicitly in the system of forms.

There are a number of improvements known for some small values of k and R ; see Davenport and Lewis [28], Cook [16–19, 21–23], Vaughan [38], Baker and Brüdern [7], Brüdern [12], Atkinson, Brüdern and Cook [2]. For analogous results on inequalities, see the results of Pitman [34] and Nadesalingam [33].

In all published results to date the dependence of N on R and k is still as bad as in (4) and (5) although recent work of Low, Pitman and Wolff [32] strongly indicates that N can be taken growing linearly in R . Their simple handling of certain combinatorial problems which are a familiar occurrence in the subject, and recent improvements on the Hardy–Littlewood method due to Vaughan [40] actually permit us to prove results of this strength by a fairly simple argument. We first state a typical result of this kind in the case of equations.

Let M be an $r \times t$ matrix over a field Γ . For $0 \leq d \leq r$ let $\mu(d) = \mu(d, M)$ be the maximal number of columns from M generating a linear space of dimension d in Γ^r . It is obvious that $\mu(d)$ is invariant under row operations applied to M .

THEOREM 1. *Suppose that F_i ($1 \leq i \leq R$) are given by (3) with $\lambda_{ij} \in \mathbb{Z}$. Let $N > n_0 R$ where $n_0 = n_0(k)$ is given by Table 1 when k is small, and where $n_0(k) = 2k(\log k + O(\log \log k))$ when k is large. Suppose that Λ contains an $R \times (n_0 R + 1)$ submatrix Λ^* satisfying*

$$(7) \quad \mu(d, \Lambda^*) \leq n_0 d \quad (0 \leq d \leq R - 1).$$

Suppose that (2) has a non-singular solution in the p -adic field, for all primes p including infinity. Then the number Θ of integer solutions to (2) in the box $\max |x_j| \leq P$ satisfies

$$(8) \quad \Theta \gg P^{N-Rk}.$$

Table 1

k	2	3	4	5	6
$n_0(k)$	4	8	12	18	30

Of course the existence of p -adic solutions is a necessary proviso in the theorem. It might be of interest to note that the rank condition (7) cannot be removed completely (although it certainly can be relaxed somewhat). When k is even this has been observed by Davenport and Lewis [29] (compare their remarks following Theorem 2). However, for odd k , there are indications that some condition like (7) is necessary as well. Cassels and Guy [15] found an example of a diagonal cubic equation,

$$(9) \quad a_1 x_1^3 + a_2 x_2^3 + a_3 x_3^3 + a_4 x_4^3 = 0$$

with integer coefficients, which has non-trivial solutions in all p -adic fields, but no non-trivial integer solution. Now consider the system

$$\begin{aligned} a_1x_1^3 + a_2x_2^3 + a_3x_3^3 + a_4x_4^3 &= 0, \\ b_1x_1^3 + b_2x_2^3 + \dots + b_Nx_N^3 &= 0. \end{aligned}$$

Any solution in integers must have $x_1 = x_2 = x_3 = x_4 = 0$. We may choose b_5, \dots, b_N all non-zero so that there are non-singular p -adic solutions for all p , but the number of integer solutions Θ with $|x_i| \leq P$ is of order $P^{(N-4)-3}$ for $N > 13$, say, contrary to (8). This follows by a straightforward application of the Hardy–Littlewood method to the second equation.

On the other hand, many of the hypotheses in Theorem 1 are redundant. For example when k is odd, it is an elementary exercise to show that a non-singular real solution always exists. However, for even k this condition is vital. Moreover, Dörner [31] has shown, in a more general context, that there are non-trivial p -adic solutions provided $N > 2Rk$ and p is sufficiently large. Atkinson, Brüdern and Cook [3] have proved the same result for all primes

$$(10) \quad p > k^{2R+2}.$$

Combining this with the argument used to prove Theorem 4 of Davenport and Lewis [29] we deduce that for all primes satisfying (10) the diagonal system (2) has a non-singular p -adic solution providing (7) and $N > n_0R$ hold. Note that this restriction on N is much stronger than $N > 2Rk$. It is therefore almost obvious that the bound (10) can be reduced further, a topic to which we hope to return on another occasion. But there will remain a set of small primes for which there are no solutions to (2) in the p -adic field. Davenport and Lewis [28] gave an example with $k = 3$, $R = 2$ and $p = 7$.

In the light of (4) and (5) it is also of considerable interest to obtain better conditions on N such that (2) always has non-trivial p -adic solutions, for all primes p . Low, Pitman and Wolff [32] proved this for

$$(11) \quad N \geq 48Rk^3 \log(3Rk^2).$$

For odd k they also obtained the bound

$$(12) \quad N \geq (2 + o(1))R^2k \log k$$

as R tends to infinity. There are a number of similar results in the literature which in some cases are somewhat sharper than (11) or (12); see the papers of Schmidt [37] and Low, Pitman and Wolff [32] for further references. On combining this with the ideas from Theorem 4 of Davenport and Lewis [29] and Lemma 3 of Low, Pitman and Wolff [32] it is easy to deduce from Theorem 1 the following

COROLLARY. *Suppose N, R, k satisfy (11) or (12). Then the diagonal system (2) always has a non-trivial solution.*

We spare the reader the details. Note that for fixed k , the bound (11) grows as $R \log R$ whereas in both (10) and Theorem 1 the bound on N depends on R linearly. This raises the question whether the $\log R$ can be removed from (11).

In the case of diagonal diophantine inequalities p -adic conditions should not show up since we consider the arithmetic properties of F_1, \dots, F_R with respect to the archimedean valuation only. In particular, the lower bound on N in Theorem 1 alone should be enough to ensure the existence of a solution to (1). Theorem 2 below shows that in fact this is true for a large class of forms. To state the result we require some notation. With the matrix A we associate linear forms in $R + 1$ variables. Let $\mathcal{J} \subset \{1, \dots, N\}$ with $|\mathcal{J}| = R + 1$. Let $A_{\mathcal{J}}$ be the $(R + 1) \times (R + 1)$ matrix formed with the columns

$${}^t(\lambda_{1,j}, \dots, \lambda_{R,j}, \xi_j), \quad j \in \mathcal{J},$$

and let

$$(13) \quad L_{\mathcal{J}}(\boldsymbol{\xi}_{\mathcal{J}}) = \det A_{\mathcal{J}}$$

where $\boldsymbol{\xi}_{\mathcal{J}} = (\xi_j)_{j \in \mathcal{J}}$.

DEFINITION. Let $L(\boldsymbol{\xi})$ be a linear form over the reals in t variables. The supremum of all real numbers σ such that the inequalities

$$|L(\boldsymbol{x})| < U^{-\sigma}, \quad 0 < \max |x_i| \leq U$$

have an integer solution for all sufficiently large U is called the *order* of L .

For a brief discussion of linear forms of finite order see Cook [20]. We remark that by a well known result of Schmidt [35] linear forms with algebraic coefficients are of finite order.

THEOREM 2. *Suppose that $\lambda_{ij} \in \mathbb{R}$, that $N = n_0(k)R + 1$, and that at least one of the linear forms (13) is of finite order and has coefficients linearly independent over the rationals. Suppose that $\mu(d) \leq n_0 d$ holds for $0 \leq d \leq R - 1$. Then the number Θ_2 of solutions to the diagonal inequalities (1) in integers with $|x_i| \leq P$ satisfies*

$$\limsup_{P \rightarrow \infty} P^{Rk-N} \Theta_2 > 0.$$

It should be noted that the hypotheses in Theorems 1 and 2 have been designed to allow a particularly simple proof. At the cost of combinatorial difficulties it is possible to reduce the number of variables slightly. However, the improvements are significant for smaller values of R and k only. We illustrate this in the following result which shows, roughly speaking, that in Theorem 2 one may take $n_0(3) = 7$.

THEOREM 3. *Let $k = 3$ and $N = 7R + 1$. Suppose that at least one of the linear forms (13) is of finite order and has coefficients linearly independent over the rationals. Suppose that $\mu(d, A) \leq 11d/2$ for $1 \leq d \leq R - 1$. Then, with Θ_2 as in Theorem 2, we have $\limsup_{P \rightarrow \infty} P^{-4R-1} \Theta_2 > 0$.*

Again we have put the strong rank condition $\mu(d) \leq 11d/2$ to keep the argument reasonably simple. It should be possible to reduce this to $\mu(d) \leq 7d$, and we conjecture that $\Theta_2 > 0$ holds without any rank condition. When $R = 1$ the rank condition is void, and when $R = 2$ it can in fact be removed (see the result in Brüdern and Cook [13]). The next case is $R = 3$, and this case of Theorem 3 might be the most interesting one since the number of variables now is $N = 22$ as in the recent work of Atkinson, Brüdern and Cook [2] on the cognate problem of three additive cubic equations. At the end of the paper we shall present combinatorial arguments to reduce the rank condition in Theorem 3 to $\mu(d) \leq 7d$ when $R = 3$. A detailed discussion of this subject matter is postponed to Section 7.

It is clear that there is a result corresponding to Theorem 3 for cubic equations. However, in this specific situation even further refinements are possible by yet another method which is to be presented in a forthcoming publication [14].

2. Mean value estimates. All the results announced so far are in fact consequences of a single principle which depends on mean value estimates for modified Weyl sums. We write $e(\alpha) = \exp(2\pi i\alpha)$ and introduce the sequence

$$(14) \quad \mathcal{A}(P, Q) = \{x \leq P : p \mid x \Rightarrow p \leq Q\}.$$

Let

$$(15) \quad g_k(\alpha) = g(\alpha) = \sum_{x \in \mathcal{A}(P, P^\eta)} e(\alpha x^k)$$

where $\eta > 0$ is a small positive constant to be determined later. One would expect that for $s \geq 2k$ the inequality

$$(16) \quad \int_0^1 |g(\alpha)|^s d\alpha \ll P^{s-k},$$

or at least the slightly weaker bound

$$(17) \quad \int_0^1 |g(\alpha)|^s d\alpha \ll P^{s-k+\varepsilon},$$

holds. This is a weakened form of Hardy and Littlewood's conjecture K on average.

PROPOSITION. *Suppose that $n_0 > 4k$ and that (17) holds for $s \geq n_0$. Then the statements of Theorems 1 and 2 hold for $N > Rn_0 + 1$. This is also true if $n_0 > 2k$ and (16) holds for $s \geq n_0$.*

When $k \geq 4$ it is known that (17) holds for $s \geq n_0$ where n_0 is as in Theorem 1, and η is sufficiently small in terms of ε (see Vaughan [40, §4] and the remarks preceding Theorem 1 of Vaughan and Wooley [41]). Moreover, we have $n_0 > 4k$ for $k \geq 6$. When $k = 4$ we know from Lemma 5.2 of Vaughan [40] that (16) holds for $s \geq 12$ (consider the underlying diophantine equations). Similarly (16) holds when $k = 5$ and $s = 18$; this follows from the results of [41]. Therefore when $k \geq 4$ Theorems 1 and 2 follow from the Proposition.

Note that when $k = 3$ we know that (17) holds for $s = 8$ by Hua's Lemma (Vaughan [39], Lemma 2.5). This remark also shows that there is actually no need to use the somewhat complicated sums (15) in this case. It would suffice to use the standard Weyl sums

$$(18) \quad f(\alpha) = \sum_{\eta P < x \leq P} e(\alpha x^3)$$

and their quadratic analogue. Thus we shall exclude the cases $k = 2$, $k = 3$ from the considerations for Theorems 1 and 2, partly because there is a simple proof based on (18), and mainly since there are more effective methods for the case $k = 3$, as we have pointed out in the introduction. It is of course possible to deduce the cases $k = 2$, $k = 3$ from the Proposition if a uniform treatment is wished.

Unfortunately when $k = 3$, $s = 7$, (17) is not known at present. The little extra which is required to prove Theorem 3 is motivated by the bound

$$(19) \quad \int_0^1 |f(\alpha)g_3(\nu\alpha)^6| d\alpha \ll P^{4+\varepsilon}$$

which we proved in Brüdern and Cook [13] for any fixed real $\nu \neq 0$. Thus (19) may be considered as a surrogate for (17) when $k = 3$, $s = 7$. It is the presence of two different exponential sums in this integral which causes extra complications of a combinatorial nature in the proofs of Theorem 3 in Section 6 and its refinement in Section 7.

3. The circle method. The methods for counting solutions of diophantine equations and inequalities are slightly different but based on a common principle, and large parts of the analysis are much the same. We remark that it suffices to count solutions of (1) or (2) in positive integers. This is obvious when k is even, and when k is odd this requires a preliminary transformation of some x_i into $-x_i$ (see Pitman [34]).

Let

$$(20) \quad \Lambda_j = \lambda_{1,j}\alpha_1 + \dots + \lambda_{R,j}\alpha_R \quad (1 \leq j \leq N)$$

and let \mathfrak{U} denote a cube in \mathbb{R}^R with sides parallel to the coordinate hyperplanes, and of volume 1. We define

$$(21) \quad G(\boldsymbol{\alpha}) = \prod_{j=1}^N g(\Lambda_j).$$

In the case of equations we have

$$(22) \quad \Theta_1 \geq \int_{\mathfrak{U}} G(\boldsymbol{\alpha}) d\boldsymbol{\alpha}$$

since the integral equals the number of solutions of (2), (3) subject to $x_i \in \mathcal{A}(P, P^\eta)$.

To count the solutions of the inequalities (1), (3) in a similar manner we choose, for a given $\xi > 1$, a kernel $K : \mathbb{R} \rightarrow [0, \infty)$ with $K(\alpha) = K(-\alpha)$, $K(\alpha) \ll \min(1, |\alpha|^{-\xi})$ such that its Fourier transform \widehat{K} satisfies

$$\widehat{K}(\alpha) = \begin{cases} 0 & (|\alpha| > \tau), \\ 1 & (|\alpha| \leq \frac{1}{3}\tau), \end{cases}$$

and $0 \leq \widehat{K}(\alpha) \leq 1$ for all real α . By Lemma 1 of Davenport [25], such a function exists. Writing

$$K(\boldsymbol{\alpha}) = K(\alpha_1) \dots K(\alpha_r),$$

we now have

$$(23) \quad \Theta_2 \geq \int_{\mathbb{R}^R} G(\boldsymbol{\alpha}) K(\boldsymbol{\alpha}) d\boldsymbol{\alpha}.$$

The object is to show that $\Theta_j \gg P^{N-Rk}$ when $N = n_0R + 1$ and $j = 1$ or 2; this implies the Proposition.

Before we go on to discuss these integrals in detail we reduce Θ_2 to a finite integral. Let $\mathfrak{B}(X)$ be the box $\max |\alpha_i| \leq X$, let $\zeta > 0$ be small, and put $\mathfrak{t} = \mathbb{R}^R \setminus \mathfrak{B}(P^\zeta)$. Then if we choose ξ sufficiently large in terms of $1/\zeta$, it is obvious that

$$(24) \quad \int_{\mathfrak{t}} |G(\boldsymbol{\alpha})| K(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \ll 1.$$

The next lemma is required for the treatment of the minor arcs in both problems.

LEMMA 1. *Let $\delta > 0$, and let \mathcal{E} be a set contained in a cube of type \mathfrak{U}*

such that for $\alpha \in \mathcal{E}$ there is a $j \leq N$ with $|g(\Lambda_j)| < P^{1-2\delta}$. Then

$$\int_{\mathcal{E}} |G(\alpha)| d\alpha \ll P^{N-Rk-\delta}.$$

Let \mathcal{F} be a set contained in a cube of type \mathfrak{A} such that for $\alpha \in \mathcal{F}$ there is a j with $|g(\Lambda_j)| < P(\log P)^{-\delta}$. Then

$$\int_{\mathcal{F}} |G(\alpha)| d\alpha \ll P^{N-Rk}(\log P)^{-\delta}.$$

In the proof of this lemma we require the following combinatorial result. An $r \times rn$ matrix over a field is called *partitionable* if its columns can be grouped into n disjoint blocks, with each block forming a non-singular $r \times r$ matrix.

LEMMA 2 (Aigner's criterion). *An $r \times nr$ matrix M is partitionable if and only if for $0 \leq d \leq r-1$ one has $\mu(d, M) \leq nd$.*

For a proof see Aigner [1] or Low, Pitman and Wolff [32].

We shall require the following results on modified Weyl sums.

LEMMA 3. *Let $\alpha \in \mathbb{R}$. There is a $\gamma = \gamma(k) > 0$ such that either $|g(\alpha)| < P^{1-\gamma}$, or there are coprime integers a, q , and*

$$g(\alpha) \ll q^{\varepsilon-1/(2k)} P \left(1 + P^k \left| \alpha - \frac{a}{q} \right| \right)^{-1/(2k)} (\log P)^3.$$

This is a weakened version of Theorem 1.8 of Vaughan [40] combined with Lemma 7.1 of Vaughan and Wooley [41]. The next lemma is a special case of Lemma 8.5 of Vaughan and Wooley [41].

LEMMA 4. *Let $A > 0$. Suppose that $(a, q) = 1$ and $q \leq (\log P)^A$, $|q\alpha - a| \leq (\log P)^A P^{-k}$. Then*

$$g(\alpha) \ll_{A,\varepsilon} P q^{\varepsilon-1/k} \left(1 + P^k \left| \alpha - \frac{a}{q} \right| \right)^{-1/k}.$$

Proof of Lemma 1. Let \mathcal{E}_j be the set of all $\alpha \in \mathcal{E}$ such that $|g(\Lambda_j)| < P^{1-2\delta}$. Since \mathcal{E} is the union of all \mathcal{E}_j it suffices to prove the lemma with \mathcal{E}_j in place of \mathcal{E} . By Lemma 2 and (7) the $R \times n_0 R$ submatrix of Λ where the j th column is omitted, is partitionable. Hence we can choose disjoint sets $\mathcal{L}(1), \dots, \mathcal{L}(n_0)$ from $\{1, 2, \dots, j-1, j+1, \dots, N\}$, of R elements each such that for all $t \leq n_0$ the matrices $A_{\mathcal{L}(t)}$ are non-singular; here $A_{\mathcal{L}(t)}$ denotes the $R \times R$ matrix formed with the l th columns of Λ , $l \in \mathcal{L}(t)$. Now, for $\alpha \in \mathcal{E}_j$ we have

$$(25) \quad G(\alpha) \ll (G_1 + \dots + G_{n_0}) \max_{\alpha} |g(\Lambda_j)|$$

where

$$G_t = \prod_{l \in \mathcal{L}(t)} |g(\Lambda_l)|^{n_0}.$$

Since $A_{\mathcal{L}(t)}$ is non-singular a change of variable and (17) yield

$$\int_{\mathfrak{U}} G_t d\alpha \ll \int_0^1 \dots \int_0^1 |g(\alpha_1) \dots g(\alpha_R)|^{n_0} d\alpha \ll P^{R(n_0-k+\varepsilon)}.$$

The first part of the lemma now follows from (25). Plainly, the second part of the lemma follows in the same way providing (16) holds.

To prove the second part if we only have (17), we may suppose that $n_0 > 4k$. We apply the first part of the lemma with $2k\delta < \gamma$ where γ is as in Lemma 3. It now suffices to estimate the contribution from the set $\mathcal{F} \setminus \mathcal{E} = \mathcal{H}$, say. By Lemma 3, we may assume that the second alternative of that lemma holds for all $g(\Lambda_j)$ whenever $\alpha \in \mathcal{H}$. In particular, $\Lambda_j \pmod{1} \in \mathcal{M}$ where \mathcal{M} is the union of all intervals $|q\alpha - a| \leq P^{\gamma-k}$ subject to $1 \leq a \leq q \leq P^\gamma$, $(a, q) = 1$. By Lemmata 3 and 4, when $n_0 > 4k$ we easily find that

$$\int_{\mathcal{M}} |g(\alpha)|^{n_0} d\alpha \ll P^{n_0-k}.$$

Hence a change of variable gives

$$\int_{\mathcal{H}} G_t d\alpha \ll \int_{\mathcal{M}} \dots \int_{\mathcal{M}} |g(\alpha_1) \dots g(\alpha_R)|^{n_0} d\alpha \ll P^{R(n_0-k)},$$

and the second part of the lemma also follows via (25).

4. The problem on equations. The endgame techniques differ considerably as to whether we deal with equations or inequalities. We now restrict our attention to equations, and suppose in this section that $\lambda_{ij} \in \mathbb{Z}$.

Let $0 < A \leq 1$ be a constant to be determined later, and put $L = (\log P)^A$. The major arcs \mathfrak{M} are defined as the union of all boxes

$$\mathfrak{M}(q, \mathbf{a}) = \{\alpha : |q\alpha_j - a_j| \leq LP^{-k} \quad (1 \leq j \leq R)\}$$

subject to $1 \leq a_j \leq q \leq L$ and $(q, a_1, \dots, a_R) = 1$. In (22) we choose \mathfrak{U} as the cube $LP^{-k} < |\alpha_j| \leq 1 + LP^{-k}$ and denote the complement of \mathfrak{M} in \mathfrak{U} by \mathfrak{m} .

Now choose $\delta > 0$ so small that $2kR\delta < A$. We proceed to show that

$$(26) \quad \int_{\mathfrak{m}} |G(\alpha)| d\alpha \ll P^{N-Rk} (\log P)^{-\delta}.$$

By Lemma 1 it suffices to show that for $\alpha \in \mathfrak{m}$ there is a j such that $|g(\Lambda_j)| < P(\log P)^{-\delta}$. If this were not the case we see from Lemmata 3 and 4 that for all $j \leq N$ there are coprime integers a_j, q_j such that $|q_j \Lambda_j - a_j| \leq$

$(\log P)^{2\delta k} P^{-k}$ and $q_j \leq P^{2\delta k}$. After relabelling indices we may suppose that $\Lambda_1, \dots, \Lambda_R$ are linearly independent forms. Now it follows easily that there is a $q \ll q_1 \dots q_R$ and a'_1, \dots, a'_R such that $|q\alpha_i - a'_i| \ll (\log P)^{2R\delta k}$. In particular, $\boldsymbol{\alpha} \in \mathfrak{M}$. This proves (26).

It remains to treat the major arcs. This is fairly straightforward. We begin by introducing further notation. Let $\varrho(u)$ denote Dickman's function which is defined by the relations

$$\begin{aligned} \varrho(u) &= 0 & (u < 0), \\ \varrho(u) &= 1 & (0 \leq u \leq 1), \\ u\varrho'(u) &= -\varrho(u-1) & (u > 1), \end{aligned}$$

and the request for a differentiable solution for $u > 1$ which is continuous at $u = 1$. Now we introduce

$$(27) \quad w(\beta) = \frac{1}{k} \int_0^{P^k} x^{1/k-1} \varrho\left(\frac{\log x}{k \log P}\right) e(\beta x) dx$$

and

$$(28) \quad S(q, a) = \sum_{x \leq q} e\left(\frac{ax^k}{q}\right).$$

In much the same way as in Brüdern [11], §13, it is shown that

$$(29) \quad g\left(\frac{a}{q} + \beta\right) = q^{-1} S(q, a) w(\beta) + O\left(\frac{P}{\log P} (q + P^k |\beta|)\right).$$

Note that a, q need not be coprime here.

Let $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$. Then $\Lambda_j = b_j/q + \beta_j$ where

$$b_j = \sum_{i=1}^R \lambda_{ij} a_i, \quad \beta_j = \sum_{i=1}^R \lambda_{ij} \left(\alpha_i - \frac{a_i}{q}\right).$$

By (29),

$$g(\Lambda_j) = q^{-1} S(q, b_j) w(\beta_j) + O(P(\log P)^{A-1}).$$

The measure of \mathfrak{M} is $\ll L^{R+1} P^{-Rk}$, and it follows that

$$\begin{aligned} \int_{\mathfrak{M}} G(\boldsymbol{\alpha}) d\boldsymbol{\alpha} &= \sum_{q \leq L} \sum_{\mathbf{a}(q)} q^{-N} S(q, b_1) \dots S(q, b_N) \int_{\mathfrak{M}(q, \mathbf{a})} w(\beta_1) \dots w(\beta_N) d\boldsymbol{\alpha} \\ &\quad + O(P^{N-Rk} (\log P)^{(R+2)A-1}). \end{aligned}$$

Here $\sum_{\mathbf{a}(q)}$ denotes an R -fold sum over a_1, \dots, a_R running over a system of residues modulo q with $(q, a_1, \dots, a_R) = 1$. The error term here is acceptable if we choose $A = 1/(2R + 4)$.

We now write $b_j/q = c_j/q_j$ in lowest terms. By Theorem 4.2 of Vaughan [39] we have $|S(q, b_j)| \ll q_j^{(k-1)/k}$. Hence

$$(30) \quad \sum_{\mathbf{a}(q)} q^{-N} |S(q, b_1) \dots S(q, b_N)| \ll \sum_{\mathbf{a}(q)} (q_1 \dots q_N)^{-1/k} \ll q^{-2-1/k+\varepsilon}$$

as we shall now show. Note that after relabelling the variables we may assume that for $j = 0, \dots, n_0 - 1$ the forms $\Lambda_{jR+1}, \dots, \Lambda_{(j+1)R}$ are linearly independent. By Hölder's inequality, dropping an excess q_N , it suffices to show that

$$\sum_{\mathbf{a}(q)} (q_{jR+1} \dots q_{(j+1)R})^{-n_0/k} \ll q^{-2-1/k+\varepsilon}.$$

We have $n_0/k \geq 2 + 1/k$, and the proposed inequality follows by an obvious modification of the argument used to prove formula (93) of Davenport and Lewis [29].

By partial integration,

$$w(\beta) \ll P(1 + P^k|\beta|)^{-1/k}.$$

Therefore we can use (30) and an obvious variant of Lemma 22 of Davenport and Lewis [29] to show that

$$(31) \quad \int_{\mathfrak{M}} G(\boldsymbol{\alpha}) d\boldsymbol{\alpha} = \mathfrak{S}I + O(P^{N-Rk}(\log P)^{-\delta})$$

where

$$\begin{aligned} \mathfrak{S} &= \sum_{q=1}^{\infty} \sum_{\mathbf{a}(q)} q^{-N} S(q, b_1) \dots S(q, b_N), \\ I &= \int_{\mathbb{R}^R} w(\Lambda_1) \dots w(\Lambda_N) d\boldsymbol{\alpha}. \end{aligned}$$

It is readily shown that $I \gg P^{N-Rk}$ (see §13 of Brüdern [11] for details in the case $k = 3, R = 2$, the methods extend to the present situation; compare also §12 of Davenport and Lewis [29]). Moreover, the singular series of course converges absolutely. In this case the singular series has been studied in detail by Davenport and Lewis [29]. Their Lemma 31 asserts that $\mathfrak{S} > 0$ provided that the corresponding system of equations has a non-singular p -adic solution for all primes p . Theorem 1 is now available from (31) and (26).

5. The problem on inequalities. In various aspects the problem on inequalities is much simpler. For example, there is only one “major arc” which we now define as $\mathfrak{M} = \mathfrak{B}(P^{\eta-k})$, and no singular series. For the minor arcs \mathfrak{m} , which in the present context are defined as the complement of \mathfrak{M} in $\mathfrak{B}(P^\zeta)$, we will obtain the crucial saving by a method of Cook [20].

For a given $\delta > 0$ suppose that $\alpha \in \mathfrak{m}$ and $|g(\Lambda_j)| > P^{1-\delta}$ holds for all j . If δ is sufficiently small we find from Lemma 3 that there are coprime integers a_j, q_j such that

$$(32) \quad q_j \ll P^{3k\delta}, \quad |q_j \Lambda_j - a_j| < P^{3k\delta-k}, \quad a_j \ll q_j P^\zeta.$$

We may assume that the linear form associated with the first $R+1$ variables has coefficients linearly independent over the rationals, and is of finite order. Now the method of Cook [20] is readily adopted to derive a contradiction to the conditions (32) providing δ is sufficiently small (in terms of the order), and P is restricted to a certain infinite set \mathcal{P} of integers. We assume $P \in \mathcal{P}$ for the rest of this section. For these P , and $\alpha \in \mathfrak{m}$ we have proved that $g(\Lambda_j) \ll P^{1-\delta}$ for at least one j .

The box $\mathfrak{B}(P^\zeta)$ can be covered by $O(P^{R\zeta})$ cubes of volume 1. Hence, by Lemma 1,

$$(33) \quad \int_{\mathfrak{m}} |G(\alpha)| K(\alpha) d\alpha \ll P^{N-Rk-\delta/2}$$

providing we choose $\zeta > 0$ such that $R\zeta < \delta/2$.

It remains to show that

$$(34) \quad \int_{\mathfrak{M}} |G(\alpha)| K(\alpha) d\alpha \gg P^{N-Rk},$$

the proposition and Theorem 2 then follow from (33) and (24).

By Lemma 7.1 of Vaughan and Wooley [41], for $\alpha \in \mathfrak{M}$,

$$(35) \quad g(\Lambda_j) \ll P(1 + P^k |\Lambda_j|)^{-1/(2k)} (\log P)^3.$$

The presence of the logarithmic factor is a nuisance, and we proceed to prune the major arc in two steps. Let $Q = (\log P)^{6N} P^{-k}$. The first step is to show that

$$(36) \quad \int_{\mathfrak{M} \setminus \mathfrak{B}(Q)} |G(\alpha)| K(\alpha) d\alpha \ll P^{N-Rk} (\log P)^{-1}.$$

We use the trivial estimate for $g(\Lambda_N)$. The remaining Rn_0 columns of Λ form a partitionable matrix. Hence, by Hölder's inequality, it suffices to show that

$$\int_{\mathfrak{M} \setminus \mathfrak{B}(Q)} |g(\Lambda_{j(1)}) \dots g(\Lambda_{j(R)})|^{n_0} d\alpha \ll P^{R(n_0-k)} (\log P)^{-1}$$

where $\Lambda_{j(1)}, \dots, \Lambda_{j(R)}$ are any R linearly independent forms. The mapping $\alpha \rightarrow \mathbf{A} = (\Lambda_{j(1)}, \dots, \Lambda_{j(R)})$ is a non-singular linear transformation, so that

by (35) the left hand side in the above inequality is

$$\ll (\log P)^{3Rn_0} P^{Rn_0} \int_M \prod_{t=1}^R (1 + P^k |\Lambda_{j(t)}|)^{-n_0/(2k)} d\mathbf{A}.$$

Here M is the image of $\mathfrak{M} \setminus \mathfrak{B}(Q)$. We now observe that for $\mathbf{A} \in M$ we have $\max(|\Lambda_{j(t)}|) \gg (\log P)^{6N} P^{-k}$, and recall that $n_0/(2k) > 1$. The required bound follows easily.

Now let $R = (\log P)^{1/2} P^{-k}$. We replace (35) by the estimate from Lemma 4 to see in the same way that

$$(37) \quad \int_{\mathfrak{B}(Q) \setminus \mathfrak{B}(R)} |G(\boldsymbol{\alpha})| K(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \ll P^{N-Rk} (\log P)^{-1}.$$

It is now straightforward, using (35) and standard arguments, to show that

$$(38) \quad \int_{\mathfrak{B}(R)} G(\boldsymbol{\alpha}) K(\boldsymbol{\alpha}) d\boldsymbol{\alpha} = I + O(P^{N-Rk} (\log P)^{-\delta})$$

where $\delta > 0$ and

$$I = \int_{\mathbb{R}^R} w(\Lambda_1) \dots w(\Lambda_N) K(\boldsymbol{\alpha}) d\boldsymbol{\alpha}.$$

Moreover, the usual method involving Fourier's inversion formula (see Pitman [34] for more details) immediately leads to the lower bound

$$I \gg P^{N-Rk}.$$

By (36)–(38) we see that (34) holds.

6. Cubic inequalities. We shall now describe the refinements in the method which are needed to deduce Theorem 3. Whenever we refer to earlier sections of this paper we are always concerned with the special case $k = 3$.

There is already a demand for a different generating function involving the exponential sum (18). It is convenient to redefine $G(\boldsymbol{\alpha})$ as

$$(39) \quad G(\boldsymbol{\alpha}) = f(\Lambda_1) \dots f(\Lambda_{R+1}) \prod_{t=R+2}^{7R+1} g(\Lambda_t)$$

where we may assume that the linear form associated with the first $R + 1$ variables has coefficients linearly independent over the rationals, and is of finite order.

With $G(\boldsymbol{\alpha})$ now defined as in (39) we still have (23) and (24). As before, the goal is to show that $\Theta_2 \gg P^{4R+1}$. We define the major arc as the box $\max |\alpha_i| \leq P^{-9/4}$. For $\boldsymbol{\alpha} \in \mathfrak{M}$ we then have the standard estimate

$$f(\Lambda_i) \ll P(1 + P^3 |\Lambda_i|)^{-1}.$$

This strong bound makes the pruning of the major arc much simpler than in the general case where no Weyl sums (18) were present. We may therefore leave it to the reader to prove the lower bound (34) in the present context. It might be worth mentioning that the use of Lemmata 3 and 4 can be avoided here if wished.

Defining again \mathfrak{m} as the complement of \mathfrak{M} in $\mathfrak{B}(P^\zeta)$ it remains to establish the estimate

$$(40) \quad \int_{\mathfrak{m}} |G(\boldsymbol{\alpha})|K(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \ll P^{4R+1-\delta}$$

for some $\delta > 0$.

The next lemma is a multi-dimensional version of (19).

LEMMA 5. *Let $A = (a_{ij})$ be an $r \times 7r$ matrix of real numbers such that $\mu(d, A) \leq 11d/2d$ for $d = 0, 1, \dots, r-1$. Suppose that for $0 \leq l \leq 6$ the matrices (a_{ij}) , $1 \leq i \leq r$, $lr < j \leq (l+1)r$ are non-singular. Let*

$$\gamma_j = \sum_{i=1}^r a_{ij}\alpha_i \quad (1 \leq j \leq 7r)$$

and suppose that \mathfrak{U} is a cube as in Lemma 1. Then

$$\int_{\mathfrak{U}} |f(\gamma_1) \dots f(\gamma_r)g(\gamma_{r+1}) \dots g(\gamma_{7r})| d\boldsymbol{\alpha} \ll P^{4r+\varepsilon}.$$

The implied constant depends on a_{ij} and ε only.

PROOF. Let $\delta = \varepsilon^2$. For any subset $\mathcal{J} \subset \{1, \dots, r\}$ let

$$U(\mathcal{J}) = \{\boldsymbol{\alpha} \in \mathfrak{U} : |f(\gamma_j)| \geq P^{3/4+\delta} \ (j \in \mathcal{J}), \\ |f(\gamma_l)| < P^{3/4+\delta} \ (l \leq r, l \notin \mathcal{J})\}.$$

We estimate the contribution from any of these sets separately. Here we recall the bounds

$$(41) \quad \int_0^1 |g(\beta)|^6 d\beta \ll P^{13/4+\delta},$$

$$(42) \quad \int_0^1 |g(\beta)|^8 d\beta \ll P^{5+\delta}$$

((41) is Theorem 4.4 of Vaughan [40] when η is sufficiently small, and (42) follows from Hua's Lemma, on considering the underlying diophantine equation). We also need the well known fact that

$$(43) \quad \int_{\mathcal{F}} |f(\beta)|^4 d\beta \ll P^{1+\delta}$$

where \mathcal{F} is the set of all $\beta \in [0, 1]$ where $|f(\beta)| \geq P^{3/4+\delta}$ (see e.g. Brüderin [9]).

The contributions from $U(\emptyset)$ and $U(\{1, \dots, r\})$ are easy to estimate. We have

$$|g(\gamma_{r+1}) \dots g(\gamma_{7r})| \leq \sum_{l=1}^6 |g(\gamma_{lr+1}) \dots g(\gamma_{(l+1)r})|^6$$

so that a simple change of variable and (41) give

$$\int_{\mathfrak{A}} |g(\gamma_{r+1}) \dots g(\gamma_{7r})| d\alpha \ll P^{(13/4+\delta)r}.$$

Hence the contribution from $U(\emptyset)$ is $\ll P^{(3/4+\delta)r} P^{(13/4+\delta)r} \ll P^{4r+\varepsilon}$ providing ε is sufficiently small. For $U(\{1, \dots, r\})$ we use Hölder's inequality, (42) and (43) to find in much the same way that

$$\begin{aligned} & \int_{U(\{1, \dots, r\})} |f(\gamma_1) \dots f(\gamma_r) g(\gamma_{r+1}) \dots g(\gamma_{7r})| d\alpha \\ & \ll \left(\int_{\mathcal{F}^r} |f(\gamma_1) \dots f(\gamma_r)|^4 d\gamma_1 \dots d\gamma_r \right)^{1/4} \left(\int_{\mathfrak{A}} |g(\gamma_{r+1}) \dots g(\gamma_{7r})|^{4/3} d\alpha \right)^{3/4} \\ & \ll (P^{r(1+\delta)})^{1/4} (P^{r(5+\delta)})^{3/4} \ll P^{4+\varepsilon} \end{aligned}$$

as required.

The other sets can be treated by a hybrid version of the two arguments just given. The problems which now arise are mainly of a notational character. Therefore we first observe that for symmetry reasons it suffices to estimate the contribution from the sets $U(\{1, \dots, t\}) = U_t$, say, where $1 \leq t \leq r-1$.

Let \mathbf{a}_j denote the j th column of A . We split the indices $r+1, r+2, \dots, 7r$ into two disjoint sets \mathcal{V}, \mathcal{W} such that $|\mathcal{V}| = 6t$, $|\mathcal{W}| = 6r - 6t$. Then we write

$$\begin{aligned} (44) \quad & |f(\gamma_1) \dots f(\gamma_t) g(\gamma_{r+1}) \dots g(\gamma_{7r})| \\ & = (|f(\gamma_1)|^{1/6} \dots |f(\gamma_t)|^{1/6})^6 \left(\prod_{v \in \mathcal{V}} |g(\gamma_v)|^{1/3} \right)^3 \left(\prod_{w \in \mathcal{W}} |g(\gamma_w)|^{1/4} \right)^4. \end{aligned}$$

The right hand side is a product of $6t + 18t + 4(6r - 6t) = 24r$ factors, and in this product $24r$ linear forms γ_i occur to which there corresponds an $r \times 24r$ matrix

$$(45) \quad \mathbf{A}_t = (\mathbf{a}_j \mathbf{a}_j \mathbf{a}_j \mathbf{a}_j \mathbf{a}_j \mathbf{a}_j \mathbf{a}_v \mathbf{a}_v \mathbf{a}_v \mathbf{a}_v \mathbf{a}_w \mathbf{a}_w \mathbf{a}_w \mathbf{a}_w)_{1 \leq j \leq t, v \in \mathcal{V}, w \in \mathcal{W}}.$$

This matrix is partitionable. To see this let $1 \leq d \leq r-1$ and suppose that there are exactly K columns in \mathbf{A}_t lying in a subspace of dimension d . Since \mathbf{A}_t consists of repetitions of columns from A , there is a corresponding set of κ columns from A in the same subspace. We can write $\kappa = \kappa_1 + \kappa_2 + \kappa_3$

where κ_1 is the number of these columns with index $j \leq t$, where κ_2 is the number of columns with index in \mathcal{V} , and κ_3 is the number of columns with index in \mathcal{W} . By (45),

$$K = 6\kappa_1 + 3\kappa_2 + 4\kappa_3 \leq 2\kappa_1 + 4\kappa.$$

Note that $\kappa \leq \mu(d, A) \leq 11d/2$. The forms $\gamma_1, \dots, \gamma_t$ are linearly independent so that $\kappa_1 \leq \min(d, t)$. It follows that $K \leq 24d$. By Aigner's criterion, \mathbf{A}_t is indeed partitionable.

Now we can group the $24r$ factors in (44) into 24 blocks of r factors each such that to any block there corresponds a non-singular $r \times r$ matrix. Then, by Hölder's inequality,

$$\int_{U(t)} |f(\gamma_1) \dots f(\gamma_t) g(\gamma_{r+1}) \dots g(\gamma_{7r})| d\boldsymbol{\alpha} \ll (I_1 \dots I_{24})^{1/24}$$

where any factor I_j is of one of the forms

$$\int_{U(t)} |f_{\nu(1)} \dots f_{\nu(\sigma)}|^4 |g(\gamma_{\nu(\sigma+1)}) \dots g(\gamma_{\nu(\varrho)})|^8 |g(\gamma_{\nu(\varrho+1)}) \dots g(\gamma_{\nu(r)})|^6 d\boldsymbol{\alpha},$$

and where $1 \leq \nu(l) \leq t$ ($1 \leq l \leq \sigma$), $\nu(l) \in \mathcal{V}$ ($\sigma + 1 \leq l \leq \varrho$), $\nu(l) \in \mathcal{W}$ ($\varrho + 1 \leq l \leq r$). After a change of variable in any of these integrals we have

$$\begin{aligned} (I_1 \dots I_{24})^{1/24} &\ll \left(\prod_{1 \leq l \leq t} \int_{\mathcal{F}} |f(\gamma_l)|^4 d\gamma_l \right)^{1/4} \\ &\quad \times \left(\prod_{v \in \mathcal{V}} \int_0^1 |g(\gamma_v)|^8 d\gamma_v \right)^{1/8} \left(\prod_{w \in \mathcal{W}} \int_0^1 |g(\gamma_w)|^6 d\gamma_w \right)^{1/6}. \end{aligned}$$

Therefore, by (41)–(43) and the definition of U_t it follows that the contribution from this set to the integral in question is

$$\ll P^{(r-t)(3/4+\delta)} P^{t(1/4+\delta)} P^{6t(5+\delta)/8} P^{6(r-t)(13/4+\delta)/6} \ll P^{4r+2r\delta}.$$

This proves the lemma.

For a given $\delta > 0$ let \mathcal{E}_t be the set of all $\boldsymbol{\alpha} \in \mathfrak{m}$ where $|f(\Lambda_t)| \leq P^{1-2\delta}$. Then, by Lemma 5, since $\mathfrak{B}(P^\zeta)$ is covered by $O(P^{r\zeta})$ copies of \mathfrak{A} ,

$$\int_{\mathcal{E}_t} |G(\boldsymbol{\alpha})| K(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \ll P^{4r+r\zeta+\varepsilon-2\delta} \ll P^{4r-\delta}$$

providing ζ and ε are small compared with δ . However, by Weyl's inequality (Lemma 2.4 of Vaughan [39]), $|f(\Lambda_t)| \geq P^{1+2\delta}$ implies that there are integers a, q such that $q \ll P^{9\delta}$, $|q\Lambda_t - a| \ll P^{9\delta-3}$ providing δ is small. This can be used to show, as in Section 5, that for P in a certain infinite set of integers, the union of all \mathcal{E}_t covers \mathfrak{m} . This proves (40).

7. Relaxing hypotheses. The rank condition in Theorem 3 can certainly be relaxed. As was pointed out in the introduction, the results and proofs of Theorems 1 and 2 suggest that it should be possible to deal with systems satisfying $\mu(d) \leq 7d$. Actually, a careful inspection of Section 6 shows that our argument applies to such systems apart from the treatment of the sets U_t when $1 \leq t \leq r - 1$, in the proof of Lemma 5. However, in principle it should be possible to choose the sets \mathcal{V} , \mathcal{W} so that the relevant matrix \mathbf{A}_t remains partitionable even on the weaker assumption that $\mu(d) \leq 7d$ although the combinatorial problems associated with this might be awkward. In view of our recent work on three additive equations in 22 variables [2] it might be interesting to give the details of such a refinement when $r = 3$. The key is the following improvement of Lemma 5.

LEMMA 6. *Suppose that $r = 3$. Suppose that the matrix $(a_{ij})_{1 \leq i \leq r, 1 \leq j \leq 7r}$ partitions into $(a_{ij})_{1 \leq i \leq r, r+l < j \leq (l+1)r}$ for $0 \leq l \leq 6$. Then the claim of Lemma 5 holds.*

Proof. Inspecting the proof of Lemma 5, we see that we only have to show that for $t = 1$ and $t = 2$ there exist appropriate choices of $\mathcal{V} = \mathcal{V}_t$ and $\mathcal{W} = \mathcal{W}_t$ such that the matrix \mathbf{A}_t defined in (45) is partitionable.

Consider the matrix $(\mathbf{a}_l)_{4 \leq l \leq 21}$. This matrix is partitionable. Define $\nu(d)$ as the maximal number of columns \mathbf{a}_l , $4 \leq l \leq 21$, generating a linear space of dimension d . By Aigner's criterion, $\nu(d) \leq 6d$.

Let $V_t = \text{lin}(\mathbf{a}_1, \dots, \mathbf{a}_t)$. This space has dimension t . Let \mathcal{G} denote the set of all indices l with $4 \leq l \leq 21$ and $\mathbf{a}_l \in V_t$, and put $\gamma = |\mathcal{G}|$. Since $\text{lin}(\mathbf{a}_l : l \in \mathcal{G}) \subset V_t$ we have $\gamma \leq \nu(t) \leq 6t$. Hence we can choose \mathcal{V}_t such that $\mathcal{G} \subset \mathcal{V}_t$. We now show that in many cases this choice of \mathcal{V} is already appropriate.

Let $d = 1$ or $d = 2$, and let X be a linear space of dimension d , generated by $D = D(d, X)$ columns from \mathbf{A}_t . Then there must be 6ϱ such columns \mathbf{a}_j , $1 \leq j \leq t$, and 3σ such columns \mathbf{a}_v , $v \in \mathcal{V}$, and also 4τ such columns \mathbf{a}_w , $w \in \mathcal{W}$. Then

$$(46) \quad D = 6\varrho + 3\sigma + 4\tau.$$

By Aigner's criterion, in order to prove that \mathbf{A}_t is partitionable we have to show that $D \leq 24d$.

Note that the $\sigma + \tau$ columns $\mathbf{a}_v, \mathbf{a}_w$ just considered span a space of dimension $\leq d$. Hence, by Aigner's criterion, $\sigma + \tau \leq \nu(d) \leq 6d$. Hence, if $\varrho = 0$, by (46) we have $D \leq 4\nu(d) \leq 24d$ as required.

Thus we may now suppose that $\varrho \geq 1$. On the other hand, it is clear that $\varrho \leq \min(t, d)$. Now, when $d \leq t$ and $\varrho = d$ then by construction $X \subset V_t$, and this implies $\tau = 0$ by choice of \mathcal{V} . Now $\sigma \leq 6d$, and by (46) it follows that $D \leq 6\varrho + 18d \leq 24d$, which is acceptable.

Now we may suppose that $\varrho \leq d - 1$ if $d \leq t$, and that $\varrho \leq t$ if $d > t$. In the present context this implies $\varrho = 1, d = 2$.

We consider first the situation $t = 1$. We have to estimate $D(2, X)$ where X is a 2-dimensional space with $\varrho = 1$, that is, $\mathbf{a}_1 \in X$. Furthermore, for any such X we have $\mathbf{a}_j \in X$ for $j \in \mathcal{G}$. Let X_1, \dots, X_s be all the distinct 2-spaces spanned by \mathbf{a}_1 and another \mathbf{a}_l . This splits the indices $j \in \mathcal{H} = \{4, 5, \dots, 21\} \setminus \mathcal{G}$ into s disjoint sets $\mathcal{X}_1, \dots, \mathcal{X}_s$ such that $l \in \mathcal{X}_\lambda$ if and only if $\mathbf{a}_l \in X_\lambda$. We put $h_\lambda = |\mathcal{X}_\lambda|$, and may suppose $h_1 \geq h_2 \geq \dots \geq h_s > 0$. Moreover,

$$(47) \quad h_1 + h_2 + \dots + h_s = 18 - \gamma.$$

Note that X_λ contains \mathbf{a}_l for $l \in \mathcal{G} \cup \mathcal{X}_\lambda$. By Aigner's criterion,

$$(48) \quad h_\lambda + \gamma \leq \nu(2) \leq 12.$$

The trivial bound for D is now

$$D(2, X_\lambda) \leq 6 + 3\gamma + 4h_\lambda \leq 42 + h_\lambda$$

(compare (46), (48)). Hence, whenever $h_\lambda \leq 6$ we have $D(2, X_\lambda) \leq 48$ as required, irrespective of a more specific choice of \mathcal{V} .

Hence we may now suppose that $h_1 \geq \dots \geq h_L \geq 7 > h_{L+1}$. By (47) we have $7L \leq 18 - \gamma$. In particular, $L \leq 2$. We wish to put $h_\lambda - 6$ indices from \mathcal{X}_λ into \mathcal{V} , for $1 \leq \lambda \leq L$. This is always possible since we then have specified

$$\gamma + \sum_{\lambda \leq L} (h_\lambda - 6) \leq 6$$

elements from \mathcal{V} so far; here we used (48) when $L = 1$ and (47) when $L = 2$. But now, by construction, at most six columns $\mathbf{a}_l \in X_\lambda$ have $l \in \mathcal{W}$. Therefore, using (48) once again, we have

$$D(2, X_\lambda) \leq 6 + 3(\gamma + h_\lambda - 6) + 4 \times 6 \leq 48.$$

Hence, when $t = 1$ there is a choice for \mathcal{V} such that \mathbf{A}_1 is partitionable.

Now consider the case $t = 2$. We imitate the procedure used when $t = 1$ but the situation is more complicated. The 2-spaces we still have to consider are those which contain either \mathbf{a}_1 or \mathbf{a}_2 but not both. Thus let $X_{1,1}, \dots, X_{1,s_1}$ be the distinct 2-spaces containing \mathbf{a}_1 but not \mathbf{a}_2 . Similarly define $X_{2,1}, \dots, X_{2,s_2}$ as the distinct 2-spaces containing \mathbf{a}_2 but not \mathbf{a}_1 . As before, when $u = 1$ or 2 , the $X_{u,1}, \dots, X_{u,s_u}$ give rise to a splitting of \mathcal{H} into disjoint sets $\mathcal{X}_{u,\lambda}$, $1 \leq \lambda \leq s_u$, such that $\mathbf{a}_l \in X_{u,\lambda}$ if and only if $l \in \mathcal{X}_{u,\lambda}$. We put $h_{u,\lambda} = |\mathcal{X}_{u,\lambda}|$ and have, as in (47),

$$(49) \quad h_{u,1} + \dots + h_{u,s_u} = 18 - \gamma.$$

Let ϑ_u be the number of columns \mathbf{a}_l with $l \in \mathcal{G}$ which lie on the line generated

by \mathbf{a}_u . Then $\vartheta_u \leq \nu(1) \leq 6$ and $\vartheta_1 + \vartheta_2 \leq \gamma$. As in (48) we have

$$(50) \quad h_{u,\lambda} + \vartheta_u \leq \nu(2) \leq 12.$$

The simplest bound now is

$$D(2, X_{u,\lambda}) \leq 6 + 3\vartheta_u + 4h_{u,\lambda} \leq 42 + h_{u,\lambda}.$$

Hence, as in the case $t = 1$ we already obtain the required bound $D(2, X_{u,\lambda}) \leq 48$ for all u, λ with $h_{u,\lambda} \leq 6$.

We may now assume that $h_{u,1} \geq \dots \geq h_{u,L_u} \geq 7$, and all other $h_{u,l} \leq 6$. As before we see that $L_u \leq 2$. Without loss of generality we may therefore suppose that

$$(51) \quad 0 \leq L_2 \leq L_1 \leq 2, \quad L_1 \geq 1.$$

When $h_{u,\lambda} \geq 7$ we wish to put $h_{u,\lambda} - 6$ indices from $\mathcal{X}_{u,\lambda}$ into \mathcal{V} . If this were possible for all $u = 1, 2$, $1 \leq \lambda \leq L_u$, then we would be left with at most 6 indices from $\mathcal{X}_{u,\lambda}$ to be put into \mathcal{W} so that by (50)

$$D(2, X_{u,\lambda}) \leq 6 + 3(\vartheta_u + h_{u,\lambda} - 6) + 4 \times 6 \leq 48.$$

It remains to show that we can actually put that many indices into \mathcal{V} . If we choose $h_{u,\lambda} - 6$ indices from $\mathcal{X}_{u,\lambda}$ at random for putting these into \mathcal{V} we have determined at most

$$(52) \quad \gamma + \sum_{u=1}^2 \sum_{\lambda=1}^{L_u} (h_{u,\lambda} - 6) = \gamma + \sum_{u,\lambda} h_{u,\lambda} - 6(L_1 + L_2)$$

elements from \mathcal{V} , and it remains to show that (52) does not exceed 12. As a crude argument we first use (49) with $u = 1$, and then either (49) or (50) with $u = 2$, and deduce that (52) is

$$\leq 18 + \min(18, 12L_2) - 6(L_1 + L_2).$$

For L_1, L_2 subject to (51) this is ≤ 12 unless $L_1 = L_2 = 1$. In this remaining case we refine the argument as follows. We observe that (52) now becomes

$$(53) \quad \gamma + h_{1,1} + h_{2,1} - 12.$$

If $h_{1,1} + h_{2,1} \leq 18 - \gamma$ then (53) is ≤ 6 , which is more than required. If $h_{1,1} + h_{2,1} > 18 - \gamma$ then recall that $\mathcal{X}_{u,1}$ is a subset of the set \mathcal{H} with $18 - \gamma$ elements. Hence $\mathcal{X}_{1,1} \cap \mathcal{X}_{2,1}$ contains at least $h_{1,1} + h_{2,1} + \gamma - 18$ elements. In other words, there are that many columns $\mathbf{a}_l \in X_{1,1} \cap X_{2,1}$. This intersection is a line, whence

$$h_{1,1} + h_{2,1} + \gamma - 18 \leq \nu(1) \leq 6.$$

Therefore (53) does not exceed 12 as required. This completes the proof of the lemma.

The reader should have no difficulty to reduce the rank condition in Theorem 3 to $\mu(d) \leq 6d$ when $r = 3$, using Lemma 6 alone. However, for a further reduction yet another argument is required.

LEMMA 7. *Suppose $A = (\mathbf{a}_l)_{1 \leq l \leq 22}$ is a 3×22 matrix satisfying $\mu(d) \leq 7d$ for $d = 0, 1, 2, 3$. Then the columns can be renumbered so that $(\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3)$ is non-singular, and the maximal number $\mu^*(d)$ of columns \mathbf{a}_l , $4 \leq l \leq 22$, generating a linear space of dimension d satisfies $\mu^*(d) \leq 6d$ ($d = 1, 2$).*

PROOF. If $\mu(d) \leq 6d$ for $d = 1, 2$ then the lemma is trivial. Hence we may suppose that either $\mu(1) = 7$ or $13 \leq \mu(2) \leq 14$.

First consider the case $\mu(2) \geq 13$. Suppose that there are s different 2-spaces X_1, \dots, X_s , say, which contain at least 13 columns \mathbf{a}_j . Suppose further that there are t different lines, Y_1, \dots, Y_t , which contain 7 columns \mathbf{a}_j . We note that no \mathbf{a}_j can lie in two of the Y_j , which implies $7t \leq 22$, that is, $0 \leq t \leq 3$. Furthermore, it is clear that either $Y_j \subset X_l$, or $Y_j \cap X_l = \{0\}$, and that at most two lines Y_j can lie in the same X_l .

Let $s = 1$ first. If X_1 contains two Y_l then we take a column from each of these two lines, and denote them by $\mathbf{a}_1, \mathbf{a}_2$. If X_1 contains one Y_l , we take a column lying in this line and denote it by \mathbf{a}_1 . Then we take a column lying in X_1 which is linearly independent of \mathbf{a}_1 and denote it by \mathbf{a}_2 . If X_1 contains no Y_l we take two linearly independent columns in X_1 and denote these by $\mathbf{a}_1, \mathbf{a}_2$. Now since X_1 contains at least 13 columns there can be at most one Y_l which does not lie in X_1 . If such a Y_j exists we take a column lying in it and denote it by \mathbf{a}_3 . If not we take an arbitrary column not in X_1 and denote it by \mathbf{a}_3 . In any case, $\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \neq 0$ and $\mu^*(d) \leq 6d$ by construction.

Suppose $s \geq 2$. Now $X_{ij} = X_i \cap X_j$ is a line and therefore contains at most 7 columns. It follows that at most 3 columns do not lie in one of the X_j . In particular, if there are any Y_l then these must lie in some X_j .

Now consider the case $s = 2$. The number of columns in both X_1, X_2 is at least 13 so that X_{12} contains at least 4 columns. Therefore the number of columns in X_j but not in X_{12} does not exceed 10. It follows that if X_j contains two lines Y_l then one of these must be X_{12} . Therefore we first choose a column in X_{12} and denote it by \mathbf{a}_3 . Then we take a column $\mathbf{a}_j \in X_j$ which is not in X_{12} . If X_j contains a line $Y_l \neq X_{12}$ then we arrange that \mathbf{a}_j is from this line. It is clear that $\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \neq 0$. We also have $\mu^*(d) \leq 6d$ since we have used one column from any exceptional line Y_l and two columns from any exceptional 2-space X_j for the choice of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

Now let $s \geq 3$. We first observe that no three X_j can intersect in a line. To see this suppose that X_1, X_2, X_3 have one line, L say, in common. Let λ be the number of columns in L and H_j be the number of columns in X_j

but not in L . Then

$$\lambda + h_1 + h_2 + h_3 \leq 22$$

and

$$13 \leq \lambda + h_j \leq 22.$$

The last inequality implies $3\lambda + h_1 + h_2 + h_3 \geq 39$. Combining this with the first inequality we deduce that $2\lambda \geq 17$. On the other hand, it is clear that $\lambda \leq \mu(1) \leq 7$, contrary to the previous inequality.

Thus we may suppose that all the intersection lines X_{ij} are distinct. We define h_{ij} to be the number of columns \mathbf{a}_l in X_{ij} , and let h_l be the number of columns in X_l which do not lie in any X_{ij} . Then

$$(54) \quad \sum_{i=1}^s h_i + \sum_{1 \leq i < j \leq s} h_{ij} \leq 22,$$

and, for any fixed l ,

$$(55) \quad 14 \geq h_l + \sum_{\substack{1 \leq i \leq s \\ i \neq l}} h_{il} \geq 13$$

since the sum here is the number of columns in X_l . We sum this over l and obtain

$$\sum_{i=1}^s h_i + 2 \sum_{1 \leq i < j \leq s} h_{ij} \geq 13s.$$

From (54) it follows that

$$(56) \quad \sum_{1 \leq i < j \leq s} h_{ij} \geq 13s - 22.$$

On the other hand, by (54) again, the left hand side here is ≤ 22 , which gives $s \leq 3$. Thus we may suppose that $s = 3$. In this case, by (56) and (54), $h_1 + h_2 + h_3 \leq 5$. In particular, $h_j \leq 5$. Hence, if there are any exceptional lines Y_l these must coincide with some X_{ij} . Now $h_{ij} \leq \mu(1) \leq 7$, and from (55) and $h_j \leq 5$ we deduce that $h_{ij} > 0$ for any pair $1 \leq i < j \leq 3$. Thus we may take one column from each X_{ij} and denote these by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. It is clear that these are linearly independent, and that $\mu^*(d) \leq 6d$.

It remains to consider the case where $\mu(2) \leq 12$ but $\mu(1) = 7$. If there were two exceptional lines Y_1, Y_2 , these would span a 2-space containing 14 columns, which is impossible. Hence we have exactly one exceptional line Y_1 . We take a column in Y_1 and denote it by \mathbf{a}_1 . Then we take two further columns such that $\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \neq 0$. This completes the proof of the lemma.

Now it is time to reconsider the situation in Theorem 3 when $R = 3$. We suppose first that the matrix of coefficients satisfies $\mu(1) \leq 7$ and $\mu(2) \leq 14$.

We rearrange columns according to Lemma 7 and define, using the notation introduced in Section 6,

$$G(\boldsymbol{\alpha}) = f(\Lambda_1)f(\Lambda_2)f(\Lambda_3) \prod_{l=4}^{22} g(\Lambda_l).$$

The only real difficulty in adopting the method from Section 6 is in the treatment of the set \mathfrak{m} as defined there. Let $5\zeta < \delta < 10^{-5}$ and consider the sets

$$\mathcal{G}_t = \{\boldsymbol{\alpha} \in \mathfrak{m} : |g(\Lambda_t)| < P^{1-\delta}\}.$$

Then

$$\int_{\mathcal{G}_t} |G(\boldsymbol{\alpha})|K(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \ll P^{1-\delta} \int_{\mathfrak{B}(P^\zeta)} |f(\Lambda_1)f(\Lambda_2)f(\Lambda_3)| \prod_{\substack{l=4 \\ l \neq t}}^{22} |g(\Lambda_l)| d\boldsymbol{\alpha}.$$

Now Lemma 6 can be used to estimate the integral on the right by $\ll P^{12+\varepsilon+3\zeta}$, which leads to an acceptable bound for the contribution from \mathcal{G}_t . If P is restricted to a certain infinite set of integers then the minor arc \mathfrak{m} is covered by the union of all \mathcal{G}_t , $4 \leq t \leq 22$. This can be shown as on earlier occasions in this paper. For the major arcs we may refer to Section 6. Thus we have proved the following result which we announce in a notation slightly digressing from earlier usage.

THEOREM 4. *Consider the inequalities*

$$\begin{aligned} |a_1x_1^3 + \dots + a_{22}x_{22}^3| &< \varepsilon, \\ |b_1x_1^3 + \dots + b_{22}x_{22}^3| &< \varepsilon, \\ |c_1x_1^3 + \dots + c_{22}x_{22}^3| &< \varepsilon, \end{aligned}$$

with real coefficients such that at least one of the quaternary linear forms associated with it has coefficients linearly independent over the rationals, and is of finite order. Suppose that the matrix of coefficients satisfies $\mu(1) \leq 7$ and $\mu(2) \leq 14$. Then the inequalities have a simultaneous non-trivial integer solution.

This should be compared with the corresponding result on equations in Atkinson, Brüdern and Cook [2].

Obviously the rank condition can be relaxed further but at present it is unlikely that it can be removed completely. If $\mu(1) \geq 8$ then after taking linear combinations we may suppose that $b_l = c_l = 0$ for $1 \leq l \leq 8$. We put $x_9 = \dots = x_{22} = 0$ and solve the single inequality

$$|a_1x_1^3 + \dots + a_8x_8^3| < \varepsilon.$$

This can be done using the results of Davenport [26] and Davenport and Roth [30], irrespective of the coefficients a_i (for more recent results on this

particular problem see Baker [6] and Brüdern [10]).

If $\mu(2) \geq 16$ then by a similar argument we can reduce the system to a pair of inequalities

$$(57) \quad \begin{aligned} |a_1x_1^3 + \dots + a_{16}x_{16}^3| &< \varepsilon, \\ |b_1x_1^3 + \dots + b_{16}x_{16}^3| &< \varepsilon. \end{aligned}$$

Here a combination of the methods of Pitman [34] and Vaughan [38] can be used to show this pair to be non-trivially soluble.

There remains the case $\mu(2) = 15$, $\mu(1) \leq 7$. Then, after taking linear combinations we can arrange that the system becomes

$$(58) \quad \begin{aligned} |a_1x_1^3 + \dots + a_{16}x_{16}^3 + \dots + a_{22}x_{22}^3| &< \varepsilon, \\ |b_1x_1^3 + \dots + b_{16}x_{16}^3 + \dots + b_{22}x_{22}^3| &< \varepsilon, \\ |c_{16}x_{16}^3 + \dots + c_{22}x_{22}^3| &< \varepsilon, \end{aligned}$$

where $c_{16} \dots c_{22} \neq 0$. The natural method now would be to put $x_{16} = \dots = x_{22} = 0$. Then one is in the situation of (57) but with 15 variables in place of 16. We have recently found [13] that such a pair has solutions providing one of the *ternary* linear forms associated with it has coefficients which are linearly independent over the rationals, and is of finite order. However, this is not always true. For example, all $a_i, b_i, 1 \leq i \leq 15$, may be integers. Then the question reduces to finding a simultaneous non-trivial *zero* of a pair of additive cubic forms in 15 variables with integer coefficients. There are examples due to Davenport and Lewis [28] where such a zero does not exist.

The above method can be refined. Suppose there is a solution to

$$(59) \quad |c_{16}x_{16}^3 + \dots + c_{22}x_{22}^3| < \varepsilon,$$

for any $\varepsilon > 0$. Then an idea from Davenport and Lewis [28] can be adopted to the present situation to reduce (58) to a pair of inequalities in 16 rather than 15 variables which can be solved by (57). Although it is conjectured that (59) always has a non-trivial integer solution this is known at present only in the case where the cubic form is a multiple of a form with integer coefficients only, by recent work of Baker [5]. Thus there remains an exceptional case which denies treatment. It transpires that the key is the long awaited proof that (59) is soluble.

References

- [1] M. Aigner, *Combinatorial Theory*, Springer, Berlin 1979.
- [2] O. D. Atkinson, J. Brüdern and R. J. Cook, *Three additive cubic equations*, Acta Arith. 60 (1991), 29–83.

- [3] O. D. Atkinson, J. Brüdern and R. J. Cook, *Simultaneous additive congruences to a large prime modulus*, *Mathematika*, to appear.
- [4] R. C. Baker, *Diagonal cubic equations, I*, in: *Théorie des Nombres, Number Theory*, J.-M. De Koninck and C. Levesque (eds.), de Gruyter, Berlin 1989, 15–28.
- [5] —, *Diagonal cubic equations, II*, *Acta Arith.* 53 (1989), 217–250.
- [6] —, *Diagonal cubic equations, III*, *Proc. London Math. Soc.* (3) 58 (1989), 495–518.
- [7] R. C. Baker and J. Brüdern, *On pairs of additive cubic equations*, *J. Reine Angew. Math.* 391 (1988), 157–180.
- [8] B. J. Birch, *Homogeneous forms of odd degree in a large number of variables*, *Mathematika* 4 (1957), 102–105.
- [9] J. Brüdern, *Additive diophantine inequalities with mixed powers I*, *ibid.* 34 (1987), 124–130.
- [10] —, *Cubic diophantine inequalities*, *ibid.* 35 (1988), 51–58.
- [11] —, *On pairs of diagonal cubic forms*, *Proc. London Math. Soc.* (3) 61 (1990), 273–343.
- [12] —, *On Waring’s problem for fifth powers and some related topics*, *ibid.* 457–479.
- [13] J. Brüdern and R. J. Cook, *On pairs of cubic diophantine inequalities*, *Mathematika* 38 (1991), 250–263.
- [14] —, —, *Systems of cubic equations*, in preparation.
- [15] J. W. S. Cassels and M. J. T. Guy, *On the Hasse principle for cubic surfaces*, *Mathematika* 13 (1966), 111–120.
- [16] R. J. Cook, *Simultaneous quadratic equations*, *J. London Math. Soc.* (2) 4 (1971), 319–326.
- [17] —, *A note on a lemma of Hua*, *Quart. J. Math. Oxford* 23 (1972), 287–288.
- [18] —, *Pairs of additive equations*, *Michigan Math. J.* 19 (1972), 325–331.
- [19] —, *Simultaneous quadratic equations II*, *Acta Arith.* 25 (1973), 1–5.
- [20] —, *Simultaneous quadratic inequalities*, *ibid.* 25 (1974), 337–346.
- [21] —, *Pairs of additive equations II. Large odd degree*, *J. Number Theory* 17 (1983), 80–92.
- [22] —, *Pairs of additive equations III: quintic equations*, *Proc. Edinburgh Math. Soc.* (2) 26 (1983), 191–211.
- [23] —, *Pairs of additive equations IV. Sextic equations*, *Acta Arith.* 43 (1984), 227–243.
- [24] —, *Pairs of additive congruences: quintic congruences II*, in: *Computers in Mathematical Research*, N. M. Stephens and M. P. Thorne (eds.), Clarendon Press, Oxford 1988, 93–117.
- [25] H. Davenport, *Indefinite quadratic forms in many variables*, *Mathematika* 3 (1956), 81–101.
- [26] —, *Cubic forms in thirty-two variables*, *Philos. Trans. Roy. Soc. London Ser. A* 251 (1959), 193–232.
- [27] H. Davenport and D. J. Lewis, *Homogeneous additive equations*, *Proc. Roy. Soc. London Ser. A* 274 (1963), 443–460.
- [28] —, —, *Cubic equations of additive type*, *Philos. Trans. Roy. Soc. London Ser. A* 261 (1966), 97–136.
- [29] —, —, *Simultaneous equations of additive type*, *ibid.* 264 (1969), 557–595.
- [30] H. Davenport and K. F. Roth, *The solubility of certain Diophantine inequalities*, *Mathematika* 2 (1955), 81–96.
- [31] E. Dörner, *Simultaneous diagonal equations over certain p -adic fields*, *J. Number Theory* 36 (1990), 1–11.
- [32] L. Low, J. Pitman and A. Wolff, *Simultaneous diagonal congruences*, *ibid.* 29 (1988), 31–59.

- [33] T. Nadesalingam, *Diophantine inequalities in many variables*, Ph.D. thesis, Univ. of Adelaide, South Australia, 1981.
- [34] J. Pitman, *Pairs of diagonal inequalities*, in: Recent Progress in Analytic Number Theory, Vol. 2, H. Halberstam and C. Hooley (eds.), Academic Press, London 1981, 183–215.
- [35] W. M. Schmidt, *Linear forms with algebraic coefficients. I*, J. Number Theory 3 (1971), 253–277.
- [36] —, *Diophantine inequalities for forms of odd degree*, Adv. in Math. 38 (1980), 128–151.
- [37] —, *The solubility of certain p -adic equations*, J. Number Theory 19 (1984), 63–80.
- [38] R. C. Vaughan, *On pairs of additive cubic equations*, Proc. London Math. Soc. (3) 34 (1977), 354–364.
- [39] —, *The Hardy–Littlewood Method*, Cambridge University Press, 1981.
- [40] —, *A new iterative method in Waring’s problem*, Acta Math. 162 (1989), 1–71.
- [41] R. C. Vaughan and T. Wooley, *On Waring’s problem: some refinements*, Proc. London Math. Soc. (3) 63 (1991), 35–68.

MATHEMATISCHES INSTITUT
BUNSENSTRASSE 3–5
3400 GÖTTINGEN, GERMANY

DEPARTMENT OF PURE MATHEMATICS
HICKS BUILDING
HOUNSFIELD ROAD
SHEFFIELD S3 7RH, ENGLAND

Received on 4.10.1991

(2175)