SMALL ZEROS OF ADDITIVES FORMS IN MANY VARIABLES. II(¹)

BY

WOLFGANG M. SCHMIDT

University of Colorado
Boulder, Colorado
U.S.A.

1. Introduction

It is a well known consequence of the Hardy-Littlewood Circle Method that a diophantine equation

$$a_1 x_1^k + \ldots + a_s x_s^k = 0 \tag{1.1}$$

has a nontrivial solution in nonnegative integers $x_1, \ldots, x_s$, provided only that $s \geq c_1(k)$ and that the coefficients $a_1, \ldots, a_s$ are not all of the same sign. In the first paper [4] under the present title, the author proved that if $\varepsilon > 0$, and if at least $c_2(k, \varepsilon)$ of the coefficients are positive and at least $c_2(k, \varepsilon)$ are negative, then the equation has a nontrivial solution in nonnegative integers with

$$|x_i| \leq A^{(1/k)+\varepsilon} \quad (i = 1, \ldots, s) \tag{1.2}$$

where

$$A = \max(1, |a_1|, \ldots, |a_s|). \tag{1.3}$$

In the equation $b_1(x_1^k + \ldots + x_t^k) - b_2(x_{t+1}^k + \ldots + x_{2t}^k) = 0$ where b_1, b_2 are coprime and positive, every nontrivial solution in nonnegative $x_1, \ldots, x_{2t}$ has some $x_i \geq (B/t)^{1/k}$ where $B = \max(b_1, b_2)$. This shows that the exponent in (1.2) is essentially best possible.

In particular, it follows that if k is odd, if $s \geq 2c_2(k, \varepsilon)$ and if $a_1, \ldots, a_s$ have arbitrary signs, then there is a nontrivial solution of (1.1) in integers $x_1, \ldots, x_s$ (not necessarily nonnegative) with (1.2). This latter result had also been shown by Birch [1]. But much more is true. We will show that if k is odd and if $s \geq c_3(k, \varepsilon)$ where $\varepsilon > 0$, then (1.1) has a nontrivial solution in integers $x_1, \ldots, x_s$ with

$$|x_i| \leq A^{\varepsilon} \quad (i = 1, \ldots, s). \tag{1.4}$$

(¹) Written with partial support from NSF grant NSF-MCS 78-01770.

It is well known (see the remark in [1]) that this result has applications to diophantine inequalities involving forms of odd degree with real coefficients; more about these applications will be said in subsequent work.

The example given above shows that a similar result cannot be true if k is even. The trouble is that the values of x^k cannot be negative in this case. To help such k overcome their handicap, we replace powers x^k by σx^k where σ may be 1 or -1. We then have the

THEOREM. *Suppose k, s are natural numbers with $s \geq c_4(k, \varepsilon)$ where $\varepsilon > 0$. Then given integers $a_1, \ldots, a_s$, the equation*

$$\sigma_1 a_1 x_1^k + \ldots + \sigma_s a_s x_s^k = 0 \tag{1.5}$$

has a solution in numbers $\sigma_1, \ldots, \sigma_s$, $x_1, \ldots, x_s$, where each σ_i is 1 or -1, and where the x_i are integers, not all zero, with (1.4).

Our proof employs the Circle Method but is no straightforward application of this method. It is similar to the proof in the first paper [4]. We will again use a result of Pitman [3], but with the expection of two lemmas the present paper is independent of [4]. Our method allows in principle to compute explicit values for $c_4(k, \varepsilon)$, but the values so obtained would be extremely large.

2. Preliminaries

We are dealing with additive forms

$$\mathcal{A} = \mathcal{A}(\mathbf{x}) = \mathcal{A}(x_1, \ldots, x_s) = a_1 x_1^k + \ldots + a_s x_s^k$$

with integer coefficients in vectors $\mathbf{x} = (x_1, \ldots, x_s)$. If $\mathcal{A}$ is not identically zero, put

$$\mathcal{A}' = (a_1/d) x_1^k + \ldots + (a_s/d) x_s^k,$$

where $d > 0$ is the greatest common divisor of $a_1, \ldots, a_s$, and if $\mathcal{A}$ is identically zero, put $\mathcal{A}' = \mathcal{A}$. Put

$$|\mathcal{A}| = \max(1, |a_1|, \ldots, |a_s|),$$

and denote the number of variables of $\mathcal{A}$ by $s(\mathcal{A})$.

When k is odd set $X = \mathbf{Z}$, the ring of integers. When k is even, let X be the set of products $u\zeta$ where $u \in \mathbf{Z}$ and where ζ is a $(2k)$-th root of unity. In either case we see that $x^k = |x|^k$ or $x^k = -|x|^k$ for each $x \in X$, and both possibilities actually do occur. X is closed under multiplication. Let X^s consist of vectors $\mathbf{x} = (x_1, \ldots, x_s)$ with components in X; for such $\mathbf{x}$ set

$$|\mathbf{x}| = \max(|x_1|, \ldots, |x_s|).$$

For $\mathbf{x} \in X^s$, $\mathcal{A}(\mathbf{x})$ is always a rational integer. We say that $\mathcal{A}$ *represents* an integer z if there is a nonzero $\mathbf{x} \in X^s$ with $\mathcal{A}(\mathbf{x}) = z$. We write $\mathcal{A} \to z$ in this case, and we put

$$\psi(\mathcal{A}|z) = \min |\mathbf{x}|,$$

where the minimum is taken over nonzero $\mathbf{x} \in X^s$ with $\mathcal{A}(\mathbf{x}) = z$. It is clear that $\mathcal{A} \to 0$ is equivalent to $\mathcal{A}' \to 0$ and that

$$\psi(\mathcal{A}|0) = \psi(\mathcal{A}'|0). \tag{2.1}$$

Our theorem may now be formulated as follows.

If $\mathcal{A}$ is a form with $s(\mathcal{A}) \geqslant c_4(k, \varepsilon)$, then

$$\psi(\mathcal{A}|0) \leqslant |\mathcal{A}|^{\varepsilon}. \tag{2.2}$$

Put $\mathbf{x} \wedge \mathbf{u}$ if $x_i u_i = 0$ for $i = 1, \ldots, s$. We say that $\mathcal{A}$ *represents* a form $\mathcal{B} = \mathcal{B}(y_1, \ldots, y_t)$ if there are $\mathbf{x}_1, \ldots, \mathbf{x}_t$ in X^s with $\mathbf{x}_i \neq \mathbf{0}$ $(1 \leqslant i \leqslant s)$ and $\mathbf{x}_i \wedge \mathbf{x}_j$ $(1 \leqslant i < j \leqslant s)$ such that

$$\mathcal{B}(y_1, \ldots, y_t) = \mathcal{A}(y_1 \mathbf{x}_1 + \ldots + y_t \mathbf{x}_t). \tag{2.3}$$

This equation means that

$$\mathcal{B}(y_1, \ldots, y_t) = b_1 y_1^k + \ldots + b_t y_t^k \tag{2.4}$$

where $b_i = \mathcal{A}(\mathbf{x}_i)$ $(i = 1, \ldots, t)$. Whenever $\mathcal{A} \to \mathcal{B}$ put

$$\psi(\mathcal{A}|\mathcal{B}) = \min (\max (|\mathbf{x}_1|, \ldots, |\mathbf{x}_t|)),$$

where the minimum is over t-tuples $\mathbf{x}_1, \ldots, \mathbf{x}_t$ as described above which have (2.3). If $\mathcal{A} \to \mathcal{B}$ and $\mathcal{B} \to z$ then $\mathcal{A} \to z$, and in fact

$$\psi(\mathcal{A}|z) \leqslant \psi(\mathcal{A}|\mathcal{B}) \psi(\mathcal{B}|z). \tag{2.5}$$

3. Reductions

In all that follows, k will be fixed and we will not explicitly express the dependency of constants or of sets on k. Let Λ be the set of numbers $\mu > 0$ such that there is a $c_5 = c_5(\mu)$ with the property that every form $\mathcal{A}$ with $s(\mathcal{A}) \geqslant c_5$ has

$$\psi(\mathcal{A}|0) \leqslant |\mathcal{A}|^{\mu}. \tag{3.1}$$

By the work of Pitman [3], Λ is not empty. Let λ be the greatest lower bound of Λ. By [1] or [4], $\lambda \leqslant 1/k$. Our goal here will be to show that

$$\lambda = 0. \tag{3.2}$$

We will suppose that $\lambda > 0$ and we will reach a contradiction.

The polynomial $g(\varrho)=\lambda+k\lambda^2-k\lambda\varrho-k^2\lambda^2\varrho-\varrho$ has $g(\lambda)=-k^2\lambda^3<0$. Hence we can pick ϱ with

$$0<\varrho<\lambda \tag{3.3}$$

and $g(\varrho)<0$, i.e. with

$$\lambda+k\lambda^2-k\lambda\varrho-k^2\lambda^2\varrho<\varrho. \tag{3.4}$$

Pick $\nu>0$ so small that

$$\text{(i)}\quad \varrho+8\lambda\nu<\lambda, \tag{3.5}$$

$$\text{(ii)}\quad \nu<1/5,$$

$$\text{(iii)}\quad \nu<\varrho/10.$$

Finally pick μ with

$$\max(\varrho+8\lambda\nu,\ \lambda-\tfrac{1}{2}\lambda\nu)<\mu<\lambda. \tag{3.6}$$

We will show that $\mu\in\Lambda$, and this will be the desired contradiction. We will show that (3.1) holds whenever $s(\mathcal{A})$ is large. We clearly may suppose that no coefficient of $\mathcal{A}$ is zero.

Suppose we can show that (3.1) holds whenever both $|\mathcal{A}|$ and $s(\mathcal{A})$ are large. A short reflection shows that (3.1) is true when $|\mathcal{A}|$ is under a fixed bound and when $s(\mathcal{A})$ is large. Hence it then follows that (3.1) is true if just $s(\mathcal{A})$ is very large. Thus it will suffice to show the validity of (3.1) when both $|\mathcal{A}|$ and $s(\mathcal{A})$ are large.

Pick τ with

$$\max(\varrho+8\lambda\nu,\ \lambda-\tfrac{1}{2}\lambda\nu)<\tau<\mu \tag{3.7}$$

and choose $\delta>0$ so small that

$$(1+\delta)\tau+(2\delta/k)<\mu. \tag{3.8}$$

Divide the interval $0\leqslant x\leqslant 1$ into a finite number of subintervals I of length not exceeding δ. If s is large, one of these subintervals will be such that many of the coefficients a_i will have $|a_i|=|\mathcal{A}|^{\alpha_i}$ with $\alpha_i\in I$. We may suppose that the first coefficients $a_1, \ldots, a_t$ have $|a_i/a_j|\leqslant|\mathcal{A}|^{\delta}$ $(1\leqslant i, j\leqslant t)$ where t is large. Put $A^*=|\mathcal{A}|^{\delta}\max(|a_1|, \ldots, |a_t|)$. Let $p_1, \ldots, p_t$ be the largest integers with

$$|a_i|p_i^k\leqslant A^*.$$

Now $A^*/|a_i|\geqslant|\mathcal{A}|^{\delta}$ $(i=1, \ldots, t)$, and if $|\mathcal{A}|$ is large (which we may suppose), then $p_i\geqslant 2^{-1/k}(A^*/|a_i|)^{1/k}$, so that

$$\tfrac{1}{2}A^*\leqslant|a_ip_i^k|\leqslant A^*\quad (i=1, \ldots, t). \tag{3.9}$$

We have $\mathcal{A}\to a_1p_1^ky_1^k+\ldots+a_tp_t^ky_t^k=\mathcal{B}$, say, with

$$\psi(\mathcal{A}\,|\,\mathcal{B})\leqslant\max(p_1, \ldots, p_t)\leqslant|\mathcal{A}|^{2\delta/k}\quad\text{and}\quad|\mathcal{B}|\leqslant A^*\leqslant|\mathcal{A}|^{1+\delta}.$$

If we can show that

$$\psi(\mathcal{B}|0) \leqslant |\mathcal{B}|^{\tau},$$

then

$$\psi(\mathcal{A}|0) \leqslant \psi(\mathcal{A}|\mathcal{B})\psi(\mathcal{B}|0) \leqslant |\mathcal{A}|^{(2\delta/k)+(1+\delta)\tau} \leqslant |\mathcal{A}|^{\mu}$$

by (3.8), which is what we want.

What is special about $\mathcal{B}$ is that by (3.9) each of its coefficients has absolute value at least equal to $\frac{1}{2}|\mathcal{B}|$. Hence it will suffice to show that *if* $\mathcal{A} = a_1 x_1^k + \ldots + a_s x_s^k$ *is a form such that*

$$\tfrac{1}{2}|\mathcal{A}| \leqslant |a_i| \leqslant |\mathcal{A}| \quad (i = 1, \ldots, s), \tag{3.10}$$

and if $s = s(\mathcal{A}) \geqslant c_6$, *then*

$$\psi(\mathcal{A}|0) \leqslant |\mathcal{A}|^{\tau}. \tag{3.11}$$

Of course c_6 depends on k and τ, but since $k, \lambda, \varrho, \nu, \mu, \tau$ will be fixed, we will not indicate the dependency of c_6 (and of subsequent constants) on these parameters.

Proposition. *If* $s(\mathcal{A}) \geqslant c_7$ *and if* (3.10) *holds, then either* (3.11) *is true or there is a* z *with*

$$\mathcal{A} \to z, \quad |z| \leqslant |\mathcal{A}|^{4\nu} \quad \textit{and} \quad \psi(\mathcal{A}|z) \leqslant |\mathcal{A}|^{\varrho}. \tag{3.12}$$

This proposition appears to be too weak, but in fact is all that we need. For note that $2\lambda > \lambda$ and that $c_5(2\lambda)$ is defined; in fact we may suppose it to be an integer, and similarly we may take c_7 to be an integer. Now if $s(\mathcal{A}) \geqslant c_7 c_5(2\lambda)$, then we may write

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}_1(\mathbf{x}_1) + \ldots + \mathcal{A}_t(\mathbf{x}_t)$$

where $t = c_5(2\lambda)$ and where $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_t)$ and each $\mathbf{x}_i$ has c_7 coordinates, so that $s(\mathcal{A}_i) = c_7$ $(i = 1, \ldots, t)$. If some $\mathcal{A}_i$ has $\psi(\mathcal{A}_i|0) \leqslant |\mathcal{A}_i|^{\tau} \leqslant |\mathcal{A}|^{\tau}$, then we are done. Otherwise, the proposition tells us that $\mathcal{A}_i \to z_i$ $(i = 1, \ldots, t)$ with (3.12) for each i. Thus $\mathcal{A} \to z_1 y_1^k + \ldots + z_t y_t^k = \mathcal{B}$, say, where

$$|\mathcal{B}| \leqslant |\mathcal{A}|^{4\nu}, \quad \psi(\mathcal{A}|\mathcal{B}) \leqslant |\mathcal{A}|^{\varrho} \quad \text{and} \quad s(\mathcal{B}) = t = c_5(2\lambda).$$

It follows that $\psi(\mathcal{B}|0) \leqslant |\mathcal{B}|^{2\lambda}$, whence we get

$$\psi(\mathcal{A}|0) \leqslant \psi(\mathcal{A}|\mathcal{B})\psi(\mathcal{B}|0) \leqslant |\mathcal{A}|^{\varrho}|\mathcal{B}|^{2\lambda} \leqslant |\mathcal{A}|^{\varrho + 8\lambda\nu} \leqslant |\mathcal{A}|^{\tau}$$

by (3.7).

We will now proceed to prove the proposition.

4. The Circle Method

We may suppose without loss of generality that s is even and that half of the coefficients of $\mathcal{A}$ are positive and half are negative. For a given form $\mathcal{A}$ we put

$$A = |\mathcal{A}|; \tag{4.1}$$

then (3.10) may be rewritten as

$$\tfrac{1}{2}A \leqslant |a_i| \leqslant A \quad (i = 1, \ldots, s). \tag{4.2}$$

Let N, H be the integer parts of A^{ϱ}, $A^{4\nu}$, respectively. Then

$$\tfrac{1}{2}A^{\varrho} < N \leqslant A^{\varrho}, \quad \tfrac{1}{2}A^{4\nu} < H \leqslant A^{4\nu} \tag{4.3}$$

if $A = |\mathcal{A}|$ is sufficiently large. The proposition will certainly be true for $\mathcal{A}$ if we can solve the equation

$$a_1 x_1^k + \ldots + a_s x_s^k - z = 0 \tag{4.4}$$

in integers $x_1, \ldots, x_s$, z subject to

$$1 \leqslant x_i \leqslant N \quad (i = 1, \ldots, s) \quad \text{and} \quad 1 \leqslant z \leqslant H. \tag{4.5}$$

The number Z of such solutions is given by

$$Z = \int_0^1 f(\alpha)\, d\alpha \tag{4.6}$$

where

$$f(\alpha) = \sum_{x_1=1}^{N} \cdots \sum_{x_s=1}^{N} \sum_{z=1}^{H} e(\alpha(a_1 x_1^k + \ldots + a_s x_s^k - z)) \tag{4.7}$$

and where $e(x) = e^{2\pi i x}$. We are finished if we can show that $Z > 0$.

We define the *Major Arcs* to be the intervals modulo 1 of the type

$$\mathfrak{M}_{qu}: \left|\alpha - \frac{u}{q}\right| < A^{-1+\nu} N^{-k}, \tag{4.8}$$

where

$$1 \leqslant q < A^{\nu} \quad \text{and} \quad \text{g.c.d. } (q, u) = 1. \tag{4.9}$$

These arcs do not overlap, at least when A is large, since their centers have mutual distances $\geqslant A^{-2\nu} > 2A^{-1+\nu}$ by (3.5 ii). The complement of the major arcs constitutes the *Minor Arcs*.

For later reference we state the following

LEMMA 1. *Suppose that* $\eta > 0$, *that* $N \geqslant c_8(\eta) = c_8(k, \eta)$ *and that* $C \geqslant N^{1-(1/K)+\eta}$ *where* $K = 2^{k-1}$. *If* α *is such that*

$$\left| \sum_{x=1}^{N} e(\alpha x^k) \right| \geqslant C,$$

then there is a natural

$$q \leqslant (N/C)^K N^\eta \quad \textit{with} \quad \|\alpha q\| \leqslant (N/C)^K N^{\eta - k},$$

where $\|\cdot\|$ *denotes the distance to the nearest integer.*

Proof. This is the corollary to Lemma 1 of [4]. It is an easy consequence of the "Weyl Inequality".

5. The Minor Arcs

LEMMA 2. *Suppose* $s \geqslant c_9$, *and suppose* α *lies in a Minor Arc. Then either*

$$|f(\alpha)| < HN^{s-k}A^{-2}, \tag{5.1}$$

or $\psi(\mathcal{A}|0) \leqslant A^\tau$, i.e. (3.11) *holds.*

Proof. We may suppose that $0 \leqslant \alpha \leqslant 1$. Choose η with

$$0 < \eta < c_{10}, \tag{5.2}$$

where c_{10} is a constant (depending on $k, \lambda, \varrho, \nu, \mu, \tau$) to be determined later. The quantity $c_5(\lambda + \eta)$ is well defined and may be taken to be an integer. Set

$$n = c_5(\lambda + \eta), \quad h = n^2. \tag{5.3}$$

Choose c_9 so large that $s \geqslant c_9$ implies

$$\left(k + \frac{4}{\varrho}\right) \Big/ (s - h + 1) < \eta.$$

Since by (4.3), $A < N^{2/\varrho}$ if A is large, we have

$$(N^k A^2)^{1/(s-h+1)} < N^{(k+(4/\varrho))/(s-h+1)} < N^\eta. \tag{5.4}$$

Now if (5.1) fails to hold, then the sums

$$S_i(\alpha) = \sum_{x=1}^{N} e(\alpha a_i x^k) \quad (i = 1, \dots, s) \tag{5.5}$$

satisfy

$$|S_1(\alpha) \dots S_s(\alpha)| \geqslant N^{s-k}A^{-2}. \tag{5.6}$$

If, say, $|S_1(\alpha)| \geqslant \ldots \geqslant |S_s(\alpha)|$, then the left hand side of (5.6) is bounded by $|S_h(\alpha)|^{s-h+1}N^{h-1}$, and $|S_h(\alpha)|$ and therefore $|S_i(\alpha)|$ for $i=1, \ldots, h$ satisfy

$$\begin{aligned}|S_i(\alpha)| &\geqslant N^{(s-k-h+1)/(s-h+1)}A^{-2/(s-h+1)}\\ &= N(N^kA^2)^{-1/(s-h+1)} > N^{1-\eta}\end{aligned}$$

by (5.4). The hypotheses of Lemma 1 are satisfied by $C = N^{1-\eta}$, since $N^{1-\eta} > N^{1-(1/K)+\eta}$ by (5.2), if c_{10} is small enough. Lemma 1 yields the existence of natural numbers $q_1, \ldots, q_h$ with

$$q_i \leqslant N^{2K\eta} \quad \text{and} \quad \|\alpha a_i q_i\| \leqslant N^{-k+2K\eta} \quad (i=1, \ldots, h). \tag{5.7}$$

It follows that

$$\|\alpha a_i q_i^k\| \leqslant N^{-k+2kK\eta} \quad (i=1, \ldots, h).$$

There are integers $u_1, \ldots, u_h$ with

$$|\alpha a_i q_i^k - u_i| \leqslant N^{-k+2kK\eta} \quad (i=1, \ldots, h). \tag{5.8}$$

We obtain

$$\begin{aligned}|a_i q_i^k u_j - a_j q_j^k u_i| &\leqslant |(\alpha a_j q_j^k - u_j) a_i q_i^k| + |(\alpha a_i q_i^k - u_i) a_j q_j^k|\\ &\leqslant 2N^{-k+2kK\eta}AN^{2kK\eta} \quad (1 \leqslant i, j \leqslant h).\end{aligned}$$

Thus the integer vectors

$$\mathbf{a}_i = (a_i q_i^k, u_i) \quad (i=1, \ldots, h) \tag{5.9}$$

satisfy

$$|\det(\mathbf{a}_i, \mathbf{a}_j)| \leqslant 2AN^{-k+4kK\eta} \quad (1 \leqslant i, j \leqslant h). \tag{5.10}$$

Write $\mathbf{a}_1 = r\mathbf{b}$ where $\mathbf{b}$ is primitive, i.e. a vector with coprime integer components; say

$$\mathbf{b} = (q, u) \quad \text{with} \quad q > 0 \quad \text{and} \quad \text{g.c.d.}\,(q, u) = 1. \tag{5.11}$$

Now (5.8) yields $|u_1| \leqslant 2|a_1|q_1^k$, so that $|u| \leqslant 2q$ and $|\mathbf{b}| \leqslant 2q$, which in turn yields

$$|r| = |\mathbf{a}_1|/|\mathbf{b}| \geqslant A/(2|\mathbf{b}|) \geqslant A/(4q). \tag{5.12}$$

Choose $\mathbf{c}$ such that $\mathbf{b}, \mathbf{c}$ becomes a basis for the integer vectors. Then $|\det(\mathbf{b}, \mathbf{c})| = 1$ and each $\mathbf{a}_i$ may be written as

$$\mathbf{a}_i = v_i\mathbf{b} + w_i\mathbf{c} \quad (i=1, \ldots, h)$$

with integers v_i, w_i. In view of (5.10) and (5.12) we have

$$\begin{aligned}|w_i| = |\det(\mathbf{a}_i, \mathbf{b})| &= |r|^{-1}|\det(\mathbf{a}_i, \mathbf{a}_1)|\\ &\leqslant |r|^{-1}\cdot 2AN^{-k+4kK\eta}\\ &\leqslant 8qN^{-k+4kK\eta} = M, \quad (i=1, \ldots, h)\end{aligned} \tag{5.13}$$

say.

6. The Minor Arcs, continued

We now distinguish two cases (I) and (II).

(I) $M \geq 1$. This is the fun case. Recall from (5.3) that $h = n^2$. We now replace the indices $i = 1, \ldots, h$ by double indices j, l where $1 \leq j, l \leq n$. So, for example, $a_1, \ldots, a_h$ are now written as $a_{11}, \ldots, a_{1n}, \ldots, a_{n1}, \ldots, a_{nn}$. Introduce the forms

$$\mathcal{A}_j = \mathcal{A}_j(x_{j1}, \ldots, x_{jn}) = w_{j1} x_{j1}^k + \ldots + w_{jn} x_{jn}^k \quad (j = 1, \ldots, n).$$

We have $|\mathcal{A}_j| \leq M$ by (5.13) and since $M \geq 1$. Further since $n = c_5(\lambda + \eta)$ by (5.3), we have

$$\psi(\mathcal{A}_j | 0) \leq |\mathcal{A}_j|^{\lambda+\eta} \leq M^{\lambda+\eta} \quad (j = 1, \ldots, n). \tag{6.1}$$

Choose nonzero vectors $\mathbf{x}_j = (x_{j1}, \ldots, x_{jn}) \in X^n$ with $A_j(\mathbf{x}_j) = 0$ and $|\mathbf{x}_j| = \psi(\mathcal{A}_j | 0)$ $(j = 1, \ldots, h)$. Then the two dimensional vectors

$$\mathfrak{b}_j = x_{j1}^k \mathfrak{a}_{j1} + \ldots + x_{jn}^k \mathfrak{a}_{jn} \quad (j = 1, \ldots, n)$$

are integer multiples of $\mathfrak{b}$, and hence the first coordinate b_j of each $\mathfrak{b}_j$ is divisible by q. We observe that

$$b_j = a_{j1} q_{j1}^k x_{j1}^k + \ldots + a_{jn} q_{jn}^k x_{jn}^k \quad (j = 1, \ldots, n), \tag{6.2}$$

whence it follows that $\mathcal{A} \to \mathcal{B}$ where

$$\mathcal{B} = b_1 y_1^k + \ldots + b_n y_n^k.$$

We note that

$$\psi(\mathcal{A} | \mathcal{B}) \leq \max_{1 \leq j, l \leq n} |q_{jl} x_{jl}| \leq N^{2K\eta} M^{\lambda+\eta} \tag{6.3}$$

by (5.7), (6.1) and our choice of the $\mathbf{x}_j$. In view of (6.2) it is clear that

$$|\mathcal{B}| \leq nA(N^{2K\eta} M^{\lambda+\eta})^k = nAN^{2kK\eta} M^{k\lambda+k\eta}. \tag{6.4}$$

Observe again that $n = c_5(\lambda + \eta)$, so that $\mathcal{B} \to 0$ and

$$\psi(\mathcal{B} | 0) = \psi(\mathcal{B}' | 0) \leq |\mathcal{B}'|^{\lambda+\eta} \leq (\max(1, |\mathcal{B}|/q))^{\lambda+\eta}. \tag{6.5}$$

This is true if $\mathcal{B} = \mathcal{B}' = 0$ and $|\mathcal{B}'| = 1$, and also if $\mathcal{B}' \neq 0$, since each coefficient of $\mathcal{B}$ is divisible by q and therefore $|\mathcal{B}'| \leq |\mathcal{B}|/q$ in this case. Combining (6.3) and (6.5) we obtain

$$\psi(\mathcal{A} | 0) \leq N^{2K\eta} (\max(M, M|\mathcal{B}|/q))^{\lambda+\eta}. \tag{6.6}$$

Now q, being a divisor of $a_1 q_1^k$, satisfies

$$q \leq AN^{2kK\eta} \tag{6.7}$$

by (5.7). Thus from (5.13),

$$M \leqslant 8AN^{-k+6kK\eta}. \tag{6.8}$$

Since by (6.7), q does not exceed the right hand side of (6.4), we have

$$\max (M, M|\mathcal{B}|/q) \leqslant MnAN^{2kK\eta}M^{k\lambda+k\eta}/q,$$

and by (5.13) this is

$$\leqslant 8N^{-k+4kK\eta}nAN^{2kK\eta}M^{k\lambda+k\eta}$$
$$= 8nAN^{-k+6kK\eta}M^{k\lambda+k\eta}.$$

Observing (6.8) we obtain

$$\max (M, M|\mathcal{B}|/q) < 8nAN^{-k+6kK\eta}8^{k\lambda+k\eta}A^{k\lambda+k\eta}N^{-k^2\lambda+6kK\eta(k\lambda+k\eta)}$$
$$< A^{1+k\lambda+k\eta}N^{-k-k^2\lambda+7kK\eta(1+2k\lambda)}$$

if A is large and if $\eta<\lambda$. But $\eta<\lambda$ can be made true by choosing the constant c_{10} in (5.2) sufficiently small. If we substitute this into (6.6) we get

$$\psi(\mathcal{A}|0) < A^{\lambda+k\lambda^2}N^{-k\lambda-k^2\lambda^2}A^{c_{11}\eta}$$

with a certain constant c_{11} independent of η. In view of (4.3) we have

$$\psi(\mathcal{A}|0) < |\mathcal{A}|^{\lambda+k\lambda^2-k\lambda\varrho-k^2\lambda^2\varrho+2c_{11}\eta}. \tag{6.9}$$

Now if the constant c_{10} in (5.2) is sufficiently small, the exponent in (6.9) is less than ϱ by (3.4), hence is less than τ by (3.7). So we get $\psi(\mathcal{A}|0) \leqslant |\mathcal{A}|^{\tau}$, i.e. the desired (3.11).

(II) $M<1$. This case resembles the situation in [4]. We revert to the original notation with indices $i=1, ..., h$. We have $w_i=0$ by (5.13), and hence each vector $\mathbf{a}_i$ $(i=1, ..., h)$ is a multiple of $\mathbf{b}$. Therefore q divides each $a_iq_i^k$ $(i=1, ..., h)$. We have $\mathcal{A}\to\mathcal{B}$ where

$$\mathcal{B} = a_1q_1^ky_1^k + \dots + a_hq_h^ky_h^k,$$

and

$$\psi(\mathcal{A}|\mathcal{B}) \leqslant N^{2K\eta}, \quad |\mathcal{B}| \leqslant AN^{2kK\eta} \tag{6.10}$$

by (5.7). We have $s(\mathcal{B})=h=n^2\geqslant n=c_5(\lambda+\eta)$ by (5.3), and

$$\psi(\mathcal{B}|0) = \psi(\mathcal{B}'|0) \leqslant |\mathcal{B}'|^{\lambda+\eta} \leqslant (|\mathcal{B}|/q)^{\lambda+\eta},$$

since each coefficient of $\mathcal{B}$ is divisible by q. Thus from (6.10) and (4.3),

$$\psi(\mathcal{A}|0) \leqslant \psi(\mathcal{A}|\mathcal{B})\psi(\mathcal{B}|0) \leqslant N^{2K\eta}(|\mathcal{B}|/q)^{\lambda+\eta}$$
$$\leqslant N^{2K\eta}(AN^{2kK\eta})^{\lambda+\eta}q^{-\lambda} \leqslant A^{\lambda+\eta}N^{2K\eta(1+4k\lambda)}q^{-\lambda}$$
$$\leqslant A^{\lambda+\eta+2K\varrho\eta(1+4k\lambda)}q^{-\lambda} \leqslant A^{\lambda+(\nu\lambda/2)}q^{-\lambda}$$

if η is sufficiently small by (5.2). Now if $q \geqslant A^{\nu}$, then

$$\psi(\mathcal{A}|0) \leqslant |\mathcal{A}|^{\lambda-(\nu\lambda/2)} \leqslant |\mathcal{A}|^{\tau}$$

by (3.7). We may thus suppose that $q < A^{\nu}$, so that (4.9) holds. (5.8) yields

$$\left|\alpha - \frac{u}{q}\right| = \left|\alpha - \frac{u_1}{a_1 q_1^k}\right| \leqslant 2A^{-1}|\alpha a_1 q_1^k - u_1|$$
$$\leqslant 2A^{-1}N^{-k+2kK\eta} < A^{-1+\nu}N^{-k}$$

if η is small and A is large. So α lies in a Major Arc. We have shown that if (5.1) is false then either (3.11) holds or α lies in a Major Arc. Lemma 2 follows.

7. The Major Arcs

From here on $s \geqslant c_9$ will be fixed. We will employ the O-notation, with explicit constants which may depend on $k, \lambda, \mu, \ldots, s$ only, but not on A. We will assume A to be large. We will suppose that (3.11) is false, so that by Lemma 2 we have (5.1) unless α lies in a Major Arc. We obtain from (4.6) that

$$Z = \sum_{q<A^{\nu}} \sum_{\substack{u=1 \\ (u,q)=1}}^{q} \int_{\mathfrak{M}_{qu}} f(\alpha)\,d\alpha + O(HN^{s-k}A^{-2}). \tag{7.1}$$

LEMMA 3. *For* $\alpha = (u/q) + \beta \in \mathfrak{M}_{qu}$ *we have*

$$S_i(\alpha) = q^{-1} S_i\left(\frac{u}{q}\right) I_i(\beta) + O(A^{2\nu}) \quad (i = 1, \ldots, s) \tag{7.2}$$

where

$$S_i\left(\frac{u}{q}\right) = \sum_{y=1}^{q} e\left(\frac{a_i u}{q} y^k\right) \quad \textit{and} \quad I_i(\beta) = \int_0^N e(a_i \beta \xi^k)\,d\xi. \tag{7.3}$$

Proof. Write $x = qz + y$. Then

$$S_i(\alpha) = \sum_{y=1}^{q} e\left(\frac{a_i u}{q} y^k\right) \sum_z e(a_i \beta (qz + y)^k), \tag{7.4}$$

where the sum over z is over integers z in $1 \leqslant qz + y \leqslant N$. We endeavour to approximate the sum over z by the integral of $e(a_i \beta (q\zeta + y)^k)$ with respect to ζ in the interval determined by $0 \leqslant q\zeta + y \leqslant N$. The function

$$g(\zeta) = e(a_i \beta (q\zeta + y)^k)$$

has

$$|g'(\zeta)| \leqslant 2\pi |a_i \beta| kqN^{k-1}, \quad |g(\zeta)| \leqslant 1$$

in this interval, which is of length N/q. Therefore

$$\left|\sum e(a_i\beta(qz+y)^k)-\int e(a_i\beta(q\zeta+y)^k)\,d\zeta\right|$$
$$\leqslant (N/q)(2\pi kq|a_i\beta|N^{k-1})+3\leqslant 2\pi kN^kA|\beta|+3$$
$$\leqslant 2\pi kA^{\nu}+3=O(A^{\nu}),$$

since $|\beta|\leqslant A^{-1+\nu}N^{-k}$. Taking the sum over y in (7.4) we obtain

$$S_i(\alpha)=\sum_{y=1}^{q}e\left(\frac{a_iu}{q}y^k\right)\int e(a_i\beta(q\zeta+y)^k)\,d\zeta+O(A^{2\nu}).$$

The change of variables $\xi=q\zeta+y$ yields the desired result.

Let $\mathfrak{J}(\gamma)$ be the "singular integral" defined by

$$\mathfrak{J}(\gamma)=\int_{|\beta|<\gamma}\prod_{i=1}^{s}\left(\int_0^1 e(\chi_i\xi_i^k\beta)\,d\xi_i\right)d\beta,$$

where

$$\chi_i=a_i/A \quad (i=1,\ldots,s). \tag{7.5}$$

LEMMA 4.

$$\int_{\mathfrak{M}_{qu}}f(\alpha)\,d\alpha=N^{s-k}A^{-1}q^{-s}\mathcal{S}_1\left(\frac{u}{q}\right)\ldots\mathcal{S}_s\left(\frac{u}{q}\right)\left(\sum_{z=1}^{H}e\left(-\frac{u}{q}z\right)\right)\mathfrak{J}(A^{\nu})+O(HN^{s-k-1}A^{-1+3\nu}).$$

Proof. Since $|S_i(\alpha)|\leqslant N$, the preceding lemma shows that for $\alpha=(u/q)+\beta\in\mathfrak{M}_{qu}$,

$$S_1(\alpha)\ldots S_s(\alpha)=q^{-s}\mathcal{S}_1\left(\frac{u}{q}\right)\ldots\mathcal{S}_s\left(\frac{u}{q}\right)I_1(\beta)\ldots I_s(\beta)+O(N^{s-1}A^{2\nu}).$$

For $1\leqslant z\leqslant H\leqslant A^{4\nu}$ we have $|\beta z|\leqslant A^{-1+\nu}N^{-k}A^{4\nu}\leqslant A^{\nu}N^{-1}$ by (3.5 ii), so that $|e(\beta z)-1|=2|\sin\pi\beta z|\leqslant 2\pi|\beta z|<A^{2\nu}N^{-1}$, whence

$$\left|e(-\alpha z)-e\left(-\frac{u}{q}z\right)\right|<A^{2\nu}N^{-1}$$

and

$$S_1(\alpha)\ldots S_s(\alpha)e(-\alpha z)=q^{-s}\mathcal{S}_1\left(\frac{u}{q}\right)\ldots\mathcal{S}_s\left(\frac{u}{q}\right)e\left(-\frac{u}{q}z\right)I_1(\beta)\ldots I_s(\beta)+O(N^{s-1}A^{2\nu}).$$

Taking the sum over z we obtain

$$f(\alpha)=\sum_{z=1}^{H}S_1(\alpha)\ldots S_s(\alpha)e(-\alpha z)$$
$$=q^{-s}\mathcal{S}_1\left(\frac{u}{q}\right)\ldots\mathcal{S}_s\left(\frac{u}{q}\right)\left(\sum_{z=1}^{H}e\left(-\frac{u}{q}z\right)\right)I_1(\beta)\ldots I_s(\beta)+O(HN^{s-1}A^{2\nu}).$$

Since $\mathfrak{M}_{qu}$ is of length $2A^{-1+\nu}N^{-k}$ we infer that

$$\int_{\mathfrak{M}_{qu}} f(\alpha)\,d\alpha = q^{-s}S_1\left(\frac{u}{q}\right)\dots S_s\left(\frac{u}{q}\right)\left(\sum_{z=1}^{H} e\left(-\frac{u}{q}z\right)\right)\mathcal{K} + O(HN^{s-k-1}A^{-1+3\nu}),$$

where

$$\mathcal{K} = \int_{|\beta|<A^{-1+\nu}N^{-k}} I_1(\beta)\dots I_s(\beta)\,d\beta.$$

Put $\xi_i = N\xi_i'$ $(i=1, \dots, s)$, $\beta = A^{-1}N^{-k}\beta'$. Then

$$a_i\beta\xi_i^k = (a_iN^k/AN^k)\beta'\xi_i'^k = \chi_i\beta'\xi_i'^k \quad (i=1, \dots, s).$$

We now have $|\beta'| \leq A^\nu$, and if $\xi = \xi_i$ in the definition (7.3) of $I_i(\beta)$ ranged in $0 \leq \xi_i \leq N$, then ξ_i' ranges in $0 \leq \xi_i' \leq 1$. Thus after a change of notation we see that

$$\mathcal{K} = N^{s-k}A^{-1}\mathcal{J}(A^\nu).$$

8. Conclusion

Recall that at the beginning of § 4 we made the convention that s be even and that half of the coefficients a_i be positive, the other half negative. Hence half of the χ_i are positive, half are negative. Moreover we have

$$\tfrac{1}{2} \leq |\chi_i| \leq 1 \quad (i=1, \dots, s) \tag{8.1}$$

by (4.2) and (7.5).

Lemma 5. *Under the conditions just stated, and assuming $s>k$, the limit of $\mathcal{J}(\gamma)$ as $\gamma \to \infty$ exists; denote this limit by $\mathcal{J}(\infty)$. Here $\mathcal{J}(\gamma)$ and $\mathcal{J}(\infty)$ depend on $\chi_1, \dots, \chi_s$, but the convergence to the limit is uniform in $\chi_1, \dots, \chi_s$ subject to* (8.1). *Moreover,*

$$\mathcal{J}(\infty) \geq c_{12}(k, s) > 0.$$

Proof. This was shown in [4, § 7], which in turn had a reference to [2].(1)

Since the number of summands on the right hand side of (7.1) is $< A^{2\nu}$, Lemma 4 yields

$$Z = N^{s-k}A^{-1}\mathfrak{S}\mathcal{J}(A^\nu) + O(HN^{s-k}A^{-2} + HN^{s-k-1}A^{-1+5\nu}), \tag{8.2}$$

where $\mathfrak{S}$ is the "singular series"

(1) *Added in proof.* There is a minor mistake in [4]. The integral in formula (7.3) of [4] should be replaced by $\int_\alpha^\beta \Omega(u)\,(\sin 2\pi\omega u/\pi u)\,du$, where $\alpha = -\sum_\nu \sigma_\nu$, $\beta = \sum_\nu \varrho_\nu$. Two lines below, $\Omega(\omega)$ should be $\Omega(u)$.

$$S = S(A^\nu, H) = \sum_{z=1}^{H} \sum_{q<A^\nu} \sum_{\substack{u=1\\(u,q)=1}}^{q} \sum_{y_1=1}^{q} \dots \sum_{y_s=1}^{q} q^{-s} e\left(\frac{u}{q}(a_1 y_1^k + \dots + a_s y_s^k - z)\right). \tag{8.3}$$

The summands $q=1$ give the contribution H to the multiple sum on the right hand side. When $q>1$,

$$\left|\sum_{z=1}^{H} e\left(-\frac{u}{q}z\right)\right| < q,$$

so that the summands with fixed $q>1$ contribute $O(q^2)$. Taking the sum over q in $1<q<A^\nu$ we get a total contribution $O(A^{3\nu})$, which is of smaller order of magnitude than H by (4.3). Hence if A is sufficiently large,

$$|S| > \tfrac{1}{2}H.$$

On the other hand by Lemma 5,

$$|\mathfrak{J}(A^\nu)| \geq \tfrac{1}{2}c_{12}$$

if A is large. Hence the main term in (8.2) will be

$$> (c_{12}/4)HN^{s-k}A^{-1}.$$

This is for large A of a greater order of magnitude than the error term, since

$$HN^{s-k-1}A^{-1+5\nu} = O(HN^{s-k-1}A^{-1}N^{5\nu/\varrho}) = O(HN^{s-k-(1/2)}A^{-1})$$

by (4.3) and (3.5 iii). Thus $Z>0$ if A is sufficiently large. Our proof of the proposition and hence of the theorem is complete.

References

[1]. Birch, B. J,, Small zeros of diagonal forms of odd degree in many variables. *Proc. London Math. Soc.* (*3*), 21 (1970), 12–18.

[2]. Davenport, H., Analytic methods for diophantine equations and diophantine inequalities. Lecture Notes, Univ. Michigan, 1962.

[3]. Pitman, J., Bounds for solutions of diagonal equations. *Acta Arith.*, 19 (1971), 223–247.

[4]. Schmidt, W. M., Small zeros of additive forms in many variables. *Trans. Amer. Math. Soc.*, 248 (1979), 121–133.

Received March 8, 1979