Conspectus materiae tomi XXXIV, fasciculi 2

La revue est consacrée à la Théorie des Nombres
The journal publishes papers on the Theory of Numbers
Die Zeitschrift veröffentlicht Arbeiten aus der Zahlentheorie
Журнал посвящен теории чисел

L'adresse de la Rédaction et de l'échange
Address of the Editiorial Board and of the exchange
Die Adresse der Schriftleitung und des Austausches
Адрес редакции и книгообмена

ACTA ARITHMETICA

ul. Śniadeckich 8, 00-950 Warszawa

Les auteurs sont priés d'envoyer leurs manuscrits en deux exemplaires
The authors are requested to submit papers in two copies
Die Autoren sind gebeten um Zusendung von 2 Exemplaren jeder Arbeit
Рукописи статей редакция просит предлагать в двух экземплярах

PRINTED IN POLAND

W R O C Ł A W S K A D R U K A R N I A N A U K O W A

ACTA ARITHMETICA
XXXIV (1978)

On the fractional parts of cubic forms

by

R. J. Cook (Sheffield)

1. Introduction. In 1948 Heilbronn [10] proved that for every $\varepsilon > 0$ and $N > 1$ and every real θ there is an integer x such that

$$(1.1)\qquad 1 \leqslant x \leqslant N \quad \text{and} \quad \|\theta x^2\| < C(\varepsilon) N^{-\frac{1}{2}+\varepsilon},$$

where $C(\varepsilon)$ depends only on ε and $\|\alpha\|$ is the difference between α and the nearest integer, taken positively. The result is uniform in θ and so analogous to Dirichlet's theorem for the fractional parts of $n\theta$. Danicic [4] obtained a similar result for the fractional parts of $n^k\theta$, where k is a positive integer, the proof is more readily accessible in [9].

Danicic [5] has shown that if $\varepsilon > 0$, $N > 1$ and $Q(x)$ is a real quadratic form in n variables then there exists an integer vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ such that $0 < |\boldsymbol{x}| \leqslant N$, where $|\boldsymbol{x}| = \max |x_i|$, and

$$(1.2)\qquad \|Q(\boldsymbol{x})\| < CN^{-n/(n+1)+\varepsilon},$$

where C depends only on n and ε. In particular, we observe that the exponent of N in (1.2) tends to -1 as $n \to \infty$. Danicic also obtained results for simultaneous approximations to two quadratic forms [6].

Davenport [8] proved that a cubic form with rational coefficients represents zero non-trivially provided that the form contains at least 16 variables; an expository account of a similar result for cubic forms in 17 variables is contained in [7]. Pitman [13] proved that a real cubic form in sufficiently many variables takes small values; the method used depends on showing that a real cubic form in sufficiently many variables represents a cubic form which is almost an additive cubic form and then using the corresponding result for additive forms. We shall use similar methods to obtain results for cubic forms analogous to the results for quadratic forms obtained by Danicic [5], [6].

The distribution (mod 1) of one additive form was considered in [3].

Theorem 1. *For every $\varepsilon > 0$, $N > 1$ and additive form*

$$f(\boldsymbol{x}) = \theta_1 x_1^k + \ldots + \theta_s x_s^k,$$

of degree $k \geqslant 2$, in s variables, where $1 \leqslant s \leqslant K = 2^{k-1}$, there exist integers $x_1, \ldots, x_s$, not all zero, such that

$$\max_i |x_i| \leqslant N \quad \text{and} \quad \|f(\boldsymbol{x})\| < C(\varepsilon, k) N^{-s/K+\varepsilon} \tag{1.3}$$

where $C(\varepsilon, k)$ depends on ε and k only.

In particular, when $k = 3$ and $s = 4$ there are integers $x_1, \ldots, x_4$, not all zero, with

$$\max |x_i| \leqslant N \quad \text{and} \quad \|\theta_1 x_1^3 + \ldots + \theta_4 x_4^3\| < C(\varepsilon) N^{-1+\varepsilon}. \tag{1.4}$$

Liu ([11], [12]) obtained results for simultaneous approximations to several additive forms but some of his results have recently been improved upon by Baker and Gajraj [1].

THEOREM 2 (Baker, Gajraj, Liu). *Let*

$$f_i(\boldsymbol{x}) = \sum_{j=1}^{s} \theta_{ij} x_j^k, \quad i = 1, \ldots, R,$$

be additive forms in s variables. For any $N > 1$ and $\varepsilon > 0$ there are integers $x_1, \ldots, x_s$, not all zero, such that

$$\max_i \|f_i(\boldsymbol{x})\| < CN^{-1/g(k,R,s)+\varepsilon} \tag{1.5}$$

where $C = C(\varepsilon, k, R, s)$ and, writing K for 2^{k-1},

$$g(k, R, s) = \begin{cases} 2K+1 & \text{if} \quad R = 2,\ s = 2, \\ 2^{R+1}-1 & \text{if} \quad k = 2,\ R \geqslant 2,\ s = 1, \\ 2^{R-2}(3K+k^{-1}+1)-1 & \text{if} \quad k \geqslant 3,\ R \geqslant 2,\ s = 1. \end{cases} \tag{1.6}$$

Liu proved this result when $R = 2$ and $s = 1$, the remaining cases are due to Baker and Gajraj. In particular we see that $g(3, 2, 2) = 9$.

We apply these results for additive forms to obtain results for general cubic forms.

THEOREM 3. *Let $C(\boldsymbol{x})$ be a real cubic form in n variables $x_1, \ldots, x_n$. For any $\varepsilon > 0$ and $N > 1$ there exist integers $x_1, \ldots, x_n$, not all zero, such that*

$$\max_j |x_j| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| < AN^{-1/\varrho(n)+\varepsilon} \tag{1.7}$$

where A depends only on n and ε and $\varrho(n)$ is defined by

$$\varrho(n) = \begin{cases} 4 & \text{for} \quad n \leqslant 111, \\ 17/5 & \text{for} \quad 112 \leqslant n \leqslant 172, \\ 2+7\left([(2n)^{1/3}]-2\right)^{-1} & \text{for} \quad 173 \leqslant n < 2\cdot 12^7, \\ 4/3+5\left([(n/2)^{1/7}]-2\right)^{-1} & \text{for} \quad 2\cdot 12^7 \leqslant n < 3^8\cdot 2^{60}, \\ 1+4[(3^{-8}n)^{1/15}]^{-1} & \text{for} \quad n \geqslant 3^8 \cdot 2^{60}, \end{cases} \tag{1.8}$$

where $[\alpha]$ denotes the integer part of α.

In particular, as $n \to \infty$

$$1 - \frac{1}{\varrho(n)} \sim 4\cdot 3^{8/15} n^{-1/15}. \tag{1.9}$$

THEOREM 4. *For any $\varepsilon > 0$, $N > 1$ and any two real cubic forms $C^1(\boldsymbol{x})$, $C^2(\boldsymbol{x})$ in n variables there exist integers $x_1, \ldots, x_n$, not all zero, such that*

$$\max_j |x_j| \leqslant N \quad \text{and} \quad \max\left(\|C^1(\boldsymbol{x})\|, \|C^2(\boldsymbol{x})\|\right) < AN^{-1/\eta(n)+\varepsilon} \tag{1.10}$$

where A depends only on n and ε and $\eta(n)$ is defined by

$$\eta(n) = \begin{cases} 12+1/3 & \text{for} \quad n < 4\cdot 11^3, \\ 9+28/([(n/4)^{1/3}]-2) & \text{for} \quad n \geqslant 4\cdot 11^3. \end{cases} \tag{1.11}$$

In particular, as $n \to \infty$

$$\frac{1}{9} - \frac{1}{\eta(n)} \sim \left(\frac{28}{81}\right) 4^{1/3} n^{-1/3}. \tag{1.12}$$

The ideas used in proving Theorems 3 and 4 can also be used to deduce results for $R > 2$ cubic forms from the corresponding result for R additive cubic forms, and since it requires little extra work we include details of this in Section 5. However, at present the results known for simultaneous approximations to $R > 2$ additive forms in s variables are no better than the results obtained when $s = 1$. We use Vinogradov's $\ll$-notation where the implicit constants depend on n, R and ε.

I am indebted to the referee for many useful comments on the original draft of this paper.

2. Reduction to almost diagonal form. Let $C^1(\boldsymbol{x}), \ldots, C^R(\boldsymbol{x})$ be cubic forms in n variables, say

$$C^l(\boldsymbol{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \gamma_{ijk}^l x_i x_j x_k \quad \text{for} \quad l = 1, \ldots, R \tag{2.1}$$

where the real coefficients γ_{ijk}^l are symmetrical in the three suffices. With the cubic forms $C^l(\boldsymbol{x})$ we associate trilinear forms

$$T^l(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \gamma_{ijk}^l x_i y_j z_k \quad \text{for} \quad l = 1, \ldots, R. \tag{2.2}$$

LEMMA 1. *Let $L_1, \ldots, L_t$ be t real linear forms in $x_1, \ldots, x_n$, say*

$$L_i = \sum_{j=1}^{m} \gamma_{ij} x_j \quad \text{for} \quad i = 1, \ldots, t.$$

Then for any $P > 1$ there exist integers $x_1, \ldots, x_m$, not all zero, such that

(2.3) $$|x_j| \leqslant P \quad \textit{for} \quad j = 1, \ldots, m,$$

and

(2.4) $$\|L_i(\boldsymbol{x})\| \leqslant P^{-m/t} \quad \textit{for} \quad i = 1, \ldots, t.$$

This is Theorem VI of Chapter 1 of Cassels [2].

If $\boldsymbol{a} = (a_1, \ldots, a_n)$ then we write $|\boldsymbol{a}|$ for $\max|a_i|$. If $m \leqslant n$ let $V(m)$ denote the subspace of $\boldsymbol{R}^n$ consisting of those vectors $\boldsymbol{a}$ for which $a_i = 0$ for $m < i \leqslant n$. For any integer $s > 1$ we choose non-zero integral vectors $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s \in \boldsymbol{R}^n$ and put

(2.5) $$\boldsymbol{x} = u_1\boldsymbol{a}^1 + \ldots + u_s\boldsymbol{a}^s.$$

Then for $l = 1, \ldots, R$

(2.6) $$C^l(\boldsymbol{x}) = \sum_{i=1}^{s}\sum_{j=1}^{s}\sum_{k=1}^{s}\lambda_{ijk}^l u_i u_j u_k = D^l(\boldsymbol{u})$$

say, where the coefficients λ_{ijk}^l are symmetrical in the three suffices and are given by

$$\lambda_{ijk}^l = T^l(\boldsymbol{a}^i, \boldsymbol{a}^j, \boldsymbol{a}^k).$$

Let

(2.7) $$n \geqslant n_1 > n_2 > \ldots > n_s > n_{s+1} = 0,$$

then we shall choose $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s$ in turn such that $\boldsymbol{a}^i \in V(n_i)$, $\|T^l(\boldsymbol{a}^i, \boldsymbol{a}^i, \boldsymbol{a}^j)\|$ is small for $j > i$ and $\|T^l(\boldsymbol{a}^i, \boldsymbol{y}, \boldsymbol{z})\|$ is small for all suitably bounded $\boldsymbol{y}$ and $\boldsymbol{z}$ in $V(n_{i+1})$. Then $\|T^l(\boldsymbol{a}^i, \boldsymbol{a}^j, \boldsymbol{a}^k)\|$ is small whenever i, j and k are not all equal and so the non-diagonal terms of $D^l(\boldsymbol{u})$ are nearly integers. The result obtained by this procedure is given in the following lemma which is based on Lemma 2 of Pitman [13].

LEMMA 2. *Let $P > 1$, $Z > 1$ and $n_1, \ldots, n_{s+1}$ be integers satisfying* (2.7). *We put*

(2.8) $$\mu_j = \frac{n_j}{R\{(j-1) + \frac{1}{2}n_{j+1}(n_{j+1}+1)\}}$$

for $j = 1, \ldots, s$. If

(2.9) $$\mu_j \geqslant Z+2 \quad \textit{for} \quad j = 1, \ldots, s-1, \quad \mu_s \geqslant Z,$$

then there exist non-zero integral vectors $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s$ such that

(2.10) $$|\boldsymbol{a}^j| \leqslant P \quad \textit{and} \quad \boldsymbol{a}^j \in V(n_j) \quad \textit{for} \quad j = 1, \ldots, s$$

and if i, j and k are not all equal then

(2.11) $$\|T^l(\boldsymbol{a}^i, \boldsymbol{a}^j, \boldsymbol{a}^k)\| \leqslant n^2P^{-Z} \quad \textit{for} \quad l = 1, \ldots, R.$$

Proof. In order to choose $\boldsymbol{a}^1$ we see that if $j > 1$, $k > 1$ and $|\boldsymbol{a}^i| \leqslant P$ for $i = 1, \ldots, s$ then for $l = 1, \ldots, R$

$$\|T^l(\boldsymbol{a}^1, \boldsymbol{a}^j, \boldsymbol{a}^k)\| = \Big\|\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{r=1}^{n}\gamma_{pqr}^l a_p^1 a_q^j a_r^k\Big\| \leqslant \sum_{q=1}^{n_2}\sum_{r=1}^{n_2}\Big\|\sum_{p=1}^{n_1}\gamma_{pqr}^l a_p^1\Big\|P^2$$

since $V(n_j)$ and $V(n_k)$ are both contained in $V(n_2)$. Thus it is sufficient to choose $\boldsymbol{a}^1$ satisfying (2.10) and

(2.12) $$\Big\|\sum_{p=1}^{n_1}\gamma_{pqr}^l a_p^1\Big\| \leqslant P^{-Z-2},$$

for $l = 1, \ldots, R$ and $1 \leqslant q, r \leqslant n_2$. This is a set of at most $\frac{1}{2}n_2(n_2+1)R$ distinct linear inequalities in n_1 variables, and

$$\mu_1 = \frac{n_1}{\frac{1}{2}n_2(n_2+1)R} > Z+2.$$

Therefore, by Lemma 1, we can choose a non-zero integral vector $\boldsymbol{a}^1$ which satisfies (2.10) and (2.12).

When $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^{j-1}$ have been chosen we use Lemma 1 to choose a non-zero integral vector $\boldsymbol{a}^j$ satisfying (2.10) and such that for $l = 1, \ldots, R$,

(2.13) $$\Big\|\sum_{p=1}^{n_j}\gamma_{pqr}^l a_p^j\Big\| \leqslant P^{-Z-2} \quad \text{for} \quad 1 \leqslant q, r \leqslant n_{j+1}$$

and

(2.14) $$\|T^l(\boldsymbol{a}^i, \boldsymbol{a}^i, \boldsymbol{a}^j)\| \leqslant P^{-Z-2} \quad \text{for} \quad 1 \leqslant i < j.$$

Since at most

$$R\{(j-1) + \tfrac{1}{2}n_{j+1}(n_{j+1}+1)\}$$

of these inequalities are distinct and

$$\mu_j = \frac{n_j}{R\{(j-1) + \frac{1}{2}n_{j+1}(n_{j+1}+1)\}} \geqslant Z+2$$

it follows from Lemma 1 that we can choose $\boldsymbol{a}^j$ satisfying (2.10), (2.13) and (2.14). Then for any $\boldsymbol{y}$, $\boldsymbol{z}$ in $V(n_{j+1})$ with $|\boldsymbol{y}| \leqslant P$ and $|\boldsymbol{z}| \leqslant P$ we have

$$\|T^l(\boldsymbol{a}^j, \boldsymbol{y}, \boldsymbol{z})\| = \Big\|\sum_{p=1}^{n_j}\sum_{q=1}^{n_{j+1}}\sum_{r=1}^{n_{j+1}}\gamma_{pqr}^l a_p^j y_q z_r\Big\| \leqslant \sum_{q=1}^{n_{j+1}}\sum_{r=1}^{n_{j+1}}\Big\|\sum_{p=1}^{n_j}\gamma_{pqr}^l a_p^j\Big\|P^2 \leqslant n^2P^{-Z}.$$

When $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^{s-1}$ have been chosen we can choose $\boldsymbol{a}^s$ which satisfies (2.10) and which also satisfies the $R(s-1)$ linear inequalities

$$\|T^l(\boldsymbol{a}^i, \boldsymbol{a}^i, \boldsymbol{a}^s)\| \leqslant P^{-Z}$$

for $i = 1, \ldots, s-1$, $l = 1, \ldots, R$ since $\mu_s \geqslant Z$. Then $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s$ have the required properties.

3. Proof of Theorem 3. We now take $R = 1$ and $s = 4$. Having reduced $C(\boldsymbol{x})$ to a form

(3.1) $$D(\boldsymbol{u}) = \sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{k=1}^{4}\lambda_{ijk}u_iu_ju_k$$

whose off-diagonal coefficients are nearly integers, we apply Theorem 1 to the additive form

(3.2) $$D_0(\boldsymbol{u}) = \lambda_{111}u_1^3+\ldots+\lambda_{444}u_4^3.$$

For any $U > 1$, $\varepsilon > 0$ there exist integers $u_1, \ldots, u_4$, not all zero, such that

(3.3) $$\max_i |u_i| \leqslant U \quad \text{and} \quad \|D_0(\boldsymbol{u})\| \ll U^{-1+\varepsilon}.$$

From Lemma 2 we have

(3.4) $$\|D(\boldsymbol{u})-D_0(\boldsymbol{u})\| \ll U^3P^{-Z}$$

so that the substitution

(3.5) $$\boldsymbol{x} = u_1\boldsymbol{a}^1+\ldots+u_4\boldsymbol{a}^4$$

gives $|\boldsymbol{x}| \leqslant 4UP$ and

(3.6) $$\|C(\boldsymbol{x})\| = \|D(\boldsymbol{u})\| \ll U^{-1+\varepsilon}+U^3P^{-Z}.$$

We choose U and P so that $4UP = N$ and $U^{-1} = U^3P^{-Z}$, that is

(3.7) $$U = (N/4)^{Z/(Z+4)} \quad \text{and} \quad P = (N/4)^{4/(Z+4)}.$$

If $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^4$ are linearly independent we have

(3.8) $$0 < |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| \ll N^{-Z/(Z+4)+\varepsilon}.$$

If $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^4$ are linearly dependent then for some j

(3.9) $$\alpha_1\boldsymbol{a}^1+\ldots+\alpha_j\boldsymbol{a}^j = 0$$

where $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^{j-1}$ are linearly independent and $\alpha_1, \ldots, \alpha_j$ are integers. Then

$$\beta_1\boldsymbol{a}^1+\ldots+\beta_{j-1}\boldsymbol{a}^{j-1} = \boldsymbol{a}^j$$

where $\beta_i = -\alpha_i/\alpha_j$. Applying Cramer's rule to a suitable subset of $j-1$ of these equations in $\beta_1, \ldots, \beta_{j-1}$ we have

$$\beta_i = \Delta_i/\Delta \quad \text{for} \quad i = 1, \ldots, j-1$$

where Δ is the determinant of a non-singular $(j-1)\times(j-1)$ integral matrix with components of absolute value at most P so

$$1 \leqslant \Delta \leqslant (j-1)P^{j-1}$$

and, similarly, for $i = 1, \ldots, j-1$

$$\Delta_i \ll P^{j-1} \quad \text{and so} \quad \beta_i \ll P^{j-1}.$$

Hence we can take $\alpha_j = \Delta \leqslant (j-1)P^{j-1}$ and then

$$\alpha_i = -\beta_i\alpha_j \ll P^{2(j-1)} \quad \text{for} \quad i = 1, \ldots, j-1.$$

Since $\|n\theta\| \leqslant |n|\,\|\theta\|$ when n is an integer we have

(3.10) $$\|C(\alpha_j\boldsymbol{a}^j)\| = \|T(\alpha_j\boldsymbol{a}^j, \alpha_j\boldsymbol{a}^j, \alpha_j\boldsymbol{a}^j)\| = \left\|\alpha_j^2\sum_{i=1}^{j-1}(-\alpha_i)T(\boldsymbol{a}^i, \boldsymbol{a}^j, \boldsymbol{a}^j)\right\|$$
$$\ll P^{4(j-1)}P^{-Z} \ll P^{12-Z} \ll N^{4(12-Z)/(Z+4)} \ll N^{-Z/(Z+4)}$$

provided that $Z \geqslant 16$. Then, as $Z \geqslant 16$,

$$|\alpha_j\boldsymbol{a}^j| \leqslant (j-1)P^{j-1}\cdot P \leqslant 3P^4 < N.$$

For $Z \geqslant 4$ the inequalities (2.9) will be satisfied if

(3.11) $$n_4 \geqslant 3Z \quad \text{and} \quad n_j \geqslant n_{j+1}^2 Z \quad \text{for} \quad j = 1, 2, 3$$

since we then have $n_{j+1} \geqslant \max(12, Z)$ for $j = 1, 2, 3$ and so

$$\begin{aligned} n_j &\geqslant n_{j+1}^2Z \geqslant n_{j+1}^2(\tfrac{1}{2}Z+2) \geqslant \tfrac{1}{2}n_{j+1}^2Z+n_{j+1}(n_{j+1}+\tfrac{1}{2}n_{j+1}+6)\\ &\geqslant \tfrac{1}{2}n_{j+1}^2Z+n_{j+1}(n_{j+1}+1)+Z(\tfrac{1}{2}n_{j+1}+5) = \tfrac{1}{2}n_{j+1}(n_{j+1}+1)(Z+2)+5Z\\ &> \tfrac{1}{2}n_{j+1}(n_{j+1}+1)(Z+2)+(Z+2)(j-1). \end{aligned}$$

From (3.11) we deduce that $n \geqslant n_1 \geqslant 3^8Z^{15}$. Therefore if $n \geqslant 3^8 2^{60}$ we take $Z = [(3^{-8}n)^{1/15}]$, where $[x]$ denotes the integer part of x, and choose n_1, n_2, n_3 and n_4 to be defined by

$$n_4 = 3Z \quad \text{and} \quad n_j = n_{j+1}^2Z \quad \text{for} \quad j = 1, 2, 3.$$

If $n < 3^8 2^{60}$ we take $s = 3$ and apply Theorem 1 to

$$\lambda_{111}u_1^3+\lambda_{222}u_2^3+\lambda_{333}u_3^3.$$

In place of (3.5) we use the substitution

(3.12) $$\boldsymbol{x} = u_1\boldsymbol{a}^1+u_2\boldsymbol{a}^2+u_3\boldsymbol{a}^3$$

to obtain

(3.13) $$\|C(\boldsymbol{x})\| \ll U^{-3/4+\varepsilon}+U^3P^{-Z}.$$

Taking

(3.14) $$U = (N/3)^{4Z/(4Z+15)} \quad \text{and} \quad P = (N/3)^{15/(4Z+15)}$$

we have

(3.15) $$3UP = N \quad \text{and} \quad U^{-3/4} = U^3P^{-Z}.$$

If $\boldsymbol{a}^1, \boldsymbol{a}^2$ and $\boldsymbol{a}^3$ are linearly independent then

(3.16) $$0 < |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| \ll N^{-3Z/(4Z+15)+\varepsilon}.$$

If $\boldsymbol{a}^1, \boldsymbol{a}^2$ and $\boldsymbol{a}^3$ are linearly dependent then, in place of (3.10) we see that for some j there is an integer a_j such that $|a_j \boldsymbol{a}^j| \leqslant 2P^3 < N$ and

$$\|C(a_j \boldsymbol{a}^j)\| \ll P^{8-Z} \ll N^{-3Z/(4Z+15)}$$

provided that $Z \geqslant 10$. The inequalities (2.9) will be satisfied if

$$n_3 = 2Z, \quad n_2 = 2(Z+1)^3$$

and

$$n_1 \geqslant (Z+2)\{(Z+1)^3(2(Z+1)^3+1)\} = 2(Z+1)^6(Z+2)+(Z+1)^3(Z+2)$$

so we take $n_1 = 2(Z+2)^7$ where Z is an integer, $Z \geqslant 10$. Thus for $n \geqslant 2\cdot 12^7$ we can take $Z = [(n/2)^{1/7}]-2$ and find an integral vector $\boldsymbol{x}$ satisfying (3.16).

If $n < 2\cdot 12^7$ we take $s = 2$,

(3.17) $$U = (N/2)^{2Z/(2Z+7)} \quad \text{and} \quad P = (N/2)^{7/(2Z+7)}.$$

The substitution $\boldsymbol{x} = u_1\boldsymbol{a}^1 + u_2\boldsymbol{a}^2$ gives an integral vector $\boldsymbol{x}$ satisfying

(3.18) $$|\boldsymbol{x}| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| \ll N^{-Z/(2Z+7)+\varepsilon}.$$

If $\boldsymbol{a}^1$ and $\boldsymbol{a}^2$ are linearly dependent then, as $\boldsymbol{a}^1 \neq 0$, in place of (3.10) we obtain

$$|a_2\boldsymbol{a}^2| \leqslant P^2 < N \quad \text{provided that} \quad Z > 7/2,$$

and

$$\|C(a_2\boldsymbol{a}^2)\| \ll P^{4-Z} \ll N^{7(4-Z)/(2Z+7)} \ll N^{-Z/(2Z+7)}$$

provided that $Z \geqslant 14/3$. The inequalities (2.9) become

$$n_2 \geqslant Z \quad \text{and} \quad n_1 \geqslant (Z+2)\big(1+\tfrac{1}{2}n_2(n_2+1)\big).$$

If $2\cdot 12^7 > n \geqslant 173$ we take $n_2 = Z = [(2n)^{1/3}]-2 \geqslant 5$ to obtain an integral vector $\boldsymbol{x}$ satisfying

(3.19) $$0 < |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| \ll N^{-Z/(2Z+7)+\varepsilon}.$$

If $172 \geqslant n \geqslant 112$ we take $n_2 = Z = 5$ to obtain an integral vector $\boldsymbol{x}$ satisfying

(3.20) $$0 < |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| \ll N^{-5/17+\varepsilon}.$$

If $n \leqslant 111$ we consider the section $C(x_1, \boldsymbol{0})$ and apply a result of Danicic [4] to obtain an integral vector $\boldsymbol{x}$ satisfying

(3.21) $$0 < |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \|C(\boldsymbol{x})\| \ll N^{-1/4+\varepsilon},$$

which completes the proof of Theorem 3.

It is possible to make some improvement in this result when n is large. If $n_1, \ldots, n_4$ are all sufficiently large and $\varphi > 0$ then the inequalities (2.9) will be satisfied if

(3.22) $$n_4 \geqslant 3Z \quad \text{and} \quad n_j \geqslant \tfrac{1}{2}(1+\varphi)n_{j+1}^2 Z \quad \text{for} \quad j = 1, 2, 3,$$

so that

$$n_1 \geqslant \big(\tfrac{1}{2}(1+\varphi)\big)^7 3^8 Z^{15}.$$

Therefore, for all sufficiently large n, we may take $\varrho(n)$ to satisfy

(3.23) $$-1/\varrho(n) = -1+4(1+\psi)(3^{-8}2^7 n)^{-1/15}$$

where ψ is an arbitrary positive constant and the formula is valid for $n > n_0(\psi)$.

4. Proof of Theorem 4. We take $R = 2$ and $s = 2$. We reduce $C^l(x)$, $l = 1, 2$, to forms

(4.1) $$D^l(\boldsymbol{u}) = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \lambda_{ijk}^l u_i u_j u_k$$

whose off-diagonal coefficients are nearly integers. We apply Theorem 2, with $k = 3$, $R = s = 2$, to the additive forms

(4.2) $$D_0^l(\boldsymbol{u}) = \lambda_{111}^l u_1^3 + \lambda_{222}^l u_2^3, \quad l = 1, 2.$$

For any $\varepsilon > 0$, $U > 1$ there exist integers u_1 and u_2, not both zero, such that

(4.3) $$\max_i |u_i| \leqslant U \quad \text{and} \quad \max_l \|D_0^l(\boldsymbol{u})\| \ll U^{-1/9+\varepsilon}.$$

From Lemma 2 we have

(4.4) $$\|D^l(\boldsymbol{u}) - D_0^l(\boldsymbol{u})\| \ll U^3 P^{-Z} \quad \text{for} \quad l = 1, 2,$$

so that the substitution

(4.5) $$\boldsymbol{x} = u_1\boldsymbol{a}^1 + u_2\boldsymbol{a}^2$$

gives $|\boldsymbol{x}| \leqslant 2UP$ and, for $l = 1, 2$,

(4.6) $$\|C^l(\boldsymbol{x})\| = \|D^l(\boldsymbol{u})\| \ll U^{-1/9+\varepsilon} + U^3P^{-Z}.$$

We choose U and P so that $2UP = N$ and $U^3P^{-Z} = U^{-1/9}$, that is

(4.7) $$U = (N/2)^{Z/(Z+28/9)} \quad \text{and} \quad P = (N/2)^{28/9(Z+28/9)}.$$

If $\boldsymbol{a}^1$ and $\boldsymbol{a}^2$ are linearly independent we have

(4.8) $$0 < |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \max_l \|C^l(\boldsymbol{x})\| \ll N^{-1/(9+28/Z)+\varepsilon}.$$

If $\boldsymbol{a}^1$ and $\boldsymbol{a}^2$ are linearly dependent then there is an integer $a_2 \ll P$ such that $|a_2\boldsymbol{a}^2| \leqslant P^2 < N$ and for $l = 1, 2$

$$\|C^l(a_2\boldsymbol{a}^2)\| = \|a_2^2(-a_1)T(\boldsymbol{a}^l, \boldsymbol{a}^2, \boldsymbol{a}^2)\| \ll P^{4-Z} \ll N^{28(4-Z)/(9Z+28)} \ll N^{-1/9}$$

for $Z \geqslant 5$. The inequalities (2.9) will be satisfied if $n_2 \geqslant 2Z$ and $n \geqslant (Z+2)n_2(n_2+1)$. For $n \geqslant 4\cdot 11^3$ we take

$$n_2 = 2Z \quad \text{and} \quad Z = [(n/4)^{1/3}] - 2 \geqslant 9,$$

so that $9 + 28/Z < 12 + 1/3$, and we obtain a non-zero integral vector satisfying (4.8). If $n < 4\cdot 11^3$ then Theorem 4 follows from the case $s = 1$ of Theorem 2.

5. The case of $R > 2$ cubic forms. Suppose that we have a result of the following form:

PROPOSITION. *For any $\varepsilon > 0$, $U > 1$ and any R additive cubic forms $f^l(u)$ in $s \geqslant 2$ variables there exists an integral vector $\boldsymbol{u}$ such that*

$$(5.1) \qquad 0 < |\boldsymbol{u}| \leqslant U, \quad \max_{l} \|f^l(\boldsymbol{u})\| < C(\varepsilon, R, s)\, U^{-\sigma+\varepsilon}$$

where $\sigma = \sigma(R, s)$ satisfies $0 < \sigma \leqslant 1$.

Given R cubic forms $C^l(\boldsymbol{x})$ in n variables, where n is so large that we can choose integers $n_1, \ldots, n_s$ satisfying (2.7), (2.8) and (2.9), then we can choose integral vectors $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s$ so that the substitution

$$(5.2) \qquad \boldsymbol{x} = u_1\boldsymbol{a}^1 + \ldots + u_s\boldsymbol{a}^s$$

gives $C^l(\boldsymbol{x}) = D^l(\boldsymbol{u})$ for $l = 1, \ldots, R$ where

$$(5.3) \qquad \|D^l(\boldsymbol{u}) - D_0^l(\boldsymbol{u})\| \ll U^3P^{-Z} \quad \text{for} \quad l = 1, \ldots, R,$$

and $D_0^l(\boldsymbol{u})$ is the additive form corresponding to $D^l(\boldsymbol{u})$. Applying the hypothesis we obtain an integral vector $\boldsymbol{u}$ such that

$$(5.4) \qquad 0 < |\boldsymbol{u}| \leqslant U, \quad \max_l \|D_0^l(\boldsymbol{u})\| \ll U^{-\sigma+\varepsilon}.$$

Then $|\boldsymbol{x}| \leqslant sUP$ and

$$(5.5) \qquad \max_l \|C^l(\boldsymbol{x})\| \ll U^{-\sigma+\varepsilon} + U^3P^{-Z}.$$

We take

$$(5.6) \qquad U = (N/s)^{Z/(Z+\sigma+3)} \quad \text{and} \quad P = (N/s)^{(\sigma+3)/(Z+\sigma+3)}$$

so that

$$(5.7) \qquad |\boldsymbol{x}| \leqslant N \quad \text{and} \quad \max_l \|C^l(\boldsymbol{x})\| \ll N^{-\sigma(1+(\sigma+3)/Z)^{-1}+\varepsilon}.$$

If $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s$ are linearly independent this provides a suitable integral vector $\boldsymbol{x}$. If $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^s$ are linearly dependent then an argument similar to that in § 3 shows that for some $\boldsymbol{a}^j$ there is an integer a_j such that for $l = 1, \ldots, R$

$$(5.8) \qquad \|C^l(a_j\boldsymbol{a}^j)\| \ll P^{4(j-1)}P^{-Z} \ll P^{4s-Z} \ll N^{-\sigma},$$

provided that

$$(5.9) \qquad Z > 4s + \sigma + (\sigma^2 + 4s\sigma)/3 = \tau, \text{ say},$$

and $|a_j\boldsymbol{a}^j| \leqslant sP^s \leqslant N$, which will be satisfied when $Z > (s-1)(\sigma+3)$ and this last condition is weaker than (5.9).

THEOREM 5. *Suppose that the proposition holds for some σ. If Z satisfies* (5.9) *and there exist integers $n_1, \ldots, n_s$ satisfying*

$$(5.10) \qquad n_j \geqslant R\{(j-1) + \tfrac{1}{2}n_{j+1}(n_{j+1}+1)\}(Z+2) \quad \text{for} \quad j = 1, \ldots, s-1$$

and

$$(5.11) \qquad n_s \geqslant R(s-1)Z$$

then for any $\varepsilon > 0$, $N > 1$ and R cubic forms $C^l(\boldsymbol{x})$ in $n \geqslant n_1$ variables there exists a non-zero integral vector $\boldsymbol{x}$ satisfying

$$(5.12) \qquad \max_j |x_j| \leqslant n \quad \textit{and} \quad \max_l \|C^l(\boldsymbol{x})\| \ll N^{-\varrho+\varepsilon}$$

where

$$(5.13) \qquad \varrho = \frac{\sigma Z}{Z+\sigma+3} = \sigma - \frac{\sigma(\sigma+3)}{Z+\sigma+3},$$

and the implicit constant in (5.12) *depends only on ε, R, Z and s.*

The inequalities (5.10) and (5.11) will be satisfied if

$$(5.14) \qquad n_s \geqslant R(s-1)Z \quad \text{and} \quad n_j \geqslant RZn_{j+1}^2 \quad \text{for} \quad j = 1, \ldots, s-1,$$

since (5.9) implies that $Z > 8$. Then

$$n_{s-r} \geqslant (RZ)^{2^{r+1}-1}(s-1)^{2^r} \quad \text{for} \quad r = 0, \ldots, s-1$$

and in particular

$$(5.15) \qquad n_1 \geqslant (RZ)^{2^s-1}(s-1)^{2^{s-1}}.$$

Write

$$(5.16) \qquad f(n) = \left[R^{-1}\left(n/(s-1)^{2^{s-1}}\right)^{1/(2^s-1)}\right].$$

If $f(n) \geqslant \tau$ we may take $Z = f(n)$ in Theorem 5.

Corollary. *Suppose that the proposition holds for some σ and that $f(n) \geqslant \tau$, where $f(n)$ and τ are as above. Then for any $\varepsilon > 0$, $N > 1$ and R cubic forms in n variables there exists a non-zero integral vector $\boldsymbol{x}$ satisfying*

$$\max_j |x_j| \leqslant N \quad \text{and} \quad \max_l \|C^l(\boldsymbol{x})\| \ll N^{-\varrho+\varepsilon} \tag{5.17}$$

where

$$\varrho = \sigma\{1+O(n^{-1/(2^8-1)})\} \quad \text{as} \quad n \to \infty \tag{5.18}$$

and the implicit constants depend only on ε, R, n and s.

References

[1] R. C. Baker and J. Gajraj, *On the fractional parts of certain additive forms*, Proc. Camb. Phil. Soc. 79(1976), pp. 463–467.
[2] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge University Press, 1957.
[3] R. J. Cook, *The fractional parts of an additive form*, Proc. Camb. Phil. Soc. 72 (1972), pp. 209–212.
[4] I. Danicic, *Ph. D. thesis*, University of London, 1957.
[5] — *An extension of a theorem of Heilbronn*, Mathematika 5 (1958), pp. 30–37.
[6] — *The distribution* (mod 1) *of pairs of quadratic forms with integer variables*, J. London Math. Soc. 42 (1967), pp. 618–623.
[7] H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, Ann Arbor, Michigan, 1962.
[8] — *Cubic forms in 16 variables*, Proc. Roy. Soc. A 272 (1963), pp. 285–303.
[9] — *On a theorem of Heilbronn*, Quart. J. Math. (Oxford) (2) 18 (1967), pp. 339–344.
[10] H. Heilbronn, *On the distribution of the sequence $n^2\theta$* (mod 1), ibid. 19 (1948), pp. 249–256.
[11] M. C. Liu, *Simultaneous approximation of two additive forms*, Proc. Camb. Phil. Soc. 75 (1974), pp. 77–82.
[12] — *Simultaneous approximation of additive forms*, Trans. Amer. Math. Soc. 206(1975), pp. 361–373.
[13] Jane Pitman, *Cubic inequalities*, J. London Math. Soc. 43 (1968), pp. 119–126.

THE UNIVERSITY, Sheffield, England

Received on 27. 2. 1975
and in revised form on 6. 9. 1975 (758)

ACTA ARITHMETICA
XXXIV (1978)

Darstellungsmaße binärer quadratischer Formen über totalreellen algebraischen Zahlkörpern

von

Horst Pfeuffer (Mainz)

Von den lokalen Faktoren in der Maßformel des Hauptsatzes der Siegelschen analytischen Theorie der quadratischen Formen über algebraischen Zahlkörpern sind die *dyadischen Darstellungsdichten* am umständlichsten zu bestimmen, jedoch hängt der Aufwand sehr von der Berechnungsmethode ab. Die Darstellungsdichte einer *binären, lokal nicht modularen Form* durch sich selbst, die von Körner ([6], Satz 3) für quadratische Zahlkörper bestimmt wurde, läßt sich ohne Mehraufwand auch für beliebige Zahlkörper angeben, wenn man geeignete Invarianten von O'Meara [3] benutzt.

Die $\mathfrak{p}$-adische Darstellungsdichte eines Gitters hängt nur von seiner Lokalisierung nach $\mathfrak{p}$ ab. Es genügt daher, $\mathfrak{o}$-*Gitter G auf einem zweidimensionalen Raum mit nichtausgearteter quadratischer Form über einem $\mathfrak{p}$-adischen Zahlkörper mit Ganzheitsbereich* $\mathfrak{o}$ zu betrachten. Ein nicht modulares Gitter ist bis auf Skalarfaktoren an der Form von der Gestalt

$$G = \mathfrak{o}x_1 \perp \mathfrak{o}x_2 \cong \begin{pmatrix} 1 & 0 \\ 0 & \pi^s a \end{pmatrix}$$

mit $a \in \mathfrak{u}$, $\mathfrak{s}_2 = \mathfrak{p}^s = \mathfrak{d}G$ und $dG \underset{2}{\cong} \pi^s a$. Die Bezeichnungen sind aus [3] entnommen oder in [7] erklärt.

Mit den Hilfssätzen 2, 4, 5 und 7 sowie (28) aus [7] folgt

$$d_{\mathfrak{p}}(G, G) = \mathfrak{N}\mathfrak{p}^{s+e} d_{\mathfrak{p}}(G, \mathfrak{o}x_1),$$

worin

$$d_{\mathfrak{p}}(G, \mathfrak{o}x_1) = \frac{A_{\mathfrak{p}^\nu}(G, \mathfrak{o}x_1)}{\mathfrak{N}\mathfrak{p}^\nu}$$

für genügend großes ν und $A_{\mathfrak{p}^\nu}(G, \mathfrak{o}x_1)$ die Anzahl der modulo $\mathfrak{p}^\nu G$ verschiedenen Vektoren $x = \xi x_1 + \eta x_2$ aus G mit

$$(x, x) = \xi^2 + \pi^s a\eta^2 \equiv 1 \pmod{\mathfrak{p}^\nu} \tag{$*$}$$