Downloaded from rsta.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Small Fractional Parts of Additive Forms

Glyn Harman

Phil. Trans. R. Soc. Lond. A 1993 **345**, 327-338
doi: 10.1098/rsta.1993.0132

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. A* go to:
http://rsta.royalsocietypublishing.org/subscriptions

This journal is © 1993 The Royal Society

Downloaded from rsta.royalsocietypublishing.org

Small fractional parts of additive forms

By Glyn Harman

School of Mathematics, University of Wales College of Cardiff, Senghennydd Road, Cardiff CF2 4AG, U.K.

We show how the methods of Vaughan & Wooley, which have proved fruitful in dealing with Waring's problem, may also be used to investigate the fractional parts of an additive form. Results are obtained which are near to best possible for forms with algebraic coefficients. New results are also obtained in the general case. Extensions are given to make several additive forms simultaneously small. The key ingredients in this work are: mean value theorems for exponential sums, the use of a small common factor for the integer variables, and the large sieve inequality.

1. Introduction

We write $\|x\|$ to denote the least distance from x to an integer. Also, given real numbers $\alpha_1, \ldots, \alpha_s$, and a positive integer k, we put

$$F(\boldsymbol{x}) = \alpha_1 x_1^k + \ldots + \alpha_s x_s^k \quad \text{for} \quad \boldsymbol{x} = (x_1, \ldots, x_s),$$

and we also write $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_s)$. In this paper we shall be concerned with the problem of making $\|F(\boldsymbol{x})\|$ 'small' for $\boldsymbol{x} \in \mathbb{Z}^s, \boldsymbol{x} \neq \boldsymbol{0}$. In fact we will take all the coordinates of x to be non-negative (if k is odd and the coordinates of x can take values of either sign the problem is somewhat different (see ch. 14 of Baker (1986)). Given a positive integer N we write

$$\boldsymbol{N} = \{\boldsymbol{n} \in \mathbb{Z}^s : 0 \leqslant n_j \leqslant N, \max_j n_j \geqslant 1\}.$$

The natural question to ask then is: for what function $a(s,k)$ is it true that

$$\min_{\boldsymbol{n} \in \boldsymbol{N}} \|F(\boldsymbol{n})\| < N^{-a(s,k)}, \tag{1}$$

at least for all large N? For almost all $\boldsymbol{\alpha} \in \mathbb{R}^s$ one can take $a(s,k) = s$, at least for infinitely many N, and this is best possible (see §6, ch.1 of Sprindzuk 1979). There is no difficulty in showing that in general one must have $a(s,k) \leqslant k$, by considering $\alpha_j = \sqrt{2}$ for all j, for example. By Dirichlet's theorem $a(1,1) = 1$, so we may henceforth suppose that $k \geqslant 2$. Cook (1972) obtained $a(s,k) = s2^{1-k}$ for $s \leqslant 2^{k-1}$, while Schlickewei (1979) established $a(s,k) = k - c(s,k)$ with $c(s,k) \to 0$ as $s \to \infty$. Unfortunately $c(s,k)$ tends to zero very slowly (like $(\log s)^{-\frac{1}{2}}$). Much better results are known for the case $k = 2$ (see Baker 1983; Baker & Harman 1982; Heath-Brown 1991) in all of which $c(s,2) = O(1/s)$). In view of the apparently considerable difficulty in improving $c(s,k)$ for $k \geqslant 3$ by Schlickewei's method, it seems worthwhile

Phil. Trans. R. Soc. Lond. A (1993) **345**, 327–338
Printed in Great Britain © 1993 The Royal Society

13-2

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to *Philosophical Transactions: Physical Sciences and Engineering.*
www.jstor.org

Downloaded from rsta.royalsocietypublishing.org

to relax the question somewhat. Here we shall consider what happens for $\boldsymbol{\alpha} \in \mathbb{A}^s$ ($\mathbb{A}$ being the set of real algebraic numbers) and what happens for $\boldsymbol{\alpha}$ in general if one is only interested in infinitely many solutions to (1).

Our method relies heavily on an important auxiliary result from the Hardy–Littlewood circle method, namely

$$\int_0^1 \left| \sum_{\substack{n=1 \\ n \in \mathscr{B}}}^{N} e(yn^k) \right|^R \mathrm{d}y \ll N^{R-k+\epsilon}. \tag{2}$$

Here $\mathscr{B}$ is a relatively dense set of integers, R is a suitably large (in terms of k) even integer, and $e(x) = \exp(2\pi i x)$. Even with Hua's inequality for (2) (see Lemma 2.5 of Vaughan 1981) we would obtain strong results by our present approach. We are able to use the much more powerful results given by Vaughan & Wooley, however, and so prove theorems which tie up quite well with what is known for Waring's problem.

Theorem 1. *Let $k \geqslant 2$ be given, and define $R = R(k)$ by*

k:	2	3	4	5	6	7	8	9	10
R:	4	8	12	18	24	34	42	52	60

For $k \geqslant 11$ let R be the smallest even integer not less than the value for $G(k)$ given by Vaughan & Wooley (this volume) and Wooley (1992). Let $\boldsymbol{\alpha} \in \mathbb{A}^R$ and suppose that $\epsilon > 0$ is given. Then for all $N > N(\boldsymbol{\alpha}, \epsilon)$ there are solutions to (1) with $a(s,k) = k - \epsilon$.

Remarks. By Wooley (1992) for large k we have $R \sim k \log k$. With some extra work R can be taken as an odd integer and this would reduce the values given by one for certain k. Theorem 1 improves the work of Cook (1989) where, for example, $2k^2$ variables yield only $a(s,k) \approx 0.132\,k$. Our result is best possible, apart from the ϵ, as we have already remarked, although one would hope to require fewer variables. We now show that the ϵ may be removed with an additional constraint imposed.

Theorem 2. *Let $k \geqslant 2$ be given, and define $T = T(k) = R(k) + 1$. Let $\boldsymbol{\alpha} \in \mathbb{A}^T$, and suppose that each coordinate of $\boldsymbol{\alpha}$ is a quadratic irrational. Then for all $N > N(\boldsymbol{\alpha})$ there are solutions to*

$$\min_{\boldsymbol{n} \in N} \|F(\boldsymbol{n})\| < K(\alpha) N^{-k}. \tag{3}$$

Here $K(\boldsymbol{\alpha})$ is an effectively computable constant depending on α.

Remarks. The constant $N(\boldsymbol{\alpha})$ might not be effectively computable given our current state of knowledge. If we drop the constraint that each α_j is a quadratic irrational then (3) is true for infinitely many N as we shall explain later.

Theorem 3. *Let $k \geqslant 2$ and $s \geqslant R(k)$ be given, and suppose that $\boldsymbol{\alpha} \in \mathbb{R}^s$. Then, given $\epsilon > 0$, there are infinitely many solutions to (1) with $a(s,k) = k - c(s,k) - \epsilon$, where*

$$c(s,k) = k(1 - (2k-1)^{-1/t}) \quad \text{with} \quad t = [(s - \tfrac{1}{2})/(R-1)] \tag{4}$$

$$= (k(R-1)/s) \log(2k-1) + O(s^{-2}). \tag{5}$$

Here $R = R(k)$ is defined as in Theorem 1.

Remark. We have $c(s,k) \sim C_k/s$ with $C_2 \approx 6.6$, $C_3 \approx 33.8$, $C_4 \approx 85.6$, and $C_k \sim k^2(\log k)^2$ as $k \to \infty$. Better results are known for the case $k = 2$ (see Baker 1983).

Downloaded from rsta.royalsocietypublishing.org

In the case $k=2$ by Theorem 20 B of Schmidt (1977) we know that if $\alpha_1, \dots, \alpha_s$ form a not very well approximable set, in the sense that

$$\prod_{j=1}^{s} \|\alpha_j n\| < n^{-1-\delta} \tag{6}$$

has only finitely many solutions in n for every $\delta > 0$, then one can take $a(s,2) \sim \sqrt{(2s)}$ as $s \to \infty$. Examples of such sets include sets of algebraic numbers where $1, \alpha_1, \dots, \alpha_s$ are linearly independent over $\mathbb{Q}$. Almost all sets of s reals are also not very well approximable. Chapter 3 of Harman (1982) generalizes Schmidt's result to $a(s,k) \sim \sqrt{(ks)}$ as $s \to \infty$. We now improve these results as follows.

Theorem 4. *Let $k \geqslant 2$ and $\epsilon > 0$ be given. Let $\mathscr{A} = \{\alpha_1, \dots, \alpha_s\}$ be a set of reals. Suppose that there are R disjoint subsets $\mathscr{A}_j$ of $\mathscr{A}$, each containing t members, and such that each $\mathscr{A}_j$ is a not very well approximable set. Then, for $N > N(\epsilon, \boldsymbol{\alpha})$, there are solutions to (1) with $a(s,k) = tk - \epsilon$. In particular, if $\mathscr{A}$ is a not very well approximable set, then one obtains $a(s,k) = k[s/R] - \epsilon$.*

Our final result concerns the generalization of Theorem 1 to simultaneous approximation (see Baker & Harman (1981) for what is known without the restriction $\boldsymbol{\alpha} \in \mathbb{A}^s$). For a vector $\boldsymbol{b} = (b_1, \dots, b_h) \in \mathbb{R}^h$ we write $\|\boldsymbol{b}\|$ to denote $\max_j \|b_j\|$. Given hs real numbers α_{ij}, $i = 1, \dots, h$; $j = 1, \dots, s$, we put

$$F_i(\boldsymbol{x}) = \sum_{j=1}^{s} \alpha_{ij} x_j^k,$$

and write $\boldsymbol{F}(\boldsymbol{x}) = (F_1(\boldsymbol{x}), \dots, F_h(\boldsymbol{x}))$. We then have the following result.

Theorem 5. *Let $k, h \geqslant 2$ be given, and let $\alpha_{ij} \in \mathbb{A}$ be given for $i = 1, \dots, h$; $j = 1, \dots, s$ where*

$$s = R(\tfrac{1}{2}h(h+1) - 1). \tag{7}$$

Then, given $\epsilon > 0$, for $N > N(\epsilon, (\alpha_{ij}))$ we have

$$\min_{\boldsymbol{n} \in N} \|\boldsymbol{F}(\boldsymbol{n})\| < N^{-k/h+\epsilon}.$$

Remark. The exponent $-k/h$ is best possible as can be seen by taking $\alpha_{ij} = \alpha_i$ where $1, \alpha_1, \dots, \alpha_h$ are linearly independent over $\mathbb{Q}$.

2. Preparatory results

We begin with a familiar type of lemma linking Diophantine inequalities to bounds for exponential sums. The result follows by modifying the work of ch. 2 in Baker (1986). These modifications appear explicitly in ch. 1 of Harman (1982).

Lemma 1. *Let y_n be a sequence of real numbers, and a_n a non-negative real sequence. Then for all N, $L \geqslant 1$, if*

$$\sum_{u=1}^{L} \left| \sum_{n=1}^{N} a_n e(u y_n) \right| < \frac{1}{6} \sum_{n=1}^{N} a_n \tag{9}$$

we have

$$\min_{1 \leqslant n \leqslant N} \|y_n\| < L^{-1}.$$

Downloaded from rsta.royalsocietypublishing.org

We shall be estimating the exponential sums which arise *via* the large-sieve inequality. The following is a multidimensional version of the inequality which occurs as Lemma 5.3 in Vaughan (1981).

Lemma 2. *Let $\mathscr{G}$ be a finite set of distinct points in $\mathbb{R}^h$ where $h \geqslant 1$. Write*

$$\delta = \min_{\substack{\boldsymbol{g} \neq \boldsymbol{h} \\ \boldsymbol{g}, \boldsymbol{h} \in \mathscr{G}}} \|\boldsymbol{g} - \boldsymbol{h}\|.$$

Let $a(\boldsymbol{n})$ be a sequence of complex numbers for $\boldsymbol{n} \in \mathbb{Z}^h$. Then

$$\sum_{g \in \mathscr{G}} \left| \sum_{n \in N} a(\boldsymbol{n})\, e(\boldsymbol{g}\boldsymbol{n}) \right|^2 \ll (\delta^{-1} + N)^h \sum_{n \in N} |a(\boldsymbol{n})|^2. \tag{10}$$

Lemma 3. *Let $k \geqslant 2$ and $\eta > 0$ be given. Define $R(k)$ as in Theorem 1. Then, given P, there exists a set $\mathscr{B}$ such that*

$$\mathscr{B} \subseteq \mathbb{Z} \cap [1, P], \quad |\mathscr{B}| \gg P, \tag{11}$$

and

$$\int_0^1 \left| \sum_{n \in \mathscr{B}} e(yn^k) \right|^R \mathrm{d}y \ll P^{R-k+\eta}. \tag{12}$$

Remarks. The implied constants in (11) and (12) depend at most on η. In actual fact, the η is only required for certain k.

Proof. For $k = 2$ or 3 the bound (12) follows from Hua's inequality with $\mathscr{B} = [1, P] \cap \mathbb{Z}$ (see Lemma 2.5 of Vaughan 1981). For $k = 4$ the result follows from §4 of Vaughan (1989) with

$$\mathscr{B} = \{n \in \mathbb{Z} : 1 \leqslant n \leqslant P, p \,|\, n \Rightarrow p < N^{\delta}\}$$

for some $\delta = \delta(\eta)$ (from the work later in Vaughan's paper it is clear that the η can be dispensed with). For $5 \leqslant k \leqslant 15$ see (Vaughan & Wooley, this volume), which gives the result with the same type of set $\mathscr{B}$ (only in the case $k = 6$ is the η required). For $k \geqslant 16$ see Wooley (1992). In this last paper integrals of the form

$$\int_0^1 \left| \sum_{n \in \mathscr{B}} e(yn^k) \right|^u \left| \sum_{n \in \mathscr{A}} e(yn^k) \right|^v \mathrm{d}y$$

are estimated, but the set $\mathscr{A}$ (whose members have a large prime factor) can be replaced by $\mathscr{B}$ if Theorem 1.8 is appealed to in place of Theorem 1.5 in Vaughan (1989).

Lemma 4. *Let $k \geqslant 2$ be given, and write $H(k) = R(k) + 1$. Then, for every set $\{a_1, \ldots, a_H\}$ of non-zero integers, there are integers s, b ($s \neq 0$) such that for all sufficiently large r the equation*

$$rs + b = a_1 n_1^k + \ldots + a_H n_H^k \tag{13}$$

can be solved in integers $n_1, \ldots, n_H$ with

$$1 \leqslant n_j \leqslant 2\,|rs + b|^{1/k}.$$

Proof. This follows using the Hardy–Littlewood circle method as in ch. 7 of Davenport (1962) with Lemma 3. Of course, for $k \geqslant 3$ one can reduce $H(k)$ to $G(k)$, and for $k = 2$ one can use the theory of quadratic forms (see ch. 8 of Jones (1950)) but the result as stated suffices for our applications here.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

3. Proof of Theorems 1 to 4

Proof of Theorem 1. Put $\eta = \epsilon/(10k)$, $L = N^{k-\epsilon}$, $M = N^{3\eta}$, $P = N/M$. We will show that for all large N there are solutions to

$$\|m^k(n_1^k\alpha_1 + \ldots + n_R^k\alpha_R)\| < L^{-1}$$

with

$$m \leqslant M, \quad 1 \leqslant n_j \leqslant P, \quad n_j \in \mathscr{B} \quad (\mathscr{B} \text{ as given by Lemma 3}).$$

The result is trivial if any α_j is rational, so we henceforth suppose each α_j to be irrational. By Lemma 1 it suffices to show that

$$\sum_{u \leqslant L}\ \sum_{m \leqslant M}\ \sum_{j \leqslant R} \left|\sum_{n_j \in \mathscr{B}} e(um^k n_j^k \alpha_j)\right| < \tfrac{1}{6}[M]\,|\mathscr{B}|^R.$$

This will follow from Hölder's inequality for all large N once we establish that, for each j,

$$\sum_{u \leqslant L}\ \sum_{m \leqslant M} \left|\sum_{n \in \mathscr{B}} e(um^k\alpha_j n^k)\right|^R = o(P^R M). \tag{14}$$

We write $a(n)$ for the number of solutions to

$$n = r_1^k + \ldots + r_s^k, \quad r_j \in \mathscr{B}, \quad s = \tfrac{1}{2}R.$$

We note that the number of solutions to $v = um^k$ for fixed v is $\ll P^\eta$. The left hand side of (14) is thus

$$\ll P^\eta \sum_{u \leqslant LM^k} \left|\sum_{n \leqslant P^k} a(n)\, e(un\alpha_j)\right|^2.$$

By Roth's Theorem (Roth 1955, or see ch. 5 of Schmidt 1980) since α_j is an algebraic irrational we have

$$\|u\alpha_j\| > (LM^k)^{-1-\eta}$$

for $0 < u \leqslant LM^k$ provided that LM^k is sufficiently large. Hence, by Lemma 2 with $h = 1$, we have

$$\sum_{u \leqslant LM^k} \left|\sum_{n \leqslant P^k} a(n)\, e(un\alpha_j)\right|^2 \ll ((LM^k)^{1+\eta} + P^k) \sum_{n \leqslant P^k} a(n)^2 \ll P^k P^{R-k+\eta}$$

by Lemma 3. This establishes (14) and so completes the proof.

The reader should note the vital role played by the variable m above. Without it the application of the large-sieve inequality would have yielded an unsuitable bound. This use of a small common factor is characteristic of most of our proofs here. The reader should note that Theorem 1 remains valid when the variables are restricted in certain ways. For example, we could have made the variables square-free, or sums of two squares. We could not have taken the variables to be prime by working as above. We shall show elsewhere, however, that weaker results (in terms of the number of variables required to obtain the best possible exponent) can be obtained even in this case. In order to take R as an odd integer we need to work directly from the Sobolev–Gallagher inequality in place of the large sieve, and modify the work of Vaughan & Wooley to obtain bounds of the type given by Lemma 3 for

$$\int_0^1 \left|\sum_{n \in \mathscr{B}} (n/P)^k\, e(yn^k)\right|^R \mathrm{d}y.$$

Downloaded from rsta.royalsocietypublishing.org

Proof of Theorem 2. Suppose first that $1, \alpha_j, \alpha_k$ are linearly dependent over $\mathbb{Q}$ for any j, k with $1 \leqslant j, k \leqslant T$. Then there are integers A_j, B_j, C_j, such that

$$B_j \alpha_j = C_j + A_j \alpha_1, \quad j = 2, \ldots, H, \quad B_j > 0, \quad A_j \neq 0.$$

We therefore need only prove that

$$\|\alpha_1(n_1^k + d_2 n_2^k + \ldots + d_H n_H^k)\| < KN^{-k}$$

with $1 \leqslant \max n_j \leqslant N/(\max\ B_j)\,(= X \text{ say})$ and $d_j = A_j B_j^{k-1}$, since then

$$\|\alpha_1 n_1^k + \alpha_2(B_2 n_2)^k + \ldots + \alpha_H(B_H n_H)^k\| < KN^{-k}.$$

Now let b, s be the integers given by Lemma 4 for the set $\{1, d_2, \ldots, d_H\}$ and let a/q be the convergent to α_1 with largest denominator less than $X^k/|s|$. Then we can find an integer $d \equiv b \pmod s$ with $|d| \leqslant |sq|$,

$$\|\mathrm{d}\alpha_1\| < |\mathrm{d}s|/qX^k + b/q,$$

and
$$d = n_1^k + d_2 n_2^k + \ldots + d_H n_H^k$$

has a solution with $1 \leqslant \max n_j < 2\,|d|^{1/k} \leqslant X$. Since

$$q > c(\alpha_1) X^k \tag{15}$$

(because α_1 is quadratic) we have

$$\|\alpha_1 n_1^k + \alpha_2(B_2 n_2)^k + \ldots + \alpha_H(B_H n_H)^k\| = \|d\alpha_1\| < s^2/X^k + b/c(\alpha_1) X^k < KN^{-k}$$

as desired. This part of the argument fails if α_1 is not quadratic, since we would then only know $q > c(\alpha_1, \eta) X^{k(1-\eta)}$ in place of (15). We may obtain infinitely many solutions, however, by only considering those N for which $X^k \geqslant qs \geqslant (X-1)^k$.

If we are not in the above situation, then, without loss of generality, $\{1, \alpha_1, \alpha_2\}$ is a linearly independent set over $\mathbb{Q}$. Hence, by Theorem 1B of ch. 6 in Schmidt (1980), $\{\alpha_1, \alpha_2\}$ is a not very well approximable set. Thus

$$\min_{1 \leqslant v \leqslant V} \max\,(\|\alpha_1 v\|, \|\alpha_2 v\|) > V^{-\frac{2}{3}} \tag{16}$$

for all sufficiently large V. We will establish Theorem 2 by finding solutions to

$$\|m^k(\alpha_1 n_1^k + \alpha_2 n_2^k) + \alpha_3 n_3^k + \ldots + \alpha_T n_T^k\| < L^{-1}$$

with $L = N^k$, $1 \leqslant m \leqslant N^{1/10}$, $n_1, n_2 \in \mathscr{B}_1$ (the set from Lemma 3 with $P = N^{9/10}$, $\eta = 1/(100Rk)$), $n_j \in \mathscr{B}_2$ for $j \geqslant 3$ ($\mathscr{B}_2$ being the set from Lemma 3 with η as $\mathscr{B}_1$, but $P = N$). These restrictions will henceforth be denoted by *. In view of Lemma 1, we therefore need to demonstrate that

$$\sum_{u \leqslant L} \left| \sum^{*} e(u(m^k(\alpha_1 n_1^k + \alpha_2 n_2^k) + \alpha_3 n_3^k + \ldots + \alpha_T n_T^k)) \right| = o(N^{R+9/10}). \tag{17}$$

By (16) and Lemma 2 with $h = 2$ we have

$$\sum_{\substack{u \leqslant L \\ m \leqslant N^{1/10}}} \left| \sum_{n_j \in \mathscr{B}_1} e(um^k(\alpha_1 n_1^k + \alpha_2 n_2^k)) \right|^R \ll (N^{9k/10} + N^{11k/15})^2 N^{9(R-k)/5 + 3\eta}$$

$$\ll N^{9R/5 + 3\eta}. \tag{18}$$

Downloaded from rsta.royalsocietypublishing.org

Also, by Lemma 2 with $h = 1$, for $j \geqslant 3$ we have

$$\sum_{u \leqslant L} \Big| \sum_{n_j \in \mathscr{B}_2} e(u\alpha_j n_j^k) \Big|^R \ll (N^k + N^{k(1+\eta)}) N^{R-k+\eta}$$
$$\ll N^{R+(k+1)\eta}.$$

Thus, by Hölder's inequality, the left-hand side of (17) is

$$\ll (N^{1/10} N^{R+(k+1)\eta})^{1-1/R} (N^{9R/5+3\eta})^{1/R}$$
$$= o(N^{R+9/10})$$

as desired, by our choice of η. This establishes (17) and so completes the proof.

Proof of Theorem 3. Let $t = [(s-\frac{1}{2})/(R-1)]$ and define θ by

$$(1+\theta)^t = 2k-1 \quad (\text{so } \theta = (2k-1)^{1/t}-1).$$

Now let Y be a large number. By Dirichlet's theorem, for each α_j there is a q_j with

$$\|q_j \alpha_j\| < Y^{1-2k} \quad \text{and} \quad 1 \leqslant q_j \leqslant Y^{2k-1}. \tag{19}$$

We suppose that the smallest q_j satisfying (19) is chosen. If $q_j \leqslant Y$ for some j then $\|\alpha_j q_j^k\| < Y^{-k}$ so we get a solution of the required type (indeed a much better one!) with only one variable non-zero. If there is a solution of this type for infinitely many Y this proves the result (if any α_j is rational the result is trivial, otherwise $q_j \to \infty$ as $Y \to \infty$).

Now suppose that $q_j > Y$ for each j. By the pigeon-hole principle there must be R or the q_j (say with $j \in \mathscr{A}$) in a range of the form $[H, H^{1+\theta}]$ with $Y \leqslant H$, $H^{1+\theta} \leqslant Y^{2k-1}$. Define P, L, M, N by

$$P^k = H^{1+\theta}, \quad LM^k = H, \quad N = MP, \quad M = P^{3\eta}, \quad \eta = \epsilon/10k.$$

It then follows, as in the proof of Theorem 1, for each $j \in \mathscr{A}$, that we have

$$\sum_{u \leqslant L} \sum_{m \leqslant M} \Big| \sum_{n \in \mathscr{B}} e(\alpha_j u m^k n^k) \Big|^R \ll P^{R+2\eta},$$

since

$$\|\alpha_j u\| > (2H^{1+\theta})^{-1} \quad \text{for} \quad u \leqslant H,$$

by the definition of the q_j. Hence, if we put the variables $n_j = 0$ for $j \notin \mathscr{A}$ we obtain a solution to

$$\min_{\boldsymbol{n} \in N} \|F(\boldsymbol{n})\| < L^{-1}.$$

Since $L > N^{k(1+\theta)^{-1}-\epsilon}$ and $k(1+\theta)^{-1} = k - c(s,k)$ this completes the proof.

When s is not very large an improvement in the exponent can be obtained by defining θ by

$$(1+\theta)^t = k(1+(1+\theta)^{-1})-1,$$

and choosing q_j with

$$\|q_j \alpha_j\| < Y^{-(1+\theta)^t} \quad \text{and} \quad 1 \leqslant q_j \leqslant Y^{(1+\theta)^t}.$$

For large s this only improves the $O(1/s^2)$ term in $c(s,k)$. However, for $s = R$ this produces a notable improvement. In this case $P = H = Y$ above. Professor R. C. Baker has pointed out that this leads to the following result which is uniform in N and improves the work of Cook (1972) for $k \geqslant 6$.

Downloaded from rsta.royalsocietypublishing.org

Theorem 6. *Let $k \geqslant 6, N > C(k, \epsilon)$, and $R = R(k)$ be as given in Theorem 1. Let $\alpha_1, \ldots, \alpha_R$ be real numbers. Then there are solutions to (1) with $a(R, k) = 1 - \epsilon$.*

Proof of Theorem 4. Clearly we need only prove that

$$\sum_{u \leqslant L} \sum_{m \leqslant M} \left| \sum_{n_j \in \mathscr{B}} e(um^k(\alpha_1 n_1^k + \ldots + \alpha_t n_t^k)\right|^R = o(P^{tR} M), \tag{20}$$

where $\alpha_1, \ldots, \alpha_t$ are not very well approximable, $L = N^{tR-\epsilon}$, $M = N^{3t\eta}$, $\eta = \epsilon/(10kt)$, $P = N/M$. We combine um^k into a single variable as before with the loss of a P^η factor. We can then bound the left-hand side of (20) with Lemmas 2 and 3. The estimate we obtain is then

$$\ll P^\eta (P^k + P^k)^t \, (P^{R-k+\eta})^t = o(P^{tR-\eta} M)$$

since, by (6), there are no solutions to

$$\max_j \|\alpha_j r\| < P^{-k}$$

if $r < P^{tk-\eta}$ and P is sufficiently large.

4. A simultaneous approximation lemma

In this section we shall establish an analogue of Lemma 1 for simultaneous approximation. Lemma 7.4 of Baker (1986) would have sufficed for our application (with the addition of weights a_n) but it has an extra N^η factor on the summation length. One could add weights to the results of Cochrane (1988) and so obtain an analogue of Lemma 1. Here we present an alternative approach and establish a slightly more general result (in that we speak of an arbitrary lattice Λ), and find that the calculation of the constants is very straightforward. The constants which arise from Cochrane (1988) are slightly better, however, We write $|\boldsymbol{u}|$ for max $|u_j|$ where $\boldsymbol{u} = (u_1, \ldots, u_h) \in \mathbb{R}^h$. We also write $\mathscr{B}_h$ for the hypercube $[-1, 1]^h \subset \mathbb{R}^h$. We then have the following result.

Lemma 5. *Let $h > 1$ and let Λ be a lattice in $\mathbb{R}^h$. Suppose that $\boldsymbol{\alpha}_n \in \mathbb{R}^h$ and a_n is a sequence of non-negative reals. Then the inequality*

$$\sum_{\substack{\boldsymbol{u} \in \Pi \\ \boldsymbol{u} \neq 0 \\ |\boldsymbol{u}| < h}} \left| \sum_{n \leqslant N} a_n \, e(\boldsymbol{u}\boldsymbol{\alpha}_n)\right| < \frac{1}{4h^2 - 1} \sum_{n \leqslant N} a_n \tag{21}$$

implies that a solution exists to

$$\boldsymbol{\alpha}_n \in \mathscr{B}_h + \boldsymbol{y} \quad \text{where} \quad \boldsymbol{y} \in \Lambda \quad \text{and} \quad 1 \leqslant n \leqslant N. \tag{22}$$

Here Π denotes the polar lattice to Λ.

Corollary. *Let $\boldsymbol{\alpha}_n$ be a sequence in $\mathbb{R}^h$ and a_n a sequence of non-negative reals. Then, for all $N, L \geqslant 1$, the inequality*

$$\sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^h \\ \boldsymbol{u} \neq 0 \\ |\boldsymbol{u}| \leqslant L}} \left| \sum_{n \leqslant N} a_n \, e(\boldsymbol{u}\boldsymbol{\alpha}_n)\right| < \frac{1}{4h^2 - 1} \sum_{n \leqslant N} a_n \tag{23}$$

Downloaded from rsta.royalsocietypublishing.org

implies that a solution exists to

$$\|\boldsymbol{\alpha}_n\| < h/L \quad \textit{with} \quad 1 \leqslant n \leqslant N.$$

The corollary follows by putting $\Lambda = (L\mathbb{Z}/h)^h$ (so $\Pi = (h\mathbb{Z}/L)^h$), while replacing $\boldsymbol{\alpha}_n$ by $L\boldsymbol{\alpha}_n/h$.

We shall prove Lemma 5 by constructing a lower bound approximation to the characteristic function of $\mathscr{B}_h$. We write $\chi^0(x)$ for the characteristic function of the interval $[-1, 1]$. We note that by Lemma 2.5 of Baker (1986), for any $w > 0$, there are continuous functions $\chi^+(x)$ and $\chi^-(x)$ belonging to $L^1(\mathbb{R})$ such that

$$\chi^-(x) \leqslant \chi^0(x) \leqslant \chi^+(x), \quad \hat{\chi}^+(t) = \hat{\chi}^-(t) = 0 \quad \text{for} \quad |t| > w, \tag{24}$$

and
$$\int_{-\infty}^{\infty} (\chi^0(x) - \chi^-(x))\, \mathrm{d}x = \int_{-\infty}^{\infty} (\chi^+(x) - \chi^0(x))\, \mathrm{d}x = w^{-1}. \tag{25}$$

Here
$$\hat{f}(t) = \int_{-\infty}^{\infty} f(x)\, e(-xt)\, \mathrm{d}x.$$

The fact that $\chi^-(x)$ can be negative causes problems in the present context, since we cannot form a product of such functions and still obtain a lower bound. We circumvent this difficulty by employing the following lemma.

Lemma 6. *Let $\chi_{\mathscr{B}}(\boldsymbol{x})$ denote the characteristic function of $\mathscr{B}_h$. Then, for any choice of w above, we have*

$$\chi_{\mathscr{B}}(\boldsymbol{x}) \geqslant \sum_{k=1}^{h} \prod_{j=1}^{h} \chi^{\sigma(i,j)}(x_j) - (h-1) \prod_{j=1}^{h} \chi^+(x_j), \tag{26}$$

where
$$\sigma(j,k) = \begin{cases} - & \text{if} \quad j = k, \\ + & \text{if} \quad j \neq k. \end{cases}$$

Proof. This lemma is essentially proved in a different context by J. Brüdern and E. Fouvrey in work currently being prepared on 'Le crible vectoriel', and I thank J. Brüdern in particular for drawing my attention to this result. This has considerably improved the present section of this paper from its original version. We first note that

$$\chi_{\mathscr{B}}(\boldsymbol{x}) \geqslant \sum_{k=1}^{h} \prod_{j=1}^{h} \chi^{\tau(j,k)}(x_j) - (h-1) \prod_{j=1}^{h} \chi^+(x_j), \tag{27}$$

where
$$\tau(j,k) = \begin{cases} 0 & \text{if} \quad j = k, \\ + & \text{if} \quad j \neq k. \end{cases}$$

To see this, first note that the right-hand side of (27) is clearly non-positive if one of $\chi^0(x_j)$ is zero. If every $\chi^0(x_j)$ is 1 then the right-hand side can be shown not to exceed 1 by induction on h. Clearly (26) follows from (27).

Proof of Lemma 5. Let $g(\boldsymbol{x})$ denote the right-hand side of (26) with $w = h$. Then, if

$$\hat{g}(\boldsymbol{t}) = \int_{\mathbb{R}^h} g(\boldsymbol{x})\, e(-\boldsymbol{xt})\, \mathrm{d}t, \tag{28}$$

we have
$$\hat{g}(\boldsymbol{t}) = 0 \quad \text{for} \quad |\boldsymbol{t}| \geqslant h. \tag{29}$$

Also
$$\begin{aligned} \hat{g}(\mathbf{0}) &= h(2+1/h)^{h-1}(2-1/h) - (h-1)(2+1/h)^h \\ &= h^{-1}(2+1/h)^{h-1}. \end{aligned}$$

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

Downloaded from rsta.royalsocietypublishing.org

Now let
$$T(\boldsymbol{x}) = \sum_{\boldsymbol{y}\in\Lambda} g(\boldsymbol{x}+\boldsymbol{y}),$$
and write
$$c_{\boldsymbol{p}} = (\det \Lambda)^{-1}\int_{\mathscr{F}} T(\boldsymbol{x})\, e(-\boldsymbol{p}\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \quad \text{for} \quad \boldsymbol{p}\in\Pi,$$
where $\mathscr{F}$ denotes a fundamental paralleliped of Λ. Then, by (28) and (29), $c_{\boldsymbol{p}} = 0$ for $|\boldsymbol{p}| \geqslant h$. Hence
$$\sum_{\boldsymbol{p}\in\Pi} |c_{\boldsymbol{p}}|$$
converges, so
$$T(\boldsymbol{x}) = \sum_{\boldsymbol{p}\in\Pi} c_{\boldsymbol{p}}\, e(\boldsymbol{p}\boldsymbol{x})$$
(see Lemma 7.2 of Baker (1986) for example).

Now suppose that no solution exists to $\boldsymbol{\alpha}_n \in \mathscr{B}_h + \boldsymbol{y}$ for $\boldsymbol{y}\in\Lambda$. Then, by (26),
$$\sum_{n\leqslant N} a_n\, T(\boldsymbol{\alpha}_n) \leqslant 0.$$
Hence
$$\sum_{\substack{\boldsymbol{p}\in\Pi\\ \boldsymbol{p}\neq 0}} |c_{\boldsymbol{p}}| \left|\sum_{n\leqslant N} a_n\, e(\boldsymbol{p}\boldsymbol{\alpha}_n)\right| \geqslant c_0 \sum_{n\leqslant N} a_n. \tag{30}$$
Since the transform of any product $\Pi\chi^{\sigma(j,k)}(x_j)$ is, in modulus, bounded by $(2+1/h)^h$ (using (25) h times), we obtain, when $c_{\boldsymbol{p}} \neq 0$,
$$\frac{|c_0|}{|c_{\boldsymbol{p}}|} \geqslant \frac{h^{-1}(2+1/h)^{h-1}}{(2h-1)\,(2+1/h)^h} = \frac{1}{4h^2-1}.$$
This completes the proof of (21).

5. Proof of Theorem 5

Given α_{ij} as in the statement of the theorem, write $\boldsymbol{\alpha}_j = (\alpha_{1j}, \ldots, \alpha_{hj})$. We need to be careful of linear dependencies between the coordinates of $\boldsymbol{\alpha}_j$. For example, if the set $\{1, \alpha_{ij}, \alpha_{uj}\}$ is linearly dependent over $\mathbb{Q}$ for many j, then we are essentially in an $(h-1)$-dimensional situation. To facilitate the discussion we write
$$\Lambda_j = \{\boldsymbol{n}\in\mathbb{Z}^h : \boldsymbol{n}\boldsymbol{\alpha}_j \in \mathbb{Q}\}. \tag{31}$$
Clearly each Λ_j is a lattice, and $\Lambda_j = \{\boldsymbol{0}\}$ if $1, \alpha_{1j}, \ldots, \alpha_{hj}$ form a linearly independent set over $\mathbb{Q}$.

Lemma 7. *Let α_j be as above, and let $\eta > 0$ be given. Suppose that $L, M, P \geqslant 1$ satisfy $(LM^k)^{h+\eta} \leqslant P^k$. Suppose also that for some set $\mathscr{A}$ of positive integers we have*
$$\bigcap_{j\in\mathscr{A}} \Lambda_j = \{\boldsymbol{0}\}. \tag{32}$$
Then
$$\sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^h\\ 0<|\boldsymbol{u}|\leqslant L}} \sum_{m\leqslant M} \left|\sum_{\substack{n_j\in\mathscr{B}\\ j\in\mathscr{A}}} e\left(\sum_{j\in\mathscr{A}} n_j^k\, m^k\, \boldsymbol{u}\boldsymbol{\alpha}_j\right)\right|^R \ll P^{|\mathscr{A}|(R+\eta)}. \tag{33}$$

Proof. Since the α_{ij} are algebraic, for LM^k sufficiently large, either
$$\|\boldsymbol{\alpha}_j \boldsymbol{u}\| > (LM^k)^{-h-\eta} \quad \text{or} \quad \|\boldsymbol{\alpha}_j \boldsymbol{u}\| = 0$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

for $|\boldsymbol{u}| \leqslant LM^k$ (this follows from Theorem 1 D of ch. 6 in Schmidt (1980)). Since (32) holds, however, we cannot have $\|\boldsymbol{\alpha}_j \mathbf{u}\| = 0$ for every $j \in \mathscr{A}$ for any one $\boldsymbol{u}$. We may thus bound the left hand side of (33) by Lemma 2 with $h = |\mathscr{A}|$, $\delta = P^{-k}$, also using Lemma 3 as before. The required bound then follows.

Proof of Theorem 5. Suppose that the result has been proved in $h-1$ dimensions. Let

$$T = \begin{cases} R & \text{if} \quad h = 2, \\ (\tfrac{1}{2}h(h-1)-1)R & \text{if} \quad h \geqslant 2. \end{cases}$$

So T is the number of variables required to prove the result in $h-1$ dimensions. Suppose first of all that there is a set $\mathscr{A} \subset \{1, \ldots, s\}$ with $|\mathscr{A}| \geqslant T$ and

$$\bigcap_{j \in \mathscr{A}} \Lambda_j \neq \{\mathbf{0}\}.$$

Without loss of generality $\mathscr{A} = \{1, \ldots, T\}$. Then there is $\boldsymbol{D} \in \mathbb{Z}^h$ with $\boldsymbol{D} \neq \mathbf{0}$ and $\boldsymbol{D\alpha}_j \in \mathbb{Z}$ for $j = 1, \ldots, T$. Without loss of generality the hth coordinate of $\boldsymbol{D}$, say D_h, is positive. Let $\boldsymbol{\alpha}_j'$ and $\boldsymbol{D}'$ be points of $\mathbb{R}^{h-1}$ consisting of the first $h-1$ coordinates of $\boldsymbol{\alpha}_j$ and $\boldsymbol{D}$ respectively. By the inductive hypothesis we can find $n_1, \ldots, n_T$ with $1 \leqslant \max n_j \leqslant N/D_h$ and

$$\left\| \sum_{j \leqslant T} n_j^k \boldsymbol{\alpha}_j' \right\| < (N/D_h)^{-k/(h-1)+\epsilon}$$

for all large N. Here we can take $\epsilon = k(2h(h-1))^{-1}$. Now

$$\begin{aligned} \left\| \sum_{j \leqslant T} \alpha_{hj} D_h^k n_j^k \right\| &= \left\| D_h^{k-1} \left(\sum_{j \leqslant T} n_j^k (D_h \alpha_{hj}) \right) \right\| \\ &= \left\| D_h^{k-1} \sum_{j \leqslant T} n_j^k (\boldsymbol{D\alpha}_j - \boldsymbol{D}'\boldsymbol{\alpha}_j') \right\| \\ &\leqslant D_h^{k-1} \sum_{j=1}^{h-1} |D_j| \left\| \sum_{j \leqslant T} n_j^k \boldsymbol{\alpha}_j' \right\|. \end{aligned}$$

Hence

$$\begin{aligned} \left\| \sum_{j \leqslant T} (D_h n_j)^k \boldsymbol{\alpha}_j \right\| &\leqslant D_h^{k-1} \sum_{j \leqslant h} |D_j| (N/D_h)^{-k/(h-1)+\epsilon} \\ &< N^{-k/h} \end{aligned}$$

for all large N. This proves the result in this case.

Now we suppose that (32) holds for all subsets of $\{1, \ldots, s\}$ having at least T members. If $\mathscr{A}$ is such a subset then there must be a subset $\mathscr{D}$ with $|\mathscr{D}| \leqslant h$ and

$$\bigcap_{j \in \mathscr{D}} \Lambda_j = \{0\}.$$

Since $s = T + hR$ for $h \geqslant 3$, we can split $\{1, \ldots, s\}$ into R subsets $\mathscr{D}_v$ with this property, plus some remainder. In the case $h = 2$ it is evident that $1, \ldots, 2R$ can be split into R such sets (here, of course, each $\mathscr{D}_v$ has one or two members). The proof is completed by an appeal to Lemmas 5 and 7.

I thank Professor R. C. Vaughan and Professor T. D. Wooley for making available their results before publication. I also thank Professor R. C. Baker for pointing out Theorem 6.

Downloaded from rsta.royalsocietypublishing.org

References

Baker, R. C. 1983 Small solutions of congruences. *Mathematika* **20**, 164–188.

Baker, R. C. 1986 *Diophantine inequalities*. Oxford: Clarendon Press.

Baker, R. C. & Harman, G. 1981 Small fractional parts of quadratic and additive forms. *Math. Proc. Camb. phil. Soc.* **90**, 5–12.

Baker, R. C. & Harman, G. 1982 Small fractional parts of quadratic forms. *Proc. Edinb. math. Soc.* **25**, 269–277.

Cochrane, T. 1988 Trigonometric approximation and uniform distribution modulo one. *Proc. Am. math. Soc.* **103**, 695–702.

Cook, R. J. 1972 The fractional parts of an additive form. *Math. Proc. Camb. phil. Soc.* **72**, 209–212.

Cook, R. J. 1989 Small fractional parts of additive forms. *Proc. Indian Acad. Sci.* (*Math. Sci.*) **99**, 147–153.

Davenport, H. 1962 *Analytic methods for Diophantine equations and inequalities*. Ann Arbor Publishers.

Harman, G. 1982 Diophantine approximation and prime numbers. Ph.D. thesis, University of London.

Heath-Brown, D. R. 1991 Small solutions of quadratic congruences. II. *Mathematika* **38**, 264–284.

Jones, B. W. 1950 *The arithmetic theory of quadratic forms*. New York: Wiley.

Roth, K. F. 1955 Rational approximations to algebraic numbers. *Mathematika* **2**, 1–20.

Schlickewei, H. P. 1979 On indefinite diagonal forms in many variables. *J. reine angew. Math.* **307/308**, 279–294.

Schmidt, W. M. 1977 *Small fractional parts of polynomials*. Regional conference series 32. Providence: AMS.

Schmidt, W. M. 1980 *Diophantine approximation*. Lecture Notes in Mathematics 785. Berlin: Springer.

Sprindzuk, V. G. 1979 *Metric theory of Diophantine approximations* (transl. R. A. Silverman). New York: Winston/Wiley.

Vaughan, R. C. 1981 *The Hardy–Littlewood method*. Cambridge University Press.

Vaughan, R. C. 1989 A new iterative method in Waring's problem. *Acta Mathematica* **162**, 1–71.

Wooley, T. D. 1992 Large improvements in Waring's problem. *Ann. Math.* **135**, 131–164.