ON THE SPATIAL DISTRIBUTION OF SOLUTIONS
OF DECOMPOSABLE FORM EQUATIONS

G. EVEREST, I. GAÁL, K. GYŐRY, AND C. RÖTTGER

ABSTRACT. We study the distribution in space of the integral solutions to an integral decomposable form equation, by considering the images of these solutions under central projection onto a unit ball. If we think of the solutions as stars in the night sky, we ask what constellations are visible from the earth (the unit ball). Answers are given for a large class of examples which are then illustrated using the software packages KANT and Maple. These pictures highlight the accuracy of our predictions and arouse interest in cases not covered by our results. Within the range of applicability of our results lie solutions to norm form equations and units in abelian group rings. Thus our theory has a lot to say about where these interesting objects can be found and what they look like.

0. INTRODUCTION

Let $F(\underline{x}) = F(x_1, \ldots, x_n) \in \mathbf{Z}[x_1, \ldots, x_n]$ denote a *decomposable form.* This is a homogeneous polynomial with coefficients in $\mathbf{Z}$ which factorises over $\mathbf{C}$ as a product of linear forms. It is known that there are $q \in \mathbf{Q}^*$, finite extension fields $M_1, \ldots, M_t$ of $\mathbf{Q}$ and linear forms $\phi_i(\underline{x})$ with coefficients in $M_i, i = 1, \ldots, t$, such that

$$(0.1) \qquad F(\underline{x}) = q \prod_{i=1}^{t} N_{M_i|\mathbf{Q}}(\phi_i(\underline{x})).$$

In (0.1), $N_{M_i|\mathbf{Q}} : M_i \to \mathbf{Q}, i = 1, \ldots, t$, denotes the field norm. Given a nonzero $a \in \mathbf{Z}$, the *decomposable form equation*

$$(0.2) \qquad F(\underline{x}) = a, \quad \underline{x} \in \mathbf{Z}^n,$$

is a very general equation with many important examples.

Examples of decomposable form equations.

(1) If $d > 0$ is a nonsquare integer, then *Pell's Equation*

$$x_1^2 - dx_2^2 = 1$$

is a decomposable form equation, with $t = 1$ and $M_1 = \mathbf{Q}(\sqrt{d})$.

(2) A generalisation of Pell's equation, also with $t = 1$, is the *norm form equation*. Here $q = 1, \phi_1(\underline{x}) = \sum_i a_i x_i$ and the a_i lie in the ring of integers of M_1. See [Sc1], [Sc2] and [Sc3] for background to this equation. Schmidt made some fundamental breakthroughs in the study of the norm form equation, using powerful techniques from diophantine approximation. In this case, when $a = \pm 1$ and the a_i form a **Z**-basis for the ring of integers, the solutions correspond to units of the number field M_1. Our results are new, even in this special case.

(3) A less well known example of a decomposable form equation arises with the study of units in abelian group rings. Let Γ denote a finite abelian group with **Z**Γ denoting the integral group ring. This is the set of all expressions

$$\sum_{\gamma \in \Gamma} x_\gamma \gamma, \quad x_\gamma \in \mathbf{Z}.$$

This set forms a ring with component-wise addition, and with multiplication respecting both the operation in Γ and the distributive law. There is considerable interest in the group of units of this ring; see [K] and [Se] for details. In [EG], we showed that the group of units can be identified with the integral solutions to two decomposable form equations. Our methods give refined information about where the units in group rings lie and what they look like.

This paper is about the internal structure of the set of solutions of the decomposable form equation under the following assumptions (A):

(A1) F contains n linearly independent forms among its factors;

(A2) equation (0.2) has infinitely many solutions $\underline{x} \in \mathbf{Z}^n$.

Let $|\underline{x}|$ denote the "max"-norm defined by $|\underline{x}| = \max_{1 \le i \le n}\{|x_i|\}$. Given any positive real number T, let $F(a, T)$ denote the set

$$(0.3) \qquad F(a, T) = \{\underline{x} \in \mathbf{Z}^n : \underline{x} \text{ satisfies } (0.2) \text{ and } |\underline{x}| < T\}.$$

Since it is clear that $F(a, T)$ is a finite set, the two questions that follow are natural:

Q1 What is the asymptotic behaviour of $|F(a, T)|$, the cardinality of the set, for large T?

Q2 For every $\underline{x} \in F(a, T)$, let $c(\underline{x}) = \underline{x}/|\underline{x}|$ denote the central projection of $\underline{x}$ onto the unit ball centred at $\underline{0}$. We ask what is the asymptotic distribution of the images of the elements of $F(a, T)$ under this projection? In other words, can this set of points be described when T is large? One imagines the solutions of the equation as corresponding to the stars in the night sky. Standing upon the earth (the unit ball) and looking up, what constellations would be visible?

Write $P(T)$ for the cardinality of $F(a, T)$. In [EG], we showed there is a two-term asymptotic formula for $P(T)$ and we specified a large class of examples where a three-term asymptotic formula holds. This class of examples will now be defined. We say F is of *CM type* if the M_i in (0.1) are totally real fields or totally imaginary quadratic extensions of totally real fields and none of them has a (not necessarily proper) subfield of unit rank 1. If $n = \sum_{i=1}^t [M_i : \mathbf{Q}]$, then the condition on the subfields of $M_1, \dots, M_t$ can be omitted. This condition is very slightly broader than the one in Theorem 2 of [EG]. There we insisted that the unit ranks all be greater than 1 but this is not necessary. It is only rank equal to 1 that we wish to avoid.

Theorem A ([EG]). *Under assumptions (A):*

(i) (*See also* [EvG].) *There is a positive integer* r, *defined by (1.2), and a constant* $\rho_1 > 0$ *depending on* F *and* a *such that*

$$(0.4) \qquad P(T) = \rho_1(\log T)^r + O((\log T)^{r-1}), \quad T \to \infty.$$

(ii) *If* F *is of CM type, then* $r > 1$ *and there are constants* $\rho_1 > 0$, ρ_2 *depending on* F *and* a *such that*

$$(0.5) \qquad P(T) = \rho_1(\log T)^r + \rho_2(\log T)^{r-1} + o((\log T)^{r-1}), \quad T \to \infty.$$

Thus Q1 is answered fairly successfully. Of particular note is the three-term formula (0.5) in the CM case. Question Q2 was posed to try to understand better the implications of this three-term formula in the CM case and because our curiosity was aroused as to what can happen in the non-CM case. The method of proof of Theorem A has already implicit within it statements about distribution of the kind in Q2. We are now going to bring these to the fore in Theorem 1.

Let S denote the surface of the unit ball, $S = \{\underline{x} \in \mathbf{R}^n : |\underline{x}| = 1\}$. If $R \subset S$, write

$$P_R(T) = \#\{\underline{x} \in F(a, T) : c(\underline{x}) \in R\}.$$

Given any point Q on S and $\epsilon > 0$, let $Q(\epsilon)$ denote the ϵ-neighbourhood of Q on S. If $R \subset S$, let $R(\epsilon)$ denote the union of the $Q(\epsilon)$ for $Q \in R$.

Theorem 1. *Under assumptions (A), if* F *is of CM type:*

(i) *There is a finite set of points* $V = \{Q_1, \dots, Q_m\} \subset S$ *such that for any* $\epsilon > 0$, *with* ρ_1 *and* r *as in (0.4),*

$$(0.6) \qquad P_{V(\epsilon)}(T) = \rho_1(\log T)^r + O((\log T)^{r-1}), \quad T \to \infty.$$

(ii) *Let* W *denote the union of the projections to* S *of the straight lines joining the* $Q_i, i = 1, \dots, m$. *Then for any* $\epsilon > 0$, *with* ρ_2 *as in (0.5),*

$$(0.7) \qquad P_{W(\epsilon)}(T) = \rho_1(\log T)^r + \rho_2(\log T)^{r-1} + o((\log T)^{r-1}), \quad T \to \infty.$$

Formulae (0.6) and (0.7) say that most of the images of the solutions cluster around the lines comprising W and most of these cluster more densely around the points in V. In astronomical terms, the formulae posit the existence of finitely many "Milky Ways" which contain finitely many brighter clusters of stars.

Our second theorem gives, in the CM case, the distribution of the projections of solutions of (0.2) which fall outside $V(\epsilon)$, in terms of a "potential". Examples 1, 2 and 4 in Section 3 suggest that the distribution is not uniform around the set W. For any $R \subset S$, write R' for $S - R$, the set-theoretic complement.

Theorem 2. *Under assumptions (A) suppose* F *is of CM type. There is a continuous monotone decreasing function* $f : (0, \infty) \to [0, \infty)$ *depending on* F *and* a, *such that*

$$(0.8) \qquad P_{V(\epsilon)'}(T) = f(\epsilon)(\log T)^{r-1} + o((\log T)^{r-1}), \quad T \to \infty.$$

For $\epsilon \sim 0$, $f(\epsilon) \sim \rho_3 \log(1/\epsilon)$ *for a constant* $\rho_3 > 0$ *and* f *vanishes for all large* ϵ.

Note. Theorems 1 and 2 can be proved with other choices of norm in (0.3). For example, the Euclidean norm $|\underline{x}| = (\sum x_i^2)^{\frac{1}{2}}$ saves us from a flat earth but is messier to work with.

Geometric explanation of cluster regions. A natural question to ask is for the relation between the set V and the form F. In the sequel, we work mainly with

forms dual to the factors of F so the answer to the question is not obvious. Suppose first that all of the fields $M_i, i = 1, \ldots, t$, are totally real. Our analysis in Section 1 (see (1.8)) shows that, for most of the solutions of (0.2), one factor of F is about as large as it could be and these solutions lie close to the hyperplanes determined by the remaining factors. The intersections of these hyperplanes with the unit ball then determine the points in V. When at least one of the fields $M_i, i = 1, \ldots, t$, has complex embeddings, the set V is potentially larger. When F is of CM type, the set V is finite because at most two factors of F dominate and the real and imaginary parts of the remaining factors yield enough hyperlanes to intersect in a finite set of points on S. In the non-CM case, this argument can break down. Example 3 in Section 3 is a cubic norm form equation corresponding to a number field with two complex conjugate embeddings. When the two complex conjugate forms dominate, the real form yields a hyperplane which intersects S in a line, along which the projections are uniformly distributed. When the real form dominates, the real and imaginary parts of the complex forms yield two hyperplanes which intersect on S in two points. In this example, formula (0.6) holds with V consisting of one line and two points (see Figure 2 in Section 3). More generally, in the non-CM case, the distribution of the solutions under projection depends in quite subtle ways upon the arithmetic of the fields $M_1, \ldots, M_t$ (see Figures 2, 4, 5 and 6 in Section 3).

To help understand the issues in the non-CM case, note first that there is some laxity in the definitions of V and W, because we can add arbitrary points to V and lines to W without changing the formulae in (0.6) and (0.7). If R is a subset of V or W but the formulae do not change when it is removed, we say R is *virtual*, otherwise *actual*. For arbitrary F, formula (0.4) always holds and formula (0.6) holds for certain subsets $V \subset S$. One can ask what is the actual subset V, that is, the smallest subset—assuming it exists—of S for which (0.6) holds. Potentially there will be examples where formula (0.5) holds. For these examples, it is a challenge to describe the actual regions V and W. Formula (0.5) does not always hold; an example is given by the equation

$$(0.9) \qquad\qquad (x_1^2 - d_1 x_2^2)(x_3^2 - d_2 x_4^2) = 1,$$

where d_1 and d_2 are positive nonsquare integers. Example 5 in Section 3 (see (3.4)) is a norm form equation which does not satisfy (0.5). In examples like these, probably the best formula is one of the shape

$$(0.10) \qquad\qquad P(T) = P_{W(\epsilon)}(T) + o((\log T)^{r-1}), \quad T \to \infty.$$

One can always take W to be the convex hull of V; then it is a challenge to describe the actual region W. The reader can verify that, for equation (0.9), the actual V is a finite set of points. The actual W is a finite set of lines unless d_1/d_2 is a rational square. In the latter case, the actual W is an infinite set of points lying on finitely many lines with only a finite set of limit points, and the set of limit points is the actual V. To see the comparison, refer to Figures 5 and 6 in Section 3.

It was proved in [G] that the set of solutions of (0.2) is the union of finitely many *families* of solutions (to be defined in Section 1). In [EG] this was combined with a new variant of the Hardy-Littlewood method to give a very accurate formula in answer to Q1: Theorem A above. The diophantine input to this method is Schmidt's Subspace Theorem, which is a powerful generalisation of Roth's Theorem (see [Sc2], [Sc3]). In Section 1, the results from [EG] will be recalled then recast in geometric terms to prove Theorem 1. We suggest that Section 1 is read in tandem

with Section 3 at the end of the paper. Here computational results are presented in the form of pictures, which give a convincing account of the phenomena in [EG]. Examples are also included of cases not covered by our results. Besides their aesthetic appeal, we found these pictures inspired both our curiosity and our understanding. Theorem 2 is an explanation of some of the pictures on view in Section 3 and it is proved in Section 2, using the theory of uniform distribution.

We are indebted to the referee for many helpful comments on an earlier version of this paper.

1. Proof of Theorem 1

In order to prove Theorem 1, we will need to go into the background to the proof of Theorem A. The solutions of (0.2) lie in a finite number of classes which are orbits of unit groups. The technical term for a class is *family of solutions* and we begin by defining this term. Let A denote the algebra

$$A = M_1 \oplus \cdots \oplus M_t.$$

This is the $\mathbf{Q}$-algebra direct sum of the number fields $M_1, \ldots, M_t$ formed with componentwise operations. Thus, $1_A = (1, \ldots, 1)$ is the unity of A, and A^*, the multiplicative group of invertible elements of A, is $\{(\alpha_1, \ldots, \alpha_t) \in A : \alpha_1 \ldots \alpha_t \neq 0\}$. The norm $N_{A|\mathbf{Q}}(\alpha)$ of $\alpha = (\alpha_1, \ldots, \alpha_t) \in A$ is defined to be the usual algebra norm, that is, the determinant of the $\mathbf{Q}$-linear map $x \mapsto \alpha x$ from A to itself. The norm is multiplicative and

$$N_{A|\mathbf{Q}}(\alpha) = \prod_{i=1}^{t} N_{M_i|\mathbf{Q}}(\alpha_i).$$

Therefore rewrite equation (0.2) as

$$(1.1) \qquad\qquad q N_{A|\mathbf{Q}}(c) = a, \qquad c \in \mathfrak{M},$$

where $\mathfrak{M}$ is defined to be $\mathfrak{M} = \{c = (\phi_1(\underline{x}), \ldots, \phi_t(\underline{x})) \in A : \underline{x} \in \mathbf{Z}^n\}$. Now $\mathfrak{M}$ is a finitely generated $\mathbf{Z}$-module. Let $V = \mathbf{Q}\mathfrak{M}$ denote the $\mathbf{Q}$-vector space generated by $\mathfrak{M}$. For any subalgebra B of A with $1_A \in B$, denote by O_B the integral closure of $\mathbf{Z}$ in B and by O_B^* the multiplicative group of invertible elements of O_B. Let

$$V^B = \{v \in V : vB \subseteq V\} \quad \text{and} \quad \mathfrak{M}^B = V^B \cap \mathfrak{M}.$$

Obviously V^B is closed under multiplication by elements of B. Now define

$$U_{\mathfrak{M},B} = \{u \in O_B^* : u\mathfrak{M}^B = \mathfrak{M}^B, \quad N_{A|\mathbf{Q}}(u) = 1\}.$$

This is a subgroup of finite index in O_B^*. If $c \in \mathfrak{M}^B$ is a solution of (1.1), so is every element of $cU_{\mathfrak{M},B}$. Such an orbit is called an $(\mathfrak{M}, B)$-*family of solutions* of (1.1), and hence of (0.2) as well. It is a fundamental result in this subject (see [G]) that the set of solutions of (1.1) is a union of finitely many families of solutions.

The group O_B^* is finitely generated; let r_B denote the torsion-free rank. Use r to denote the maximum of the r_B,

$$(1.2) \qquad\qquad r = \max_B \{r_B\},$$

taken over all $\mathbf{Q}$-subalgebras B of A with $1_A \in B$ for which (1.1) has an $(\mathfrak{M}, B)$-family of solutions. Assumption (A1) guarantees that $r > 0$ and the CM assumption guarantees that $r > 1$. Any $(\mathfrak{M}, B)$-family with $r = r_B$ is called a *maximal family*. In the CM case, we may replace $U_{\mathfrak{M},B}$ by a subgroup $\overline{U}_{\mathfrak{M},B}$ of finite index consisting

of elements $(u_1, \ldots, u_t)$ with totally real and totally positive units $u_1, \ldots, u_t$ from $M_1, \ldots, M_t$, respectively. When this is done, we refer to *real families* and *maximal real families* with the obvious abuse of language. (A real family does not necessarily consist of real numbers, but rather, of numbers which are the orbit of a group consisting of real numbers.) Note, in the CM case, that the solutions of (0.2) are contained in a union of finitely many real families. It is therefore sufficient to do any counting within a fixed, maximal real family of solutions.

Let $\mathfrak{F}$ denote a fixed maximal real family as above and write U for the associated $\overline{U}_{\mathfrak{M},B}$. For each $M_i, i = 1, \ldots, t$, let $\sigma_{ij} : M_i \to \mathbf{C}, j = 1, \ldots, [M_i : \mathbf{Q}]$ denote the distinct embeddings into $\mathbf{C}$. Write $\phi_{ij}(\underline{x})$ for the conjugates of the forms $\phi_i(\underline{x})$, $i = 1, \ldots, t$. We are assuming that these $\sum_{i=1}^t [M_i : \mathbf{Q}]$ forms contain n linearly independent forms. By (1.1) and the definition of the algebra norm, there are algebraic numbers b_{ij} such that for all $\underline{x} \in \mathfrak{F}$

$$(1.3) \qquad \phi_{ij}(\underline{x}) = b_{ij} u_{ij}, \quad i = 1, \ldots, t; \ j = 1, \ldots, [M_i : \mathbf{Q}],$$

with algebraic units $u_{ij} = \sigma_{ij}(u_i)$. Write $(u_{ij}) = (u_k)_{1 \le k \le m}$ for the vector of the $u_{ij}, i = 1, \ldots, t, j = 1, \ldots, [M_i : \mathbf{Q}]$, where $m = \sum_{i=1}^t [M_i : \mathbf{Q}]$, and similarly for $(b_{ij}) = (b_k)$. Then we obtain a system of m linear equations

$$(1.4) \qquad \Phi \underline{x} = (b_k u_k),$$

where the coefficients of the $m \times n$ matrix Φ are those of the linear forms ϕ_{ij}. The system in (1.4) is (left) invertible by the assumption being made about the linear factors of F. Writing Ψ for the left inverse of Φ gives

$$(1.5) \qquad \underline{x} = \Psi(b_k u_k).$$

If $u = (u_1, \ldots, u_t) \in U$, then define

$$(1.6) \qquad H(u) = \max_{i,j}\{\sigma_{ij}(u_i)\},$$

the largest value of any conjugate of any $u_i, i = 1, \ldots, t$. Write $H^*(u)$ for the second largest element of the set in (1.6), where complex conjugate embeddings are identified. It follows from (1.4), (1.5) and the triangle inequality that $|\underline{x}|$ and $H(u)$ are commensurate; that is, they are bounded by constant multiples of each other. Theorems 1 and 2 exploit that fact, enabling the counting of solutions of (0.2) in a particular family to be effected by counting elements $u \in U$ with respect to H.

Proof of Theorem 1. Fix indices (i, j) with $H(u) = u_{ij}$. Using (1.5), there is a vector $\underline{c}$ depending on $\mathfrak{F}$ and the (i, j) only (via Ψ and $\underline{b}$) such that, for all u with $H(u) = u_{ij}$,

$$(1.7) \qquad \underline{x} = \underline{c}H(u) + O(H^*(u)),$$

where the big O term denotes a vector whose norm is $O(H^*(u))$. It follows from (1.7) that $|\underline{x}|$ will have the shape

$$|\underline{x}| = |dH(u) + O(H^*(u))|.$$

Writing $I(u) = H^*(u)/H(u)$, $c(\underline{x})$ will have the shape

$$c(\underline{x}) = \underline{e} + O(I(u)).$$

A fundamental result from [EG] (see Lemma 6(i)) is that in the CM case, for all $0 < \epsilon < 1$, asymptotically all $u \in U$ have $I(u) < \epsilon$. In other words,

$$(1.8) \qquad |\{u \in U : \epsilon \le I(u), H(u) < T\}| = O((\log T)^{r-1}).$$

Formula (0.6) in Theorem 1 follows from (1.7) and (1.8), by varying the indices (i, j) and the maximal real family $\mathfrak{F}$. Each vector $\underline{e}$ in (1.7) gives rise to an element of V and all elements of V arise in this way. Note that $\underline{e}$ is guaranteed to be real.

Formula (0.7) comes about by refining (0.6) together with a more delicate interplay between H and $|.|$. Write $H^{**}(u)$ for the third largest element of $\{\sigma_{ij}(u_i)\}$, where complex conjugate embeddings are identified. Fix indices (i, j) and (k, l) with $H(u) = u_{ij}$ and $H^*(u) = u_{kl}$. There is a vector $\underline{c}^*$ with

$$(1.9) \qquad \underline{x} = \underline{c}H(u) + \underline{c}^*H^*(u) + O(H^{**}(u)).$$

It follows from (1.9) that $|\underline{x}|$ will have the shape

$$|\underline{x}| = |dH(u) + d^*H^*(u) + O(H^{**}(u))|.$$

Writing $J(u) = H^{**}(u)/H(u)$, $c(\underline{x})$ will have the shape

$$(1.10) \qquad c(\underline{x}) = \underline{e} + \underline{e}^*I(u)/|1 + eI(u)| + O(J(u)).$$

In [EG] (see Lemma 6(ii)), we proved that, for all $0 < \epsilon < 1$,

$$(1.11) \qquad |\{u \in U : \epsilon \le J(u), H(u) < T\}| = O((\log T)^{r-2}).$$

Formula (0.7) in Theorem 1 follows from (1.10) and (1.11), by varying the indices $(i, j), (k, l)$ (and hence the vectors $\underline{e}, \underline{e}^*$) and the maximal real family $\mathfrak{F}$. Each pair of vectors $\underline{e}$ and $\underline{e}^*$ in (1.9) gives rise to a line in W and all the lines in W arise in this way. $\qquad \square$

This section closes with the explicit determination of the set V for Example 3 from Section 0. In [EG], we showed that if Γ is a finite abelian group, then the units of $\mathbf{Z}\Gamma$ yield a form of CM type if and only if the following property holds:

$$(1.12) \qquad \text{no quotient of } \Gamma \text{ is cyclic of order 5, 8 or 12.}$$

Thus, formulae (0.4) and (0.6) hold always but (0.5) and (0.7) hold only under condition (1.12). Let $\chi \in \hat{\Gamma}$ denote a character and define

$$e_\chi = |\Gamma|^{-1} \sum_{\gamma \in \Gamma} \overline{\chi}(\gamma)\gamma \in \mathbf{C}\Gamma.$$

The e_χ for $\chi \in \hat{\Gamma}$ form a system of $n = |\Gamma|$ independent, orthogonal idempotents for $\mathbf{C}\Gamma$. The results in [E1] and [E2] show that the set V in formula (0.6) comes from the elements

$$e_\chi\chi(\gamma) + e_{\overline{\chi}}\overline{\chi}(\gamma) \in \mathbf{R}\Gamma \text{ for } \chi \in \hat{\Gamma}, \gamma \in \Gamma.$$

These elements give virtual points in V if and only if $\mathbf{Q}(\chi)$, the field generated over $\mathbf{Q}$ by the values of χ, is $\mathbf{Q}$ or an imaginary quadratic extension of $\mathbf{Q}$.

2. Proof of Theorem 2

Formula (1.10) from Section 1 gives

$$|c(\underline{x}) - \underline{e}| = |\underline{e}^*|I(u)/|1 + eI(u)| + O(J(u)).$$

There are only finitely many possibilities for $\underline{e}, \underline{e}^*$ and e. From (1.11), we can regard the last term as vanishingly small. Thus, for solutions $\underline{x}$ of (0.2) lying in a fixed maximal real family, with i, j, k, l fixed as in Section 1, the condition $c(\underline{x}) \notin V(\epsilon)$ yields finitely many inequalities of the form $I(u)/|1 + eI(u)| > \kappa\epsilon$ for $\kappa > 0$. Inverting each of these inequalities locates $I(u)$ within an open interval in $[0, 1]$ and $c(\underline{x}) \notin V(\epsilon)$ corresponds to the intersection of these intervals. For all small $\epsilon > 0$,

this interval is of the form $(\epsilon/(\lambda\epsilon + \theta), 1)$ for constants λ and $\theta > 0$. Thus, it is sufficient to count units u with i, j, k, l fixed and $I(u)$ above a fixed bound. Let $C > 0$ denote a constant and suppose i, j, k, l are fixed. That is, consider those $u \in U$ with $H(u) = u_{ij}$ and $H^*(u) = u_{kl}$. Define

$$(2.1) \qquad\qquad U_C(T) = |\{u : e^{-C} < u_{kl}/u_{ij}, u_{ij} < T\}|.$$

We claim there is a positive constant ν, which depends upon i, j, k, l only, such that $U_C(T)$ satisfies the following asymptotic formula:

$$(2.2) \qquad\quad U_C(T) = C\nu(\log T)^{r-1} + o((\log T)^{r-1}), \quad \text{as } T \to \infty.$$

From (2.2), a formula like (0.8) holds with $f(\epsilon) = \nu \log(\lambda + \theta/\epsilon)$ for all small ϵ. Theorem 2 follows by summing over all maximal real families and all i, j, k, l.

To prove formula (2.2), use the notation in Section 1. Note that U is a free abelian group of rank r. Taking logarithms of the u_{ij} gives rise to a family of linear forms $L_1, \ldots, L_s$ on U. Each form corresponds to the logarithm of a conjugate of some component of u. After choosing a basis of U, we may regard the L_i, $i = 1, \ldots, s$, as linear forms on $\mathbf{Z}^r$. Assuming that forms are not counted if they are identically zero, the CM condition (in particular, the prohibition of the rank 1 case) guarantees that at least two of the coefficients of each L_i are linearly independent over $\mathbf{Q}$. The following relation is satisfied by this family of forms:

$$(2.3) \qquad\qquad L_1(\underline{y}) + \cdots + L_s(\underline{y}) = 0, \quad \text{for all } \underline{y} \in \mathbf{Z}^r.$$

This comes from the fact that the underlying quantity is a unit so the product of all the conjugates of all the components is equal to 1. Taking logarithms gives the relation in (2.3). Clearly each of the forms extends to $\mathbf{R}^r$ and the same relation (2.3) holds. Counting heights of elements $u \in U$ with $H(u) < T$ is equivalent to counting lattice points $\underline{y} \in \mathbf{Z}^r$ satisfying $L(\underline{y}) = \max_i\{L_i(\underline{y})\} < X = \log T$.

Let $L^*(\underline{y})$ denote the second largest component of the vector $(L_i(\underline{y}))_{1 \leq i \leq s}$. The inequalities defining $U_C(T)$, in (2.1), become ($X = \log T$),

$$(2.4) \qquad\qquad L(\underline{y}) < X, \quad -C < L^*(\underline{y}) - L(\underline{y}).$$

To take account that i, j, k, l are fixed, we must fix $L = L_v$ and $L^* = L_w$ for some $1 \leq v, w \leq s$. Define the following counting function:

$$(2.5) \qquad A_C(X) = |\{\underline{y} : L_v(\underline{y}) < X, \quad L_w(\underline{y}) < L_v(\underline{y}) < L_w(\underline{y}) + C\}|.$$

Formula (2.2) is a direct consequence of the following asymptotic formula:

$$(2.6) \qquad\qquad A_C(X) = C\nu X^{r-1} + o(X^{r-1}), \quad \text{as } X \to \infty,$$

where ν depends only upon $L_1, \ldots, l_s, v$ and w.

The best approach to proving (2.6) is the direct one of comparing the number of lattice points being counted with the volume of the region defined by the inequalities in (2.4). But note that the volume of the boundary of the region has the same order of magnitude as the main term of the asymptotic formula so it cannot be used as the error term. However, the boundary is of the type to allow a uniform distribution argument to estimate the error. Let $\mathfrak{S}_C(X)$ denote the region of $\mathbf{R}^r$ defined by the inequalities in (2.4) and let μ_r denote Lebesgue measure in $\mathbf{R}^r$.

Lemma 2.2. *There is a positive constant ν such that*

$$\mu_r(\mathfrak{S}_C(X)) = C\nu X^{r-1} + O(X^{r-2}).$$

Proof. This is obtained by multiple integration as follows. The region of integration subdivides according to the possible orderings on the forms. After relabelling, it is sufficient to consider the region $T(X)$ defined by $L(\underline{y}) = L_v(\underline{y}) \leq X$ and

$$L_w(\underline{y}) + C \geq L_v(\underline{y}) \geq L_w(\underline{y}) \geq \cdots \geq L_s(\underline{y}).$$

To avoid discussing trivial cases, assume that $T(X)$ has positive volume. For $0 < \gamma \leq C$, let $T_\gamma(X)$ be the region defined by $L(\underline{y}) = L_v(\underline{y}) \leq X$ and

$$L_w(\underline{y}) + \gamma = L_v(\underline{y}) \geq L_w(\underline{y}) \geq \cdots \geq L_s(\underline{y}).$$

A special property of the linear forms L_i is that any two of them are linearly dependent if and only if they are equal (disregarding forms which are identically zero). Hence we may assume that L_v and L_w are linearly independent, and this guarantees that $T_\gamma(X)$ is an $(r-1)$-dimensional polytope. The inequalities defining $T_\gamma(X)$ are such that

$$(2.7) \qquad \mu_{r-1}\left(T_\gamma(X)\right) = X^{r-1}\mu_{r-1}\left(T_{\frac{\gamma}{X}}(1)\right)$$

with μ_{r-1} denoting $(r-1)$-dimensional Lebesgue measure. Therefore we want to calculate the $(r-1)$-dimensional volume of $T_\gamma(1)$ for small γ. Again because L_1 and L_2 are independent, this volume is a differentiable function of γ in the neighbourhood of $\gamma = 0$:

$$(2.8) \qquad \mu_{r-1}\left(T_\gamma(1)\right) = \mu_{r-1}\left(T_0(1)\right) + O(\gamma).$$

Now substitute γ/X for γ in (2.8) and put this into (2.7). Integrating γ over $0 < \gamma \leq C$ gives the required estimate. $\qquad\square$

Lemma 2.3. *For every choice of L and L^*, the linear form $L - L^*$ has the property that at least two of its coefficients are linearly independent over $\mathbf{Q}$.*

Proof. First, do the case where $t = 1$. The linear forms L and L^* correspond to embeddings σ and σ^* of the field M_1. The identification of complex conjugate embeddings makes it sufficient to assume σ and σ^* differ on M_1^+, the maximal real subfield of M_1. If the allegation in Lemma 2.3 is false, then $L - L^*$ is a real multiple of an integral linear form whose integral zeros are a lattice of rank $r - 1$. There are finitely many lattices coming from units belonging to proper subfields of M_1^+ and the rank of each one is bounded by $\frac{r+1}{2} - 1$. This is strictly less than $r - 1$ because $1 < r$. There exists an integer vector $\underline{y}$ with $L(\underline{y}) = L^*(\underline{y})$ that does not belong to any of these lattices. This vector $\underline{y}$ corresponds to a unit of M_1^+ which does not lie in any proper subfield of M_1^+. Thus σ and σ^* agree on this unit and hence on M_1^+, a contradiction. The general case is entirely similar. Now σ and σ^* correspond to vectors of embeddings. Assuming they differ on one component, we can use the equation $L(\underline{y}) = L^*(\underline{y})$ to find a unit $u \in U$ upon which σ and σ^* agree on every component, a contradiction. $\qquad\square$

Now we complete the proof of (2.6). The region $\mathfrak{S}_C(X)$ is defined by inequalities involving finitely many linear forms. Thus the boundary consists of a finite union of hyperplanes. Write $\Delta\mathfrak{S}_C(X)$ for the boundary of the region. For lattice points $\underline{y} \in \mathbf{Z}^r$, write $C_{\underline{y}}$ for the unit ball centred at $\underline{y}$. Let $\mathbf{Z}_C(X)$ denote the lattice points

$\underline{y} \in \mathbf{Z}^r$ such that $C_{\underline{y}}$ has nonempty intersection with $\Delta \mathfrak{S}_C(X)$. Write

$$(2.9) \qquad\qquad S_1 = \sum_{C_{\underline{y}} \subset \mathfrak{S}_C(X)} 1,$$

$$(2.10) \qquad\qquad S_2 = \sum_{\underline{y} \in \mathbf{Z}_C(X) \cap \mathfrak{S}_C(X)} \mu_r(C_{\underline{y}} \cap \mathfrak{S}_C(X)),$$

$$(2.11) \qquad\qquad S_3 = \sum_{\underline{y} \in \mathbf{Z}_C(X) - \mathfrak{S}_C(X)} \mu_r(C_{\underline{y}} \cap \mathfrak{S}_C(X)).$$

The volume in Lemma 3.2 decomposes as follows:

$$(2.12) \qquad\qquad \mu_r(\mathfrak{S}_C(X)) = S_1 + S_2 + S_3.$$

In S_2, the boundary conditions guarantee that the distances between the $\underline{y}$ and the boundary are uniformly distributed. We sum the values of a continuous function of those distances. The function clearly has integral $1/2$. Similar remarks hold for the sum S_3. From the theory of uniform distribution (see [KN]) and the symmetry, each of S_2 and S_3 is

$$(2.13) \qquad\qquad \frac{1}{2}\left(\sum_{\underline{y} \in \mathbf{Z}_C(X) \cap \mathfrak{S}_C(X)} 1 \right) + o(X^{r-1}).$$

Thus, (2.9), (2.12) and (2.13) give

$$\mu_r(\mathfrak{S}_C(X)) = \sum_{\underline{y} \in \mathfrak{S}_C(X)} 1 + o(X^{r-1}) = A_C(X) + o(X^{r-1}).$$

Now (2.6) follows from Lemma 2.2.

3. Computational results

In this section, we will present some pictures to illustrate the clustering phenomena for some decomposable form equations in a small number of variables. Examples in both the CM and non-CM cases are included. For generating the pictures, we used the software packages KANT from the TU Berlin ([Ka]) for the calculations of number field data and Maple for plotting.

The first example, for motivation, is Pell's equation with $d = 2$, $x_1^2 - 2x_2^2 = 1$. The field is $K_1 = \mathbf{Q}(\sqrt{2})$ and the solutions (x_1, x_2) correspond to units of norm 1 in the ring $\mathbf{Z}[\sqrt{2}]$ via $(x_1, x_2) \mapsto x_1 + x_2\sqrt{2}$. The solutions all lie on a hyperbola with asymptotes $x_1 = \pm\sqrt{2}x_2$ so the distribution is obvious. The set V consists of four points $(\pm 1, \pm 1/\sqrt{2})$ on the unit square.

For cubics and quartics, we can still visualize the different types of behaviour. Consider the norm form equation for the totally real cubic K_2 of discriminant 49, which is the maximal real subfield of the cyclotomic number field generated by a primitive 7th root of unity ζ. As a field, K_2 is generated over $\mathbf{Q}$ by $\theta = \zeta + \zeta^6$, the minimal polynomial of θ is $x^3 + x^2 - 2x - 1$, and the ring of integers is $\mathbf{Z}[\theta]$. Choosing $1, \theta, \theta^2$ as the basis of $\mathbf{Z}[\theta]$, every unit of norm 1 corresponds to a solution $\underline{x} \in \mathbf{Z}^3$ of the norm form equation

$$(3.1)$$
$$x_1^3 + x_2^3 + x_3^3 - x_1^2 x_2 + 5x_1^2 x_3 - 2x_1 x_2^2 + 6x_1 x_3^2 - x_2^2 x_3 - 2x_2 x_3^2 - x_1 x_2 x_3 = 1.$$
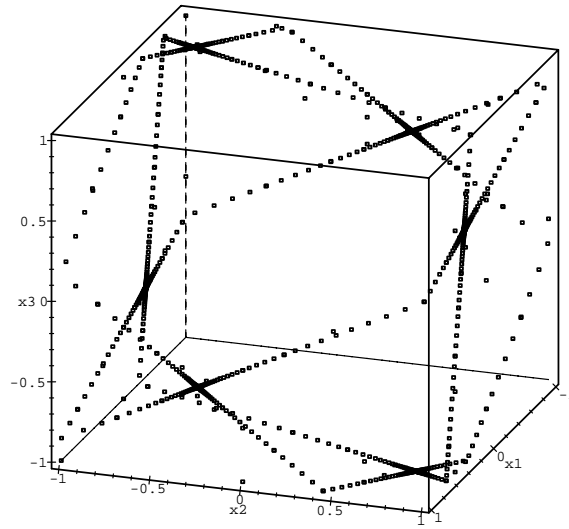
FIGURE 1.   Projection of solutions of (3.1)

After central projection, those solutions look like Figure 1. In this example, there are four maximal real families of solutions, each contributing to three of the six points in V. Theorem 2 describes precisely the distribution of points close to W.

For the third example, take $K_3 = \mathbf{Q}(2^{\frac{1}{3}})$ which has ring of integers $\mathbf{Z}[2^{\frac{1}{3}}]$. Choosing the basis $1, 2^{\frac{1}{3}}, 2^{\frac{2}{3}}$ for the ring of integers gives the non-CM, norm form equation

$$(3.2) \qquad\qquad x_1^3 + 2x_2^3 + 4x_3^3 - 6x_1x_2x_3 = 1.$$
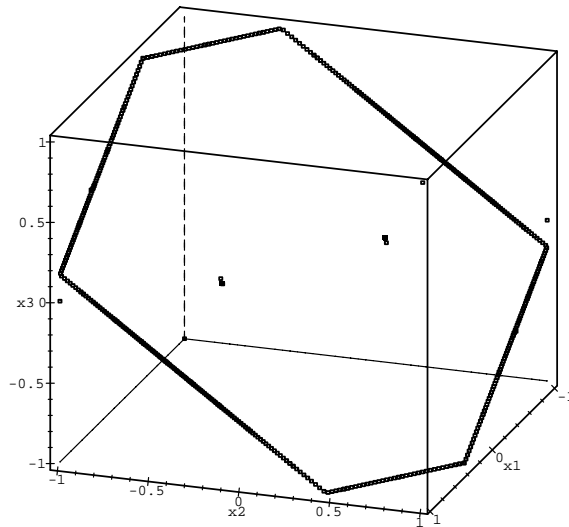
FIGURE 2.   Projection of solutions of (3.2)

Figure 2 shows the distribution of the projected solutions: essentially a line and two isolated points. Due to the limited resolution, many images are printed on top of each other—out of 800 points in the whole picture, 261 are closer than 0.01 in distance to the isolated points! Note that, this time, the projections around the line are uniformly distributed. Formula (0.6) holds with V consisting of the union of a line and two points.

Next come two quartic cases: one is CM and the other is not. For a totally real quartic, take $K_4 = \mathbf{Q}(\alpha)$, where α is a root of $x^4 - 2x^3 + 3x + 2$. The ring of integers
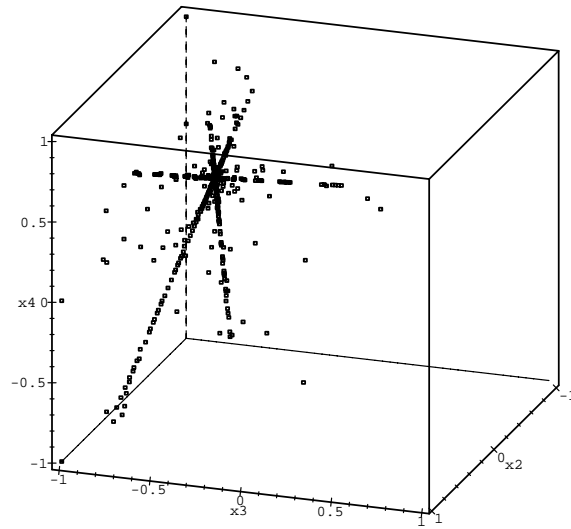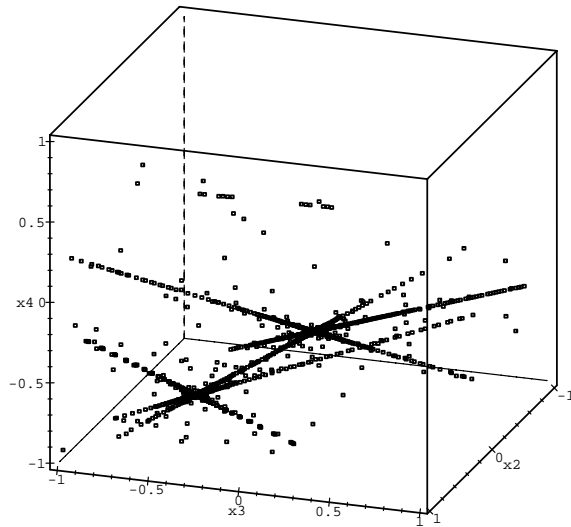

FIGURE 3A.   Projection of solutions of (3.3)

FIGURE 3B.   Projection of solutions of (3.3)—face 2

of K_4 is $\mathbf{Z}[\alpha]$ and we choose the powers of α as basis. Then we consider the norm form equation

$$(3.3) \qquad N_{K_4|\mathbf{Q}}(x_1 + x_2\alpha + x_3\alpha^2 + x_4\alpha^3) = 1.$$

Figure 3A shows one face (a 3-dimensional cube) of the corresponding 4-dimensional unit ball with the projection of the smallest 700 solutions with respect to Euclidean norm. Figure 3B shows another face, with the projection of 4217 solutions and two points of V. It is clearly visible that the images of solutions cluster densely around the lines in W, and yet more densely around the points
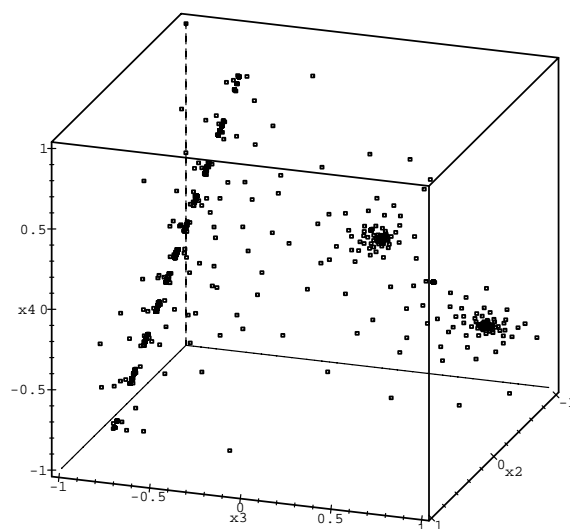

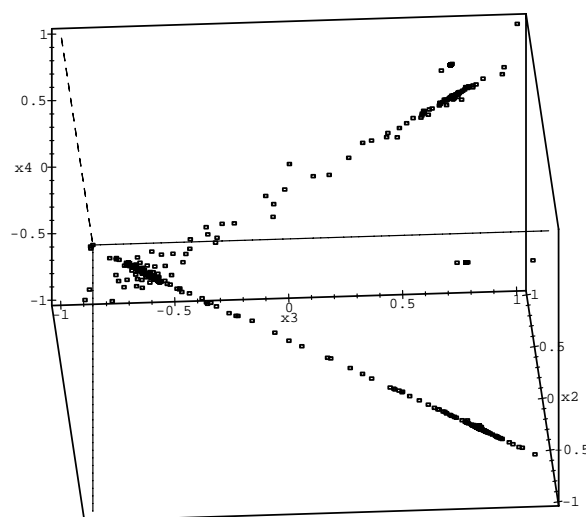FIGURE 4A.   Projection of solutions of (3.4)

FIGURE 4B.   The same face as Figure 4A, tilted upwards

in V. Once again, Theorem 2 goes beyond this in describing this phenomenon quantitatively.

For the next example, let K_5 denote the field $K_5 = \mathbf{Q}(2^{\frac{1}{4}})$. The ring of integers here is $\mathbf{Z}[2^{\frac{1}{4}}]$, and we have chosen the powers of $2^{\frac{1}{4}}$ as basis. The norm form equation is

$$(3.4) \qquad N_{K_5|\mathbf{Q}}(x_1 + x_2 2^{\frac{1}{4}} + x_3 2^{\frac{2}{4}} + x_4 2^{\frac{3}{4}}) = 1.$$

There are 3000 solutions represented in Figure 4A. This example is not CM and the clustering behaviour is different. Formula (0.6) holds with the actual V consisting of two points (both of which are captured on the face shown) and one line which appears very curiously "dashed". However, it is possible to show that the distribution around this line is uniform. This line corresponds to units where the complex conjugates dominate in absolute value. The points correspond to units where one of the real conjugates dominates. Another new phenomenon appears when we tilt the picture upwards—see Figure 4B. The solutions are nearly all confined to two planes determined by the line and one of the points. If we tilted the face a little bit more, we could shrink the line to a point on the paper, and the planes would appear as lines. Formula (0.10) holds with the actual W consisting of finitely many planes and a discrete set of points which lie on finitely many lines.

The last two examples arise from the product of two Pellian equations (see (0.9)). First

$$(3.5) \qquad (x_1^2 - 2x_2^2)(x_3^2 - 3x_4^2) = 1.$$

Figure 5 shows one 3-dimensional face of the 4-dimensional unit cube containing the projections of 4121 solutions which belong to two different families. Each family contributes two V-shapes. In many respects, the distribution is the same as in the 2nd and 4th examples. Formulae (0.5) and (0.7) do not hold but (0.6) holds with V
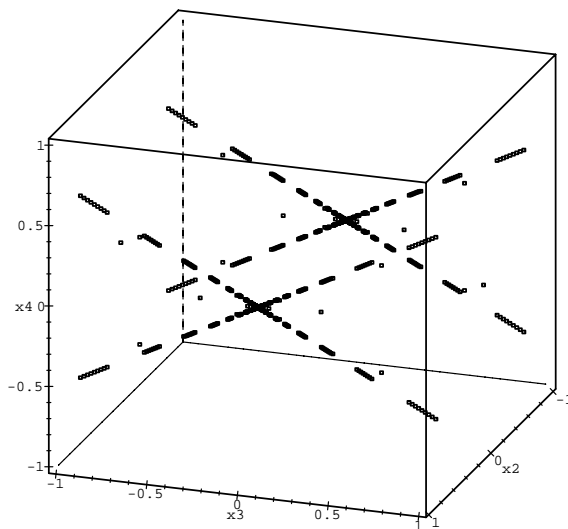

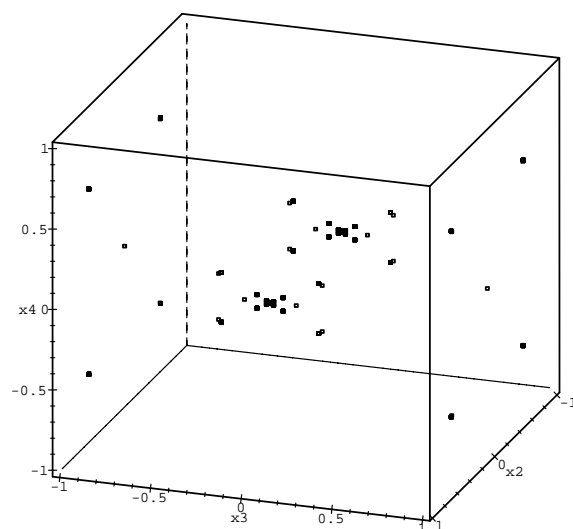FIGURE 5.   Projection of solutions of (3.5)

FIGURE 6. Projection of solutions of (3.6)

a finite set of points. Formula (0.8) holds and formula (0.10) holds with the actual W consisting of a finite union of lines.

Now consider the equation

$$(3.6) \qquad (x_1^2 - 3x_2^2)(x_3^2 - 3x_4^2) = 1.$$

Figure 6 shows one 3-dimensional face of the 4-dimensional cube. The distribution is markedly different. The projections of some 4000 solutions are on view but the picture appears to contain far fewer points due to the limited resolution. The solutions belong to two families, each one contributing two V-shapes. Formulae (0.5), (0.7) and (0.8) do not hold. Formula (0.6) holds with V a finite set of points. Formula (0.10) holds but, in contrast to the previous example, the actual W is an infinite set of points which lie on finitely many lines and have V as the only limit points.

References

[E1] G. R. Everest, *Angular distribution of units in abelian group rings-an application to Galois-module theory*, J. Reine Angew. Math. (1987), 24-41. MR **88g:**11085

[E2] G. R. Everest, *Units in abelian group rings and meromorphic functions*, Illinois J. Math. **33** (1989), 542-553. MR **90k:**20011

[EG] G. R. Everest, K. Győry, *Counting solutions of decomposable form equations*, Acta Arith. **79** (1997), 173-191. MR **98e:**11037

[EvG] J.-H. Evertse, K. Győry, *The number of families of solutions of decomposable form equations*, Acta Arith. **80** (1997), 367-394. MR **98e:**11038

[G] K. Győry, *On the number of families of solutions of systems of decomposable form equations*, Publ. Math. Debrecen **22** (1993), 65-101. MR **94e:**11027

[K] G. Karpilovsky, *Unit Groups of Classical Rings*, Oxford University Press, 1988. MR **90e:**20007

[Ka] M. Daberkow, C. Fieker, J. Klüners, M. Pohst, K. Roegner and K. Wildanger, *KANT V4*, J. Symbolic Comp. **24** (1997), 267-283. MR **99g:**11150

[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974. MR **54:**7415

[Sc1]   W. Schmidt, *Norm form equations*, Ann. of Math. **96** (1972), 526-551. MR **47:**3313

[Sc2]   W. Schmidt, *Diophantine Approximation, Lecture Notes in Mathematics vol. 785*, Springer, 1980. MR **81j:**10038

[Sc3]   W. Schmidt, *Diophantine Approximation and Diophantine Equations, Lecture Notes in Mathematics vol. 1467*, Springer, 1991. MR **94f:**11059

[Se]   S. Sehgal, *Topics in Group Rings*, Dekker, New York 1978. MR **80j:**16001

SCHOOL OF MATHEMATICS, UNIVERSITY OF EAST ANGLIA, NORWICH, NORFOLK NR4 7TJ, UNITED KINGDOM

E-mail address: `g.everest@uea.ac.uk`

INSTITUTE OF MATHEMATICS AND INFORMATICS, LAJOS KOSSUTH UNIVERSITY, H-4010 DEBRECEN, PF 12, HUNGARY

E-mail address: `igaal@math.klte.hu`

INSTITUTE OF MATHEMATICS AND INFORMATICS, LAJOS KOSSUTH UNIVERSITY, H-4010 DEBRECEN, PF 12, HUNGARY

E-mail address: `gyory@math.klte.hu`

SCHOOL OF MATHEMATICS, UNIVERSITY OF EAST ANGLIA, NORWICH, NORFOLK NR4 7TJ, UNITED KINGDOM

E-mail address: `C.Rottger@uea.ac.uk`