JOURNAL OF NUMBER THEORY 43, 228–250 (1993)

The Number of Solutions of Bounded Height to a System of Linear Equations

JEFFREY LIN THUNDER

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48103

Communicated by M. Waldschmidt

Received April 18, 1991; revised August 21, 1991

We count the number of solutions with height less than or equal to B to a system of linear equations over a number field. We give explicit asymptotic estimates for the number of such solutions as B goes to infinity, where the constants involved depend on the classical invariants of the number field (degree, discriminant, class number, etc.). The problem is reformulated as an estimate for the number of lattice points in a certain bounded domain. © 1993 Academic Press, Inc.

1. INTRODUCTION

The simplest Diophantine equations are the linear ones. Some typical questions often asked in Diophantine equations are easily answered in this case. There are other (often very important) questions which naturally arise, however, which do not have immediately apparent answers. For example, in transcendental number theory and Diophantine approximation one is often led to the construction of certain auxiliary polynomials with suitable properties. The problem then becomes either to explicitly construct them, or to at least prove the existence of such polynomials. The latter is typically reduced to finding "small" solutions (usually in rational integers) to a system of linear equations. For a homogeneous system, results of the following kind are often used.

SIEGEL'S LEMMA [11]. *Suppose l and n are positive integers with $l < n$. Let*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ln}x_n &= 0 \end{aligned} \tag{1}$$

be a system of l linear equations in n unknowns with rational integer coef-

0022-314X/93 $5.00

ficients a_{ij} satisfying $|a_{ij}| \leqslant A$ for all i and j. Then there is a non-trivial integral solution $\mathbf{x} = (x_1, \ldots, x_n)$ with

$$\max_{1 \leqslant i \leqslant n} \{|x_i|\} < 1 + (nA)^{l/(n-l)}.$$

This can be proven using Dirichlet's box principle, for example. One can also show that the exponent $l/(n-l)$ is best possible [10]. Bombieri and Vaaler in [1] tackle the more general problem of finding $n-l$ linearly independent solutions of small *height*, where the coefficients are integers in a number field K and the unknowns x_i are restricted to integers in a subfield $F \subseteq K$ (see also [10]).

Here, instead of trying to prove the existence of certain solutions of small height, we ask how many solutions with height less than or equal to a given bound one would expect to find. For instance, suppose the equations in (1) are linearly independent with coefficients a_{ij} in a number field K. The system (1) then defines an m-dimensional subspace $S \subseteq K^n$, where $m = n - l$. How many one-dimensional subspaces $T \subseteq S$ have height less than or equal to a given bound? In answer to this question, we prove the following.

THEOREM 1. *Let K be a number field of degree d and let S be an m-dimensional subspace of K^n, where $2 \leqslant m \leqslant n$. The number of one-dimensional subspaces $T \subseteq S$ with height $\leqslant B$ is asymptotically*

$$a(m, K)\frac{B^m}{H(S)} + O(B^{m-1/d})$$

as $B \to \infty$, where the constant implicit in the O notation depends only on m and K. If $m = 2$ and $d = 1$, then the error term above is replaced by $O(B \log B)$.

The value of the constant $a(m, K)$ appearing in the above theorem is explicitly given as follows. Write $d = r_1 + 2r_2$, where r_1 is the number of real embeddings of K into the complex numbers $\mathbb{C}$ and r_2 is the number of pairs of complex conjugate embeddings. Let δ be the discriminant, R be the regulator, h be the class number, and w be the number of roots of unity of K. Further, let ζ_K be the Dedekind zeta function of K and let $V(n)$ denote the volume of the unit ball in $\mathbb{R}^n$. With this notation,

$$a(m, K) = \frac{hR}{w}\left(\frac{2^{r_2}}{\sqrt{|\delta|}}\right)^m m^{r_1+r_2-1}\frac{(V(m))^{r_1}(V(2m))^{r_2}}{\zeta_K(m)}.$$

In the special case $m = n$, i.e., when $S = K^n$, the theorem estimates the number of one-dimensional subspaces of K^n, or what is the same, points in projective $(n-1)$-space $\mathbb{P}^{n-1}(K)$, of height $\leqslant B$. This has already been

done by Schanuel [8], although with a slightly different definition of height. This case is also a specific instance of [12, Theorem 1], where an estimate for the number of m-dimensional subspaces of K^n with height $\leqslant B$ is given.

In Section 4 we prove an analogous result for the inhomogeneous case.

We now give our definition of the height of a subspace. Such a definition was first given in [9] and is the height used by Bombieri and Vaaler in [1].

Let

$$\alpha \mapsto \alpha^{(i)} \qquad (1 \leqslant i \leqslant d)$$

denote the embeddings of K into the complex numbers, ordered so that the first r_1 are real and $\alpha^{(i+r_2)} = \overline{\alpha^{(i)}}$ for $r_1 + 1 \leqslant i \leqslant r_1 + r_2$, where $\bar{\alpha}$ denotes the complex conjugate of the number α. Given a vector $\mathbf{a} = (\alpha_1, \alpha_2, ..., \alpha_n) \in K^n \backslash \{\mathbf{0}\}$, denote by $[\mathbf{a}]$ the fractional ideal generated by the components. We define the *height* of $\mathbf{a}$ to be

$$H(\mathbf{a}) = N([\mathbf{a}])^{-1} \cdot \prod_{i=1}^{d} \|\mathbf{a}^{(i)}\| = N([\mathbf{a}])^{-1} \cdot \prod_{i=1}^{r_1+r_2} \|\mathbf{a}^{(i)}\|^{e_i},$$

where N denotes the norm of the ideal, $\|\cdot\|$ denotes the usual norm on $\mathbb{C}^n$:

$$\|\mathbf{a}\| = \left(\sum_{i=1}^{n} \alpha_i \overline{\alpha_i} \right)^{1/2},$$

and

$$e_i = \begin{cases} 1 & \text{if } i \leqslant r_1 \\ 2 & \text{if } i > r_1. \end{cases}$$

It will be convenient to write

$$H_\infty(\mathbf{a}) = \prod_{i=1}^{r_1+r_2} \|\mathbf{a}^{(i)}\|^{e_i},$$

so that $H(\mathbf{a}) = N([\mathbf{a}])^{-1} H_\infty(\mathbf{a})$. This height is projective, as one may easily verify. Hence, we define the *height* of the one-dimensional subspace $K\mathbf{a} \subset K^n$ by

$$H(K\mathbf{a}) = H(\mathbf{a}).$$

Now suppose S is an m-dimensional subspace of K^n, where $1 \leqslant m < n$. If $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m$ are a basis for S over K, we can form the wedge product (see [6, Chap. 7])

$$\mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \cdots \wedge \mathbf{a}_m \in K^{\binom{n}{m}},$$

where $\binom{n}{m}$ is the binomial coefficient, $n!/m!(n-m)!$. Also, we have

$$\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m \in K^*(\beta_1 \wedge \beta_2 \wedge \cdots \wedge \beta_m)$$

if and only if the α's and the β's span the same subspace. (For proofs, see [6].) Thus, the wedge product will give a one-to-one (though not generally onto) mapping of m-dimensional subspaces of K^n to one-dimensional subspaces of $K^{\binom{n}{m}}$. Any such $\mathbf{x} = \alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_m$ is called a set of *Grassmann coordinates* of S. Such an $\mathbf{x}$ is determined up to a scalar multiple. We define the *height* of the subspace S to be

$$H(S) = H(\mathbf{x}).$$

Also, define

$$H(\{\mathbf{0}\}) = H(K^n) = 1.$$

Finally, we introduce some notation from [12] which we will use. For $\mathbf{X} \in \mathbb{R}^{nr_1} \oplus \mathbb{C}^{2nr_2}$ we write

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_d),$$

where

$$\mathbf{x}_i \in \begin{cases} \mathbb{R}^n & \text{for } 1 \leqslant i \leqslant r_1, \\ \mathbb{C}^n & \text{for } r_1 < i \leqslant d. \end{cases}$$

Let $\mathbb{E}^{nd} \subset \mathbb{R}^{nr_1} \oplus \mathbb{C}^{2nr_2}$ be given by the set of points satisfying $\mathbf{x}_{i+r_2} = \overline{\mathbf{x}_i}$ for $r_1 < i \leqslant r_1 + r_2$. For $\mathbf{X}$ and $\mathbf{Y}$ in $\mathbb{E}^{nd}$, we define the inner product of $\mathbf{X}$ and $\mathbf{Y}$ to be $\mathbf{X} \cdot \bar{\mathbf{Y}}$, the usual inner product in $\mathbb{C}^{nd}$. Thus, for $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_d)$ and $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_d)$ in $\mathbb{E}^{nd}$,

$$\mathbf{X} \cdot \bar{\mathbf{Y}} = \sum_{i=1}^{r_1} \mathbf{x}_i \cdot \mathbf{y}_i + \sum_{i=r_1+1}^{r_1+2r_2} \mathbf{x}_i \cdot \overline{\mathbf{y}_i}.$$

One easily verifies that, under this inner product, $\mathbb{E}^{nd}$ is a Euclidean vector space of dimension nd. In particular, $\mathbf{X} \cdot \bar{\mathbf{Y}}$ is real and $\mathbf{X} \cdot \bar{\mathbf{Y}} = \mathbf{Y} \cdot \bar{\mathbf{X}}$.

We embed K^n into $\mathbb{E}^{nd}$ as follows. Let $\rho: K^n \to \mathbb{E}^{nd}$ be defined by

$$\rho(\alpha) = (\alpha^{(1)}, \alpha^{(2)}, ..., \alpha^{(d)}),$$

where the maps $\alpha \mapsto \alpha^{(i)}$ denote the embeddings of K into $\mathbb{C}$, ordered as above.

For S a subspace of K^n, $S^{(i)} = \{\alpha^{(i)} : \alpha \in S\}$ will be a subspace of

$$(K^{(i)})^n \subset \begin{cases} \mathbb{R}^n & \text{if } 1 \leqslant i \leqslant r_1, \\ \mathbb{C}^n & \text{if } r_1 < i \leqslant d. \end{cases}$$

For a subspace $V \subseteq \mathbb{R}^n$, let $V^\perp$ be its orthogonal complement. Similarly, for V a subspace of $\mathbb{C}^n$, let $V^\perp$ be the orthogonal complement

$$V^\perp = \{\mathbf{x} \in \mathbb{C}^n : \mathbf{x} \cdot \bar{\mathbf{y}} = 0 \text{ for all } \mathbf{y} \in V\}.$$

Let $\pi^{(i)}$ be the orthogonal projection from $\mathbb{R}^n$ or $\mathbb{C}^n$ onto $(S^{(i)})^\perp$ when $1 \leqslant i \leqslant r_1$ or $r_1 < i \leqslant d$, respectively. Define

$$\pi = \pi^{(1)} \times \pi^{(2)} \times \cdots \times \pi^{(d)},$$

so that

$$\pi(\mathbf{X}) = (\pi^{(1)}(\mathbf{x}_1), ..., \pi^{(d)}(\mathbf{x}_d))$$

for all $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_d) \in \mathbb{E}^{nd}$. Note that $\pi \circ \rho$ is linear on K^n and vanishes only on S. We remark that $(S^{(i)})^\perp$ is not necessarily defined over $K^{(i)}$. When we write π we assume the subspace S is given.

2. A Reformulation of Theorem 1

We return to the situation in Theorem 1, where $S \subseteq K^n$ is a fixed subspace of dimension $m > 1$. We think of K^{n-1} as being a subspace of K^n:

$$K^{n-1} = \{\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_{n-1}, 0) \in K^n\}.$$

Without loss of generality, we may assume that $S \not\subseteq K^{n-1}$, so that $V = S \cap K^{n-1}$ is a subspace of dimension $m-1$. We then have

$$S = K(\boldsymbol{\beta}, 1) \oplus V, \tag{2}$$

where $\boldsymbol{\beta} \in K^{n-1}$ is unique modulo V.

Let $\mathfrak{O}$ denote the ring of integers of K. Let $\mathfrak{I}(\tilde{\boldsymbol{\beta}})$ be the ideal

$$\mathfrak{I}(\tilde{\boldsymbol{\beta}}) = \{a \in K : a\boldsymbol{\beta} \in V + \mathfrak{O}^{n-1}\}.$$

For $\boldsymbol{\alpha} \in \tilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + V$, define

$$\mathfrak{I}(\boldsymbol{\alpha}) = \{a \in K : a\boldsymbol{\alpha} \in \mathfrak{O}^{n-1}\}.$$

For any ideal $\mathfrak{I}$, let $\mathfrak{I}^* = \mathfrak{I} \cap \mathfrak{O}$. Certainly, we have

$$\mathfrak{I}(\boldsymbol{\alpha}) \subseteq \mathfrak{I}(\tilde{\boldsymbol{\beta}}), \qquad \mathfrak{I}^*(\boldsymbol{\alpha}) \subseteq \mathfrak{I}^*(\tilde{\boldsymbol{\beta}}). \tag{3}$$

For such an $\boldsymbol{\alpha}$, we get a unique one-dimensional subspace $T = K(\boldsymbol{\alpha}, 1) \subset S$, and all such one-dimensional subspaces $T \subset S$, $T \not\subseteq V$ arise in this way. Since $\mathfrak{I}^*(\boldsymbol{\alpha}) = [(\boldsymbol{\alpha}, 1)]^{-1}$, we have the height $H(T) = N(\mathfrak{I}^*(\boldsymbol{\alpha})) \cdot H_\infty(\boldsymbol{\alpha}, 1)$.

Evidently, to prove Theorem 1 it will suffice to count the number of such $\alpha \in \tilde{\beta}$ with $N(\mathfrak{I}^*(\alpha)) \cdot H_\infty(\alpha, 1) \leqslant B$.

Define the sets

$$L(\mathfrak{A}, B) = \left\{ \alpha \in \tilde{\beta} : \mathfrak{I}(\alpha) \supseteq \mathfrak{A} \text{ and } H_\infty(\alpha, 1) \leqslant \frac{B}{N(\mathfrak{A})} \right\}$$

and

$$\bar{L}(\mathfrak{A}, B) = \{ \alpha \in L(\mathfrak{A}, B) : \mathfrak{I}^*(\alpha) = \mathfrak{A} \},$$

where $\tilde{\beta}$ is as above and $\mathfrak{A}$ is any fractional ideal. Note that $\bar{L}(\mathfrak{A}, B)$ is empty if $\mathfrak{A} \not\subseteq \mathfrak{I}^*(\tilde{\beta})$ by (3). Denote the cardinalities of $L(\mathfrak{A}, B)$ and $\bar{L}(\mathfrak{A}, B)$ by $\lambda(\mathfrak{A}, B)$ and $\bar{\lambda}(\mathfrak{A}, B)$, respectively. We will compute $\lambda(\mathfrak{A}, B)$ and recapture $\bar{\lambda}(\mathfrak{A}, B)$ with an inversion.

Let μ be the Möbius function on ideals:

$$\mu(\mathfrak{O}) = 1,$$

$$\mu(\mathfrak{P}^v) = \begin{cases} -1 & \text{if } \mathfrak{P} \text{ is a prime ideal and } v = 1, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mu(\mathfrak{A}\mathfrak{B}) = \mu(\mathfrak{A})\,\mu(\mathfrak{B}) \qquad \text{if } \mathfrak{A} \text{ and } \mathfrak{B} \text{ are relatively prime.}$$

One easily verifies that, as in the case for the rational integers (see [4]),

$$\sum_{\mathfrak{B} \mid \mathfrak{I}} \mu(\mathfrak{B}) = \begin{cases} 1 & \text{if } \mathfrak{I} = \mathfrak{O}, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

LEMMA 1. *For $\mathfrak{A}$ integral,*

$$\bar{\lambda}(\mathfrak{A}, B) = \sum_{\mathfrak{C} \mid \mathfrak{A}} \mu(\mathfrak{C})\, \lambda(\mathfrak{A}\mathfrak{C}^{-1}, B/N(\mathfrak{C})).$$

Proof. For $\mathfrak{A}$ integral,

$$\lambda(\mathfrak{A}, B) = \sum_{\mathfrak{C} \mid \mathfrak{A}} \bar{\lambda}\left(\mathfrak{C}, \frac{BN(\mathfrak{C})}{N(\mathfrak{A})}\right).$$

Setting $\mathfrak{B} = \mathfrak{A}\mathfrak{C}^{-1}$ gives

$$\lambda(\mathfrak{A}, B) = \sum_{\mathfrak{B} \mid \mathfrak{A}} \bar{\lambda}(\mathfrak{A}\mathfrak{B}^{-1}, B/N(\mathfrak{B})).$$

Thus,

$$\begin{aligned}
&\sum_{\mathfrak{C}\mid\mathfrak{A}} \mu(\mathfrak{C})\,\lambda(\mathfrak{A}\mathfrak{C}^{-1}, B/N(\mathfrak{C})) \\
&\quad = \sum_{\mathfrak{C}\mid\mathfrak{A}} \sum_{\mathfrak{B}\mid\mathfrak{A}\mathfrak{C}^{-1}} \mu(\mathfrak{C})\,\bar{\lambda}\left(\mathfrak{A}(\mathfrak{C}\mathfrak{B})^{-1}, \frac{B}{N(\mathfrak{C})\,N(\mathfrak{B})}\right) \\
&\quad = \sum_{\mathfrak{I}\mid\mathfrak{A}} \bar{\lambda}(\mathfrak{A}\mathfrak{I}^{-1}, B/N(\mathfrak{I})) \sum_{\mathfrak{C}\mid\mathfrak{I}} \mu(\mathfrak{C}) \\
&\quad = \bar{\lambda}(\mathfrak{A}, B),
\end{aligned}$$

by (4), where $\mathfrak{I} = \mathfrak{B}\mathfrak{C}$.

Now for the sum over ideals $\mathfrak{A} \subseteq \mathfrak{I}^*(\tilde{\boldsymbol{\beta}})$ we have

$$\begin{aligned}
\sum_{\mathfrak{A}} \bar{\lambda}(\mathfrak{A}, B) &= \sum_{\mathfrak{A}} \sum_{\mathfrak{C}\mid\mathfrak{A}} \mu(\mathfrak{C})\,\lambda(\mathfrak{A}\mathfrak{C}^{-1}, B/N(\mathfrak{C})) \qquad \text{letting } \mathfrak{A} = \mathfrak{B}\mathfrak{C} \\
&= \sum_{\mathfrak{C}} \sum_{\mathfrak{B}} \mu(\mathfrak{C})\,\lambda(\mathfrak{B}, B/N(\mathfrak{C})),
\end{aligned}$$

where the first sum is over integral ideals $\mathfrak{C}$ with $N(\mathfrak{C}) \leqslant B$ and the second is over integral ideals $\mathfrak{B} \subseteq \mathfrak{I}^*(\tilde{\boldsymbol{\beta}})$ with $N(\mathfrak{B}) \leqslant B/N(\mathfrak{C})$, since $L(\mathfrak{A}, B)$ is empty if $N(\mathfrak{A}) > B$ (obviously, $H_\infty(\boldsymbol{\alpha}, 1) \geqslant 1$) and $L(\mathfrak{A}\mathfrak{C}^{-1}, B/N(\mathfrak{C}))$ is empty if $\mathfrak{A}\mathfrak{C}^{-1} \not\subseteq \mathfrak{I}^*(\tilde{\boldsymbol{\beta}})$. Thus, to prove our Theorem 1, it suffices to prove the following.

PROPOSITION. *With the definitions above,*

$$\sum_{\mathfrak{C}} \sum_{\mathfrak{B}} \mu(\mathfrak{C})\,\lambda(\mathfrak{B}, B/N(\mathfrak{C})) = a(m, K)\frac{B^m}{H(S)} + O(B^{m-1/d}),$$

where the first sum is over integral ideals $\mathfrak{C}$ with $N(\mathfrak{C}) \leqslant B$ and the second is over integral ideals $\mathfrak{B} \subseteq \mathfrak{I}^*(\tilde{\boldsymbol{\beta}})$ with $N(\mathfrak{B}) \leqslant B/N(\mathfrak{C})$. In the case $m = 2$ and $K = \mathbb{Q}$, the error term above is replaced by $O(B \log B)$.

Let $\tilde{\boldsymbol{\beta}}$ be as above and suppose $\mathfrak{B}$ is any fractional ideal. We will write $\mathfrak{B}^{n-1}$ for the subset of K^{n-1},

$$\underbrace{\mathfrak{B} \times \mathfrak{B} \times \cdots \times \mathfrak{B}}_{n-1 \text{ times}}.$$

Let $\Gamma = (\mathfrak{B}^{-1})^{n-1}$. Then for $\gamma \in \tilde{\boldsymbol{\beta}}$, $\mathfrak{B} \subseteq \mathfrak{I}(\gamma)$ is equivalent to $\gamma \in \Gamma$.

By the definition of $\mathfrak{I}(\tilde{\boldsymbol{\beta}})$, we have $\mathfrak{I}(\tilde{\boldsymbol{\beta}})\,\boldsymbol{\beta} \subseteq V + \mathfrak{O}^{n-1}$, so that $\mathfrak{O}\boldsymbol{\beta} \subseteq V + (\mathfrak{I}(\tilde{\boldsymbol{\beta}})^{-1})^{n-1}$. This implies that there is some $\boldsymbol{\alpha} \in V$ such that $\boldsymbol{\beta} + \boldsymbol{\alpha} \in (\mathfrak{I}(\tilde{\boldsymbol{\beta}})^{-1})^{n-1}$, i.e., such that $\mathfrak{I}(\tilde{\boldsymbol{\beta}}) \subseteq \mathfrak{I}(\boldsymbol{\beta} + \boldsymbol{\alpha})$. Hence, we may assume that

$$\mathfrak{I}(\boldsymbol{\beta}) = \mathfrak{I}(\tilde{\boldsymbol{\beta}}). \tag{5}$$

Suppose $\mathfrak{B} \subseteq \mathfrak{I}(\tilde{\boldsymbol{\beta}})$ and let Γ be as above. Then $\boldsymbol{\beta} \in \Gamma$ by (5). If $\boldsymbol{\alpha} \in V \cap \Gamma$, then $\mathfrak{B} \subseteq \mathfrak{I}(\boldsymbol{\alpha}+\boldsymbol{\beta})$ since both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are in Γ. Conversely, if $\boldsymbol{\gamma} \in \tilde{\boldsymbol{\beta}}$ with $\mathfrak{B} \subseteq \mathfrak{I}(\boldsymbol{\gamma})$, then $\boldsymbol{\gamma}-\boldsymbol{\beta} \in V \cap \Gamma$. With $I(V)$ written for $V \cap \mathfrak{O}^{n-1}$, this shows that $\lambda(\mathfrak{B}, B/N(\mathfrak{C}))$ is the number of $\boldsymbol{\alpha} \in V \cap \Gamma = \mathfrak{B}^{-1} I(V)$ satisfying $H_\infty(\boldsymbol{\alpha}+\boldsymbol{\beta}, 1) \leqslant B/N(\mathfrak{B}\mathfrak{C})$. In other words, $\lambda(\mathfrak{B}, B/N(\mathfrak{C}))$ is the number of lattice points $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_d) \in \rho(\mathfrak{B}^{-1} I(V))$ satisfying

$$\prod_{i=1}^{d} (\|\mathbf{x}_i + \boldsymbol{\beta}^{(i)}\|^2 + 1)^{1/2} \leqslant \frac{B}{N(\mathfrak{B}\mathfrak{C})}.$$

Let π be the projection in Section 1 defined with respect to the subspace V, and define

$$a_i = \sqrt{\|\pi_i(\boldsymbol{\beta}^{(i)})\|^2 + 1} \qquad (1 \leqslant i \leqslant d).$$

By [12, Theorem 3]

$$H(S) = H(V)\, N(\mathfrak{I}^*(\tilde{\boldsymbol{\beta}})) \prod_{i=1}^{d} a_i = H(V)\, N(\mathfrak{I}^*(\tilde{\boldsymbol{\beta}})) \prod_{i=1}^{r_1+r_2} a_i^{e_i}. \tag{6}$$

Then $\lambda(\mathfrak{B}, B/N(\mathfrak{C}))$ is the number of lattice points $\mathbf{X} \in \rho(\mathfrak{B}^{-1} I(V))$ in some fixed (depending on $\boldsymbol{\beta}$) translation of the domain

$$\left\{ \mathbf{X} \in \mathbb{E}^{(n-1)d} : \prod_{i=1}^{d} (\|\mathbf{x}_i\|^2 + a_i^2)^{1/2} \leqslant B/N(\mathfrak{B}\mathfrak{C}) \right\}.$$

In order to directly apply the results of Section 5, we embed $\mathbb{E}^{(n-1)d}$ into $\mathbb{R}^{(n-1)d}$ as follows. Let

$$t: \mathbb{E}^{(n-1)d} \to \mathbb{R}^{(n-1)d}$$

be defined by

$$\begin{aligned} t(\mathbf{X}) &= t((\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d)) \\ &= (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{r_1}, \mathbf{x}'_{r_1+1}, \mathbf{x}'_{r_1+2}, \ldots, \mathbf{x}'_{r_1+r_2}), \end{aligned}$$

where, for $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{i(n-1)})$ and $i = r_1+1, r_1+2, \ldots, r_1+r_2$, we define

$$\mathbf{x}'_i = (\mathrm{Re}(x_{i1}), \mathrm{Im}(x_{i1}), \ldots, \mathrm{Re}(x_{i(n-1)}), \mathrm{Im}(x_{i(n-1)})).$$

One sees that the determinant of t is $2^{-r_2(n-1)}$. Denote the composition of t and ρ by t'.

For $\mathbf{Y} \in \mathbb{R}^{(n-1)d}$ we write

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{r_1+r_2}),$$

where

$$\mathbf{y}_i \in \begin{cases} \mathbb{R}^{n-1} & \text{for } 1 \leqslant i \leqslant r_1, \\ \mathbb{R}^{2(n-1)} & \text{for } r_1+1 \leqslant i \leqslant r_1+r_2. \end{cases}$$

For $0 \leqslant l \leqslant (m-1)d$ and $x \in \mathbb{R}^+$, define $V(x, l, B)$ to be the sum of the l-dimensional volumes of the projections of the domain

$$D(x/B) = \left\{ \mathbf{Y} \in \mathbb{R}^{(m-1)d} : \prod_{i=1}^{r_1+r_2} (\|\mathbf{y}_i\|^2 + a_i^2)^{e_i/2} \leqslant B/x \right\}$$

on the various coordinate axes obtained by equating $(m-1)d-l$ coordinates to zero.

LEMMA 2. *For $\mathfrak{B} \subseteq \mathfrak{I}(\tilde{\boldsymbol{\beta}})$ and $\mathfrak{C}$ any ideal,*

$$\left| \lambda(\mathfrak{B}, B/N(\mathfrak{C})) - \frac{V(N(\mathfrak{B}\mathfrak{C}), (m-1)d, B)}{H(V)} \left(\frac{N(\mathfrak{B})\, 2^{r_2}}{\sqrt{|\delta|}} \right)^{m-1} \right|$$
$$\ll \sum_{i=0}^{m-2} \sum_{j=0}^{d-1} V(N(\mathfrak{B}\mathfrak{C}), id+j, B)\, N(\mathfrak{B})^{i+j/d},$$

where the constant implicit in the $\ll$ notation depends only on m and K.

Proof. The domain $D(x)$ is certainly a coordinate domain (see Section 5). After a unitary transformation τ of $\mathbb{R}^{(m-1)d}$, $\lambda(\mathfrak{B}, B/N(\mathfrak{C}))$ is the number of lattice points $\mathbf{Y} \in \tau \circ t'(\mathfrak{B}^{-1} I(V))$ in $D(N(\mathfrak{B}\mathfrak{C})/B)$. The determinant of this lattice is

$$H(V) \left(\frac{\sqrt{|\delta|}}{2^{r_2} N(\mathfrak{B})} \right)^{m-1}.$$

(See [12, Lemma 1].) Now by [12, Lemma 10]

$$(\det(\rho(\mathfrak{B}^{-1} I(V))))^{-[(m-1-i)d-j]})^{-1} \ll N(\mathfrak{B})^{i+j/d}.$$

The lemma follows from Theorem 5 (and the remark following it).

3. PROOF OF THE PROPOSITION

We estimate

$$\sum_{\mathfrak{B}} N(\mathfrak{B})^{m-1}\, V(N(\mathfrak{B}\mathfrak{C}), (m-1)d, B), \tag{7}$$

where the sum is over integral ideals $\mathfrak{B} \subseteq \mathfrak{I}^*(\tilde{\boldsymbol{\beta}})$ with $N(\mathfrak{B}) \leqslant B/N(\mathfrak{C}) = M$. By the Dedekind–Weber theorem (see [5]), the number of such ideals is

$$\frac{h\chi M}{N(\mathfrak{I}^*(\tilde{\boldsymbol{\beta}}))} + O(M^{1-1/d}),$$

where

$$\chi = \frac{2^{r_1+r_2}\pi^{r_2}R}{w\sqrt{|\delta|}}.$$

If we set $u_i = \|\mathbf{y}_i\|$ for $1 \leqslant i \leqslant r_1 + r_2$, the volume of the domain $D(x)$ is

$$((m-1)\,V(m-1))^{r_1}\,(2(m-1)\,V(2(m-1)))^{r_2}$$
$$\times \int \cdots \int_{A(x)} \prod_{i=1}^{r_1+r_2} u_i^{(m-1)e_i - 1}\,du_i,$$

where the domain of integration $A(x)$ is given by

$$\prod_{i=1}^{r_1+r_2} (u_i^2 + a_i^2)^{e_i/2} \leqslant 1/x, \qquad u_i \geqslant 0.$$

Using summation by parts (see [12, Lemma 12]), the sum (7) is

$$\frac{h\chi 2^{r_2}(m-1)^{r_1+r_2}\,(V(m-1))^{r_1}\,(V(2(m-1)))^{r_2}}{N(\mathfrak{I}^*(\tilde{\boldsymbol{\beta}}))}$$
$$\times \int_0^M x^{m-1} \int \cdots \int_{A(x/M)} \prod_{i=1}^{r_1+r_2} u_i^{(m-1)e_i - 1}\,du_i\,dx$$
$$+ O\left(\int_0^2 x^{m-1} \int \cdots \int_{A(x/M)} \prod_{i=1}^{r_1+r_2} u_i^{(m-1)e_i - 1}\,du_i\,dx\right)$$
$$+ O\left(\int_1^M x^{m-1-1/d} \int \cdots \int_{A(x/M)} \prod_{i=1}^{r_1+r_2} u_i^{(m-1)e_i - 1}\,du_i\,dx\right), \qquad (8)$$

where the second error term is not needed if $d = 1$. We compute

$$\int_0^M x^{m-1} \int \cdots \int_{A(x/M)} \prod_{i=1}^{r_1+r_2} u_i^{(m-1)e_i - 1}\,du_i\,dx$$
$$= M^m \int_0^1 y^{m-1} \int \cdots \int_{A(y)} \prod_{i=1}^{r_1+r_2} u_i^{(m-1)e_i - 1}\,du_i\,dy$$
$$= M^m \cdots \int_0^\infty u_i^{(m-1)e_i - 1} \cdots \int_0^{\prod_{i=1}^{r_1+r_2}(u_i^2 + a_i^2)^{-e_i/2}} y^{m-1}\,dy \prod_{i=1}^{r_1+r_2} du_i$$

$$= \frac{M^m}{m} \prod_{i=1}^{r_1+r_2} \left(\int_0^\infty u_i^{(m-1)e_i - 1} (u_i^2 + a_i^2)^{-me_i/2} \, du_i \right)$$

$$= \frac{M^m}{m} \prod_{i=1}^{r_1+r_2} a_i^{-e_i} \left(\int_0^\infty v^{(m-1)-1} (v^2+1)^{-m/2} \, dv \right)^{r_1}$$

$$\times \left(\int_0^\infty v^{(m-1)2-1} (v_2+1)^{-m} \, dv \right)^{r_2}$$

$$= \frac{M^m}{m} \prod_{i=1}^{r_1+r_2} a_i^{-e_i} \left(\frac{mV(m)}{(m-1)\,V(m-1)\,2} \right)^{r_1}$$

$$\times \left(\frac{2mV(2m)}{2(m-1)\,V(2(m-1))\,2\pi} \right)^{r_2}.$$

Straightforward estimates show the error terms in (8) are $O(M^{m-1/d})$. Thus, by (6) the sum (7) is

$$\frac{a(m,K)\,\zeta_K(m)\,H(V)}{H(S)} \left(\frac{\sqrt{|\delta|}}{2^{r_2}} \right)^{m-1} \left(\frac{B}{N(\mathfrak{C})} \right)^m + O\left(\frac{B}{N(\mathfrak{C})} \right)^{m-1/d}.$$

Now

$$\sum_{\mathfrak{C}} \mu(\mathfrak{C})\, N(\mathfrak{C})^{-m} = \frac{1}{\zeta_K(m)} + O(B^{1-m}),$$

where the sum is over integral ideals $\mathfrak{C}$ with norm $\leqslant B$. Similarly, unless $m=2$ and $d=1$, we get

$$\sum_{\mathfrak{C}} \mu(\mathfrak{C})\, N(\mathfrak{C})^{1/d-m} = O(1),$$

and in the case $m=2$ and $d=1$ we get

$$\sum_{\mathfrak{C}} \mu(\mathfrak{C})\, N(\mathfrak{C})^{-1} = O(\log B).$$

As for the error terms in Lemma 2, we may compute as above and show that, unless $m=2$ and $d=1$,

$$\sum_{\mathfrak{C}} \sum_{\mathfrak{B}} \mu(\mathfrak{C})\, V(N(\mathfrak{B}\mathfrak{C}), id+j, B)\, N(\mathfrak{B})^{i+j/d} \ll B^{m-1/d}$$

for any $0 \leqslant i \leqslant m-2$ and $0 \leqslant j \leqslant d-1$. In the case $m=2$ and $d=1$, the error term from Lemma 2 is

$$\sum_{\mathfrak{C}} \sum_{\mathfrak{B}} \mu(\mathfrak{C}) \ll B \log B.$$

Thus, our proof of the proposition, and hence Theorem 1, is complete.

4. The Inhomogeneous Case

Consider the system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} &= b_1 \\ &\vdots \\ a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ln-1}x_{n-1} &= b_l, \end{aligned} \tag{9}$$

where the a_{ij}'s and b_i's are elements of K (not all $b_i = 0$). We will assume the equations are linearly independent, so that the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} - b_1x_n &= 0 \\ &\vdots \\ a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ln-1}x_{n-1} - b_lx_n &= 0 \end{aligned}$$

defines a subspace $S \subseteq K^n$ of dimension $m = n - l$. Solutions to (9) correspond to points of the form $(\boldsymbol{\alpha}, 1) \in S$, where $\boldsymbol{\alpha} \in K^{n-1}$. We will write $S = K(\boldsymbol{\beta}, 1) \oplus V$ as above in (2). Solutions to the system (9) are then of the form $\mathbf{x} = \boldsymbol{\alpha} + \boldsymbol{\beta}$, where $\boldsymbol{\alpha} \in V$. For such a solution we define the *inhomogeneous height* $H_I(\mathbf{x}) = H(\mathbf{x}, 1)$. We will count integral solutions to (9) with inhomogeneous height $\leqslant B$. We may assume that $m > 1$, since this case is trivial.

Theorem 2. *Using the notation above, the system of inhomogeneous linear equations* (9) *has an integral solution if and only if* $\mathfrak{D} \subseteq \mathfrak{I}(\tilde{\boldsymbol{\beta}})$. *If this is the case, then the number of integral solutions* $\boldsymbol{\alpha}$ *with* $H_I(\boldsymbol{\alpha}) \leqslant B$ *is asymptotically*

$$\frac{b(m-1, K)}{H(V)} B^{m-1}(\log B)^{r_1+r_2-1} + O(B^{m-1}(\log B)^{r_1+r_2-2})$$

as $B \to \infty$, *with an error term of* $O(B^{m-2})$ *if* $r_1 + r_2 = 1$. *Here the constants*

implicit in the O notation depend on K, m, and $H(S)/H(V)$. The constant $b(m-1, K)$ is defined to be

$$\frac{(m-1)^{r_1+r_2-1}}{r_1+r_2-1}(V(m-1))^{r_1}(V(2m-2))^{r_2}\left(\frac{2^{r_2}}{\sqrt{|\delta|}}\right)^{m-1}$$

if $r_1+r_2>1$, and the term r_1+r_2-1 in the denominator is replaced by 1 if $r_1+r_2=1$.

Note that $\mathfrak{I}(\mathfrak{a})\supseteq\mathfrak{O}$ is equivalent to $\mathfrak{a}$ being integral. An easy consequence of this together with (3) and (5) gives us the first statement in Theorem 2. Also, if $\mathfrak{I}(\mathfrak{a})\supseteq\mathfrak{O}$, then $\mathfrak{I}^*(\mathfrak{a})=\mathfrak{O}$ and $H_I(\mathfrak{a})=H(\mathfrak{a},1)=H_\infty(\mathfrak{a},1)$.

We will now assume that $\mathfrak{O}\subseteq\mathfrak{I}(\tilde{\boldsymbol{\beta}})$. The number of integral solutions to (9) with inhomogeneous height less than or equal to B is then simply $\lambda(\mathfrak{O}, B)$. By Lemma 2 we have the estimate

$$\left|\lambda(\mathfrak{O},B)-\frac{V(1,(m-1)d,B)}{H(V)}\left(\frac{2^{r_2}}{\sqrt{|\delta|}}\right)^{m-1}\right| \ll \sum_{i=0}^{m-2}\sum_{j=0}^{d-1} V(1,id+j,B). \tag{10}$$

LEMMA 3. *For $B\geqslant H(S)/H(V)$,*

$$V(1,(m-1)d,B)=\begin{cases} b(m-1,K)(\sqrt{|\delta|}/2^{r_2})^{m-1}B^{m-1}(\log B)^{r_1+r_2-1} \\ \quad +O(B^{m-1}(\log B)^{r_1+r_2-2}) \quad \text{if } r_1+r_2>1, \\ b(m-1,K)(\sqrt{|\delta|}/2^{r_2})^{m-1}B^{m-1}+O(B^{m-2}) \\ \quad \text{if } r_1+r_2=1. \end{cases}$$

Here the constant implicit in the O notation depends on r_1+r_2, m, and $H(S)/H(V)$.

Proof. As was shown in Section 3,

$$V(1,(m-1)d,B)=((m-1)V(m-1))^{r_1}(2(m-1)V(2(m-1)))^{r_2} \times\int\cdots\int\prod_{i=1}^{r_1+r_2}u_i^{(m-1)e_i-1}\,du_i,$$

where the domain of integration is given by

$$\prod_{i=1}^{r_1+r_2}(u_i^2+a_i^2)^{e_i/2}\leqslant B, \qquad u_i\geqslant 0$$

and the a_i's are defined as above in Section 2. Note that

$$\frac{H(S)}{H(V)} = \prod_{i=1}^{r_1+r_2} a_i^{e_i}$$

by (6). Setting $w_i = u_i/a_i$ gives

$$V(1, (m-1)d, B) = 2^{r_2}(m-1)^{r_1+r_2}\,(V(m-1))^{r_1}\,(V(2(m-1)))^{r_2} \times \prod_{i=1}^{r_1+r_2} a_i^{(m-1)e_i} \int \cdots \int \prod_{i=1}^{r_1+r_2} w_i^{(m-1)e_i-1}\, dw_i,$$

where the domain of integration is given by

$$\prod_{i=1}^{r_1+r_2} (w_i^2+1)^{e_i/2} \leqslant \frac{B}{\prod_{i=1}^{r_1+r_2} a_i^{e_i}}, \qquad w_i \geqslant 0.$$

Set $v_i = w_i^2 + 1$ and

$$X = \frac{B}{\prod_{i=1}^{r_1+r_2} a_i^{e_i}}.$$

A routine induction argument on r_1+r_2 shows that for $X \geqslant 1$ and $r_1+r_2 > 1$,

$$\int \cdots \int \prod_{i=1}^{r_1+r_2} (v_i-1)^{(e_i(m-1)-2)/2}\, dv_i = \frac{2^{r_1} X^{m-1} (\log X)^{r_1+r_2-1}}{(m-1)(r_1+r_2-1)} + O(X^{m-1}(\log X)^{r_1+r_2-2}),$$

where the domain of integration is given by

$$\prod_{i=1}^{r_1+r_2} v_i^{e_i} \leqslant X^2, \qquad v_i \geqslant 1.$$

If $r_1+r_2 = 1$, then the r_1+r_2-1 in the denominator of the main term above is replaced by 1 and the error term is replaced by $O(X^{m-2})$. This proves the lemma.

Similar arguments show that

$$V(1, id+j, B) \ll \begin{cases} B^{m-1}(\log B)^{r_1+r_2-2} & \text{if } r_1+r_2 > 1, \\ B^{m-2} & \text{if } r_1+r_2 = 1, \end{cases} \tag{11}$$

for $i < m-1$ and $j < d$. Theorem 2 then follows from Lemma 3, (10), and (11).

Our methods can be used to prove results that are more precise than that of Theorem 2. For example, we can also give estimates for the number of "primitive" integral solutions as well (solutions $\mathbf{x}$ with $[\mathbf{x}] = \mathfrak{O}$).

THEOREM 3. *Using the notation above, suppose that* $\mathfrak{O} \subseteq \mathfrak{A} \subseteq \mathfrak{I}(\tilde{\boldsymbol{\beta}})$. *The number of solutions* $\boldsymbol{\alpha}$ *to the system of linear equations* (9) *with* $\mathfrak{I}(\boldsymbol{\alpha}) = \mathfrak{A}$ *and* $H_I(\boldsymbol{\alpha}) \leqslant B$ *is asymptotically*

$$\frac{b(m-1, K)}{H(V)} (N(\mathfrak{A})\, B)^{m-1} (\log B)^{r_1+r_2-1} \sum_{\mathfrak{C}} \mu(\mathfrak{C})\, N(\mathfrak{C})^{1-m} + O(B^{m-1}(\log B)^{r_1+r_2-2})$$

as $B \to \infty$, *where the sum is over integral ideals* $\mathfrak{C} \mid \mathfrak{A}\mathfrak{I}(\tilde{\boldsymbol{\beta}})^{-1}$ *and the constants implicit in the O notation depend on* $H(S)/H(V)$, K, m, *and the ideals* $\mathfrak{A}$ *and* $\mathfrak{I}(\tilde{\boldsymbol{\beta}})$. *If* $r_1 + r_2 = 1$ *then the error term is replaced by* $O(B^{m-2})$.

Proof. For $\mathfrak{A}$ as in the statement of the theorem, let

$$L^*(\mathfrak{A}, B) = \{\boldsymbol{\alpha} \in \tilde{\boldsymbol{\beta}} : \mathfrak{I}(\boldsymbol{\alpha}) = \mathfrak{A} \text{ and } H_\infty(\boldsymbol{\alpha}, 1) \leqslant B\}$$

and denote the cardinality of this set by $\lambda^*(\mathfrak{A}, B)$. The theorem then estimates this cardinality. For $\mathfrak{O} \subseteq \mathfrak{A} \subseteq \mathfrak{I}(\tilde{\boldsymbol{\beta}})$ we have

$$\lambda(\mathfrak{A}, B) = \sum_{\mathfrak{B}} \lambda^*(\mathfrak{A}\mathfrak{B}^{-1}, B/N(\mathfrak{A})),$$

where the sum is over integral ideals $\mathfrak{B} \mid \mathfrak{A}\mathfrak{I}(\tilde{\boldsymbol{\beta}})^{-1}$. Thus,

$$\begin{aligned}\sum_{\mathfrak{C}} \mu(\mathfrak{C})\, \lambda(\mathfrak{A}\mathfrak{C}^{-1}, BN(\mathfrak{A})/N(\mathfrak{C})) &= \sum_{\mathfrak{C}} \mu(\mathfrak{C}) \sum_{\mathfrak{B}} \lambda^*(\mathfrak{A}(\mathfrak{B}\mathfrak{C})^{-1}, B) \\ &= \sum_{\mathfrak{I}} \lambda^*(\mathfrak{A}\mathfrak{I}^{-1}, B) \sum_{\mathfrak{C} \mid \mathfrak{I}} \mu(\mathfrak{C}) \\ &= \lambda^*(\mathfrak{A}, B),\end{aligned}$$

by (4), where the first and fourth sums above are over integral ideals containing $\mathfrak{A}\mathfrak{I}(\tilde{\boldsymbol{\beta}})^{-1}$ and the third is over integral ideals $\mathfrak{B}$ containing $\mathfrak{A}(\mathfrak{C}\mathfrak{I}(\tilde{\boldsymbol{\beta}}))^{-1}$. Theorem 3 then follows from Lemmas 2 and 3 and (11).

5. The Number of Lattice Points in a Bounded Domain

A fundamental problem in Diophantine approximation is to estimate the number of lattice points in a given domain in $\mathbb{R}^n$. Often the domain will be unbounded, but in many instances the domain considered will be bounded;

for example, in estimating the number of integral ideals of bounded norm in a given ideal class (see [5, Theorem 121]). If we let Λ be an n-dimensional lattice in $\mathbb{R}^n$, i.e., a $\mathbb{Z}$-module of rank n, and $D \subset \mathbb{R}^n$ be bounded, we have

$$\operatorname{card}\{\Lambda \cap D\} \approx \frac{\operatorname{Vol}(D)}{\det(\Lambda)},$$

where "card" denotes cardinality, $\operatorname{Vol}(D)$ the volume of D, and $\det(\Lambda)$ the determinant of Λ. The main question is what one means by "approximately." In the case of a homogeneously expanding domain with a "nice" boundary, one has

$$\left|\operatorname{card}\{\Lambda \cap tD\} - t^n \frac{\operatorname{Vol}(D)}{\det(\Lambda)}\right| = O(t^{n-1}),$$

where the constant implicit in the O notation depends on n, Λ, and the boundary of D. (See, for example, p. 128 of [7].) However, it may very well turn out that such a bound is inadequate. One may need an error term or terms which are more explicitly given in terms of the domain and the lattice.

If one is only concerned with counting *integral* points in a bounded domain, the following theorem due to H. Davenport gives an explicit formulation for the error term.

THEOREM (Davenport [3]). *Let D be a compact domain in $\mathbb{R}^n$ satisfying the following two conditions:*

(1) *Any line parallel to a coordinate axis intersects D in a set of points which, if not empty, is the union of at most s intervals. (A point is considered an interval.)*

(2) *The same is true (with m in place of n) for any of the m-dimensional regions obtained by projecting D onto one of the coordinate spaces defined by equating $n-m$ of the coordinates to 0; and this condition is satisfied for all $m = 1, 2, \ldots, n-1$.*

Then

$$|\operatorname{card}\{D \cap \mathbb{Z}^n\} - \operatorname{Vol}(D)| \leqslant \sum_{m=0}^{n-1} s^{n-m} V_m(D),$$

where $V_m(D)$ is the sum of the m-dimensional volumes of the projections of D on the various coordinate spaces obtained by equating $n-m$ coordinates to zero if $m > 0$, and $V_0 = 1$.

Now the above result may be applied to counting arbitrary lattice points as follows. Let ψ be a linear transformation of $\mathbb{R}^n$ taking Λ to $\mathbb{Z}^n$. Specifically, let $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$ be a basis for Λ and let A be the $n \times n$ matrix with columns $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$. We let ψ be the linear transformation represented by multiplication by A^{-1}. Then

$$\operatorname{card}\{D \cap \Lambda\} = \operatorname{card}\{\psi(D) \cap \mathbb{Z}^n\}.$$

We apply Davenport's result to $\psi(D)$. We have

$$\operatorname{Vol}(\psi(D)) = \frac{\operatorname{Vol}(D)}{\det(\Lambda)}.$$

Now consider the terms $V_m(\psi(D))$. Let σ be some m-element subset of the numbers 1 through n. Let V_σ be the space spanned by the vectors $\boldsymbol{\beta}_i$ for $i \in \sigma$, and let V_{σ^*} be the space spanned by the remaining vectors $\boldsymbol{\beta}_i$. A typical summand of $V_m(\psi(D))$ will be the m-dimensional volume of the image under ψ of the set

$$S_\sigma = \{\mathbf{x} \in V_\sigma : \mathbf{x} + \mathbf{y} \in D \text{ for some } \mathbf{y} \in V_{\sigma^*}\}.$$

We have

$$\operatorname{Vol}(\psi(S_\sigma)) = \frac{\operatorname{Vol}(S_\sigma)}{\det(\Lambda_\sigma)},$$

where Λ_σ is the sublattice of Λ spanning V_σ, since ψ takes Λ_σ to $\mathbb{Z}^m$.

What we have then is the following statement of Davenport's theorem for all lattices in $\mathbb{R}^n$:

THEOREM 0. *Let D be a compact domain in $\mathbb{R}^n$ such that any line intersects D in a set of points which, if not empty, consists of the union of at most s intervals. Let Λ be an n-dimensional lattice in $\mathbb{R}^n$ and let $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$ be any basis for Λ. Then, using the notation above, we have*

$$|\operatorname{card}\{D \cap \Lambda\} - \operatorname{Vol}(D)/\det(\Lambda)| \leqslant \sum_{m=0}^{n-1} s^{n-m} \sum_{\sigma} \frac{\operatorname{Vol}(S_\sigma)}{\det(\Lambda_\sigma)},$$

where the inner sum is over m-element subsets σ of the integers 1 through n.

The result above will most likely be useful when one is working with a particular lattice and domain which one can study. It proves useful, however, to have a result in which the error terms are given by more "generic" properties of the lattice and domain. This is the case, for instance, if one has a sum over many lattices. Our first goal is to give such a result.

Let Λ be an l-dimensional lattice in $\mathbb{R}^n$ spanning a subspace V. Let $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_l$ be the successive minima of Λ with respect to the unit ball in V. Pick linearly independent points $\mathbf{y}_i \in \Lambda$ (the choice is not necessarily unique) satisfying

$$|\mathbf{y}_i| = \lambda_i \qquad \text{for} \quad i = 1, 2, \ldots, l.$$

Define

$$\Lambda^{-i} = \Lambda \cap \bigoplus_{j=1}^{l-i} \mathbb{R}\mathbf{y}_j \qquad \text{for} \quad i = 1, 2, \ldots, l-1$$

and

$$\Lambda^{-l} = \{\mathbf{0}\}.$$

Minkowski's second convex bodies theorem asserts that

$$\frac{2^l}{l!} \det(\Lambda) \leqslant \lambda_1 \lambda_2 \cdots \lambda_l V(l) \leqslant 2^l \det(\Lambda),$$

where $V(l)$ denotes the volume of the unit ball in $\mathbb{R}^l$. Since the successive minima of Λ^{-i} are, by construction, $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_{l-i}$ for $i < l$, we have $\det(\Lambda^{-i})$ is minimal among sublattices of Λ of dimension $l-i$.

Given a domain D as above and a subspace $V \subset \mathbb{R}^n$, let $D(V)$ be the orthogonal projection of D onto V. We will write $\mathrm{Vol}(D(V))$ for the m-dimensional volume of $D(V)$, where the dimension of V is $m > 0$. Let $\mathrm{Vol}(\{\mathbf{0}\}) = 1$.

Finally, for $0 \leqslant m \leqslant n$, let

$$\widetilde{V_m(D)} = \max_V \{\mathrm{Vol}(D(V))\},$$

where the maximum is over all m-dimensional subspaces V.

THEOREM 4. *Let $D \subset \mathbb{R}^n$ be a compact domain such that any line intersects D in a set of points which, if not empty, consists of the union of at most s intervals. Let Λ be an n-dimensional lattice in $\mathbb{R}^n$. Then*

$$|\mathrm{card}\{D \cap \Lambda\} - \mathrm{Vol}(D)/\det(\Lambda)|$$
$$\leqslant \sum_{m=0}^{n-1} s^{n-m} (mc(\Lambda))^m \binom{n}{m} \frac{\widetilde{V_m(D)}}{\det(\Lambda^{-[n-m]})},$$

where $c(\Lambda)$ is a constant depending only on Λ, and one may take

$$c(\Lambda) \leqslant \frac{n!\, 2^n}{V(n)}.$$

Proof. In comparing the error terms in the statements of Theorems 0 and 4, we immediately have

$$\frac{1}{\det(\Lambda_\sigma)} \leqslant \frac{1}{\det(\Lambda^{[n\ m]})}. \tag{12}$$

We will concentrate on the volumes $\mathrm{Vol}(S_\sigma)$. We do this by comparing S_σ to $D(V_{\sigma^*}^\perp)$. Now if the basis chosen for the lattice is orthogonal, then these two sets (and their volumes) will be equal. But lattices do not have orthogonal bases, in general. So we need a basis for the lattice which is as close to being orthogonal as possible, i.e., a basis such that the product of their lengths is close to the determinant of the lattice.

Let $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ be the successive minima of Λ with respect to the unit ball in $\mathbb{R}^n$. By Cassels [2], there is a basis, $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$, of Λ satisfying

$$\lambda_i \leqslant |\mathbf{b}_i| \leqslant i\lambda_i$$

for all i. By Minkowski's theorem, we have

$$\prod_{i=1}^{n} |\mathbf{b}_i| \leqslant n! \prod_{i=1}^{n} \lambda_i \leqslant \frac{n!\, 2^n}{V(n)} \det(\Lambda) = \frac{n!\, 2^n}{V(n)} |\mathbf{b}_1 \wedge \mathbf{b}_2 \wedge \cdots \wedge \mathbf{b}_n|.$$

So we will assume a basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is given with

$$\frac{\prod_{i=1}^{n} |\mathbf{b}_i|}{|\mathbf{b}_1 \wedge \mathbf{b}_2 \wedge \cdots \wedge \mathbf{b}_n|} \leqslant \frac{n!\, 2^n}{V(n)},$$

and we let $c(\Lambda)$ be this ratio.

Suppose $\mathbf{x} \in S_\sigma$, with $\mathbf{x} + \mathbf{y} \in D$, where $\mathbf{y} \in V_{\sigma^*}$. Together, $\mathbf{x}$ and $\mathbf{y}$ determine a plane in $\mathbb{R}^n$ which intersects the space $V_{\sigma^*}^\perp$ in a line. There is a unique $\mathbf{x}' \in V_{\sigma^*}^\perp$ with

$$\mathbf{x}' + a\mathbf{y} = \mathbf{x} + \mathbf{y} \in D$$

for some $a \in \mathbb{R}$. If $V_\sigma = V_{\sigma^*}^\perp$, then $\mathbf{x}' = \mathbf{x}$. More generally, we have the map $\mathbf{x}' \mapsto \mathbf{x}$ of $D(V_{\sigma^*}^\perp)$ into V_σ is a linear transformation of $\mathbb{R}^n$ taking $V_{\sigma^*}^\perp$ to V_σ. We must obtain a bound for the determinant of this transformation.

Let $\widehat{\mathbf{b}_1}, \widehat{\mathbf{b}_2}, \ldots, \widehat{\mathbf{b}_n}$ be the dual basis, i.e.,

$$\mathbf{b}_i \cdot \widehat{\mathbf{b}_j} = \delta_{ij}$$

for all i and j, where δ is the Kronecker symbol. Then

$$\widehat{\mathbf{b}_j} = \pm \frac{\mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{j-1} \wedge \mathbf{b}_{j+1} \wedge \cdots \wedge \mathbf{b}_n}{|\mathbf{b}_1 \wedge \mathbf{b}_2 \wedge \cdots \wedge \mathbf{b}_n|},$$

and

$$|\mathbf{b}_j|\,|\widehat{\mathbf{b}_j}| \leqslant \frac{|\mathbf{b}_1|\,|\mathbf{b}_2| \cdots |\mathbf{b}_n|}{|\mathbf{b}_1 \wedge \mathbf{b}_2 \wedge \cdots \wedge \mathbf{b}_n|} = c(\Lambda)$$

for all $1 \leqslant j \leqslant n$. Let

$$\mathbf{x} = \sum_{i \in \sigma} a_i \mathbf{b}_i$$

and

$$\mathbf{x}' - \mathbf{x} = \sum_{i \notin \sigma} a_i \mathbf{b}_i.$$

Then, for $i \in \sigma$, we have

$$a_i = \mathbf{x} \cdot \widehat{\mathbf{b}_i} = \left(\mathbf{x}' - \sum_{j \notin \sigma} a_j \mathbf{b}_j\right) \cdot \widehat{\mathbf{b}_i} = \mathbf{x}' \cdot \widehat{\mathbf{b}_i}.$$

This gives

$$|a_i|\,|\mathbf{b}_i| \leqslant |\mathbf{x}'|\,|\mathbf{b}_i|\,|\widehat{\mathbf{b}_i}| \leqslant c(\Lambda)\,|\mathbf{x}'|$$

and

$$|\mathbf{x}| \leqslant mc(\Lambda)\,|\mathbf{x}'|.$$

We have thus shown

$$\mathrm{Vol}(S_\sigma) \leqslant (mc(\Lambda))^m \,\mathrm{Vol}(D(V_{\sigma^*}^{\perp})). \tag{13}$$

This, together with (12) and Theorem 0, gives Theorem 4.

The major drawback of Theorem 4 is that the volumes $\widetilde{V_m(D)}$ may be difficult to work with. Of course, for particular types of domains, much more can be said. In particular, for certain domains one may use the $V_m(D)$'s appearing in Davenport's result. In this section we concentrate on a particular type of star-convex domain, which we call a *coordinate domain.*

DEFINITION. A coordinate domain, $D \subset \mathbb{R}^n$, is a compact domain such

that if $(x_1, x_2, ..., x_n)$ is a point in D, then $(y_1, y_2, ..., y_n) \in D$, provided $|y_i| \leqslant |x_i|$ for all $i = 1, 2, ..., n$.

THEOREM 5. *Let $D \subset \mathbb{R}^n$ be a coordinate domain such that any line intersects D in a set of points which, if not empty, consists of the union of at most s intervals. Let Λ be an n-dimensional lattice in $\mathbb{R}^n$. Then*

$$|\mathrm{card}\{D \cap \Lambda\} - \mathrm{Vol}(D)/\det(\Lambda)| \leqslant \sum_{m=0}^{n-1} s^{n-m}(mc(\Lambda)\, 2^{n-m})^m \frac{V_m(D)}{\det(\Lambda^{-[n-m]})},$$

where $c(\Lambda)$ and the $V_m(D)$'s are as above.

Proof. We use Theorem 0 and (13) above. We actually show something stronger than what is needed for Theorem 5. We show, for V any m-dimensional subspace of $\mathbb{R}^n$,

$$\sum_{\sigma} \mathrm{Vol}(D(V_{\sigma^*}^{\perp})) \leqslant 2^{nm-m^2} V'_m(D), \tag{14}$$

where $V'_m(D)$ is the sum of the m-dimensional volumes of the orthogonal projections onto V of D intersected with the various m-dimensional subspaces of $\mathbb{R}^n$ obtained by equating $n-m$ coordinates to zero. Note that $V'_m(D) \leqslant V_m(D)$.

Let V be an arbitrary m-dimensional subspace of $\mathbb{R}^n$ and let $\mathbf{x} \in D$ with

$$x_{\tau(1)} = x_{\tau(2)} = \cdots = x_{\tau(t)} = 0 \qquad \text{and} \qquad x_{\tau(i)} \neq 0 \text{ for } i > t,$$

where $0 \leqslant t < n-m$ and τ is some permutation of the numbers 1 through n (if $t = 0$, none of the x_i's equal zero). We claim that there is a $\mathbf{y} \in V^{\perp}$ such that, for $\mathbf{z} = \mathbf{x} + \mathbf{y}$, we have

$$z_{\tau(1)} = z_{\tau(2)} = \cdots = z_{\tau(t)} = z_{\tau(r)} = 0$$

for some $r > t$, and such that

$$\mathbf{z} \in 2D.$$

To prove this claim, suppose first that $V^{\perp}$ is not orthogonal to any coordinate axis. Then any point $\mathbf{y} \in V^{\perp}$ is completely determined by any $n-m$ components. For example, $V^{\perp}$ is given by equations of the form

$$y_{\tau(i)} = \sum_{j=1}^{n-m} a_{ij}\, y_{\tau(j)} \qquad (n-m < i \leqslant n).$$

More generally, we have

$$y_{\tau(i)} - a_{ij}\, y_{\tau(j)} \in \left\{ \sum_{l=1}^{n-m-1} c_l\, y_{\tau(l)} : c_l \in \mathbb{R} \right\}$$

for all $n-m \leqslant i, j \leqslant n$, where the a_{ij}'s are non-zero real numbers satisfying the two conditions

$$a_{ij} = (a_{ji})^{-1}, \qquad a_{ij} = a_{il}/a_{jl}, \tag{15}$$

for all i, j, and l greater than or equal to $n-m$.

Define a'_{ij} to be $|a_{ij}|$ and b_{ij} to be $|x_{\tau(i)}/x_{\tau(j)}|$ for all $n-m \leqslant i, j \leqslant n$. Then the a'_{ij}'s and the b_{ij}'s satisfy the two conditions in (15) above. Let

$$\frac{a'_{i_0 j_0}}{b_{i_0 j_0}} \geqslant \frac{a'_{ij}}{b_{ij}}$$

for all i and j. Then we have

$$a'_{ii_0} = \frac{a'_{ij_0}}{a'_{i_0 j_0}} \leqslant \frac{b_{ij_0}}{b_{i_0 j_0}} = b_{ii_0}$$

for $n-m \leqslant i \leqslant n$.

Let $\mathbf{y} \in V^{\perp}$ be given by

$$y_{\tau(i)} = \begin{cases} 0 & \text{if} \quad i < n-m, \\ -x_{\tau(i_0)} & \text{if} \quad i = i_0, \end{cases}$$

so that $y_{\tau(i)} = -a_{ii_0} x_{\tau(i_0)}$ for $i > n-m$. Then $\mathbf{z} = \mathbf{x} + \mathbf{y}$ satisfies

$$z_{\tau(1)} = z_{\tau(2)} = \cdots = z_{\tau(t)} = z_{\tau(i_0)} = 0,$$

and moreover,

$$|z_{\tau(i)}| = |x_{\tau(i)} - a_{ii_0} x_{\tau(i_0)}| \leqslant |x_{\tau(i)}| + a'_{ii_0}\, |x_{\tau(i_0)}| \leqslant 2\, |x_{\tau(i)}|$$

for $i > n-m$, since

$$a'_{ii_0} \frac{|x_{\tau(i_0)}|}{|x_{\tau(i)}|} = \frac{a'_{ii_0}}{b_{ii_0}} \leqslant 1.$$

This gives $\mathbf{z} \in 2D$, proving the claim in this case. The general case follows as above, ignoring the components of points in $V^{\perp}$ which are orthogonal to coordinate axes, unless $V^{\perp}$ is in fact orthogonal to m coordinate axes. But this case is trivial: we simply choose $\mathbf{y} \in V^{\perp}$ with all components but one equal to zero and the $\tau(i_0)$th component equal to $x_{\tau(i_0)}$, where $i_0 > t$. (This is possible since $t < n-m = \dim_{\mathbb{R}}(V^{\perp})$.)

So suppose that $\mathbf{x} \in D$. By $n-m$ applications of the claim above, there is a $\mathbf{y} \in V^{\perp}$ such that

$$\mathbf{x} + \mathbf{y} \in 2^{n-m} D$$

and $n-m$ of the coordinates are zero. This proves (14), which, together with (12), (13), and Theorem 0, gives Theorem 5.

As a final remark, suppose that the domain D we are dealing with is translated by some vector $\mathbf{w} \in \mathbb{R}^n$. Then for Theorem 4, we certainly have $\mathrm{Vol}(D) = \mathrm{Vol}(D + \mathbf{w})$ and $\widetilde{V_m(D)} = \widetilde{V_m(D + \mathbf{w})}$ for any m. Thus, Theorem 4 remains valid for any translation of the domain as well. Similarly, the volumes in (13) above remain unchanged, so that the statement of Theorem 5 remains true for any translation of a coordinate domain.

REFERENCES

1. E. BOMBIERI AND J. VAALER, On Siegel's lemma, *Invent. Math.* **73** (1983), 11–32.
2. J. CASSELS, "An Introduction to the Geometry of Numbers," Grundlehren Math. Wiss., Vol. 99, Springer-Verlag, Berlin, 1959.
3. H. DAVENPORT, On a principle of Lipschitz, *J. London Math. Soc.* **26** (1951), 179–183.
4. G. HARDY AND E. WRIGHT, "An Introduction to the Theory of Numbers," Clarendon, Oxford, 1954.
5. E. HECKE, "Lectures on the Theory of Algebraic Numbers," Graduate Texts in Mathematics, Vol. 77, Springer-Verlag, Berlin, 1981.
6. W. HODGE AND D. PEDOE, "Methods of Algebraic Geometry," Cambridge Univ. Press, Cambridge, 1947.
7. S. LANG, "Algebraic Number Theory," Graduate Texts in Mathematics, Vol. 110, Springer-Verlag, Berlin, 1986.
8. S. SCHANUEL, Heights in number fields, *Bull. Soc. Math. France* **107** (1979), 433–449.
9. W. SCHMIDT, On heights of algebraic subspaces and Diophantine approximation, *Ann. of Math.* **85** (1967), 430–472.
10. W. SCHMIDT, *in* Diophantine Approximation, "Lecture Notes in Mathematics," Vol. 1467, Springer-Verlag, Berlin, 1991.
11. C. L. SIEGEL, Über einige Anwendungen diophantischer Approximationen, *Abh. Preuss. Akad. Wiss. Math. Phys. Kl.*, Nr. 1 (=*Gesammelten Abh.*, I) (1929), 209–266.
12. J. THUNDER, An asymptotic estimate for heights of algebraic subspaces, *Trans. Amer. Math. Soc.* **331**, No. 1 (1992), 395–424.

Printed in Belgium
Uitgever: Academic Press, Inc.
Verantwoordelijke uitgever voor België:
Hubert Van Maele
Altenastraat 20, B-8310 Sint-Kruis