TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 336, Number 2, April 1993

SMALL SOLUTIONS TO INHOMOGENEOUS LINEAR EQUATIONS OVER NUMBER FIELDS

ROBBIN O'LEARY AND JEFFREY D. VAALER

ABSTRACT. We consider a system of M independent, inhomogeneous linear equations in $N > M$ variables having coefficients in an algebraic number field k. We give a best possible lower bound on the inhomogeneous height of a solution vector in k^N and determine when a solution exists in $(\mathcal{O}_S)^N$, where $\mathcal{O}_S$ is the ring of S-integers in k. If such a system has a solution vector in $(\mathcal{O}_S)^N$, we show that it has a solution $\vec{\zeta}$ in $(\mathcal{O}_S)^N$ such that the inhomogeneous height of $\vec{\zeta}$ is relatively small. We give an explicit upper bound for this height in terms of the heights of the matrices defining the linear system. Our method uses geometry of numbers over adele spaces and local to global arguments.

1. INTRODUCTION

Let A be an $M \times N$ matrix over an algebraic number field k with $1 \leq \operatorname{rank}(A) = M < N$ and let $\vec{b} \neq \vec{0}$ be a (column) vector in k^M. We consider Diophantine inequalities which arise from the system of inhomogeneous linear equations $A\vec{x} = \vec{b}$. If S is a suitable finite set of places of k containing all Archimedean places and $\mathcal{O}_S$ is the ring of S-integers in k, we will show that there exists a solution in $(\mathcal{O}_S)^N$ having relatively small inhomogeneous height. This may be regarded as the inhomogeneous analog of the results obtained in [2]. We give precise definitions for our heights in §2 and note here only that they are identical to that used in the related work [23 and 24]. To begin with we have the following simple lower bound.

Theorem 1. *If $\vec{\xi} \in k^N$ satisfies the equation $A\vec{\xi} = \vec{b}$, then*

$$\frac{H(A\vec{b})}{H(A)} \leq H\begin{pmatrix}\vec{\xi}\\ 1\end{pmatrix}, \tag{1.1}$$

where $(A\vec{b})$ is the $M \times (N+1)$ augmented matrix.

Before considering upper bounds we note that when studying solutions to $A\vec{x} = \vec{b}$ which occur in $(\mathcal{O}_S)^N$ and initial question to be addressed is the existence of solutions. This classical problem was solved for $k = \mathbb{Q}$ and $\mathcal{O}_S = \mathbb{Z}$

Received by the editors September 19, 1990 and, in revised form, February 8, 1991.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11D04, 11D75, 11H31.
Research of the second author was supported in part by the National Science Foundation, DMS-8701396.

0002-9947/93 $1.00 + $.25 per page

by I. Heger [14]. Subsequent proofs were given by H. J. S. Smith [22] and G. Frobenius [11, pp. 171–173]. It is instructive to recall Heger's result. Suppose that A and $\vec{b}$ have entries in $\mathbb{Z}$ and we seek a solution vector in $\mathbb{Z}^N$. Let

$$F = \text{g.c.d.}\{\det A_I : I \subseteq \{1, 2, \dots, N\} \text{ and } |I| = M\},$$

$$G = \text{g.c.d.}\{\det(A\vec{b})_J : J \subseteq \{1, 2, \dots, N, N+1\} \text{ and } |J| = M\},$$

where A_I is the $M \times M$ submatrix of A having columns indexed by the elements of I. Of course $(A\vec{b})_J$ is a similarly defined submatrix of $(A\vec{b})$. Heger proved that $A\vec{x} = \vec{b}$ has a solution in $\mathbb{Z}^N$ if and only if $F = G$. In the terminology of heights, $A\vec{x} = \vec{b}$ has a solution in $\mathbb{Z}^N$ if and only if the local p-adic heights satisfy

$$H_p(A) = H_p(A\vec{b}) \tag{1.2}$$

at each (finite) prime number p. In fact (1.2) is a necessary and sufficient condition if we assume more generally that A and $\vec{b}$ have entries in $\mathbb{Q}$. The extension to the ring of S-integers of a number field k is now clear.

Theorem 2. *Let S be a finite set of places of k containing all Archimedean places. There exists a vector $\vec{\zeta} \in (\mathcal{O}_S)^N$ which satisfies $A\vec{\zeta} = \vec{b}$ if and only if*

$$H_v(A) = H_v(A\vec{b}) \tag{1.3}$$

at each place $v \notin S$.

The main problem which we consider is to establish the existence of solutions to the inhomogeneous equation $A\vec{x} = \vec{b}$ which occur in $(\mathcal{O}_S)^N$ and which have small inhomogeneous height. The results obtained in [2] for the homogeneous problem made use of the analog in the adele ring $k_{\mathbf{A}}$ of Minkowski's theorem on successive minima in the geometry of numbers. In the present paper we will also require an adelic formulation of some classical inequalities for the inhomogeneous minimum of a suitable subset of $(k_{\mathbf{A}})^L$. This method naturally leads to an inequality containing constants which depend on the field k and the codimension $L = N - M$. Of these, $c_k(L)$, which is given by (2.2) and (2.3), is completely explicit. The constant $\nu(k)$ defined by (5.3) is less so. Essentially, $\nu(k)$ is the maximum distance, in an appropriate sense, between a point of $\prod_{v|\infty} k_v$ and the nearest point in the ring $\mathcal{O}_k$ of algebraic integers in k. In order that our results be as explicit as possible we give an upper bound for $\nu(k)$ in §7.

Theorem 3. *Let S be a finite set of places of k containing all Archimedean places and suppose that the matrices A and $(A\vec{b})$ satisfy (1.3) at each place $v \notin S$. Then there exists $\vec{\zeta} \in (\mathcal{O}_S)^N$ which satisfies $A\vec{\zeta} = \vec{b}$ and*

$$H\begin{pmatrix}\vec{\zeta}\\ 1\end{pmatrix} \le \left\{\frac{H(A\vec{b})^2}{H(A)^2} + \nu(k)^2(L - 1 + c_k(L)^L H(A))^2\right\}^{1/2}, \tag{1.4}$$

where $L = N - M$.

We note that the bound (1.4) is independent of S. It applies whenever the equation $A\vec{x} = \vec{b}$ has a solution in $(\mathscr{O}_S)^N$. In view of Theorem 2, for each choice of A and $\vec{b}$ we may select S so that the complimentary set of places $\tilde{S}$ is given by

$$\tilde{S} = \{v : v \nmid \infty \text{ and } H_v(A) = H_v(A\vec{b})\}.$$

As A and $\vec{b}$ are defined over k it is obvious that $\tilde{S}$ contains all but finitely many places.

It is possible in very simple situations to demonstrate that nontrivial cases of equality may occur in each of the inequalities (1.1) and (1.4). Consider, for example, the Diophantine equation $a_1x_1 + a_2x_2 = b$, to be solved in integers x_1, x_2 under the assumption that a_1, a_2, and b are integers with g.c.d.$(a_1, a_2) = 1$. The inhomogeneous height of such a solution is then $(1 + x_1^2 + x_2^2)^{1/2}$. If ξ_1, ξ_2 represents one integer solution then every integer solution has the form

$$x_1 = \xi_1 - na_2, \quad x_2 = \xi_2 + na_1, \qquad n \in \mathbb{Z}.$$

Thus the minimum inhomogeneous height of a solution is

$$\min_{n\in\mathbb{Z}}\{1 + (\xi_1 - na_2)^2 + (\xi_2 + na_1)^2\}^{1/2}.$$

If we use the identity $a_1\xi_1 + a_2\xi_2 = b$ we find that

$$1 + (\xi_1 - na_2)^2 + (\xi_2 + na_1)^2 = \frac{a_1^2 + a_2^2 + b^2}{a_1^2 + a_2^2} + \left(n - \frac{a_2\xi_1 - a_1\xi_2}{a_1^2 + a_2^2}\right)^2 (a_1^2 + a_2^2).$$

It follows that the minimum inhomogeneous height of a solution is

$$\left\{\frac{a_1^2 + a_2^2 + b^2}{a_1^2 + a_2^2} + \left\|\frac{a_2\xi_1 - a_1\xi_2}{a_1^2 + a_2^2}\right\|^2 (a_1^2 + a_2^2)\right\}^{1/2},$$

where $\|x\|$ is the distance from the real number x to the nearest integer. If $((a_2\xi_1 - a_1\xi_2)/(a_1^2 + a_2^2))$ is an integer then there is equality in the lower bound (1.1). On the other hand, we have $\nu(\mathbb{Q}) = 1/2$ and $c_{\mathbb{Q}}(1) = 1$. Therefore, if $((a_2\xi_1 - a_1\xi_2)/(a_1^2 + a_2^2)) - 1/2$ is an integer then there is equality in the upper bound (1.4).

There have been several recent results on the existence of small integer solutions to the equation $A\vec{x} = \vec{b}$ when A and $\vec{b}$ have integer entries and some solution in $\mathbb{Z}^N$ is known to exist. In particular the papers [3–7] all deal with this problem. In some cases small positive solutions are also considered. For example, Borosh, Flahive, Rubin and Treybig [7] have established the existence of $\vec{\zeta} \in \mathbb{Z}^N$ such that $A\vec{\zeta} = \vec{b}$ and

$$\max_n |\zeta_n| \le \max_{|J|=M} |\det(A\vec{b})_J|. \tag{1.5}$$

This is not strictly comparable with Theorem 3 which bounds

$$\{1 + (\zeta_1)^2 + (\zeta_2)^2 + \cdots + (\zeta_N)^2\}^{1/2}$$

in the integer case. However, a less precise form of (1.4) would be

$$\{1+(\zeta_1)^2+(\zeta_2)^2+\cdots+(\zeta_N)^2\}^{1/2} \ll N\left\{\frac{\max_{|J|=M}|\det(A\vec{b})_J|}{\max_{|I|=M}|\det A_I|}+\max_{|I|=M}|\det A_I|\right\},$$

where the implied constant depends only on N. This is clearly sharper than (1.5) if

$$\left(\max_{|J|=M}|\det(A\vec{b})_J|\right)^{\varepsilon} \ll_{\varepsilon} \max_{|I|=M}|\det A_I| \ll_{\varepsilon} \left(\max_{|J|=M}|\det(A\vec{b})_J|\right)^{1-\varepsilon}$$

for some ε, $0<\varepsilon\le 1/2$, and $\max_{|J|=M}|\det(A\vec{b})_J|$ is sufficiently large.

The authors wish to thank the referee for several helpful suggestions concerning the organization of this paper.

2. Heights and projections

We assume throughout that the number field k has degree d over $\mathbb{Q}$. If v is a place of k we write k_v for the completion of k at v. Then $d_v=[k_v:\mathbb{Q}_v]$ denotes the local degree. If v is an infinite place we write $\|\ \|_v$ for the usual Euclidean absolute value on k_v. If v is finite then $\|\ \|_v$ denotes the unique absolute value on k_v which extends the usual p-adic absolute value on $\mathbb{Q}_p$, where $v|p$. We will also employ a second absolute value $|\ |_v$ at each place v by setting $|\ |_v=\|\ \|_v^{d_v/d}$. If $\alpha\in k$, $\alpha\neq 0$, then the product formula takes the form $\prod_v|\alpha|_v=1$. If $v\nmid\infty$ we write

$$\mathcal{O}_v=\{x\in k_v:|x|_v\le 1\} \tag{2.1}$$

for the ring of v-adic integers in k_v. Also, we extend the absolute values $\|\ \|_v$ and $|\ |_v$ to column vectors $\vec{x}$ in $(k_v)^L$ by setting

$$\|\vec{x}\|_v=\max_{1\le l\le L}\|x_l\|_v \quad \text{if } v\nmid\infty,$$

$$\|\vec{x}\|_v=\left(\sum_{l=1}^{L}\|x_l\|_v^2\right)^{1/2} \quad \text{if } v\mid\infty,$$

and $|\vec{x}|_v=\|\vec{x}\|_v^{d_v/d}$ in both cases.

We select a Haar measure β_v on the additive group of k_v by making the following normalization:

(i) If $v\nmid\infty$ we require that $\beta_v(\mathcal{O}_v)=|\mathcal{D}_v|_v^{d/2}$, where $\mathcal{D}_v$ is the local different of k at v.

(ii) If $k_v=\mathbb{R}$ then β_v is the usual Lebesgue measure on $\mathbb{R}$.

(iii) If $k_v=\mathbb{C}$ then β_v is Lebesgue measure on the complex plane multiplied by 2.

We write β for the normalized Haar measure on the adele ring $k_{\mathbf{A}}$ which is induced by the product measure $\prod_v\beta_v$. If $(k_{\mathbf{A}})^L$ is the L-fold product of adele spaces we write V for the product Haar measure β^L on $(k_{\mathbf{A}})^L$. In the geometry of numbers over $(k_{\mathbf{A}})^L$ the Haar measure V plays the role of volume

in the classical theory. Of course we will always identify k and k^L with their canonical injections in $k_{\mathbf{A}}$ and $(k_{\mathbf{A}})^L$, respectively.

If $L \geq 1$ is an integer it will be convenient to define

$$r_v(L) = \begin{cases} \pi^{-1/2}\{\Gamma(\frac{1}{2}L+1)\}^{1/L} & \text{if } v \text{ is real}, \\ (2\pi)^{-1/2}\{\Gamma(L+1)\}^{1/2L} & \text{if } v \text{ is complex}, \\ 1 & \text{if } v \nmid \infty, \end{cases} \tag{2.2}$$

and

$$c_k(L) = \left\{ 2|\Delta_k|^{1/2d} \prod_{v|\infty} (r_v(L))^{d_v/d} \right\}, \tag{2.3}$$

where Δ_k is the discriminant of k. Because of our normalization of β_v we have

$$\beta_v^L(\{\vec{x} \in (k_v)^L : \|\vec{x}\|_v < r_v(L)\}) = 1 \tag{2.4}$$

at each infinite place v of k. At the finite places we will make use of the identity

$$\prod_{v \nmid \infty} \beta_v(\mathcal{O}_v) = \prod_{v \nmid \infty} |\mathcal{D}_v|_v^{d/2} = |\Delta_k|^{-1/2}. \tag{2.5}$$

Let A be an $M \times N$ matrix over k_v with $1 \leq \operatorname{rank}(A) = M < N$. Then the *local height* $H_v(A)$ is defined on the Grassman coordinates of A as follows:

(i) If $v \nmid \infty$ then

$$H_v(A) = \max_{|I|=M} |\det A_I|_v,$$

(ii) if $v \mid \infty$ then

$$H_v(A) = \left(\sum_{|I|=M} \|\det A_I\|_v^2 \right)^{d_v/2d}.$$

In each of these formulae $I \subseteq \{1, 2, \ldots, N\}$, I is a subset of cardinality $|I| = M$, and A_I is the $M \times M$ submatrix having columns indexed by the elements in I. The set of all such I indexes the Grassmann coordinates $\{(\det A_I)\}$ of the matrix A. If A has elements in k we define the *global height* $H(A)$ by

$$H(A) = \prod_v H_v(A).$$

Of course $H_v(A) = 1$ at almost all places v. Further properties of these heights are given in [24, §2].

If X is an $N \times L$ matrix over k or k_v, $1 \leq \operatorname{rank}(X) = L < N$, we define $H(X)$ and $H_v(X)$ by applying the above definitions to the transpose of X. Now it will be convenient to write ${}_JX$ for the $L \times L$ submatrix of X having rows indexed by the elements of J, where $J \subseteq \{1, 2, \ldots, N\}$ and $|J| = L$. In particular, if $\vec{x} \in k^N$ then

$$H(\vec{x}) = \prod_v H_v(\vec{x}) = \prod_v |\vec{x}|_v$$

denotes the global height of $\vec{x}$. If $\alpha \in k$, $\alpha \neq 0$, then $H(\alpha \vec{x}) = H(\vec{x})$ by the product formula. Thus H is a *homogeneous height* on k^N—it may by regarded as a height on the projective space $\mathbb{P}_k^{N-1}$. In the present paper we require an inhomogeneous height on k^N. We obtain this by mapping $\vec{x}$ in k^N into $\binom{\vec{x}}{1}$ in k^{N+1} and then applying H on k^{N+1}. That is, the *inhomogeneous height* of $\vec{x}$ in k^N is simply $H\binom{\vec{x}}{1}$.

Suppose that X is an $N \times L$ matrix over k_v, $1 \leq \operatorname{rank}(X) = L < N$. We use X to define an $N \times N$ matrix $P_v = P_v(X)$ as follows:

(i) If $v \nmid \infty$ we select $J \subseteq \{1, 2, \ldots, N\}$ so that $|J| = L$ and

$$H_v(X) = \max_{|I|=L} |\det {}_I X|_v = |\det {}_J X|_v .$$

Then we define

$$P_v = P_v(X) = X({}_J X)^{-1} {}_J(1_N), \tag{2.6}$$

where 1_N is the $N \times N$ identity matrix.

(ii) If $v \mid \infty$ we set

$$P_v = P_v(X) = X(X^* X)^{-1} X^*, \tag{2.7}$$

where X^* is the complex conjugate transpose of X (or simply the transpose if v is a real place). Now let $\mathfrak{X}_v = \{X\vec{u} : \vec{u} \in (k_v)^L\}$ be the L-dimensional subspace of $(k_v)^N$ spanned over k_v by the columns of X. The matrix P_v acts as a projection operator onto the subspace $\mathfrak{X}_v$. If $v \mid \infty$ then P_v is the usual orthogonal projection of $(k_v)^N$ onto $\mathfrak{X}_v$, where orthogonality is defined with respect to the Hermitian inner product on $\mathbb{R}^N$ or $\mathbb{C}^N$. In either case we have $P_v \vec{y} \in \mathfrak{X}_v$ for all $\vec{y} \in (k_v)^N$ and $P_v \vec{y} = \vec{y}$ if $\vec{y} \in \mathfrak{X}_v$. These and other properties of P_v are proved in [24, §4] and [23, §2]. In particular it can be shown that P_v depends only on the subspace $\mathfrak{X}_v$ and not on the matrix X having columns which form a basis for $\mathfrak{X}_v$ over k_v. Thus we may write $P_v = P_v(\mathfrak{X}_v)$. Also, if $\vec{y} \in (k_v)^N$ and $v \mid \infty$ then

$$\|\vec{y}\|_v^2 = \|P_v \vec{y}\|_v^2 + \|(1_N - P_v)\vec{y}\|_v^2, \tag{2.8}$$

which is the familiar Pythagorean identity. If $v \nmid \infty$ and $\vec{y} \in (k_v)^N$ then the analog of (2.8) is

$$|\vec{y}|_v = \max\{|P_v \vec{y}|_v, |(1_N - P_v)\vec{y}|_v\}. \tag{2.9}$$

A proof of (2.9) is given in [24, Lemma 7].

3. A LOCAL RESULT

Throughout this section we assume that v is a fixed place of k and that $A = (a_{mn})$ is an $M \times N$ matrix over k_v with $1 \leq \operatorname{rank}(A) = M < N$. We define

$$\mathfrak{X}_v = \{\vec{x} \in (k_v)^N : A\vec{x} = \vec{0}\} \tag{3.1}$$

so that $\mathfrak{X}_v \subseteq (k_v)^N$ is an $L = N - M$ dimensional subspace. Let $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_L\}$ be a basis for $\mathfrak{X}_v$ over k_v and write

$$X = (\vec{x}_1 \vec{x}_2 \cdots \vec{x}_L) \tag{3.2}$$

for the corresponding $N \times L$ matrix over k_v.

Theorem 4. *Suppose that* $\vec{\xi} \in (k_v)^N$ *satisfies* $A\vec{\xi} = \vec{b}$ *and* $P_v = P_v(\mathfrak{X}_v)$ *is the projection onto the subspace* $\mathfrak{X}_v$ *defined by* (3.1). *Then the following identity holds*:

(i) *if* $v \nmid \infty$ *then*

$$\frac{H_v(A\vec{b})}{H_v(A)} = \max\{1, |(1_N - P_v)\vec{\xi}|_v\}, \tag{3.3}$$

(ii) *if* $v \mid \infty$ *then*

$$\frac{H_v(A\vec{b})}{H_v(A)} = \{1 + \|(1_N - P_v)\vec{\xi}\|_v^2\}^{d_v/2d}. \tag{3.4}$$

Proof. Since the columns of X span the null space of A, the Grassmann coordinates of A and X are related by the duality theorem of Brill-Gordan [12] (see also [16, Theorem I, p. 294]). Thus there exists a constant $\gamma \in k_v$, $\gamma \neq 0$, with the following property: if $I \subseteq \{1, 2, \ldots, N\}$, $|I| = M$, and $\tilde{I} = \{1, 2, \ldots, N\} \backslash I$ is the complement of I, then

$$\det A_I = (-1)^{\varepsilon(I)}\gamma \det_{\tilde{I}} X, \tag{3.5}$$

where $\varepsilon(I) = \sum_{i \in I} i$. It follows that $H_v(A) = |\gamma|_v H_v(X)$.

Now define

$$\mathscr{Y}_v = \{\vec{y} \in (k_v)^{N+1} : (A\vec{b})\vec{y} = \vec{0}\},$$

so that $\mathscr{Y}_v \subseteq (k_v)^{N+1}$ is a subspace of dimension $L + 1$. By hypothesis the columns of the $(N + 1) \times (L + 1)$ matrix

$$Z = \begin{pmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_L & \vec{\xi} \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

span $\mathscr{Y}_v$ over k_v. Again by the Brill-Gordan theorem there exists a constant $\gamma' \in k_v$, $\gamma' \neq 0$, so that

$$\det(A\vec{b})_J = (-1)^{\varepsilon(J)}\gamma' \det_{\tilde{J}} Z \tag{3.6}$$

for each subset $J \subseteq \{1, 2, \ldots, N, N + 1\}$, $|J| = M$, with

$$\tilde{J} = \{1, 2, \ldots, N, N + 1\} \backslash J.$$

By selecting $I = J \subseteq \{1, 2, \ldots, N\}$ in (3.5) and (3.6) we find that $\gamma = -\gamma'$. Of course this implies that $H_v(A\vec{b}) = |\gamma|_v H_v(Z)$ and therefore

$$\frac{H_v(A\vec{b})}{H_v(A)} = \frac{H_v(Z)}{H_v(X)}. \tag{3.7}$$

Next we observe that the Grassmann coordinates $\{\det_{\tilde{J}} Z\}$ split naturally into two disjoint classes: those for which $(N+1) \in \tilde{J}$ and those for which $(N+1) \notin \tilde{J}$. If $v \nmid \infty$ we conclude that

$$H_v(Z) = \max\{H_v(X), H_v(X\vec{\xi})\}$$

and when $v \mid \infty$ we have

$$H_v(Z) = \{H_v(X)^{2d/d_v} + H_v(X\vec{\xi})^{2d/d_v}\}^{d_v/2d}.$$

Finally we use the identity

$$H_v(X\vec{\xi}) = H_v(X)|(1_N - P_v)\vec{\xi}|_v,$$

which follows from [24, Lemma 4]. The statement of the theorem follows by combining these identities with (3.7).

4. Proof of Theorems 1 and 2

Suppose that A is an $M \times N$ matrix over k, $1 \le \operatorname{rank}(A) = M < N$, and $\vec{b} \in k^M$, $\vec{b} \neq \vec{0}$. If $\vec{\xi} \in k^N$ satisfies $A\vec{\xi} = \vec{b}$ we may apply Theorem 4 at each place v. If $v \nmid \infty$ we use (2.9) and (3.3) to conclude that

$$\frac{H_v(A\vec{b})}{H_v(A)} \le \max\{1, |\vec{\xi}|_v\} = H_v\begin{pmatrix}\vec{\xi}\\ 1\end{pmatrix}.$$

If $v \mid \infty$ then (2.8) and (3.4) imply that

$$\frac{H_v(A\vec{b})}{H_v(A)} \le (1 + \|\vec{\xi}\|_v^2)^{d_v/2d} = H_v\begin{pmatrix}\vec{\xi}\\ 1\end{pmatrix}.$$

Theorem 1 follows by taking the product over all v.

We now prove Theorem 2. If $A\vec{\zeta} = \vec{b}$ with $\vec{\zeta} \in (\mathscr{O}_S)^N$ then Theorem 4 applies at each place v. In particular, if $v \notin S$ we have $|(1_N - P_v)\vec{\zeta}|_v \le |\vec{\zeta}|_v \le 1$ by (2.9). The identity $H_v(A) = H_v(A\vec{b})$ now follows from (3.3).

Suppose that $H_v(A\vec{b}) = H_v(A)$ at each $v \notin S$. Let $\vec{\xi} \in k^N$ satisfy $A\vec{\xi} = \vec{b}$ and let $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_L\}$ be a basis for the subspace $\mathfrak{X} = \{\vec{x} \in k^N : A\vec{x} = \vec{0}\}$, where $L + M = N$. Then write $X = (\vec{x}_1 \vec{x}_2 \cdots \vec{x}_L)$ for the corresponding $N \times L$ matrix over k. Since the columns of X span the closure of $\mathfrak{X}$ in $(k_v)^N$ we have

$$\{\vec{y} \in (k_v)^N : A\vec{y} = \vec{b}\} = \{\vec{\xi} - X\vec{u} : \vec{u} \in (k_v)^L\}$$

at each v. For $v \notin S$ we define $T_v = \{\vec{u} \in (k_v)^L : |\vec{\xi} - X\vec{u}|_v \le 1\}$. Plainly each T_v is open (and closed) in $(k_v)^L$ and, as $\vec{\xi}$ and X contain finitely many entries, we have $T_v = (\mathscr{O}_v)^L$ at almost all $v \notin S$. We also claim that each T_v is not empty. To see this let $v \notin S$ be fixed and $J = J_v$ selected so that $P_v = X({}_JX)^{-1}{}_J(1_N)$, as in (2.6). Then $\vec{u} = ({}_JX)^{-1}{}_J(1_N)\vec{\xi}$ is in $(k_v)^L$ and satisfies

$$(1_N - P_v)\vec{\xi} = \vec{\xi} - X\vec{u}. \tag{4.1}$$

Our initial hypothesis (1.3) together with Theorem 4 implies that

$$|(1_N - P_v)\vec{\xi}|_v \le 1. \tag{4.2}$$

Now (4.1) and (4.2) clearly establish that T_v is not empty.

Let $k_{\mathbf{A}}^S$ denote the restricted topological product of $\{k_v : v \notin S\}$ with respect to $\{\mathcal{O}_v : v \notin S\}$. By the strong approximation theorem, k^L is dense in $(k_{\mathbf{A}}^S)^L$ (see [9, §15]). As $\prod_{v \notin S} T_v$ is open and not empty in $(k_{\mathbf{A}}^S)^L$, there exists a $\vec{z} \in k^L$ such that $|\vec{\xi} - X\vec{z}|_v \leq 1$ at all $v \notin S$. The vector $\vec{\zeta} = \vec{\xi} - X\vec{z}$ is therefore in $(\mathcal{O}_S)^N$ and satisfies $A\vec{\zeta} = \vec{b}$.

5. The inhomogeneous minimum

At each place v of k let $R_v \subseteq (k_v)^L$ be a nonempty open set. If $v \nmid \infty$ we assume also that R_v is a compact $\mathcal{O}_v$-module and that $R_v = (\mathcal{O}_v)^L$ for almost all finite v. If $v \mid \infty$ we assume that R_v is convex, symmetric, and bounded. In this case the product

$$\mathcal{R} = \prod_v R_v \subseteq (k_{\mathbf{A}})^L$$

will be called an *admissible* subset of the L-fold product of adele spaces. If $\tau > 0$ we define the dilation $\tau\mathcal{R}$ by

$$\tau\mathcal{R} = \prod_{v|\infty} (\tau R_v) \times \prod_{v\nmid\infty} R_v .$$

Obviously $\tau\mathcal{R}$ is also admissible. Next we recall the definition of the successive minima $0 < \lambda_1 \leq \cdots \leq \lambda_L < \infty$ of $\mathcal{R}$. We set

$$\lambda_l = \inf\{\lambda > 0 : \lambda\mathcal{R} \cap k^L \text{ contains } l \text{ linearly independent vectors}\}.$$

The adelic analog of Minkowski's second theorem (see [2, Theorem 3]) states that

$$(\lambda_1\lambda_2 \cdots \lambda_L)^d V(\mathcal{R}) \leq 2^{dL} . \tag{5.1}$$

Another important quantity associated with an admissible subset $\mathcal{R}$ is the *inhomogeneous minimum* $\mu = \mu(\mathcal{R})$, which we now define. As is well known, $(k_{\mathbf{A}})^L$ is a locally compact abelian group, $k^L \subseteq (k_{\mathbf{A}})^L$ is a discrete subgroup and the quotient $(k_{\mathbf{A}})^L/k^L$ is compact. It is not difficult to show that $\tau\mathcal{R}$ must contain a fundamental set for $(k_{\mathbf{A}})^L/k^L$ if $\tau > 0$ is sufficiently large. This may be proved as in [25, Proposition 6, p. 89]. We define $\mu(\mathcal{R})$ to be the infimum of positive real numbers τ such that $\tau\mathcal{R}$ contains a complete set of representatives of the cosets modulo k^L in $(k_{\mathbf{A}})^L$. That is, we define

$$\mu = \mu(\mathcal{R}) = \inf\left\{\tau > 0 : (k_{\mathbf{A}})^L \subseteq \bigcup_{\vec{\xi} \in k^L} (\tau\mathcal{R} + \vec{\xi})\right\} . \tag{5.2}$$

As in the classical situation there are simple inequalities which relate $\mu(\mathcal{R})$ and the successive minima of $\mathcal{R}$. In order to establish these in the adelic setting we need to introduce a field constant. At each *infinite* place v of k let

$$\mathcal{O}_v = \{x \in k_v : \|x\|_v < 1\},$$

(and, of course, $\mathcal{O}_v$ is defined for $v \nmid \infty$ by (2.1)). Then $\mathcal{S}_k = \prod_v \mathcal{O}_v \subseteq k_{\mathbf{A}}$ is an admissible subset. We define

$$\nu = \nu(k) = \mu(\mathcal{S}_k), \tag{5.3}$$

so that ν depends only on k. For example, $\nu(\mathbb{Q}) = 1/2$ and this amounts to the observation that the distance from a real number to the nearest integer is less than or equal to $1/2$. In our proof of Theorem 3 we require the following adelic version of a well-known inequality of Jarnik [18], (see also [8, p. 313] or [13, pp. 99–106]).

Theorem 5. *If $\mathscr{R} \subseteq (k_{\mathbf{A}})^L$ is admissible then the inhomogeneous minimum μ and the successive minima $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$ satisfy*

$$\tfrac{1}{2}\lambda_L \leq \mu \leq \nu(k)(\lambda_1 + \lambda_2 + \cdots + \lambda_L). \tag{5.4}$$

Proof. Let $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_L$ be independent vectors in k^L such that for each l, $1 \leq l \leq L$, $\{\vec{u}_1, \ldots, \vec{u}_l\} \subseteq \lambda\mathscr{R}$ whenever $\lambda_l < \lambda$. Assume that $\mu < \frac{1}{2}\lambda_L$. Let τ_1 and τ_2 satisfy $\mu < \tau_1 < \frac{1}{2}\lambda_L < \tau_2$ and $\tau_1 + \tau_2 < \lambda_L$. Then select $(\vec{w}_v) \in (k_{\mathbf{A}})^L$ by setting $\vec{w}_v = \frac{1}{2}\vec{u}_L$ if $v \mid \infty$ and $\vec{w}_v = \vec{u}_L$ if $v \nmid \infty$. Since $\mu < \tau_1$ there exists $\vec{\xi} \in k^L$ such that $(\vec{w}_v - \vec{\xi}) \in \tau_1\mathscr{R}$. That is, $\frac{1}{2}\vec{u}_L - \vec{\xi} \in \tau_1 R_v$ at each infinite place and $\vec{u}_L - \vec{\xi} \in R_v$ at each finite place.

If $v \mid \infty$ then $\tau_1 R_v$ is symmetric, hence $\vec{\xi} - \frac{1}{2}\vec{u}_L \in \tau_1 R_v$. As $\frac{1}{2}\lambda_L < \tau_2$ we also have $\frac{1}{2}u_L \in \tau_2 R_v$. Using the convexity of R_v we conclude that both

$$\vec{\xi} = (\vec{\xi} - \tfrac{1}{2}\vec{u}_L) + (\tfrac{1}{2}\vec{u}_L) \quad \text{and} \quad \vec{u}_L - \vec{\xi} = (\tfrac{1}{2}\vec{u}_L - \vec{\xi}) + (\tfrac{1}{2}\vec{u}_L))$$

are in $\tau_1 R_v + \tau_2 R_v = (\tau_1 + \tau_2)R_v$. If $v \nmid \infty$ then $\vec{u}_L$ and $\vec{u}_L - \vec{\xi}$ are both in R_v. As R_v is an $\mathscr{O}_v$-module we also have that $\vec{\xi} \in R_v$. These observations at the finite and infinite places of k imply that $\vec{\xi}$ and $\vec{u}_L - \vec{\xi}$ are both vectors in $(\tau_1 + \tau_2)\mathscr{R} \cap (k^L)$. As $\tau_1 + \tau_2 < \lambda_L$ we must have

$$\{\vec{\xi}, \vec{u}_L - \vec{\xi}\} \subseteq \operatorname{span}_k\{\vec{u}_1, \ldots, \vec{u}_{L-1}\}.$$

But this contradicts our assumption that $\vec{u}_1, \ldots, \vec{u}_{L-1}, \vec{u}_L$ are linearly independent. It follows that $\frac{1}{2}\lambda_L \leq \mu$.

If $v \mid \infty$ let $f_v : (k_v)^L \to [0, \infty)$ be the distance function

$$f_v(\vec{x}) = \inf\{\tau > 0 : \vec{x} \in \tau R_v\}.$$

We suppose that $(\vec{\alpha}_v)$ is an arbitrary point in $(k_{\mathbf{A}})^L$. As $\{\vec{u}_1, \ldots, \vec{u}_L\}$ is a basis for k^L over k it is also a basis for each $(k_v)^L$ over k_v. Thus for each place v there exists $\beta_{v,l}$ in k_v, $1 \leq l \leq L$, so that $\vec{\alpha}_v = \sum_{l=1}^{L} \beta_{v,l}\vec{u}_l$. Now for each fixed l it is clear that $(\beta_{v,l}) \in k_{\mathbf{A}}$. From the definition of $\nu = \nu(k)$ we conclude that there exists $\gamma_l \in k$ with

$$\|\beta_{v,l} - \gamma_l\|_v \leq \nu \quad \text{if } v \mid \infty, \qquad \|\beta_{v,l} - \gamma_l\|_v \leq 1 \quad \text{if } v \nmid \infty.$$

It follows that at each infinite place

$$f_v\left(\overrightarrow{\alpha}_v - \sum_{l=1}^{L} \gamma_l \overrightarrow{u}_l\right) = f_v\left(\sum_{l=1}^{L} (\beta_{v,l} - \gamma_l)\overrightarrow{u}_l\right)$$
$$\leq \sum_{l=1}^{L} \|\beta_{v,l} - \gamma_l\|_v f_v(\overrightarrow{u}_l)$$
$$\leq \nu(k) \sum_{l=1}^{L} \lambda_l .$$

At each finite place we have

$$\overrightarrow{\alpha}_v - \sum_{l=1}^{L} \gamma_l \overrightarrow{u}_l = \sum_{l=1}^{L} (\beta_{v,l} - \gamma_l)\overrightarrow{u}_l$$

in R_v since $(\beta_{v,l} - \gamma_l) \in \mathscr{O}_v$, each $\overrightarrow{u}_l \in R_v$ and R_v is an $\mathscr{O}_v$-module. We have shown that $\overrightarrow{\xi} = \sum_{l=1}^{L} \gamma_l \overrightarrow{u}_l$ is in k^L and $(\overrightarrow{\alpha}_v - \overrightarrow{\xi}) \in \tau\mathscr{R}$ for all $\tau > \nu(k) \sum_{l=1}^{L} \lambda_l$. This proves the inequality on the right-hand side of (5.4).

If $\mathscr{R} \subseteq (k_{\mathbf{A}})^L$ is admissible we may regard $V(\mathscr{R})$ as fixed and then ask how large the sum $\lambda_1 + \lambda_2 + \cdots + \lambda_L$ can be subject to the constraint (5.1) arising from Minkowski's second theorem and a *lower* bound on λ_1. If $0 < \lambda_0 \leq \lambda_1$ it is easy to prove that $\lambda_1 + \lambda_2 + \cdots + \lambda_L$ is largest when $\lambda_0 = \lambda_1 = \lambda_2 = \cdots = \lambda_{L-1}$ and $(\lambda_0)^{d(L-1)}\lambda_L^d V(\mathscr{R}) = 2^{dL}$. This argument is due to Scherk [21] and shows that

$$\mu(\mathscr{R}) \leq \nu(k)\{(L-1)\lambda_0 + 2^L \lambda_0^{1-L} V(\mathscr{R})^{-1/d}\}. \tag{5.5}$$

6. Proof of Theorem 3

Let A be an $M \times N$ matrix over k, $1 \leq \operatorname{rank}(A) = M < N$ and suppose that $\overrightarrow{b} \in k^M$, $\overrightarrow{b} \neq \overrightarrow{0}$. We define $\mathfrak{X} = \{\overrightarrow{x} \in k^N : A\overrightarrow{x} = \overrightarrow{0}\}$ and we assume that $\{\overrightarrow{x}_1, \overrightarrow{x}_2, \ldots, \overrightarrow{x}_L\}$ is a basis for $\mathfrak{X}$ over k, where $L + M = N$. We write $X = (\overrightarrow{x}_1 \overrightarrow{x}_2 \cdots \overrightarrow{x}_L)$ for the corresponding $N \times L$ matrix. Plainly $\{\overrightarrow{x}_1, \overrightarrow{x}_2, \ldots, \overrightarrow{x}_L\}$ is also a basis for

$$\mathfrak{X}_v = \{\overrightarrow{x} \in (k_v)^N : A\overrightarrow{x} = \overrightarrow{0}\}$$

over k_v at each place v. Let $P_v = P_v(\mathfrak{X}_v)$ be projection onto $\mathfrak{X}_v$ as defined by (2.6) or (2.7).

Next we let $\overrightarrow{\xi}$ be a vector in k^N which satisfies $A\overrightarrow{\xi} = \overrightarrow{b}$. As $\operatorname{rank}(A) = M < N$ such a vector clearly exists. We use $\overrightarrow{\xi}$ to construct an admissible subset of $(k_{\mathbf{A}})^L$ as follows. If $v \nmid \infty$ we set

$$R_v = \{\overrightarrow{y} \in (k_v)^L : |X\overrightarrow{y}|_v \leq \max\{1, |(1_N - P_v)\overrightarrow{\xi}|_v\}\}$$

and if $v \mid \infty$ we let

$$R_v = \{\overrightarrow{y} \in (k_v)^L : \|X\overrightarrow{y}\|_v < (1 + \|(1_N - P_v)\overrightarrow{\xi}\|_v^2)^{1/2}\}.$$

Then the product

$$\mathscr{R} = \prod_v R_v \subseteq (k_{\mathbf{A}})^L$$

is easily seen to be admissible. By using the v-adic cube slicing identity, which is (4.8) and (4.9) of [2], we find that for $v \nmid \infty$ the Haar measure of R_v is given by

$$\beta_v^L(R_v) = |\mathscr{D}_v|_v^{dL/2} H_v(X)^{-d} \max\{1, |(1_N - P_v)\vec{\xi}|_v^{dL}\}. \tag{6.1}$$

If $v \mid \infty$ then a standard calculation using (2.4) leads to

$$\beta_v^L(R_v) = r_v(L)^{-d_v L} H_v(X)^{-d} (1 + \|(1_N - P_v)\vec{\xi}\|_v^2)^{d_v L/2}. \tag{6.2}$$

When we combine (2.5), (6.1), and (6.2) we find that

$$\begin{aligned} V(\mathscr{R}) &= \prod_v \beta_v^L(R_v) \\ &= \left\{ |\Delta_k|^{1/2d} \prod_{v|\infty} r_v(L)^{d_v/d} \right\}^{-dL} H(X)^{-d} \\ &\quad \cdot \left\{ \prod_{v|\infty} (1 + \|(1_N - P_v)\vec{\xi}\|_v^2)^{d_v L/2} \right\} \\ &\quad \cdot \left\{ \prod_{v\nmid\infty} \{1, |(1_N - P_v)\vec{\xi}|_v^{dL}\} \right\}. \end{aligned}$$

Now we use (2.3) and apply Theorem 4 at each place v. In this way we establish the identity

$$V(\mathscr{R}) = H(X)^{-d} \left\{ \frac{2H(A\vec{b})}{c_k(L)H(A)} \right\}^{dL}. \tag{6.3}$$

Let $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_L < \infty$ denote the successive minima of $\mathscr{R}$. From the definition of λ_1 there exists a vector $\vec{u}_1 \neq \vec{0}$ in k^L so that $\vec{u}_1 \in \lambda\mathscr{R}$ whenever $\lambda_1 < \lambda$. It follows that

$$\begin{aligned} 1 &\le \prod_v |X\vec{u}_1|_v \\ &\le \lambda_1 \left\{ \prod_{v|\infty} (1 + \|1_N - P_v)\vec{\xi}\|_v^2)^{d_v/2d} \right\} \\ &\quad \cdot \left\{ \prod_{v\nmid\infty} \max\{1, |(1_N - P_v)\vec{\xi}|_v\} \right\} \\ &= \lambda_1 \left(\frac{H(A\vec{b})}{H(A)} \right). \end{aligned} \tag{6.4}$$

When we apply (5.5) with (6.3) and the lower bound (6.4) for λ_1 we obtain the upper bound

$$\mu(\mathscr{R}) \leq \nu(k)\{(L-1) + c_k(L)^L H(A)\} \left\{ \frac{H(A)}{H(A\vec{b})} \right\} \tag{6.5}$$

for the inhomogeneous minimum of $\mathscr{R}$.

At each place v we have $P_v\vec{\xi} \in \mathfrak{X}_v$. As the columns of X form a basis for $\mathfrak{X}_v$ over k_v, there exists an $\vec{\alpha}_v \in (k_v)^L$ such that $P_v\vec{\xi} = X\vec{\alpha}_v$. Using (2.9) we have $|X\vec{\alpha}_v|_v \leq |\vec{\xi}|_v = 1$ at almost all finite places v. Thus $(\vec{\alpha}_v)$ may be regarded as a point in $(k_{\mathbf{A}})^L$. Therefore, if $\mu = \mu(\mathscr{R}) < \tau$ there exists a point $\vec{z} \in k^L$ so that $(\vec{\alpha}_v - \vec{z}) \in \tau\mathscr{R}$. In fact, k^L is discrete in $(k_{\mathbf{A}})^L$ and the closure of each $\tau\mathscr{R}$ is compact, hence we may assume that $(\vec{\alpha}_v - \vec{z})$ is in the closure of $\mu\mathscr{R}$. In other words, we have

$$|X(\vec{\alpha}_v - \vec{z})|_v \leq \max\{1, |(1_N - P_v)\vec{\xi}|_v\} \tag{6.6}$$

if $v \nmid \infty$ and

$$\|X(\vec{\alpha}_v - \vec{z})\|_v \leq \mu(1 + \|(1_N - P_v)\vec{\xi}\|_v^2)^{1/2} \tag{6.7}$$

if $v \mid \infty$. We will now complete the proof by showing that (1.4) holds with $\vec{\zeta} = \vec{\xi} - X\vec{z}$ and that $\vec{\zeta} \in (\mathscr{O}_S)^L$. Of course $A\vec{\zeta} = A\vec{\xi} = \vec{b}$ since the columns of X span $\mathfrak{X}$.

If $v \nmid \infty$ then

$$\begin{aligned} |\vec{\zeta}|_v &= \max\{|P_v\vec{\zeta}|_v, |(1_N - P_v)\vec{\zeta}|_v\} \\ &= \max\{|X(\vec{\alpha}_v - \vec{z})|_v, |(1_N - P_v)\vec{\xi}|_v\} \\ &\leq \max\{1, |(1_N - P_v)\vec{\xi}|_v\} \end{aligned}$$

by (6.6). Thus, we have

$$\max\{1, |\vec{\zeta}|_v\} \leq \max\{1, |(1_N - P_v)\vec{\xi}|_v\} = \frac{H_v(A\vec{b})}{H_v(A)}, \tag{6.8}$$

and this shows that $\vec{\zeta} \in (\mathscr{O}_S)^L$. If $v \mid \infty$ then

$$\begin{aligned} \|\vec{\zeta}\|_v^2 &= \|P_v\vec{\zeta}\|_v^2 + \|(1_N - P_v)\vec{\zeta}\|_v^2 \\ &= \|X(\alpha_v - \vec{z})\|_v^2 + \|(1_N - P_v)\vec{\xi}\|_v^2 \\ &\leq \mu^2(1 + \|(1_N - P_v)\vec{\xi}\|_v^2) + \|(1_N - P_v)\vec{\xi}\|_v^2 \end{aligned}$$

and therefore

$$\begin{aligned} (1 + \|\vec{\zeta}\|_v^2)^{d_v/2d} &\leq (1+\mu^2)^{d_v/2d}(1 + \|(1_N - P_v)\vec{\xi}\|_v^2)^{d_v/2d} \\ &= (1+\mu^2)^{d_v/2d} \left\{ \frac{H_v(A\vec{b})}{H_v(A)} \right\}. \end{aligned} \tag{6.9}$$

From (6.8) and (6.9) we conclude that

$$H\begin{pmatrix}\vec{\zeta}\\ 1\end{pmatrix} \leq (1+\mu^2)^{1/2}\left\{\frac{H(A\vec{b})}{H(A)}\right\}. \tag{6.10}$$

Finally, the inequality (1.4) in the statement of the theorem follows from (6.5) and (6.10).

7. An estimate for $\nu(k)$

In general a precise evaluation of $\nu(k)$ in terms of the elementary invariants of k does not seem to be known. For many qualitative results it is important only to know that $\nu(k)$ is finite (note the discussion in Cassels [9, pp. 64–68]). Here we would like the inequality (1.4) to be as explicit as possible and so we provide the following estimate.

Theorem 6. *Let k be an algebraic number field but assume that k is not isomorphic to $\mathbb{Q}(\sqrt{-1})$ and k is not isomorphic to $\mathbb{Q}(\sqrt{-3})$. Then*

$$\nu(k) \leq \frac{1}{2}\left(\frac{2}{\pi}\right)^s |\Delta_k|^{1/2}, \tag{7.1}$$

where Δ_k is the discriminant of k and s is the number of complex places of k. More precisely, if $D \leq -1$ is a square free integer and $k \cong \mathbb{Q}(\sqrt{D})$ we have

(i) *if $D \equiv 1 \bmod 4$ then*

$$\nu(\mathbb{Q}(\sqrt{D})) = \tfrac{1}{4}|D|^{1/2}(1+|D|^{-1}) = \tfrac{1}{4}|\Delta_k|^{1/2}(1+|\Delta_k|^{-1}), \tag{7.2}$$

(ii) *if $D \not\equiv 1 \bmod 4$ then*

$$\nu(\mathbb{Q}(\sqrt{D})) = \tfrac{1}{2}|D|^{1/2}(1+|D|^{-1})^{1/2} = \tfrac{1}{4}|\Delta_k|^{1/2}(1+4|\Delta_k|^{-1})^{1/2}. \tag{7.3}$$

In particular we have

$$\nu(\mathbb{Q}(\sqrt{-1})) = \frac{1}{\sqrt{8}}|\Delta_{\mathbb{Q}(\sqrt{-1})}|^{1/2}, \tag{7.4}$$

$$\nu(\mathbb{Q}(\sqrt{-3})) = \frac{1}{3}|\Delta_{\mathbb{Q}(\sqrt{-3})}|^{1/2}. \tag{7.5}$$

Before proving Theorem 6 we make a few remarks. The inequality (7.1) is deduced from a more general theorem of Birch [1] and Kneser [19] together with known lower bounds for the discriminant of number fields. During the course of the argument it is necessary to directly evaluate $\nu(k)$ in four special cases and these are all imaginary quadratic fields. In fact the evaluation of $\nu(k)$ for an arbitrary imaginary quadratic field follows easily from a result of J. Heinhold [15] and we record it here as (7.2) and (7.3).

It is easy to see that $\nu(k)$ can be defined using $\mathcal{O}_k$ and only the Archimedean completions of k. Suppose that $\nu^*(k)$ is the infimum of real numbers $\tau > 0$ such that for each (γ_v) in $\prod_{v|\infty} k_v$ there exists ξ in $\mathcal{O}_k$ such that $\|\gamma_v - \xi\|_v < \tau$ at each $v \mid \infty$. By setting $\gamma_v = 0$ for $v \nmid \infty$ it is clear that $\nu^*(k) \leq \nu(k)$. Now let (α_v) be a point of $k_{\mathbf{A}}$. By the strong approximation theorem there exists $\eta \in k$ such that $\|\alpha_v - \eta\|_v \leq 1$ at each $v \nmid \infty$. If $\xi \in \mathcal{O}_k$ then $\|\alpha_v - \eta - \xi\|_v \leq 1$ continues to hold at each finite v. It follows that $\nu(k) \leq \nu^*(k)$.

Let $\sigma_1, \sigma_2, \dots, \sigma_d$ be the distinct embeddings of k into the field $\mathbb{C}$ of complex numbers. We assume that $\sigma_1, \sigma_2, \dots, \sigma_r$ are real, that $\sigma_{r+1}, \sigma_{r+2}, \dots, \sigma_{r+s}$ are complex and not real, and that $\overline{\sigma}_{r+j} = \sigma_{r+s+j}$ for $j = 1, 2, \dots, s$. Then we suppose that $\omega_1, \omega_2, \dots, \omega_d$ is an integral basis for $\mathcal{O}_k$. We write Ω for the $d \times d$ matrix $\Omega = (\sigma_i(\omega_j))$, where $i = 1, 2, \dots, d$ indexes rows and $j = 1, 2, \dots, d$ index columns. Then we define W to be the $d \times d$ matrix which is organized into blocks as

$$W = \begin{pmatrix} 1_r & 0 & 0 \\ 0 & \frac{1}{2}1_s & \frac{1}{2}1_s \\ 0 & \frac{1}{2i}1_s & \frac{-1}{2i}1_s \end{pmatrix}.$$

We note that $\det W = (-2i)^{-s}$. Next we write $C_{r,s}$ for the open, convex, symmetric subset of $\mathbb{R}^d$ defined by

$$C_{r,s} = \{\vec{x} \in \mathbb{R}^d : |x_i| < 1 \text{ if } 1 \le i \le r, \text{ and } (x_{r+j})^2 + (x_{r+s+j})^2 < 1 \text{ if } 1 \le j \le s\}.$$

A routine computation shows that $\mathrm{Vol}_d(C_{r,s}) = 2^d(\pi/4)^s$. It follows that the set $K = \{\vec{x} \in \mathbb{R}^d : W\Omega\vec{x} \in C_{r,s}\}$ is open, convex, symmetric and has

$$\mathrm{Vol}_d(K) = |\det W\Omega|^{-1} \mathrm{Vol}_d(C_{r,s}) = 2^d \left(\frac{\pi}{2}\right)^s |\Delta_k|^{-1/2}.$$

We recall that each Archimedean absolute value $\xi \to \|\xi\|_v$ on $\mathcal{O}_k$ coincides with a map $\xi \to |\sigma_j(\xi)|$, $1 \le j \le r+s$, and this correspondence is bijective. As each $\xi \in \mathcal{O}_k$ has a unique expansion

$$\xi = \sum_{j=1}^{d} u_j \omega_j, \qquad \vec{u} \in \mathbb{Z}^d, \tag{7.6}$$

it follows easily that $\nu(k) = \nu^*(k)$ is equal to the inhomogeneous minimum of K with respect to the lattice $\mathbb{Z}^d$. That is, $\nu(k) = \mu(K)$ where

$$\mu(K) = \inf\left\{\tau > 0 : \mathbb{R}^d \subseteq \bigcup_{\vec{u} \in \mathbb{Z}^d} (\tau K + \vec{u})\right\}.$$

In particular, we have $\|\xi\|_v < \tau$ at each $v \mid \infty$ if and only if $\vec{u} \in \tau K$, where ξ and $\vec{u}$ are related by (7.6). This also shows that $1 = \lambda_1 = \lambda_1(K, \mathbb{Z}^d)$, where λ_1 is the first successive minima of K with respect to the lattice $\mathbb{Z}^d$. By a result of Birch [1] and Kneser [19] (see also Cassels [8, pp. 315–317]), if

$$d \le 2^d \lambda_1^{-d} (\mathrm{Vol}_d(K))^{-1} = \left(\frac{2}{\pi}\right)^s |\Delta_k|^{1/2} \tag{7.7}$$

then $\nu(k) = \mu(K) \le \frac{1}{2}(2/\pi)^s |\Delta_k|^{1/2}$, which is (7.1). Thus it remains to determine those number fields k which satisfy (7.7).

If we use the classical lower bound $(\pi/4)^s(d^d/d!) < |\Delta_k|^{1/2}$ due to Minkowski we find that (7.7) holds whenever $5 \le d$. In order to deal with the remaining cases let $\delta(r, s)$ denote the minimum value of $|\Delta_k|$ over all number fields k having r real and s complex places. The value of $\delta(r, s)$ has been determined at least for $r + 2s = d \le 5$ and we record it here in Table 1.

TABLE 1

d	r	s	$\delta(r, s)$	$(2/\pi)^s\delta(r, s)^{1/2}$
1	1	0	1	1.0000⋯
2	0	1	3	1.1026⋯
2	2	0	5	2.2360⋯
3	1	1	23	3.0531⋯
3	3	0	49	7.0000⋯
4	0	2	117	4.3838⋯
4	2	1	275	10.5571⋯
4	4	0	725	26.9258⋯
5	1	2	1609	16.2569⋯
5	3	1	4511	42.7579⋯
5	5	0	14641	121.0000⋯

For cubic, quartic and quintic fields the values of $\delta(r, s)$ were found by H. Davenport [10], J. Mayer [20] and J. Hunter [17], respectively. Plainly they demonstrate that (7.7) holds for $1 \leq d \leq 5$ except possibly when k is an imaginary quadratic field. We now proceed to consider this exceptional case.

Let $D \leq -1$ be a square free integer. As is well known, if $D \equiv 1 \mod 4$ then $\Delta_{\mathbb{Q}(\sqrt{D})} = D$ and if $D \not\equiv 1 \mod 4$ then $\Delta_{\mathbb{Q}(\sqrt{D})} = 4D$. It is therefore easy to verify that (7.7) holds for all imaginary quadratic fields $\mathbb{Q}(\sqrt{D})$ with exactly four exceptions and these are the fields $\mathbb{Q}(\sqrt{-1})$, $\mathbb{Q}(\sqrt{-2})$, $\mathbb{Q}(\sqrt{-3})$, and $\mathbb{Q}(\sqrt{-7})$. Let $F_D(x_1, x_2)$ be the positive definite quadratic form defined by

$$F_D(x_1, x_2) = x_1^2 + x_1x_2 + \left(\frac{1-D}{4}\right)x_2^2 \quad \text{if } D \equiv 1 \mod 4$$

and

$$F_D(x_1, x_2) = x_1^2 - Dx_2^2 \quad \text{if } D \not\equiv 1 \mod 4.$$

Then we have

$$\nu(\mathbb{Q}(\sqrt{D})) = \left\{ \sup_{\vec{y} \in \mathbb{R}^2} \left(\inf_{\vec{u} \in \mathbb{Z}^2} F_D(u_1 - y_1, u_2 - y_2) \right) \right\}^{1/2}. \tag{7.8}$$

For example, if $D \equiv 1 \mod 4$ then $\{1, \frac{1}{2}(1 + \sqrt{D})\}$ is an integral basis for $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$. An arbitrary complex number z can be written as $y_1 + \frac{1}{2}(1 + \sqrt{D})y_2$ with $\vec{y} \in \mathbb{R}^2$ and an arbitrary point $\xi \in \mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ can be written as $\xi = u_1 + \frac{1}{2}(1 + \sqrt{D})u_2$ with $\vec{u} \in \mathbb{Z}^2$. It follows that $|\xi - z|^2 = F_D(u_1 - y_1, u_2 - y_2)$ and (7.8) is now obvious. The case $D \not\equiv 1 \mod 4$ is similar. Finally, the expression on the right of (7.8) has been evaluated by J. Heinhold [15, Satz 3, p. 667] and this leads immediately to (7.2) and (7.3). It is now trivial to verify that the bound (7.1) holds for $\mathbb{Q}(\sqrt{-2})$ and $\mathbb{Q}(\sqrt{-7})$ but *fails* to hold in the two cases $\mathbb{Q}(\sqrt{-1})$ and $\mathbb{Q}(\sqrt{-3})$. For these we have the identities (7.4) and (7.5) from the result of Heinhold. This completes the proof of Theorem 6.

REFERENCES

1. B. J. Birch, *A transference theorem of the geometry of numbers*, J. London Math. Soc. **31** (1956), 248–251.
2. E. Bombieri and J. Vaaler, *On Siegel's lemma*, Invent. Math. **73** (1983), 11–32.

3. I. Borosh, *A sharp bound for positive solutions of homogeneous linear Diophantine equations*, Proc. Amer. Math. Soc. **60** (1976), 19–21.
4. I. Borosh and L. B. Treybig, *Bounds on positive integral solutions of linear Diophantine equations*, Proc. Amer. Math. Soc. **55** (1976), 299–304.
5. ____, *Bounds on positive integral solutions of linear Diophantine equations*. II, Canad. Math. Bull. (3) **22** (1979), 357–361.
6. I. Borosh, M. Flahive and L. B. Treybig, *Small solutions of linear Diophantine equations*, Discrete Math. **58** (1986), 215–220.
7. I. Borosh, M. Flahive, D. Rubin and L. B. Treybig, *A sharp bound for solutions of linear Diophantine equations*, Proc. Amer. Math. Soc. **105** (1989), 844–846.
8. J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, New York, 1959.
9. ____, *Global fields*, Algebraic Number Theory (J. W. S. Cassels and A. Fröhlich, eds.), Academic Press, London 1967.
10. H. Davenport, *On the product of three homogeneous linear forms*. III, Proc. London Math. Soc. **45** (1939), 98–125.
11. G. Frobenius, *Theorie der linearen formen mit ganzen Coefficienten*, J. Reine Angew. Math. **86** (1879), 146–208.
12. P. Gordan, *Über den grössten gemeinsamen Factor*, Math. Ann. **7** (1873), 443–448.
13. P. M. Gruber and C. G. Lekkerkerker, *Geometry of numbers*, 2nd ed., North-Holland, New York, 1987.
14. I. Heger, *Sitzungsber*, Akad. Wiss. Wien (Math.) **21** (1856), 550–560.
15. J. Heinhold, *Verallgemeinerung und Verschärfung eines Minkowskischen Satzes*, Math. Z. **44** (1939), 659–688.
16. W. V. D. Hodge and D. Pedoe, *Methods of algebraic geometry*, vol. 1, Cambridge Univ. Press, 1968.
17. J. Hunter, *The minimum discriminant of quintic fields*, Proc. Glasgow Math. Assoc. **3** (1957), 57–67.
18. V. Jarnik, *Zwei Bemerkungen zur Geometrie de Zahlen*, Věstnik Krălovské České Společnosit Nauk, 1941.
19. M. Kneser, *Ein Satz über abelsche Gruppen mit Anwendungen auf die Geometrie der Zahlen*, Math. Z. **61** (1955), 429–434.
20. J. Mayer, *Die absolut-kleinsten Diskriminanten der biquadratischen Zahlkörper*, S. B. Akad. Wiss. Wien. IIa **138** (1929), 733–742.
21. P. Scherk, *Convex bodies off center*, Arch. Math. **3** (1950), 303.
22. H. J. S. Smith, *On systems of linear indeterminate equations and congruences*, Philos. Trans. Roy. Soc. London **151** (1861), 293–326 (=Collected Math. Papers, I, pp. 367–409).
23. T. Struppeck and J. D. Vaaler, *Inequalities for heights on algebraic subspaces and the Thue-Siegel principle*, Analytic Number Theory, Proc. Conf. in honor of P. T. Bateman (B. C. Berndt, H. G. Diamond, H. Halberstam, A. Hilderbrand, eds.), Birkhäuser, Boston, Mass., 1990.
24. J. D. Vaaler, *Small zeros of quadratic forms over number fields*, Trans. Amer. Math. Soc. **302** (1987), 281–296.
25. A. Weil, *Basic number theory*, Springer-Verlag, New York, 1974.

SCHOOL OF NATURAL AND MATHEMATICAL SCIENCES, SEATTLE PACIFIC UNIVERSITY, SEATTLE, WASHINGTON 98119

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT AUSTIN, AUSTIN, TEXAS 78712