ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 26, Number 3, Summer 1996

BASES OF NUMBER FIELDS WITH SMALL HEIGHT

DAMIEN ROY AND JEFFREY LIN THUNDER

ABSTRACT. For every number field viewed as a vector space over the rational numbers, we prove there exists a basis with height that is small in comparison to the absolute value of the discriminant. We get best-possible results in the case of a totally real number field and other cases as well.

0. Introduction. Let K be a number field of degree d, and suppose that $a_1, \dots, a_d$ are elements of K which are linearly independent over $\mathbf{Q}$. A more general result due to J. Silverman (Theorem 2 of [**10**]) gives, in this instance,

$$H_K(a_1, \dots, a_d) \geq d^{-d/2} |D(K)|^{1/2},$$

where H_K is a standard multiplicative field height (defined below) and $D(K)$ is the discriminant of K. This inequality has connections with Siegel's lemma over number fields (see [**1** and **8**]) and Northcott's theorem on points of bounded height in $\overline{\mathbf{Q}}^n$ (see [**9**]).

Among all bases $\{a_1, \dots, a_d\}$ of K viewed as a d-dimensional vector space over $\mathbf{Q}$, there is one with smallest height. We denote this smallest height by $B(K)$. Silverman's result implies

$$B(K) \geq d^{-d/2} |D(K)|^{1/2}.$$

In this paper we deal with upper bounds for the quantity $B(K)$. The obvious question here is whether, for all $d \geq 1$, there is a constant $c(d)$ such that $B(K) \leq c(d) |D(K)|^{1/2}$ for all number fields K of degree d. Alternatively, one may wish to consider some family of fields with infinitely many members of degree d and ask for the same type of bound for those K in the family. Also, any basis $\{a_1, \dots, a_d\}$ of a number field K of degree d viewed as a vector space over $\mathbf{Q}$ determines a fractional ideal $\mathfrak{O}_K a_1 + \cdots + \mathfrak{O}_K a_d$, where $\mathfrak{O}_K$ denotes the ring of integers in K.

First author partially supported by NSERC.
Received by the editors on October 25, 1994, and in revised form on May 15, 1995.

Another question is whether or not, for each K of degree d (or of degree d in some given family), there is a basis $\{a_1, \dots, a_d\}$ of K with $H_K(a_1, \dots, a_d)$ close to $B(K)$ and $\mathfrak{O}_K a_1 + \cdots + \mathfrak{O}_K a_d = \mathfrak{O}_K$. Here by "close" we mean within some constant multiple depending only on d.

If one restricts attention to totally real number fields, the answer to both the above questions is yes (see [**8**] for other examples). Specifically, we have the following result.

Theorem 1. *Let K be a totally real number field of degree d. Then*

$$B(K) \le C_1(d)|D(K)|^{1/2},$$

where $C_1(d) = 2^{d(3d+1)/2}$. Moreover, there is a $\mathbf{Z}$-basis $\{b_1, \dots, b_d\}$ of $\mathfrak{O}_K$ with

$$H_K(b_1, \dots, b_d) \le (d/2)^d C_1(d)|D(K)|^{1/2}.$$

For number fields with complex places these questions become more difficult. Following Masser and Wüstholz (see [**7**]), we define the *class index* of K to be the smallest positive integer $i(K)$ such that each ideal class contains an ideal of $\mathfrak{O}_K$ with norm no larger than $i(K)$. We will prove the following.

Theorem 2. *For arbitrary number fields K of degree d, we have*

$$B(K) \le C_1(d)^2 \frac{|D(K)|}{i(K)},$$

where $C_1(d)$ is as above in Theorem 1.

Fix $\varepsilon > 0$ and an integer $d \ge 1$. The Brauer-Siegel theorem (see Chapter XVI of [**4**]) shows that, for any number field K of degree d, the class number $h(K)$ and the regulator $R(K)$ of K are related to the discriminant $D(K)$ by $h(K)R(K) \gg_{d,\varepsilon} |D(K)|^{1/2-\varepsilon}$, where the implicit constant depends only on d and ε and is ineffective. On the other hand, D. Masser in [**6**] shows that $h(K) \le i(K)(1 + \log i(K))^{d-1}$ for all number fields K of degree d. By virtue of Theorem 2, this gives

Corollary. *For any number field K of degree d and any ε with $0 < \varepsilon < 1$, we have*

$$B(K) \ll_{d,\varepsilon} R(K)^{1-\varepsilon} |D(K)|^{1/2+2\varepsilon},$$

where the implicit constant depends on d and ε and is ineffective.

In particular, for imaginary quadratic number fields K we get $B(K) \ll_\varepsilon |D(K)|^{1/2+\varepsilon}$ for all $\varepsilon > 0$, where the implicit constant depends (ineffectively) only on ε. We will show in Section 3 that these fields also satisfy $B(K) \geq |D(K)|/(4i(K))$. So the upper bound given by Theorem 2 is essentially best possible for imaginary quadratic number fields. More generally, in view of Silverman's lower bound for $B(K)$ above, the inequality in Theorem 2 is sharp (up to a factor depending only on d) for families of fields K of degree d satisfying $i(K) \gg_d |D(K)|^{1/2}$ with an implicit constant depending only on d. In private communication with the authors, D. Masser has shown how to construct such families of fields for all even degrees. We give his construction explicitly in Section 3 below. In [**6**] he also constructs, for any $\varepsilon > 0$ and integer $d > 1$, infinitely many number fields K of degree d with $i(K) \gg_{d,\varepsilon} |D(K)|^{1/2-\varepsilon}$, where the implicit constant depends only on ε and d (though again ineffectively).

1. Definitions. For K a number field we let $M(K)$ denote the set of places of K. For each $v \in M(K)$, let $|\cdot|_v$ denote the absolute value on K that extends the usual absolute value on $\mathbf{Q}$ if $v|\infty$, or the usual p-adic absolute value on $\mathbf{Q}$ if $v|p$. Let n_v denote the local degree of K at v. We define a norm $||\cdot||_v$ on K^n for each place v by

$$||(x_1, \ldots, x_n)||_v = \max_{1 \leq i \leq n} \{|x_i|_v\}.$$

With this notation we define the *height* of a nonzero vector $\mathbf{x} \in K^n$ by

$$H_K(\mathbf{x}) = \prod_{v \in M(K)} ||\mathbf{x}||_v^{n_v}.$$

For K of degree d, we write $d = r_1 + 2r_2$, where r_1 is the number of real places and r_2 is the number of complex places. Denote the embeddings of K into $\mathbf{C}$ by $\alpha \mapsto \alpha^{(i)}$ and order them so that the first r_1 are real and $\alpha^{(i)}$ is the complex conjugate of $\alpha^{(r_2+i)}$ for $r_1 < i \leq r_1 + r_2$. Let $\rho : K \to \mathbf{R}^d$ be defined by

$$\rho(\alpha) = (\alpha^{(1)}, \dots, \alpha^{(r_1)}, \mathrm{Re}\,(\alpha^{(r_1+1)}), \dots, \mathrm{Re}\,(\alpha^{(r_1+r_2)}),\\ \mathrm{Im}\,(\alpha^{(r_1+1)}), \dots, \mathrm{Im}\,(\alpha^{(r_1+r_2)})),$$

where Re and Im denote the real and imaginary part, respectively. Then for $\mathfrak{A}$ any fractional ideal of K we have that $\rho(\mathfrak{A})$ is a lattice in $\mathbf{R}^d$ of determinant $2^{-r_2}N(\mathfrak{A})|D(K)|^{1/2}$ (Lemma 2, Section 2, Chapter V of [**4**]).

2. Proof of Theorems 1 and 2.

Proof of Theorem 1. Let K be a totally real number field of degree d and let $\rho : K \to \mathbf{R}^d$ be as above. Then $\Lambda = \rho(\mathfrak{O}_K)$ is a lattice of determinant $|D(K)|^{1/2}$. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_d$ be the successive minima of Λ with respect to the unit cube $[-1,1]^d$ of $\mathbf{R}^d$.

By a result due to J.H. Evertse (Lemma 3.3.5 of [**3**]), there exist linearly independent lattice points $\mathbf{x}_1 = (x_{1,1}, \dots, x_{1,d}), \dots, \mathbf{x}_d = (x_{d,1}, \dots, x_{d,d}) \in \Lambda$ and a permutation σ of $\{1, \dots, d\}$ that satisfy

$$|x_{i,j}| \leq 2^{i+\sigma(j)} \min\{\lambda_i, \lambda_{\sigma(j)}\}$$

for all i and j. Write $\mathbf{x}_i = \rho(a_i)$ with $a_i \in \mathfrak{O}_K$ for each i. We then have

$$\begin{aligned}\prod_{j=1}^{d} \max_{1\leq i\leq d}\{|a_i^{(j)}|\} &= \prod_{j=1}^{d} \max_{1\leq i\leq d}\{|x_{i,j}|\}\\ &\leq \prod_{j=1}^{d} 2^{d+\sigma(j)}\lambda_{\sigma(j)}\\ &= C_1(d)\prod_{j=1}^{d}\lambda_j.\end{aligned}$$

By Minkowski's theorem, we have

$$\prod_{j=1}^{d}\lambda_j \leq \det(\Lambda) = |D(K)|^{1/2}.$$

By the definition of height, we get $H_K(a_1, \dots, a_d) \le C_1(d)|D(K)|^{1/2}$. This proves the first part of Theorem 1.

For the second part, note that the successive minima of the convex body

$$C_\sigma = \{(y_1, \dots, y_d) \in \mathbf{R}^d : |y_j| \le 2^{d+\sigma(j)}\lambda_{\sigma(j)} \quad \text{for } 1 \le j \le d\}$$

with respect to Λ are all no greater than 1. By a result of Mahler (Lemma 8, Chapter V of [**2**]), there is a basis $\{\rho(b_1), \dots, \rho(b_d)\}$ of Λ that satisfies $\rho(b_i) \in (i/2)C_\sigma$ for each $i > 1$ and $\rho(b_1) \in C_\sigma$. Whence

$$\begin{aligned} H_K(b_1, \dots, b_d) &\le \prod_{j=1}^{d} \max_{1\le i\le d}\{|b_i^{(j)}|\} \\ &\le \prod_{j=1}^{d} (d/2)2^{d+\sigma(j)}\lambda_{\sigma(j)} \\ &= (d/2)^d 2^{d^2+d(d+1)/2} \prod_{j=1}^{d} \lambda_j \\ &\le (d/2)^d C_1(d)|D(K)|^{1/2} \end{aligned}$$

by Minkowski's theorem. This completes the proof of Theorem 1. □

Lemma. *Let K be a number field and let $\mathfrak{A}$ be an ideal of $\mathfrak{O}_K$ with smallest norm among all ideals in the same class. Then*

$$\min_{\alpha}\{|N(\alpha)|\} = 1,$$

where the minimum is over all nonzero $\alpha \in \mathfrak{A}^{-1}$.

Proof. Let $\alpha \in \mathfrak{A}^{-1}$, $\alpha \ne 0$. Let $\mathfrak{B} = (\alpha)\mathfrak{A}$. Then $\mathfrak{B}$ is an ideal of $\mathfrak{O}_K$ and $\mathfrak{B}$ is in the same ideal class as $\mathfrak{A}$. Since $N(\mathfrak{B}) = |N(\alpha)|N(\mathfrak{A})$, we get $|N(\alpha)| \ge 1$. Finally, as $1 \in \mathfrak{A}^{-1}$, we get an equality for this minimum. □

Proof of Theorem 2. Let K be a number field of degree $d = r_1 + 2r_2$. Choose an ideal $\mathfrak{A}$ of $\mathfrak{O}_K$, a representative of its ideal class with smallest norm, satisfying $N(\mathfrak{A}) = i(K)$. Let $\Lambda = \rho(\mathfrak{A}^{-1}) \subset \mathbf{R}^d$. Let $\lambda_1 \le \cdots \le \lambda_d$ be the successive minima of Λ with respect to the unit cube. Choose $\alpha \in \mathfrak{A}^{-1}$ with $||\rho(\alpha)|| = \lambda_1$, where $||\rho(\alpha)||$ denotes the maximum norm of $\rho(\alpha)$. By the lemma we have $|N(\alpha)| \ge 1$. On the other hand, we find that $|N(\alpha)| \le (\sqrt{2}\lambda_1)^d$. Thus, $\lambda_1 \ge 1/\sqrt{2}$, so that $\lambda_i \ge 1/\sqrt{2}$ for all i.

Using Evertse's result again, we choose linearly independent $\mathbf{x}_1 = \rho(a_1), \dots, \mathbf{x}_d = \rho(a_d) \in \Lambda$ and a permutation σ that satisfy

$$|x_{i,j}| \le 2^{i+\sigma(j)} \min\{\lambda_i, \lambda_{\sigma(j)}\}.$$

Using this together with the lower bound on the successive minima above, we get

$$\begin{aligned}
\prod_{v|\infty} ||(a_1, \dots, a_d)||_v^{n_v} &= \\
&\prod_{j=1}^{r_1} \max_{1\le i\le d}\{|x_{i,j}|\} \prod_{j=r_1+1}^{r_1+r_2} \max_{1\le i\le d}\{|x_{i,j}|^2 + |x_{i,j+r_2}|^2\} \\
&\le \prod_{j=1}^{r_1} (2^{d+\sigma(j)}\lambda_{\sigma(j)}) \\
&\quad \times \prod_{j=r_1+1}^{r_1+r_2} ((2^{d+\sigma(j)}\lambda_{\sigma(j)})^2 + (2^{d+\sigma(j+r_2)}\lambda_{\sigma(j+r_2)})^2) \qquad (1)\\
&\le \prod_{j=1}^{d} (2^{d+\sigma(j)}\lambda_{\sigma(j)})^2 \\
&= (2^{d^2+d(d+1)/2})^2 \left(\prod_{j=1}^{d} \lambda_j\right)^2 \\
&\le C_1(d)^2 |D(K)| N(\mathfrak{A})^{-2},
\end{aligned}$$

by Minkowski's theorem.

Let $\mathfrak{B} = \mathfrak{O}_K a_1 + \cdots + \mathfrak{O}_K a_d$. Then $\mathfrak{B} \subseteq \mathfrak{A}^{-1}$, so that

$$\prod_{v\nmid\infty} ||(a_1, \dots, a_d)||_v^{n_v} = N(\mathfrak{B})^{-1} \le N(\mathfrak{A}).$$

By this and (1) we have

$$\begin{aligned} B(K) &\leq H_K(a_1, \ldots, a_d) \\ &\leq C_1(d)^2 |D(K)| N(\mathfrak{A})^{-1} \\ &= C_1(d)^2 \frac{|D(K)|}{i(K)}. \qquad \square \end{aligned}$$

3. Imaginary quadratic fields and class indices. For the case of imaginary quadratic number fields, we get a very explicit relationship between $B(K)$ and the class index.

Theorem 3. *Let K be an imaginary quadratic number field. Then*

$$\frac{|D(K)|}{4i(K)} \leq B(K) \leq \frac{|D(K)|}{3i(K)}.$$

Proof. We prove the upper bound in Theorem 3 in a similar fashion as above. Let K be an imaginary quadratic number field and let $\rho : K \to \mathbf{R}^2$ be defined as above. Choose an ideal $\mathfrak{A}$ of $\mathfrak{O}_K$ which is a representative of its class with least norm that satisfies $N(\mathfrak{A}) = i(K)$. Let $\lambda_1 \leq \lambda_2$ be the successive minima of $\rho(\mathfrak{A}^{-1})$ with respect to the unit disk in $\mathbf{R}^2$. Note that the Euclidean norm of $\rho(a)$ is $|N(a)|^{1/2}$ for any $a \in K$, so that $\lambda_1 = 1$ by the lemma. Let $a, b \in \mathfrak{A}^{-1}$ be linearly independent and satisfy $|a| = 1$, $|b| = \lambda_2$. Since $\rho(\mathfrak{A}^{-1})$ is a two-dimensional lattice, we have a and b form a $\mathbf{Z}$-basis for $\mathfrak{A}^{-1}$. Using the known value $\gamma_2 = 2/\sqrt{3}$ of Hermite's constant (see page 318 of [**5**]), we have

$$\begin{aligned} H_K(a, b) &= N(\mathfrak{A}) \max\{|a|^2, |b|^2\} \\ &= N(\mathfrak{A}) \lambda_2^2 \\ &= N(\mathfrak{A}) \lambda_1^2 \lambda_2^2 \\ &\leq N(\mathfrak{A}) \left(\frac{2 \det(\rho(\mathfrak{A}^{-1}))}{\sqrt{3}} \right)^2 \\ &= N(\mathfrak{A}) \left(\frac{|D(K)|^{1/2} N(\mathfrak{A})^{-1}}{\sqrt{3}} \right)^2 \\ &= \frac{|D(K)|}{3i(K)}. \end{aligned}$$

As for the lower bound in Theorem 3, let $B(K) = H_K(a,b)$ and write $\mathfrak{O}_K a + \mathfrak{O}_K b = \mathfrak{A}^{-1}$. Without loss of generality, $\mathfrak{A}$ is an integral ideal of smallest norm in its ideal class, so that $N(\mathfrak{A}) \le i(K)$. We have

$$B(K) = N(\mathfrak{A}) \max\{|a|^2, |b|^2\}.$$

Let $\lambda_1 \le \lambda_2$ be the successive minima of $\rho(\mathfrak{A}^{-1})$ with respect to the unit disk. Again we have $\lambda_1 = 1$, and by Hadamard's inequality $\lambda_2 = \lambda_1\lambda_2 \ge \det(\rho(\mathfrak{A}^{-1}))$. But since a and b are linearly independent over $\mathbf{Q}$, we must have

$$\max\{|a|, |b|\} \ge \lambda_2.$$

The lower bound in Theorem 3 follows. □

We remark that the same reasoning shows that $H_K(a,b) \ge |D(K)|/4$ for any imaginary quadratic field K and any basis $\{a,b\}$ of K over $\mathbf{Q}$ with $\mathfrak{O}_K a + \mathfrak{O}_K b = \mathfrak{O}_K$. In view of the lower bound for $B(K)$ given in Theorem 3, this implies that, for any $c > 0$, there are only finitely many imaginary quadratic number fields K which admit a basis $\{a,b\}$ over $\mathbf{Q}$ satisfying both $\mathfrak{O}_K a + \mathfrak{O}_K b = \mathfrak{O}_K$ and $H_K(a,b) \le cB(K)$. This answers, for these fields, one of the questions posed in the introduction.

As noted in the introduction, the upper bound for $B(K)$ given in Theorem 2 is sharp (up to a constant multiple depending only on the degree) for all fields K satisfying $i(K) \gg |D(K)|^{1/2}$ with an implicit constant depending only on the degree of K. In private communication with the authors, D. Masser expanded on an argument in Proposition 3 of [**8**] and constructed, for all positive integers d, infinitely many number fields K of degree $2d$ with this property. He was kind enough to allow the authors to give his construction here.

Theorem 4 (D. Masser). *Let d be a positive integer. There are infinitely many number fields K (none totally real) of degree $2d$ and a constant $C_2(d) > 0$, depending only on d, with $i(K) \ge C_2(d)|D(K)|^{1/2}$.*

Proof. Fix a totally real cyclic field F of degree d with odd discriminant $D(F)$. We claim that there are infinitely many square-free positive integers m which are relatively prime to $D(F)$ and divisible by a prime number p with $\sqrt{m/2} \le p \le \sqrt{m}$, such that p generates a prime ideal $p\mathfrak{O}_F$ of $\mathfrak{O}_F$.

To see this claim, note that the Chebotarev density theorem shows that there are infinitely many prime numbers p such that $p\mathfrak{O}_F$ is a prime ideal in $\mathfrak{O}_F$, these primes being those whose Artin automorphism in F generates the Galois group of F over $\mathbf{Q}$. Choose such a $p > D(F)$ and put $m = pq$, where q is a prime number with $p < q < 2p$. Then m and its divisor p have the required properties.

For such an m, let $K_m = F(\sqrt{-m})$. Note that the discriminant of $\mathbf{Q}(\sqrt{-m})$ is $-m$ or $-4m$, which in either case is relatively prime to $D(F)$. This shows that K_m has degree $2d$ over $\mathbf{Q}$ and we have

$$|D(K_m)|^{1/2} = D(F)|D(\mathbf{Q}(\sqrt{-m}))|^{d/2} \le D(F)(4m)^{d/2}. \tag{2}$$

Let p be a divisor of m as above. Then p ramifies in $\mathbf{Q}(\sqrt{-m})$ and is inert in F, so there is a unique prime $\mathfrak{P}$ of K_m that lies above p. It has degree d and ramification index 2 over $\mathbf{Q}$, so that its norm satisfies

$$(m/2)^{d/2} \le N(\mathfrak{P}) = p^d \le m^{d/2}. \tag{3}$$

Let $\mathfrak{A}$ be an integral representative of least norm in the ideal class containing $\mathfrak{P}^{-1}$. Then $\mathfrak{A} = (\alpha)\mathfrak{P}^{-1}$ for some nonzero $\alpha \in \mathfrak{P}$ and $N(\mathfrak{A}) = |N(\alpha)|N(\mathfrak{P})^{-1}$. Now, if $\alpha \in F$, then α belongs to $p\mathfrak{O}_F$, so that $|N(\alpha)| \ge N(p) = p^{2d} \ge (m/2)^d$. On the other hand, if $\alpha \notin F$, then we may write $\alpha = (a + b\sqrt{-m})/2$ with $a, b \in \mathfrak{O}_F$ and $b \ne 0$. Arguing as in the proof of Proposition 3 of [**8**], we get $|N(\alpha)| \ge N(b/2)m^d \ge (m/4)^d$. We conclude that $N(\mathfrak{A}) \ge (m/4)^d N(\mathfrak{P})^{-1}$. Theorem 3 follows from (2) and (3), using $C_2(d) = (8^d D(F))^{-1}$. □

Acknowledgments. The authors thank M. Waldschmidt, J.H. Evertse, and especially D. Masser for their helpful comments.

REFERENCES

1. E. Bombieri and J.D. Vaaler, *On Siegel's lemma*, Invent. Math. **73** (1983), 11–32.

2. J.W.S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1959.

3. J.H. Evertse, *The subspace theorem of W.M. Schmidt*, in *Diophantine approximation and Abelian varieties*, Lecture Notes in Math. **1566** (1993), 31–50.

4. S. Lang, *Algebraic number theory*, Springer-Verlag, New York, 1986.

5. G. Lekkerkerker, *Geometry of numbers*, Wolter-Noordhoff Publishing, Groningen, 1969.

6. D.W. Masser, *A note on Siegel's lemma*, Rocky Mountain J. Math. **26** (1996), 1057–1068.

7. D.W. Masser and G. Wüstholz, *Factorization estimates for abelian varieties*, Inst. Hautes Études Sci. Publ. Math. **81** (1995), 5–24.

8. D. Roy and J.L. Thunder, *A note on Siegel's lemma over number fields*, Monatsh. Math. **120** (1995), 307–318.

9. W.M. Schmidt, *Northcott's theorem on heights* II. *The quadratic case*, Acta Arith. **70** (1995), 343–375.

10. J.H. Silverman, *Lower bounds for height functions*, Duke Math. J. **51** (1984), 395–403.

Département de Mathématiques, Université d'Ottawa, Ottawa, Ontario, Canada K1N 6N5

Department of Mathematics, Northern Illinois University, DeKalb, IL 60115