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

A NOTE ON SIEGEL'S LEMMA

D.W. MASSER

Consider a system of M homogeneous linear equations in N variables with coefficients in a number field K. If $N > M$ the system always has nontrivial solutions over K. For the purpose of finding "small" solutions there are many results in the literature with the general name of "Siegel's lemma." One of the most precise of these is due to Bombieri and Vaaler [**2**].

The results of Bombieri and Vaaler, stated below, involve the height of the system, the heights of the solutions, and the discriminant of the field K. It is well-known that the heights enter into the estimates in an optimal way. Recently it was proved by Roy and Thunder [**9**] that, in general, the discriminant must also be present. Their work covers all possibilities for M and N except $N = M + 1$. The purpose of the present note is to settle the remaining case $N = M+1$. It will turn out that in this case the discriminant enters into the estimates of Bombieri and Vaaler also in an essentially optimal way.

Let us now state precisely the results of [**2**] and [**9**] that are relevant to our discussion. It is slightly more convenient to assume that the linear equations are linearly independent, and at the same time to replace the system by its solution space. Thus, let V be a subspace of K^N (as a vector space over K) with dimension $N - M$ strictly between 0 and N. The height of V was first defined by Wolfgang Schmidt [**11**], in terms of Grassmann coordinates. We shall use the absolute ("nonlogarithmic") projective height $H'(V)$ with L^2 norms at the infinite valuations. This coincides with the definition $H(A)$ of [**2**, p. 15] for any matrix A over K with M rows and N columns defining a system S whose solution space is V. It also coincides with the definition $H'(S)$ of [**9**].

For a vector x in K^N we define similarly $H(x)$ as the absolute (nonlogarithmic) projective height using instead the L^1 norms at the infinite valuations. This coincides with the definition $h(x)$ of [**2**, p. 15]. We also define $H_1(x)$ as the absolute (nonlogarithmic) affine height; if

Received by the editors on December 28, 1994.

$x = (\xi_1, \dots, \xi_N)$ this is just $H(x_1)$ for $x_1 = (1, \xi_1, \dots, \xi_N)$ in K^{N+1}.

Finally, let $2s$ be the number of nonreal embeddings of the field K, and let Δ_K be the square root of the absolute value of the discriminant of the ring of integers $\mathcal{O}_K$ of K (note, however, that this is the $\sqrt{|\Delta_K|}$ of [**2**, p. 23]).

We can now combine Theorem 9 and Corollary 11 of [**2**] as follows.

Theorem (Bombieri-Vaaler). *Let M and N be positive integers with $N \geq M+1$, and let V be a subspace of K^N with dimension $N-M$. Then*

(i) *there is a basis $\mathcal{B}$ in K^N of V with*

$$\prod_{x \in \mathcal{B}} H(x) \leq (2^s \Delta_K / \pi^s)^{(N-M)/d} H'(V),$$

(ii) *there is a basis $\mathcal{B}$ in $\mathcal{O}_K^N$ of V with*

$$\prod_{x \in \mathcal{B}} H_1(x) \leq \Delta_K^{(N-M)/d} H'(V).$$

Next, we can combine Theorem 2 and Proposition 5 of [**9**] as follows.

Theorem (Roy-Thunder). *Let M and N be positive integers with $N \geq M+2$. Then for any integer $d \geq 2$ there are positive real constants C and γ, depending only on M, N and d, with the following property. There are infinitely many number fields K of degree d for which a subspace V of K^N with dimension $N-M$ exists such that*

(i) *for any basis $\mathcal{B}$ in K^N of V we have*

$$\prod_{x \in \mathcal{B}} H(x) \geq C^{-1} \Delta_K^{\gamma} H'(V),$$

(ii) *for any basis $\mathcal{B}$ in $\mathcal{O}_K^N$ of V we have*

$$\prod_{x \in \mathcal{B}} H_1(x) \geq C^{-1} \Delta_K^{\gamma} H'(V).$$

Of course, the discriminants Δ_K are necessarily unbounded in this theorem. Also, in the present formulation (i) implies (ii), because $H_1(x) \geq H(x)$ and H is projective. But actually the results of [**9**] are stronger in several ways. For example, if $N \geq M+3$, then their Theorem 3 gives for even d an exponent $\gamma = (2/d)[(N-M-1)/2]$ in (ii), which is rather close to the exponent $(N-M)/d$ in the upper bound (ii) of Bombieri-Vaaler. Further, if $N \geq M+3$, then their Theorem 2 supplies a subspace V in (i) for every number field K.

Now suppose that $N = M+1$. Then we have $H(x) \leq H'(V)$ for any nonzero x in V (the two heights differ only in the choice of norms), and therefore a lower bound (i) is not possible. The purpose of the present note is to establish the lower bound (ii). More precisely, we shall prove the following result.

Theorem. *For any integers $N \geq 2$, $d \geq 2$, and any real $\gamma < 1/d$ there are infinitely many number fields K of degree d for which a one-dimensional subspace V of K^N exists such that*

$$H_1(x) \geq \Delta_K^{\gamma} H'(V)$$

for any nonzero x in $\mathcal{O}_K^N$ of V.

In particular, the exponent $(N-M)/d = 1/d$ in the upper bound (ii) of Bombieri-Vaaler is best possible for $N = M+1$.

The proof of our theorem breaks into two disjoint parts. They involve the class index i_K of the ring of integers of K. This was introduced in [**8**] in the general context of orders in division algebras (and later generalized further to semisimple algebras), but in the present (commutative) situation the definition can be formulated as follows. The class index is the smallest positive integer I for which every ideal class of K contains an integral ideal $\mathcal{A}$ with norm $N(\mathcal{A}) \leq I$; equivalently, the smallest positive integer I for which every ideal $\mathcal{Z}$ contains a nonzero element ζ with $|\mathrm{norm}\,\zeta|/N(\mathcal{Z}) \leq I$.

Our theorem is now an immediate consequence of the following two propositions.

Proposition 1. *For any positive integer $N \geq 2$ and any number field K there is a one-dimensional subspace V of K^N such that for any* *nonzero x in $\mathcal{O}_K^N$ of V we have*

$$H_1(x) \geq 2^{-1/2} i_K^{1/d} H'(V),$$

where d is the degree of K.

Proposition 2. *For any integer $d \geq 2$ and any real $\varepsilon > 0$ there are infinitely many number fields K of degree d such that*

$$i_K \geq \Delta_K^{1-\varepsilon}.$$

It would be interesting to know if Proposition 1 is best possible in the sense that every one-dimensional subspace V of K^N contains nonzero x in $\mathcal{O}_K^N$ with $H_1(x) \leq C i_K^{1/d} H'(V)$ for some C depending only on N and d. If so, it would mean that Siegel's lemma in this case is essentially linked to the class index.

Certainly Proposition 2 is best possible in the sense that $i_K \leq \Delta_K$ for every number field K (see the class index lemma of [**8**]). But for even degree d, the ε can be removed; see [**10**].

Now for the proofs. I am grateful to Jeff Thunder for the following proof of Proposition 1, which greatly simplified my original argument (involving the prime ideal theorem, Minkowski's first theorem, etc.).

We start with the case $N = 2$. Consider the "worst" ideal class of K consisting of ideals $\mathcal{Z}$ for which

$$|\mathrm{norm}\,\zeta| \geq i_K N(\mathcal{Z}) \tag{1}$$

for every nonzero ζ in $\mathcal{Z}$. Pick such a $\mathcal{Z}$; like any ideal, the inverse $\mathcal{Z}^{-1}$ can be generated over $\mathcal{O}_K$ by two elements, say α and β. We choose for V the subspace of K^2 generated over K by (α, β). Then $2^{-1/2} H'(V) \leq H(\alpha, \beta)$ (again, only the choice of norms is different). The righthand side breaks into the product of factors $H_{\mathrm{fin}}(\alpha, \beta)$ and $H_{\mathrm{inf}}(\alpha, \beta)$ corresponding to the finite and infinite valuations respectively. The former is well known to be $(N(\mathcal{Z}^{-1}))^{-1/d}$ (see, for example, [**5**, p. 54]). So we get

$$2^{-1/2} H'(V) \leq (N(\mathcal{Z}))^{1/d} H_{\mathrm{inf}}(\alpha, \beta). \tag{2}$$

On the other hand, let x be any nonzero point in $\mathcal{O}_K^2$ of V. Then $x = \zeta(\alpha,\beta)$ for some nonzero ζ in $\mathcal{Z}$, and the height $H_1(x)$ breaks into factors $H_{\mathrm{fin}}(1,\zeta\alpha,\zeta\beta)$ and $H_{\mathrm{inf}}(1,\zeta\alpha,\zeta\beta)$. The former is at least 1, and the latter is at least $H_{\mathrm{inf}}(\zeta\alpha,\zeta\beta)$, which is $|\mathrm{norm}\,\zeta|^{1/d}H_{\mathrm{inf}}(\alpha,\beta)$. Using (1), we get

$$H_1(x) \geq i_K^{1/d}(N(\mathcal{Z}))^{1/d}H_{\mathrm{inf}}(\alpha,\beta),$$

and now Proposition 1 follows from this together with (2), at least if $N = 2$. We make the same proof work for $N > 2$ simply by adjoining $N-2$ zeros to the generator of V.

For the proof of Proposition 2 we shall also need the standard class number h_K of K. The general inequality at the end of Section 2 of [**8**] implies the relation $h_K \leq i_K^d$ where d is the degree of K. But in our commutative situation this can easily be improved as follows.

Lemma 1. *For a number field K, we have*

$$h_K \leq i_K(1+\log i_K)^{d-1}$$

where d is the degree of K.

Proof (compare the proof of Lemma 6.1 of [**7**]). Let $f_K(n)$ be the number of integral ideals of K with norm n. We will show that

$$f_K(n) \leq \tau_{d-1}(n) \tag{3}$$

the number of ways of writing n as an ordered product of d positive integers. First suppose that n is a power p^k of a prime p. If $N(\mathcal{A}) = n$, the integral ideal $\mathcal{A}$ must be a product $\mathcal{P}_1^{k_1}\cdots\mathcal{P}_g^{k_g}$ of all the prime ideals $\mathcal{P}_1,\ldots,\mathcal{P}_g$ of K dividing p, and we obtain the equation

$$n_1\cdots n_g = p^k \tag{4}$$

for $n_1 = p^{k_1f_1},\ldots,n_g = p^{k_gf_g}$ and the residue class degrees $f_1,\ldots,f_g$. So the number $f_K(p^k)$ of $(k_1,\ldots,k_g)$ does not exceed the number $\tau_{g-1}(p^k)$ of solutions of (4), and because $g \leq d$ this leads to (3) for prime powers. Since both sides of (3) are multiplicative functions of n, we deduce (3) in general.

Now, by definition, every ideal class contains an integral ideal of norm at most i_K. It follows from (3) that

$$h_K \leq F_{d-1}(i_K), \tag{5}$$

where $F_{d-1}(x)$ is defined for real $x > 0$ by

$$F_{d-1}(x) = \sum_{n \leq x} \tau_{d-1}(n).$$

But we easily verify the identities

$$F_d(x) = \sum_{m \leq x} F_{d-1}(x/m) \quad d \geq 1, \qquad F_0(x) = [x].$$

These lead by a straightforward induction to the inequality

$$F_{d-1}(x) \leq x \biggl(\sum_{m \leq x} 1/m \biggr)^{d-1}.$$

Now Lemma 1 follows from this and (5) together with the obvious estimate for the partial sum of the harmonic series. $\square$

We will also need the following upper bound for the regulator R_K of K. Let $t = d - s - 1$ be the rank of the unit group U_K of K, where $2s$ is the number of nonreal embeddings of K.

Lemma 2. *Suppose that $t \geq 1$. Then*

$$R_K \leq (2d)^t \prod_{\eta \in \mathcal{U}} \log H_1(\eta)$$

for any set $\mathcal{U}$ of t multiplicatively independent units of K.

Proof. Select t suitably independent infinite valuations of K and define the usual logarithmic map $\mathcal{L}$ from the nonzero elements of K into $\mathbf{R}^t$. Now the sum of the absolute values of the logarithms of all the infinite valuations (counted with multiplicity) of a unit η is

$2d \log H_1(\eta)$. So this is an upper bound for the L^1 norm of $\mathcal{L}(\eta)$. It follows easily that the matrix formed with the $\mathcal{L}(\eta)$ for η in $\mathcal{U}$ has determinant of absolute value

$$R \leq (2d)^t \prod_{\eta \in \mathcal{U}} \log H_1(\eta).$$

On the other hand, $R = IR_K \geq R_K$, where I is the index of the subgroup of U_K generated by the elements of $\mathcal{U}$ modulo torsion. This completes the proof of Lemma 2. □

Next, fix an integer $d \geq 2$, and write $e = d - 2$. To construct our special fields K we will use the polynomials $P(X,Y) = P_d(X,Y)$ defined by

$$P(X,Y) = (Y-1)\dots(Y-e)(Y^2+X) - 1, \qquad d \geq 3,$$
$$P(X,Y) = Y^2 + X - 1, \qquad d = 2.$$

Let $D(X) = D_d(X)$ denote the discriminant ("non-square root") of $P(X,Y)$ with respect to the variable Y, so that $D(X)$ is in $\mathbf{Z}[X]$.

Lemma 3. *The polynomial $D(X)$ has a nonconstant factor in $\mathbf{Z}[X]$, irreducible over $\mathbf{Q}$, with odd multiplicity.*

Proof. It clearly suffices to verify that $D(X)$ has odd degree. If $d = 2$, then $D(X) = -4(X-1)$, so we can assume that $d \geq 3$. We calculate the degree by examining the behavior of $D(x)$ as $X = x$ tends to infinity. Now

$$D(x) = \prod_{1 \leq i < j \leq d} (y_i - y_j)^2,$$

where $Y = y_1, \dots, y_d$ are the zeros of $P(x,Y)$. It is easy to see that as $x \to \infty$ these have the asymptotic values $1, \dots, e, i\sqrt{x}, -i\sqrt{x}$. A straightforward calculation shows that $D(x)$ is asymptotically cx^{2e+1} for a certain nonzero constant c. It follows that the polynomial $D(X)$ has odd degree $2e+1$, and this completes the proof. □

Lemma 4. *There is a constant c, depending only on d, with the following property. Let S_4 be the set of integers n such that some zero*

$Y = y$ of $P(n,Y)$ does not generate a number field of degree d. Then, for any $x \geq 1$ the set S_4 contains at most $cx^{1/2}$ positive integers $n \leq x$.

Proof. Since $P(X,Y)$ has degree 1 in X it is easy to see that it is irreducible over $\mathbf{C}$; and therefore also over $\mathbf{Q}(X)$ thanks to Gauss's lemma. This is the situation of the Hilbert irreducibility theorem. So the set of rational t such that $P(t,Y)$ is reducible over $\mathbf{Q}$ is contained in a thin subset of $\mathbf{Q}$ in the sense of [**12**, pp. 121–123] (or, equivalently, its complement contains a Hilbert set in the sense of [**5**, p. 225]). Now Lemma 4 follows from standard cardinality estimates for thin sets; see, for example, [**12**, p. 134] (or the much easier Corollary 2.3 of [**5**, p. 231] would also suffice for the purposes of the present note). □

Lemma 5. *Suppose that $d \geq 3$. Then there is a constant c, depending only on d, with the following property. Let S_5 be the set of integers n for which there is a zero $Y = y$ of $P(n,Y)$ such that the numbers $y-1,\dots,y-e$ are multiplicatively dependent. Then for any $x \geq 2$ the set S_5 contains at most $c(\log x)^{e^2}$ positive integers $n \leq x$.*

Proof. We write $c_1, c_2, \dots$ for constants depending only on d. Consider first the affine variety V defined by $P(X,Y) = 0$. We already noted that this polynomial is irreducible over $\mathbf{C}$, and therefore V is an irreducible curve. We claim that the functions $Y-1,\dots,Y-e$ on V are multiplicatively independent. For, suppose there is a relation $(Y-1)^{a_1}\cdots(Y-e)^{a_e} = 1$ with rational integers $a_1,\dots,a_e$. It is easy to see that $Y-1$ is not identically zero but has a zero at one of the points of V above $X = \infty$ (compare the proof of Lemma 3). It follows that the other factors in the relation are finite and nonzero at this point. But this forces $a_1 = 0$. Similarly for the other exponents; thus, our relation must be trivial and the functions are indeed multiplicatively independent.

We now use the Theorem of [**6**, p. 422] with $k = \mathbf{Q}$, V as above, and Γ generated by $Y-1,\dots,Y-e$. For $x \geq 2$ let $S_5(x)$ be the set of positive integers $n \leq x$ in S_5; that is, the set of all positive integers $n \leq x$ for which there is a zero $Y = y_n$ of $P(n,Y)$ such that $y_n - 1,\dots,y_n - e$ are multiplicatively dependent. Standard arguments (for example, the

Heights lemma of [**6**, p. 419]) yield the height estimate

$$H_1(n, y_n) \leq (\max\{2, n\})^{c_1} \leq e^h \tag{6}$$

for $h = c_1 \log x$. Applying the theorem with d and h as above, we find a polynomial $Q(X, Y)$, of degree at most $c_2 h^\kappa$ for $\kappa = \max\{0, e^2 - 1\}$, which vanishes at the points (n, y_n) for all n in $S_5(x)$ but which does not vanish identically on V. Since V is irreducible, the resultant $R(X)$ of $P(X, Y)$ and $Q(X, Y)$ with respect to Y is nonzero. It vanishes on $S_5(x)$, and therefore the cardinality of $S_5(x)$ does not exceed the degree of $R(X)$, which is at most $c_3 h^\kappa$. This completes the proof of Lemma 5. □

Lemma 6. *Suppose that $E = E(X)$ in $\mathbf{Z}[X]$ is nonconstant and irreducible over $\mathbf{Q}$. Then there is a positive constant C, depending only on E, with the following property. Let S_6 be the set of integers n such that $E(n)$ has some prime factor $p \geq n^{1/4}$ with odd multiplicity. Then, for any $x \geq C$, the set S_6 contains at least $C^{-1}x/\log x$ positive integers $n \leq x$.*

Proof. Let m be the largest positive integer such that $E(n)$ is divisible by m for all integers n. Let S_6' be the set of integers $n > m^4$ for which all prime factors p of $F(n) = E(n)/m$ satisfy $p \geq n^{1/4}$. We claim that there exists $C' > 0$, depending only on E, such that for any $x \geq C'$ the set S_6' contains at least $C'^{-1}x/\log x$ positive integers $n \leq x$.

If the polynomial $F(X) = E(X)/m$ has integer coefficients, this is an immediate deduction from Theorem 9.7 of [**3**, p. 259], together with Remark 3.8, since the condition (5.1) is satisfied. In general, one should apply Theorem 9.3 directly to the numbers $F(n)$. In that case, the verification of the conditions $(\Omega_1), (\Omega_2^*(1))$, (Ω_3), $(R(1, \alpha))$ (with $\alpha = 1$) is relatively straightforward and may be left to the reader; the main observation needed is that the congruence class of $F(n)$ modulo an integer d is now determined by the congruence class of n modulo dm_d, where m_d is the largest factor of m composed only of primes dividing d. This establishes the above claim in general.

To guarantee odd multiplicity we can use the Hilbert irreducibility theorem as in the proof of Lemma 4. It is easy to see that the polynomials $\pm F(X) - Y^2$ are irreducible over $\mathbf{Q}(X)$. Therefore, the set of rational t such that at least one of $\pm F(t)$ is a square is contained in a thin subset of $\mathbf{Q}$. So, for some C'' depending only on E, there are at most $C''x^{1/2}$ positive integers $n \leq x$ such that at least one of $\pm F(n)$ is a square. Removing these from S_6', we are left with numbers $F(n)$ having at least one prime factor p with odd multiplicity; and $p \geq n^{1/4} > m$ by definition. In particular, p does not divide m, and since $E(n) = mF(n)$ this completes the proof of Lemma 6. □

We can now start on the proof of Proposition 2. Fix an integer $d \geq 2$. By Lemma 3 we can write

$$D(X) = (E(X))^u F(X) \tag{7}$$

where u is odd, $E(X)$ in $\mathbf{Z}[X]$ is nonconstant and irreducible over $\mathbf{Q}$, and $F(X)$ in $\mathbf{Z}[X]$ is prime to $E(X)$. We will apply Lemma 6 to this $E(X)$. Let R be the resultant of $E(X)$ and $F(X)$. Thus, R is a nonzero integer.

By comparing cardinality estimates in Lemmas 4, 5 and 6, we see that there are infinitely many positive integers n in the set S_6 but not in S_4 or (if $d \geq 3$) S_5. For such an n, let $Y = y_n$ be any zero of $P(n, Y)$; then since n is not in S_4 the number field $K = K_n = \mathbf{Q}(y_n)$ has degree d. Let Δ_K be its discriminant. We start by proving that

$$\Delta_K \geq n^{1/8} \tag{8}$$

if n is sufficiently large.

Recall that $D(X)$ is the discriminant (nonsquare root) of the polynomial $P(X, Y)$ with respect to Y. It follows that the rational integers Δ_K^2 and $D(n)$ are equal modulo multiplicative squares. From (7) we get $D(n) = (E(n))^u F(n)$. Since n is in S_6, the number $E(n)$ has a prime factor $p \geq n^{1/4}$ with odd multiplicity v, say. If n is sufficiently large, this p cannot divide $F(n)$, otherwise it would divide the resultant R. It follows that $D(n)$ has the factor p with odd multiplicity uv. Therefore, p also divides Δ_K^2, and (8) is an immediate consequence.

In particular, $\Delta_K \to \infty$ as $n \to \infty$, and so there are infinitely many different number fields $K = K_n$ arising from our construction. This allows us to apply the Brauer-Siegel theorem in the form

$$\log(h_K R_K)/\log \Delta_K \longrightarrow 1 \tag{9}$$

as $n \to \infty$ (see, for example, [**4**, p. 328]).

The inequality (8) implies that the discriminant Δ_K is "polynomially large" in n. We next prove that the regulator R_K is by contrast "logarithmically small" in n. For the fields K it is clear that $s = 1$ if n is sufficiently large. In particular, $R_K = 1$ for $d = 2$. If $d \geq 3$, then $y_n - 1, \dots, y_n - e$ are units of K. Since n is not in S_5, these are multiplicatively independent. Recalling from (6) that $H_1(y_n) \leq n^c$ for $n \geq 2$, we deduce from Lemma 2 the upper bound $R_K \leq c(\log n)^e$, where c depends only on d. Now (9) together with (8) implies $\log h_K / \log \Delta_K \to 1$. Finally, this combined with Lemma 1 leads to the assertion of Proposition 2. Our theorem is thereby established. □

Added in proof. I am grateful to S. Louboutin for pointing out that my Proposition 2 is an immediate consequence of my Lemma 1 together with the results of [**1**]. In fact the fields k_n that I construct are essentially special cases of the fields constructed in [**1**] (p. 59) with $r_1 = d - 2, r_2 = 1$ and $N = n$. On the other hand it may be interesting that my discriminant bounds (8) are sharper than the corresponding bounds in Lemma 5 of [**1**].

REFERENCES

1. N.C. Ankeny, R. Brauer, and S. Chowla, *A note on the class-numbers of algebraic number fields*, Amer. J. Math. **78** (1956), 51–61.

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

3. H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, New York, San Francisco, 1974.

4. S. Lang, *Algebraic number theory*, Addison-Wesley, Reading, 1970.

5. ———, *Fundamentals of diophantine geometry*, Springer, New York, 1983.

6. D.W. Masser, *Specializations of finitely generated subgroups of abelian varieties*, Trans. Amer. Math. Soc. **311** (1989), 413–424.

7. D.W. Masser and G. Wüstholz, *Isogeny estimates for abelian varieties, and finiteness theorems*, Ann. Math. **137** (1993), 459–472.

8. ———, *Factorization estimates for abelian varieties*, Pub. I.H.E.S. **81** (1995), 5–24.

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

10. ———, *Bases of number fields with small height*, manuscript, 1994.

11. W.M. Schmidt, *On heights of algebraic subspaces and diophantine approximations*, Ann. Math. **85** (1967), 430–472.

12. J.-P. Serre, *Lectures on the Mordell-Weil theorem*, Aspects Math. **E15**, Vieweg, 1990.

Mathematisches Institut, Universität Basel, Rheinsprung 21, 4051 Basel, Switzerland
E-mail address: Masser@math.unibas.ch