BULL. AUSTRAL. MATH. SOC.
VOL. 25 (1982), 125-132.

10E05, 10F10, 10F35

BOUNDS FOR SOLUTIONS OF SYSTEMS OF LINEAR EQUATIONS

JEFFREY D. VAALER AND A.J. VAN DER POORTEN

We apply recent results of Vaaler that simultaneously bound linear forms so as to obtain a 'Siegel lemma', sharper than those that have appeared in the literature, and in a shape convenient for application in transcendence theory.

1.

In [3] Vaaler proved the following result. Let

$$\Lambda_i(x) = \sum_{j=1}^{n} b_{ij}x_j \ , \quad i = 1, \ldots, M \ ,$$

be M linear forms in n variables $x = (x_1, \ldots, x_n)$ with coefficients $b_{ij} \in \mathbb{C}$. We suppose that each form Λ_i which does not have real coefficients is accompanied by its complex conjugate form $\Lambda_{i'}$, thus $\Lambda_{i'}(x) = \sum \overline{b}_{i'j}x_j$ for some $i' \in \{1, \ldots, M\}$.

Further let B be the $M \times n$ matrix of coefficients $B = (b_{ij})$ and let $\Delta = (d_i\delta_{ij})$ be a $M \times M$ diagonal matrix with positive entries d_i subject to $d_i = d_{i'}$ when Λ_i and $\Lambda_{i'}$ are complex conjugate.

Suppose that for a positive integer k we have

$$|\det B^*\Delta^2 B| \leq k^{-2} \ ,$$

Received 20 August 1981.

where as usual B^* is the complex conjugate transpose of B. Then there are at least k distinct pairs of non-zero points $\pm x \in \mathbb{Z}^n$ such that $|\Lambda_i(\pm x)| \leq d_i^{-1}$ for Λ_i real, and $|\Lambda_i(\pm x)| \leq (2/\pi)^{\frac{1}{2}} d_i^{-1}$ if not. Next, in [4], we see that this result may be applied to bounding integer solutions of homogeneous linear systems by taking

$$\Lambda_i(x) = x_i , \quad i = 1, \ldots, n ,$$

$$\Lambda_{n+i}(x) = L_i(x) = \sum_{j=1}^{n} a_{ij} x_j , \quad i = 1, \ldots, m$$

(so $M = n + m$), and noticing that now $B^*\Delta^2 B$ is of the shape

$$B^*\Delta^2 B = \Delta_1^2 + (\Delta_2 A)^*(\Delta_2 A)$$

for a $n \times n$ diagonal matrix Δ_1, and a $m \times m$ diagonal matrix Δ_2. It remains to find a suitable upper bound for the determinant of matrices of this shape ([4], Lemma 5). In a notation compatible with that above this is

$$\det B^*\Delta^2 B \leq \left(\prod_{j=1}^{n} d_j^2 \right) \prod_{i=1}^{m} \left(1 + d_{n+i}^2 \sum_{j=1}^{n} d_j^{-2} |a_{ij}|^2 \right) .$$

2.

We shall apply these results to the case of linear forms with coefficients in an algebraic number field.

Let

$$L_i(x) = \sum_{j=1}^{n} a_{ij} x_j , \quad i = 1, \ldots, m ,$$

be m non-trivial linear forms in n variables $x = (x_1, \ldots, x_n)$ with coefficients a_{ij} in an algebraic number field $\mathbb{K}$ of degree D over $\mathbb{Q}$. To set notation it might be convenient to suppose that of the D distinct embeddings σ of $\mathbb{K}$ into $\mathbb{C}$, D_1 are real and $2D_2$ are complex (so $D = D_1 + 2D_2$).

https://doi.org/10.1017/S0004972700005104 Published online by Cambridge University Press

We define a demonimator $\operatorname{den} L_i$ for each form by the product

$$\operatorname{den} L_i = \prod_{\nu} \max(|a_{i1}|_\nu, \ldots, |a_{in}|_\nu)$$

taken over the normalised non-archimedean valuations ν of $\mathbb{K}$.

The following preliminary form of our theorem is an immediate consequence of the result cited in §1 above, and is a refinement of similar lemmata appearing in the literature.

THEOREM 1'. *Let* $n > mD$, *and write*

$$T_{i\sigma} = \begin{cases} \left(1+(\operatorname{den} L_i)^{2/D} \sum_{j=1}^{n} |\sigma(a_{ij})|^2\right)^{\frac{1}{2}} \\ \left(1+(\operatorname{den} L_i)^{2/D}(2/\pi) \sum_{j=1}^{n} |\sigma(a_{ij})|^2\right)^{\frac{1}{2}} \end{cases}$$

according as $\sigma(\mathbb{K})$ *is real or otherwise. Then there is a lattice point* $u = (u_1, \ldots, u_n) \neq 0$ *in* $\mathbb{Z}^n$ *such that* $L_i(u) = 0$ *for* $i = 1, \ldots, m$ *and*

$$|u_j| \leq \left(\prod_{i=1}^{m} \prod_{\sigma} T_{i\sigma}\right)^{1/(n-mD)}, \quad j = 1, \ldots, n .$$

In practice this result is neither in a practical nor in a natural form. The following result, though a little weaker, is more congenial.

THEOREM 1. *Let*

$$L_i(x) = \sum_{j=1}^{n} a_{ij}x_j \quad (1 \leq i \leq m) ,$$

be non-trivial linear forms with coefficients a_{ij} *is an algebraic number field* $\mathbb{K}$ *of degree* D *over* $\mathbb{Q}$. *Define a* size $\|L_i\|$ *for each form* L_i *by*

$$\|L_i\|^D = \prod_{\nu} \max(|a_{i1}|_\nu, \ldots, |a_{in}|_\nu)$$

where the product is over all *the normalised valuations (both archimedean*

https://doi.org/10.1017/S0004972700005104 Published online by Cambridge University Press

and non-archimedean) of $\mathbb{K}$ *, and write*

$$A = \left(\|L_1\|\|L_2\| \ldots \|L_m\|\right)^{1/m} .$$

Suppose $n > mD$ *. Then there is a lattice point* $u \neq 0$ *in* $\mathbb{Z}^n$ *such that* $L_i(u) = 0$ *for* $i = 1, 2, \ldots, m$ *and*

$$|u_j| \leq (\sqrt{n+1}\, A)^{mD/(n-mD)} \quad (1 \leq j \leq n) .$$

Of course we could do a little better if $D_2 > 0$. Amusingly, one did a little worse in exactly this case with arguments directly depending on the box principle (for example [5], pp. 1.10-14).

To see that Theorem 1 follows from Theorem 1' it suffices to notice that, by the product formula for valuations, certainly each $\|L_i\| \geq 1$. It seems, at first, that in defining $\|L_i\|$ one should take $\max\left(1, |\sigma a_{i1}|, \ldots, |\sigma a_{in}|\right)$ for the archimedean valuations but, as pointed out to the authors by David Masser, this inhomegeneity is not necessary in view of the homogeneous definition of the den L_i (made possible by our supposing that no one of the forms vanishes identically). Indeed we may replace each of the a_{ij} by qa_{ij} for any non-zero rational q , and then without loss of generality the 1 in the maximum becomes redundant whilst the $T_{i\sigma}$ of Theorem 1' are unchanged. Theorem 1' itself is simply a matter of applying the results of Section 1 to the mD forms $\sigma\left(L_i\right)$ and noting that

$$\left(\text{den } L_i\right) \prod_{\sigma} \sigma\left(L_i(x)\right)$$

is a rational integer for x in $\mathbb{Z}^n$. Thus $L_i(x) = 0$ if this quantity has absolute value less than 1 .

3.

In applications in transcendence theory one will not be concerned with minimising the co-ordinates, as such, of a lattice point but rather with constructing non-zero expressions

https://doi.org/10.1017/S0004972700005104 Published online by Cambridge University Press

$$L = \sum_{j=1}^{n} c_j u_j$$

so that they be small relative to the size of the $c_j u_j$. With the uniform bound U for the co-ordinates, such as that given by Theorem 1, one obtains

$$|L| \leq U \sum_{j=1}^{n} |c_j| = nU\left((1/n) \sum_{j=1}^{n} |c_j|\right) .$$

It may be preferable to attempt to select the u_j "in sympathy with their eventual use" by a suitable variant of Theorem 1; for example by viewing the given forms as

$$\sum a_{ij}x_j = \sum \left(c_j^{-1}a_{ij}\right)\left(c_j x_j\right) ,$$

and applying Theorem 1' to find a "lattice point" with "co-ordinates" $c_j x_j$. We are not aware of circumstances in which this improvement has a material effect in applications. In particular, the authors received no comfort in their forays against Lehmer's question (concerning algebraic integers near the unit circle); see Dobrowolski [1].

A real improvement in a relative bound for L may result if Conjecture 6 of Vaaler [3] is established. Theorem 7 of [3] would then become unconditional and the methods of [4] could be employed to give a generalisation of our Theorem 1' in which L is bounded directly in place of $\max|u_j|$.

4.

If α is a non-zero element of the algebraic number field $\mathbb{K}$ we define its *size* $\|\alpha\|$ by

$$\|\alpha\|^D = \prod_{\nu} \max(1, |\alpha|_\nu)$$

where the product is over all the normalised valuations ν of $\mathbb{K}$. One has $\|\alpha\| = \exp h(\alpha)$, where $h(\alpha)$ is the *absolute logarithmic height* of α .

https://doi.org/10.1017/S0004972700005104 Published online by Cambridge University Press

Frequently it is useful to solve the system $L_i(z) = 0 \quad (1 \le i \le m)$ for z in $\mathbb{K}^n$, $z \neq 0$, rather than for z in $\mathbb{Z}^n$ as we have above. In the literature it seems to be suggested that this may introduce difficult to estimate field constants, but the following strategy avoids this possibility and motivates the shape of Theorem 2 below: there is no loss of generality in supposing that the mn a_{ij} generate $\mathbb{K}$ over $\mathbb{Q}$; hence there are D quantities

$$\beta_\mu = a_{11}^{\mu(11)} a_{12}^{\mu(12)} \ \dots \ a_{mn}^{\mu(mn)}$$

which generate $\mathbb{K}$ as a $\mathbb{Q}$-vector space. Here the $\mu(ij)$ are non-negative integers with sum less than D. We presume a sensible choice for the β_μ. If we suppose $\|a_{ij}\| < A_0$, then it follows that for each μ, $\|\beta_\mu\| \le A_0^{D-1}$, and if we now view the m forms $L_i(z) = \sum a_{ij} z_j$ as forms in nD variables $x_{\mu j}$ in $\mathbb{Z}$,

$$L_i(x) = \sum_{j=1}^{n} \sum_{\mu} a_{ij} \beta_\mu x_{\mu j} ,$$

then Theorem 1 yields a non-trivial solution u with

$$|u_{\mu j}| \le \left(\sqrt{nD+1}\, A_0^{D-1} A \right)^{m/D(n-m)} ,$$

already provided that $n > m$. Plainly it is always efficient to seek solutions in $\mathbb{K}^n$ in this sense, but one may have to take care that expressions L that one now constructs

$$L = \sum_{j=1}^{n} \sum_{\mu} \beta_\mu u_{\mu j} c_j ,$$

do not vanish identically. (For example, above, one could not once again suppose that the $x_{\mu j}$ may be chosen in $\mathbb{K}$.)

The following form for Theorem 1 is due to Mignotte and Waldschmidt [2].

THEOREM 2. *Let $\theta_1, \dots, \theta_r$ be non-zero algebraic numbers in a number field of degree D over $\mathbb{Q}$ and let*

https://doi.org/10.1017/S0004972700005104 Published online by Cambridge University Press

$$P_{ij}(X_1, \ldots, X_r) \in \mathbb{Z}[X_1, \ldots, X_r]$$

be polynomials of degree at most $N_{i,k}$ *in* X_k $(1 \le i \le m,\ 1 \le j \le n,\ 1 \le k \le r)$. *Denote by* $L(P_{ij})$ *the length of the polynomial* P_{ij} *(that is, the sum of the absolute value of its coefficients). Write*

$$a_{ij} = P_{ij}(\theta_1, \ldots, \theta_r)$$

and denote by B *the geometric mean of the* m *quantities*

$$B_i = \sum_{j=1}^{n} L(P_{ij}) \prod_{k=1}^{r} \|\theta_k\|^{N_{i,k}} .$$

If $n > mD$ *there exist rational integers* $u_1, \ldots, u_n$ *not all zero such that*

$$\sum_{j=1}^{n} a_{ij}u_j = 0 \quad (1 \le i \le m)$$

and

$$|u_j| \le B^{mD/(n-mD)} \quad (1 \le h \le n) .$$

Of course this is an immediate corollary of Theorem 1. We could (somewhat pointlessly) improve on the B_i by replacing the sum by the square root of 1 plus the sum of the squares of the summands. Our point is that Vaaler's results yield a tidy bound, in contrast to the bound

$$2 + (2B)^{mD/(n-mD)}$$

in effect obtained in [2]. This is at least psychologically useful in applications. Moreover, an anonymous adviser has kindly pointed out to us that there are circumstances in which a precise bound can be applied to good purpose (and this circumstance is relevant to [1]): *Let* θ *be an algebraic number of degree at most* D *and with size* $\|\theta\|$ *satisfying* $\|\theta\|^D < 2$. *Then there is a polynomial* $P(X) \ne 0$ *whose coefficients are* 0 *or* ± 1 *such that* $P(\theta) = 0$ *and* P *has degree at most* N, *if*

$$N/\mathrm{Log}(2\sqrt{N+2}) > D/\mathrm{Log}(2/\|\theta\|^D) .$$

https://doi.org/10.1017/S0004972700005104 Published online by Cambridge University Press

To see this, apply Theorem 1 to the single equation $x_0\theta^N + x_1\theta^{N-1} + \ldots + x_N = 0$, and ask for a solution $x \neq 0$ in $\mathbb{Z}^{N+1}$ with $|x_j| < 2$, for all j.

References

[1] E. Dobrowolski, "On a question of Lehmer and the number of irreducible factors of a polynomial", *Acta Arith.* 34 (1978-1979), 391-401.

[2] Maurice Mignotte and Michel Waldschmidt, "Linear forms in two logarithms and Schneider's method", *Math. Ann.* 231 (1978), 241-267.

[3] Jeffrey D. Vaaler, "A geometric inequality with applications to linear forms", *Pacific J. Math.* 83 (1979), 543-553.

[4] Jeffrey D. Vaaler, "On linear forms and diophantine approximation", *Pacific J. Math.* 90 (1980), 475-482.

[5] Michel Waldschmidt, *Nombres transcendants* (Lecture Notes in Mathematics, 402. Springer-Verlag, Berlin, Heidelberg, New York, 1974).

Department of Mathematics,
University of Texas,
Austin,
Texas 78712,
USA;

School of Mathematics and Physics,
Macquarie University,
North Ryde,
New South Wales 2113,
Australia.