On Siegel's lemma outside of a union of varieties

Lenny Fukshansky
Claremont McKenna College & IHES

Universität Magdeburg
November 9, 2010

Thue and Siegel

Let

$$A\boldsymbol{x} = \mathbf{0} \tag{1}$$

be an $M \times N$ linear system of rank $M < N$ with integer entries. Define the **height** of a vector $\boldsymbol{x} \in \mathbb{Z}^N$ to be

$$|\boldsymbol{x}| = \max_{1 \le i \le N} |x_i|,$$

and similarly let the height of the matrix

$$A = (a_{ij})_{1 \le i \le M, 1 \le j \le N}$$

be

$$|A| = \max\{|a_{ij}| : 1 \le i \le M, 1 \le j \le N\}.$$

Question 1. *What is the smallest height of a non-trivial integral solution to (1)?*

Indeed, it is natural to expect that there must exist a solution vector $\boldsymbol{x}$ with $|\boldsymbol{x}|$ not too large compared with $|A|$.

In 1929 Carl Ludwig Siegel proved that there exists a non-trivial integral solution $\boldsymbol{x}$ to (1) with

$$|\boldsymbol{x}| \leq (1 + N|A|)^{\frac{M}{N-M}}. \tag{2}$$

The proof uses Dirichlet box principle. In fact, a similar result was at least informally observed by Axel Thue as early as 1909. This result is best possible in the sense that the exponent $\frac{M}{N-M}$ in (2) cannot be improved.

Results of this sort are known under the general name of **Siegel's lemma**, and are very important in transcendence. In the recent years Siegel's lemma was studied by many authors in Diophantine approximations for its own sake as well: it can be thought of as the simplest case of an **effective** existence result for rational points on varieties.

Indeed, since there are only finitely many integral vectors $\boldsymbol{x}$ satisfying (2), one can easily test all of them to find a solution to (1).

Bombieri-Vaaler

A bound like (2) however depends on the choice of a specific matrix A in (1), which is a weakness: if (1) is multiplied on the left by a matrix $U \in \mathrm{GL}_M(\mathbb{Z})$, the solution space is unchanged, but $|UA|$ can be quite different from $|A|$.

In 1983 Enrico Bombieri and Jeffrey Vaaler proved that there exists a non-zero vector $\boldsymbol{x} \in \mathbb{Z}^N$ satisfying (1) such that

$$|\boldsymbol{x}| \leq \left(D^{-1} \sqrt{|\det(AA^t)|} \right)^{\frac{1}{N-M}}, \qquad (3)$$

where D is greatest common divisor of the determinants of all $M \times M$ minors of A. Notice that the quantity $D^{-1}\sqrt{|\det(AA^t)|}$, unlike $|A|$, is invariant under left-multiplication of A by elements of $\mathrm{GL}_M(\mathbb{Z})$.

In fact, the full power of Bombieri-Vaaler result gives a full small-height basis for the null-space of A, and extends to much more general situations. For this we need additional notation.

Absolute values

Throughout this talk, K will be either a number field (finite extension of $\mathbb{Q}$), a function field, or algebraic closure of one or the other; in any case, we write $\overline{K}$ for the algebraic closure of K, so it may be that $K = \overline{K}$. In fact, until further notice assume that $K \neq \overline{K}$.

By a function field we will always mean a finite algebraic extension of the field $\mathfrak{K} = \mathfrak{K}_0(t)$ of rational functions in one variable over a field $\mathfrak{K}_0$, where $\mathfrak{K}_0$ can be any *perfect* field.

When K is a number field, clearly $K \subset \overline{K} = \overline{\mathbb{Q}}$; when K is a function field, $K \subset \overline{K} = \overline{\mathfrak{K}}$, the algebraic closure of $\mathfrak{K}$. In the number field case, we write $d = [K : \mathbb{Q}]$ for the global degree of K over $\mathbb{Q}$; in the function field case, the global degree is $d = [K : \mathfrak{K}]$.

There are infinitely many **absolute values** on K: those that satisfy the triangle inequality

$$|a+b| \leq |a| + |b|,$$

but not the ultrametric inequality

$$|a+b| \leq \max\{|a|, |b|\},$$

are called **archimedean**, and those that satisfy the ultrametric inequality are called **non-archimedean**. We can define an equivalence relation on absolute values: $|\ |_1$ and $|\ |_2$ are said to be equivalent if there exists a real number θ such that

$$|a|_1 = |a|_2^\theta$$

for all $a \in K$. Equivalence classes of absolute values are called **places**, and we write $M(K)$ for the set of all places of K. For each place $v \in M(K)$ we pick representatives $|\ |_v$ and we write $v|\infty$ if v is archimedean, and $v \nmid \infty$ otherwise. We also write K_v for the completion of K at v and let d_v be the local degree of K at v, which is $[K_v : \mathbb{Q}_v]$ in the number field case, and $[K_v : \mathfrak{K}_v]$ in the function field case.

In any case, for each place u of the ground field, be it $\mathbb{Q}$ or $\mathfrak{K}$, we have

$$\sum_{v \in M(K), v|u} d_v = d. \tag{4}$$

If K is a number field, then for each place $v \in M(K)$ we define the absolute value $| \ |_v$ to be the unique absolute value on K_v that extends either the usual absolute value on $\mathbb{R}$ or $\mathbb{C}$ if $v|\infty$, or the usual p-adic absolute value on $\mathbb{Q}_p$ if $v|p$, where p is a prime.

If K is a function field, then all absolute values on K are non-archimedean. For each $v \in M(K)$, let $\mathfrak{O}_v$ be the valuation ring of v in K_v and $\mathfrak{M}_v$ the unique maximal ideal in $\mathfrak{O}_v$. We choose the unique corresponding absolute value $| \ |_v$ such that:

(i) if $1/t \in \mathfrak{M}_v$, then $|t|_v = e$,

(ii) if an irreducible polynomial $p(t) \in \mathfrak{M}_v$, then $|p(t)|_v = e^{-\deg(p)}$.

In both cases, for each non-zero $a \in K$ the **Artin-Whaples product formula** reads

$$\prod_{v \in M(K)} |a|_v^{d_v} = 1. \qquad (5)$$

Example: Let $K = \mathbb{Q}(t)$, and let

$$f(t) = \frac{t-1}{t-2}.$$

Let $| \ |_1$ be the absolute value, corresponding to the ideal $(t-1)$ and $| \ |_2$ be the absolute value corresponding to the ideal $(t-2)$. Then

$$|f(t)|_1 = e^{-1}, \ |f(t)|_2 = e^1,$$

and $|f(t)| = e^0$ for every absolute value $| \ |$ different from $| \ |_1$ and $| \ |_2$. Thus:

$$\prod_{v \in M(K)} |f(t)|_v^{d_v} = \frac{1}{e} \times e = 1.$$

Height functions

We can define local norms on each K_v^N by

$$|\boldsymbol{x}|_v = \max_{1 \leq i \leq N} |x_i|_v,$$

and for all archimedean places v also define

$$\|\boldsymbol{x}\|_v = \left(\sum_{i=1}^{N} |x_i|_v^2 \right)^{1/2},$$

for each $\boldsymbol{x} = (x_1, ..., x_N) \in K_v^N$. Then define a **projective height function** on K^N by

$$H(\boldsymbol{x}) = \prod_{v \in M(K)} |\boldsymbol{x}|_v^{d_v/d}$$

for each $\boldsymbol{x} \in K^N$. This product is convergent because only finitely many of the local norms for each vector $\boldsymbol{x} \in K^N$ are different from 1. Moreover, because of the normalizing power $1/d$ in the definition, H is *absolute*, i.e. does not depend on the field of definition. H is called projective because it is well defined on the projective space $\mathbb{P}^{N-1}(K)$, i.e.

$$H(a\boldsymbol{x}) = H(\boldsymbol{x}),\ \forall\ 0 \neq a \in K,\ \boldsymbol{x} \in K^N,$$

which is true by the product formula.

We also define the **inhomogeneous height** on K^N by

$$h(\boldsymbol{x}) = H(1, \boldsymbol{x}),$$

for all $\boldsymbol{x} \in K^N$. It is easy to see that

$$h(\boldsymbol{x}) \geq H(\boldsymbol{x}) \geq 1,$$

for all non-zero $\boldsymbol{x} \in K^N$.

While the advantage of H is its projective nature, h is more sensitive and refined when measuring the "size" and "arithmetic complexity" of a specific vector, not just the corresponding projective point.

A very important property that both of these heights satisfy over number fields is

Northcott's theorem: *If K is a number field, then for every $B \in \mathbb{R}_{>}0$ the sets*

$$\{\boldsymbol{x} \in \mathbb{P}^{N-1}(K) : H(\boldsymbol{x}) \leq B\}$$

and

$$\{\boldsymbol{x} \in K^N : h(\boldsymbol{x}) \leq B\}$$

are finite.

Northcott's theorem is also true for function fields whose field of constants $\mathfrak{K}_0$ is finite.

We can also talk about height of subspaces of K^N. Let $V \subseteq K^N$ be an L-dimensional subspace, and let $\boldsymbol{x}_1, ..., \boldsymbol{x}_L$ be a basis for V. Then

$$\boldsymbol{y} := \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_L \in K^{\binom{N}{L}}$$

under the standard embedding. Define

$$\mathcal{H}(V) := \prod_{v \nmid \infty} |\boldsymbol{y}|_v^{d_v/d} \times \prod_{v \mid \infty} \|\boldsymbol{y}\|_v^{d_v/d}.$$

This definition is legitimate, i.e. does not depend on the choice of the basis. Hence we have defined a height on points of a Grassmanian over K.

Northcott's theorem, when it works, has the following most important consequence.

Suppose we want to find a point satisfying some arithmetic condition, and assume that we can prove the existence of a point of height $\leq B$ satisfying this condition. But there are only finitely many such points. This suggests a search algorithm, and so B is a **search bound**.

Moreover, height measures arithmetic complexity, and so a point of relatively small height is "arithmetically simple", which makes it even more interesting.

We are now ready to apply this machinery.

Generalized Siegel's lemma

Theorem 1. *Let K be a number field, a function field, or the algebraic closure of one or the other. Let $V \subseteq K^N$ be an L-dimensional subspace, $1 \leq L \leq N$. Then there exists a basis $\boldsymbol{v}_1, \dots, \boldsymbol{v}_L$ for V over K such that*

$$\prod_{i=1}^{L} H(\boldsymbol{v}_i) \leq C_K(L)\mathcal{H}(V), \tag{6}$$

where $C_K(L)$ is an explicit field constant. In fact, if K is a number field or $\overline{\mathbb{Q}}$, then the basis $\boldsymbol{v}_1, \dots, \boldsymbol{v}_L$ as above satisfies the stronger inequality

$$\prod_{i=1}^{L} h(\boldsymbol{v}_i) \leq C_K(L)\mathcal{H}(V). \tag{7}$$

If, on the other hand, K is a function field of genus g (i.e. K is the field of rational functions on a smooth projective curve of genus g over a perfect coefficient field $\mathfrak{K}_0$), then there exists a basis $\boldsymbol{u}_1, \dots, \boldsymbol{u}_L$ for V over K such that

$$\prod_{i=1}^{L} h(\boldsymbol{u}_i) \leq e^{gL} C_K(L)\mathcal{H}(V). \tag{8}$$

Inequality (6) of this general version of Siegel's lemma was obtained by Bombieri and Vaaler (1983) if K is a number field, by Jeffrey Thunder (1995) if K is a function field, and by Damien Roy and Jeffrey Thunder (1996) if K is the algebraic closure of one or the other; (7) is a fairly direct corollary of (6). On the other hand, (8) (F., 2010) required more work.

An immediate consequence of Theorem 1 is the existence of a nonzero point $\boldsymbol{v}_1 \in V$ such that

$$H(\boldsymbol{v}_1) \leq \left(C_K(L)\mathcal{H}(V)\right)^{1/L}. \qquad (9)$$

The bounds of (6) - (9) are sharp in the sense that the exponents on $\mathcal{H}(V)$ are smallest possible.

Faltings' version

In 1992 Gerd Faltings proved a refinement of Siegel's lemma, which guaranteed the existence of a small-height point in a vector space outside of a proper subspace, all over $\mathbb{Q}$. Here is our first generalization of Faltings' result.

Theorem 2 (F., 2006). *Let K be a number field of degree d, let $N \geq 2$ be an integer, and let $V \subseteq K^N$ be an L-dimensional subspace, $1 \leq L \leq N$. Let $U_1, \dots, U_M$ be nonzero subspaces of K^N such that $V \nsubseteq \bigcup_{i=1}^M U_i$. Let $J = \max_{1 \leq i \leq M}\{\dim_K(U_i)\}$. Then there exists a point $\boldsymbol{x} \in V \setminus \bigcup_{i=1}^M U_i$ with coordinate in algebraic integers such that*

$$\begin{aligned} H(\boldsymbol{x}) \;\; &\leq \;\; B_K(N,L,J)\mathcal{H}(V)^d \times \\ &\times \;\; \left\{ \left(\sum_{i=1}^{M} \frac{1}{\mathcal{H}(U_i)^d} \right)^{\frac{1}{(L-J)d}} + M^{\frac{1}{(L-J)d+1}} \right\}, \end{aligned}$$

where $B_K(N,L,J)$ is an explicit field constant.

More generally...

A sharper version of the bound of Theorem 2, again depending of $\mathcal{H}(V)$, $\mathcal{H}(U_i)$, and M was recently obtained by Éric Gaudron (2009). On the other hand, here is a more general result of similar nature.

Theorem 3 (F., 2010)**.** *Let K be a number field, function field, or $\overline{\mathbb{Q}}$. Let $N \geq 2$ be an integer, and let V be an L-dimensional subspace of K^N, $1 \leq L \leq N$. Let $\mathcal{Z}_K$ be a union of algebraic varieties defined over K such that $V \not\subseteq \mathcal{Z}_K$, and let M be sum of degrees of these varieties. Then there exists a basis $\boldsymbol{x}_1, \dots, \boldsymbol{x}_L \in V \setminus \mathcal{Z}_K$ for V over K such that for each $1 \leq n \leq L$,*

$$H(\boldsymbol{x}_n) \leq h(\boldsymbol{x}_n) \leq A_K(L, M)\mathcal{H}(V), \qquad (10)$$

where $A_K(L, M)$ is an explicit field constant.

The exponent 1 on $\mathcal{H}(V)$ in the bound of (10) is sharp in general.

Sketch of the proof of Theorem 3

• Reduction to the case of one polynomial

• Combinatorial Nullstellensatz on a subspace

• Siegel's lemma (Theorem 1) with inhomogeneous heights

• Inhomogeneous height inequality:

$$h\left(\sum_{i=1}^{L} \xi_i \boldsymbol{v}_i\right) \leq L^{\delta} h(\boldsymbol{\xi}) \prod_{i=1}^{L} h(\boldsymbol{x}_i), \tag{11}$$

where $\boldsymbol{\xi} \in K^L$, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_L \in K^N$, and

$$\delta = \begin{cases} 1 & \text{if } K \text{ is a number field or } \overline{\mathbb{Q}} \\ 0 & \text{otherwise.} \end{cases}$$

It should be remarked that the inequality (11) no longer holds if the inhomogeneous height h in the upper bound is replaced with the projective height H, which is why we need Siegel's lemma with inhomogeneous heights.

• Assuming we have a bound on $h(\boldsymbol{\xi})$, we can combine (11) with Siegel's lemma to finish the proof.

We want to construct a set $S \subseteq K$ with $|S| > M$ so that $h(\boldsymbol{\xi})$ is small for every $\boldsymbol{\xi} \in S^L$.

If K is a number field with the number of roots of unity $\omega_K > M$, $\overline{\mathbb{Q}}$, or function field with either an infinite field of constants or a finite field of constants $\mathbb{F}_q$ so that $q > M$, then there exists such a set S with $h(\boldsymbol{\xi}) = 1$ for every $\boldsymbol{\xi} \in S^L$.

The main difficulty arises if K is a number field with $\omega_K \leq M$ or if K is a function field over a finite field $\mathbb{F}_q$ with $q \leq M$.

In both cases the construction of S comes from a certain lattice in Euclidean space. In the number field case, this lattice is the image of the ring of algebraic integers O_K under the standard embedding of K into $\mathbb{R}^d$.

In the function field case, this lattice is the image of the ring of rational functions with all zeros and poles on the curve, over which K is defined, under the principal divisor map.

Lattice point counting estimates are then used to construct S.

Algebraic integers of small height

As a corollary of the proof of Theorem 3, we produce a uniform lower bound on the number of algebraic integers of bounded height in a number field K. The subject of counting *algebraic numbers* of bounded height has been started by the famous asymptotic formula of Schanuel. Some explicit upper and lower bounds have also been produced later, for instance by Schmidt. Recently a new sharp upper bound has been given by Loher and Masser. We produce the following estimate for the number of *algebraic integers*.

Corollary 4 (F., 2010)**.** *Let K be a number field of degree d over $\mathbb{Q}$ with discriminant $\mathcal{D}_K$ and r_1 real embeddings. Let O_K be its ring of integers. For all $R \geq (2^{r_1}|D_K|)^{1/2}$,*

$$(2^{r_1}|\mathcal{D}_K|)^{-1/2} R^d < |\{x \in O_K \ : \ h(x) \leq R\}| .$$