Chapter 2

Siegel Lemma

The goal of this chapter is to introduce the Siegel Lemma in different grades. In general, Siegel Lemma concerns finding small non-zero solutions to a linear system. We will explain:

- Basic version in §2.1, with the goal of finding *one* such small solution.
- Bombieri–Vaaler's version in §2.3, with the goal of finding a *basis* of such small solutions. However, we do not distinguish different solutions in this basis. In some way, we are finding linearly independent solutions whose "average" is small.
- Faltings's version in §2.4, with the goal of finding *successively* such small solutions.

2.1 Basic version

We start with the very basic version of Siegel's Lemma.

Lemma 2.1.1. *Let $a_{ij} \in \mathbb{Z}$ with $i = 1, \dots, M$ and $j = 1, \dots, N$. Assume that a_{ij} are not all 0 and $|a_{ij}| \leq B$ for all i and j.*

If $N > M$, then the homogeneous linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N &= 0 \end{aligned}$$

has a non-zero solution $(x_1, \dots, x_N) \in \mathbb{Z}^N$ with

$$\max_j |x_j| \leq \lfloor (NB)^{\frac{M}{N-M}} \rfloor.$$

In practice, it is more convenient to denote by $A = (a_{ij})_{1 \leq i \leq M,\ 1 \leq j \leq N}$ which is a non-zero $M \times N$-matrix with entries in $\mathbb{Z}$. The upshot of this lemma is that the linear system $A\mathbf{x} = \mathbf{0}$ has a *small* non-zero solution provided that $N > M$. Here *small* means that the height of this non-zero solution is bounded in terms of N, M and $h(A)$.[1] It should be understood that *M is the number of equations and $N - M$ is the dimension of the space of solutions.*

[1] Here $h(A)$ is defined to be the height of $[a_{ij}]_{i,j}$ viewed as a point in $\mathbb{P}^{MN-1}(\overline{\mathbb{Q}})$.

Proof. We may and do assume that no row of A is identically 0. Thus $M \geq 1$. For a positive integer k, consider the set

$$T := \{\mathbf{x} \in \mathbb{Z}^N : 0 \leq x_j \leq k,\ j = 1, \dots, N\}.$$

Then $\#T = (k+1)^N$.

Next, for each $i \in \{1, \dots, M\}$, denote by S_i^+ the sum of the positive entries in the i-th row of A, and by S_i^- the sum of the negative entries. Then

$$\text{For } \mathbf{x} \in T \text{ and } \mathbf{y} := A\mathbf{x}, \text{ we have } kS_i^- \leq y_i \leq kS_i^+ \quad \text{for each } i. \tag{2.1.1}$$

Next, set

$$T' := \{\mathbf{y} \in \mathbb{Z}^N : kS_i^- \leq y_i \leq kS_i^+ \text{ for each } i\}.$$

Then for $B_i := \max_j |a_{ij}|$, we have $S_i^+ - S_i^- \leq NB_i$ and we can conclude that $\#T' \leq \prod_{i=1}^M (NkB_i + 1)$.

Now take $k := \lfloor \prod_{i=1}^M (NB_i)^{1/(N-M)} \rfloor$. Then $NkB_i + 1 < NB_i(k+1)$ because $N \geq M > 1$, and hence

$$\prod_{i=1}^M (NkB_i + 1) < \prod_{i=1}^M NB_i(k+1) = (k+1)^M \prod_{i=1}^M NB_i.$$

On the other hand, $\prod_{i=1}^M (NB_i)^{1/(N-M)} \leq k+1$. So

$$\prod_{i=1}^M (NkB_i + 1) < (k+1)^M (k+1)^{N-M} = (k+1)^N = \#T.$$

We have seen that $\#T'$ is bounded above by the left hand side. So $\#T' < \#T$. By the Pigeonhole Principle and (2.1.1), there exist two different points $\mathbf{x}', \mathbf{x}'' \in T$ such that $A\mathbf{x}' = A\mathbf{x}''$.

Now $\mathbf{x} := \mathbf{x}' - \mathbf{x}''$ is a non-zero solution of the linear system in question such that $\max_j |x_j| \leq k = \lfloor \prod_{i=1}^M (NB_i)^{1/(N-M)} \rfloor \leq \lfloor (NB)^{M/(N-M)} \rfloor$. □

This basic version self-improves to a version for number fields.

Lemma 2.1.2. *Let $K \subseteq \mathbb{C}$ be a number field of degree d, and let $|\cdot|$ be the usual absolute value on $\mathbb{C}$. Let $M, N \in \mathbb{Z}$ with $0 < M < N$. Then there exist positive integers C_1 and C_2 such that the following property holds true: For any non-zero $M \times N$-matrix A with entries $a_{mn} \in \mathcal{O}_K$, there exists $\mathbf{x} \in \mathcal{O}_K^N \setminus \{\mathbf{0}\}$ with $A\mathbf{x} = \mathbf{0}$ and*

$$H(\mathbf{x}) \leq C_1 (C_2 NB)^{\frac{M}{N-M}},$$

where $B := \max_{\sigma,m,n} |\sigma(a_{mn})|$ with σ running over all the embeddings $K \hookrightarrow \mathbb{C}$.

The constants C_1 and C_2 depend only K (and hence d), M and N, but they are independent of the choice of the matrix A.[2] By the Fundamental Inequality (Proposition 1.2.10), B can be bounded by $H(A)$ with A viewed as a point $(a_{mn})_{m,n} \in \overline{\mathbb{Q}}^{MN}$.

Proof. Let $\omega_1, \dots, \omega_d$ be a $\mathbb{Z}$-basis of $\mathcal{O}_K$. The entries of A may be written as

$$a_{mn} = \sum_{j=1}^d a_{mn}^{(j)} \omega_j, \quad a_{mn}^{(j)} \in \mathbb{Z}. \tag{2.1.2}$$

[2] In fact by the proof, one can see that C_2 depends only on K.

For each $\mathbf{x} = (x_1, \dots, x_N) \in \mathcal{O}_K^N$, using $x_n = \sum_{k=1}^d x_n^{(k)} \omega_k$ we get

$$(A\mathbf{x})_m = \sum_{n=1}^N \sum_{j,k=1}^d a_{mn}^{(j)} \omega_j \omega_k x_n^{(k)} = \sum_{l=1}^d \sum_{n=1}^N \sum_{j,k=1}^d a_{mn}^{(j)} b_{jk}^{(l)} x_n^{(k)} \omega_l,$$

where $\omega_j \omega_k = \sum_{l=1}^d b_{jk}^{(l)} \omega_l$. Set A' to be the $(Md) \times (Nd)$-matrix

$$A' := \left(\sum_{j=1}^d a_{mn}^{(j)} b_{jk}^{(l)} \right)$$

with rows indexed by (m, l) and columns indexed by (n, k). Write $\mathbf{y} \in \mathbb{Z}^{Nd}$ for the vector $(x_n^{(k)})$.

Apply the basic version of Siegel's Lemma, Lemma 2.1.1, to A'. Then we obtain a non-zero integer solution $\mathbf{y}$ with $A'\mathbf{y} = \mathbf{0}$ such that

$$H(\mathbf{y}) \leq \left(Nd^2 \max_{m,n,j} |a_{mn}^{(j)}| \max_{j,k,l} |b_{jk}^{(l)}| \right)^{\frac{M}{N-M}}.$$

As $x_n = \sum_{k=1}^d x_n^{(k)} \omega_k$ for each n, we then obtain a constant C_1 such that $H(\mathbf{x}) \leq C_1 H(\mathbf{y})$.

Next we wish to bound $\max_j |a_{mn}^{(j)}|$ in terms of $\max_{\sigma,m,n} |\sigma(a_{mn})|$. Let σ run over the $d = [K : \mathbb{Q}]$ different embeddings $K \hookrightarrow \mathbb{C}$. Apply each σ to (2.1.2). It is known from Algebraic Number Theory that the $d \times d$-matrix $(\sigma(\omega_j))_{\sigma,j}$ is invertible.[3] So we obtain a constant C_2' such that

$$\max_j |a_{mn}^{(j)}| \leq C_2' \max_\sigma |\sigma(a_{mn})|.$$

Thus we can conclude by taking $C_2 := C_2' d^2 \max_{j,k,l} |b_{jk}^{(l)}|$. □

Next, we also have the following *relative version* of Siegel's Lemma.

Lemma 2.1.3 (Relative version of Siegel's Lemma, basic version)**.** *Let K be a number field of degree d. Then there exists a positive number C such that the following property holds true For any $M, N \in \mathbb{Z}$ with $0 < dM < N$ and any non-zero $M \times N$-matrix A with entries $a_{mn} \in \mathcal{O}_K$, there exists $\mathbf{x} \in \mathbb{Z}^N \setminus \{0\}$ with $A\mathbf{x} = \mathbf{0}$ and*

$$H(\mathbf{x}) \leq \lfloor (CNB)^{\frac{dM}{N-dM}} \rfloor$$

where $B := \max_{\sigma,m,n} |\sigma(a_{mn})|$ with σ running over all the embeddings $K \hookrightarrow \mathbb{C}$.

Again, by the Fundamental Inequality (Proposition 1.2.10), B can be bounded by $H(A)$ with A viewed as a point $(a_{mn})_{m,n} \in \overline{\mathbb{Q}}^{MN}$. We emphasize that the constant C depends only on the field K.

Proof. Let $\omega_1, \dots, \omega_d$ be a $\mathbb{Z}$-basis of $\mathcal{O}_K$. For the entries of $A = (a_{mn})$, we have

$$a_{mn} = \sum_{j=1}^d a_{mn}^{(j)} \omega_j \tag{2.1.3}$$

[3] $\deg(\sigma(\omega_j))_{\sigma,j}^2 = \mathrm{Disc}(K/\mathbb{Q}) \neq 0$.

for uniquely determined $a_{mn}^{(j)} \in \mathbb{Z}$. Consider the $M \times N$-matrix $A^{(j)} = (a_{mn}^{(j)})$ for each $j \in \{1, \dots, d\}$. Then for $\mathbf{x} \in \mathbb{Q}^N$, the equation $A\mathbf{x} = \mathbf{0}$ is equivalent to the system of equations $A^{(j)}\mathbf{x} = \mathbf{0}$ for all $j = 1, \dots, d$. This new system has dM equations and N unknowns. Write A' for the $dM \times N$-matrix $\begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(d)} \end{pmatrix}$. Since $dM < N$, we can apply Lemma 2.1.1 to find a non-zero solution $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N$ with

$$\max_i |x_i| \leq \lfloor (N \max_{m,n,j} |a_{mn}^{(j)}|)^{\frac{dM}{N-dM}} \rfloor.$$

It remains to compare $\max_{m,n,j} |a_{mn}^{(j)}|$ and $\max_{\sigma,m,n} |\sigma(a_{mn})|$. We use the same argument as for Lemma 2.1.2. Let σ run over the $d = [K : \mathbb{Q}]$ different embeddings $K \hookrightarrow \mathbb{C}$. Apply each σ to (2.1.3). It is known from Algebraic Number Theory that $\deg(\sigma(\omega_j))_{\sigma,j}^2 = \mathrm{Disc}(K/\mathbb{Q}) \neq 0$. So we obtain a constant C such that $\max_j |a_{mn}^{(j)}| \leq C \max_\sigma |\sigma(a_{mn})|$. Hence we are done. □

2.2 Arakelov height of matrices

While the basic versions of Siegel's Lemma are sufficient for many applications, we state and prove a generalized version. Its proof, which is by the Geometry of Numbers and in particular uses the adelic version of Minkowski's second main theorem, is of particular importance.

Theorem 2.2.1. *Let A be an $M \times N$-matrix of rank M with entries in a number field K of degree d. Then the K-vector space of solutions of $A\mathbf{x} = \mathbf{0}$ has a basis $\mathbf{x}_1, \dots, \mathbf{x}_{N-M}$, contained in $\mathcal{O}_K^N$, such that*

$$\prod_{l=1}^{N-M} H(\mathbf{x}_l) \leq |D_{K/\mathbb{Q}}|^{\frac{N-M}{2d}} H_{\mathrm{Ar}}(A),$$

where $D_{K/\mathbb{Q}}$ is the discriminant of K over $\mathbb{Q}$.

There are several things to be explained for this statement. First, $H(\mathbf{x}) = \exp(h(\mathbf{x}))$ is the multiplicative homogeneous height with $\mathbf{x}$ considered as a point in $\mathbb{P}^{N-1}(K)$; thus we may assume $\mathbf{x} \in \mathcal{O}_K^N$ because we can replace any solution by a non-zero scalar multiple and this does not change its height. Second, we need to define the *Arakelov height* $H_{\mathrm{Ar}}(A)$ of the matrix A; this is what we will do in this section.

Moreover, there is also a relative version for this generalized version. See Theorem 2.3.3.

2.2.1 Arakelov height on $\mathbb{P}^N$

Recall the Weil height which we defined before. For a point $\mathbf{x} = [x_0 : \cdots : x_N] \in \mathbb{P}^N(K)$, we have

$$[K : \mathbb{Q}]h(\mathbf{x}) = \sum_{v \in M_K^0} \log \max_j \|x_j\|_v + \sum_{v|\infty} \log \max_j \|x_j\|_v = \sum_{v \in M_K^0} \log \max_j \|x_j\|_v + \sum_{v|\infty} [K_v : \mathbb{R}] \log \max_j |x_j|_v.$$

There are other choices for the height function on $\mathbb{P}^N(\overline{\mathbb{Q}})$. In Arakelov theory, a more natural choice is to replace the L^∞-norm $\max_j |x_j|_v$ at the *archimedean* place by the L^2-norm $\left(\sum_{j=0}^N |x_j|_v^2\right)^{1/2}$. In other words, we define:

Definition 2.2.2. *For* $\mathbf{x} = [x_0 : \cdots : x_N] \in \mathbb{P}^N(\overline{\mathbb{Q}})$ *with each* $x_j \in K$*, define*

$$h_{\mathrm{Ar}}(\mathbf{x}) := \frac{1}{[K:\mathbb{Q}]} \left(\sum_{v \in M_K^0} \log \max_j \|x_j\|_v + \sum_{v|\infty} [K_v : \mathbb{R}] \log \left(\sum_{j=0}^{N} |x_j|_v^2 \right)^{1/2} \right).$$

One can check that $h_{\mathrm{Ar}}(\mathbf{x})$ is independent of the choice of the homogeneous coordinates (by the Product Formula) and of the choice of the number field K.

To ease notation, we introduce the following definition.

Definition 2.2.3. *For* $\mathbf{x} = [x_0 : \ldots : x_N] \in \mathbb{P}^N(K)$ *and* $v \in M_K$*, set*

$$H_v(\mathbf{x}) := \begin{cases} \max_j \|x_j\|_v = \max_j |x_j|_v^{[K_v:\mathbb{Q}_p]} & \text{if } v \text{ is non-archimedean,} \\ \left(\sum_{j=0}^{N} |x_j|_v^2 \right)^{1/2 \cdot [K_v:\mathbb{R}]} & \text{if } v \text{ is archimedean.} \end{cases}$$

With this definition, the following holds true. For $\mathbf{x} = [x_0 : \cdots : x_N] \in \mathbb{P}^N(\overline{\mathbb{Q}})$ with each $x_j \in K$, we have

$$h_{\mathrm{Ar}}(\mathbf{x}) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log H_v(\mathbf{x}). \tag{2.2.1}$$

The following lemma will be proved in the Exercise class.

Lemma 2.2.4. *On* $\mathbb{P}^N(\overline{\mathbb{Q}})$*, the height functions* h *and* h_{Ar} *differ from a bounded function.*

Thus in view of the Height Machine, h_{Ar} is in the class represented by $h_{\mathbb{P}^N, \mathcal{O}(1)}$.

2.2.2 Height of matrices

We start by defining a height function on the Grassmannians. Let W be an M-dimensional subspace of $\overline{\mathbb{Q}}^N$. Then $\wedge^M W$ is a 1-dimensional subspace of $\wedge^M \overline{\mathbb{Q}}^N \simeq \overline{\mathbb{Q}}^{\binom{N}{M}}$. Thus we may view W as a point P_W of the projective space $\mathbb{P}(\wedge^M \overline{\mathbb{Q}}^N)$.

Definition 2.2.5. *The* ***Arakelov height*** *of* W *is defined to be* $h_{\mathrm{Ar}}(W) := h_{\mathrm{Ar}}(P_W)$*. We also define the multiplicative Arakelov height* $H_{\mathrm{Ar}}(W) := \exp(h_{\mathrm{Ar}}(P_W))$*.*

Now we are ready to define the Arakelov height of a matrix A.

Definition 2.2.6. *Let* A *be an* $N \times M$*-matrix with entries in* $\overline{\mathbb{Q}}$*.*

(i) Assume $\mathrm{rk}A = M$*. Then* $h_{\mathrm{Ar}}(A)$ *is defined as* $h_{\mathrm{Ar}}(W)$*, where* W *is the subspace of* $\overline{\mathbb{Q}}^N$ *spanned by the columns of* A*.*[4]

(ii) Assume $\mathrm{rk}A = N$*. Then* $h_{\mathrm{Ar}}(A) := h_{\mathrm{Ar}}(A^t)$ *with* A^t *the transpose of* A*.*

We also define the multiplicative Arakelov height $H_{\mathrm{Ar}}(A) := \exp(h_{\mathrm{Ar}}(P_W))$*.*

In general, A may not have the full rank. We then consider the subspace spanned by the columns or by the rows. This will lead to $h_{\mathrm{Ar}}^{\mathrm{col}}$ and $h_{\mathrm{Ar}}^{\mathrm{row}}$. We omit the definitions here but the idea will show up in the discussion of the generalized Siegel's Lemma in the next section.

We start with the following lemma, which makes the two parts of Definition 2.2.6 more "symmetric".

[4] Notice that A defines a linear map $A\colon \mathbb{R}^M \to \mathbb{R}^N$. The subspace W is precisely the image of this map. The assumption $\mathrm{rk}A = M$ is equivalent to the map A being injective.

Lemma 2.2.7. *Let A be an $N \times M$-matrix with entries in $\overline{\mathbb{Q}}$. Assume* $\mathrm{rk}A = N$. *Then $h_{\mathrm{Ar}}(A)$ equals the Arakelov height of the subspace of $\overline{\mathbb{Q}}^M$ spanned by the rows of A.*

Proof. Consider the transpose A^t of A. It can be easily seen that A^t is an $M \times N$-matrix of rank N, and hence defines an injective linear map $\overline{\mathbb{Q}}^N \to \overline{\mathbb{Q}}^M$, which by abuse of notation we still denote by A^t. Part (i) of Definition 2.2.6 (applied to A^t) says that $h_{\mathrm{Ar}}(A^t)$ equals $h_{\mathrm{Ar}}(W)$ with $W \subseteq \overline{\mathbb{Q}}^M$ the subspace spanned by the columns of A^t. Notice that $W = \mathrm{Im}(A^t)$.

The matrix A defines a linear map $A\colon (\overline{\mathbb{Q}}^M)^* \to (\overline{\mathbb{Q}}^N)^*$ which is the dual of A^t. Consider the subspace $\mathrm{Ker}(A)$ of $(\overline{\mathbb{Q}}^M)^*$. Its annihilator $\mathrm{Ker}(A)^\perp$ in $((\overline{\mathbb{Q}}^M)^*)^* = \overline{\mathbb{Q}}^M$ then equals $\mathrm{Im}(A^t) = W$ by Linear Algebra. It is known that $\mathrm{Ker}(A)^\perp$ is spanned by the rows of A, and so is W. Hence we are done because $h_{\mathrm{Ar}}(A) = h_{\mathrm{Ar}}(A^t) = h_{\mathrm{Ar}}(W)$. □

Proposition 2.2.8. *Let W be an M-dimensional subspace of $\overline{\mathbb{Q}}^N$ and let $W^\perp$ be its annihilator in the dual $(\overline{\mathbb{Q}}^N)^* \simeq \overline{\mathbb{Q}}^N$. Then $h_{\mathrm{Ar}}(W^\perp) = h_{\mathrm{Ar}}(W)$.*

This proposition has the following immediate corollary.

Corollary 2.2.9. *Let A be an $N \times M$-matrix with* $\mathrm{rk}A = N$ *and with entries in $\overline{\mathbb{Q}}$. Then the Arakelov height of the space of solutions of $A\mathbf{x} = \mathbf{0}$ equals $h_{\mathrm{Ar}}(A)$.*

Proof. We have $h_{\mathrm{Ar}}(A) = h_{\mathrm{Ar}}(A^t) = h_{\mathrm{Ar}}(\mathrm{Im}(A^t))$. But $\mathrm{Im}(A^t) = \mathrm{Ker}(A)^\perp$. So $h_{\mathrm{Ar}}(A) = h_{\mathrm{Ar}}(\mathrm{Ker}(A)^\perp)$, which then equals $h_{\mathrm{Ar}}(\mathrm{Ker}(A))$ by Proposition 2.2.8. Hence we are done. □

Proof of Proposition 2.2.8. Write $V = \overline{\mathbb{Q}}^N$. Any element $x \in \wedge^M V$ defines a linear map $\psi(x)\colon \wedge^{N-M} V \to \wedge^N V$, $y \mapsto x \wedge y$, and thus an element $\varphi(x) \in \wedge^N V \otimes \wedge^{N-M}(V^*)$. In other words, we obtained a map

$$\varphi\colon \wedge^M V \to \wedge^N V \otimes \wedge^{N-M}(V^*).$$

Then φ is an isomorphism and (better) each element of the canonical basis of $\wedge^M V$ is mapped to an element of the canonical basis of $\wedge^N V \otimes \wedge^{N-M}(V^*)$ up to a sign.

Notice that $\wedge^N V$ is a line. So it is easy to check that for any non-zero $x \in \wedge^M W$ (which is a line), the image of $\psi(x)$ is $\wedge^N V$ and the kernel of $\psi(x)$ is the subspace of $\wedge^{N-M} V$ generated by the elements of the form $w \wedge z$ with $w \in W$ and $z \in \wedge^{N-M-1} V$. Thus $\varphi(\wedge^M W) = \wedge^N V \otimes \wedge^{N-M}(W^\perp)$. Hence the coordinates of $\wedge^M W$ in $\mathbb{P}(\wedge^M V)$ are, up to a sign, equal to the coordinates of $\wedge^{N-M}(W^\perp)$ in $\mathbb{P}(\wedge^{N-M}(V^*))$. This proves the proposition. □

We finish this section by the following explicit formula for the definition of $h_{\mathrm{Ar}}(A)$. Let A be an $N \times M$-matrix with entries in $\overline{\mathbb{Q}}$.

For simplicity *we only consider the case* $\mathrm{rk}A = M$. Let $I \subseteq \{1, \dots, N\}$ with $|I| = M$. Denote by A_I the $M \times M$-submatrix of A formed with the i-th rows, $i \in I$, of A. Then the point in $\mathbb{P}(\wedge^N \overline{\mathbb{Q}}^M)$ corresponding to $\mathrm{Im}(A)$ is given by the coordinates $\det(A_I)$, where I ranges over all subsets of $\{1, \dots, N\}$ of cardinality M.

Let $K \subseteq \overline{\mathbb{Q}}$ be a number field which contains all entries of A. For each $v \in M_K$, set

$$H_v(A) := \begin{cases} \max_I |\det(A_I)|_v^{[K_v:\mathbb{Q}_p]} = \max_I \|\det(A_I)\|_v & \text{if } v \text{ is non-archimedean,} \\ \left(\sum_I |\det(A_I)|_v^2\right)^{1/2\cdot[K_v:\mathbb{R}]} = |\det(A^*A)|_v^{1/2\cdot[K_v:\mathbb{R}]} = \|\det(A^*A)\|_v^{1/2} & \text{if } v \text{ is archimedean.} \end{cases} \tag{2.2.2}$$

Here $A^* = \overline{A}^t$ is the adjoint of A, and $\sum_I |\det(A_I)|_v^2 = |\det(A^*A)|_v$ at the archimedean places by the *Binet Formula*.

Under this convention, we have

$$h_{\mathrm{Ar}}(A) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log H_v(A). \tag{2.2.3}$$

An immediate corollary of this explicit formula is:

Corollary 2.2.10. *Let G be an invertible $M \times M$-matrix. Then $h_{\mathrm{Ar}}(AG) = h_{\mathrm{Ar}}(A)$.*

Another application of this explicit formula is:

Corollary 2.2.11. *Let B and C be two complementary submatrices of A of type $N \times M_1$ and $M \times M_2$ respectively. Then $h_{\mathrm{Ar}}(A) \leq h_{\mathrm{Ar}}(B) + h_{\mathrm{Ar}}(C)$.*

Proof. We only give a sketch. It suffices to prove $H_v(A) \leq H_v(B)H_v(C)$ for each $v \in M_K$. If v is non-archimedean, it follows from Laplace's expansion. If v is archimedean, it follows from Fischer's inequality

$$\det \begin{pmatrix} B^*B & B^*C \\ C^*B & C^*C \end{pmatrix} \leq \det(B^*B)\det(C^*C).$$

Alternatively, this corollary is an immediate consequence of the important theorem of Schmidt (independently of Struppeck–Vaaler) $h_{\mathrm{Ar}}(V+W) + h_{\mathrm{Ar}}(V \cap W) \leq h_{\mathrm{Ar}}(V) + h_{\mathrm{Ar}}(W)$ for any subspaces V, W of $\overline{\mathbb{Q}}^M$. □

2.3 Generalized Siegel Lemma by Bombieri–Vaaler

The goal of this section is to have a deeper discussion of the generalized Siegel's Lemma by Bombieri–Vaaler (Theorem 2.2.1); in particular we give its proof. We repeat the statement here.

Theorem 2.3.1. *Let A be an $M \times N$-matrix of rank M with entries in a number field K of degree d. Then the K-vector space of solutions of $A\mathbf{x} = \mathbf{0}$ has a basis $\mathbf{x}_1, \ldots, \mathbf{x}_{N-M}$, contained in $\mathcal{O}_K^N$, such that*

$$\prod_{l=1}^{N-M} H(\mathbf{x}_l) \leq |D_{K/\mathbb{Q}}|^{\frac{N-M}{2d}} H_{\mathrm{Ar}}(A),$$

where $D_{K/\mathbb{Q}}$ is the discriminant of K over $\mathbb{Q}$.

As said below Theorem 2.2.1, there is no deep information about the $\mathbf{x}_i$'s being contained in $\mathcal{O}_K^N$.

In practice, we may not always assume that A has maximal rank M. This can be obviated. We hereby state a corollary of Theorem 2.3.1, which bounds the heights of the solutions by the (multiplicative) Weil height instead of the Arakelov height.

Corollary 2.3.2. *Let A be an $M \times N$-matrix of rank R with entries in a number field K of degree d. Then there exists a basis $\mathbf{x}_1, \ldots, \mathbf{x}_{N-R}$ of the kernel $\mathrm{Ker}(A)$, contained in $\mathcal{O}_K^N$, such that*

$$\prod_{l=1}^{N-R} H(\mathbf{x}_l) \leq |D_{K/\mathbb{Q}}|^{\frac{N-R}{2d}} \left(\sqrt{N} H(A)\right)^R.$$

Here $H(A)$ is the multiplicative Weil height of the point $[a_{ij}]_{i,j}$ viewed as a point in $\mathbb{P}^{MN-1}(K)$, with a_{ij} the entries of A.

In particular, there is a non-zero solution $\mathbf{x} \in \mathcal{O}_K^N$ of $A\mathbf{x} = \mathbf{0}$ with

$$H(\mathbf{x}) \leq |D_{K/\mathbb{Q}}|^{\frac{1}{2d}} \left(\sqrt{N} H(A)\right)^{\frac{R}{N-R}}.$$

2.3.1 Proof of Corollary 2.3.2 assuming Theorem 2.3.1

The "In particular" part follows clearly from the main part. So we will focus on proving the main part.

As $\mathrm{rk}A = R$, there is an $R \times N$-submatrix A' of A with $\mathrm{rk}A' = R$. Applying Theorem 2.3.1 to the matrix A', we get a basis $\mathbf{x}_1, \ldots, \mathbf{x}_{N-R}$ of $\mathrm{Ker}(A)$ such that

$$\prod_{l=1}^{N-R} H(\mathbf{x}_l) \leq |D_{K/\mathbb{Q}}|^{\frac{N-R}{2d}} H_{\mathrm{Ar}}(A'). \tag{2.3.1}$$

On the other hand, if we denote by A_m the m-th row of A, then Corollary 2.2.11 implies that

$$H_{\mathrm{Ar}}(A') \leq \prod_m H_{\mathrm{Ar}}(A_m),$$

where m runs over the R rows of A'. Furthermore, the following inequality clearly holds true by definition

$$H_{\mathrm{Ar}}(A_m) \leq \sqrt{N} H(A).$$

Now, the two inequalities above yield $H_{\mathrm{Ar}}(A') \leq (\sqrt{N}H(A))^R$. So we can conclude by (2.3.1). □

2.3.2 Relative Version

As for Lemma 2.1.3 with respect to Lemma 2.1.1, we also have the following relative version of this generalized form of Siegel's Lemma.

Theorem 2.3.3. *Let K be a number field of degree d and F/K be a finite extension with $[F : K] = r$. Let A be an $M \times N$-matrix with entries in F.*

Assume $rM < N$. Then there exists $N - rM$ K-linearly independent vectors $\mathbf{x}_l \in \mathcal{O}_K^N$ such that $A\mathbf{x}_l = 0$ for each $l \in \{1, \ldots, N - rM\}$ and

$$\prod_{l=1}^{N-rM} H(\mathbf{x}_l) \leq |D_{K/\mathbb{Q}}|^{\frac{N-rM}{2d}} \prod_{i=1}^{M} H_{\mathrm{Ar}}(A_i)^r,$$

where A_i is the i-th row of A.

The proof follows the guideline set up in Lemma 2.1.3.

Proof. Let $\omega_1, \ldots, \omega_r$ be a basis of F/K. For the entries of $A = (a_{mn})$, we have

$$a_{mn} = \sum_{j=1}^{r} a_{mn}^{(j)} \omega_j$$

for uniquely determined $a_{mn}^{(j)} \in K$. Let $A^{(j)}$ be the $M \times N$-matrix with entries $a_{mn}^{(j)}$. Then for $\mathbf{x} \in K^N$, the equation $A\mathbf{x} = \mathbf{0}$ is equivalent to te system of equations $A^{(j)}\mathbf{x} = \mathbf{0}$ for all $j = 1, \ldots, r$. Write A' for the $rM \times N$-matrix $\begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(r)} \end{pmatrix}$. Denote by $R := \mathrm{rk}A'$.

It is attempting to apply Theorem 2.3.1 to A'. But we need to do one more step. Let $\sigma_1, \dots, \sigma_r$ be the distinct embeddings of F into $\overline{K}$ over K. Let Ω be the $rM \times rM$-matrix built up by r^2 blocks of $M \times M$-matrices $\Omega_{ij} = \sigma_i(\omega_j) I_M$. By construction of A', we have

$$A'' := \begin{pmatrix} \sigma_1 A \\ \vdots \\ \sigma_r A \end{pmatrix} = \Omega A'.$$

From Algebraic Number Theory, it is known that $D_{F/K} = \det(\sigma_i(\omega_j))^2$. Thus Ω is invertible, and its inverse is again formed by r^2 blocks of multiples of I_M. In particular, $\mathrm{rk} A'' = \mathrm{rk} A' = R$ and $\mathrm{Ker}(A'') = \mathrm{Ker}(A')$.

There exists an $R \times N$-submatrix A''' of A' with $\mathrm{rk} A''' = R$. Applying Theorem 2.3.1 to A''', we get a basis $\mathbf{x}_1, \dots, \mathbf{x}_{N-R}$ of $\mathrm{Ker}(A''') = \mathrm{Ker}(A')$, contained in $\mathcal{O}_K$, such that

$$\prod_{l=1}^{N-R} H(\mathbf{x}_l) \leq |D_{K/\mathbb{Q}}|^{\frac{N-R}{2d}} H_{\mathrm{Ar}}(A'''),$$

If we denote by A_m the m-th row of A'', then Corollary 2.2.11 implies that

$$H_{\mathrm{Ar}}(A'') \leq \prod_m H_{\mathrm{Ar}}(A''_m),$$

where m runs over the R rows of A''. Thus if we rearrange our basis $\mathbf{x}_l$ by increasing height, we have

$$\prod_{l=1}^{N-rM} H(\mathbf{x}_l) \leq \left(\prod_{l=1}^{N-R} H(\mathbf{x}_l) \right)^{\frac{N-rM}{N-R}} \leq |D_{K/\mathbb{Q}}|^{\frac{N-rM}{2d}} \left(\prod_m H_{\mathrm{Ar}}(A''_m) \right)^{\frac{N-rM}{N-R}}. \tag{2.3.2}$$

By definition of the Arakelov height, we have H_{Ar} takes value in $[1, \infty)$. Thus $(\prod_m H_{\mathrm{Ar}}(A''_m))^{\frac{N-rM}{N-R}} \leq \prod_{i=1}^{rM} H_{\mathrm{Ar}}(A''_i)$. Now the conclusion follows because H_{Ar} is invariant under each σ_i. □

2.4 Faltings's version of Siegel's Lemma

In his famous paper `Diophantine approximation on abelian varieties` (*Annals of Math.* **133**:549–576, 1991), Faltings proved a fancier Siegel's Lemma. It plays a fundamental role for his proof of the Mordell–Lang Conjecture. In this section, we discuss about this.

2.4.1 Background and statement

Recall the following basic version of Siegel's Lemma, Lemma 2.1.1.

Lemma 2.4.1. *Let $A = (a_{ij})$ be an $M \times N$-matrix with entries in $\mathbb{Z}$. Set $B = \max_{i,j} |a_{ij}|$. If $N > M$, then* $\mathrm{Ker}(A)$ *contains a non-zero vector $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N$ such that*

$$\max_j |x_j| \leq (NB)^{\frac{M}{N-M}}.$$

Let us digest this lemma in the following way. The matrix A defines a linear map $\alpha \colon \mathbb{R}^N \to \mathbb{R}^M$ such that $\alpha(\mathbb{Z}^N) \subseteq \mathbb{Z}^M$, *i.e* α maps the lattice $\mathbb{Z}^N$ into the lattice $\mathbb{Z}^M$. If $N > M$, then we are able to find a non-trivial lattice point *of small norm* in $\mathrm{Ker}(\alpha)$. As we said before, $N - M$

should be understood to be $\dim \mathrm{Ker}(A)$ (although in the current formulation they may not be the same).

Faltings's fancier version looks not for only one, but for an *arbitrary number of linearly independent lattice points* in $\mathrm{Ker}(\alpha)$. To say that these lattice points are *of small norm*, we use the *successive minima*. Moreover, it is more natural to work with arbitrary normed real vector spaces.

Let $(V, \|\cdot\|)$ be a finite dimensional normed real vector space, and let Λ be a lattice (a discrete subgroup of V which spans V). Denote by $B(V)$ the unit ball $\{x \in V : \|x\| \leq 1\}$ in V.

Definition 2.4.2. *The n-th successive minimum of $(V, \|\cdot\|, \Lambda)$ is*

$$\begin{aligned}\lambda_n(V, \|\cdot\|, \Lambda) :&= \inf\{t > 0 : \ \Lambda \text{ contains } n \text{ linearly independent vectors of norm} \leq t\} \\ &= \inf\{t > 0 : tB(V) \text{ contains } n \text{ linearly-independent vectors of } \Lambda\}.\end{aligned}$$

Next for two normed real vector spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$, the *norm* of a linear map $\alpha\colon V \to W$ is defined to be

$$\|\alpha\| := \sup\left\{\frac{\|\alpha(x)\|_W}{\|x\|_V} : x \neq 0\right\}. \tag{2.4.1}$$

We are ready to state Faltings's version of Siegel's Lemma.

Theorem 2.4.3. *Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two finite dimensional normed real vector spaces, let Λ_V be a lattice in V and Λ_W be a lattice in W.*

Let $\alpha\colon V \to W$ be a linear map with $\alpha(\Lambda_V) \subseteq \Lambda_W$. Assume furthermore that there exists a real number $C \geq 2$ such that

(i) $\|\alpha\| \leq C$,

(ii) Λ_V *is generated by elements of norm* $\leq C$,

(iii) every non-zero element of Λ_V *and of* Λ_W *has norm* $\geq C^{-1}$.

Then for $U := \mathrm{Ker}(\alpha)$ with the induced norm $\|\cdot\|_U$ (the restriction of $\|\cdot\|_V$ on U) and the lattice $\Lambda_U := \Lambda_V \cap U$, we have

$$\lambda_{n+1}(U, \|\cdot\|_U, \Lambda_U) \leq \left(C^{3\dim V} \cdot (\dim V)!\right)^{1/(\dim U - n)}$$

for each $0 \leq n \leq \dim U - 1$.

Notice that the hypotheses (i)–(iii) can always be achieved by enlarging C.

The basic version of Siegel's Lemma (Lemma 2.4.1), up to changing the constant, follows from Theorem 2.4.3 with $n = 0$.

2.4.2 Proof of Theorem 2.4.3

The proof of Theorem 2.4.3 uses Minkowski's Second Theorem.

Let $(V, \|\cdot\|_V, \lambda_V)$ be a finite dimensional normed real vector space with a lattice. Set $d_V := \dim V$. For simplicity, denote by

$$V/\Lambda_V := \{v \in V : v = \sum_{j=1}^{d_V} \lambda_j v_j, \ 0 \leq \lambda_j < 1\}$$

where $\{v_1, \dots, v_{d_V}\}$ is a basis of Λ_V. Notice that V/Λ_V depends on the choice of the basis.

We can endow V with a Lebesgue measure μ_V as follows. Fix an isomorphism $\psi\colon V \simeq \mathbb{R}^{d_V}$ and use μ to denote the standard Lebesgue measure on $\mathbb{R}^{d_V}$. Then set for any Lebesgue measurable $A \subseteq \mathbb{R}^{d_V}$

$$\mu_V(\psi^{-1}(A)) = \mu(A). \tag{2.4.2}$$

Up to a constant, there is only one Lebesgue measure on V. Thus the quantity

$$\mathrm{Vol}(V) = \mathrm{Vol}(V, \|\cdot\|_V, \Lambda_V) := \frac{\mu_V(B(V))}{\mu_V(V/\Lambda_V)} \tag{2.4.3}$$

does not depend on the choice of μ_V; it clearly does not depend on the choice of the basis of Λ_V in the definition of V/Λ_V.

Theorem 2.4.4 (Minkowski's Second Theorem)**.** *With the notation above, we have*

$$\frac{2^{d_V}}{d_V!} \le \prod_{n=1}^{d_V} \lambda_n(V, \|\cdot\|_V, \Lambda_V) \cdot \mathrm{Vol}(V) \le 2^{d_V}.$$

Here we used the fact that the unit ball $B(V)$ is convex and symmetric (*i.e.* $B(V) = -B(V)$).

To apply Minkowski's Second Theorem to prove Theorem 2.4.3, we need one last preparation on the *quotient norm*. More precisely, on V/U, we consider the norm

$$\|\overline{v}\|_{V/U} := \inf\{\|v + u\|_V : u \in U\}$$

for each $v \in V$. Having this norm, we can define the unit ball $B(V/U)$. Moreover, $\alpha(\Lambda_V)$ is a lattice in $\alpha(V)$, which can then be viewed as a lattice in V/U by the natural isomorphism $V/U \simeq \alpha(V)$. So we can define $\mathrm{Vol}(V/U) := \mathrm{Vol}(V/U, \|\cdot\|_{V/U}, \alpha(\Lambda_V))$. Recall the notation from Theorem 2.4.3; we naturally have the quantity $\mathrm{Vol}(U) := \mathrm{Vol}(U, \|\cdot\|_U, \Lambda_U)$.

Lemma 2.4.5. $\mathrm{Vol}(V) \le 2^{\dim U}\mathrm{Vol}(U)\mathrm{Vol}(V/U)$.

Proof of Theorem 2.4.3 assuming Lemma 2.4.5. We will identify $V/U \simeq \alpha(V)$ in the proof. Take $w \in \alpha(\Lambda_V) \setminus \{0\}$. Write $w = \alpha(v)$ for some $v \in \Lambda_V$. Then

$$\|w\|_{V/U} = \inf_{u \in U} \|v + u\|_V \ge \frac{\|\alpha(v)\|_W}{\|\alpha\|} \ge C^{-2};$$

here the last inequality follows from hypotheses (i) and (iii). In particular, this implies that $\lambda_1(V/U, \|\cdot\|_{V/U}, \alpha(\Lambda_V)) \ge C^{-2}$.

Write $d_V := \dim V$ and $d_U := \dim U$. Minkowski's Second Theorem (applied to V/U) yields $\lambda_1(V/U, \|\cdot\|_{V/U}, \alpha(\Lambda_V))^{\dim V/U} \cdot \mathrm{Vol}(V/U) \le 2^{\dim V/U}$. Thus from the paragraph above, we get $\mathrm{Vol}(V/U) \le (2C^2)^{d_V - d_U}$.

Next, by hypothesis (ii), we have $\lambda_{d_V}(V, \|\cdot\|_V, \Lambda_V) \le C$. Thus Minkowski's Second Theorem (applied to V) yields $\mathrm{Vol}(V) \ge 2^{d_V} C^{-d_V}/d_V!$.

Apply Lemma 2.4.5 and the volume estimates above. Then we get

$$\mathrm{Vol}(U)^{-1} \le C^{3d_v - 2d_U} \cdot d_V!. \tag{2.4.4}$$

We apply another time Minkowski's Second Theorem (to U). For each $0 \le n \le d_U - 1$, we then get $\lambda_1(U, \|\cdot\|_U, \Lambda_U)^n \cdot \lambda_{n+1}(U, \|\cdot\|_U, \Lambda_U)^{d_U - n} \cdot \mathrm{Vol}(U) \le 2^{d_U}$. But $\lambda_1(U, \|\cdot\|_U, \Lambda_U) \ge C^{-1}$

by hypothesis (iii). So we obtain

$$\begin{aligned}\lambda_{n+1}(U, \|\cdot\|_U, \Lambda_U) &\le \left(2^{d_U}\mathrm{Vol}(U)^{-1}C^n\right)^{1/(d_U-n)} \\ &\le \left(2^{d_U}C^{n+3d_v-2d_U}\cdot d_V!\right)^{1/(d_U-n)} \qquad \text{by (2.4.4)} \\ &\le \left(C^{3d_V}\cdot d_V!\right)^{1/(d_U-n)}.\end{aligned}$$

Hence we are done. □

Proof of Lemma 2.4.5. Write $d_U := \dim U$ and $d_V := \dim V$.

Let μ_V and μ_U be the Lebesgue measures on V and U, respectively. On V/U we have a unique Lebesgue measure $\mu_{V/U}$ determined as follows: For any μ_V-measurable subset $E \subseteq V$, we have

$$\mu_V(E) = \int_{V/U} f_E(\overline{v})\mathrm{d}\mu_{V/U}(\overline{v})$$

where $f_E(\overline{v}) := \mu_U(\{u \in U : u + v \in E\})$; here $f_E(\overline{v})$ is independent of the representative v because μ_U is translation invariant.

We compute $f_{B(V)}(\overline{v})$ for $\overline{v} \in V/U$. If $\overline{v} \notin B(V/U)$, then $\|v\|_V > 1$. So $v \notin B(V)$ for $v + u$ for all $u \in U$. Thus $f_{B(V)}(\overline{v}) = 0$ in this case. If $\overline{v} \in B(V/U)$, then $v + u \in B(V)$ for some $u \in U$. Thus $\|u\|_U \le \|u + v\|_V + \|v\|_V \le 2$. So $f_{B(V)}(\overline{v}) \le \mu_U(2B(U)) = 2^{d_U} \cdot \mu_U(B(U))$ in this case. In either case, we have

$$\mu_V(B(V)) \le 2^{d_U} \cdot \mu_U(B(U)) \cdot \mu_{V/U}(B(V/U)). \tag{2.4.5}$$

Next we turn to $f_{V/\Lambda_V}(\overline{v})$. Let $\{u_1, \ldots, u_{d_U}\}$ be a basis of $\Lambda_U = \Lambda_V \cap U$ and expand it to a basis $\{u_1, \ldots, u_{d_U}, v_1, \ldots, v_{d_V-d_U}\}$ of Λ_V. Then $\{\overline{v}_1, \ldots, \overline{v}_{d_V-d_U}\}$ is a basis of $\alpha(\Lambda_V)$. For each $\overline{v} \in (V/U)/\alpha(\Lambda_V)$, we have

$$f_{V/\Lambda_V}(\overline{v}) = \mu_U(\{u \in U : u + v \in V/\Lambda_V\}) = \mu_U(U/\Lambda_U).$$

Otherwise $f_{V/\Lambda_V}(\overline{v}) = 0$. So

$$\mu_V(V/\Lambda_V) = \mu_U(U/\Lambda_U) \cdot \mu_{V/U}((V/U)/\alpha(\Lambda_V)). \tag{2.4.6}$$

Now the conclusion follows from the definition of the volumes $\mathrm{Vol}(V) = \mu_V(B(V))/\mu_V(V/\Lambda_V)$ *etc.* □

2.5 Reading material: Proof of Bombieri–Vaaler's Siegel Lemma

2.5.1 Adelic version of Minkowski's Second Theorem

The proof of Theorem 2.3.1 uses geometry of numbers over the adèles and Minkowski's Second Theorem. In this subsection, we introduce/recall these prerequisites.

Let K be a number field, $v \in M_K$ and K_v be the completion of K with respect to v. It is known that K_v is a locally compact group.

The **ring of adèles** of K is the subring

$$\mathbb{A}_K := \{\mathbf{x} = (x_v) \in \prod_{v \in M_K} K_v : x_v \in R_v \text{ up to finitely many } v\}.$$

of $\prod_{v\in M_K} K_v$.

One should be careful with the *topology* on $\mathbb{A}_K$. It is *not* induced by the product topology on $\prod_{v\in M_K} K_v$! Rather, we consider for each finite subset $S \subseteq M_K$ containing all archimedean places the product

$$H_S := \prod_{v\in S} K_v \times \prod_{v\notin S} R_v.$$

The product topology makes each such H_S into a locally compact topological group. The topology which we put on $\mathbb{A}_K$ is *the unique topology such that the groups H_S are open topological subgroups of $\mathbb{A}_K$*. In fact, this makes $\mathbb{A}_K$ a locally compact topological ring.

It is known that the diagonal map $K \to \mathbb{A}_K$, $x \mapsto (x_v)_{v \in M_K}$, makes K into a *discrete closed subgroup* of $\mathbb{A}_K$. Moreover $\mathbb{A}_K/K$ is compact.

Let $v|p \in M_{\mathbb{Q}}$. Then K_v is a locally compact group with Haar measure uniquely determined up to a scalar. We normalize this Haar measure as follows:

(a) if v is non-archimedean, β_v denotes the Haar measure on K_v normalized so that

$$\beta_v(R_v) = |D_{K_v/\mathbb{Q}_p}|_p^{1/2}$$

where R_v is the valuation ring of K_v and $D_{K_v/\mathbb{Q}_p}$ is the discriminant;

(b) if $K_v = \mathbb{R}$, then β_v is the usual Lebesgue measure;

(c) if $K_v = \mathbb{C}$, then β_v is twice the usual Lebesgue measure.

For each finite subset $S \subseteq M_K$ containing all archimedean places, the product measure $\beta_S := \prod_{v \in S} \beta_v \times \prod_{v \notin S} \beta_v|_{R_v}$ is then a Haar measure on the open topological subgruop H_S of $\mathbb{A}_K$. The measures β_S fit together to give a Haar measure β on $\mathbb{A}_K$.[5]

Let N be a positive integer. For each (archimedean) $v|\infty$, let S_v be a non-empty convex, symmetric, open subset of K_v^N; here "symmetric" means $S_v = -S_v$. For each (non-archimedean) $v \in M_K^0$, let S_v be a K_v-lattice in K_v^N, *i.e.* a non-empty compact open R_v-submodule of K_v^N. *Assume that $S_v = R_v^N$ for all but finitely many v.* Then the set

$$\Lambda := \{\mathbf{x} \in K^N : \mathbf{x} \in S_v \text{ for all } v \in M_K^0\}$$

is a K-lattice in K^N, *i.e.* a finitely generated $\mathcal{O}_K$-module which generates K^N as a vector space. Moreover, the image Λ_∞ of Λ under the canonical embedding $K^N \hookrightarrow E_\infty := \prod_{v|\infty} K_v^N$ is an $\mathbb{R}$-lattice in E_∞.[6]

Definition 2.5.1. *The n-**th successive minimum** of the non-empty convex symmetric open subset $S_\infty := \prod_{v|\infty} S_v$ of E_∞ with respect to the lattice Λ_∞ is*

$$\lambda_n := \inf\{t > 0 : tS_\infty \text{ contains } n \text{ } K\text{-linearly independent vectors of } \Lambda_\infty\}.$$

Now we are ready to state (the adelic version of) Minkowski's Second Theorem.

Theorem 2.5.2 (Minkowski's Second Theorem, adelic form)**.** *The successive minima defined above satisfy*

$$(\lambda_1 \cdots \lambda_N)^d \prod_{v \in M_K} \beta_v(S_v) \leq 2^{dN}.$$

Here, the product $\prod_{v \in M_K} \beta_v(S_v)$ should be understood to be the *volume* of S with respect to the Haar measure on $\mathbb{A}_K$ defined by the β_v's at each $v \in M_K$.

[5] With this in hand, we can shortly explain why we take the normalizations above. The Haar measure β on $\mathbb{A}_K$ induces a Haar measure $\beta_{\mathbb{A}_K/K}$ on the compact group $\mathbb{A}_K/K$, and the normalization above makes the volume of $\mathbb{A}_K/K$ to be 1.

[6] This is the familiar notion of a lattice, namely Λ_∞ is a discrete subgroup fo the $\mathbb{R}$-vector space E_∞ and that E_∞/Λ_∞ is compact.

2.5.2 Setup for the application of Minkowski's Second Theorem

For the purpose of proving Siegel's Lemma in the form of Theorem 2.3.1, we do the following preparation.

For the sets S_v: First, let Q_v^N be the unit cube in K_v^N of volume 1 with respect to the Haar measure β_v. More explicitly, $\mathbf{x} = (x_1, \dots, x_N) \in Q_v^N$ if and only if

$$\begin{cases} \max_n \|x_n\|_v < \frac{1}{2} & \text{if } v \text{ is real} \\ \max_n \|x_n\|_v < \frac{1}{2\pi} & \text{if } v \text{ is complex} \\ \max_n \|x_n\|_v \le 1 & \text{if } v \text{ is non-archimedean.} \end{cases}$$

Let A be an $N \times M$-matrix with entries in K such that $\mathrm{rk} A = M$. Set

$$S_v := \{\mathbf{y} \in K_v^M : A\mathbf{y} \in Q_v^N\}. \tag{2.5.1}$$

If v is archimedean, then S_v is a non-empty convex symmetric bounded open subset of K_v^N; indeed, under the injective linear map $\mathbf{x} \mapsto A\mathbf{x}$, the image of S_v is a linear slice of the cube Q_v^N. If v is non-archimedean, then one can show that S_v is a K_v-lattice in K_v^M and that $S_v = R_v^M$ for all but finitely many v; in fact in this case we have the following more precise result.

Proposition 2.5.3. *Let $v \in M_K^0$ lying over the prime number p. Then S_v is a K_v-lattice in K_v^M and $S_v = R_v^M$ for all but finitely many v. Moreover, we have*

$$\beta_v(S_v) = |D_{K_v/\mathbb{Q}_p}|_p^{M/2} \left(\max_I \|\det(A_I)\|_v\right)^{-1},$$

where I runs over all subsets of $\{1, \dots, N\}$ of cardinality M, and A_I is the $M \times M$-matrix formed by the i-th rows of A with $i \in I$.

Proof. Choose a subset $J \subseteq \{1, \dots, N\}$ of cardinality M such that $\|\det(A_J)\|_v = \max_I \|\det(A_I)\|_v$. Without loss of generality, we may assume $J = \{1, \dots, M\}$. Then $W := AA_J^{-1}$ is of the form

$$W = \begin{pmatrix} I_M \\ W' \end{pmatrix}.$$

For any subset $I \subseteq \{1, \dots, N\}$ of cardinality M, we have $\|\det(W_I)\|_v \le 1$ by choice of J. In particular, taking $I = \{1, \dots, l-1, l+1, \dots, M, M+j\}$ we get

$$\|w_{M+j,l}\|_v = \|\det(W_I)\|_v \le 1.$$

Thus all entries of W are in the valuation ring R_v and this proves

$$A_J S_v = \{\mathbf{y} \in K_v^M : W\mathbf{y} \in Q_v^N\} = R_v^M. \tag{2.5.2}$$

This proves that S_v is a K_v-lattice in K_v^M and that $S_v = R_v^M$ for all but finitely many v

It remains to compute $\beta_v(S_v)$. It is known that under the linear transformation $\mathbf{y} \mapsto A_J^{-1}\mathbf{y}$ on K_v^M, the volume transforms by the factor $\|\det(A_J)\|_v^{-1}$. Thus

$$\beta_v(S_v) = \|\det(A_J)\|_v^{-1} \beta_v(R_v^M) = \|\det(A_J)\|_v^{-1} |D_{K_v/\mathbb{Q}_p}|_p^{M/2}$$

which is what we desire. □

We also need to bound $\beta_v(S_v)$ from below for v archimedean. For this purpose, we have

Proposition 2.5.4. *Let $v \in M_K$ with $v|\infty$. Then*

$$\beta_v(S_v) \ge \|\det(A^* A)\|_v^{-1/2}$$

where $A^* = \overline{A}^t$ is the adjoint of A.

Proof. The proof uses *Vaaler's cube-slicing theorem*, which we state here without proof.

Vaaler's cube-slicing theorem. Let $N = n_1 + \cdots + n_r$ be a partition. Let $Q_N := B_{\rho(n_1)} \times \cdots \times B_{\rho(n_r)}$, where each $B_{\rho(n_j)}$ is the closed ball of volume 1 in $\mathbb{R}^{n_j}$ centered at 0.[7] For a real $N \times M$-matrix B of rank M, we have

$$\det(B^t B)^{-1/2} \le \mathrm{Vol}\left(\{\mathbf{y} \in \mathbb{R}^M : B\mathbf{y} \in Q_N\}\right). \tag{2.5.3}$$

An easier way to understand this volume bound is as follows. Let $L := \mathrm{Im}(B) \subseteq \mathbb{R}^N$ which is an M-dimensional subspace. Then (2.5.3) is equivalent to $1 \le \mathrm{Vol}(Q_N \cap L)$, *i.e.* the volume of a slice through the center of a product of balls of volume 1 is bounded below by 1.

Now we go back to the proof of Proposition 2.5.4. If $K_v = \mathbb{R}$, then this is (2.5.3) for $r = N$ and $n_1 = \cdots = n_N = 1$. Assume $K_v = \mathbb{C}$. Write $A = U + \sqrt{-1}V$ and $\mathbf{y} = \mathbf{u} + \sqrt{-1}\mathbf{v}$ for real $U, V, \mathbf{u}, \mathbf{v}$. Thus $K_v^M \simeq \mathbb{R}^{2M}$, $\mathbf{y} \mapsto (\mathbf{u}, \mathbf{v})$. Similarly we have $K_v^N \simeq \mathbb{R}^{2N}$. Now, the linear map $\mathbf{y} \mapsto A\mathbf{y}$ is given by the real $2N \times 2M$-matrix

$$A' = \begin{pmatrix} U & -V \\ V & U \end{pmatrix}$$

and

$$Q_v^N = \left\{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{2N} : u_j^2 + v_j^2 < \frac{1}{2\pi}\right\}.$$

By (2.5.3) for $n_1 = \cdots = n_N = 2$, we then have

$$\beta_v(S_v) \ge \det(A'^t A')^{-1/2}.$$

Since $A \mapsto A'$ is a ring homomorphism from the complex $N \times M$-matrices to the real $2N \times 2M$-matrices, we have $\det(A'^t A') = \det((A^* A)') = \det(A^* A)^2$. Hence we can conclude. □

2.5.3 Proof of Theorem 2.3.1

With the preparation from last subsection, we prove Bombieri–Vaaler's Siegel Lemma in this subsection. We start with:

Proposition 2.5.5. *Let A be an $N \times M$-matrix of rank M with entries in K. Then the image of A has a basis $\mathbf{x}_1, \ldots, \mathbf{x}_M$ with*

$$\prod_{m=1}^{M} H(\mathbf{x}_m) \le \left(\frac{2}{\pi}\right)^{\frac{Ms}{d}} |D_{K/\mathbb{Q}}|^{\frac{M}{2d}} H_{\mathrm{Ar}}(A)$$

where s is the number of complex places of K and $d = [K : \mathbb{Q}]$.

Proof. By Proposition 2.5.3 and Proposition 2.5.4, we have

$$\prod_{v \in M_K} \beta_v(S_V) \ge \prod_{v \in M_K^0} |D_{K_v/\mathbb{Q}_p}|_p^{M/2} \left(\prod_{v \in M_K^0} \max_I \|\det(A_I)\|_v \cdot \prod_{v|\infty} \|\det(A^* A)\|_v^{1/2} \right)^{-1}.$$

By (2.2.3), this becomes

$$\prod_{v \in M_K} \beta_v(S_V) \ge \left(\prod_{v \in M_K^0} |D_{K_v/\mathbb{Q}_p}|_p^{M/2} \right) H_{\mathrm{Ar}}(A)^{-d}.$$

It is known, from Algebraic Number Theory, that $|D_{K/\mathbb{Q}}|_p = \prod_{v|p} |D_{K_v/\mathbb{Q}_p}|_p$ for each prime number p. Thus the Product Formula implies $|D_{K/\mathbb{Q}}|^{-1} = \prod_{v \in M_K^0} |D_{K_v/\mathbb{Q}_p}|_p$. So the inequality above becomes

$$\prod_{v \in M_K} \beta_v(S_V) \ge |D_{K/\mathbb{Q}}|^{-M/2} H_{\mathrm{Ar}}(A)^{-d}.$$

[7] So the radius of $B_{\rho(n_j)}$ is $\rho(n_j) = \pi^{-1/2} \Gamma(n_j/2 + 1)^{1/n_j}$.

Thus, Minkowski's Second Theorem, Theorem 2.5.2, yields

$$\lambda_1 \cdots \lambda_M \le 2^M |D_{K/\mathbb{Q}}|^{M/2d} H_{\mathrm{Ar}}(A). \tag{2.5.4}$$

It remains to use the successive minima find the desired basis. For the specific sets S_v constructed in (2.5.1), recall the K-lattice $\Lambda = \{\mathbf{x} \in K^N : \mathbf{x} \in S_v \text{ for all } v \in M_K^0\}$ which is identified with its image Λ_∞ under the canonical embedding $K^N \hookrightarrow E_\infty = \prod_{v|\infty} K_v^N$. Let $\mathbf{y} \in K^M$ be a lattice point in λS_∞ for some $\lambda > 0$ and let $\mathbf{x} = A\mathbf{y}$. Then the definition of $S_\infty = \prod_{v|\infty} S_v$ yields $\max_n \|x_n\|_v < \lambda/2$ if v is real, $\max_n \|x_n\|_v < \lambda^2/2\pi$ if v is complex, and $\max_n \|x_n\|_v \le 1$ if $v \in M_K^0$. Thus we have

$$H(A\mathbf{y}) < \frac{\lambda}{2} \left(\frac{2}{\pi}\right)^{s/d}. \tag{2.5.5}$$

By the definition of successive minima, there are linearly independent lattice points $\mathbf{y}_1, \dots, \mathbf{y}_M \in K^M$ such that $\mathbf{y}_m \in \lambda_m S_\infty$ for each $m \in \{1, \dots, M\}$. Then we obtain the desired basis from (2.5.4) and (2.5.5), with $\mathbf{x}_m = A\mathbf{y}_m$. □

Proof of Theorem 2.3.1. For the $M \times N$-matrix A of rank M, its transpose A^t is an $N \times M$-matrix of rank M. It is attempting to apply Proposition 2.5.5 directly to A^t, but we need to do more.

We wish to find a basis of $\mathrm{Ker}(A)$ of small height. To do this, we first of all take an arbitrary basis $\mathbf{y}_1, \dots, \mathbf{y}_{N-M}$ of $\mathrm{Ker}(A)$, and let $A' := \begin{pmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_M \end{pmatrix}$. Then A' is an $N \times (N-M)$-matrix with rank $N - M$, and $\mathrm{Im}(A') = \mathrm{Ker}(A)$. Recall that $h_{\mathrm{Ar}}(A) = h_{\mathrm{Ar}}(\mathrm{Ker}(A))$ by Corollary 2.2.9. Hence $h_{\mathrm{Ar}}(A') = h_{\mathrm{Ar}}(A)$.

Apply Proposition 2.5.5 to A'. Then we get a basis $\mathbf{x}_1, \dots, \mathbf{x}_{N-M}$ of $\mathrm{Im}(A') = \mathrm{Ker}(A)$ such that

$$\prod_{l=1}^{N-M} H(\mathbf{x}_l) \le \left(\frac{2}{\pi}\right)^{(N-M)s/d} |D_{K/\mathbb{Q}}|^{\frac{N-M}{2d}} H_{\mathrm{Ar}}(A').$$

But $2/\pi < 1$. So we are done because $H_{\mathrm{Ar}}(A') = H_{\mathrm{Ar}}(A)$. □