Hermite, Mordell-Weil, Siegel

Hugues RANDRIAM

April 5, 2004

The aim of this paper is to present three finiteness or finite-generation theorems:

- Hermite’s theorem about number fields with given degree and ramification
- the Mordell-Weil theorem, concerning the group of rational points on abelian varieties
- Siegel’s theorem about integral points on curves of non-zero genus.

Our presentation follows mainly the one given in [2].

1 Hermite’s theorem

We begin with a weak form of Hermite’s finiteness theorem.

**Proposition 1.1.** There exists only finitely many number fields of given degree and discriminant.

**Proof.** Let \( n \) and \( D \) be two integers, and \( K \) a number field of degree \( n \) and discriminant \( D \) admitting \( r_1 \) real embeddings \( \sigma_1, \ldots, \sigma_{r_1} \) and \( 2r_2 \) complex embeddings \( \sigma_{r_1+1}, \ldots, \sigma_{r_1+r_2}, \overline{\sigma_{r_1+1}}, \ldots, \overline{\sigma_{r_1+r_2}} \), so that \( \mathcal{O}_K \) can be viewed as a lattice in \( \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \) via \( (\sigma_1, \ldots, \sigma_{r_1+r_2}) \).

Suppose first \( r_1 \neq 0 \). Then by Minkowski’s theorem there exists a constant \( C \) (depending explicitly on \( n \) and \( D \)) such that there exists \( \alpha \in \mathcal{O}_K \) with

\[
|\sigma_1(\alpha)| \leq C, \quad |\sigma_2(\alpha)| \leq 1/2, \ldots, |\sigma_{r_1+r_2}(\alpha)| \leq 1/2.
\]

Then by the product formula one also has \( |\sigma_1(\alpha)| \geq 1 \), so \( |\sigma_1(\alpha)| \neq |\tau(\alpha)| \) and hence \( \sigma_1(\alpha) \neq \tau(\alpha) \) for any embedding \( \tau \neq \sigma_1 \). This implies \( K = \mathbb{Q}(\alpha) \).
Now the (integral) coefficients of the minimal polynomial of such an \( \alpha \) can be bounded in terms of \( n \) and \( C \), so they can take only finitely many values. This in turn implies the finiteness result announced.

The case \( r_1 = 0 \) can be treated in the same way by considering \( \alpha \in \mathcal{O}_K \) verifying \( |\Re \sigma_1(\alpha)| \leq 1/2 \), \( |\Im \sigma_1(\alpha)| \leq C \), and \( |\sigma_j(\alpha)| \leq 1/2 \) for \( j \geq 2 \) (remark that \( \Im \sigma_1(\alpha) \) is then non-zero, so that \( \sigma_1(\alpha) \neq \overline{\sigma_1(\alpha)} \)). \( \square \)

**Proposition 1.2 (Hensel).** Let \( A \) be a Dedekind ring, \( K \) its quotient field, \( L \) a finite extension of \( K \), \( B \) the integral closure of \( A \) in \( L \), \( \mathfrak{P} \) a non-zero prime ideal of \( B \), and \( v_\mathfrak{P} \) the associated discrete valuation. Let \( p = \mathfrak{P} \cap \mathcal{O}_K \) and \( e = v_\mathfrak{P}(p) \) the corresponding ramification index. Suppose furthermore that the residual extension is separable. Then the exponent of \( \mathfrak{P} \) in the different \( \mathcal{D}_{L/K} \) can be bounded as follows:

\[
(2) \quad v_\mathfrak{P}(\mathcal{D}_{L/K}) \leq e - 1 + v_\mathfrak{P}(e).
\]

**Proof.** We recall the proof from [1]. Without loss of generality one can suppose \( A \) and \( B \) complete local rings. After replacing \( K \) by its maximal unramified extension in \( L \), one can also suppose \( L \) totally ramified over \( K \), so \( [L : K] = e \). In this situation there exists \( \pi \in B \) that satisfies an Eisenstein equation \( f(\pi) = 0 \), with

\[
(3) \quad f(X) = \sum_{i=0}^{e} a_i X^i, \quad a_e = 1, \quad a_{e-1}, \ldots, a_0 \in p, \quad a_0 \notin p^2,
\]

and such that \( B = A[\pi] \). This last assertion implies that the different is generated by \( f'(\pi) = \sum_{i=1}^{e} i a_i \pi^{i-1} \). Remark then that the \( e - 1 \) terms of this sum all have distinct valuation, so that

\[
(4) \quad v_\mathfrak{P}(\mathcal{D}_{L/K}) = \inf_{1 \leq i \leq e} v_\mathfrak{P}(i a_i \pi^{i-1}) \leq v_\mathfrak{P}(e a_e \pi^{e-1}) = e - 1 + v_\mathfrak{P}(e)
\]

since \( a_e = 1 \). \( \square \)

**Theorem 1.3 (Hermite’s theorem).** Let \( K \) be a number field, \( n \) an integer and \( S \) a finite set of places of \( K \). Then \( K \) admits only finitely many extensions of degree \( n \) unramified out of \( S \).

**Proof.** If \( L \) is such an extension, proposition 1.2 gives a bound for the different \( \mathcal{D}_{L/K} \) at each element of \( S \), and as the absolute discriminant of \( L \) can be expressed in terms of this different and of the absolute discriminant of \( K \), the finiteness result follows from proposition 1.1. \( \square \)
2 The Mordell-Weil theorem

Theorem 2.1 (Chevalley-Weil). Let $K$ be a number field, $X$ and $Y$ two smooth projective varieties over $K$, and $h : X \to Y$ an étale morphism. Then there exists a finite extension $L$ of $K$ such that $Y(K)$ is contained in $h(X(L))$.

Proof. Using the openness of the étale locus and the projectivity hypothesis, one gets a finite set $S$ of places of $K$, projective models $\mathfrak{X}$ and $\mathfrak{Y}$ of $X$ and $Y$ over $\mathcal{O}_S$, and an étale morphism $\mathfrak{h} : \mathfrak{X} \to \mathfrak{Y}$ that extends $h$. Then if $\Sigma \subset \mathfrak{Y}$ is the Zariski closure of $Q \in Y(K)$, and if $P$ is the generic point of an irreducible component of the fiber product $\mathfrak{h}^{-1}(\Sigma) = \Sigma \times_\mathfrak{Y} \mathfrak{X}$, the field $K(P)$ is an extension of $K$ of degree less than the degree of $h$ unramified out of $S$. By Hermite’s theorem there exists only finitely many such extensions, so they are all contained in their compositum. This compositum is the $L$ we were looking for. \qed

Remarks.

1. One in fact has proved the stronger result: $h^{-1}(Y(K)) \subset X(L)$.

2. For non algebraic-geometry oriented readers, the preceding construction can be made a little more concrete by considering equations over $K$ defining $Y$ as a subvariety of some $\mathbb{P}^N$, and $X$ as a subvariety of some $\mathbb{P}^M$ over $Y$, the last ones with non-vanishing jacobian determinant. Then the set $S$ of places to throw away are those occuring in the denominators of these equations and jacobian.

Proposition 2.2 (Weak Mordell-Weil theorem). Let $K$ be a number field, $A$ an abelian variety over $K$. Then for any integer $m \geq 1$, the group $A(K)/m.A(K)$ is finite.

Proof. Applying the Chevalley-Weil theorem with $X = Y = A$ and $h$ the multiplication-by-$m$ map, one gets a finite extension $L$ of $K$ such that $m.A(L)$ contains $A(K)$. Without loss of generality one can suppose that $L$ is Galois over $K$ with group $G$. By construction the group of Galois invariants of $m.A(L)$ is then the whole of $A(K)$:

\begin{equation}
(m.A(L))^G = A(K).
\end{equation}

This and the long exact sequence in cohomology associated to the short exact sequence of $G$-modules

\begin{equation}
0 \to A_m \to A(L) \xrightarrow{m} m.A(L) \to 0
\end{equation}

3
gives an injection of $A(K)/mA(K)$ into the finite set $H^1(G, A_m)$. This proves the proposition. \qed

Remark. The preceding cohomological argument can be made a little more explicit as follows. Without loss of generality one can suppose that $A(K)$ contains $A_m$. For any $Q \in A(K)$, choose an $m$-th root $P$ of $Q$ in $A(L)$. Then for any $\sigma \in G$, the difference $P^\sigma - P$ lies in $A_m$ and does not depend on the choice of $P$. The map $\sigma \mapsto P^\sigma - P$ is then a group homomorphism from $G$ to $A_m$, and one easily checks that the map sending $Q$ to this homomorphism gives a group homomorphism from $A(K)$ to $\text{Hom}(G, A_m)$ with kernel $mA(K)$, so that $A(K)/mA(K)$ embeds in the finite group $\text{Hom}(G, A_m)$.

From now on we will suppose known the basic properties of the Néron-Tate height $\tilde{h}$ associated to a symmetric ample line bundle on $A$, in particular the fact that $\tilde{h}$ comes from a positive definite symmetric bilinear form on the (for the moment possibly infinite dimensional) real vector space $\overline{A(K)} \otimes_{\mathbb{Z}} \mathbb{R}$.

**Theorem 2.3 (Mordell-Weil).** Let $K$ be a number field, $A$ an abelian variety over $K$, and $\Gamma = A(K)$ its group of rational points. Then $\Gamma$ is finitely generated.

**Proof.** Choose any symmetric ample line bundle on $A$, and define a norm $|.|$ on $\Gamma \otimes_{\mathbb{Z}} \mathbb{R}$ by the formula $|P|^2 = \tilde{h}(P)$, where $\tilde{h}$ is the corresponding Néron-Tate height. For any real $t$, put

$$
\Gamma_t = \{ P \in \Gamma \mid |P| \leq t \},
$$

which is a finite set by Northcott’s theorem. Choose $P_1, \ldots, P_n \in \Gamma$ representatives of $\Gamma/2\Gamma$ (which is finite by the weak Mordell-Weil theorem) and put

$$
C = \max\{|P_1|, \ldots, |P_n|, 1\}.
$$

We claim that $\Gamma$ is generated by the finite set $\Gamma_{2C}$. As $\Gamma$ is the union of the $\Gamma_{kC}$ for $k \in \mathbb{N}$, all we have to prove is that the subgroup $<\Gamma_{2C}>$ generated by $\Gamma_{2C}$ contains all the $\Gamma_{kC}$. This is obviously true for $k = 2$. Now we proceed by induction, supposing that $<\Gamma_{2C}>$ contains $\Gamma_{(k-1)C}$, and taking $P \in \Gamma_{kC}$, where $k \geq 3$. Now there is a unique $P_1$ having the same class as $P$ in $\Gamma/2\Gamma$, so that there exists $Q \in \Gamma$ with $P = 2Q + P_1$. One then has

$$
|Q| = \frac{1}{2}|P - P_1| \leq \frac{1}{2}(|P| + |P_1|) \leq \frac{k+1}{2}C \leq (k-1)C.
$$

Thus $Q \in \Gamma_{(k-1)C} < \Gamma_{2C}$, and $P = 2Q + P_1 < \Gamma_{2C} + \Gamma_C = < \Gamma_{2C}>$, which proves the claim. \qed
Remark. In fact, a careful analysis of the proof would lead to the fact that one even has $\Gamma = < \Gamma_{C_0} >$ with $C_0 = \max\{|P_1|, \ldots, |P_n|\}$.

3 Siegel’s Theorem

To begin with, we quote the following geometric reformulation of Roth’s approximation theorem:

Theorem 3.1. Let $V$ be a smooth projective variety over a number field $K$, and $h$ a height function associated to an ample line bundle $\mathcal{L}$ on $V$. Fix an embedding $\sigma : K \hookrightarrow \mathbb{C}$, choose a Riemannian metric on $V_{\sigma}(\mathbb{C})$, and denote by $d_\sigma$ the corresponding distance. Then there exists $\delta > 0$ such that for any $\alpha \in V(K)$ and for any $C > 0$, there exists only finitely many $\omega \in V(K)$ with

$$d_\sigma(\alpha, \omega) \leq Ce^{-\delta h(\omega)}. \quad (10)$$

Remark that the particular case $V = \mathbb{P}^1$ (or more generally $V = \mathbb{P}^N$) with $\mathcal{L} = \mathcal{O}(1)$ gives precisely the usual version of Roth’s theorem, where any $\delta > 2$ is then convenient. The general case follows by using some power of $\mathcal{L}$ to embed $V$ in some $\mathbb{P}^N$.

In the following theorem, we show that if $V = A$ is an abelian variety, $\delta$ can be taken arbitrarily small.

Theorem 3.2. Let $A$ be an abelian variety over a number field $K$, and $h$ a height function associated to an ample line bundle $\mathcal{L}$ on $A$. Fix an embedding $\sigma : K \hookrightarrow \mathbb{C}$, choose a Riemannian metric on $A_{\sigma}(\mathbb{C})$, and denote by $d_\sigma$ the corresponding distance. Then for any $\varepsilon > 0$, for any $\alpha \in A(K)$ and for any $C > 0$, there exists only finitely many $\omega \in A(K)$ with

$$d_\sigma(\alpha, \omega) \leq Ce^{-\varepsilon h(\omega)}. \quad (11)$$

Proof. One can suppose that $\mathcal{L}$ is symmetric, that $h$ is the associated Néron-Tate height, and that the metric is invariant by translation. Let $\delta > 0$ be the constant given by the preceding theorem, and choose an integer $m$ such that $\varepsilon m^2 > \delta$. Proceeding by contradiction, suppose there is an infinite sequence $(\omega_n)_{n \in \mathbb{N}}$ of distinct elements of $\Gamma = A(K)$ satisfying (11). By Northcott’s theorem the $h(\omega_n)$ tend to infinity, so that the $\omega_n$ converge to $\alpha$ (for the complex topology on $A_{\sigma}(\mathbb{C})$).

By the weak Mordell-Weil theorem, there are infinitely many $\omega_n$ having the same class modulo $m\Gamma$, and after a translation one can assume this class is zero. Thus after extracting a subsequence one can suppose there
exists $\omega'_n$ in $\Gamma$ such that $\omega_n = m\omega'_n$ for all $n$. Extracting a subsequence again one can suppose the $\omega'_n$ converge to some $\alpha'$ in $A_\sigma(\mathbb{C})$. One then has $\alpha = m\alpha'$, so that $\alpha'$ lies in fact in $A(\mathbb{K})$. Now, since for all big enough $n$, $d_\sigma(\alpha', \omega'_n) = \frac{1}{m}d_\sigma(\alpha, \omega_n)$ and $h(\omega'_n) = \frac{1}{m}\sigma h(\omega_n)$, one finds

$$d_\sigma(\alpha', \omega'_n) = \frac{1}{m}d_\sigma(\alpha, \omega_n) \leq \frac{C}{m}e^{-\varepsilon h(\omega_n)} = \frac{C}{m}e^{-\varepsilon m^2 h(\omega'_n)} \leq \frac{C}{m}e^{-\delta h(\omega'_n)},$$

which contradicts theorem 3.1. \hfill \Box

Now we explain how this theorem of Diophantine approximation on abelian varieties can be used to get finiteness results for integral points on curves.

**Theorem 3.3 (Siegel).** Let $K$ be a number field, $C$ a smooth projective curve over $K$ of genus $g \geq 1$, and $z : C \to \mathbb{P}^1_K$ a rational function on $C$ of degree $D \geq 1$. Then there exists only finitely many $P \in C(K)$ with $z(P) \in \mathcal{O}_K$.

**Proof.** Suppose by contradiction that there is an infinite sequence $P_1, P_2, \ldots$ of distinct elements of $C(K)$ with $z_i = z(P_i) \in \mathcal{O}_K$ for all $i$. This last condition implies that the height of $z_i$ can be expressed as

$$h(z_i) = \sum_{\tau : K \hookrightarrow \mathbb{C}} \log^+ |\tau(z_i)|,$$

and then for each $i$ there is a $\tau$ with $\log |\tau(z_i)| \geq \frac{1}{|K: \mathbb{Q}|} h(z_i)$. Thus after extracting a subsequence, one can suppose there is an embedding $\sigma : K \hookrightarrow \mathbb{C}$ such that for all $i$,

$$\log |\sigma(z_i)| \geq \frac{1}{|K: \mathbb{Q}|} h(z_i).$$

By compactness, after extracting again, one can suppose the $P_i$ converge to some $P$ in $C_\sigma(\mathbb{C})$. Remark that Northcott’s theorem implies that the terms in (14) tend to infinity, so that $P$ is a pole of $z$, and hence lies in $C(\mathbb{K})$. Let $e$ be the order of this pole.

Now choose an embedding of $C$ in its Jacobian $J$, $\mathcal{L}$ a symmetric ample line bundle on $J$ with associated Néron-Tate height $h_\mathcal{L}$, and $d_\sigma$ a distance on $J_\sigma(\mathbb{C})$, as in theorem 3.2. Then one has

$$\log |\sigma(z_i)| = -e \log d_\sigma(P_i, P) + O(1)$$

6
and for any $\varepsilon > 0$,

$$
(16) \quad h(z_i) \geq \left( \frac{D}{\deg L_i} - \varepsilon \right) h_{L_i}(P_i) + O(1).
$$

Combining this with (14) one finds $\kappa > 0$ such that

$$
(17) \quad -\log d_\sigma(P_i, P) \geq \kappa h_{L_i}(P_i)
$$

for all $i \gg 0$, which contradicts theorem 3.2. \qed

*Remark.* In theorems 3.1 and 3.2 we restricted our attention to an archimedean place. However, analogous results still hold in a non-archimedean setting. Thus Siegel’s finiteness theorem generalizes to points with coordinates in $O_S$ for any finite set of places $S$.

**References**
