Annali della
Scuola Normale Superiore di Pisa
Classe di Scienze

Enrico Bombieri

The Mordell conjecture revisited

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 17, nº 4 (1990), p. 615-640
<http://www.numdam.org/item?id=ASNSP_1990_4_17_4_615_0>

The Mordell Conjecture Revisited

ENRICO BOMBIERI

1. - Introduction

Let C be an algebraic curve defined over a number field k. A well-known theorem of Faltings states that if C has genus at least 2 then C has only finitely many points with coordinates in k, as originally conjectured by Mordell. Quite recently, a completely different proof was found by Vojta, by a method strongly reminiscent of familiar techniques in diophantine approximation, in the so-called Thue-Siegel-Roth theory. Vojta's proof is rather difficult, depending on Arakelov's arithmetic intersection theory and on the arithmetic Riemann-Roch theorem of Gillet and Soulé for arithmetic threefolds. A very important generalization and simplification of Vojta's proof was then found by Faltings, who proved Lang's conjecture on the finiteness of rational points on subvarieties of abelian varieties not containing any translate of an abelian variety. This new proof by Faltings avoids the use of the difficult arithmetic Riemann- Roch theorem, but still uses arithmetic intersection theory and a sophisticated notion of height.

In this paper we shall give a reasonably self-contained proof of Faltings' theorem for curves, based on Vojta's approach but substituting the elementary theory of heights in place of intersection theory. As in Faltings, the arithmetic Riemann-Roch is replaced by an appropriate use of Siegel's Lemma, although our treatment is somewhat different.

Algebraic geometry is kept to a minimum, namely the easy geometric Riemann-Roch inequality on algebraic surfaces, and classical facts about projective embeddings and endomorphisms of abelian varieties. A technical difficulty with the use of derivations is resolved by an appeal to a local version of Eisenstein's theorem on denominators of coefficients of Taylor series of algebraic functions.

The main ideas in this paper are already in the fundamental papers of Vojta [V] and Faltings [F]; our contributions, if any, are technical in nature. The price paid for a more elementary exposition is losing the beauty of arithmetic intersection theory and arithmetic Riemann-Roch, now replaced by simple-

Pervenuto alla Redazione il 25 Settembre 1990.

minded arguments with polynomials and Dirichlet's box principle. Thus we do not view this paper as superseding previous work; rather, we hope that it will make these results readily accessible to a larger audience.

As in all preceding papers, our proof of Faltings' theorem is ineffective and no bound can be given for the height of rational points on C. On the other hand, it will be clear from our proof that our arguments belong entirely to constructive algebraic geometry and therefore all constants appearing in this paper are effectively and explicitly computable; in particular, one can give explicit bounds for the number of rational points on C. We believe this to be of some interest and hope to return to these aspects of the proof in the future.

The author wishes to thank A. Granville, P. Vojta, W. Schmidt and H. Lenstra for several useful comments, and Rutgers University and the Università di Pisa for providing financial support during the preparation of this paper.

2. - The elementary theory of heights

We normalize the absolute values $|\ \ |_v$ of number fields k as follows. If $k = \mathbf{Q}$ then they coincide with the usual p- adic and euclidean absolute values, and in general they are consistent by field extension. Thus for every finite extension K of k and every $x \in k$, $x \neq 0$ we want

$$\log |x|_v = \sum_{w|v} \log |x|_w,$$

where $w|v$ means that w runs over all places of K lying over the place v of k. The point of this normalization is to make sure that all global heights are independent of the field of definition of the objects involved. This allows us to not worry too much over fields of definition and to extend heights to all of $\overline{\mathbf{Q}}$. For the rest of this paper, k will be an algebraic number field and, unless stated otherwise, a field of definition for all varieties, sections and polynomials we shall consider.

Let X be a smooth projective variety over $\overline{\mathbf{Q}}$ and let ϕ be a projective embedding $\phi : X \to \mathbf{P}^n$ of X into projective space $\mathbf{P}^n$, with standard co-ordinates $x_0, \cdots, x_n$.

DEFINITION. *The height of a point $p \in X(\overline{\mathbf{Q}})$, relative to the embedding $\phi : X \to \mathbf{P}^n$ is*

$$h_\phi(p) = \sum_v \max_j \log |x_j|_v$$

with $x = \phi(p)$ and where the sum is over all places v of any field of definition for X, p, ϕ.

The preceding definition does not depend of the choice of representatives for x, because of the product formula in k.

By linearity, the notion of height relative to a projective embedding is extended to the free group $\mathrm{Emb}(X)$ generated by all projective embeddings of X, and for $\gamma = \sum n_i\phi_i \in \mathrm{Emb}(X)$ we denote by h_γ the corresponding height $h_\gamma = \sum n_i h_{\phi_i}$.

The simplest example is $X = \mathbf{P}^n$, the n-dimensional projective space with standard co-ordinates $(x_0, \cdots, x_n)$, taking ϕ to be the identity map. The height of the point $p = (x_0, \cdots, x_n)$ is (we omit the reference to ϕ here):

$$h(p) = \sum_v \max_j \log |x_j|_v.$$

This gives the well-known elementary definition of height of an algebraic point in projective space and, by restriction, a corresponding height in affine space, again denoted by h.

There is an obvious homomorphism

$$cl : \mathrm{Emb}(X) \to \mathrm{Pic}(X)$$

obtained by associating to a projective embedding the linear equivalence class of hyperplane sections of X in this embedding. Since every divisor can be written as a difference of two hyperplane sections in suitable embeddings, cl is a surjective homomorphism. Let also $cl(D)$ denote the class of a divisor D modulo linear equivalence. The key results on these heights, due to Weil [W], are summarized as follows.

WEIL'S THEOREM. *The height h_γ has the following properties*:

(i) *it is additive in* $\mathrm{Emb}(X)$, *and h_γ is bounded on $X(\bar{k})$ if $cl(\gamma) = 0$;*

(ii) *if $f : X \to Y$ is a morphism and γ, γ' are two pairs such that $cl(\gamma') = f^*cl(\gamma)$, then $h_{\gamma'} - h_\gamma \circ f$ is a bounded function on $X(\bar{k})$;*

(iii) *if $A > 0$ is effective and $cl(\alpha) = cl(A)$ then h_α is bounded below on* $X(\bar{k})$ – (base locus of $|A|$)*;*

(iv) *if ϕ is a projective embedding, subsets of $X(\bar{k})$ with bounded degree and bounded h_ϕ height are finite.*

In the special case in which X is an abelian variety, one can do things in a more precise fashion, as shown by Néron and Tate, because now any equivalence class of heights has a distinguished representative. Let A be an abelian variety defined over a number field k. Let D be a divisor on A and let $h = h_D$ be any height associated to D. For any x, y in A one considers $B_n(x,y) = 4^{-n}\bigl(h(2^n x + 2^n y) - h(2^n x) - h(2^n y)\bigr)$; by using properties of endomorphisms of abelian varieties one sees that $B_n(x,y)$ is a Cauchy sequence, and that $B(x,y) = \lim B_n(x,y)$ is a bilinear form in x, y. One then proves in a similar way that $L(x) = \lim 2^{-n}\bigl(h(2^n x) - 4^n \frac{1}{2} B(x,x)\bigr)$ exists and is a linear function in x. This shows that $h_D(x) = \frac{1}{2} B(x,x) + L(x) + O(1)$ and that the bilinear form B and the linear function L depend only on the divisor class

$cl(D)$ of D. The height $\hat{h}_{cl(D)}(x) = \frac{1}{2} B(x, x) + L(x)$ is the Néron-Tate height of the divisor class $cl(D)$.

Let $[m]$ denote the homomorphism $x \to mx$ on A, and for a divisor class c let $c^- = [-1]^*c$. The class c is said to be even if $c^- = c$, and odd if $c^- = -c$. This yields a decomposition, in $Pic(A) \otimes \mathbf{Q}$, of a class c into $c_{\text{even}} = \frac{1}{2}(c + c^-)$ and $c_{\text{odd}} = \frac{1}{2}(c - c^-)$. Then we have $\hat{h}_c(mx) = m^2 \hat{h}_{c_{\text{even}}}(x) + m\hat{h}_{c_{\text{odd}}}(x)$. For complete proofs and further details, we refer to [L], Ch. 5.

One final point about heights. Most of our calculations will be with heights of points in projective or affine spaces, and the distinction between them should always be clear from the context. Also, if P is a polynomial we shall denote by $h(P)$ the height of the homogeneous vector of coefficients of P, and a similar notation will apply to vectors of polynomials. In case we want to consider the affine height of a value of a polynomial P at a point ξ, we shall always write $h(P(\xi))$, so that no confusion should arise from our notation.

3. - The Vojta divisor

Let C be a projective non-singular curve of genus $g \geq 2$, defined over a number field k. Let P be a divisor of degree 1 on C defined over k, which we fix once for all. Let Δ denote the diagonal of $C \times C$ and let $\Delta' = \Delta - P \times C - C \times P$. The divisor

$$V = d_1 P \times C + d_2 C \times P + d\Delta',$$

is said to be a Vojta divisor if d_1, d_2, d are positive integers with $gd^2 < d_1 d_2 < g^2 d^2$.

We are interested in expressing the divisor V as a difference of two well-chosen divisors on $C \times C$, in order to calculate an associated height. Hence we begin by choosing a positive integer s such that $B = sP \times C + sC \times P - \Delta'$ is linearly equivalent to a hyperplane section in some embedding $\phi_B : C \times C \to \mathbf{P}^m$, and we define $h_B = h_{\phi_B}$. For sufficiently large d every section of $\mathcal{O}(dB)$ is the pull-back by ϕ_B^* of a section of $\mathcal{O}_{\mathbf{P}^m}(d)$. Geometrically, this means that for large d effective divisors linearly equivalent to dB are complete intersections of $C \times C$ with hypersurfaces of degree d in the ambient projective space. In other words, every regular section s of $\mathcal{O}(dB)$ is the restriction to $C \times C$ of a homogeneous polynomial of degree d in the co-ordinates $y_0, \cdots, y_m$ of $\mathbf{P}^m$. An elementary short proof is for instance in Mumford [M2], Ch.6, (6.10) Theorem, p. 102.

In a similar way, we choose a large integer N such that NP is linearly equivalent to a hyperplane section of C in some embedding $\phi_{NP} : C \to \mathbf{P}^n$, and we define $h_{NP} = h_{\phi_{NP}}$. We also note that the embedding ϕ_{NP} is determined only up to a projective automorphism of $\mathbf{P}^n$, i.e. up to a choice of a basis of sections of $\mathcal{O}(NP)$. We shall use this extra freedom to ensure that certain geometric constructions (mainly projections into linear subspaces) are sufficiently generic, usually without proof.

The embedding ϕ_{NP} gives a product embedding $\psi : C \times C \to \mathbf{P}^n \times \mathbf{P}^n$. Again, if δ_1 and δ_2 are sufficiently large integers then every section of $\mathcal{O}(\delta_1(NP) \times C + \delta_2 C \times (NP))$ is the pull back by ψ^* of a section of $\mathcal{O}_{\mathbf{P}^n \times \mathbf{P}^n}(\delta_1, \delta_2)$. Geometrically, this means that, for large δ_1 and δ_2, divisors on $C \times C$ linearly equivalent to $\delta_1(NP) \times C + \delta_2 C \times (NP)$ are complete intersections of $C \times C$ with hypersurfaces of bidegree (δ_1, δ_2) in $\mathbf{P}^n \times \mathbf{P}^n$. In other words, every regular section of $\mathcal{O}(\delta_1(NP) \times C + \delta_2 C \times (NP))$ is the restriction to $C \times C$ of a bihomogeneous polynomial of bidegree (δ_1, δ_2) in the co-ordinates $(x_0, \cdots, x_n, x'_0, \cdots, x'_n)$ of $\mathbf{P}^n \times \mathbf{P}^n$.

Suppose that $\delta_1 = (d_1 + sd)/N$ and $\delta_2 = (d_2 + sd)/N$ are integers, and let $V = d_1 P \times C + d_2 C \times P + d\Delta'$, hence

$$V = \delta_1(NP) \times C + \delta_2 C \times (NP) - dB.$$

Let ψ_{δ_1,δ_2} be the Segre embedding determined by all monomials of bidegree (δ_1, δ_2) in $x_0, \cdots, x_n$ and $x'_0, \cdots, x'_n$, let ϕ_{dB} be the Segre embedding determined by all monomials of degree d in $y_0, \cdots, y_m$ and let $\gamma = \psi_{\delta_1,\delta_2} - \phi_{dB}$. Then it is easily verified that $cl(\gamma) = cl(V)$, $h_{\psi_{\delta_1,\delta_2}}(z, w) = \delta_1 h_{NP}(z) + \delta_2 h_{NP}(w)$ and $h_{\phi_{dB}}(z, w) = d h_B(z, w)$, therefore we have an associated height on $C \times C$:

$$h_V(z, w) = h_\gamma(z, w) = \delta_1 h_{NP}(z) + \delta_2 h_{NP}(w) - d h_B(z, w).$$

4. - A first idea for a proof, and an outline of the paper

As early as 1965, Mumford showed that the height h_Δ on $C \times C$ could be expressed, up to bounded quantities, in terms of Néron-Tate heights on the Jacobian of C. It then follows that

$$h_\Delta(z, w) = (\text{quadratic form}) + (\text{linear form}) + O(1)$$

by the quadraticity of heights on abelian varieties. Since Δ is an effective curve on $C \times C$, the left-hand side of this equation is bounded below away from the diagonal by a constant. On the other hand, one sees directly that the quadratic form in the right-hand side of this equation is indefinite, if the genus g is at least 2. This puts strong restrictions on the pair (z, w) because it means that z and w, considered as points in the Mordell-Weil group of the Jacobian, can never be nearly parallel with respect to the positive definite inner product determined by the Néron pairing. A simple geometric argument now shows that the heights of rational points on C, arranged in increasing order, grow at least exponentially. This is in sharp contrast with the quadratic growth one encounters on elliptic curves, and shows that rational points on curves of genus 2 or more are much harder to come by.

The diagonal Δ is a Vojta divisor, except for the fact that now $d_1 = 1$, $d_2 = 1$, $d = 1$ so that the inequalities characterizing a Vojta divisor are not satisfied. However, it is easy to see that Mumford's method applies generally to any divisor linear combination of $P \times C$, $C \times P$ and Δ. Now the condition $d_1 d_2 < g^2 d^2$ simply expresses the fact that the associated quadratic form is indefinite, and again we get an useful result if we can show that the height is bounded below. As in Mumford, it suffices to have an effective curve in the linear equivalence class of the divisor, and now the other condition $d_1 d_2 > g d^2$ for a Vojta divisor assures, by Riemann-Roch, that multiples of V have effective representatives.

The advantage in this generalization of Mumford's result is that now we have a two-parameter family of indefinite quadratic forms at our disposal, instead of just one. Thus one is tempted, given (z, w), to choose a quadratic form such that its value at (z, w) is negative, which would yield a contradiction unless z and w have bounded height.

The new problem one faces here is the fact that the choice of the quadratic form depends on the ratio of the heights of z and w, and therefore we need show not only that h_V is bounded below, but also that the lower bound has a sufficiently good uniformity with respect to the quadratic form. This is where arithmetic algebraic geometry has been used: arithmetic Riemann-Roch, for finding a good effective representative for V defined by equations with "small" coefficients, and arithmetic intersection theory for a precise control of the bounded terms arising in the elementary theory of heights.

As Vojta's paper clearly shows, this idea is overly simple and there is one more big obstacle to overcome. The argument used in obtaining a lower bound for h_V fails if the effective representative for V goes through the point (z, w) we are studying. By an appropriate use of derivations, one sees that this is not a too serious difficulty unless the representative of the divisor V goes through (z, w) with very high multiplicity. On the other hand, this representative must be defined by equations with small coefficients and there is very little room for moving it away from (z, w), and one cannot exclude a priori that this divisor has very high multiplicity at (z, w).

This situation is reminiscent of a familiar difficulty in transcendence theory, namely the non-vanishing at specific points of functions arising from auxiliary constructions. In the classical case, various independent techniques have been devised for this purpose: Roth's Lemma, which is arithmetic in nature, the algebro-geometric Dyson's Lemma and the Zero Estimates of Masser and Wüstholz.

Vojta, by proving a suitable generalization of Dyson's Lemma, shows that if $d_1 d_2$ is sufficiently close to $g d^2$ then any effective representative for V does not vanish much at (z, w), thereby completing the proof.

Faltings proceeds in a different way, using a new geometric tool, the Product Theorem. He is able to show that the difficulty with high multiplicity can be eliminated, except perhaps for a set of "bad" points (z, w) which is contained in a product subvariety of $C \times C$. It should be noted that Faltings' result applies not only to $C \times C$ but in fact to a product of an arbitrary number of varieties, thus providing a tool for handling higher dimensional varieties by induction on the dimension.

Here we solve this problem in classical fashion, namely by a direct application of Roth's Lemma.

The rest of this paper is roughly divided as follows. We begin by applying Mumford's method, expressing the height h_V, V a Vojta divisor, in terms of the standard Néron-Tate bilinear form of the Jacobian of C. Mumford's theorem also follows, yielding the analogue of the so-called "strong gap principle" in the Thue-Siegel-Roth theory.

Next, given an effective representative for V which does not go through (z, w), we show how to obtain rather simply an explicit lower bound for h_V, in terms of an elementary norm on the space of sections of $\mathcal{O}(V)$; this allows us to avoid arithmetic algebraic geometry altogether. A simple explicit use of derivations along C shows how the lower bound has to be modified in case the section vanishes at (z, w).

The third step consists in finding a section of $\mathcal{O}(V)$ with small norm. This is done by an application of Siegel's Lemma, perhaps the most standard tool in transcendence theory.

The fourth step is a direct application of Roth's Lemma to control the vanishing of the small section, by proving that the Roth index at (z, w) of a section of $\mathcal{O}(V)$ with small norm is also small.

The fifth, and final step, combines the upper and lower bounds for $h_V(z, w)$ and the bound for the index, concluding the proof of the Mordell conjecture.

5. - Mumford's method and an upper bound for the height

Let P be the divisor of degree 1 on C introduced in section 3. We embed the curve C of genus $g \geq 2$ into its Jacobian $A = \mathrm{Pic}_0(C)$ by associating to a point $Q \in C$ the divisor class $x = cl(Q - P)$, which has degree 0, we identify C with a subvariety of A and denote by $j : C \to A$ the inclusion. In order to distinguish between points of $j(C)$ and divisors of C we denote divisors by capital letters and points on A by small letters.

Given C and A as above the Θ divisor on A is the sum of $g - 1$ copies of C, therefore

$$\Theta = \{x \in A \mid x = x_1 + \cdots + x_{g-1},\ x_i \in C\}.$$

Let $\theta = cl(\Theta)$ be the associated divisor class. Then

$$< x, y > = \hat{h}_\theta(x + y) - \hat{h}_\theta(x) - \hat{h}_\theta(y)$$

is a symmetric bilinear form, the Néron form on A. The associated quadratic form $\frac{1}{2}|x|^2 = \frac{1}{2} < x, x >$ is the quadratic part of the height $\hat{h}_\theta$ and therefore $|x|^2 = \hat{h}_\theta(x) + \hat{h}_{\theta^-}(x)$. Since θ is an ample class it follows that the sets in $A(k)$ with bounded height are finite, which also implies that $|x|^2$ is a positive definite quadratic form on $A(k)/\mathrm{tors}$.

We can also view $<x,y>$ as a height on $A \times A$. Consider the maps $s_1(x,y) = x$, $s_2(x,y) = y$, $s_{12}(x,y) = x+y$ from $A \times A$ to A. Let $\delta = s_1^*\theta + s_2^*\theta - s_{12}^*\theta$, which is a divisor class on $A \times A$; then

$$\hat{h}_\delta(x,y) = - <x,y> .$$

LEMMA 1. *We have* $j^*(\theta^-) = g\,cl(P)$ *and* $(j \times j)^*\delta = cl(\Delta')$.

PROOF. Except for the notation, this is in Mumford [M]; see also Lang [L], Ch. 5, Theorem 5.8 and Proposition 5.6.

Now we can obtain an upper bound for h_V, $V = d_1 P \times C + d_2 C \times P + d\Delta'$.

LEMMA 2. *There are positive constants* c_1 *and* c_2, *depending on* C, P, ϕ_{NP}, ϕ_B *such that if* $(z,w) \in (C \times C)(\bar{k})$ *we have*

$$h_V(z,w) \le \frac{d_1}{2g}|z|^2 + \frac{d_2}{2g}|w|^2 - d<z,w>$$
$$+ c_1(|d_1||z| + |d_2||w|) + c_2(|d_1| + |d_2| + |d|).$$

PROOF. By Weil's Theorem, (ii) applied to the embedding $j : C \to J$, the equations between divisor classes of Lemma 1 translate into the approximate equality of heights

$$h_{NP} = \frac{N}{g}(\hat{h}_{\theta^-}) \circ j + O(1)$$

and

$$h_{\Delta'} = \hat{h}_\delta + O(1).$$

Also $h_{\Delta'}$ can be defined by

$$h_B(z,w) = \frac{s}{N}h_{NP}(z) + \frac{s}{N}h_{NP}(w) - h_{\Delta'}(z,w),$$

which combined with the definition of h_V gives

$$h_V(z,w) = \frac{d_1}{g}\hat{h}_{\theta^-}(z) + \frac{d_2}{g}\hat{h}_{\theta^-}(w) - d<z,w> + O(|d_1| + |d_2| + |d|).$$

Now $\hat{h}_{\theta^-}(z) = \frac{1}{2}|z|^2 - \frac{1}{2}\hat{h}_{\theta-\theta^-}(z)$ and $\theta - \theta^-$ is an odd class, so that $\hat{h}_{\theta-\theta^-}(z)$ is linear. In particular, there is a bound $\hat{h}_{\theta-\theta^-}(z) = O(|z|)$, and the statement of Lemma 2 follows.

We also note that the constants c_1, c_2 are effectively computable. The only thing we need worry about are the $O(1)$ terms arising in comparing heights for linearly equivalent divisors in a same divisor class. If we look at Weil's proof of this fact, we see that it depends on Hilbert's Nullstellensatz over the field k, for which various effective versions are available.

As a corollary, we obtain

MUMFORD'S THEOREM. *There is a positive constant* c_3 *such that*

$$\frac{1}{2g}\,|z|^2 + \frac{1}{2g}\,|w|^2 - <z,w> \geq -c_3(1+|z|+|w|).$$

for every $(z,w) \in (C \times C)(\bar{k}) - \Delta$.

PROOF. We apply Lemma 2 with $d_1 = 1$, $d_2 = 1$, $d = 1$ and note that $V = \Delta$ is effective, hence h_V is bounded from below outside Δ by Weil's Theorem.

6. - A lower bound for the height

Let $V = \delta_1(NP) \times C + \delta_2 C \times (NP) - dB$ be as in the preceding section. According to the results in section 3, we may assume that δ_1, δ_2, d are so large that global sections of $\mathcal{O}(\delta_1(NP)\times C+\delta_2 C\times(NP))$ and $\mathcal{O}(dB)$ are the restriction to $C \times C$ of global sections of $\mathcal{O}_{\mathbf{P}^n\times\mathbf{P}^n}(\delta_1,\delta_2)$ and $\mathcal{O}_{\mathbf{P}^m}(d)$.

Let s be a global section of $\mathcal{O}(V)$ and let s_1 be any global section of $\mathcal{O}(dB)$. Then $s_1 s$ is a global section of $\mathcal{O}(\delta_1(NP) \times C + \delta_2 C \times (NP))$; by our preceding remarks, s_1 is the restriction to $C \times C$ of a homogeneous form $G(y)$ of degree d in the variables $y = (y_0, \cdots, y_m)$ and similarly $s_1 s$ is the restriction to $C \times C$ of a bihomogeneous form $F(x,x')$ of bidegree (δ_1, δ_2) in the variables $x = (x_0, \cdots, x_n)$ and $x' = (x'_0, \cdots, x'_n)$.

We do this choosing for s_1 the sections induced by y_i^d, $i = 0, \cdots, m$ and obtain forms $F_i(x,x')$ of bidegree (δ_1,δ_2) such that

$$s = (F_i(x,x')/y_i^d)|_{C\times C}$$

for $i = 0, \cdots, m$.

Conversely, assume that $F_i(x,x')$, $i = 0, \cdots, m$ are $m+1$ forms of bidegree (δ_1, δ_2) such that

$$F_i(x,x')/y_i^d = F_j(x,x')/y_j^d$$

on $C\times C$, for every i, j. Then $s = (F_i(x,x')/y_i^d)|_{C\times C}$ is independent of i and is a global section of $\mathcal{O}(V)$. In fact, $s = (F_i(x,x')/y_i^d)|_{C\times C}$ is regular except possibly at the divisor $y_i = 0$, and the y_i's have no common zero, which shows that s is regular everywhere.

The following result gives a lower bound for $h_V(z,w)$ if we have a global section of $\mathcal{O}(V)$ which is non-zero at (z,w). We state it only for illustrating a basic procedure.

LEMMA 3. *Let* s *be a global section of* $\mathcal{O}(V)$ *which does not vanish at* (z,w) *and let* $\mathcal{F} = \{F_i\}$ *be an associated collection of forms with* $s = F_i(x,x')/y_i^d|_{C\times C}$, $i = 0, \cdots, m$.

Then we have

$$h_V(z,w) \geq -h(\mathcal{F}) - n\,\log((\delta_1+n)(\delta_2+n)).$$

PROOF. Let $x = \phi_{NP}(z)$, $x' = \phi_{NP}(w)$, $y = \phi_B(z,w)$. By our definitions,

$$h_V(z,w) = \delta_1 \sum_v \max_j \log |x_j|_v + \delta_2 \sum_v \max_{j'} \log |x'_{j'}|_v - d \sum_v \max_i \log |y_i|_v$$

which we rewrite as

$$h_V(z,w) = \sum_v \min_i \max_{j,j'} \log \left| \frac{x_j^{\delta_1} x_{j'}'^{\delta_2}}{y_i^d} \right|_v .$$

Since $s(z,w) \neq 0$, we have $\sum_v \log |s(z,w)|_v = 0$ by the product formula. This gives

$$\begin{aligned} h_V(z,w) &= \sum_v \min_i \max_{j,j'} \log \left| \frac{x_j^{\delta_1} x_{j'}'^{\delta_2}}{y_i^d s(z,w)} \right|_v \\ &= \sum_v \min_i \max_{j,j'} \log \left| \frac{x_j^{\delta_1} x_{j'}'^{\delta_2}}{F_i(x,x')} \right|_v \\ &= -\sum_v \max_i \min_{j,j'} \log \left| F_i \left(\frac{x}{x_j}, \frac{x'}{x'_{j'}} \right) \right|_v . \end{aligned}$$

Now any choice of j, j' at any place v will yield a lower bound for h_V. The careful choice j_v, j'_v for j, j' at the place v is that determined by $|x_{j_v}|_v = \max_j |x_j|_v$ and $|x'_{j'_v}|_v = \max_{j'} |x'_{j'}|_v$, so that $|x_j/x_{j_v}|_v \leq 1$ and $|x'_{j'}/x'_{j'_v}|_v \leq 1$.

It follows that

$$\min_{j,j'} \log \left| F_i \left(\frac{x}{x_j}, \frac{x'}{x'_{j'}} \right) \right|_v \leq \max \log |\text{coefficients of } F_i|_v$$

at all finite places, and for the infinite places there is a similar estimate, with an additional contribution of $\log |\text{number of coefficients of } F_i|_v$. The number of monomials of bidegree (δ_1, δ_2) is $\binom{\delta_1+n}{n}\binom{\delta_2+n}{n} \leq (\delta_1+n)^n(\delta_2+n)^n$, and our result follows.

In practice, we must allow the possibility that s vanishes at (z,w) and the usual procedure is to use derivations to obtain a quantity which does not vanish there. Although the idea is simple, there are complications arising from the fact that we need some uniformity with respect to the order of derivatives.

Let ς be a local uniformizing parameter for C at z, that is a rational function on C with a simple zero at z and let $\xi_{ij} = (x_i/x_j)|_C$, considered as a function of ς. In similar way, we denote by ς', $\xi'_{ij} \cdots$ the similar construction for an uniformizer at w and we write $\eta_{ij} = (y_i/y_j)|_{C\times C}$, viewed again as a function of (ς, ς'). We also write ξ_j for the vector with components ξ_{ij}, $i = 0, \cdots, n$ and similarly for $\xi'_{j'}$.

Let $s = (F_i/y_i^d)|_{C\times C}$ be a global section of $\mathcal{O}(V)$. Suppose that ξ_{j0}, $\xi'_{j'0}$, η_{i0} do not have zeros or poles at z, w, (z,w). Then

$$g = s/(x_0^{\delta_1} x_0'^{\delta_2} y_0^{-d}) = F_i(\xi_0, \xi_0')/\eta_{i0}^d$$

is a rational function on $C \times C$ regular in a neighborhood of (z,w), which we can differentiate with respect to the uniformizers ς, ς'. We abbreviate

$$\partial_i = \frac{1}{i!}\left(\frac{d}{d\varsigma}\right)^i, \quad \partial_i' = \frac{1}{i!}\left(\frac{d}{d\varsigma'}\right)^i$$

and pick any pair (i_1^*, i_2^*) such that

$$b = \partial_{i_1^*}\partial'_{i_2^*} g(z,w) \neq 0$$

and which is admissible in the sense that

$$\partial_{i_1}\partial'_{i_2} g(z,w) = 0$$

whenever $i_1 \le i_1^*$, $i_2 \le i_2^*$ and $(i_1, i_2) \neq (i_1^*, i_2^*)$. By the admissibility property of (i_1^*, i_2^*) we see that

$$b = \eta_{i0}^{-d}(0,0)\xi_{j0}^{\delta_1}(0)\xi_{j'0}'^{\delta_2}(0)\partial_{i_1^*}\partial'_{i_2^*} F_i(\xi_j, \xi'_{j'})(0,0) \neq 0.$$

LEMMA 4. *Let s and $\mathcal{F} = \{F_i\}$ be as before.*
Then we have

$$\begin{aligned} h_V(z,w) \ge &- h(\mathcal{F}) - n \log((\delta_1+n)(\delta_2+n)) \\ &- \sum_v \max\left(\sum_\lambda \max_\nu \log|\partial_{i_\lambda}\xi_{\nu j}(0)|_v\right) \\ &- \sum_v \max\left(\sum_\lambda \max_\nu \log|\partial_{i'_\lambda}\xi'_{\nu j'}(0)|_v\right) \\ &- (\delta_1 + \delta_2 + i_1^* + i_2^*), \end{aligned}$$

where the last two maxima run over all partitions $\{i_\lambda\}$ and $\{i'_\lambda\}$ of i_1^* and i_2^* and where $j = j_v$, $j' = j'_v$ are such that $|\xi_{\nu j}(0)|_v \le 1$ and $|\xi'_{\nu j'}(0)|_v \le 1$ for every ν.

PROOF. We proceed as in the proof of Lemma 3.

$$\begin{aligned} h_V(z,w) &= \sum_v \min_i \max_{j,j'} \log\left|\frac{x_j^{\delta_1} x_{j'}'^{\delta_2}}{y_i^d}(z,w)\right|_v \\ &= \sum_v \min_i \max_{j,j'} \log\left|\frac{x_j^{\delta_1} x_{j'}'^{\delta_2}}{y_i^d b}(z,w)\right|_v \end{aligned}$$

$$= -\sum_v \max_i \min_{j,j'} \log \left| \partial_{i_1^*}\partial'_{i_2^*}F_i(\xi_j, \xi'_{j'})(0,0)\,\frac{y_0^d}{x_0^{\delta_1}x_0'^{\delta_2}}\,(z,w)\right|_v$$

$$= -\sum_v \max_i \min_{j,j'} \log \left| \partial_{i_1^*}\partial'_{i_2^*}F_i(\xi_j, \xi'_{j'})(0,0)\right|_v$$

$$\geq -h(\mathcal{F}) - n\,\log((\delta_1+n)(\delta_2+n)) - \sum_v \min_j \max_l \log|\partial_{i_1^*}(\xi_j^l)(0)|_v$$

$$- \sum_v \min_{j'} \max_{l'} \log|\partial'_{i_2^*}(\xi_{j'}'^{l'})(0)|_v$$

where ξ_j^l and $\xi_{j'}'^{l'}$ run over all monomials of degrees δ_1 and δ_2.

Let us consider $\log|\partial_{i_1^*}(\xi_j^l)(0)|_v$. By Leibnitz's rule, we have

$$\partial_{i_1^*}(\xi_j^l) = \sum \prod_{\nu=0}^{n}\prod_{\mu=1}^{l_\nu} \partial_{i_{\mu\nu}}\xi_{\nu j}$$

where $\sum_{\mu\nu} i_{\mu\nu} = i_1^*$. Since the total number of pairs $\mu\nu$ is δ_1, the number of possibilities for $i_{\mu\nu}$ does not exceed $\binom{\delta_1+i_1^*}{i_1^*} \leq 2^{\delta_1+i_1^*}$. Also, as in the proof of Lemma 3, we choose j such that $|\xi_{\nu j}|_v \leq 1$ for every ν. This gives the bound

$$\log|\partial_{i_1^*}(\xi_j^l)(0)|_v \leq \max\left(\sum_\lambda \max_\nu \log|\partial_{i_\lambda}\xi_{\nu j}(0)|_v\right) + \epsilon_v\,\log\left|2^{\delta_1+i_1^*}\right|_v$$

with $\epsilon_v = 1$ if $v|\infty$ and $\epsilon_v = 0$ otherwise, and where $\{i_\lambda\}$ runs over all partitions of i_1^*.

An entirely analogous estimate holds for the sum involving $\xi'_{j'}$ and Lemma 4 follows.

7. - Estimates for derivatives

In this section we consider the problem of estimating the coefficients of Taylor series expansions of algebraic functions of one variable. Let C, z, ς be a non-singular algebraic curve, a point on it and an uniformizer there. The degree of the rational map $\varsigma : C \to \mathbf{P}^1$ is the degree of the divisor of zeros of the function ς. Let k be a field of definition for C, z, ς; then the function field $k(C)$ is a finite extension of $k(\varsigma)$ of the same degree. It follows that for every rational function $\xi \in k(C)$ there is a polynomial $p(\xi,\varsigma)$ with coefficients in k such that $p(\xi,\varsigma) = 0$.

The following result is a local version of Eisenstein's theorem on denominators of coefficients of Taylor expansions of algebraic functions. The statement and proof follow a suggestion of A. Granville.

LEMMA 5. *Let $p(\xi,\varsigma)$ be a polynomial in two variables with coefficients in k, and let $\xi = \xi(\varsigma)$ be an algebraic function of ς such that $p(\xi,\varsigma) = 0$. Suppose* *further that the discriminant $D(\varsigma)$ of p with respect to ξ does not vanish at $\varsigma = 0$ and that $\xi(0) \in k$.*

Let v be a finite place of k and suppose that $|\xi(0)|_v \leq 1$. Then we have for every $l \geq 1$:

$$|\partial_l \xi(0)|_v \leq \left(\frac{|p|_v}{|p_\xi(\xi(0), 0)|_v} \right)^{2l-1},$$

with $p_\xi = \partial p / \partial \xi$ and $|p|_v = \max |\text{coefficients of } p|_v$.

If instead v is an infinite place, we have with the same hypotheses:

$$|\partial_l \xi(0)|_v \leq |2 \deg p|_v^{5l} \left(\frac{|p|_v}{|p_\xi(\xi(0), 0)|_v} \right)^{2l-1}.$$

PROOF. Suppose first that v is a finite place. By Leibnitz's formula,

$$\partial_l p = \sum_{ij} \sum p_{ij} \partial_{l_0} \varsigma^j \partial_{l_1} \xi \cdots \partial_{l_i} \xi,$$

where the inner sum is over all solutions of $l_0 + \cdots + l_i = l$; the p_{ij}'s are the coefficients of the polynomial p. The sum of the terms with $l_\lambda = l$ and $\lambda \neq 0$ is simply $p_\xi \partial_l \xi$; since $\partial_l p$ is identically 0 and v is ultrametric we get

$$|p_\xi(\xi(0), 0)|_v |\partial_l \xi(0)|_v \leq |p|_v \max |\partial_{l_1} \xi(0)|_v \cdots |\partial_{l_i} \xi(0)|_v,$$

where $|p|_v$ is the maximum of the coefficients of p and max runs over $l_1 + \cdots + l_i \leq l$ with each $l_\lambda < l$. On the other hand, the hypothesis about the discriminant implies that $p_\xi(\xi(0), 0) \neq 0$ and we also have $|\xi(0)|_v \leq 1$. This means that in the last displayed inequality we need only consider products in which each l_λ is at least 1; noting that $|p_\xi(\xi(0), 0)|_v \leq |p|_v$ we get Lemma 5 in case v is a finite place.

We treat the case in which v is infinite in a different fashion. Let us abbreviate $\xi^{(l)} = \left(\frac{d}{d\varsigma} \right)^l \xi$. By induction on l we establish that there is a polynomial $q_l(\xi, \varsigma)$ such that $q_l + (p_\xi)^{2l-1} \xi^{(l)} = 0$ for $l \geq 1$, where $q_1 = p_\varsigma$ and

$$q_{l+1} = (q_l)_\varsigma p_\xi^2 - (q_l)_\xi p_\xi p_\varsigma + (2l-1) q_l (p_{\xi\xi} p_\varsigma - p_{\xi\varsigma} p_\xi).$$

If $d = \deg p$ then q_l has degree at most $(2l-1)(d-1)$. By induction again, we estimate the height $|q_l|_v$ as

$$|q_{l+1}|_v \leq |(2l-1) 16 d^5|_v |p|_v^2 |q_l|_v,$$

while clearly $|q_1|_v \leq |d|_v |p|_v$. This yields $|q_l|_v \leq |(l-1)! 32^{l-1} d^{5l-4}|_v |p|_v^{2l-1}$ and the required estimate for $\partial_l \xi(0) = \xi^{(l)}(0)/l!$ follows easily from $\xi^{(l)} = -q_l / p_\xi^{2l-1}$, completing the proof of Lemma 5.

Let us consider the embedding $\phi_{NP} : C \to \mathbf{P}^n$ and the functions $\xi_{\nu j} = (x_\nu / x_j)|_C$ associated to it. For each νj, $\nu \neq j$, $\nu j \neq 10$, $\nu j \neq 01$ we have a polynomial relation $g_{\nu j}(\xi_{\nu j}, \xi_{10}) = 0$, of degree at most $2N^2$, and by choosing ϕ_{NP} generic we may ensure that the plane curve $g_{\nu j} = 0$ is a birational model of C; in other words, we have $k(C) = k(\xi_{\nu j}, \xi_{10})$. Since we are in characteristic zero, this implies firstly that for $z \in C$ the function $\varsigma = \xi_{10} - x_1(z)/x_0(z)$ is a local uniformizer at z on C, except possibly for finitely many points $z \in C(\overline{k})$, and secondly it implies that the discriminant $D_{\nu j}(\xi_{10})$ of $g_{\nu j}$ with respect to $\xi_{\nu j}$ is not identically 0.

LEMMA 6. *There is a constant c_4 depending only on C and ϕ_{NP} and a finite subset Z of $C(\overline{k})$ such that the following holds.*

Let s, $\mathcal{F} = \{F_i\}$ be as in Lemma 4 and let (i_1^*, i_2^*) be admissible for s at a point $(z, w) \in (C \times C)(\overline{k})$ with z, $w \notin Z$. Then we have

$$h_V(z, w) \geq -h(\mathcal{F}) - c_4(i_1^*|z|^2 + i_2^*|w|^2 + i_1^* + i_2^*) - (\delta_1 + \delta_2) - o(\delta_1 + \delta_2).$$

PROOF. For every $z \in C$ with $x_0(z) \neq 0$ the function $\varsigma = \xi_{10} - x_1(z)/x_0(z)$ vanishes at z and is an uniformizer at z except possibly for z in a finite set Z; in a similar way, $\varsigma' = \xi'_{10} - x'_1(w)/x'_0(w)$ is an uniformizer at w, except possibly for $w \in Z$.

Now we proceed with the calculation of an upper bound for the quantity

$$\sum_v \max\left(\sum_\lambda \max_\nu \log|\partial_{i_\lambda} \xi_{\nu j}(0)|_v\right)$$

which appears in Lemma 4. We apply Lemma 5 to the polynomials $p_{\nu j}(\xi, \varsigma) = g_{\nu j}(\xi, \varsigma + x_1(z)/x_0(z))$, choosing j such that $|\xi_{\nu j}(0)|_v \leq 1$ for every ν. We obtain, writing for simplicity $x_i = x_i(z)$, $g_\xi = \partial g/\partial \xi$, $d_0 = \max_{\nu j} \deg p_{\nu j}$:

$$\begin{aligned}
&\sum_v \max\left(\sum_\lambda \max_\nu \log|\partial_{i_\lambda} \xi_{\nu j}(0)|_v\right)\\
&\leq 2\sum_v \max \sum_\lambda \max_\nu i_\lambda \log\left(\frac{|p_{\nu j}|_v}{|(g_{\nu j})_\xi(x_\nu/x_j, x_1/x_0)|_v}\right)\\
&+ \sum_{v \text{ infinite}} \sum_\lambda 5i_\lambda \log|2d_0|_v\\
&\leq 2i_1^* \sum_{\nu j} \sum_v (\log^+|\text{coefficients of } p_{\nu j}|_v + \log^+|1/(g_{\nu j})_\xi(x_\nu/x_j, x_1/x_0)|_v)\\
&+ 5i_1^* \log(2d_0)\\
&\leq c_5 i_1^* h(x_1/x_0) + c_6 i_1^* + 2i_1^* \sum_{\nu j} h(1/(g_{\nu j})_\xi(x_\nu/x_j, x_1/x_0))\\
&= c_5 i_1^* h(x_1/x_0) + c_6 i_1^* + 2i_1^* \sum_{\nu j} h((g_{\nu j})_\xi(x_\nu/x_j, x_1/x_0))\\
&\leq c_7 i_1^* h_{NP}(z) + c_7 i_1^* \leq c_8 i_1^* |z|^2 + c_9 i_1^*.
\end{aligned}$$

The same considerations apply to the sum with $'$, and Lemma 6 follows.

8. - Application of Siegel's Lemma

In this section, we prove the existence of small sections s of $\mathcal{O}(V)$, V a Vojta divisor, by finding representatives $\{F_i\}$ with small height $h(\mathcal{F})$.

LEMMA 7. *Let $\gamma > 0$ and let d_1, d_2, d be sufficiently large integers such that $d_1d_2 - gd^2 > \gamma d_1 d_2$. Let V be the Vojta divisor determined by d_1, d_2, d.*

Then there exists a non-trivial global section s of $\mathcal{O}(V)$ admitting a representative $\mathcal{F} = \{F_i\}$ with

$$h(\mathcal{F}) \leq c_{10}(d_1 + d_2)/\gamma + o(d_1 + d_2).$$

PROOF. We have seen in section 6 that global sections s of $\mathcal{O}(V)$ can be written in the form

$$s = (F_i(x, x')/y_i^d)|_{C\times C}, \quad i = 0, \cdots, m$$

where the F_i's are bihomogeneous polynomials of bidegree (δ_1, δ_2) in the projective co-ordinates $x = (x_0, \cdots, x_n)$ and $x' = (x_0', \cdots, x_n')$ of the embedding $\phi_{NP} \times \phi_{NP} : C \times C \to \mathbf{P}^n \times \mathbf{P}^n$, and where $y = (y_0, \cdots, y_m)$ are projective co-ordinates of the embedding $\phi_B : C \times C \to \mathbf{P}^m$.

Conversely, if we have forms F_i, $i = 0, \cdots, m$ such that

$$(F_i(x, x')/y_i^d)|_{C\times C} = (F_j(x, x')/y_j^d)|_{C\times C}$$

for all i, j then $s = (F_i/y_i^d)|_{C\times C}$ is a global section of $\mathcal{O}(V)$.

Let us abbreviate $\xi_j = (x_j/x_0)|_{C\times C}$ and $\eta_i = (y_i/y_0)|_{C\times C}$. We want to write the basic system as a system of linear equations in the coefficients of the polynomials F_i. It is simpler to work with affine co-ordinates, for example $\xi_j = x_j/x_0$, $\xi_{j'}' = x_{j'}'/x_0'$. By choosing ϕ_{NP} sufficiently generic, we may assume that for every $a \neq 0$ the image of C in the embedding $\phi_{NP} : C \to \mathbf{P}^n$ does not intersect the linear subspace of codimension 2 of $\mathbf{P}^n$ given by $x_0 = 0$, $x_a = 0$; this means that the projection $\pi_a : \mathbf{P}^n \to \mathbf{P}$ given by $\pi_a(x_0, \cdots, x_n) = (x_0, x_a)$ is defined everywhere on C and has degree N. Furthermore, we may suppose that for $j \neq 0, a$, the projection $\phi : \mathbf{P}^n \to \mathbf{P}^2$ given by $\phi(x_0, \cdots, x_n) = (x_0, x_a, x_j)$ is a birational morphism from C to its image in $\mathbf{P}^2$, which is a plane curve of degree N, and ξ_j is integral over $k[\xi_a]$, of precise degree N. The projection $\pi_{ab} : C \times C \to \mathbf{P}^1 \times \mathbf{P}^1$ given by $\pi_{ab}(x, x') = (x_0, x_a, x_0', x_b')$ is a finite morphism of degree N^2 and a product map of two finite morphisms of degree N.

Since $\phi_{NP} \times \phi_{NP}$ is a birational embedding, we have $k(C \times C) = k(\xi_1, \xi_2, \xi_1', \xi_2')$ and therefore we can write

$$\eta_i = P_i(\xi_1, \xi_2, \xi_1', \xi_2')/Q_i(\xi_1, \xi_2, \xi_1', \xi_2')$$

for some polynomials P_i, Q_i with coefficients in k. Now the system of equations characterizing global sections of $\mathcal{O}(V)$ becomes

$$((P_iQ_j)^d F_j)|_{C\times C} = ((P_jQ_i)^d F_i)|_{C\times C}$$

for every i, j.

The idea of the proof of Lemma 7 is to view the preceding system of equations as a linear system for the coefficients of the F_i's and solve it by the well-known Siegel Lemma. The height of the coefficients of $(P_iQ_j)^d$ is bounded by $c_{11}d$, the space of $F_i|_{C\times C}$ has dimension asymptotic to $(m+1)N^2\delta_1\delta_2$ and the space of solutions to the system is $H^0(\mathcal{O}(V))$, which has dimension at least $d_1d_2 - gd^2 - O(d_1 + d_2)$. Thus we expect a non-trivial solution with $h(\mathcal{F}) < (1 + o(1))(c_{11}d)(m + 1)N^2\delta_1\delta_2/(d_1d_2 - gd^2)$, according to the philosophy of Siegel's Lemma. In practice, it is convenient to work in a space of forms F of restricted type, in order to be able to write the conditions of restriction to $C \times C$ in a simple way, without increasing the height of the coefficients of the associated linear system.

Let $R = k[\xi_1, \xi_2, \xi_1', \xi_2']$ be the co-ordinate ring of $C \times C$ over the open set $x_0 \neq 0$, $x_0' \neq 0$. Then the monomials $\xi_2^i\xi_2'^{i'}$, $0 \le i, i' \le N - 1$ form a basis $\{b_\nu\}$, $\nu = 1, \cdots, N^2$ of $k(C \times C)$ over $k(\xi_1, \xi_1')$. As noted before, ξ_2 is integral over $k[\xi_1]$, and we have an equation on C:

$$\xi_2^N = A_0(\xi_1) + A_1(\xi_1)\xi_2 + \cdots + A_{N-1}(\xi_1)\xi_2^{N-1}$$

with suitable polynomials A_i with $\deg A_i \le N-i$. A similar result of course holds for ξ_2'. Now an easy induction shows that for every monomial $\xi^l\xi'^{l'} = \xi_1^{l_1}\xi_2^{l_2}\xi_1'^{l_1'}\xi_2'^{l_2'}$ we have on $C \times C$

$$\xi_1^{l_1}\xi_2^{l_2}\xi_1'^{l_1'}\xi_2'^{l_2'} = \sum q_{ll'\nu}(\xi_1, \xi_1')b_\nu$$

where the polynomials q satisfy the bound

$$\deg_\xi q_{ll'\nu} \le |l|, \quad \deg_{\xi'} q_{ll'\nu} \le |l'|, \quad h(q_{ll'\nu}) \le c_{12}(|l| + |l'|) + O(1),$$

and where we have abbreviated $|l| = l_1 + l_2$.

We proceed to rewrite our basic system $(P_iQ_j)^dF_j|_{C\times C} = (P_jQ_i)^dF_i|_{C\times C}$. Let U_0 be the k-vector space of functions $F_i(\xi, \xi')|_{C\times C}$. By the Riemann-Roch theorem, it has dimension

$$\dim U_0 = (m + 1)(N^2\delta_1\delta_2 - N(g - 1)(\delta_1 + \delta_2) + (g - 1)^2) \le (m + 1)N^2\delta_1\delta_2$$

if δ_1, δ_2 are sufficiently large, which we suppose. The space U_0 contains the subspace U_1 generated by F_i's which are linear combinations of monomials of type

$$\xi_1^l\xi_1'^{l'}b_\nu, \quad l \le \delta_1 - \deg_{\xi_2} b_\nu, \quad l' \le \delta_2 - \deg_{\xi_2'} b_\nu,$$

with $\{b_\nu\}$ the preceding basis of $k(C \times C)$ over $k(\xi, \xi_1')$, and U_1 has dimension

$$\dim U_1 = (m+1)\left(N^2\delta_1\delta_2 - N^2\frac{(N-3)}{2}(\delta_1+\delta_2) + \left(\frac{N(N-3)}{2}\right)^2\right),$$

because these monomials are linearly independent on $C \times C$. Now the space U_2 of solutions to our basic system is isomorphic to $H^0(C \times C, \mathcal{O}(V))$ and therefore has dimension at least

$$\dim U_2 \geq d_1d_2 - gd^2 - O(d_1+d_2+d) \geq \gamma d_1 d_2 - O(\delta_1+\delta_2),$$

again by the Riemann-Roch theorem. It follows that

$$\dim U_2 \cap U_1 \geq \dim U_2 - \operatorname{codim} U_1 \geq \gamma d_1 d_2 - O(\delta_1+\delta_2).$$

We can write

$$b_\lambda b_\mu = \sum g_{\lambda\mu\nu}(\xi_1, \xi_1')b_\nu$$

with $g_{\lambda\mu\nu} \in k(\xi_1, \xi_1')$. Let r be a common denominator for all $g_{\lambda\mu\nu}$'s. Then we see that for $\{F_i\} \in U_1$ we have

$$r((P_iQ_j)^dF_j - (P_jQ_i)^dF_i)|_{C\times C} = \sum L_{ijll'\nu}\xi_1^l\xi_1'^{l'}b_\nu$$

where the $L_{ijll'\nu}$ are linear forms in the coefficients of the F_i's, with height

$$h(L_{ijll'\nu}) \leq (h(\{P_iQ_j\}) + 2c_{12}\deg(\{P_iQ_j\}))d + o(d) \leq c_{13}d + o(d).$$

By construction, the basic system restricted to the subspace U_1 is equivalent to the linear system

$$L_{ijll'\nu} = 0$$

in the coefficients of $\{F_i\} \in U_1$, and we have shown that the linear forms $L_{ijll'\nu}$ have height bounded by $c_{13}d + o(d)$. The number of unknowns is bounded by $\dim U_1 \leq \dim U_0 \leq (m+1)N^2\delta_1\delta_2$ and the space of solutions has dimension $\dim U_2 \cap U_1 \geq \gamma d_1d_2 - O(\delta_1+\delta_2)$. Now Siegel's Lemma in its simplest version shows that there is a non-trivial solution to our basic system, with height at most

$$h(\mathcal{F}) \leq c_{13}d\,\frac{(m+1)N^2\delta_1\delta_2}{\gamma d_1 d_2} + o(\delta_1+\delta_2).$$

Also $\delta_i = (d_i + sd)/N$ and $d_1d_2 > gd^2 \geq 2d^2$ and we get Lemma 7 with for example $c_{10} = c_{13}(m+1)(s+1)^2$.

9. - Application of Roth's Lemma

In this section we obtain an upper bound for the admissible vanishing (i_1^*, i_2^*) of the section $s = F_i/y_i^d$ at (z, w).

Let $\xi_j = x_j/x_0$, $\xi'_{j'} = x'_{j'}/x'_0$, $\eta_i = y_i/y_0$ and let ξ, ξ', η be the corresponding vectors. Then

$$g = s/(x_0^{\delta_1} x_0'^{\delta_2}/y_0^d) = F_i(\xi, \xi')/\eta_i^d$$

is a rational function on $C \times C$. Taking $i = 0$ we get $g = F_0(\xi, \xi')$, which shows that g has poles only along the divisors $x_0 = 0$ and $x'_0 = 0$.

Let $\xi(0)$ and $\xi'(0)$ be the affine co-ordinates of z and w. We have

$$h_{NP}(z) = h(\xi(0)) \leq \sum_{j=1}^{n} h(\xi_j(0))$$

and therefore there is a suffix a such that $h(\xi_a(0)) \geq h_{NP}(z)/n$. In the same way, there is b such that $h(\xi'_b(0)) \geq h_{NP}(w)/n$.

As noted in the preceding section, the projection $\pi_{ab} : C \times C \to \mathbf{P}^1 \times \mathbf{P}^1$ defined by $\pi_{ab}(x, x') = (x_0, x_a, x'_0, x'_b)$ is defined everywhere and is a finite morphism of degree N^2, the product of two finite morphisms of degree N.

The function g is an element of the function field $k(C \times C)$ and its norm defines an element

$$G_0 = \mathrm{Norm}_{k(C\times C)/k(\xi_a,\xi'_b)} g = \mathrm{Norm}_{k(C\times C)/k(\xi_a,\xi'_b)}(F_i(\xi, \xi')/\eta_i^d)$$

in $k(\xi_a, \xi'_b)$. Now , as already remarked in the preceding section, ξ_j and ξ'_j are integral over $k[\xi_a]$ and $k[\xi'_b]$ by means of an equation of total degree N, and it follows that

$$G_i = \mathrm{Norm}_{k(C\times C)/k(\xi_a,\xi'_b)} F_i(\xi, \xi')$$

is a polynomial of bidegree at most $(N^2\delta_1, N^2\delta_2)$ in the variables ξ_a and ξ'_b.

The norm $\mathrm{Norm}_{k(C\times C)/k(\xi_a,\xi'_b)}\eta_i$ is an element of $k(\xi_a, \xi'_b)$, say U_i/V_i, for suitable polynomials U_i and V_i, hence

$$\mathrm{Norm}_{k(C\times C)/k(\xi_a,\xi'_b)} g = G_0 = G_i(V_i/U_i)^d$$

for $i = 0, \cdots, m$. Since $k[\xi_a, \xi'_b]$ is a factorial ring, we deduce that U_i^d divides G_i and that $T_i = G_i/U_i^d$ is a polynomial which divides G_0.

In order to compute the bidegree of T_i we proceed as follows. The divisor D_i of zeros of y_i is linearly equivalent to B, while $B \cdot (C \times P) = B \cdot (P \times C) = s$. Looking at $D_i \cap (\pi_a^{-1}(\xi_a) \times C)$ and $D_i \cap (C \times \pi_b^{-1}(\xi'_b))$ we see that U_i has precise degree Ns in both variables ξ_a and ξ'_b, for $i \neq 0$. Since G_i has bidegree at most $(N^2\delta_1, N^2\delta_2)$ in ξ_a, ξ'_b, it follows that T_i is a polynomial of degree at most $N^2\delta_1 - Nds = Nd_1$ in ξ_a and at most $N^2\delta_2 - Nds = Nd_2$ in ξ'_b. We also note that since $y_0, \cdots, y_m$ have no common zeros, there is a suffix i_0 such that

$U_{i_0}(\xi_a(0), \xi'_b(0)) \neq 0$, thus $T = T_{i_0}$ vanishes at $(\xi_a(0), \xi'_b(0))$ at least as much as the section s at (z, w).

By Lemma 7, G_0 is a polynomial with height bounded by $c_{15}N^2(d_1 + d_2)/\gamma + o(d_1 + d_2)$ and of total degree at most $N^2\delta_1 + N^2\delta_2$. Since T divides G_0, we deduce that

$$h(T) \leq c_{15}N^2(d_1 + d_2)/\gamma + 2(N^2\delta_1 + N^2\delta_2) + o(d_1 + d_2).$$

We can see this last point using the following result.

PROPOSITION. *Let $P_1, \cdots, P_k$ be polynomials over $\overline{\mathbf{Q}}$ of degrees at most $D_1, \cdots, D_k$, in m variables. Then we have*

$$\sum_{i=1}^{k} h(P_i) \leq h(P_1 \cdots P_k) + m(D_1 + \cdots + D_k).$$

PROOF. Let $h_v(P) = \max \log|\text{coefficients of } P|_v$. By Gauss' Lemma we know that $h_v(P_1P_2) = h_v(P_1)+h_v(P_2)$ at all finite places. Thus it suffices to prove the result for polynomials with complex coefficients and the ordinary absolute value. In this form it is a result of Gelfond [G], Ch. 3, Lemma II, p. 135. See also [L], Ch. 3, Proposition 2.12.

We have shown that the norm of the section s yields a polynomial T in ξ_a, ξ'_b of bidegree at most (Nd_1, Nd_2), height at most

$$h(T) \leq c_{15}N^2(d_1 + d_2)/\gamma + 2N(d_1 + d_2 + 2sd) + o(d_1 + d_2),$$

and vanishing at $(\xi_a(0), \xi'_b(0))$ at least as much as the section s at (z, w).

In what follows we abbreviate $I = (i_1, \cdots, i_m)$ and

$$\partial_I = \frac{1}{i_1! \cdots i_m!} \left(\frac{\partial}{\partial \varsigma_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial \varsigma_m}\right)^{i_m}.$$

We have

ROTH'S LEMMA. *Let $P(\varsigma_1, \cdots, \varsigma_m)$ be a polynomial in m variables, of degree at most r_μ in ς_μ, with algebraic coefficients and not identically* 0. *Let $\varsigma^0 = (\varsigma_1^0, \cdots, \varsigma_m^0)$ be an algebraic point and let*

$$Z = \min_{\mu} r_\mu h(\varsigma_\mu^0).$$

Suppose also that $\epsilon > 0$ is such that

$$r_{\mu+1}/r_\mu \leq \epsilon^{2^{m-1}}, \quad \mu = 1, \cdots m - 1$$

and

$$Z \geq \epsilon^{-2^{m-1}}(h(P) + 2mr_1).$$

Then there is $I^* = (i_1^*, \cdots, i_m^*)$ *such that*

$$\partial_{I^*} P(\varsigma^0) \neq 0$$

and

$$\sum_{\mu=1}^{m} \frac{i_\mu^*}{r_\mu} \leq 2m\epsilon.$$

PROOF. We give only a brief sketch of proof, since results of this type have become quite standard (see e.g. [L], Ch.7, Proposition 3.2); the prototype is of course in Roth [R]. The constant $2m$ appearing in the conclusion of the theorem is not optimal and its actual value has little importance in this paper, which uses only the case $m = 2$.

The proof is by induction on m.

If $m = 1$, the preceding Proposition proves that

$$i_1^* h(\varsigma_1^0) \leq h(P) + r_1,$$

which implies the result with the better bound ϵ in place of 2ϵ.

Now suppose that $m \geq 2$. One defines the index of a polynomial Q with respect to $(r_1, \cdots, r_m)$ at the point $\varsigma^0 = (\varsigma_1^0, \cdots, \varsigma_m^0)$ to be the quantity

$$\operatorname{ind}(Q) = \min_I \left(\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m} \mid \partial_I Q(\varsigma^0) \neq 0 \right).$$

It is clear that

$$\text{(i)} \quad \operatorname{ind}(\partial_I Q) \geq \operatorname{ind}(Q) \frac{i_1}{r_1} - \cdots - \frac{i_m}{r_m};$$

$$\text{(ii)} \quad \operatorname{ind}(Q_1 + Q_2) \geq \min(\operatorname{ind}(Q_1), \operatorname{ind}(Q_2));$$

$$\text{(iii)} \quad \operatorname{ind}(Q_1 Q_2) = \operatorname{ind}(Q_1) + \operatorname{ind}(Q_2).$$

Next, one writes P in the form

$$P = \sum_0^s \phi_j(\varsigma_1, \cdots, \varsigma_{m-1}) \psi_j(\varsigma_m)$$

where $s \leq r_m$ and where the ϕ_j's, and similarly the ψ_j's, are linearly independent polynomials, defined over a number field k. It then follows that the Wronskians

$$U(\varsigma_1, \cdots, \varsigma_{m-1}) = \det(\partial_{I'_\mu} \phi_j)_{\mu, j=0,\cdots,s}$$

and

$$V(\varsigma_m) = \det(\partial_\nu \psi_j)_{\nu, j=0,\cdots,s}$$

are not identically 0, for suitably chosen $I'_\mu = (i_{\mu 1}, \cdots, i_{\mu,m-1})$ with $\sum_l i_{\mu l} \le s$. We multiply the two determinants and obtain

$$\det(\partial_{I'_\mu}\partial_\nu P) = W(\varsigma_1, \cdots, \varsigma_m) = U(\varsigma_1, \cdots, \varsigma_{m-1})V(\varsigma_m).$$

The idea of the proof consists in comparing a lower bound for ind(W), obtained directly in terms of ind(P) from the definition of W as a determinant, with an upper bound, obtained by an induction assumption on the number of variables. The details are as follows.

We note first that:

(A) we may assume $r_{\mu+1}/r_\mu \le \frac{1}{2}$ for every μ, since otherwise the conclusion of Roth's Lemma is trivial, and in particular we have $r_1 + \cdots + r_m \le 2r_1$;

(B) the degrees of U and V are bounded by $((s+1)r_1, \cdots, (s+1)r_{m-1})$ and $(s+1)r_m$;

(C) the heights of U and V are bounded by

$$h(U) + h(V) \le (s+1)(h(P) + 2r_1).$$

Only (C) requires some explanation but it follows by noting that $h(U) + h(V) = h(W)$ because $UV = W$ and U, V are polynomials in different sets of variables. We estimate $h(W)$ by expanding the determinant and obtain

$$h(U) + h(V) = h(W) \le (s+1)(h(P) + (r_1 + \cdots + r_m)\log 2) + \log(s+1)!\,;$$

the required estimate follows using $r_1 + \cdots + r_m \le 2r_1$, $s \le r_m \le \frac{1}{2} r_1$, $\log(s+1) \le s$ and $2\log 2 + \frac{1}{2} < 2$.

An upper bound for ind(W) follows from the equation

$$\operatorname{ind}(W) = \operatorname{ind}(U) + \operatorname{ind}(V).$$

We also obtain a lower bound for ind(W) by expanding the determinant for W, using properties (i), (ii) and (iii) of the index to estimate from below the index of W in terms of the index of a typical term in the expansion. Since

$$\operatorname{ind}(\partial_{I'_\mu}\partial_\nu P) \ge \operatorname{ind}(P) - \frac{i_{\mu 1}}{r_1} - \cdots - \frac{i_{\mu,m-1}}{r_{m-1}} - \frac{\nu}{r_m} \ge \operatorname{ind}(P) - \frac{r_m}{r_{m-1}} - \frac{\nu}{r_m}$$

we obtain

$$\begin{aligned}\operatorname{ind}(W) &\ge \sum_{\nu=0}^{s} \max\left(\operatorname{ind}(P) - \frac{\nu}{r_m}, 0\right) - (s+1)\frac{r_m}{r_{m-1}} \\ &\ge (s+1)\min\left(\frac{1}{2}\operatorname{ind}(P), \frac{1}{2}\operatorname{ind}(P)^2\right) - (s+1)\frac{r_m}{r_{m-1}}.\end{aligned}$$

We compare this lower bound with the upper bound and get

$$\min\left(\frac{1}{2}\operatorname{ind}(P),\ \frac{1}{2}\operatorname{ind}(P)^2\right) \le \frac{1}{s+1}\operatorname{ind}(U) + \frac{1}{s+1}\operatorname{ind}(V) + \epsilon^{2^{m-1}}.$$

Suppose we have proved Roth's Lemma for polynomials in l variables, for $l < m$; note that this is so if $m = 2$, with a bound which is better by a factor $\frac{1}{2}$. Also we may assume that $\epsilon < 1$, otherwise the result is trivial. We apply the inductive assumption of Roth's Lemma to U and V using $\epsilon(U) = \epsilon^2$ and $\epsilon(V) = \epsilon^{2^{m-1}}$ in place of ϵ and $(s+1)r_\mu$ in place of r_μ. In view of the bounds obtained for $h(U)$ and $h(V)$ the hypotheses of Roth's Lemma are satisfied, and we get

$$\frac{1}{s+1}\operatorname{ind}(U) \le 2(m-1)\epsilon^2,$$

$$\frac{1}{s+1}\operatorname{ind}(V) \le \epsilon^{2^{m-1}}.$$

We substitute these two inequalities in the bound for $\operatorname{ind}(P)$, note that $\operatorname{ind}(P) \le m$, and obtain

$$\frac{1}{2m}\operatorname{ind}(P)^2 \le 2(m-1)\epsilon^2 + 2\epsilon^{2^{m-1}} \le 2m\epsilon^2.$$

This completes the induction step and the proof of Roth's Lemma.

LEMMA 8. *There is a constant* $c_{14} > 0$ *depending only on* C, ϕ_{NP}, ϕ_B *with the following property.*

Suppose that $0 < \epsilon < 1$ *and*

$$d_2/d_1 \le \epsilon^2, \quad \min(d_1 h_{NP}(z), d_2 h_{NP}(w)) \ge \frac{c_{14}}{\gamma\epsilon^2} d_1.$$

Then for every section s *of* $\mathcal{O}(V)$ *as in Lemma* 7 *there is an admissible pair* (i_1^*, i_2^*) *such that*

$$\frac{i_1^*}{d_1} + \frac{i_2^*}{d_2} \le 4N\epsilon.$$

PROOF. We apply Roth's Lemma with $m = 2$ to the polynomial $T(\xi_a, \xi_b')$ constructed before, i.e. the norm of the section s, at the point $(\xi_a(0), \xi_b'(0))$. We take $r_1 = Nd_1$, $r_2 = Nd_2$ and note that

$$h(T) + 4r_1 \le c_{15}N^2(d_1 + d_2)/\gamma + 2N(3d_1 + d_2 + 2sd) + o(d_1 + d_2) \le c_{16}d_1/\gamma.$$

We need verify the condition $d_2/d_1 \le \epsilon^2$, which we suppose, and

$$\min(Nd_1 h(\xi_a(0)), Nd_2 h(\xi_b'(0)) \ge \epsilon^{-2}c_{16}d_1/\gamma.$$

Using the inequalities $h(\xi_a(0)) \geq h_{NP}(z)/n$ and $h(\xi'_b(0)) \geq h_{NP}(w)/n$ we have already obtained, we see that the last condition is a consequence of

$$\min(d_1 h_{NP}(z), d_2 h_{NP}(w)) \geq \epsilon^{-2} c_{14} d_1/\gamma$$

with $c_{14} = nc_{16}$, and the proof of Lemma 8 is complete.

As remarked by Vojta, the lower bounds for $h(\xi_a(0))$ and $h(\xi'_b(0))$ we have used are much weaker than what is actually true. In fact, since $\pi_a : C \to \mathbf{P}^1$ is a morphism, Weil's Theorem, (ii) shows that the function $h_{NP} - h \circ \pi_a$ is bounded on $C(\overline{k})$ and therefore $h(\xi_a(0)) = h_{NP}(z) + O(1)$ for every $a \neq 0$. However, our procedure leads immediately to explicit estimates, without the need to get a bound for the quantity hidden in the $O(1)$ term.

10. - Proof of the Mordell conjecture

The proof of Mordell's conjecture is now easy.

THEOREM 1. *Let C be a projective non-singular curve of genus $g \geq 2$, defined over a number field k. Then we can effectively determine a constant $\gamma(C) > 1$ with the following property:*

For every pair of points $z, w \in C(\overline{k})$ satisfying

$$\gamma(C) \leq |z|, \quad \gamma(C)|z| \leq |w|$$

we have

$$< z, w > \leq \frac{3}{4} |z||w|.$$

PROOF. We note that since the exceptional set Z of Lemma 6 is finite and effectively determinable, we may assume that the constant $\gamma(C)$ is so large that the conditions on $|z|$, $|w|$ automatically imply $z, w \notin Z$. Moreover, we may suppose that $d, \delta_1, \delta_2, d_2$ are majorized by d_1. We also assume that $d_1 d_2 - gd^2 \geq \gamma d_1 d_2$ and that d_1, d_2, d are sufficiently large.

We combine the inequalities of Lemma 2, Lemma 6 and Lemma 7 and obtain, after absorbing several terms together by adjusting the constant c_{17}, the main inequality:

$$\begin{aligned} &- c_4(i_1^*|z|^2 + i_2^*|w|^2) - c_{17}(d_1 + i_1^* + i_2^*) \\ &- c_{17} d_1/\gamma \leq \frac{d_1}{2g} |z|^2 + \frac{d_2}{2g} |w|^2 - d < z, w > + c_1(d_1|z| + d_2|w|). \end{aligned}$$

By Lemma 8, if we also assume $h_{NP}(z) \geq c_{14}/(\gamma\epsilon^2)$, $d_2 h(w) \geq c_{14} d_1/(\gamma\epsilon^2)$ we may also ensure that

$$\frac{i_1^*}{d_1} + \frac{i_2^*}{d_2} \leq 4N\epsilon.$$

Now we choose

$$d_1 = \sqrt{g+\gamma_0}\,\frac{D}{|z|^2} + O(1), \quad d_2 = \sqrt{g+\gamma_0}\,\frac{D}{|w|^2} + O(1), \quad d = \frac{D}{|z||w|} + O(1)$$

where D is large, and eventually will go to ∞, and where the $O(1)$ terms are for small adjustments so that d_1, d_2, d, δ_1, δ_2 are all large positive integers; note also that the condition $d_1 d_2 - g d^2 \geq \gamma d_1 d_2$ is now satisfied with $\gamma = \gamma_0/(g+\gamma_0) + o(1)$, and the condition $d_2/d_1 \leq \epsilon^2$ becomes $|z|/|w| \leq \epsilon + o(1)$, where the term implicit in $o(1)$ tends to 0 as D tends to ∞.

We substitute these values in the main inequality and use the upper bounds

$$i_1^* \leq 4N\epsilon d_1, \quad i_2^* \leq 4N\epsilon d_2$$

to derive

$$\begin{aligned} &- c_4\sqrt{g+\gamma_0}\,8N\epsilon D - c_{18}\,\frac{D}{\gamma_0|z|^2} \leq \frac{\sqrt{g+\gamma_0}}{g}\,D - \frac{<z,w>}{|z||w|}\,D \\ &+ c_{19}\left(\frac{1}{|z|} + \frac{1}{|w|}\right) D + o(D), \end{aligned}$$

provided $h_{NP}(z) \geq c_{14}/(\gamma\epsilon^2)$, $d_2 h_{NP}(w) \geq c_{14} d_1/(\gamma\epsilon^2)$ and $|z|/|w| \leq \epsilon + o(1)$.

Now $h_{NP}(z) = N|z|^2/2g + O(1)$, $h_{NP}(w) = N|w|^2/2g + O(1)$, thus $d_2 h_{NP}(w)/d_1 = N|z|^2/2g + O(1)$ and the conditions on $h_{NP}(z)$, $h_{NP}(w)$ are implied by $|z| \geq c_{20}/\epsilon$ for a suitably large constant c_{20}, depending on γ. We divide the last displayed inequality by D and let D go to ∞ and we find, after rearranging terms:

$$\frac{<z,w>}{|z||w|} \leq \frac{\sqrt{g+\gamma_0}}{g} + c_4\sqrt{g+\gamma_0}\,8N\epsilon + c_{19}\left(\frac{1}{|z|} + \frac{1}{|w|}\right) + \frac{c_{18}}{\gamma_0|z|^2}$$

provided $|z| \geq c_{20}/\epsilon$ and $|z|/|w| \leq \epsilon$.

If we choose γ_0 sufficiently small and then ϵ even smaller we get Theorem 1 because $\sqrt{2}/2 < .75$.

The next result is Vojta's theorem [V].

THEOREM 2. *Let C be a projective non-singular curve of genus $g \geq 2$ defined over a number field k.*

Then $C(k)$ is a finite set. More precisely, for every finite extension K of k the points of $C(K)$ either have bounded height $\gamma(C)$, or they belong to a finite set of cardinality at most

$$|\mathrm{tors}(A(K))|\,7^{\rho}(1 + \log\gamma(C)/\log 2),$$

where ρ is the rank of the Mordell-Weil group $A(K)$ over K of the Jacobian variety of C and $\gamma(C)$ is the constant appearing in Theorem 1.

PROOF. We follow Mumford's argument. The inner product $<z,w>$ gives a structure of euclidean space to $\mathbf{R}^\rho = (A(K)/\text{tors}) \otimes \mathbf{R}$. We can cover the unit sphere $S^{\rho-1}$ with not more than 7^ρ spherical caps of sufficiently small radius so to have $<z,w>> 3/4$ for any pair of points z, w in a same spherical cap; this can be seen by covering $S^{\rho-1}$ with balls of radius $2\sin\left(\frac{1}{4}\arccos(3/4)\right)$ centered at every point of $S^{\rho-1}$, and recalling the general fact that if we cover a set B in euclidean space with a family of balls of the same radius centered at every point of B, then there is a subfamily which still covers B and such that if we reduce the radii of the balls by a factor $\frac{1}{2}$ then they become disjoint (it suffices to consider a maximal family of points on B, with mutual distance at least the radius).

Let Γ be a cone in $\mathbf{R}^\rho$ over such a spherical cap and let z, w be two points in $C(K)$ with image, modulo the torsion group of $A(K)$, in the cone Γ and with $\gamma(C) \leq |z| \leq |w|$.

Then Theorem 1 shows that we must have

$$|z| \leq |w| \leq \gamma(C)|z|.$$

Now we apply Mumford's Theorem

$$\frac{1}{2g}|z|^2 + \frac{1}{2g}|w|^2 - <z,w> \geq -c_3(1+|z|+|w|)$$

to the sequence of points $z_0 = z$, $z_1, \cdots z_m$ with $|z_i| \geq |z| \geq \gamma(C)$, arranged by increasing height. If $\gamma(C)$ is sufficiently large, as we may suppose, Mumford's Theorem implies $|z_{i+1}| \geq 2|z_i|$ for every i and therefore $|z_m| \geq 2^m|z|$. In view of the bound for $|w|$ this implies $2^m \leq \gamma(C)$ and $m \leq \log\gamma(C)/\log 2$. Thus the cone Γ contains only the finitely many points with $|z| \leq \gamma(C)$ and at most $1 + \log\gamma(C)/\log 2$ other points, completing the proof of Theorem 2.

We conclude with the remark that effective upper bounds for the rank of the Mordell-Weil group can be obtained by performing the first 2-descent, reducing the problem to Dirichlet's theorem on units in number fields. Hence Theorem 2 indeed leads to effective bounds for the number of rational points over K of the curve C.

REFERENCES

[F] G. FALTINGS, *Diophantine approximation on abelian varieties*. To appear in Annals of Math. (2) **133** (1991).

[G] A.O. GELFOND, *Transcendental and Algebraic Numbers*. (English translation by L. F. Boron.) Dover Publications, New York 1960.

[L] S. LANG, *Fundamentals of Diophantine Geometry*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo 1983.

[M] D. MUMFORD, *A remark on Mordell's conjecture.* Amer. J. Math. **87** (1965), pp. 1007-1016.

[M2] D. MUMFORD, *Algebraic Geometry I. Complex Projective Varieties.* Springer-Verlag, Berlin, Heidelberg, New York 1976.

[R] K.F. ROTH, *Rational approximations to algebraic numbers.* Mathematika **2** (1955), pp. 1-20.

[V] P. VOJTA, *Siegel's theorem in the compact case.* To appear in Annals of Math. (2) **133** (1991).

[W] A. WEIL, *Arithmetic on algebraic varieties.* Annals of Math. (2) **53** (1951), pp. 412-444.

Institute for Advanced Study
Princeton, N.J. 08540
USA