Annals of Mathematics, 133 (1991), 549–576

Diophantine approximation on abelian varieties

By Gerd Faltings

1. Introduction

The theory of diophantine approximation has been the classical method to treat diophantine equations. Let us just mention the names Thue, Siegel, Dyson, Gel'fond, Roth, Schmidt. Classically one deals with the projective line or at least a projective space, and approximates a subvariety by rational points. If there are many approximates which are too good, one obtains a contradiction as follows: Construct a polynomial f (in several variables) which vanishes to a high order at the variety we want to approximate. Then if the approximating rational points are very close to it the polynomial takes very small values in them, and must vanish because of integrality conditions. On the other hand (and this usually is the hard part) one shows that it does not vanish, which gives a contradiction. The overview [S] of W. Schmidt gives a good picture of the classical results. Compared with this we use several new ideas:

The first one is due to P. Vojta ([V]). He has discovered that on products of abelian varieties there are more line bundles than on products of projective spaces, which gives a new degree of freedom in choosing f.

The difficult results about non-vanishing have been circumvented by a new method (the product theorem). Either the assertion holds, or our points lie in a smaller subvariety to which we can apply induction. The proof of this uses methods similar to the zero-estimates (Nestorenko, Masser-Wüstholz) of transcendental number theory.

Furthermore we have made some minor improvements in technicalities. That is, we use a refined version of Siegel's lemma, and we avoid the difficult Arakelov theory in Vojta's paper.

Now let us describe our results. Assume that A is an abelian variety over a number field k. The following had been conjectured by A. Weil and also by S. Lang:

Theorem 1. *Assume $X \subset A$ is a subvariety such that X does not contain any translate of a positive dimensional abelian subvariety. Then X contains only finitely many k-rational points.*

As an example one may take for A the Jacobian of a hyperelliptic curve C of genus $g > 3$, and for X the image of $C \times C$ in A (see Example 4.5).

For the next result, recall that for a k-rational point $x \in A$ one can define its (big) height $H(x)$ (after choice of an ample line bundle on A). Furthermore if w denotes a place of k and $E \subset A$ a k-subvariety, one defines naturally the w-adic distance $d_w(x, E)$ from x to E. For an infinite place it is just the usual distance. We then can show:

THEOREM 2. *For any $\kappa > 0$, there exist only finitely many k-rational points $x \in A - E$ for which $d_w(x, E) < H(x)^{-\kappa}$.*

In fact this even holds if instead of $d_w(x, E)$ we use the maximum of the w-adic norms of $l_i(x)$, where the l_i run through a system of (local) equations defining E. As an immediate application we get a conjecture of S. Lang, namely, that for any ample divisor E the complement $A - E$ contains only finitely many integral points.

One might ask for effectivity, that is, for explicit bounds for the heights of those points whose number is assured to be finite. However, usually this is not possible in diophantine approximation, and as far as I can see everything here is ineffective beyond hope.

The proof of Theorem 1 follows Vojta's method, with a couple of technical improvements. For Theorem 2 we need some more arithmetic, namely a good definition of the height of a subvariety. This can be obtained most conveniently if one uses Arakalov theory, but only the theory in [GS1] and [GS2] (not involving analytic torsion, the index theorem etc.). The advantage of this theory is that heights are honest functions, not just defined up to bounded functions. In the next chapter we give these technicalities. After that follows the product theorem, which is our main technical innovation. Finally we first prove Theorem 1 and then Theorem 2. In both of them the product theorem plays a vital role, however, for Theorem 1 more on the geometric side and for Theorem 2 more on the arithmetic.

It is my pleasure to thank P. Deligne and G. Wüstholz for many interesting discussions. Of course I am indebted to P. Vojta whose wonderful paper [V] inspired the present work. Finally I thank the NSF and the Guggenheim foundation for their support, and the ETH Zürich for its hospitality.

2. Preliminaries

a) *Degrees*. We need estimates for the degrees of varieties which occur in various constructions. The precise estimates are not that important; we only need the fact that applied to varieties with bounded degrees these procedures lead again to varieties of bounded degrees. Let $\mathbf{P}^n$ denote the projective space over a field k, and $\mathscr{L} = \mathscr{O}(1)$ the canonical ample line bundle on it. For a subvariety $X \subset \mathbf{P}^n$ of pure dimension d, its degree $\deg(X)$ is defined as the intersection number $\deg(X) = \mathscr{L}^d \cdot X$.

Let us consider projections. Assume that $X \subset \mathbf{P}^n$ has pure dimension d, and that $L \subset \mathbf{P}^n$ is a linear subspace disjoint from X, of codimension $t + 1$. This defines a projection $\pi\colon X \subset \mathbf{P}^n - L \to \mathbf{P}^t = P$, and $\dim(\pi(X)) = d$ (π has affine, hence finite, fibres), $\deg(\pi(X)) \leq \deg(X)$. If we choose $t = d + 1$ we obtain that $\pi(X)$ is a hypersurface of degree $\leq \deg(X)$. Furthermore for any point $x \in \mathbf{P}^n - X$, we can find a π such that $\pi(x)$ lies not in $\pi(X)$; hence we derive:

Proposition 2.1. *As an algebraic set X can be defined by homogeneous polynomials of degree $\leq \deg(X)$. Moreover if X is irreducible these polynomials can be chosen to be irreducible too.*

Now let us choose $t = d$, assume that k has characteristic zero and X is irreducible. Then $\pi\colon X \to P = \mathbf{P}^d$ is finite and generically étale. We want to find a nowhere dense subvariety $Y \subset X$ of bounded degree, such that π is étale outside of Y. For this we proceed as follows:

First use a projection $\pi_0\colon X \to \pi_0(X) = V(F) \subset \mathbf{P}^{d+1}$ as before, with centre L, to obtain an irreducible polynomial F, in $d + 2$ variables and of degree $\leq \deg(X)$, which vanishes on X. If we project from a point $x \in \mathbf{P}^n - L$ such that $\pi_0(x)$ is not in $\pi_0(X)$ to obtain $\pi\colon Y \subset \mathbf{P}^n - \{x\} \to P = \mathbf{P}^{n-1}$, then some non-trivial derivative of F (in the direction of π) annihilates $\Omega_{X/P}$. As this derivative does not vanish identically on $V(F)$ and hence also not on X, we have obtained a hypersurface $Y \subset \mathbf{P}^n$ with $\deg(Y) \leq \deg(X)$, such that π is étale outside $X \cap Y$. By induction we derive:

Proposition 2.2. *Suppose X is irreducible, and that k has characteristic zero. There exist a projection $\pi\colon X \to P = \mathbf{P}^d$ and a hypersurface $Y \subset \mathbf{P}^n$ not containing X, with $\deg(Y) \leq (n - d) \cdot \deg(X)$, such that the ideal of Y annihilates $\Omega_{X/P}$. Especially, π is étale on $X - (X \cap Y)$.*

Also $Z = \pi(X \cap Y) \subset P$ is a hypersurface of degree $\leq (n - d) \cdot \deg(X)^2$ whose ideal annihilates $\Omega_{X/P}$.

We also need to estimate the degrees of irreducible components of intersections of hypersurfaces. For this we consider an m-fold product $P = \mathbf{P}^{n_1} \times \mathbf{P}^{n_2} \times \cdots \times \mathbf{P}^{n_m}$, and on it the line bundles $\mathscr{L}_i = \mathrm{pr}_i^*(\mathscr{O}(1))$. For any irreducible subvariety $X \subset P$ we obtain various degrees, indexed by sets of integers $e_1, \ldots, e_m$ with $e_1 + e_2 + \cdots e_m = \dim(X)$: Consider the intersection number $\mathscr{L}_1^{e_1} \cdot \mathscr{L}_2^{e_2} \cdot \cdots \cdot \mathscr{L}_m^{e_m} \cdot X$. They are always ≥ 0.

Proposition 2.3. *Suppose $Y \subset P$ is a subscheme which is an intersection of hypersurfaces of multidegree $(d_1, \dots, d_m)$, that is, intersection of the zero-sets of multihomogeneous polynomials of these degrees. If the X_j are irreducible components of Y with multiplicities m_j, and of the same codimension t, then*

$$\sum m_j \cdot \left(\mathscr{L}_1^{e_1} \cdot \mathscr{L}_2^{e_2} \cdot \cdots \mathscr{L}_m^{e_m} \cdot X_j\right)$$
$$\leq \mathscr{L}_1^{e_1} \cdot \mathscr{L}_2^{e_2} \cdot \cdots \mathscr{L}_m^{e_m} \cdot \left(d_1 \mathscr{L}_1 + d_2 \mathscr{L}_2 + \cdots + d_m \mathscr{L}_m\right)^t.$$

Proof. One easily finds $t = \operatorname{codim}(X)$ polynomials $F_1, \dots, F_t$, of multidegree $(d_1, \dots, d_n)$ and vanishing on Y, such that each X_j is an irreducible component of their common set of zeros. One then uses induction to bound the sum (with multiplicities) of the degrees $\mathscr{L}_1^{e_1} \cdot \mathscr{L}_2^{e_2} \cdot \cdots \mathscr{L}_m^{e_m} \cdot Z$ of those irreducible components Z of $V(F_1, \dots, F_{t-1})$ which have codimension $t - 1$; intersecting them with $V(F_t)$ (which is a proper intersection) gives the result.

b) *Heights*. There is a similar theory for heights; that is, we try to measure the arithmetic complexity of a subvariety of projective space. Now assume that k is a number field, and R its ring of integers. Places of k are denoted by v, and for a finite place v we let q_v denote the order of the residue field at v. Consider the projective space $\mathbf{P}^n$ over the integers $\mathbf{Z}$, or its base-extension to R. It admits the canonical line bundle $\mathscr{L} = \mathscr{O}(1)$, and everything is homogeneous under the algebraic group $\mathrm{GL}(n + 1)$. Over the complex numbers we endow $\mathscr{O}(1)$ with its hermitian metric as a quotient of $\mathscr{O}^{n+1}$, using the standard inner product on $\mathbf{Z}^{n+1}$ and $\mathbf{C}^{n+1}$. This is invariant under the unitary group $\mathrm{U}(n + 1, \mathbf{C})$. The curvature h of this metric defines a Kähler-metric on $\mathbf{P}^{n+1}(\mathbf{C})$, and harmonic differential forms coincide with $\mathrm{U}(n + 1, \mathbf{C})$-invariant forms, and are all multiples of powers of h. Especially, the product of harmonic forms is harmonic. H. Gillet and C. Soulé have defined arithmetic Chow groups $\widehat{CH}^p(\mathbf{P}^n)$, as follows (see [GS1], [GS2]): Consider $P = \mathbf{P}^n \times \operatorname{Spec}(R)$ as a regular scheme over $\operatorname{Spec}(\mathbf{Z})$, and let $P_\infty = P \otimes \mathbf{C}$ denote the associated complex variety. P_∞ is a disjoint union of complex projective spaces, and has the previously defined hermitian (Kähler) metric. The objects of $\widehat{CH}^p(\mathbf{P}^n)$ consist of pairs (X, g), where X is a cycle of pure codimension p in P, and g is a $(p - 1, p - 1)$-current on P_∞ such that $\delta_X - (1/\pi i)\bar{\partial}\partial g$ is C^∞. Also g must satisfy some reality condition. This has to be divided by the subgroup of pairs $(\operatorname{div}(f), -\log|f| + \partial u + \bar{\partial} v)$, where f is a meromorphic function on a $(p - 1)$-cycle and u and v are currents. If one chooses $\mathbf{Q}$-coefficients, one obtains an intersection product $\widehat{CH}^p(\mathbf{P}^n) \times \widehat{CH}^q(\mathbf{P}^n) \to \widehat{CH}^{p+q}(\mathbf{P}^n)$, as follows: If the two cycles X_1 and X_2 intersect transversally in the generic fibre, then after suitable modification of g_1 and g_2 one has

$$(X_1, g_1) \cdot (X_2, g_2) = \left(X_1 \cap X_2, g_1 \cdot \delta_{X_2} + \left(\delta_{X_1} - (1/\pi i)\bar{\partial}\partial g_1\right) \cdot g_2\right)$$

(one has to show that this makes sense). In general one uses a moving lemma. If $\mathscr{L}$ is a line bundle on P such that the induced $\mathscr{L}_\infty$ on P_∞ has a hermitian metric, there exists a well-defined class $\hat{c}_1(\mathscr{L}) \in \widehat{CH}^1(\mathbf{P}^n)$: If f is a non-zero meromorphic section of $\mathscr{L}$, $\hat{c}_1(\mathscr{L})$ can be represented by $(\operatorname{div}(f), -\log\|f\|)$. In fact we shall use the intersection product only to multiply cycles with classes $\hat{c}_1(\mathscr{L})$. This is much simpler than the general product and does not require the most involved considerations in [GS2], nor $\mathbf{Q}$-coefficients.

Definition 2.4. Suppose $X \subset P$ is an irreducible subvariety of codimension p. We define a class $\hat{X} \in \widehat{CH}^p(\mathbf{P}^n)$, as follows: If X lies in the fibre over a finite place v of k then $\hat{X} = (X, 0)$. Otherwise let $h_X = \deg(X_k) \cdot h^p$ denote the harmonic (p,p)-form on P_∞ representing X. Then there exists a current g_X such that i) $1/(\pi i)\bar{\partial}\partial g_X = \delta_X - h_X$, ii) $\int g_X \cdot \omega = 0$ for any harmonic form ω. Now g_X is unique up to terms $\partial u + \bar{\partial} v$, so that $\hat{X} = (X, g_X) \in \widehat{CH}^p(\mathbf{P}^n)$ is well-defined.

Also there is a degree map deg: $\widehat{CH}^{n+1}(\mathbf{P}^n) \to \mathbf{R}$, which sends (X, g) to $\deg(X) + \int_{P_\infty} g$, where $\deg(X) = \log(\text{order of } \Gamma(X, \mathscr{O}_X))$ (X is supported on finitely many closed points), and the analytic term is now the sum of the integrals over the components of P_∞. This defines an intersection-pairing $\widehat{CH}^p(\mathbf{P}^n) \times \widehat{CH}^{n+1-p}(\mathbf{P}^n) \to \mathbf{R}$, similar for several factors.

As an example consider $\mathscr{L} = \mathscr{O}(d)$ with its canonical metric, and suppose f is a global section of $\mathscr{L}$. Then $\operatorname{div}(f)\hat{\ } = \hat{c}_1(\mathscr{L}) + (0, \int_{P_\infty} \log\|f\| \cdot h^n)$. In the second term the integral denotes the locally constant function which on any component of P_∞ is given by the integral over this component. Also if X is an irreducible subvariety of codimension p, flat over R, such that f does not vanish identically on X, one checks that $\hat{c}_1(\mathscr{L}) \cdot \hat{X} = \operatorname{div}(f|X)\hat{\ } - (0, h^p \cdot \int_{X_\infty} \log\|f\| \cdot h^{n-p})$, where the meaning of the integral is as above.

Definition 2.5. Let X denote a cycle of pure codimension p on P, $\mathscr{L} = \mathscr{O}(1)$ with its canonical metric. Then $h(X) = \deg(\hat{X} \cdot \hat{c}_1(\mathscr{L})^{n+1-p}) \in \mathbf{R}$.

For example if X is contained in the fibre over a finite place v, then $h(X) = \log(q_v) \cdot \deg(X)$ which is always ≥ 0. This persists in general.

PROPOSITION 2.6. *For any effective cycle X, $h(X) \geq 0$.*

Proof. We use decreasing induction over p. If $p = n + 1$ the assertion holds, and it is also true for cycles supported on special fibres. So we may assume that X is irreducible and flat over R. Let $\{T_0, \ldots, T_n\}$ denote the standard-sections of $\mathscr{L} = \mathscr{O}(1)$. One of them, say f, does not vanish identically on X. Also f has norm ≤ 1 at the infinite places. It follows that

$$\hat{X} \cdot \hat{c}_1(\mathscr{L}) = \operatorname{div}(f|X)\hat{\ } - \left(0, h^p \cdot \int_{X_\infty} \log\|f\| \cdot h^{n-p}\right).$$

The integral in the second term on the right derives from the orthogonality of Green's functions to harmonic forms. It follows that $h(X) = h(\operatorname{div}(f|X)) - \int_{X_\infty} \log\|f\| \cdot h^{n-p}$.

As $\log\|f\| \leq 0$ the second term on the right is nonnegative, while we can apply induction to the first term.

Similarly, if f, for some d, is an integral nontrivial section of $\mathscr{L}^d$ over X (assumed to be irreducible), then $d \cdot \hat{X} \cdot \hat{c}_1(\mathscr{L})$ is represented by $\operatorname{div}(f)\hat{\ } - (0, h^p \cdot \int_{X_\infty} \log\|f\| \cdot h^{n-p})$. As the height of $\operatorname{div}(f)\hat{\ }$ is non-negative, we derive the following estimate:

Proposition 2.7. *Suppose X is irreducible and flat over R, $f \in \Gamma(X, \mathscr{L}^d)$ a nontrivial integral section. Then $\sum_v \sup(\log\|f\|_v) \geq -d \cdot h(X)/\deg(X)$. The sum is over the sup-norms at the various complex embeddings of k, or over all infinite places v of k where the complex places are counted with multiplicity two.*

Remark 2.8. i) Applied to $X =$ closure of a k-rational point $x \in P$, the above proof shows that $h(X)$ coincides with the usual logarithmic height of x, that is, the degree (as a metricised line-bundle; see [FW, Ch. II]) of the fibre of $\mathscr{L}$ in x. A similar remark applies to rational points in finite extensions of k.

ii) Assume $0 \neq f \in \Gamma(P, \mathscr{L}^d)$ is a global integral section. It follows from the above that up to terms bounded by constant $\cdot\, d$, the height of the divisor of f is given by $\int_{P_\infty} \log\|f\| \cdot h^n \in \mathbf{R}$ (which this time denotes the sum of the integrals over the components).

iii) For f as above there are various ways to measure the size of f at any infinite place v of k: One can use the supremum-norm $\sup(\|f\|)$, the L^2-norm or just the supremum of the coefficients of f. It follows from standard facts that any of these norms can be bounded in terms of any other, with a factor of the form constant^d. So if one is only interested in estimates up to such factors, these norms are all equivalent. It is a little bit less trivial that the same holds for the expression $\exp(\int_{P_\infty} \log\|f\|)$ occurring in ii) above:

Lemma 2.9. *There exists a constant $c > 0$, such that for any complex homogeneous polynomial f of degree d on $P = \mathbf{P}^n(\mathbf{C})$ (using $U(n+1, \mathbf{C})$-*

invariant norms and measures):

$$\exp\left(\int_P \log\|f\|\right) \le \sup(\|f\|) \le c^d \cdot \exp\left(\int_P \log\|f\|\right).$$

Proof. The first inequality is trivial, and it is easy to show the second one for $\mathbf{P}^1(\mathbf{C})$. In general we show the second inequality for some probability measure on P, independent of f. If we average over $\mathrm{U}(n+1, \mathbf{C})$, it then also holds for the invariant measure. The symmetric functions give a covering $\pi\colon \mathbf{P}^1(\mathbf{C})^n \to \mathbf{P}^n(\mathbf{C})$ which shows that $\mathbf{P}^n(\mathbf{C})$ is the n-symmetric power of $\mathbf{P}^1(\mathbf{C})$. Also $\pi^*(\mathscr{O}(1)) = \mathscr{O}(1, 1, \dots, 1)$, however, with a change in metrics. Now if we integrate $\log\|\pi^*(f)\|$ over $\mathbf{P}^1(\mathbf{C})^n$, we may first integrate over the last factor and get (by the $\mathbf{P}^1$-case) a lower bound for this by the supremum of $\log\|\pi^*(f)\|$ on $\mathbf{P}^1(\mathbf{C})^{n-1} \times \{z\}$, for any $z \in \mathbf{P}^1(\mathbf{C})$. If we choose an appropriate z and treat the other factors as above we get an upper bound for $\sup(\log\|\pi^*(f)\|)$ in terms of the integral of $\log\|\pi^*(f)\|$ over $\mathbf{P}^1(\mathbf{C})^n$, which implies the assertion.

After these generalities, we can start to prove the arithmetic analogues of the previous results about degrees. In the following we usually assume that the degrees of the varieties involved are already bounded a priori. We then obtain bounds on their heights. All these bounds will depend on k, and on the degrees of the varieties involved.

Let us first consider projections. These are in general defined over k. Assume $X_k \subset \mathbf{P}_k^n$ is irreducible. We want to project from a k-rational point $x \in \mathbf{P}_k^n - X$ to obtain $\pi(X_k) \subset \mathbf{P}_k^{n-1}$. To do the necessary estimates at infinity we need the facts that there the distance from x to X is not too small, and that x has bounded height. To satisfy the latter condition we choose a point x whose homogeneous coordinates are all integers of absolute value $\le \deg(X_k) \cdot [k:\mathbf{Q}]$. The projection of X_k to $\mathbf{P}_{\mathbf{Q}}^n$ (induced from $\mathrm{Spec}(k) \to \mathrm{Spec}(\mathbf{Q})$) has degree $\le \deg(X_k) \cdot [k:\mathbf{Q}]$; so there exists a nontrivial homogeneous polynomial F of that degree, with coefficients in $\mathbf{Q}$, such that F vanishes on the projection. As F cannot vanish identically on all the points x above (whose coordinates are bounded by $\deg(X_k) \cdot [k:\mathbf{Q}]$) we find one x not in X. Furthermore the norm of F can be bounded in terms of the norm of its values at these points, which easily implies that we can find an x such that for each infinite place v the distance $d_v(x, X)$ is bounded below by a positive constant only depending on $\deg(X_k)$. A projection, with centre such a point x, is called a good projection. Note that for fixed $\deg(X_k)$ there exist only a finite number of good projections, as there are only finitely many points x with the exhibited bound on their coordinates.

Suppose $\pi\colon X_k \subset \mathbf{P}_k^n - \{x\} \to \mathbf{P}_k^{n-1}$ is such a good projection. On $\mathbf{P}_k^{n-1}$ we choose an integral structure and norms by identifying it with some hyperplane of

$\mathbf{P}^n_{\mathbf{Z}}$ which does not meet the closure of $\{x\}$. This amounts to choosing a complement to a line in $\mathbf{Z}^{n+1}$, and it can be done such that the norms at infinity of $\mathscr{O}(1)|X_k$ and $\pi^*(\mathscr{O}(1))$ are mutually bounded by a fixed constant.

If we try to extend π to the integers R we find that in general it cannot be defined on all of X (= closure of X_k). More precisely x extends to an R-point of $\mathbf{P}^n$, which intersects X in a closed subscheme isomorphic to $\mathrm{Spec}(R/I)$, for a non-trivial ideal $I \subset R$, and π becomes regular after blowing up this subscheme of X. We denote by $\deg(\pi)$ the degree of π.

Proposition 2.10. *Assume π is a good projection. There exists a constant c (depending on* $\deg(X)$) *such that* $\deg(\pi) \cdot h(\pi(X)) \leq h(X) + c$.

Proof. Let $P = \mathbf{P}^n$, $P_1 = \mathbf{P}^{n-1}$, X_1 = closure of $\pi(X)$. We also may apply Arakelov theory to $P \times P_1$ (using the product metric). If p and p_1 denote the two projections onto the factors, we have direct image maps p_* and p_{1*} for cycles and currents. These preserve harmonic forms, and one derives that for closed subschemes $Z \subset P \times P_1$ the formation of $\hat{Z}$ commutes with direct images. We apply this to

$$Z = \tilde{X} = \text{closure of the graph of } \pi = \text{blow-up of } X \text{ in } X \cap \{x\},$$

so that $\tilde{X}^\wedge$ projects to $\hat{X}$, respectively, $\deg(\pi) \cdot \hat{X}_1$. If $\mathscr{L}$, respectively $\mathscr{L}_1$, denote the pullbacks of the hyperplane bundles on the factors (with hermitian metrics), it follows that $h(X) = \hat{c}_1(\mathscr{L})^{d+1} \cdot \tilde{X}^\wedge$, $\deg(\pi) \cdot h(X_1) = \hat{c}_1(\mathscr{L}_1)^{d+1} \cdot \tilde{X}^\wedge$, where $d = \dim(X_k) = \dim(X) = 1$. Hence their difference is equal to the sum, for i running from 0 to d, of $\hat{c}_1(\mathscr{L})^i \cdot \hat{c}_1(\mathscr{L}_1)^{d-i} \cdot \hat{c}_1(\mathscr{L} - \mathscr{L}_1) \cdot \tilde{X}^\wedge$. But over $\tilde{X}$ we have an injection $\alpha\colon \mathscr{L}_1 \subset \mathscr{L}$, whose divisor is the exceptional divisor E (and thus supported in fibres over finite places v of k), and such that the norm of $\alpha^{\pm 1}$ at the infinite places is universally bounded. It follows that $\hat{c}_1(\mathscr{L} - \mathscr{L}_1) \cdot \tilde{X}^\wedge$ can be represented by the pair $(E, \delta_{\tilde{X}} \cdot \log\|\alpha\|)$. Now $\mathscr{L}$ is trivial on E, and the intersection of E with $\hat{c}_1(\mathscr{L}_1)^d$ is a degree and non-negative. The analytic terms are universally bounded because $\log\|\alpha\|$ is.

Remark 2.11. $\tilde{X}$ is the blow-up of X in $X \cap \{x\} = \mathrm{Spec}(R/I)$, for an ideal $I \subset A$. In a moment we shall find a rational integer $0 \neq r \in I \cap \mathbf{Z}$ bounded by $\exp(c_1 \cdot h(X) + c_2)$, for c_1 and c_2 suitable constants. It follows that if f is a global integral section of $\mathscr{L}^{\otimes d} = \mathscr{O}(d)$ over X or even over the normalisation of X, the norm (under π) of $r \cdot f$ is a regular global section of $\mathscr{L}_1^{\otimes \deg(\pi) d} = \mathscr{O}(\deg(\pi) \cdot d)$ over the normalisation of X_1.

In the following three corollaries $X_k \subset \mathbf{P}^n$ denotes an irreducible subvariety over k, with $\deg(X_k)$ uniformly bounded, and X its closure. Also, c, c_1 and c_2 are suitable positive constants.

COROLLARY 2.12. *There exists a nontrivial homogeneous polynomial F of degree* $\deg(X_k) \cdot [k:\mathbf{Q}]$, *whose coefficients are rational integers bounded in size by* $\exp(h(X)+c)$, *such that F vanishes on X.*

This follows by projection down to a hypersurface.

COROLLARY 2.13. *There exists a good projection with centre $\{x\}$ such that over R the intersection $X \cap \{x\}$ is isomorphic to* $\operatorname{Spec}(R/I)$, *with the order of R/I bounded by* $\exp(c_1 \cdot h(X) + c_2)$, *for suitable constants c_1, c_2.*

Choose F as in 2.12 and x such that F does not vanish in x.

COROLLARY 2.14. *There exist a composition of good projections $\pi\colon X_k \to P_1 = \mathbf{P}_k^d$, $d = \dim(X_k)$, and a homogeneous polynomial F of degree $\leq (n-d) \cdot \deg(X_k) \cdot [k:\mathbf{Q}]$, whose coefficients are rational integers bounded in size by* $\exp(c_1 \cdot h(X) + c_2)$, *such that F does not vanish identically in X_k, but annihilates Ω_{X/P_1}. There also exists a hypersurface $Z \subset P_1$, of degree $< (n-d) \cdot \deg(X_k)^2 \cdot [k:\mathbf{Q}]^2$, defined by a polynomial G with coefficients in $\mathbf{Z}$ and bounded by* $\exp(c_1 \cdot h(X) + c_2)$, *such that G also annihilates Ω_{X/P_1}.*

Copy the proof of Proposition 2.2. For G, take the π-norm of $r \cdot F|X, r$ as in 2.13.

Remark 2.15. i) In Corollary 2.14 we need to know what happens over the integers. Then π is only well-defined on a modification $\tilde{X} \to X$ obtained by blowing up the intersections of X with centers of projections. One checks by a local calculation that the polynomials F and G constructed above also annihilate the relative differentials of $\tilde{X}$ over P_1.

ii) We need to estimate the heights of preimages of varieties. Assume $X_k \to \mathbf{P}_k^d$ is a composition of good projections. If $Y_k \subset \mathbf{P}_k^d$ is a proper subvariety it is contained in the zero-set of a polynomial whose coefficients are bounded in terms of $h(Y)$. This polynomial defines also a subvariety of $\mathbf{P}^n$ whose height is bounded by a multiple of $h(Y)$, and the height of its intersection with X is bounded by a multiple of $h(X) + h(Y)$ (use the equation $h(\operatorname{div}(f|X)) = h(X) + \int_{X_\infty} \log\|f\| \cdot h^{n-p}$).

At last we need an arithmetic companion to Proposition 2.3. We first remark that our theory extends to finite products $P = \mathbf{P}^{n_1} \times \mathbf{P}^{n_2} \times \cdots \times \mathbf{P}^{n_m}$. Now an irreducible subvariety $X \subset P$ has several heights, indexed by integers $e_1, \ldots, e_m$ with $e_1 + \cdots + e_m = \dim(X)$, namely the Arakelov intersection numbers $\hat{X} \cdot \hat{c}_1(\mathscr{L}_1)^{e_1} \cdot \cdots \hat{c}_1(\mathscr{L}_m)^{e_m}$. If $X = X_1 \times \cdots \times X_m$ is a product of subvarieties $X_i \subset \mathbf{P}^{n_i}$, flat over R, then this vanishes except if for one i, $e_i = \dim(X_i)$, and for the others, $e_j = \dim(X_j) - 1 = \dim(X_{j,k})$. In this case it is equal to the product of $h(X_i)$ and the degrees of the $X_{j,k}$, for j different from i.

For integers $d_1, \ldots, d_m$ let us consider the line-bundle $\mathscr{L} = \mathscr{O}(d_1, \ldots, d_m)$, and let $d = d_1 + \cdots + d_m$. We can use $\mathscr{L}$ to measure heights and degrees, and we indicate this by writing $h_{\mathscr{L}}(X)$ and $\deg_{\mathscr{L}}(X)$. We are interested in estimates which are uniform in the degrees d_i. This is a little bit more delicate than before, because now the degrees of the varieties involved are no longer bounded a priori. For any integral section f of $\mathscr{L}$ we have $\hat{c}_1(\mathscr{L}) = \operatorname{div}(f)^{\wedge} + (0, -\int_{P_\infty} \log\|f\|)$, and similarly for the restriction of f to subvarieties X. Furthermore $\mathscr{L}$ is generated by those of its global sections which are monomials in the coordinates, which have sup-norm ≤ 1, and for them the integral in the second term is non-positive. We derive as before that for any effective cycle X

$$h_{\mathscr{L}}(X) \geq 0.$$

Also assume that X is irreducible, the section f of $\Gamma(X, \mathscr{L})$ does not vanish identically, and for any infinite place v of k its restriction to X has norm $\leq C_v$. Then

$$(*) \qquad h_{\mathscr{L}}(\operatorname{div}(f|X)) \leq h_{\mathscr{L}}(X) + \deg_{\mathscr{L}}(X) \cdot \sum_v \log(C_v).$$

Again, complex infinite places are counted with multiplicity two. We use this as follows, to obtain the arithmetic counterpart to Propositions 2.7 and 2.3:

Proposition 2.16. $\sum_v \log(C_v) \geq -h_{\mathscr{L}}(X_k)/\deg_{\mathscr{L}}(X)$.

Proposition 2.17. *Assume given a family $\{f_\alpha\}$ of integral global sections of $\mathscr{L}$, such that at each infinite place v of k their sup-norm is uniformly bounded by C_v. Let Y denote their common set of zeroes, and assume that X_j are irreducible components of Y, with multiplicities m_j, all of the same codimension t in P, and flat over R. Then for some constant c*

$$\sum_j m_j \cdot h_{\mathscr{L}}(X_j) \leq t \cdot \deg_{\mathscr{L}}(P_k) \cdot \Big(c \cdot d + \sum_v \log(C_v)\Big).$$

Note that $\deg_{\mathscr{L}}(P_k) = (n_1 + \cdots + n_m)!/(n_1! \cdot \cdots \cdot n_m!) \cdot d_1^{n_1} \cdot \cdots \cdot d_m^{n_m} \leq d^{n_1 + \cdots + n_m}$.

Proof. We first find global sections $f_1, \ldots, f_t$, which are linear combinations of the f_α, such that the X_j are irreducible components of their common set of zeroes. For f_1 we may choose any non-zero f_α. By Proposition 2.3 the number of irreducible components of $V(f_1)$ flat over R is bounded by $\deg_{\mathscr{L}}(V(f_1)) = \deg_{\mathscr{L}}(P_k)$. So we can choose $\deg_{\mathscr{L}}(P_k)\, f_\alpha$ such that for each such irreducible component Z at least one of them does not vanish there, except if Z is also an irreducible component of Y. By an easy combinatorial argument we find a linear combination f_2 of these f_α, with coefficients integers of absolute value $\leq \deg_{\mathscr{L}}(P_k)$, such that f_2 does not vanish on any component Z, except if Z is also an irreducible component of Y (estimate which fraction of such linear combinations can vanish on a given Z). Now we make the same argument with $V(f_1, f_2)$, considering components of height 2. This gives an f_3, which again is a linear combination of $\leq \deg_{\mathscr{L}}(P_k)\, f_\alpha$'s with coefficients bounded by $\deg_{\mathscr{L}}(P_k)$, etc. This way we obtain $f_1, \dots, f_t$ as above, and the sup-norm of the f_i at the infinite place v is bounded by $\deg_{\mathscr{L}}(P_k)^2 \cdot C_v$. We now apply repetitively the inequality $(*)$. The term constant $\cdot\, d$ is used to absorb the logarithm of $\deg_{\mathscr{L}}(P_k)$.

c) *Siegel's lemma*. We need some refinement of the usual lemma. Assume first that V is a finite-dimensional $\mathbf{R}$-vector space with a norm $\| \ \|$, of dimension b, and that $M \subset V$ is a $\mathbf{Z}$-lattice. We normalise the volume on V such that the quotient V/M has volume 1, and denote by $\mathrm{vol}(V)$ the volume of the unit ball. If we define the parameter λ_i as the smallest numbers λ such that in M there exist i independent vectors of norm $\leq \lambda$, then by Minkowski's theorem (see [GL, p. 59, Th. 1]), $2^b/b! \leq \lambda_1 \cdot \cdots \cdot \lambda_b \cdot \mathrm{vol}(V) \leq 2^b$. Thus an upper bound for λ_b gives a lower bound for $\mathrm{vol}(V)$, and a lower bound for λ_1 an upper bound for $\mathrm{vol}(V)$.

Now assume we have given two such spaces V and W, with lattices M and N, and a linear map $\alpha\colon V \to W$ with $\alpha(M) \subset N$. Assume V has dimension b, and $U = \mathrm{Ker}(\alpha)$ has dimension a. We want to find nontrivial elements of small norm in the lattice $L = U \cap M$.

To do this assume that for some constant $C \geq 2$ the following holds: α has norm $\leq C$, M is generated by elements of norm $\leq C$, and any nontrivial element of M or N has norm $\geq C^{-1}$. If $\alpha(V)$ denotes the image of α with the quotient-norm and the quotient-lattice, it follows that $\lambda_1(\alpha(V)) \geq C^{-2}$, so that $\mathrm{vol}(\alpha(V)) \leq 2^{b-a} \cdot C^{2(b-a)}$. Also $\lambda_b(V) \leq C$, so that $\mathrm{vol}(V) \geq 2^b \cdot C^{-b}/b!$. Finally one checks easily that $\mathrm{vol}(V) \leq 2^a \cdot \mathrm{vol}(U) \cdot \mathrm{vol}(\alpha(V))$, and so $\mathrm{vol}(U) \geq C^{2a-3b}/b!$. As $\lambda_1(U) \geq C^{-1}$, it follows that for any i between 0 and $a-1$, $\lambda_{i+1}(U)^{a-i} \leq 2^a \cdot C^{3b-2a+i} \cdot b!$, so that:

Proposition 2.18. *In this situation*, $\lambda_{i+1} \leq (C^{3b} \cdot b!)^{1/a-i}$.

It follows that for any subspace $U_0 \subset U$ of dimension $\leq i$ we can find an element of L which is not in U_0, and whose norm is bounded by the right-hand side of the equation above. We shall apply this in situations where the dimensions a and b go to infinity, such that $b/(a-i)$ remains bounded. Then the right-hand side grows like a fixed power of C times a power of b.

d) *Leading terms*. We list some facts about leading terms in Taylor expansions, which are trivial but will be useful later on. If X is a scheme and D a differential operator on X, we cannot in general apply D to sections of a line-bundle $\mathcal{L}$ on X. That is, if we locally identify $\mathcal{L}$ with $\mathcal{O}_X$, choosing local generators, then the induced action of D depends on the choices of these generators also. However if $Y = V(\mathcal{I}) \subset X$ is a subscheme and $f \in \Gamma(X, \mathcal{L})$ vanishes to order $m \geq \deg(D)$ on Y, that is, $f \in \Gamma(X, \mathcal{I}^m \cdot \mathcal{L})$, then the restriction of $D(f)$ to Y is well-defined. We can do even better: If all associated primes of $\mathcal{O}_Y$ have characteristic 0 then for derivations $\partial_1, \dots, \partial_\llcorner$ on X, and integers $e_1, \dots, e_\llcorner$ such that $\sum_i e_i = m$, the restriction $D(f)|Y$ is integral for the differential operator $D = \prod_i \partial_i^{e_i}/e_i!$; i.e., $(\prod_i \partial_i^{e_i}(f))|Y$ is divisible by $\prod_i e_i!$: Embed into affine space, so that we may assume that $X = \mathbf{A}^n$. The assertion is obvious if $\partial_i = \partial/\partial z_i$ are the standard derivatives, and in general it follows by writing each ∂_i as a linear combination of such. $D(f)|Y$ does not depend on the ordering of factors in the product.

There is an analogue at infinity: Assume X is a complex manifold, $Y = \{y\} \subset X$ a point, $\mathcal{L}$ a line-bundle with hermitian metric, f an analytic section of order $\geq m$ at y, D a differential operators of order m. Now $D(f)(y)$ gives a term of degree m in the Taylor expansion of f at y. We can estimate its norm using Cauchy's formula: Choose a local generator of $\mathcal{L}$, thus identifying f with a function. Furthermore embed into X a small polydisk Δ of radius r centered at y, with coordinates z_i. If $\partial_i = \partial/\partial z_i$ and $D = \prod_i \partial_i^{e_i}/e_i!$ as above, $D(f)(y)$ is a term in the Taylor expansion of f at y, thus is bounded by $r^{-m} \cdot \sup(|f||\Delta)$. For general ∂_i's one obtains a similar estimate by expressing them as linear combinations of $\partial/\partial z$'s. The bound would be much worse if in the definition of D we did not divide by the factorials $e_i!$. Of course this is only an outline of the method used, but to be more specific requires knowledge of three special situations where we use such estimates; so this is postponed until later. Finally all this will be used in a slightly more general setup: The ambient scheme X will be a product $X = X_1 \times \cdots \times X_n$, and we will use weighted degrees, with weights $1/d_i$; that is, a derivation in the i-th direction will have degree $1/d_i$. However all the remarks above easily generalize.

3. The product theorem

This result will be our main technical tool in the following. The setup is: Suppose $P = \mathbf{P}^{n_1} \times \mathbf{P}^{n_2} \times \cdots \times \mathbf{P}^{n_m}$ is a product of projective spaces over a field k *of characteristic zero*, $\mathcal{L} = \mathcal{O}(d_1, \dots, d_m)$ with positive integers d_i, and f a global section of $\mathcal{L}$ over P. For a point $x \in P$ the index $i(x, f)$ of f at x is the maximal rational number σ such that the following holds:

Choose a local trivialisation of $\mathscr{L}$ near x, so that locally f can be identified with a function. Then for any set of integers $j_1, \ldots, j_m$ with $j_1/d_1 + j_2/d_2 + \cdots + j_m/d_m < \sigma$, and any choice of differential operators D_i of degree $\leq j_i$, on the i-th factor $\mathbf{P}^{n_i}$, $D_1 \cdot D_2 \cdot \cdots \cdot D_m(f)$ vanishes in x. The index is finite unless f vanishes identically. Let $Z_\sigma \subset \mathbf{P}$ denote the subset where the index is $\geq \sigma$. Obviously, Z_σ has a natural structure of a closed subscheme of P (locally defined by suitable derivatives of f). Also suppose for the moment that k is algebraically closed.

THEOREM 3.1 (product theorem). *Choose a positive number $\varepsilon > 0$. There exists an r (depending on ε) such that the following holds: Suppose that $d_1/d_2 \geq r, d_2/d_3 \geq r, \ldots, d_{m-1}/d_m \geq r$. Then if Z is an irreducible component of $Z_{\sigma+\varepsilon}$ which is also an irreducible component of Z_σ:*

(i) $Z = Z_1 \times \cdots \times Z_m$ *is a product of closed subvarieties* $Z_i \subset \mathbf{P}^{n_i}$.

(ii) *The degrees* $\deg(Z_i)$ *are bounded by some constant only depending on* ε.

Proof. Let $I_\sigma \subset \mathscr{O}_P$ denote the ideal of Z_σ. Then $D_1 \cdot \cdots \cdot D_m(I_\sigma) \subset I_{\sigma+\varepsilon}$, if the D_i are differential operators on the i-th factor, and the weighted (by $1/d_i$) sum of their degrees is $\leq \varepsilon$. Choose a generic point $z \in Z$, and let T denote the tangent space of Z at z. We obtain an increasing filtration $(0) = T_0 \subset T_1 \subset \cdots \subset T_m = T$, by choosing T_i = intersection of T with the tangent space of the first i factors $\mathbf{P}^{n_1} \times \cdots \times \mathbf{P}^{n_i} \subset \mathbf{P}$. If $\delta_i = \dim(T_i/T_{i-1})$, one then checks that the multiplicity m_z of Z in I_σ is bounded below by $c \cdot \prod d_i^{n_i - \delta_i}$, c a positive constant depending on ε. This follows from counting the number of differential operators transversal to Z which map I_σ into the ideal of Z. On the other hand Z_σ is defined by sections of $\mathscr{L} = \mathscr{O}(d_1, \ldots, d_m)$, obtained by applying invariant differential operators to f (P is a homogeneous space, and $\mathscr{L}$ a homogeneous line-bundle). By Proposition 2.3 we can bound the degree of Z, or more precisely the degrees (using $\mathscr{L}_i$ = pullback of $\mathscr{O}(1)$ from the i-th factor)

$$Z \cdot \mathscr{L}_1^{e_1} \cdot \cdots \cdot \mathscr{L}_m^{e_m}, \sum e_i = \dim(Z) = \sum \delta_i,$$

and obtain that

$$Z \cdot \mathscr{L}_1^{e_1} \cdot \cdots \cdot \mathscr{L}_m^{e_m} \leq (1/m_Z) \cdot \mathscr{L}^{\operatorname{codim}(Z)} \cdot \mathscr{L}_1^{e_1} \cdot \cdots \cdot \mathscr{L}_m^{e_m}.$$

Using the lower bound for m_Z we see that the right-hand side is bounded by

$$\text{constant} \cdot \prod d_i^{\delta_i - e_i} = \prod (d_i/d_{i-1})^{\eta_i},$$

with $\eta_m = \delta_m - e_m$, $\eta_{m-1} = (\delta_m + \delta_{m-1}) - (e_m + e_{m-1})$, etc. We also know (considering the differentials at the generic point z) that the dimension of the projection of Z to the last factor is $\leq \delta_m$, to the last two factors is $\leq \delta_m + \delta_{m-1}$, etc. It follows that $Z \cdot \mathscr{L}_1^{e_1} \cdot \cdots \cdot \mathscr{L}_m^{e_m}$ vanishes unless $e_m \leq \delta_m$, $e_m + e_{m-1} \leq \delta_m + \delta_{m-1}$, etc., that is, unless all $\eta_i \geq 0$. In this case $Z \cdot \mathscr{L}_1^{e_1} \cdot \cdots \cdot \mathscr{L}_m^{e_m}$ is bounded above by constant $\cdot\, r^{-\eta}$, $\eta = \sum \eta_i \geq 0$, so that for big r it can be a positive integer only if all η_i vanish, that is, if $e_i = \delta_i$. It follows that the projection of Z onto the i-th factor has dimension $\leq \delta_i$; so Z must be a product $Z = Z_1 \times \cdots \times Z_m$, and the product of the degrees of the Z_i is $Z \cdot \mathscr{L}_1^{e_1} \cdot \cdots \cdot \mathscr{L}_m^{e_m}$, thus bounded.

Remark 3.2. If k is not algebraically closed there is a slight difficulty because an irreducible component may not be geometrically irreducible. Assuming that f (and thus Z_σ) is defined over k, we obtain that the degree $[k_1 : k]$ of the field of constants k_1 of Z is also bounded (applying the degree estimate to all conjugates of Z), and that over k_1, Z is a product as before. However, in our applications, the smooth locus of Z will contain a k-rational point, and thus $k = k_1$.

Finally we also shall need estimates for the heights of the Z_i, in case k is a number-field. For this we let $d = d_1 + \cdots + d_m$, and assume that Z meets the affine product $\mathbf{A}^{n_1} \times \cdots \times \mathbf{A}^{n_m}$. This is easily done by choosing suitable hyperplanes at infinity. Now sections of $\mathscr{L}$ are identified with polynomials, and we assume that all coefficients of f lie in the integers R, and that at each infinite place v of k their norm is bounded by a constant C_v. If t is a coordinate of some factor $\mathbf{A}^{n_i}$, the differential operator $(\partial/\partial_t)^l/l!$ acts on $\Gamma(P, \mathscr{L})$, preserves integral polynomials, and has norm $\leq 2^{d_i}$ at the infinite places. From the definition of the index it follows that Z_σ is defined by polynomials which are obtained by applying products of such differential operators to f, so that they are integral and their coefficients bounded in v by $C^d \cdot C_v$, C some constant. By Proposition 2.17 we derive that $m_Z \cdot h_{\mathscr{L}}(Z)$ (we identify the height of a k-variety with that of its closure over R) is bounded by

$$\operatorname{codim}(Z) \cdot \left(\sum_v \log(C_v) + c \cdot d \right) \cdot \deg_{\mathscr{L}}(P),\ c \text{ a suitable constant.}$$

If $Z = Z_1 \times \cdots \times Z_m$ is a product, with $\dim(Z_i) = e_i$, it follows $h_{\mathscr{L}}(Z)$ is bounded by a constant multiple of $(\sum_v \log(C_v) + C \cdot d) \cdot \prod_i d_i^{e_i}$. On the other hand, by the binomial theorem, $h_{\mathscr{L}}(Z)$ is equal to a sum (indexed by j) of terms (binomial coefficient) $\cdot \prod_i d_i^{e_i} \cdot \prod_{i \neq j} \deg(Z_i) \cdot (d_j \cdot h(Z_j))$, which immediately gives a bound for $\sum_i d_i \cdot h(Z_i)$. In general Z is a product over some extension field k_1 of bounded degree, and this assertion holds over k_1. We have shown:

THEOREM 3.3. *Assume the hypotheses of Theorem* 3.1, *and that in addition f has integral coefficients which at each infinite place v of k are bounded by C_v.*

Then Z is geometrically irreducible over some extension k_1 of k of bounded degree (depending on ε), and over k_1, Z is a product $Z = Z_1 \times \cdots \times Z_m$. The degrees of the Z_i are bounded (depending on ε), and for suitable constants c_1, c_2 (depending on ε),

$$\sum_i d_i \cdot h(Z_i) \leq c_1 \cdot \left(\sum_v \log(C_v) \right) + c_2 \cdot d.$$

Remark 3.4. The product theorem will be used as follows: Suppose f has index $\geq \sigma$ at a point $x = (x_1, \ldots, x_n)$. Choose an integer $N > \dim(P)$. Then there exists a decreasing sequence $P \neq Z_1, Z_2, \ldots, Z_N \ni z$, such that each Z_i is an irreducible component of $Z_{i\sigma/N}$. It follows that two of these have the same dimension, so they must be products.

4. Proof of Theorem I: Geometry

Suppose k is an algebraically closed field of characteristic zero, A an abelian variety over k, and $X \subset A$ an irreducible subvariety which does not contain any translate of an abelian subvariety $B \subset A$ of positive dimension. We use this hypothesis as follows:

LEMMA 4.1. *For m big enough, the map α_m: $X^m \to A^{m-1}$ defined by $\alpha_m(x_1, \ldots, x_m) = (2 \cdot x_1 - x_2, 2 \cdot x_2 - x_3, \ldots, 2 \cdot x_{m-1} - x_m)$ is finite.*

Proof. The projection onto the first factors $X^{m+1} \to X^m$ induces closed immersions from fibres of α_{m+1} into fibres of α_m. It follows that the maximum of dimensions of fibres of α_m decreases with m. Let d denote its limit, so that we have to show that $d = 0$. So assume that it is positive, and that for $m \geq m_0$ all fibres of α_m have dimension $\leq d$. The number and the maximum of the degrees (measured using some ample line bundle on A) of d-dimensional irreducible components of fibres of α_m, $m \geq m_0$, are also decreasing, so become constant for big m. It follows that we can find a d-dimensional irreducible $Z \subset A$ such that for any $m \geq 0$ there exists $x_m \in A$ with $2^m \cdot Z + x_m \subset X$. It follows that for any $n \geq 0$, $2^n \cdot Z$ is contained in a fibre of f_m for all positive m, so that its degree is bounded uniformly in n. If $G \subset A$ denotes the algebraic subgroup such that $g \in G$ if and only if $g + Z = Z$, then the degree of $2^n \cdot Z$ is equal $\deg(Z) \cdot 4^{nd}/$(number of 2^n-torsion points in G). It follows that the order of 2^n-torsion in g grows like 4^{nd}, so that G has dimension $\geq d$ and Z is the translate of an abelian subvariety of A.

In the following, choose an integer m as above, and a very ample symmetric line bundle $\mathscr{L}$ on A, embedding $A \subset \mathbf{P}^n$. If add: $A \times A \to A$ denotes the addition, we let $\mathscr{P} = \text{add}^*(\mathscr{L}) - \text{pr}_1^*(\mathscr{L}) - \text{pr}_2^*(\mathscr{L})$ denote the corresponding Poincaré-bundle. For simplicity we use additive notation for the tensor product in the Picard group. In the following we also use freely **Q**-rational linear combinations of line-bundles, because ampleness makes perfect sense for such combinations, and our goal is to investigate this notion. [H] is an excellent reference.

Suppose $\varepsilon, s_1, \ldots, s_m$ are positive rational numbers. The **Q**-line-bundle $\mathscr{L} = \mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ on A^m is defined as

$$\mathscr{L}(-\varepsilon, s_1, \ldots, s_m) = -\varepsilon \cdot \sum_i s_i^2 \cdot \text{pr}_i^*(\mathscr{L}) + \sum_i (s_i \cdot x_i - s_{i+1} \cdot x_{i+1})^*(\mathscr{L}).$$

Here $(s_i \cdot x_i - s_{i+1} \cdot x_{i+1})^*(\mathscr{L})$ denotes the following: Choose an integer n such that $n \cdot s_i$ and $n \cdot s_{i+1}$ are integral, so that sending $(x_1, \ldots, x_m) \in A^m$ to $ns_i \cdot x_i - ns_{i+1} \cdot x_{i+1}$ is a morphism. Then $(s_i \cdot x_i - s_{i+1} \cdot x_{i+1})^*(\mathscr{L})$ is defined to be $n^{-2} \cdot$ (pullback of $\mathscr{L}$ under this morphism). $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ is a **Q**-linear combination of $\mathscr{L}_i = \text{pr}_i^*(\mathscr{L})$ and Poincaré-bundles $\mathscr{P}_{i,j}$. For fixed ε the coefficients of the $\mathscr{L}_i$ are proportional to s_i^2, and those of $\mathscr{P}_{i,j}$ to $s_i \cdot s_j$.

LEMMA 4.2. *Suppose $Y = Y_1 \times \cdots \times Y_m$ is a product subvariety of A^m. Then as a function of the s_i the intersection-product $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\dim(Y)} \cdot Y$ is proportional to $\prod_i s_i^{2\dim(Y_i)}$.*

Proof. The intersection number is a linear combination of terms $\prod_i \mathscr{L}_i^{e_i} \cdot \prod_{i,j} \mathscr{P}_{i,j}^{e_{i,j}} \cdot X$, with coefficients proportional to $\prod_i s_i^{2e_i} \cdot \prod (s_i s_j)^{e_{i,j}}$. Here the sum of all the exponents e_i and $e_{i,j}$ is $\dim(Y)$, and the intersection numbers vanish unless for each i we have $2 \cdot e_i + \sum_j e_{i,j} \leq \dim(Y_i)$. As the sum of the terms on the left equals that of terms on the right, we must have equality for each i, which implies the assertion.

COROLLARY 4.3. *There exists a positive ε_0, such that for any $\varepsilon \leq \varepsilon_0$ and for any product variety $Y \subset X^m$ the intersection number $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\dim(Y)} \cdot Y$ is positive.*

Proof. For $s_i = 2^{m-i}$ this follows from Lemma 4.1 ($\mathscr{L}(0, s_1, \ldots, s_m)$ is the pullback by α_m of an ample line bundle, hence itself is ample, and so is $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ for small ε), and Lemma 4.2 implies the rest.

We now can formulate the main result.

THEOREM 4.4. *Choose m as in Lemma 4.1 and ε_0 as in Corollary 4.3. For any $\varepsilon < \varepsilon_0$ there exists a real number s, such that $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ is ample on X^m if $s_1/s_2 \geq s$, $s_2/s_3 \geq s, \ldots, s_{m-1}/s_m \geq s$.*

Proof. We show by induction that for any integer N, and for any product variety $Y = Y_1 \times \cdots \times Y_m$ of X^m with $\deg(Y_i) \leq N$, there exists an s_0 (depending on ε and N) such that $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ is ample on Y where $s \geq s_0$. The induction is over $\dim(Y)$. Especially we may assume that $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ is ample if we replace one Y_i by a hyperplane section, so that the restriction of $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ to some ample divisor in Y is ample. Note that by Kleiman's theorem ([H, Ch. I, Th. 6.1]), $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ is already in the closure of the ample cone if it has nonnegative degree on any irreducible curve $C \subset Y$. Decreasing ε a little bit then gives an ample line bundle. So choose such a C.

First we set up the geometry: Choose projections $\pi_i\colon Y_i \to \mathbf{P}^{n_i} = P_i$ as in Proposition 2.2, defining $\pi\colon Y \to P = P_1 \times \cdots \times P_m$, with $\deg(\pi)$ the product of the degrees of the Y_i. Furthermore we find hypersurfaces $Z_i \subset P_i$, of bounded degree, such that the ideal of Z_i annihilates the relative differentials (Prop. 2.2). It follows that any derivation ∂ on P_i (there are many such because P_i is homogeneous) pulls back to a meromorphic derivation on Y_i, with polar divisor less than or equal to $\pi^*(Z_i)$. If C is contained in the preimage of some Z_i we are finished by induction. Otherwise let $D = \pi(C)$. The projection $\pi\colon C \to D$ is generically étale. We first show that for d big (and sufficiently divisible) $\Gamma(Y, \mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ is nontrivial. This follows as for big d, the Euler-characteristic $\chi(Y, \mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ is positive (Corollary 4.3), and as $h^i(Y, \mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ vanishes for $i \geq 2$ (the restriction of $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ to an ample divisor in Y is ample). Choose a nontrivial section f in $\Gamma(Y, \mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d})$.

In the equation (of line-bundles on A^m)

$$(s_i \cdot x_i - s_{i+1} \cdot x_{i+1})^*(\mathscr{L}) + (s_i \cdot x_i + s_{i+1} \cdot x_{i+1})^*(\mathscr{L})$$

$$= 2s_i^2 \cdot \mathscr{L}_i + 2 \cdot s_{i+1}^2 \cdot \mathscr{L}_{i+1}$$

both terms (or rather suitable multiples) on the left are generated by their global sections. It follows that there are injections of $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d}$ into $4d \cdot \sum_i s_i^2 \cdot \mathscr{L}_i = \sum_i d_i \cdot \mathscr{L}_i$ without a common zero on A^m. Here $d_i = 4s_i^2 \cdot d$. We choose such an injection ρ which does not vanish identically on C.

We define the index $i(C, f)$ of f along C as $i(x, f)$, x a generic point of C, using weights d_i on the factors. This coincides with the index of $\rho(f) \in \Gamma(Y, \pi^*(\mathscr{O}(d_1, \ldots, d_m))$. If $g = \operatorname{Norm}_\pi(f) \in \Gamma(P, \mathscr{O}(d_1, \ldots, d_m)^{\deg(\pi)})$ denotes the norm of $\rho(f)$, the index of $i(D, g)$ is bounded below by $i(C, f)/\deg(\pi)$. We choose a sufficiently small number $\sigma > 0$ (depending on $\dim(Y)$) and distinguish two cases:

i) $i(D, g) \geq \sigma$: Consider the subschemes of P where g has index $\geq \sigma \cdot i/\dim(P)$, $i = 1, 2, \ldots, \dim(P)$. We obtain (as in Remark 3.4) a decreasing sequence of irreducible components B_i of these schemes, such that each B_i contains D and thus has dimension strictly between 0 and $\dim(P)$. Thus some B_i must coincide with B_{i+1}, and by the product Theorem 3.1 it is a product of subvarieties of bounded degree, provided s is big enough. Thus C is contained in a product variety of smaller dimension, and the assertion follows by induction.

ii) $i(D, g) < \sigma$: Here $i(C, f) < \sigma \cdot \deg(\pi)$, so that f has small order of vanishing at C. Intuitively this means that some derivative of low order does not vanish there.

As each P_i is a homogeneous space it has many invariant derivations ∂_i. Also they extend to derivations on Y_i with simple poles along $\pi_i^*(Z_i)$. It follows that for any (non-commutative) multi-homogeneous polynomial $H(\partial_i)$ in such derivations, such that the weighted (with $1/d_i$) degree of H is $\leq i(C, f) = i(C, \rho(f))$, the restriction of $H(\partial_i)(f)$ to C is a global section of $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d}$ with poles along the $\pi_i^*(Z_i)$, that is, a global section of $\Gamma(Y, \mathscr{L}(c \cdot i(C, f) - \varepsilon, s_1, \ldots, s_m)^d)$, for some constant c. Furthermore for suitable choice of H this global section does not vanish identically. It is one of the leading terms in the Taylor expansion of f along C. It follows that $\mathscr{L}(c \cdot i(C, f) - \varepsilon, s_1, \ldots, s_m)$ has non-negative degree on C.

Now given $\varepsilon < \varepsilon_0$ we choose σ such that $c \cdot \deg(\pi) \cdot \sigma + \varepsilon$ is still smaller than ε_0, then s big enough as prescribed by induction. The assertion follows.

Example 4.5. Assume C is a curve of genus g, and $A = J(C)$ its Jacobian. $C \subset A$ satisfies the hypotheses of Theorem 4.4 if $g > 1$. As the next simplest example consider $X = C + C \subset A$. In general X may contain an elliptic curve even for big g; i.e., choose a double cover $C \to B$ of an elliptic curve B (these exist for any g). So we need some hypotheses on C, for example, that it is hyperelliptic. If $g = 3$ there even exist hyperelliptic covers $C \to B$, but for $g > 3$ and C hyperelliptic X satisfies the assumption:

Assume $B \subset X = \mathrm{Image}(C \times C \to A)$ is an elliptic curve. By pullback we obtain a quadratic cover $B_1 \to B$, with involution σ_1 on B_1, and a map $(\alpha, \alpha \cdot \sigma_1)$: $B_1 \to C \times C$ such that $\sigma \cdot \alpha \neq \alpha \cdot \sigma_1$, σ the hyperelliptic involution on C. The image of the map $(\alpha, \alpha \cdot \sigma_1)$: $B_1 \to C \times C$ meets properly the diagonal Δ as well as the image of $(1, \sigma)$: $C \to C \times C$. The degree of the pullback of Δ is at least equal to the number of fixed points of σ_1, that is, to $2 \cdot \mathrm{genus}(B_1) - 2$. On the other hand on $C \times C$ the divisor $\Delta + (1, \sigma)(\Delta)$ is linearly equivalent to $2 \cdot F_1 + 2 \cdot F_2$, F_i fibres of projections onto the factors. It follows that $2 \cdot \mathrm{genus}(B_1) - 2 \leq 4 \cdot \deg(\alpha)$. On the other hand, by Hurwitz, it is not smaller than $(2g - 2) \cdot \deg(\alpha)$, so that $g \leq 3$.

5. Proof of Theorem I: Arithmetic

We now assume that k is a number field, R its ring of integers, A is a normal irreducible projective R-scheme whose generic fibre A_k is an abelian variety, and $X \subset A$ a closed irreducible subscheme such that X_k does not contain any translate of an abelian subvariety. Choose m as in Lemma 4.1, a very ample line-bundle $\mathscr{L}$ on A which is symmetric on the generic fibre, and a proper normal modification $B \to A^m$, trivial over k, such that the Poincaré-bundles $\mathscr{P}_{i,j}$ on B_k extend to line bundles (called $\mathscr{P}_{i,j}$ again) on B. We also assume that B contains as an open subset the Neron-model $\mathscr{A}^m$ of A_k^m. It is easily checked that this can be reconciled with the earlier assumptions.

We know that over k for integers a, b the line-bundle $a^2 \cdot \mathscr{L}_i + b^2 \cdot \mathscr{L}_j + ab \cdot \mathscr{P}_{i,j} = (a \cdot x_i + b \cdot x_j)^*(\mathscr{L})$ is generated by its global sections, by pulling back global sections f of $\mathscr{L}$. Over the integers we can bound the pole order of $(a \cdot x_i + b \cdot x_j)^*(f)$, considered as a section of $a^2 \cdot \mathscr{L}_i + b^2 \cdot \mathscr{L}_j + ab \cdot \mathscr{P}_{i,j}$.

LEMMA 5.1. *Let f_ι run through a finite system of generators of $\Gamma(A_k, \mathscr{L})$. There exists a constant c such that for any choice of a, b, i, j there exists a positive integer $r < \exp(c \cdot (1 + a^2 + b^2))$ with:*

(i) *$r \cdot (a \cdot x_i + b \cdot x_j)^*(f_\iota)$ is an integral section of $a^2 \cdot \mathscr{L}_i + b^2 \cdot \mathscr{L}_j + ab \cdot \mathscr{P}_{i,j}$.*

ii) *On $\mathscr{A}^m$ the fractional ideal generated by the $r \cdot (a \cdot x_i + b \cdot x_j)^*(f_\iota)$ contains r^2.*

Proof. On $\mathscr{A}^m$ everything is easy, since we use that the theorem of the cube holds up to a finite error, and integrate twice. It remains to bound the pole order of $(a \cdot x_i + b \cdot x_j)^*(f_\iota)$ in the finitely many divisors of B supported on $B - \mathscr{A}^m$. Let V denote the local ring of B in the generic point of one such divisor. V is a discrete valuation-ring, and x_i and x_j define V-valued points of A. We may use base change $R \to V$ and change models, to reduce to the previous case.

Over k the f_ι define maps $\rho_\iota: d \cdot \mathscr{L}(-\varepsilon, s_1, \ldots, s_m) \to \Sigma_i d_i \cdot \mathscr{L}_i$, with the notation of the previous section. If we extend $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ to B (using the same linear combination of $\mathscr{L}_i$ and $\mathscr{P}_{i,j}$ as over the generic fibre), we see that the denominator of these maps is bounded by $\exp(c \cdot \Sigma_i d_i)$, c a suitable positive constant. Such a bound also holds for the norms of these maps at infinite places, equipping all line bundles with their canonical metric (with translation-invariant curvature).

Dually, the order of their common zeros also admits such a bound on the open subset $\mathscr{A}^m$; that is, if for a section of $d \cdot \mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ over k all

$\rho_\iota(f)$ are integral, then there the denominator of f is bounded by $\exp(c \cdot \Sigma_i d_i)$. Especially this holds in an open neighbourhood of the closure of any k-rational point of A^m. The analogue at infinity claims that the sup-norm of f can be estimated from the largest sup-norm of the $\rho_\iota(f)$, up to such a factor.

In general if one has a line-bundle $\mathscr{L}$ and a family of injections $\rho_\iota\colon \mathscr{L} \to \mathscr{M}$ into another bundle, without common zero, one obtains an exact sequence, given by the start of a Koszul-complex:

$$0 \to \mathscr{L} \to \bigoplus_i \mathscr{M} \to \bigoplus_{i,j} (\mathscr{M}^{\otimes 2} \otimes \mathscr{L}^{\otimes -1}).$$

In our case we may (assuming $\varepsilon \leq 1$) embed $\mathscr{M}^{\otimes 2} \otimes \mathscr{L}^{\otimes -1}$ into $\mathscr{M}^{\otimes 3}$, by the same method as above. We now restrict to the closure Y of X^m in B, pass to global sections, clear denominators and return to multiplicative notation, to obtain:

Proposition 5.2. *There exists an exact sequence*

$$0 \to \Gamma\left(X_k^m, \mathscr{L}(-\varepsilon, s_1, \dots, s_m)^{\otimes d}\right) \to \Gamma\left(X_k^m, \bigotimes_i \mathscr{L}_i^{d_i}\right)^a \to \Gamma\left(X_k^m, \bigotimes_i \mathscr{L}_i^{3d_i}\right)^b,$$

such that norms of maps and the difference between the norm on

$$\Gamma\left(Y, \mathscr{L}(-\varepsilon, s_1, \dots, s_m)^{\otimes d}\right)$$

and that induced on it by $\Gamma(Y, \otimes_i \mathscr{L}_i^{d_i})^a$ are both bounded by $\exp(c \cdot \Sigma_i d_i)$, *$c$ a suitable constant.*

Furthermore if a section f of $\mathscr{L}(-\varepsilon, s_1, \dots, s_m)^{\otimes d}$ over X_k^m maps to $\Gamma(Y, \otimes_i \mathscr{L}_i^{d_i})^a$, then on the open subset $Y^\circ = Y \cap \mathscr{A}^m$ of Y (which contains the closure of all k-rational points), the denominator of f is bounded by $\exp(c \cdot \Sigma_i d_i)$.

This proposition will replace all the difficult Arakelov theory in Vojta's paper. As it stands it obviously invites us to apply Siegel's lemma, that is, Proposition 2.18: The norms of integral global sections of $\Gamma(Y, \otimes_i \mathscr{L}_i^{d_i})^a$ and $\Gamma(Y, \otimes_i \mathscr{L}_i^{3d_i})^b$ are bounded below by $\exp(-c \cdot \Sigma_i d_i)$, by Proposition 2.16. Also using the $\mathscr{L}_i$ to embed X into $\mathbf{P}^n$ we see that the direct sum (over all d_i) of $\Gamma(Y, \otimes_i \mathscr{L}_i^{d_i})$ is finitely generated over the ring of multihomogeneous polynomials; so it has generators whose norm is bounded by $\exp(c \cdot \Sigma_i d_i)$. Finally the dimension of $\Gamma(Y, \otimes_i \mathscr{L}_i^{d_i})$ grows like $\prod_i d_i^{\dim(X_k)}$. It follows that if we can find a subspace of $\Gamma(X_k^m, \mathscr{L}(-\varepsilon, s_1, \dots, s_m)^{\otimes d})$ whose codimension has the same order of magnitude, then we can find global sections of $\Gamma(Y^\circ, \mathscr{L}(-\varepsilon, s_1, \dots, s_m)^{\otimes d})$ not in this subspace, whose norms at the infinite places are bounded by $\exp(\text{constant} \cdot \Sigma_i d_i)$. To find such a subspace choose a point $x = (x_1, \ldots, x_m)$ of X^m with values in some extension field of k. We assume that X_k is smooth at the x_i. If $\varepsilon < \varepsilon_0$ and the ratios s_i/s_{i+1} are sufficiently big (depending on ε) then $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)$ is ample on X_k, so that there exists a section of $\mathscr{L}(-\varepsilon, s_1, \ldots, s_m)^{\otimes d}$ not vanishing in x. It follows that for any positive σ we can produce sections of $\Gamma(X^m, \mathscr{L}(\sigma - \varepsilon, s_1, \ldots, s_m)^{\otimes d})$ having prescribed Taylor expansion at x, up to order a fixed multiple of $d \cdot \sigma$. Thus the set of sections having index $\geq \sigma$ in x has codimension bounded below by a positive multiple (depending only on $\dim(A)$, m and σ) of $\prod_i d_i^{\dim(X_k)}$.

THEOREM 5.3. *Assume $x = (x_1, \ldots, x_m)$ is a point of the smooth locus of X_k^m, that $0 < \sigma < \varepsilon < \varepsilon_0$, and that the ratios s_i/s_{i+1} are sufficiently big. Then there exists for big d an element $f \in \Gamma(Y^\circ, \mathscr{L}(\sigma - \varepsilon, s_1, \ldots, s_m)^{\otimes d})$ which has index $< \sigma$ in x, and whose norm at the infinite places is bounded by $\exp(c \cdot \sum_i d_i)$, c a constant only depending on σ and ε.*

Finally we can prove Theorem I: Suppose X contains infinitely many k-rational points. For each i choose a projection $\pi_i\colon X \to P_i = \mathbf{P}^{\dim(X)}$ as in Proposition 2.2, and a hypersurface $Z_i \subset P_i$, defined by a homogeneous polynomial with integral coefficients G_i, whose ideal annihilates $\Omega_{X/P}$. We can use the same choice for each i, but later on we have to distinguish the projections on different factors. We also let $\pi\colon X^m \to P = P_1 \times \cdots \times P_m$ denote the product.

By induction on $\dim(X)$ we may assume that there are infinitely many k-rational points not in any $\pi_i^{-1}(Z_i)$. Also $A(k) \otimes \mathbf{R}$ is a finite dimensional vector-space, with an inner product given by Néron-Tate heights, and we find infinitely many $x \in X(k)$ such that their directions $x/\|x\|$ lie in a small region of its unit sphere.

Choose ε and s as in Theorem 4.4. We can find points $x_1, \ldots, x_m$, with heights $h_1, \ldots, h_m$, such that

(i) h_1 is very big (how big will be determined later),
(ii) $h_2/h_1, h_3/h_2, \ldots, h_m/h_{m-1}$ are all $> s^2$,
(iii) $\langle x_i, x_{i+1} \rangle \geq (1 - \varepsilon/2) \cdot \|x_i\| \cdot \|x_{i+1}\|$,
(iv) $\pi_i(x_i)$ is not in Z_i.

We choose rational numbers s_i very close to $h_i^{-1/2}$, so that d_i will be about $4d/h_i$. The degree (as metricised line-bundle; see [FW], Ch. II) of $\mathscr{L}(\sigma - \varepsilon, s_1, \ldots, s_m)^{\otimes d}$ at x is then bounded above by $d \cdot (\sigma - \varepsilon/2) \cdot m +$ constant $\cdot \sum_i d_i$, the second term arising from differences between usual heights and Néron-Tate heights. It is important that the second term is of size d/h_1, so that the whole expression becomes negative if h_1 is sufficiently big. As the norm at the infinite places of the global section f of $\mathscr{L}(\sigma - \varepsilon, s_1, \ldots, s_m)^{\otimes d}$ is suitable bounded, we see that for h_1 big f will vanish in x. A refinement gives a lower bound for the index of f which is bigger than σ, and we have obtained a contradiction.

To obtain a lower bound for the index requires estimating values at x of derivatives of f. Let us first explain how to define these derivatives. Over k this is easy: If ∂_i is a vector field on P_i, we can lift it via π_i to a vector field on X_i with a simple pole along $\pi^*(Z_i)$. If $i(x, f)$ is the index of f at x (with weights $1/d_i$), and if $H(\partial_i)$ is a noncommutative polynomial in derivations on the various factors P_i, of total (weighted) degree $\leq i(x, f)$, then $H(\partial_i)(f)$ has a well-defined value at x. Furthermore for some H this value is nonzero.

Now we have to study what happens over the integers R, and at the infinite places. First the π_i are well-defined over $\tilde{X}_i$ which is obtained from X by blowing up some exceptional set. We also blow up Y to obtain a $\tilde{Y}$ dominating the product of the $\tilde{X}_i$. Assume that we start with integral derivations ∂_i on some factor P_i of P. If we lift to $\tilde{X}_i$ we obtain a denominator, which disappears after multiplication by G_i (Remark 2.15). Furthermore lifting from X^m to Y requires a new factor N, an integer which annihilates the relative differentials of B over A^m. If H is a product of terms $\partial_i^a/a!$, of weighted degree $\leq i(x, f)$, then the product of powers $(N \cdot G_i(\pi_i(x_i)))^{e_i}$ (e_i the partial degrees of H) times $H(\partial_i)(f)$ is integral at x, and non-zero for some H. Also the number of factors $\partial_i^a/a!$ can be chosen to be less than or equal to the dimension on P.

There is an analogous theory at infinity: Cover P_i by balls on which we trivialise $\mathscr{O}(1)$, that is, where we choose a local generator. Furthermore we assume that the norm of this generator lies everywhere between 1 and 2. To obtain estimates we use these trivialisations (and their pullbacks to X via π_i) to identify sections of $\mathscr{O}(d_i)$ with functions. By Cauchy's formula the value of the a-th derivative of a function at a point is bounded by $a! \cdot r^{-a} \cdot$ the supremum of this function on a disk of radius r centered at the point. In our situation, fix one infinite place v. The distance r_i from $\pi_i(x_i)$ to Z_i is bounded below by a multiple of the norm of G_i at $\pi_i(x_i)$, and on the ball of radius r_i around $\pi_i(x_i)$, π_i has an inverse; that is, we can lift this ball to a ball around x_i. It then follows from Cauchy's theorem that for a section $g \in \Gamma(X^m, \pi^*\mathscr{O}(d_1, \ldots, d_m))$ with index $i(x, g)$ at x, the value of $H(\partial_i)$ at x (H as above) has norm bounded by $\|f\| \cdot C^{d_1 + d_2 + \cdots + d_m} \cdot \prod_i r_i^{-e_i}$, where C is some constant and e_i the partial degrees of H (the e_i are bounded; $e_i \leq i(x, g) \cdot d_i < \sigma \cdot d_i$). The r_i may be replaced by the v-norms of $G_i(\pi_i(x_i))$.

This holds for the pullback-metric on $\pi^*\mathscr{O}(d_1, \ldots, d_m) = \otimes_i \mathscr{L}_i^{d_i}$. It follows that such an assertion also holds for the metric induced by embedding A into projective space, and also for the canonical metric (with translation-invariant curvature). However we may have to enlarge the constant C for this.

Finally there exists an injection ρ: $\mathscr{L}(\sigma - \varepsilon, s_1, \ldots, s_m)^{\otimes d} \to \otimes_i \mathscr{L}_i^{d_i}$ with norm bounded by $\exp(c \cdot \Sigma_i d_i)$, and whose norm at x is bounded below by $\exp(-c \cdot \Sigma_i d_i)$. Applying the estimate to $g = \rho(f)$ gives the same type of estimate for the norm of $H(\partial_i)(f)$.

Combining the finite and the infinite places we derive the following: The value at x of the product $\prod_i (N \cdot G_i(x_i))^{e_i} \cdot H(\partial_i)(f)$ is an integral element of the fibre at x of $\mathscr{L}(\sigma - \varepsilon, s_1, \ldots, s_m)^{\otimes d} \otimes \otimes_i \mathscr{L}_i^{e_i \cdot \deg(G_i)}$. Its norm at the infinite places is bounded by $\exp(c \cdot \Sigma_i d_i)$ with c a suitable constant. Furthermore for some choice of $H(\partial_i)$ it does not vanish. It follows that the metricised line-bundle above has degree bounded below by $-c \cdot \Sigma_i d_i$. On the other hand its degree is $\leq d \cdot (c_1 \cdot \sigma - \varepsilon/2) \cdot m + c_2 \cdot \Sigma_i d_i$, with c_1 and c_2 suitable positive constants (c_1 is essentially given by the degrees of the G_i). If we choose σ small enough such that $c_1 \cdot \sigma \leq \varepsilon/4$, and h_1 big enough so that $\Sigma_i d_i$ becomes sufficiently small compared to d, we get a contradiction. Thus X has only finitely many rational points, and Theorem I is proved.

6. Proof of Theorem II

Suppose again that A is an abelian variety over a number field k, and $E \subset A$ a closed subvariety. We intend to show:

THEOREM II. *For any place w of k and any positive κ the number of k-rational points $x \in A - E$, for which the w-local distance $d_w(x, E)$ from x to E is less than $H(x)^{-\kappa}$, is finite.*

Here $H(x)$ denotes the (big) height of x. Writing E as an intersection of divisors we may assume that E has codimension one, and it is irreducible. Also there exists a surjection $A \to A'$ such that E is a pullback of an ample divisor on A', which obviously allows us to assume that E is an ample irreducible divisor. We choose a very ample line-bundle $\mathscr{L}$ equivalent to a positive multiple of E, so that there exists a section $l \in \Gamma(A, \mathscr{L})$ with divisor some positive multiple of E. We also assume that $\mathscr{L}$ is symmetric (extend k and translate E by a k-rational point). Suppose $X_1, \ldots, X_m$ are irreducible subvarieties of A not contained in E, and $d_1, \ldots, d_m$ are integers. We say that a locally defined function on $X = X_1 \times \cdots \times X_m$ has index $\geq \delta$ at $E \times \cdots \times E$, if locally it is contained in the ideal generated by monomials $l_1^{j_1} \cdot l_2^{j_2} \cdot \cdots \cdot l_m^{j_m}$, with $j_1/d_1 + \cdots + j_m/d_m \geq \delta$. Here $l_i = \mathrm{pr}_i^*(l)$. This also makes sense for sections of line-bundles.

Now let us explain our strategy: Assume that the theorem is false. Then we choose a fixed positive rational number $\delta < 1$, say $\delta = 1/2$, an integer m and a positive rational $\varepsilon < 1$ with $3m \cdot \varepsilon < \kappa \cdot \delta$, such that

$$(*) \qquad 2 \cdot \delta^m / m! < \varepsilon^{\dim(A)} \cdot \dim(A)^{-m} \cdot 5^{-m \cdot \dim(A)}.$$

We find points $x_1, \ldots, x_m$ with $d_w(x_i, E) < H(x_i)^{-\kappa}$, with logarithmic heights h_i, such that h_1 is big, $h_2/h_1 > s^2, \ldots, h_m/h_{m-1} > s^2$, where s denotes a sufficiently large real number. We also assume that their directions in $A(k) \otimes \mathbf{R}$ are close, in the sense that $\langle x_i, x_j \rangle \geq (1 - \varepsilon) \cdot \|x_i\| \cdot \|x_j\|$, $\|x\| = h(x)^{1/2}$, the Néron-Tate height. For rational s_i very close to $h_i^{-1/2}$ we define the ample $\mathbf{Q}$-line-bundle

$$\mathscr{L}(\varepsilon, s_1, \ldots, s_m) = \varepsilon \cdot \sum_i s_i^2 \cdot \mathscr{L} + \sum_i (s_i \cdot x_i - s_{i+1} \cdot x_{i+1})^* \mathscr{L} \quad \text{on} \quad A^m.$$

Note that now all the terms in the sum are at least in the boundary of the ample cone. Also the degree (as metricised linebundle) of $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)$ in x is $< 3m \cdot \varepsilon$.

Next we define step by step a sequence of product varieties $X = X_1 \times \cdots \times X_m \subset A^m$, starting with A^m. The X_i will be irreducible over k, they contain x_i, the degrees are bounded, and the heights $h(X_i)$ are bounded by a multiple of h_i/h_1. However at each step these bounds may increase. We also need a positive σ satisfying (with a suitable constant to be determined later) $m \cdot (3\varepsilon + \text{constant} \cdot \sigma) < \kappa \cdot (\delta - \sigma)$. In each new step σ may decrease and X may be reducible, as the X_i are not required to be geometrically irreducible.

Given such an X we find a nontrivial section $f \in \Gamma(X, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$, with index $\geq \delta$ along $E \times \cdots \times E$, such that the norm of f is suitably bounded. It then follows from the hypotheses that f must vanish at x, and even better that f has index $\geq \sigma$ at x. By the product theorem we obtain that x is contained in a new product X, of smaller dimension. However, the dimensions cannot drop forever, and we obtain a contradiction. Now let us execute this plan. Let d denote a big sufficiently divisible integer, and $d_i = 5 \cdot d \cdot s_i^2$. We first study the situation over k:

As before $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d}$ injects into $\mathscr{M} = \otimes_i \mathscr{L}_i^{d_i}$, and $\mathscr{M} - d \cdot \mathscr{L}(\varepsilon, s_1, \ldots, s_m)$ is ample. Suppose we are given $X = X_1 \times \cdots \times X_m$ as above. The dimension of $\Gamma(X, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ is for big d asymptotically equal to $((d \cdot \mathscr{L}(\varepsilon, s_1, \ldots, s_m))^{\dim(X)} \cdot X)/\dim(X)!$. Now $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)$ is defined as a sum of terms which are at least in the boundary of the ample cone, and using the binomial theorem we can bound the dimension from below by

$$d^{\dim(X)} \Big/ \Big(\prod_i \dim(X_i)! \Big) \cdot \big(\varepsilon \cdot s_1^2 \cdot \mathscr{L}_1\big)^{\dim(X_1)} \cdot (s_1 \cdot x_1 - s_2 \cdot x_2)^* \mathscr{L}^{\dim(X_2)}$$

$$\cdot \cdots \cdot (s_{m-1} \cdot x_{m-1} - s_m \cdot x_m)^* \mathscr{L}^{\dim(X_m)} \cdot X,$$

that is, by $\varepsilon^{\dim(X_1)} \cdot \prod_i (d_i/5)^{\dim(X_i)} \cdot \deg(X_i)/\dim(X_i)!$.

On the other hand, let $Y_i = V(g_i) \subset X_i$, $Y = Y_1 \times \cdots \times Y_m$. The ideal $\mathscr{I}_\delta \subset \mathscr{O}_X$ which defines the condition "index $\geq \delta$" is generated by monomials $l_1^{j_1} \cdot \cdots \cdot l_m^{j_m}$ of weighted degree $\geq \delta$, with weights $1/d_i$; the l_i form a regular sequence on X. It follows that $\mathscr{O}_X/\mathscr{I}_\delta$ can be filtered such that the successive quotients are isomorphic to $\mathscr{O}_{Y^\otimes} \otimes_i \mathscr{L}_i^{\otimes -j_i}$, for all integers j_i, with $\Sigma j_i/d_i < \delta$. The number of such tuples is about $(\delta^m/m!) \cdot \prod_i d_i$. When we embed $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d}$ into $\mathscr{M}$ and estimate dimensions as above, the number of conditions on $f \in \Gamma(X, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ imposed by requiring index $\geq \delta$ is asymptotically bounded by the product of the number of tuples $(j_1, \ldots, j_m)$ and of

$$(\mathscr{M}^{\dim(Y)} \cdot Y)/\dim(Y)! = \prod_i d_i^{\dim(Y_i)} \cdot \deg(Y_i)/\dim(Y_i)!.$$

From $(*)$ above one derives that this is at most half of the dimension of $\Gamma(X, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$. It follows that the dimension of the space $\Gamma_\delta(X, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ of sections having index $\geq \delta$ is bounded below by $c \cdot \dim(\Gamma(X, \mathscr{M}))$, c a positive constant independent of X.

The same of course applies to $\mathscr{M}$. Moreover an easy ampleness argument (using $\delta < 1$) implies that for d big and sufficiently divisible $\Gamma_\delta(X, \mathscr{M})$ is generated by elements of the form $g \cdot \prod_i l_i^{j_i}$, $g \in \Gamma(X, \otimes_i \mathscr{L}_i^{d_i - j_i})$, where $\Sigma_i j_i/d_i = \delta$.

Now introduce integral structures. Everything works as in the previous chapter: Extend A to a projective scheme over R, $\mathscr{L}$ to a very ample line-bundle on it, and choose a modification $B \to A^m$ such that $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)$ extends to B. Furthermore X extends by taking closures, and $X^\circ = X \cap \mathscr{A}^m$. We also find injections $\rho_\downarrow$: $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d} \to \mathscr{M}$ allowing us to apply Siegel's lemma to find integral sections of $\mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d}$. To get integral sections with index δ we use the $\rho_\downarrow$ to inject $\Gamma_\delta(X, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ into $\Gamma_\delta(X, \mathscr{M})^a$. However on $\Gamma_\delta(X, \mathscr{M})$ we do not use the norm as a subspace of $\Gamma(X, \mathscr{M})$ but as the quotient of a direct sum of $\Gamma(X, \otimes_i \mathscr{L}_i^{d_i - j_i})$, for $\Sigma_i j_i/d_i = \delta$, mapping to it via multiplication by the monomial $l_1^{j_1} \cdot \cdots \cdot l_m^{j_m}$. Also as its integral structure we use the quotient lattice.

Again we find integral generators of small norm, so that the volume of $\Gamma_\delta(X, \mathscr{M})$ can be bounded from above by Minkowski's theorem. We also need a lower bound for the norm of integral sections of m, which by Proposition 2.16 can be chosen as $\exp(-c \cdot \Sigma_i d_i \cdot h(X_i))$, c a suitable constant depending only on the degrees of the X_i. Using as before a complex $0 \to \Gamma_\delta(X^\circ, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d}) \to \Gamma_\delta(X, \mathscr{M})^a \to \Gamma(X, \mathscr{M}^{\otimes 3})^b$, we derive:

LEMMA 6.1. *There exists a nontrivial integral section* $f \in \Gamma_\delta(X^\circ, \mathscr{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d})$ *with norm* (*at the infinite places*) *bounded by* $\exp(c \cdot \Sigma_i d_i \cdot h(X_i))$, *c a constant depending only on the degrees* $\deg(X_i)$. *Even* *better, the images* $g_\iota = \rho_\iota(f) \in \Gamma_\delta(X, \mathcal{M})$ *can be written as a sum of terms* $g_j \cdot \prod_i l_i^{j_i}$, $g_j \in \Gamma(X, \otimes_i \mathcal{L}_i^{d_i - j_i})$, *for* $\sum_i j_i/d_i = \delta$, *such that even all the* g_j*'s are integral and bounded in norm by* $\exp(c \cdot \sum_i d_i \cdot h(X_i))$.

Now let us assume that the degrees of the X_i as well as $\sum_i d_i \cdot h(X_i)/\sum_i d_i$ are bounded by a constant. We now study the index $i(x, f)$ of f at x. First assume that $i(x, f) \geq \sigma$: Choose over k (products of) good projections π_i: $X_i \to P_i = \mathbf{P}^{\dim(X_i)}$, and extend over R to π_i: $\tilde{X}_i \to P_i$, where $\tilde{X}_i$ is a blow-up whose size is controlled by $h(X_i)$. Furthermore choose non-trivial integral homogeneous polynomials G_i of bounded degrees, defining $Z_i \subset P_i$, which annihilate relative differentials. The norm of G_i can be bounded by $\exp(c \cdot h(X_i))$. Let π: $\tilde{X}_1 \times \cdots \times \tilde{X}_m \to P = P_1 \times \cdots \times P_m$ denotes the product of the π_i. If some $\pi_i(x_i)$ is contained in Z_i we replace X_i by a component of $\pi_i^{-1}(Z_i)$, and start again (the degree is bounded by a constant, and the height by a multiple of $h(X_i)$). Otherwise after we clear denominators of size $\leq \exp(c \cdot \sum_i d_i)$, $g = \mathrm{Norm}_\pi(\rho(f))$ has norm bounded by $\exp(c \cdot \sum_i d_i)$, and its index at $\pi(x)$ is $\geq \sigma/\deg(\pi)$. Also each X_i is geometrically irreducible, as its smooth locus contains the rational point x_i.

If s is big enough by the Product Theorems 3.1, 3.3 (and the same considerations as before), $\pi(x)$ is contained in a proper subvariety of P which is a product, such that the degree of its i-th factor is bounded, and its height is bounded by a constant multiple of $(\sum_j d_j \cdot h(X_j))/d_i$ or of $(\sum_j d_j)/d_i$. Again we may replace X by a smaller product and start again. So, finally assume that $i(x, f)$ is smaller than σ. As previously we apply to f a differential operator $H(\partial_i)$ of weighted degree $\sum_i e_i/d_i = i(x, f)$, to obtain the leading term $H(\partial_i)(f)(x)$ in the Taylor expansion of f at x. Also at the finite places the result becomes integral after multiplication by $\prod_i (N \cdot G_i(\pi_i(x_i)))^{e_i}$, and at the infinite places we obtain a similar estimate: The only difference is that now at infinity the distance from $\pi(x_i)$ to Z_i is bounded below by the norm of $G_i(\pi_i(x_i))$, divided by the norm of G_i. In the estimates this leads to an additional factor $\prod_i \|G_i\|^{e_i}$, which however is as usual bounded by $\exp(c \cdot \sum_i d_i)$. Thus $\prod_i (N \cdot G_i(\pi_i(x_i)))^{e_i} \cdot H(\partial_i)(f)(x)$ is a nontrivial integral element in the fibre of $\mathcal{L}(\varepsilon + \text{constant} \cdot \sigma, s_1, \ldots, s_m)^{\otimes d}$ at x, whose norm at infinite places is bounded by $\exp(c \cdot \sum_i d_i)$.

Now finally we use the information about the index to obtain a better bound for the w-norm of $H(\partial_i)(f)(x)$. Intuitively this norm should be small as x is close to E^m, and as f vanishes to a high order on E^m. We formulate the proof for an infinite place. For finite w the argument is the same, replacing cogrlex analytic multidisks by rigid multidisks. Choosing a suitable ρ: $\mathcal{L}(\varepsilon, s_1, \ldots, s_m)^{\otimes d} \to \mathcal{M}$ we may replace f by $g = \rho(f)$. A small ball B_i of radius r_i around $\pi_i(x_i)$ lifts to X_i, where r_i is about $\|G_i(\pi_i(x_i))\|/\|G_i\|$. We also may choose generators for $\mathscr{L}_i|B_i$, of norms between 1 and 2, so that on $B = B_1 \times \cdots \times B_m$, g and the h_i become holomorphic functions. Furthermore we have an expansion $g = \sum_J l^J \cdot g_J$, where $l^J = \prod_i l_i^{j_i}$, g_J is holomorphic of norm bounded by $\exp(c \cdot \sum_i d_i)$, and $J = (j_1, \ldots, j_m)$ satisfies $\sum_i j_i/d_i = \delta$.

As before $H(\partial_i)$ is a product (with at most $\dim(P)$ factors) of operators $\partial_i^a/a!$. Writing ∂_i as a linear combination of derivatives $\partial/\partial z$ along the coordinates on B_i we may assume that they are already such standard derivations. Furthermore in the sum $g = \sum_J l^J \cdot g_J$ the number of terms is bounded by a polynomial in d; so it suffices to estimate the value of $H(\partial_i)$ applied to the individual terms. Because of the bounds on the g_J it suffices to estimate $H(\partial)(l^J)$, or to estimate the value of $\partial_i^a(l_i^b)/a!$ at x_i, where ∂_i is the derivative in some coordinate direction in B_i.

We intend to show that this is bounded by $C^{a+b} \cdot r_i^{-a} \cdot |l_i(x)|^{b-a}$, where C is a suitable constant. This follows from the next remark:

LEMMA 6.2. *Suppose $\varphi(z)$ is a holomorphic function in the unit disk, of* sup-norm ≤ 1. *Then for $a \leq b$,*

$$|(\partial^a \varphi^b/\partial z^a)(0)/a!| \leq 2^{a+b} \cdot |\varphi(0)|^{b-a}.$$

Proof. Write $\varphi(z) = \sum_n a_n \cdot z^n$, with $|a_n| \leq 1$. Then

$$a_0^{a-b} \cdot (\partial^a \varphi^b/\partial z^a)(0)/a!$$

is a polynomial in the a_n, with positive coefficients. It is thus bounded by its value for $a_0 = a_1 = \cdots = 1$, that is for $\varphi(z) = 1/(1-z)$. But this φ has absolute value ≤ 2 on the disk of radius $1/2$, etc.

Finally, we see that the previous estimate for the w-norm of $H(\partial_i)(h^J)(x)$ can be improved by a factor $\prod_i \|l_i(x_i)\|_w^{j_i - e_i}$. Now, as at w, x_i is supposed to be close to E; that is, $d_w(x_i, H) < \exp(-\kappa \cdot h(x_i))$ and the w-norm of $l_i(x_i)$ is bounded by a multiple of $\exp(-\kappa \cdot h_i)$. It follows that we have gained a factor $\exp(-\kappa \cdot \sum_i (j_i - e_j) \cdot h_i) \leq \exp(-\kappa \cdot d \cdot (\delta - \sigma))$.

Thus $\prod_i (N \cdot G_i(\pi_i(x_i)))^{e_i} \cdot H(\partial_i)(f)(x)$ is a nontrivial integral element in the fibre of $\mathscr{L}(\varepsilon + \text{constant} \cdot \sigma, s_1, \ldots, s_m)^{\otimes d}$ at x, whose norm at infinite places is bounded by $\exp(c \cdot \sum_i d_i)$, and at w by an additional factor $\exp(-\kappa \cdot d \cdot (\delta - \sigma))$. Now the degree of $\mathscr{L}(\varepsilon + \text{constant} \cdot \sigma, s_1, \ldots, s_m)^{\otimes d}$ at x is bounded below by $\kappa \cdot d \cdot (\delta - \sigma) - c \cdot \sum_i d_i$, c a suitable constant. On the other hand this degree is less than $m \cdot d \cdot (3\varepsilon + \text{constant} \cdot \sigma)$, so that $\kappa \cdot (\delta - \sigma) - m(3\varepsilon + \text{constant} \cdot \sigma) < c \cdot \sum_i h_i^{-1}$. But the left-hand side is positive by our choice of constants, and for big h_1 the right-hand side becomes arbitrarily small. This contradiction proves our claim.

Corollary 6.2 (S. Lang's conjecture). *Assume $E \subset A$ is an ample divisor. Then $A - E$ has only finitely many integral points.*

Proof. Assume as before that l is an equation for E. The height $H(x)$, for $x \in A - E$ an integral point, is essentially the inverse of the product of the v-norms of $l(x)$, v running through the infinite places of k. Our proof actually gives that all these are bounded below by multiples of $H(x)^{-\kappa}$, so that $H(x)$ must be bounded.

Princeton University, Princeton, New Jersey

References

[FW] G. Faltings and G. Wüstholz, *Rational Points*, Vieweg-Verlag, Braunschweig (1984).
[GS1] H. Gillet and C. Soulé, Classes caractéristiques en théorie d'Arakelov, C.R. Acad. Sc. Paris, **301** (1985), 439–442.
[GS2] ______, Arithmetic intersection theory, preprint, IHES (1988).
[GL] P. M. Gruber and C. M. Lekkerkerker, *Geometry of Numbers*, North Holland, Amsterdam (1987).
[H] R. Hartshorne, *Ample Subvarieties of Algebraic Varieties*, Springer Lecture Notes 156 (1970).
[S] W. M. Schmidt, Approximation to algebraic numbers, L'Ens. Mathém. **17** (1971), 187–253.
[V] P. Vojta, Siegel's theorem in the compact case, Ann. of Math. **133** (1991), prior paper, this issue.

(Received October 4, 1989)