SMALL POINTS AND ARAKELOV THEORY

SHOU-WU ZHANG

ABSTRACT. In this talk, I will explain the recent applications of Arakelov theory to the Bogomolov conjecture.

1991 Mathematics Subject Classification: 11, 14
Keywords and Phrases: Néron Tate heights, Bogomolov conjecture, Arakelov theory

Height of a solution (resp. point) of a diophantine system (resp. variety) will measure the complexity of the solution (resp. point). For an abelian variety, one can define heights for its algebraic points to respect its group structure. For example, the torsion points will be only those who have zero heights. Such a normalization is called Néron-Tate height. The solutions with big heights, zero heights, or near zero heights are all interesting and important in Diophantine geometry. The Arakelov theory, an intersection theory on arithmetic varieties, has played an important role in the study of the Néron-Tate heights in recent years. In this talk, I will explain the recent applications of Arakelov theory to the Bogomolov conjecture on small points. For more details about the proof of this conjecture using Arakelov theory, one should see [Ab, U, Zh5]. (For other recent developments in Arakelov theory, see [So].)

1. NÉRON-TATE HEIGHTS AND BOGOMOLOV CONJECTURE

Let \( \mathbb{Q} \) denote the algebraic closure of \( \mathbb{Q} \) in \( \mathbb{C} \). For each place \( p = \infty, 2, 3, 5, \ldots, \) let \( \| \|_p \) denote a \( p \)-adic norm over \( \mathbb{Q} \) normalized by \( \| p \|_p = 1/p \) if \( p \neq \infty \) and \( \| \|_\infty \) is the usual absolute value on \( \mathbb{C} \). For a point \( x = (x_0, \ldots, x_n) \in \mathbb{P}^n(\bar{\mathbb{Q}}) \), the naive height \( h_{naive}(x) \) of \( x \) is defined by

\[
h_{naive}(x) = \frac{1}{[K: \mathbb{Q}]} \sum_{\sigma \in \text{Gal}(K/\mathbb{Q})} \sum_{p} \max \log \| \sigma(x_0) \|_p, \ldots, \| \sigma(x_n) \|_p
\]

where \( K \) is a number field in \( \mathbb{C} \) containing \( x_i \), and \( \sigma \) are embeddings from \( K \) into \( \mathbb{C} \), and \( p \) are places of \( \mathbb{Q} \). If \( x \) is a rational point represented by an \( (n+1) \)-tuple of integers \( (x_0, \ldots, x_n) \) with no common divisor, then \( h_{naive}(x) \) is \( \log \max |x_0|, \ldots, |x_n| \). If we define the complexity \( c(x) \) of \( x \) as the maximum of numbers of digits of \( x_i \) which measures the time spent to write a number down, then \( h_{naive}(x) - c(x)/\log 10 \) is bounded on the set of rational points of \( \mathbb{P}^n \). A basic property of \( h_{naive} \) is the following Northcott Theorem: for any given number \( D \) and \( H \), the set of points in \( A \) with height \( \leq H \) and degree \( \leq D \) is finite.
Let $A \to \mathbb{P}^n$ be an abelian variety embedded in $\mathbb{P}^n$ defined over $\mathbb{Q}$. Assume that the embedding is symmetric; this means that there is an automorphism $\phi$ of $\mathbb{P}^n$ such that $\phi(A) = A$ and $\phi_*[A] = [-1]_*A$. The Néron-Tate height $\hat{h}(x)$ of a point $x$ in $A(\mathbb{Q})$ is defined by the formula:

$$\hat{h}(x) = \lim_{m \to \infty} \frac{h_{\text{naive}}(m^2)}{m^2}.$$ 

There are two properties of $\hat{h}$ besides Northcott’s Theorem:
1. $\hat{h}(x) \geq 0$, and $\hat{h}(x) = 0$ if and only if $x$ is a torsion point;
2. the induced function $\hat{h}$ on the $\mathbb{Q}$-vector space $A(\mathbb{Q})/A_{\text{tor}}$ is quadratic.

**Theorem A (Bogomolov conjecture).** Let $X$ be an irreducible, closed subvariety of an abelian variety $A$ defined over $\mathbb{Q}$. Let $h : A(\mathbb{Q}) \to \mathbb{R}$ be a Néron-Tate height function (with respect to a symmetric projective embedding of $A$). Assume that $X$ is not a translation of an abelian subvariety by a torsion point. Then there is a positive number $\epsilon$ such that the set

$$\{x \in X(\mathbb{Q}) : h(x) < \epsilon\}$$

is not Zariski dense in $X$.

**Remarks:**
1. As the above set contains all torsion points in $X$, the above theorem implies a theorem of Raynaud [Ra] on Lang’s conjecture.
2. The original Bogomolov conjecture stated in [Bo] p. 70 has the following form: Assume $A$ is the Jacobian of a curve $C$ of genus $\geq 2$. For each $x \in C$ define an embedding $C \to A$ by sending $p$ to the class of $x - p$. Let $r(x)$ denote the maximal number $r$ such that there are only finitely many $p \in C$ such that $x - p$ has height less than $r$. Define $R(C)$ as the infimum of $r(x)$. Then he conjectured that $R(C) > 0$ and that $R(C)$ should be a certain height function on the moduli space of curves $C$. Theorem A implies that $r(x) > 0$ in general and that $R(C) > 0$ if the subvariety

$$X = \{x - y \in A : x, y \in C\}$$

does not contain any translations of elliptic curves by torsion points. In the next section, we will explain a variant of our theorem which will implies that $R(C) > 0$ and that $R(C)$ is certainly a height function on the moduli space of $C$’s.

2. Arithmetic ampleness and the theorem of successive minima

Using Arakelov theory [A, F1, GS1, Zh4], we can express the Néron-Tate heights as the degrees of hermitian lines on arithmetic curves. We illustrate the idea in the case that $A$ is an abelian variety defined over a number field $F$ and can be extended to an abelian scheme $\pi : A \to B = \text{Spec} \mathcal{O}_F$. Let $\mathcal{L}$ be a line bundle on $A$ which extends the restriction of $\mathcal{O}(1)$ on $A$. Replacing $\mathcal{L}$ by $\mathcal{L} \otimes [-1]^*\mathcal{L}$, we may...
assume that \( \mathcal{L} \) is symmetric: \([-1]^*\mathcal{L} = \mathcal{L} \). Also replacing \( \mathcal{L} \) by \( \mathcal{L} \otimes \pi^*\varepsilon^*(\mathcal{L}^{-1}) \) we may fix a rigidification \( r : e^*\mathcal{L} \simeq \mathcal{O}_e \) where \( e \) denote the unit section of \( \mathcal{A} \). On the complex bundle \( \mathcal{L}(\mathbb{C}) \) over \( \mathcal{A}(\mathbb{C}) \) we choose a hermitian metric \( || \cdot || \) such that
1. The curvature of \( || \cdot || \) is an invariant form on \( \mathcal{A}(\mathbb{C}) \);
2. The map \( r_\sigma : e^*_\sigma \mathcal{L}(\mathbb{C}) \to \mathbb{C} \) is isometric for each archimedean place \( \sigma \).

Let \( \mathcal{L} \) denote the pair \((\mathcal{L},|| \cdot ||)\). Let \( x \) be a point in \( \mathcal{A}(\mathbb{C}) \). Then the Zariski closure \( \bar{x} \) of \( x \) has a normalization \( f : \text{Spec}\mathcal{O}_K \to \mathcal{A} \) where \( \mathcal{O}_K \) is the ring of integers of some number field \( K \). The \( \mathcal{O}_K \) invertible module \( \mathcal{N} := f^*\mathcal{L} \) is equipped with hermitian metric on \( \mathcal{N} \otimes_\sigma \mathbb{C} \) for each embedding \( \sigma : K \to \mathbb{C} \). Then we define the degree of \( \mathcal{L} \) on \( \bar{x} \) by

\[
\deg_\mathcal{L} \bar{x} = \log \frac{\#\mathcal{N}/n\mathcal{O}_K}{\prod_{\sigma K \to \mathbb{C}} ||n \otimes_\sigma 1||}
\]

where \( n \) is any nonzero element in \( \mathcal{N} \). One can show that

\[
\hat{h}(x) = \frac{1}{[K: \mathbb{Q}]} \deg_\mathcal{L} \bar{x}.
\]

One immediate advantage of Arakelov theory is to extend the definition of the degree linearly to arbitrary cycles \( Z \) of \( \mathcal{A} \) by dimension induction:
1. if \( Z \) is a closed point of \( \mathcal{A} \) then \( \deg_\mathcal{L}(Z) = \log \#\kappa(Z) \) where \( \kappa(Z) \) is the residue field of \( Z \);
2. if \( Z \) is a closed subvariety of \( \mathcal{A} \), then

\[
\deg_\mathcal{L}(Z) = \deg_\mathcal{L}(\text{div}(\ell)|_Z) - \int_{Z(\mathbb{C})} \log ||\ell|| c_1 (\mathcal{L}_C)^{\dim Z}
\]

where \( \ell \) is a section of \( \mathcal{L} \) which is nonzero on \( Z \) and \( c_1 (\mathcal{L}_C) \) is the curvature form of \( \mathcal{L} \) which at any point where \( \ell \neq 0 \) can be given by

\[
c_1(\mathcal{L}_C) = \frac{\partial \bar{\mathcal{L}}_C}{\partial \ell} \log ||\ell||.
\]

If \( X \) is a closed subvariety of \( \mathcal{A} \), then the (Néron-Tate) height \( \hat{h}(X) \) is defined by the formula

\[
\hat{h}(X) = \frac{\deg_\mathcal{L}(X)}{(\dim X + 1) \deg_\mathcal{L}(X)}
\]

where \( \mathcal{X} \) is the Zariski closure of \( X \) in \( \mathcal{A} \). As for the Néron-Tate heights for points, \( \hat{h}(X) \) in general will be nonnegative.

As an example, let us consider the case that \( X \) is a curve over a number field \( F \) of genus \( g \geq 2 \) with a smooth model \( \mathcal{X} \) over \( B = \text{Spec}\mathcal{O}_F \) and that \( A \) is the Jacobian of \( X \), and that the embedding is \( \phi_D : x \to \text{class}(x - D) \), where \( D \) is a divisor of degree \( 1 \). Then one can show that

\[
\hat{h}(X) = \frac{1}{8(g - 1)[F: \mathbb{Q}]} c_1(\Omega_X^1/B)^2 + \left( 1 - \frac{1}{g} \right) h(x - D)
\]
where $\Omega_{X/B}^1$ is equipped with the Arakelov metric. In this case, \( \hat{h}(X) \geq 0 \) as \( c_1(\Omega_{X/B}^1)^2 \geq 0 \) is proved by Faltings.

Now assume that \( (2g - 2)D - c_1(\Omega_X^1) \) is a torsion point in \( A \). The first breakthrough step which brings the Bogomolov conjecture into the context of Arakelov theory is the following observation of L. Szpiro [Sz 1-3]: if \( c_1(\Omega_{X/B}^1)^2 > 0 \), then one can deduce from Faltings’ Riemann-Roch theorem [F1] that some positive power of \( \Omega_{X/B}^1 \) will have a section \( \ell \) with norm \( ||\ell||_{\text{sup}} < 1 \). Using this section to compute the height then one obtains that the points of \( X(\mathbb{Q}) \backslash \text{div}\ell \) will have height bigger than \(-\log ||\ell||_{\text{sup}}\). Szpiro also noticed that the truth of the Bogomolov conjecture would imply the positivity \( c_1(\Omega_{X/B}^1)^2 > 0 \), if one had a Nakai-Moishezon type criterion for ampleness of \( \Omega_{X/B}^1 \). In [K1-2], M. Kim proved a Nakai-Moishezon type result and deduced the equivalence of Bogomolov conjecture and \( c_1(\Omega_{X/B}^1)^2 > 0 \) in this case.

In [Z1, 3], using the arithmetic Hilbert-Samuel formula of Gillet and Soulé [GS 2-3, AB], a general Nakai-Moishezon type theorem for arithmetic variety has been proved. One immediate consequence is the following relation between \( \hat{h}(X) \) and the heights of points:

**THEOREM OF SUCCESSIVE MINIMAS.**

\[
\frac{1}{\dim X + 1} \sum_{i=1}^{\dim X+1} e_i(X) \leq \hat{h}(X) \leq c_1(X)
\]

where

\[
e_i = \sup_{Y \subseteq X, \text{codim } Y = i} \inf_{x \in X \backslash Y} \hat{h}(x).
\]

It follows that \( \hat{h}(X) = 0 \) if and only if \( c_1(X) = 0 \), or, the set of small points is dense. In particular, if \( X \) is the translate \( T + x \) of an abelian subvariety \( T \) by a torsion point \( x \) then \( \hat{h}(X) = 0 \). Assuming that \( X \) is not the translate of an abelian subvariety by a torsion point, then following three are equivalent:

1. the Bogomolov conjecture for \( X \);
2. \( c_1(X) > 0 \);
3. \( \hat{h}(X) > 0 \).

One immediate consequence is the Bogomolov conjecture for the embedding \( \phi_D : X \to A \) defined by a divisor \( D \) such that the class \( \Omega_X^1 - (2g - 2)D \) is not torsion. Going back to Bogomolov’s original conjecture, we have

\[
\kappa_1 c_1(\Omega_{X/B}^1)^2 \leq R(X) \leq \kappa_2 c_1(\Omega_{X/B}^1)^2
\]

where \( \kappa_1, \kappa_2 \) are two positive constants. All these results can be generalized to the case where \( X \) and \( A \) may have bad reduction by introducing adelic metrics, and admissible relative sheaf \( \omega_h \). See [Zh 2, 4] for more details.
3. **Equidistribution Theorems**

The question of whether \( c_1(\Omega_{X/B})^2 > 0 \) was very challenging in Arakelov theory because in the geometric case it is proved by deformation theory and it measures how far it differs from the constant fibre. The first example with \( c_1(\Omega_{X/B})^2 > 0 \) when \( X \) has good reduction is given by J.-F. Burnol [Bu]. He proved this positivity for curves whose Jacobians have complex multiplication by a CM-field of degree \( 2g(X) \). He uses two properties of the Weierstrass divisors \( W_d \) (\( d \in \mathbb{N} \)) for powers of \( \Omega_{X/B}^1 

1. The set \( W_d(C) \) of Weierstrass points of fixed degree \( d \) in \( X(\mathbb{C}) \) has the uniform probability measure converges to the Arakelov measure on \( X(\mathbb{C}) \) as \( d \) goes to infinity.

2. The Weierstrass divisor \( W_d \) will contain a vertical component wih \( d \gg 0 \), if the Jacobian of \( X \) has complex multiplication.

The results of Burnol are generalized in [Zh4] to arbitrary subvariety \( X \) of \( A \) such that \( A \) is generated by \( \{ x-y : x, y \in X \} \) and that the morphism \( \text{NS}(A) \to \text{NS}(X) \) is not injective. In this case, we can prove the Bogomolov conjecture by applying the Faltings' Hodge index theorem. In curves case, this will imply the positivity \( c_1(\Omega_{X/B}^1)^2 \) when \( \text{End}(\text{Jac}(X)) \to \mathbb{R} \) is not a division algebra. For example all modular curves of genus \( \geq 2 \) will satisfy this condition.

If \( \text{Jac}(X) \) does not have complex multiplication, Burnol's proof implies the following important fact: if \( c_1(\Omega_{X/B}^1)^2 = 0 \), then Weierstrass points will produce small points whose probability measure converges to the Arakelov measure. This turns to be a general property of small points [SUZ; Zh5] and can be easily deduced from the Theorem of Successive Minima:

**Equidistribution Theorem.** Let \( X \) be a subvariety of \( A \) defined over a number field \( K \) and let \( x_n \) be a sequence of points \( X \) which converges to the generic point of \( X \) and such that \( h(x_n) \to 0 \). Then the uniform probability measure of the Galois orbit \( O(x_n) \) tends to the measure \( dx := c_1(\mathcal{L}^\dim X/X/\text{deg}(X)) \) in the following sense: for any continuous function \( f \) on \( X(\mathbb{C}) \)

\[
\lim_{n \to \infty} \frac{1}{\#O(x_n)} \sum_{y \in O(x_n)} f(y) = \int_{X(\mathbb{C})} f(x) dx.
\]

To prove this, one just applies the right side of the inequality of successive minima

\[ h_\lambda(X) \leq e_{1,\lambda}(X) \]

where \( h_\lambda(X) \) and \( e_{1,\lambda}(X) \) are defined in the same way as \( h(X) \) and \( e_1(X) \) but with metric \( ||\cdot|| \) replaced by

\[ ||\cdot||_\lambda = ||\exp(\lambda f)\].

The final step for the proof of \( c_1(\Omega_{X/B}^1)^2 > 0 \) for general curve \( X \) is due to E. Ullmo [U]. His marvelous idea is to use the equidistribution theorem twice which will produce two different metrics and therefore produce a contradiction.
His construction is as follows: for the canonical embedding $X \to A = \text{Jac}(X)$, consider the induced map $\phi : X^g \to A$. If $c_1(\Omega^1_{X/B})^2 = 0$ then $X$ has a sequence $(x_n, n \in \mathbb{N})$ of distinct points such that $h(x_n) \to 0$, and the Galois orbits of these points will have probability measures converging to the Arakelov measure. Then one can produce a sequence $(y_n, n \in \mathbb{N})$ of $X^g$ such that

1. $y_n$ converges to the generic point of $X^g$;
2. $h(y_n)$ converges to 0;
3. The set $\{y_n : n \in \mathbb{N}\}$ is invariant under permutation action on $X^g$.

Then again the Galois orbits of $y_n$ will have probability measures converge to the product of the Arakelov measure on $X^g$. However, the sequence $\phi(y_n)$ in $A$ satisfies the condition of the Equidistribution Theorem, therefore the corresponding probability measure converges to the Haar measure on $A$. It follows that the product measure of Arakelov measure on $X^g$ is the pullback of the Haar measure on $A$. This is impossible: as the map $\phi$ is non-smooth, the pullback of the Haar measure as a differential form will vanishes along the singular locus of $\phi$.

Ulmo's idea is generalized to prove the general Bogomolov conjecture in [Zh5] by a modified Faltings' construction in [F2]: first it is easy to reduce the Bogomolov conjecture to the case that $X$ has the trivial Ueno fibration:

$$\{x \in A : x + X = X\}$$

is finite. Then for any positive integer $m$ we consider the map:

$$\alpha_m : X^m \to A^{m-1}$$

$$\alpha_m(x_1, \ldots, x_m) = (x_1 - x_2, \ldots, x_{m-1} - x_m).$$

Then one can show that for $m$ large, $\phi_m$ will induce a birational but not smooth map $X^m \to \alpha_m(X^m)$. Now we apply the equidistribution theorem to maps: $X^m \to A^m$ and $\alpha_m(X^m) \to A^{m-1}$ for sequences of small points $(y_n, n \in \mathbb{N})$ and $(\alpha_m(y_n), n \in \mathbb{N})$. Then we obtain the equality of two forms $\alpha$ and $\beta$ on $X^m$ induced respectively from the map $X^m \to A^m$ and $X^m \to A^{m-1}$. But this is impossible as $\beta$ vanishes along the singular locus of $\alpha_m$.

Combining the Bogomolov conjecture and the equidistribution theorem, one obtains the following stronger statement about small points in probability measure rather than Zariski topology:

**Theorem B.** Let $(x_n, n \in \mathbb{N})$ be a sequence of points in $A(\mathbb{Q})$ such that the following conditions are verified:

1. There is no subsequence of $(x_n, n \in \mathbb{N})$ contained in a translation of proper abelian subvariety by a torsion point;
2. $\lim_{n \to \infty} h(x_n) = 0$.

Then the probability measures of the Galois orbits of $x_n$ converge to the Haar measure of $A(\mathbb{C})$.

One can use this to show that the set of torsion points on $A$ over the maximal totally real fields is finite; this has been previously proved by Zarhin [Za] by using Faltings' theorem on Tate's conjecture.
REMARKS:
1. There are different approaches to the Bogomolov conjecture other than Arakelov theory. Notably, the diophantine approximation method used in David and Philippon [DP] and Bombieri and Zannier [BZ] produce the lower bound for $\lambda(X)$ effectively in terms of the degree of $X$.
2. The Bogomolov conjecture also has analogues for multiplicative groups $\mathbb{G}_m^n$ (or even a dynamical system [Zh4]). For $\mathbb{G}_m^n$, the Bogomolov conjecture is proved in [Zh3] and the equidistribution theorem is proved by Bui [Bl]. His approach is very original.

4. A CONJECTURE

Let $A \to C$ be a family of abelian varieties over a curve (may be open) over $\mathbb{Q}$. Let $\Lambda$ be a finitely generated torsion free subgroup of $A(C)$. Let $\mathcal{L}$ be a relatively ample symmetric line bundle on $A$ which induces Neron-Tate height pairings $< >$ on $A_{\mathbb{Q}}(\mathbb{Q})$ for each point $x \in C(\mathbb{Q})$. We define the number $h_{\Lambda}(x)$ for each $x \in C(\mathbb{Q})$ by the formula:

$$h_{\Lambda}(x) = \det(< t_i(x), t_j(x) >)$$

where $\{t_1, t_2, \cdots \}$ is a basis of $\Lambda$ and $t_i(x)$ is the specialization of $t_i$ in $A_{\mathbb{Q}}$.

CONJECTURE. Assume that the generic fiber $A_{\eta}$ of $A$ over $C$ is geometrically simple and has dimension $\geq 2$. Then there is a positive number $\epsilon$ such that the set

$$\{s \in C(\overline{\mathbb{Q}}) : h_{\Lambda}(s) < \epsilon\}$$

is finite.

REMARKS:
1. If $A = A_0 \times C$ is constant family with fiber $A_0$ and $\Lambda$ is generated by the graphs of one embedding $C \to A_0$ then the above conjecture is the Bogomolov conjecture for the embedding $C \to A_0$.
2. If $A = A_0 \times C$ is a constant family, $e_1$ is the graph of an embedding, but $e_i (i > 1)$ are graphs of constant maps $C \to x_i \in A_0$ whose images $a_i$ generate $A(K)$ modulo torsion, where $K$ is a number field over which $C \to A$ is defined, then the above conjecture implies the Mordell-Lang Conjecture for $e_1(C) \subset A$, as

$$h_{\Lambda}(x)^{1/2} = \text{distance}(x, \Gamma \otimes \mathbb{Z} / \mathbb{R}) / \text{volume}(\Gamma)$$

where the distance is taken in $A_{\mathbb{Q}}(\mathbb{Q}) \otimes \mathbb{R}$ with respect to the Neron-Tate height, and $\Gamma$ is the lattice generated by $a_i$. A related conjecture has been formulated independently by Poonen [P] and proved by him in some special cases.
3. The dimension assumption in the Conjecture is necessary as Poonen showed to me the following argument: if $A \to C$ has relative dimension one and $\Lambda$ is generated by one section $s$ then $s \cap A[N]$ will have a lot of intersection as $N \to \infty$, unless either $s(q)$ is a torsion point at the generic fiber, or $A \to C$ is a constant family and $s$ is a constant section. Of course, the simplicity condition in the conjecture could be removed if we require that for any geometrically
simple component $B$ of $A$, either $\dim B_n \geq 2$ or $B$ is a constant family and $B \cap \Lambda$ consists of constant sections.

4. Besides the case considered as above, the next two special cases are also interesting:
   a. $\phi : A \to C$ is not a constant family and $\Lambda$ is generated by one section; The conjecture implies that if $\phi(q)$ is not torsion, then there are only finitely many points $x \in C$ such that $\phi(x)$ is torsion.
   b. $A = A_0 \times C$ but $\Lambda$ is generated by the graphs of two embeddings $\phi_1 : C \to A_0$. The conjecture implies that if $\phi_1(q)$ and $\phi_2(q)$ are linearly independent, then there are at most finitely many $x \in C$ such that $\phi_1(x)$ and $\phi_2(x)$ are linearly dependent. A different formulation is as follows: the wedge product $\phi_1 \wedge \phi_2$ defines a map

   $$C(\mathbb{Q}) \to \wedge^2 A(\mathbb{Q})$$

   $$x \to \phi_1(x) \wedge \phi_2(x).$$

   Then the height $h_\Lambda$ is induced by the norm on $\wedge^2 A(\mathbb{Q})$. So we have a Bogomolov type conjecture for small points in $\wedge^2 A(\mathbb{Q})$.

5. The height $h_\Lambda$ is unlikely the Weil height for some positive line bundle on $C$, but the Northcott type theorem follows from some works of Silverman [Si] on the specializations of heights in the case that $A$ over $C$ has no fixed part.

REFERENCES


Shou-Wu Zhang
Department of Mathematics
Columbia University
New York City, New York 10027
U. S. A.