Proceedings of the International Congress of Chinese Mathematicians

Geometry of algebraic points

Shouwu Zhang

Introduction

In this note, I will try to give an elementary introduction to the subject of Diophantine geometry for people who are not familiar with either number theory or algebraic geometry. My basic goal is to explain first of all, the important question and their current solutions, and secondly how geometry (both in language and technique) is used to address and answer these questions. This note consists of three parts: rational points, algebraic points, and the ABC and discriminant conjectures. We now informally discuss the contends of each of these sections.

The language used in the first part is completely elementary. Thus we don't assume any advanced knowledge of algebraic geometry. The basic question is how to solve Diophantine equations in two variables in integers? Equivalently, in the language of algebraic geometry, how to find rational poits on a plane algebraic curve? It turns out that the nature of the solution sets really depends on the degree of the equation. In the degree 1 and 2 case, once one solution is found, then all other solutions can be found by using a *ruler*. This ruler method has an obstruction for curves with higher degree which are precisely given by a new object, *the Jacobian variety*, and the abelian group structure on the rational solution on the Jacobian variety. This Jacobian variety is of fundamental importance in the study of Diophantine questions related to curves. The *Mordell-Weil Theorem* and *Faltings Theorem* give the finiteness of rational points on these Jacobian varieties and the original curves respectively. Besides algebraic geometry, the proof of these theorems uses the *theory of heights*. The heights are quantities that measure the complexity (= logarithmic size) of points and varieties. A crucial question which remains today is the effectivity of solutions: *what is the maximal size of the solutions?* We will present so called effective Mordell conjecture at the end of this section.

In the second part, we will study certain sets of *algebraic points* in the *abelian varieties* (a generalized notion of Jacobian varieties). First of all, we restrict ourselves to either torsion points, or division points which is a mixed notion of torsion points and rational points. (Notice that the structure of the rational solutions is the subject in the first part.) *Raynaud's Theorem* shows that Faltings' theorem can be extended to these division points. Secondly, we consider to points which

Research supported by a grant from the National Science Foundation, and a Fellowship from the Sloan Foundation.

are close to division points, which we call *almost division points.* The new phenomenon is that the Galois orbits of these points in abelian varieties are *uniformly distributed, or equidistributed* with respect to the invariant probability measure on the associated complex torus. Of course the property to be equidistributed is much stronger than the property to be *Zariski dense.* (In the curves case, this means to be *infinite.*) Thus, the equidistribution property implies Raynaud's result. (Both Raynaud's result and equidistribution result extend Faltings' theorem but do not imply it as their proof use Faltings' theorem.) Like the study of rational points, the study of almost division points relies on a renovated theory of heights: *Arakelov geometry.* In this theory, the *Grothendieck's theory of schemes* is naturally combined with the *differential geometry of Kähler manifolds* to provide a very fine tool to study algebraic points with small heights and integral sections with small norms.

In the last part, we will discuss the *ABC – conjecture* which provides effective bound on the heights of rational points for a variety of equations, including diagonal equations like the Fermat equation. Moreover, the work of Szpiro, Moret-Bailly, and Elkies suggests that some strong form of the ABC conjecture is actually equivalent to the effective Mordell conjecture discussed in the first part of this note. The ABC conjecture originated in *Szpiro's discriminant conjecture* via Frey's construction. The analog of the discriminant conjecture in function field is a theorem of Szpiro. The original proof of this theorem makes essential use of the differentials of the function field. Thus it is impossible to move to the number field case. In this last part of the note, we will provide a different proof for the case where the function field has characteristic 0. (The original ideal of this proof is due to Bogomolov etc., at least for case where the base curves are rational.) Our proof does not use the differential of the function field but make essential use of the fact that the fundamental group of the base acts on the cohomology of the generic fiber with coefficients in *integers.* Such an action does exist in number fields (or function fields of positive characteristic) but with coefficients in *ℓ-adic integers.* Thus our proof extends to neither the number field nor function field case of positive characteristic.

1. Rational Points

One of the major goals of number theory is to study the solution set of a system of Diophantine equations. The simplest case is when this system is given by a single equation in two variables:

$$f(x,y)=0 \tag{1.1}$$

where $f(x,y)=\sum_{i,j}a_{i,j}x^iy^j$ is an polynomial with integer coefficients $a_{i,j}\in\mathbb{Z}$ in the variables x and y. Thus the main problem is to describe the set

$$\left\{(x,y)\in\mathbb{Z}^2 \mid f(x,y)=0\right\}.$$

To describe what we know about this question, we consider the homogenized form

$$F(x_0,x_1,x_2):=\sum_{i,j}a_{i,j}x^iy^jz^{n-i-j}=0, \tag{1.2}$$

where $n=\deg f$, and assume for simplicity that

$$\frac{\partial F}{\partial x}=\frac{\partial F}{\partial y}=\frac{\partial F}{\partial z}=0 \tag{1.3}$$

has only the trivial solution $(x, y, z) = (0, 0, 0)$ in $\mathbb{C}^3$. Now the equation (1.2) defines a curve X in the projective plane $\mathbb{P}^2$. Here, for any natural number n, the n-dimensional projective space is defined by

$$\mathbb{P}^n(\mathbb{C}) := \mathbb{C}^{n+1} - \{0\}/\sim, \quad \text{where} \quad v_1 \sim v_2 \Longleftrightarrow \mathbb{C}v_2 = \mathbb{C}v_1.$$

Condition (1.3) says that X is a smooth curve. Now the question becomes *how to describe the set $X(\mathbb{Q})$ of rational points on X?*

Here $X(\mathbb{Q})$ stands for points with rational coordinates. The answer to this question depends on the degree n of F.

Case where $n = 1$. In this case, X is defined by a linear equation

$$ax + by + cz = 0$$

with one nonzero coefficient, say a. Thus we can define a map

$$X(\mathbb{Q}) \xrightarrow{\sim} \mathbb{P}^1(\mathbb{Q}), \quad (x, y, z) \mapsto (y, z)$$

Case where $n = 2$. In this case, $X(\mathbb{Q})$ could be empty. For example, if X is defined by the equation $x^2 + y^2 + z^2 = 0$. Assume that $X(\mathbb{Q})$ is not empty and choose an element $O = (\alpha, \beta, \gamma)$. Without loss of generality, we assume that $\alpha \neq 0$. Then we have a map

$$X(\mathbb{Q}) \xrightarrow{\sim} \{\text{rational lines} \ni O\}, \qquad P \mapsto \text{line } OP,$$

where OP stands for the line passing through both P and O if $P \neq O$, and the tangent line passing through O if $P = O$. Notice that each line passing through O is defined by an equation

$$L_{a,b,c} : ax + by + cz = 0 \quad \text{with} \quad a\alpha + b\beta + c\gamma = 0.$$

Since $\alpha \neq 0$, we have a map

$$\{\text{rational lines} \ni O\} \xrightarrow{\sim} \mathbb{P}^1(\mathbb{Q}), \qquad L_{a,b,c} \mapsto (b, c).$$

Thus in case $n = 2$ we have the following conclusion:

$$\text{either } X(\mathbb{Q}) = \emptyset \text{ or } X(\mathbb{Q}) \xrightarrow{\sim} \mathbb{P}^1(\mathbb{Q}).$$

Hasse's principal. When does $X(\mathbb{Q})$ contain a rational point? By Hasse's principal, to see if $X(\mathbb{Q}) \neq \emptyset$ one needs only check if $X(\mathbb{R}) \neq \emptyset$, and if $X(\mathbb{Q}_p) \neq \emptyset$ for every *odd prime p* dividing the discriminant of the quadratic form $F(x_0, x_1, x_2)$. More precisely, after some linear transformation with coefficients in $\mathbb{Q}$, we may assume that X is defined by an equation:

$$ax^2 + by^2 = z^2. \tag{1.4}$$

where a and b are two square free integers. For every odd prime p, we define the Hilbert symbol $(a, b)_p \in \{\pm 1\}$ which equals 1 if and only if (1.4) has solution for every power of p. Then $X(\mathbb{Q}) \neq \emptyset$ if and only if

1. one of a and b is positive,
2. $(a, b)_p = 1$ for all odd prime number p dividing ab.

For $a = \alpha p^m$ and $b = \beta p^n$, one has the formula

$$(a, b)_p == \left(\frac{\alpha}{p}\right)^m \left(\frac{\beta}{p}\right)^n \left(\frac{-1}{p}\right)^{mn}.$$

Where $\left(\frac{x}{p}\right) = \pm 1$ is the residue quadratic symbol, which equals 1 if and only if x is a square modulo p. Using the quadratic reciprocity law, one can compute $(a, b)_p$ quite effectively. (The reason I don't care about the prime 2is because of the product formula for Hilbert symbols. Thus above two conditions together will imply that $(a, b)_2 = 1$.)

For more details about quadratic forms, we refer to Serre's book [**4**].

Case where $n = 3$. Again $X(\mathbb{Q})$ could be empty, for example in the case where X is defined by equation

$$3x^3 + 4y^3 + 5z^3 = 0.$$

Assume $X(\mathbb{Q}) \neq 0$ and pick up a point $O \in X(\mathbb{Q})$ inside, then $X(\mathbb{C})$ has a unique algebraic group structure such that O is the unit element. (For example, after certain transformation, we may assume that O is a *reflection point* on X- the tangent line at O does not meet other points on X, then we may define the group rule so that $P + Q + R = 0$ whenever P, Q, R are three collinear points in $X(\mathbb{C})$.) The curve X together with the point O is called an elliptic curve. We let E denote X with this distinguished point O.

As a Lie group, $E(\mathbb{C})$ has dimension 2 and is isomorphic to $(\mathbb{R}/\mathbb{Z})^2$. But as an abstruct group, $E(\mathbb{C})$ is very large. Indeed,

- the subgroup $E(\mathbb{C})_{\text{tor}}$ of elements of finite order is isomorphic to $(\mathbb{Q}/\mathbb{Z})^2$,
- the non torsion part $E(\mathbb{C})/E(\mathbb{C})_{\text{tor}}$ is an $\mathbb{Q}$-vector space of infinite dimension.

It is not difficult to show that $E(\mathbb{Q})$ is closed under the group operation, and thus defines a subgroup of $E(\mathbb{C})$. The following theorem tells us the structure of this group:

THEOREM 1.1 (Mordell). *The group $E(\mathbb{Q})$ is finitely generated, or equivalently, $E(\mathbb{Q})_{\text{tor}}$ is finite and*

$$E(\mathbb{Q})/E(\mathbb{Q})_{\text{tor}} \simeq \mathbb{Z}^r$$

where r is a nonnegative integer.

Infinite decent. The proof this theorem uses *infinite decent*, a technique used in Fermat's own proof of his last theorem for the exponent is 4. In our case, this technique is a combination of Kummer's theory and Neron-Tate height theory. More precisely, the proof has two steps:

1. $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite,
2. $E(\mathbb{Q})_{\text{tor}}$ is finite and $E(\mathbb{Q})$ is *discrete* in $E(\mathbb{Q}) \otimes \mathbb{R}$.

The second step implies that the rank of $E(\mathbb{Q})/2E(\mathbb{Q})$ over $\mathbb{Z}/2\mathbb{Z}$ is equal to or great than the rank of $E(\mathbb{Q})$ over $\mathbb{Z}$.

To prove the first step, one embeds $E(\mathbb{Q})/2E(\mathbb{Q})$ into the Galois cohomology group $H^1(\mathbb{Q}, E[2])$ using Kummer's sequence,

$$0 \longrightarrow E[2] \longrightarrow E \xrightarrow{2\times} E \longrightarrow 0$$

where $E[2]$ is the subgroup of 2-torsion points in $E(\mathbb{C})$ which is actually algebraic and thus admits an action from the Galois group over $\mathbb{Q}$. Then one finds that the image is unramified except for a finite set of primes. Thus the image is finite, so is $E(\mathbb{Q})/2E(\mathbb{Q})$.

For the second step, let us assume that O is a reflection point on X. Then we define the height of a point $p \in E(\mathbb{Q})$ by

$$h(P) = \log\max\{|a_i| : \quad i = 0, 1, 2\}$$

where P has homogeneous coordinates a_0, a_1, a_2 which are relatively prime integers. Then one can show that this function on $E(\mathbb{Q})$ is almost quadratic: the function $h(2p) - 4h(p)$ on $E(\mathbb{Q})$ is bounded. Thus one can normalize this height by taking limit:

$$\widehat{h}(p) := \lim_{n\to\infty} 4^{-n} h(2^n p).$$

Now it is easy to show that $\widehat{h}$ on $E(\mathbb{Q})$ satisfies the following properties:

1. for any positive integer H, the set of points $p \in E(\mathbb{Q})$ with height $h(p) \leq H$ is finite;
2. $\widehat{h}(x)$ is nonnegative on $E(\mathbb{Q})$, and $h(x) = 0 \Longleftrightarrow x \in E(\mathbb{Q})_{\mathrm{tor}}$;
3. $\widehat{h}(p)$ is quadratic on $E(\mathbb{Q})$.

By these properties, one can show that $E(\mathbb{Q})_{\mathrm{tor}}$ is finite and $\widehat{h}$ induces a positive and quadratic norm on $E(\mathbb{Q}) \otimes \mathbb{R}$. Again, the first property implies that $E(\mathbb{Q})/E(\mathbb{Q})_{\mathrm{tor}}$ can't have limit point in $E(\mathbb{Q}) \otimes \mathbb{R}$. Thus $E(\mathbb{Q})$ is discrete in $E(\mathbb{Q}) \otimes \mathbb{R}$.

Tate-Shafarevich group. When is $X(\mathbb{Q})$ nonempty? Unlike the case where $n = 2$, Hasse's principal is false in the case $n = 3$. Indeed, by a theorem of Hasse, every X has solution in $\mathbb{R}$ and $\mathbb{Q}_p$ for every prime number p. There is even an analogue of Galois theory which gives a group that allows one to measure how difficult for X to have a solution. Indeed, one may assign to X an elliptic curve E, the Jacobian of X, which acts on X. This makes X a principal homogeneous space. Thus E has a rational point over $\mathbb{Q}$ and E is isomorphic to X over $\mathbb{C}$. Let Ш denote the group of isomorphic classes of principal homogeneous spaces of E over $\mathbb{Q}$ which have rational points over $\mathbb{R}$ and $\mathbb{Q}_p$ for every prime p. Then X defines an element in Ш. The group Ш is torsion and is called the *Tate-Shafarevich group*. It may be nontrivial but is conjectured to be finite.

For more details about elliptic curves, we refer to Silverman's book [**5**].

Case where $n \geq 4$. In this case, $X(\mathbb{C})$ is not a Lie group any more. But it can be embedded into a complex Lie group J of dimension $g = (n-1)(n-2)/2$ which is called the Jacobian of X. Thus, as a real Lie group,

$$J(\mathbb{C}) \simeq (\mathbb{R}/\mathbb{Z})^{2g}, \qquad J(\mathbb{C})_{\mathrm{tor}} \simeq (\mathbb{Q}/\mathbb{Z})^{2g}.$$

The group J is an *abelian variety*, this means that

1. J is defined by some equations with rational coefficients in a higher dimensional projective space,
2. the origin of J has coordinates in $\mathbb{Q}$,
3. the addition operator is defined by functions with coefficients in $\mathbb{Q}$.

Thus, we can talk about the set $J(\mathbb{Q})$ rational points, which is again a subgroup of $J(\mathbb{C})$. The natural generalization of Mordell's theorem is the following:

THEOREM 1.2 (Weil). *The group $J(\mathbb{Q})$ is finitely generated. More generally, for any number field F, $J(F)$ is finitely generated.*

The set $X(\mathbb{Q})$ could be empty or nonempty. If it is nonempty and contains a point O then we can define an algebraic morphism $X \to J$ which takes O to be the unit element. This morphism takes rational points to rational points. In this way we have a picture:

$$X(\mathbb{Q}) = X(\mathbb{C}) \cap J(\mathbb{Q}) \tag{1.5}$$

THEOREM 1.3 (Faltings). *The set $X(\mathbb{Q})$ is finite. More generally, for any subvariety Y of an abelian variety A defined over a number field F which is not a translate of an abelian subvariety, $Y(F)$ is not Zariski dense.*

The Jacobian variety J as a generalization of an elliptic curve, is a very important companion of the curve X. There are two proofs of Faltings' theorem for the curves case. In this case, it was called *Mordell's conjecture*.

The first proof which works only for the case of curves uses the fact that J is faithful to X in the sense that if two curves have the same Jacobian then these two curves should be same. Thus, using some construction of Parshin, Mordell's conjecture is reduced to *Shafarevich conjecture* for abelian varieties. This can be further reduced to *Tate's conjecture* and Weil's theorem for the *Riemann-Hypothesis for Abelian varieties over finite fields.* Many steps of the proof has been reduced to some boundedness of *Faltings' heights* for Abelian varieties with semiabelian reductions. For more details, we refer to Cornell and Silverman's book [**2**] for the first proof and background material.

The second proof, originally due to Vojta for the case of curves and extended to the general case by Faltings, uses the fact that J includes X as a subvariety, (i.e., the formula (1.5)), and performs the Diophantine approximation method on J. The basic technique includes finding small sections of a hermitian line bundle which is not very vanishing at a given point. (This will implies that the bundle has small degree at the given point.) Of course, as Roth did in the classic case, the lower vanishing index can be reached only by considering the product of varieties. Again the heights for sub varieties of an abelian variety has been defined and used in the whole proof. For the second proof, we refer to Faltings' original paper [**3**]

Effectivity. How to effectively find $X(\mathbb{Q})$? To make this question more precise for $n \geq 4$, we define the heights of X and points P in $\mathbb{P}^2(\mathbb{Q})$ as follows:

$$\begin{aligned} h(X) = &\text{complexity of } X \\ = &1 + \log \max\{|a_{i,j}| : \quad i, j \geq 0\}, \end{aligned}$$

It is the general belief that the following should be true:

CONJECTURE 1.4 (Effective Mordell). *For each natural number $n \geq 4$, there exists a positive number $c(n)$ such that all points in $X(\mathbb{Q})$ have heights bounded by $c(n)h(X)$.*

We refer to Vojta's book [**7**] for a detailed discussion for conjectured bounds of rational points.

2. Algebraic points

Let $\bar{\mathbb{Q}}$ be the field of algebraic numbers in $\mathbb{C}$. For a projective space $\mathbb{P}^N$, let $\mathbb{P}^N(\bar{\mathbb{Q}})$ denote the points P with some homogeneous coordinates $(a_0, \cdots, a_N)$ in $\bar{\mathbb{Q}}^{N+1}$. Similarly we may define $Y(\bar{\mathbb{Q}})$ for sub varieties Y of $\mathbb{P}^N$. Let Y be a subvariety of an abelian variety A defined over a number field F in $\bar{\mathbb{Q}}$. For example,

Y is a curve of genus $g \geq 2$ and A is the Jacobian variety of Y. *What can we say about $Y(\bar{F})$ and $A(\bar{F})$?*

The first answer to this question, which is almost trivial, is that both sets are very large. First of all, $Y(\bar{F})$ is infinite. For example, when Y is defined by $f(x, y) = 0$, one may solve x in $f(x, y) = 0$ for any given $y \in \bar{F}$. Secondly, since the operations on A are defined by algebraic equations, and since the origin is an algebraic point we see that all torsion points of $A(\mathbb{C})$ are in $A(\bar{F})$. Thus

$$A(\bar{F})_{\mathrm{tor}} = A(\mathbb{C})_{\mathrm{tor}} \simeq (\mathbb{Q}/\mathbb{Z})^{2g}.$$

Finally, one may show that $A(\bar{F})/A(\bar{F})_{\mathrm{tor}} = A(\bar{F}) \otimes \mathbb{Q}$ is infinite dimensional. To give some less trivial answer to this question, we have to restrict ourselves to some special class of points in $A(\bar{F})$.

Division points. A point P in $A(\bar{F})$ is called a division point if some positive multiple of it nP is in $A(F)$. Let $A(\bar{F})_{\mathrm{div}}$ denote the subgroup of $A(\bar{F})$ of division points, then we have an exact sequence

$$0 \to A(\bar{F})_{\mathrm{tor}} \to A(\bar{F})_{\mathrm{div}} \to A(F)_F \to 0$$

where $A(F)_{\mathbb{Q}}$ stands for $A(F) \otimes \mathbb{Q}$. A natural generalization of Faltings theorem is the following

Theorem 2.1 (Raynaud). *If Y is a subvariety of A defined over a number field F which is not a translate of an abelian variety by a torsion point, then $Y(\bar{F}) \cap A(\bar{F})_{\mathrm{div}}$ is not Zariski dense in Y.*

Corollary 2.2. *In particular, the set $Y(\bar{F}) \cap A(\bar{F})_{\mathrm{tor}}$ is not Zariski dense.*

It is obvious that Raynaud's theorem is more general than Faltings' theorem and the above Corollary. But Raynaud actually first proved the above corollary and then used some Galois theory argument to obtain his theorem from this corollary and Faltings' theorem.

Almost division points. Now we want to treat points which are very close to division points with distance defined by the Neron-Tate height pairing on $A(\bar{F})$. Recall that the Neron-Tate height pairing on $A(\bar{F})$ is a bilinear, symmetric pairing:

$$\langle \cdot, \cdot \rangle : \quad A(\bar{F}) \times A(\bar{F}) \to \mathbb{R} \tag{2.1}$$

such that

1. $\langle x, x \rangle \geq 0$,
2. $\langle x, x \rangle = 0$ if and only if $x \in A(\bar{F})_{\mathrm{tor}} = (\mathbb{Q}/\mathbb{Z})^{2g}$.

Thus the pairing in (2.2) defines a pre-Hilbert space structure on the $\mathbb{Q}$-vector space $A(\bar{F})/A(\bar{F})_{\mathrm{tor}} = A(\bar{F}) \otimes \mathbb{Q}$. Now for any two points x and y in $A(\bar{F}) \otimes \mathbb{Q}$ we can define the distance

$$||x - y|| = \langle x - y, x - y \rangle^{1/2}.$$

In particular for $x \in A(\bar{F})$, we can define its distance to be a division point:

$$d(x) = \inf \left\{ ||x - y|| : \quad y \in A(\bar{F})_{\mathrm{div}} \right\}.$$

So $d(x) = 0$ if and only if $x \in A(\bar{F})_{\mathrm{div}}$.

Raynaud's theorem can be generalized to the following

THEOREM 2.3 (Poonen, Zhang). *There is positive number $\epsilon > 0$ such that the subset*

$$\{x \in Y(\bar{F}) : \quad d(x) < \epsilon\}$$

is not Zariski dense in Y.

COROLLARY 2.4 (Bogomolov conjecture). *There is positive number $\epsilon > 0$ such that the subset*

$$\{x \in Y(\bar{F}) : \quad \|x\| < \epsilon\}$$

is not Zariski dense in Y.

Again, Theorem 2.3 clearly implies Faltings' theorem and the above Corollary 2.4. But all proofs of Theorem 2.3 we know actually make use of Faltings theorem and Bogomolov's conjecture. To prove Bogomolov's conjecture, one first develops a theory of positive line bundles in Arakelov geometry, then deduces an equidistribution theorem about Galois orbits of small points. We should mention that in some cases of curves, Szpiro makes first link between Bogomolov conjecture and Arakelov geometry, and then Ullmo finds a clever use of some equidistribution theorem. We refer to our paper [**8**] for some historic remarks about the proof of Bogomolov's conjecture.

Distribution of almost division points. If X contains one point $x \in A(\bar{F})$ then it contains every conjugate x^σ ($\sigma \in \mathrm{Gal}(\bar{F}/F)$). One ideal to prove Corollary 2.2 and 2.4 is that the Galois orbits of torsion points or small points tends to be uniformally distributed with respect to the Haar measure on $A(\mathbb{C})$. Thus X can't contain infinitely many such points. In the following we want to explain this ideal more precisely.

Let x_n ($n = 1, 2, \cdots$) be a sequence of points in $A(\bar{F})$. We say x_n is a sequence of *almost division points* if $d(x_n)$ converges to 0 as $n \to \infty$. We say that the Galois orbits

$$O(x_n) := \{x_n^\sigma : \quad \sigma \in \mathrm{Gal}(\bar{F}/F)\}$$

are *equidistributed* with respect to a measure $d\mu$ on $A(\mathbb{C})$, if the uniform probability measure

$$\frac{1}{\#O(x_n)} \sum_{y \in O(x_n)} \delta_y$$

converges to $d\mu$ as $n \to \infty$. Here, δ_y is the Dirac measure at y. Equivalently, this means that for any continuous function f on $A(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{\#O(x_n)} \sum_{y \in O(x_n)} f(y) = \int_{A(\mathbb{C})} f(y) d\mu(y).$$

THEOREM 2.5. *Let $(x_n, n \in \mathbb{N})$ be a sequence of almost division points. There is a subsequence $(y_n, n \in \mathbb{N})$ of $(x_n, n \in \mathbb{N})$, an abelian subvariety B of A defined over F, and a finite subset T of $A(\mathbb{C})$ such that the Galois orbits of y_n are equidistributed to the $B(\mathbb{C})$-invariant uniform measure on $T + B(\mathbb{C})$. Moreover, $B = 0$ if and only if there is a finite extension F' of F such that all x_n are rational over F'.*

Now it is easy to see the following

- Theorem 2.5 + Faltings' theorem $\Longrightarrow$ Theorem 2.3.
- Theorem 2.5 $\Longrightarrow$ Corollary 2.4.

But to prove Theorem 2.5, one must prove Bogomolov's conjecture first. See our paper [**9**] for more details.

Scatteredness of big points. What can we say about the points which are far away from division points? For a point $x \in A(\bar{F})$, instead of the distance $d(x)$ to $A(\bar{F})_{\text{div}}$ we consider the angle $\theta(x)$ to $A(\bar{F})_{\text{div}}$:

$$\theta(x) = \sin^{-1}(d(x)/||x||).$$

Then from Faltings' proof of Theorem 1.3 for higher dimensional case, we have

THEOREM 2.6. *Assume that Y is a subvariety of A defined over F which is not a translate of an abelian subvariety. Then there is a positive number ϵ such that the subset*

$$\left\{x \in Y(\bar{\mathbb{Q}}) : \quad \theta(x) < \epsilon\right\}$$

is not Zariski dense.

3. The ABC conjecture and the discriminant Conjecture

When discussing Diophantine equations, it is unavoidable to mention the *ABC-Conjecture:*

CONJECTURE 3.1 (Masser, Osterlé). *Let A, B, C be relative prime positive integer satisfying $A + B = C$, then for any $\epsilon > 0$,*

$$C \leq \kappa(\epsilon) \left(\prod_{p|ABC} p \right)^{1+\epsilon}$$

where $\kappa(\epsilon)$ is a number depending only on ϵ.

For application purposes, it suffices to have a *weaker form where $1+\epsilon$ is replaced by any fixed constant.*

There are two ways to link this conjecture to Diophantine equations. First of all this conjecture applies directly to the diagonal (or generalized Fermat's) equation:

$$aX^n + bY^n = cZ^n \tag{3.1}$$

where a, b, c are positive and relative prime integers, and n is a positive integer. The conjecture implies that there is a constant $C(a, b, c)$ such that the equation has no nontrivial solution if

$$n \geq C(a, b, c).$$

The second link is that this conjecture can be generalized to a form which holds for an arbitrary number field. Then the generalized conjecture is equivalent to a certain effective Mordell conjecture like Conjecture 1.4. This follows from some work of Szpiro, Moret-Bailly, and Elkies. We won't discuss this matter here.

We want to discuss a third link that relates the ABC-conjecture to an early conjecture of Szpiro on discriminants of elliptic curves.

Szpiro's conjecture. Recall that an elliptic curve E over $\mathbb{Q}$ is a projective curve of genus 1 with a rational point O. Any elliptic curve E can be defined by a Weierstrass equation

$$E_{\mathbb{Z}} : \qquad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.2}$$

where $a_i \in \mathbb{Z}$. The origin O is at infinity. One may use various affine transformations to obtain different form of Weierstrass equation. For such an equation one can define a discriminant Δ which a positive integer with the property that $p \mid \Delta$ if and only if equation (3.2) modulo p defines a singular curve. We say a Weierstrass equation is minimal if Δ is minimal. In this case, we call $E_{\mathbb{Z}}$ the minimal integral model of E and write Δ by Δ_E.

Recall also that an elliptic curve $E_{\mathbb{Z}}$ given by a minimal Weierstrass equation $f(x,y) = 0$ over $\mathbb{Z}$ has stable reduction at a prime p, if this equation modulo p defines a smooth curve (which we called a good reduction), or a curve with node (which we called a bad reduction). In the stable but bad reduction case, in a strict henselian neighborhood node, the definition equation of $E_{\mathbb{Z}}$ has the form $xy = p^{n_p}$, where n_p is a positive integer. If $E_{\mathbb{Z}}$ has stable reduction everywhere, the discriminant of $E_{\mathbb{Z}}$ has the form $\prod_p p^{n_p}$ where p runs through the set of bad reductions. Write $N = \prod_{p|\Delta} p$.

Conjecture 3.2 (Szpiro). *Let E be an elliptic curve defined over $\mathbb{Q}$ whose minimal model $E_{\mathbb{Z}}$ has stable reduction everywhere. Then for any $\epsilon > 0$,*

$$\Delta_E \leq \kappa(\epsilon)(N_E)^{6+\epsilon}$$

where $\kappa(\epsilon)$ is a constant depending only on ϵ.

Again for application, it suffices to have a *weaker form where $6+\epsilon$ is replaced by any fixed constant ϵ.* The link between the ABC-conjecture and Szpiro's conjecture is given by the following construction of Frey.

Frey curve. Let A, B, C be three relative prime, positive integers such that $A+B = C$, and such that $16 \mid A$ and $B \equiv -1 \pmod 4$. Consider the curve $E_{A,B,C}$ defined by the equation

$$y^2 = x(x+A)(x-B).$$

Frey proves the following:

1. The minimal Weierstrass equation of $E_{A,B,C}$ is given by
$$y^2 + xy = x^3 + \frac{A-B-1}{4}x^2 - \frac{AB}{4}x.$$
2. The above minimal model of $E_{A,B,C}$ has stable reduction modulo every prime p.
3. Under the hypothesis of (1), the discriminant of $E_{A,B,C}$ is equal to $(ABC/16)^2$.

If you apply Szpiro's conjecture then you immediately obtain the ABC-conjecture, at least in their weak forms.

Szpiro's Theorem. Szpiro proposed this based on his theorem over a function field. Let $f : E \to C$ be a proper and flat morphism from a surface E, smooth over algebraically closed field k, to a curve C, smooth and projective of genus g and geometrically connected over k. Suppose the generic fiber of f is a smooth and geometrically connected elliptic curve over the function field of C. Suppose in addition that f is not iso trivial and the degenerate fibers are stable. Then we have a non constant morphism

$$j : C \to \mathbb{P}^1_k$$

by taking j-invariants.

Let Δ_E denote the discriminant divisor of f, and let S denote the set of points on C over which f is not smooth. Let p is the characteristic of k and p^e is the degree of inseparability of the morphism j. (If k has characteristic 0, we write $p^e = 1$ for convenience).

Theorem 3.3 (Szpiro).

$$\deg \Delta_E \leq 6p^e(2g - 2 + \#S)$$

Szpiro's proof makes essential use of the derivative map induced by j on $C - S$. We refer to Szpiro's paper [**6**] for his conjecture and his theorem.

In the following we want to present a proof using monodromy action on homology. We will assume that $k = \mathbb{C}$, because we will make essential use of integral coefficients in homology. When the base C is a Riemann sphere, this proof is due to J. Amorós, F. Bogomolov, L. Katzarkov, and T. Pantev [**1**].

Monodromy action on homology. We let $f : E \to C$ be as above with $k = \mathbb{C}$. Fix a point p on C then we have a monodormy action

$$\rho : \qquad \pi_1(C - S, p) \longrightarrow \mathrm{SL}(H_1(E_p, \mathbb{Z}))$$

where the fundamental group means topological fundamental group and SL means automorphisms with determinant 1. Let U be a simply connected neighborhood in $C(\mathbb{C})$ of p in C containing S. For each $s \in S$, let $c_s \in \pi_1(C - S, p)$ be an element which can be represented by a simple loop around only s in U with positive orientation. There are a_i, b_j $(i, j = 1, \cdots, g)$ in $\pi_1(C - S, p)$ such that $\pi_1(C - S, p)$ is generated by a_i, b_j, c_s with a single relation

$$\prod_{i=1}^{g} [a_i, b_j] \cdot \prod_{s \in S} c_s = 1.$$

Recall that the intersection pairing gives an alternative pairing

$$\langle \cdot, \cdot \rangle : \qquad H_1(E_p, \mathbb{Z}) \otimes H_1(E_p, \mathbb{Z}) \longrightarrow \mathbb{Z}.$$

Fix a basis e_1 and e_2 such that $\langle e_1, e_2 \rangle = 1$. This defines an isomorphism $\mathrm{SL}(H_1(E_p, \mathbb{Z})) \simeq \mathrm{SL}_2(\mathbb{Z})$ which is unique up to conjugation by elements in $\mathrm{SL}_2(\mathbb{Z})$ (not only $\mathrm{GL}_2(\mathbb{Z})$!) Now we have the following equality in $\mathrm{SL}_2(\mathbb{Z})$:

$$\prod_{i=1}^{g} [\rho(a_i), \rho(b_j)] \cdot \prod_{s \in S} \rho(c_s) = 1. \tag{3.3}$$

We will actually show that Szpiro's inequality follows from (3.3) and the following monodromy theorem:

Theorem 3.4. *For each $s \in S$, $\rho(c_s)$ is conjugate in $\mathrm{SL}_2(\mathbb{Z})$ to*

$$\begin{pmatrix} 1 & n_s \\ 0 & 1 \end{pmatrix}$$

where $n_s = \mathrm{ord}_s(\Delta_E)$.

Proof: This is a well known theorem but we provide here a proof for the convenience of reader. Let q be the local parameter of C at s. We may enlarge q such that $D := \{q \in \mathbb{C}, |q| \leq 1/e\}$ is embedded in C. Let D^* denote $D - \{0\}$. Then the restriction of E on D^* is a Tate curve with the form

$$E^* \simeq D^* \times \mathbb{C}^\times / q^{n_s \mathbb{Z}}. \tag{3.4}$$

Obviously, the theorem does not depend on the choice of p in C. Thus we may assume that $p = 1/e$ and c_s is a simple loop in D^* with counter clockwise orientation:

$$c_s(t) = e^{2\pi i t - 1} \qquad (0 \leq t < 1).$$

By (3.4), the fiber E_t over $c_s(t)$ has the form

$$E_t = \mathbb{C}^* / e^{2\pi i t - 1} = \mathbb{C}/(\mathbb{Z} + (i+t)n_s\mathbb{Z})$$

Thus if we use $e_1 = 1$ and $e_2 = in_s$ as a base of $H_1(E_p, \mathbb{Z})$, then they are moved to 1 and $(i+t)n_s$, and they have limits 1 and $(i+1)n_s$ as $t \to 1$. Thus, we have shown that

$$c_s \cdot (e_1, e_2) = (e_1, n_s e_1 + e_2) = (e_1, e_2) \cdot \begin{pmatrix} 1 & n_s \\ 0 & 1 \end{pmatrix}.$$

Groups $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ and $\widetilde{\mathrm{SL}}_2(\mathbb{R})$. The action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{R}^2$ induces a topological action on

$$\mathbb{R}/2\pi\mathbb{Z} \simeq \mathbb{R}^2/\mathbb{R}^+$$

where $t \in \mathbb{R}/2\pi\mathbb{Z}$ corresponds to $(\cos t, -\sin t) \cdot \mathbb{R}^+$ and $\mathbb{R}^+$ is the group of positive numbers. Let $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ be the group of topological homeomorphism $\widetilde{\gamma}$ of $\mathbb{R}$ which induces an element γ in $\mathrm{SL}_2(\mathbb{R})$. Thus for all $t \in \mathbb{R}$,

$$\widetilde{\gamma}(t) \mod 2\pi = \gamma(t \mod 2\pi).$$

Lemma 3.5. *The homomorphism from $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ to $\mathrm{SL}_2(\mathbb{R})$ induced by $\widetilde{\gamma} \to \gamma$ is surjective with kernel a free group generated by z^2, where $z(t) = t + \pi$. Thus we have an exact sequence:*

$$1 \to < z^2 > \to \widetilde{\mathrm{SL}}_2(\mathbb{R}) \to \mathrm{SL}(\mathbb{R}) \to 1.$$

Moreover $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ has center generated by z.

Proof: We need only prove that the homomorphism $\widetilde{\mathrm{SL}}_2(\mathbb{R}) \to \mathrm{SL}(\mathbb{R})$ is surjective. Let $\gamma \in \mathrm{SL}_2(\mathbb{R})$. We want to explicitely construct a *canonical lifting* $\widetilde{\gamma}$ which commutes with z.

Case where $\gamma = -I$. We take an obvious lifting $\widetilde{\gamma}(t) = z$.

Case where γ is parabolic. This means that γ has a fixed point θ on $\mathbb{R}/2\pi\mathbb{Z}$ or equivalently, γ has a positive eigenvalue. In this case, γ has a unique lifting $\widetilde{\gamma}$ fixing pointwise all points in the preimage of θ.

General case. We claim that γ is a product of two elements in the above cases. This is clear if γ is either upper triangular or lower triangular. If this is not the case, then there is a unique x such that

$$\det(\gamma - u_x) = 0, \quad \text{where} \quad u_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Thus $\gamma = \gamma_1 \cdot u_x$ where γ_1 has an eigenvalue equal to 1. So $\widetilde{\gamma_1}\widetilde{u_x}$ gives a lifting for γ.

It is well known that $\mathrm{SL}_2(\mathbb{Z})$ is generated by the following elements and relations:

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad v = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \qquad (uv)^6 = I, \qquad uvu = vuv$$

Let $\widetilde{A}$ and $\widetilde{B}$ be canonical liftings of A and B constructed in the proof of Lemma 3.5.

Lemma 3.6. *Let $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$ be the inverse image of $\mathrm{SL}_2(\mathbb{Z})$ in $\widetilde{\mathrm{SL}}_2(\mathbb{R})$. Then*

1. *$\widetilde{\mathrm{SL}}_2(\mathbb{Z})$ is generated by $\widetilde{u}$ and $\widetilde{v}$ with the relation $\widetilde{u}\widetilde{v}\widetilde{u} = \widetilde{v}\widetilde{u}\widetilde{v}$.*

2. $z = (\tilde{u}\tilde{v})^3$.

Proof: It can be checked directly that

$$0 \le \tilde{u}x - x < \frac{\pi}{3}, \qquad 0 \le \tilde{v}x - x < \frac{\pi}{3} \tag{3.5}$$

Thus

$$0 \le (\tilde{u}\tilde{v})^3 x - x < \frac{\pi}{3} \times 6 = 2\pi. \tag{3.6}$$

$$0 \le \tilde{u}\tilde{v}\tilde{u}x - x < \pi, \qquad 0 \le \tilde{v}\tilde{u}\tilde{v}x - x < \pi \tag{3.7}$$

As $(\tilde{u}\tilde{v})^3$ is a liftings of $-I$, then $(\tilde{u}\tilde{v})^3$ must be equal to kz with k an odd integer. By (3.6), $k = 1$, or equivalently $z = (\tilde{u}\tilde{v})^3$. Same reasoning gives the equality $\tilde{u}\tilde{v}\tilde{u} = \tilde{v}\tilde{u}\tilde{v}$.

As the subgroup $< \tilde{u}, \tilde{v} >$ contains the center $< z >$ and maps surjectively to $\mathrm{SL}_2(\mathbb{Z})$, this subgroup must be $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$. Since $\tilde{u}\tilde{v}$ has infinite order in $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$, $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$ is generated by $\tilde{u}$ and $\tilde{v}$ with only one relation $\tilde{u}\tilde{v}\tilde{u} = \tilde{v}\tilde{u}\tilde{v}$.

Proof of Szpiro's inequality. We start with a lifting of equation (3.3) to $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$. Let α_i, β_j, and γ_s be canonical liftings of some liftings of $\rho(a_i)$ and $\rho(b_j)$ and $\rho(c_s)$ respectively in $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$ as constructed in the proof of Lemma 3.5. Then we have a new equation in $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$:

$$\prod_{i=1}^{g} [\alpha_i, \beta_j] \cdot \prod_{s \in S} \gamma_s = z^{2m} \tag{3.8}$$

where m is some integer.

Lets determine m. By Lemma 3.6, there is a degree homomorphism

$$\deg : \quad \widetilde{\mathrm{SL}}_2(\mathbb{Z}) \longrightarrow \mathbb{Z} \quad \text{such that} \quad \deg \tilde{u} = \deg \tilde{v} = 1.$$

Applying this degree map to equation (3.8), we obtain $12m = \sum_{s \in S} \deg \gamma_i$. By Theorem 3.4, $\rho(c_s)$ is conjugate to u^{n_s}. As γ_i is the canonical lifting of $\rho(c_i)$, γ_s is conjugate to $\tilde{u}^{n_s}$ in $\widetilde{\mathrm{SL}}_2(\mathbb{Z})$. Thus $\deg \gamma_s = n_s$. So we obtain

$$12m = \sum_{s \in S} n_s = \deg \Delta_E. \tag{3.9}$$

For any $\alpha \in \widetilde{\mathrm{SL}}_2(\mathbb{R})$, define its *length* by

$$\ell(\alpha) = \sup_{x \in \mathbb{R}} |\alpha(x) - x| \in \mathbb{R}^+.$$

(The right hand side is finite, because $\alpha z = z\alpha$, or equivalently $\alpha(x+\pi) = \alpha(x) + \pi$.) This length function has properties:

$$\ell(\alpha\beta) \le \ell(\alpha) + \ell(\beta), \qquad \ell(z^n) = \pi n.$$

Thus equations (3.8) gives us

$$2m\pi \le \sum_{i=1}^{g} \ell\left([\alpha_i, \beta_i]\right) + \sum_{s \in S} \ell(\gamma_s). \tag{3.10}$$

LEMMA 3.7.

$$\ell(\gamma_s) < \pi, \qquad \ell\left([\alpha_i, \beta_i]\right) \le 2\pi.$$

Proof: The first one is easy, as γ_s has fixed points on $\mathbb{R}$ with period π. For the second one, from the construction in the proof of Lemma 3.5, upto a factor z which does not change the bracket, α_i is a product of two elements conjugate in $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ to elements $\widetilde{u}_x$ and $\widetilde{u}_y$. Thus $[\alpha_i, \beta_i] = \alpha_i \cdot (\beta_i \alpha_i \beta_i^{-1})$ is a product of four elements conjugate to $\widetilde{u}_x$, $\widetilde{u}_{-x}$, $\widetilde{u}_y$, $\widetilde{u}_{-y}$. Notice that two of these four elements will move points in positive direction, and two of them will move points in negative direction. All of them have length bounded by π. Thus $[\alpha_i, \beta_i]$ has length bounded by 2π.

By (3.9) and (3.10) we have $2m \leq 2g - 1 + s$. Combining with (3.8), we have

$$\deg \Delta_E \leq 6(s + 2g - 1) \tag{3.11}$$

This inequality is not as sharp as Szpiro's. But for $g > 0$, then one may apply this equality to an unramified base change

$$E' = E \times_C C' \to C'$$

where $\pi : C' \to C$ is an unramified extension. Then we have equalities

$$\begin{aligned} \deg \Delta_{E'} &= \deg \pi \cdot \deg \Delta_E, \\ 2g(C') - 2 &= \deg \pi \cdot (2g - 2), \\ s' &= \deg \cdot \pi s \end{aligned}$$

where s' is the number of bad fibers of $E' \to C'$. Now (3.11) for $E' \to C'$ gives

$$\deg \Delta_E \leq 6(s + 2g - 2 + 1/\deg \pi).$$

As π can be chosen such that $\deg \pi$ arbitrary large, one obtains Szpiro's inequality. However this argument doesn't apply to the case where $g = 0$.

References

[1] J. Amorós, F. Bogomolov, L. Katzarkov, and T. Pantev, *Symplectic Lefschetz fibration with arbitrary fundamental groups,* Preprint, Math.GT/9810042 v2 25 Nov 1998.

[2] G. Cornell and J. H. Silverman (eds), *Arithmetic geometry*, Springer-Verlag, 1986

[3] G. Faltings, *Diophantine approximation on abelian varieties,* Ann. Math. **133** (1991), pp. 549-576.

[4] J. -P. Serre, *A course in arithmetic,* GTM **7**, Springer-Verlag, New York, 1973

[5] J. H. Silverman, *The arithmetic of elliptic curves,* GTM **106**, Springer-Verlag, New York, 1986

[6] L. Szpiro, *Discriminant et conducteur des courbes elliptiques,* In: *Séminaire sur les pinceaux de courbes elliptiques: à la recherche de Mordell effectif.* Astérisque **183**, 1990.

[7] P. Vojta, *Diophantine approximations and value distribution theory,* LNM **1239**, Springer-Verlag, Berlin, 1987

[8] S. Zhang, *Small points and Arakelov theory,* Proceedings of ICM, Berlin 1998, Vol II, pp. 217-225.

[9] ———, *Distribution of almost division points,* to appear in Duke Math. J.

Department of Mathematics, Columbia University, New York, New York 10027
E-mail address: szhang@@math.columbia.edu