Point-counting and applications

Jonathan Pila

Lecture Notes

Arizona Winter School, 2023

Introduction

These lectures discuss point-counting and applications to diophantine problems. Though we don't get there until Lectrure 3, the main objective is to describe the point-counting approach to the unlikely intersection problem of a curve in a product of modular curves.

Lecture 1

Synopsis. The basic point-counting result (for "definable sets in an o-minimal structure", but deferring a discussion of this notion) and its simplest application to an "unlikely intersection" problem: describing the distribution of torsion points on a subvariety of $(\mathbb{C}^\times)^n$, a problem known as "Multiplicative Manin-Mumford".

Introduction. Diophantine geometry studies the distribution of rational points (and more generally points defined over number fields) on algebraic varieties. For curves one has good (geometric) criteria for finiteness of such points (Faltings proof of the Mordell conjecture). This result is quantitative but not effective: one can bound the number of rational points but not their height).

In higher dimensions one has very strong conjectures (Bombieri-Lang: an algebraic variety is "mordellic" outside its (geometrically defined) "special set"), but results are sparse. Some results assert that suitable algebraic varieties have "very few" rational points beyond the"obvious" ones. As such results do not assert finiteness, they are framed in terms of counting points up to some give *height* bound H.

1.1. Definition. The *height* of a rational number $q = a/b$ in lowest terms is $H(q) = \max\{|a|, |b|\}$.

For example, conjecturally, no positive integer can be written as a sum of two fifth powers in two different ways. That is, all positive integer solutions to the diophantine equation $w^5 + x^5 = y^5 + z^5$ are **trivial** in the sense that $\{w, x\} = \{y, z\}$. Up to height H there are $H^2 + O(H)$ trivial integer solutions.

1.2. Theorem. ([17]) *For $\epsilon > 0$ and $H \geq 1$ there are $\ll_\epsilon H^{13/8+\epsilon}$ non-trivial solutions to $w^5 + x^5 = y^5 + z^5$ in positive integers up to H.*

We will start by considering analogous results for rational points on non-algebraic (but suitable) sets. Here one needs some notion of "suitable" as one can hardly hope to prove meaningful things about rational points on arbitrary sets.

We will give a provisional notion of "suitable set" and defer a precise description of these sets "definable in an o-minimal structure over the real field" to the second lecture.

Counting result for curves. The basic one-dimension result is the following.

For a set $X \subset \mathbb{R}^n$ we define

$$X(\mathbb{Q}, H) = \{x \in X \cap \mathbb{Q}^n : H(x) \leq H\}$$

and the counting function

$$N(X, H) = \#X(\mathbb{Q}, H).$$

1.3. Theorem. ([16, 36]) *Let $f(x)$ be a non-algebraic function that is real analytic on an open neighbourhood of $[0,1]$, and let $X \subset \mathbb{R}^2$ be the graph of $f : [0,1] \to \mathbb{R}$. Let $\epsilon > 0$. Then there is a constant $c(f, \epsilon)$ such that*

$$N(X, H) \leq c(f, \epsilon) H^{\epsilon}.$$

The proof proceeds by showing that the points in question reside on "few" algebraic curves of degree $d = d(\epsilon)$. Here "few" means $\ll H^{\epsilon}$. This relies on a mean value theorem of H. A. Schwarz: all the points in a small subinterval of $[0,1]$ indeed lie on one such algebraic curve. But not too small: $[0,1]$ is covered by $\ll H^{\epsilon}$ such subintervals. As f is non-algebraic, the intersection $X \cap C$ is finite and of size uniformly bounded for all curves C of degree d. This gives the result.

Higher-dimensional sets. Moving to higher dimensional sets, one must set out some reasonable class of sets, as one cannot hope to get good estimates for arbitrary sets. Our theorem will apply to sets which are 'definable in an o-minimal structure over the real field' (henceforward for brevity we will call such a set "definable") but provisionally consider the image $X \subset \mathbb{R}^n$ of a function

$$f : [0,1]^k \to \mathbb{R}^n$$

that is real analytic on an open neighbourhood of $[0,1]^k$. Such a set we will see is definable, though not all definable sets are of this form.

Also, a higher dimensional (definable) set X may contain positive-dimensional real semi-algebraic sets (see below) even if X itself is non-algebraic.

1.4. Definition. A *semi-algebraic set* in $\mathbb{R}^n$ is a finite union of sets each of which is defined by finitely many equations and inequations between polynomials with real coefficients.

For example, the graph $X \subset \mathbb{R}^3$ of $z = x^y$ on $x, y \in [1,2]$, which is of the above-mentioned type. Such semi-algebraic subsets may contain "many" rational points.

1.5. Definition. Let $X \subset \mathbb{R}^n$. The *algebraic part* of X, denoted X^{alg} is the union $\bigcup A$ of all connected positive-dimensional semi-algebraic sets $A \subset X$. The *transcendental part* of X, denoted X^{trans}, is the complement in X of X^{alg}.

1.6. Theorem. ([43]) *Let $X \subset \mathbb{R}^n$ be definable in an o-minimal structure over the real field. Let $\epsilon > 0$. Then there is constant $c(X, \epsilon)$ such that*

$$N(X^{\text{trans}}, H) \leq c(X, \epsilon) H^{\epsilon}.$$

This is the basic point-counting result. It can be elaborated in various ways, for example to count algebraic points of some fixed degree rather than rational points. In general one cannot improve the bound $\ll_{\epsilon} H^{\epsilon}$ or make the constant $C(X, \epsilon)$ effective. But under more restrictive hypotheses one can hope to do either or both, and various results are known. See e.g. [8, 10].

Diophantine applications. Here we sketch the very simplest application of the counting theorem to a diophantine problem. It is part of a wider collection of result and conjectures.

Warning. Above we discussed sets in real Eudlicdean space and "dimension" refers to real dimension. Below we discuss complex algebraic varieties and "dimension" for them will be complex dimension. Further below we interact these pictures viewing $\mathbb{C}$ as $\mathbb{R}^2$ and considering complex analytic sets as real sets. So beware that "dimension" can refer to real or complex dimension depending on context!

The Mordell conjecture (1922) has already been mentioned. Proved by Faltings (1983), it asserts that a smooth projective curve V of genus $g \geq 2$, such as smooth plane quartic curve, has only finitely many rational points.

This conjecture fits into a more general conjectural framework, including the Modell-Lang conjecture (proved in work of a number of people starting with Faltings), and a much wider and very much open Zilber-Pink conjecture.

In the course of considering the conjectural picture, Lang considered the following problem around 1960. Let F be a Laurent polynomial in two variables (polynomial in X, X^{-1}, Y, Y^{-1}) and let

$$V = \{(x, y) \subset (\mathbb{C}^{\times})^2 : F(x, y) = 0\}.$$

We want to consider points on V that are roots of unity. Roots of unity are the torsion points in the multiplicative group $\mathbb{C}^{\times}$, hence it is natural to take the ambient variety to be the group $(\mathbb{C}^{\times})^2$ rather than $\mathbb{C}^2$.

1.7. Theorem. (Proved by Ihara-Serre-Tate) *The number of such points is finite except in the case that F is of the form $x^n y^m = \eta$ where $n, m \in \mathbb{Z}$ are not both zero and η is a root of unity.*

The number of such points is infinite in the exceptional cases. Such V is a coset by a torsion point of an algebraic subgroup $x^n y^m = 1$; such a set will be called a *torsion coset.*

Let us now consider an algebraic subvariety $V \subset X = (\mathbb{C}^{\times})^n$. Then X is an algebraic group, and denote by X_{tors} its set of torsion points (points whose coordinates are all roots of unity).

The algebraic subgroups of X are all defined by multiplicative conditions: some number of equations of the form $x_1^{k_1} \dots x_n^{k_n} = 1$, where $k_i \in \mathbb{Z}$, and the *torsion cosets*

are then the components of subvarieties defined by some number of equations of the form $x_1^{k_1} \dots x_n^{k_n} = \eta$ with η a root of unity. If zero-dimensional such a point is a torsion point.

The following theorem is a special case of the 'Multiplicative Mordell-Lang conjecture' proved by Laurent (1984); it also follows from earlier results of Mann on linear relations between roots of unity.

1.8. Theorem. *Let $V \subset X$ be an algebraic variety. There are finitely many torsion cosets $X_i \subset V$ that account for all the torsion points of X that are in V.*

The point-counting approach. This follows a strategy proposed by Zannier, implemented initially in [44, 31].

Consider the (modified) exponential map $e(z) = \exp(2\pi i z)$ and its n-fold cartesian power (which we will also denote e):

$$e : \mathbb{C}^n \to (\mathbb{C}^\times)^n, \quad e(z_1, \dots, z_n) = (e(z_1), \dots, e(z_n)).$$

The pre-images of torsion points under e are precisely rational points: studying torsion points on V can be approached by studying rational points on $e^{-1}(V) \subset \mathbb{C}^n$.

Now $e^{-1}(V)$ is not a definable set (where we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ using real and imaginary parts), due to the periodicity of e. The map e is invariant under the action of $\mathbb{Z}^n$ acting on $\mathbb{C}^n$ by translation. A fundamental domain for this action is the set

$$F = \{(z_1, \dots, z_n) \in \mathbb{C}^n : 0 \leq \mathrm{Re}(z_i) < 1, i = 1, \dots, n\}$$

The graph of the restriction of e to F is a definable set (in the structure known as $\mathbb{R}_{\text{an exp}}$; see Lecture 2) and (hence) so is

$$Z = e^{-1} \cap V.$$

Moreover, every point in V has a pre-image (and indeed a unique pre-image) in Z. So studying torsion points in V is the same as studying rational points in Z. A crucial fact will be that torsion points are algebraic points of quite high degree and hence with many Galois conjugates.

Note: Here the pre-images of torsion points are rational points and lie on the real line inside F. In general the pre-images of "special" points are dense in F.

By the order $N(\eta)$ of a root of unity η we mean its minimal order. By the complexity of a torsion point $(\eta_1, \dots, \eta_n)$ we mean the maximum order of its coordinates. The degree of a root of unity is $\phi(N)$, where ϕ is the Euler ϕ-function and N is its order. This is quite large: all the primitive roots are conjugate e.g. if $N = p$ is a prime number the degree is $p - 1$.

1.9. Theorem. (see e.g. Hardy and Wright) *Let $\delta > 0$. Then there is constant $c(\delta)$ such that a root of unity η has at least $c(\delta)N^{1-\delta}$.*

It will be necessary to have a description of the algebraic part of Z. Suppose Z contains some positive dimensional semi-algebraic subset A in the real coordinates. Then in fact $e^{-1}(V)$ contains a complex algebraic subset W containing A. This is by

analytic continuation because the map e is complex analytic. Hence Z^{alg} consists of (positive dimensional components of) $F \cap e^{-1}(V)^{\text{complex alg}}$.

And what is $e^{-1}(V)^{\text{complexalg}}$? We need to understand when we can have $e(W) \subset V$. The exponential map is highly transcendental, and usually $e(W)$ is Zariski dense in $(\mathbb{C}^\times)^n$. E.g. $e(z)$ and $e(z^2)$ are algebraically independent functions. But if e.g. $W \subset \mathbb{C}^2 : z_2 = 2z_1 + 3$ then $e(W)$ is contained in $x_2 = e^3 x_1^2$ and is not Zariski dense.

A theorem of Ax [3] proves (as a special case) that these are the only kind of exceptions. The complex algebraic part of $e^{-1}(V)$ is just the union of translates of positive-dimensional rational subspaces of $\mathbb{C}^n$ contained in $e^{-1}(V)$, which is just the pre-image of cosets contained in V.

Moreover, one can show that the cosets $T \subset V$ are translates of finitely many algebraic subgroups. Hence the union of such cosets (as V is closed) is some algebraic subvariety $S \subset V$.

Proof of Theorem 1.8. via point-counting. First, we can assume that V is defined over $\overline{\mathbb{Q}}$. Indeed, for any V, there is a subvariety $W \subset V$ defined over $\overline{\mathbb{Q}}$ that contains all the algebraic points of V. Say V is defined over a Galois number field K of degree d over $\mathbb{Q}$. Then S is defined over K too.

Choose $\delta = 1/2$ say, so that a torsion point P of complexity N has at least $c(1/2)N^{1/2}$ Galois conjugates over $\mathbb{Q}$ and hence $c(1/2)N^{1/2}/d$ conjugates over K.

On the other hand, Z has at most $c(X, 1/4)$ rational points outside its algebraic part. Therefore, if V contains a torsion point P of sufficiently high complexity then the pre-images of most of its conjugates over K must lie in Z^{alg}, and so their images (the conjugates of P) must lie in S. Hence they all lie in S.

Now we can conclude the proof by induction. As S consists of translates of finitely many algebraic groups T_i, asking which translates of $T = T_i$ lie fully in V is asking for torsion points on the quotient space X/T. □

Several other problems can be approached in the same way. One has a quasi-projective algebraic variety X with some countable collection of algebraic points that are "special" (above: torsion). One has a map $u : U \to X$ from some complex domain U, invariant under some group action with fundamental domain F.

The crucial elements of the proof are: that the graph of $u|_F$ is definable, that the the pre-images of special points are algebraic points of some bounded degree; that special points themselves have high degree (a positive power of some natural complexity measure that controls the height of a pre-image in F): and a description of the algebraic part matching the description of exceptional subvarieties (cosets that contain a torsion point).

Lecture 2

Synopsis. An introduction to definable sets in o-minimal structures, examples, and refinements of point-counting to count algebraic points of bounded degree. This encounters the situation when the "basic" statement can become trivial, but the proof of the counting theorem still yields a useful statement. This will be needed in the application in Lecture 3.

Mathematical structures and model theory. (Also covered in the PAWS notes of Ronnie Nagloo.)

Algebraic structures are often defined as consisting of a set with some specified kind of "additional structure".

A prime example is a *field*. It is a set K endowed with two binary operations $+$ and $\times$ and two elements 0 and 1. Such a structure, whether or not it is a field, we would (in model theory) write as $(K, +, \times, 0, 1)$. Examples are $(\mathbb{R}, +, \times, 0, 1)$ and $(\mathbb{C}, +, \times, 0, 1)$ etc.

There is a corresponding "first-order" language with symbols $\dot{+}, \dot{\times}, \dot{0}, \dot{1}$ in addition to the logical symbols and quantifiers (and $=$ and brackets).

A structure as above is indeed a field if various properties hold (associativity, commutativity, distributivity etc.). All these properties can be expressed in the corresponding language: for example, every non-zero element has to have a multiplicative inverse, and addition is associative:

$$\forall x(\neg x = \dot{0} \to \exists y x \dot{\times} y = \dot{1}), \quad \forall x \forall y \forall z \big(x \dot{+} (y \dot{+} z) = (x \dot{+} y) \dot{+} z\big).$$

These statements use only the operation symbols $(\dot{+}, \dot{\times})$ and constant symbols $(\dot{0}, \dot{1})$ in the language, logical operations $(\neg, \to))$, and quantifiers $(\forall, \exists)$ that run over the set K (and $=$ and brackets).

Another kind of algebraic structure is an *ordered field* $(K, <, +, \times, 0, 1)$ that carries a strict total order. Here one requires that the order interact well with the field operations (e.g. multiplying by a positive number preserves inequalities). I won't list them (see [22]), but e.g. $(\mathbb{R}, <, +, \times, 0, 1)$ and $(\mathbb{Q}, <, +, \times, 0, 1)$ are ordered fields.

More generally, a structure one can have any number of functions, of given arities, on it, and any number of relations, of specified arities, and specified constants. There is a corresponding first-order language. To indicate such a general structure $\mathcal{M}$ on a set M we would write

$$\mathcal{M} = (M, \ldots).$$

For example, $\mathcal{M} = (M, <)$ is a set with a binary relation. If suitable axioms (which can be written in the language) are satisfied it will be a strictly totally ordered set. If one wants to talk about a structure $\mathcal{M}'$ on M consisting of the order $<$ and some other (unspecified) structure, one refers to this as an *expansion* of $\mathcal{M} = (M, <)$ and writes

$$\mathcal{M}' = (M, <, \ldots).$$

Formally one distinguishes the symbols in the language from their interpretation is a structure (e.g. by a dot as above); in practice one ignores this distinction.

Definable sets. Given a structure $\mathcal{M} = (M, \ldots)$, a *definable set* is a set $A \subset M^n, n \geq 1$, whose membership can be described by a sentence ϕ in the first-order language corresponding to $\mathcal{M}$, i.e.

$$A = \{(x_1, \ldots, x_n) \in M^n : \phi(x_1, \ldots, x_n)\}.$$

Strictly, only constants that are distinguished in the structure (and si represented in the language) are permitted in ϕ. A broader notion of *definable with parameters* allows the use of any constants from M. In o-minimality, "definable" nearly always means "with parameters" and we will adopt this convention.

Minimal and o-minimal structures. A non-zero polynomial has only finitely many roots in a field. A consequence is that a subset of $\mathbb{C}$ that is definable (with parameters) in $(\mathbb{C}, +, \times, 0, 1)$ is either finite or cofinite. A structure with this property is called *minimal*, as such sets are definable (with parameters) when there is no structure at all, just using $=$. (The structure is called *strongly minimal* if all elementarily equivalent (same theory) structures are minimal.) This property plays a very important role in model theory, being enjoyed by the "nicest" structures.

Now consider a ordered structure such as $(\mathbb{R}, <)$. It is not minimal as an interval and its complement are both infinite. O-minimality is the analogue of minimality for a structure $\mathcal{M} = (M, <, \ldots)$ expanding a dense linear order without endpoints.

2.1. Definition. A structure $\mathcal{M} = (M, <, \ldots)$ expanding a dense linear order without endpoints is *o-minimal* if the definable subsets of M are no more than the subsets definable in $(M, <)$. Namely, finite unions of points and open intervals (including intervals of the form (a, ∞) and $(-\infty, b)$).

One can consider expansions of more general orders, but in fact the most interesting examples arise as expansions of an ordered field. If o-minimal, the field must be real closed and we will stick to expansions $(\mathbb{R}, <, +, \times, 0, 1, \ldots)$ of the real field.

While sometimes described as fulfilment of Grothendieck's vision of a "tame topology", the idea of o-minimality arose out of the study of the model theory of the real exponential, specifically the structure $(\mathbb{R}, <, +, \times, 0, 1, e^x)$, in work of van den Dries [21], prompted by a question of Tarski, and was developed (and named) in analogy with minimality in a series of papers by Knight, Pillay, and Steinhorn ([29, 45, 46, 47]).

Properties. O-minimal structure have remarkable properties. For example, a function definable in an o-minimal structure (meaning its graph is a definable set) must be continuous except at finitely many points and even (in an expansion of a field) differentiable except at finitely many points.

One also has strong uniformity properties. A *definable family* in a structure $\mathcal{M} = (M, \ldots)$ is a definable subset X of some $M^k \times M^n$ considered as the family of fibres $X_y = \{x \in M^n : (x, y) \in X\}$ parameterized by $y \in M^k$. (Some fibres might be empty.)

If $\mathcal{M} = (M, <, \ldots)$ is an o-minimal structure and $X \subset M^k \times M^n$ is a definable family such that the fibres X_y are all finite, then there must be a uniform bound on their size. (Hence the uniform bound on $\#X \cap C$ for C of degree d in the proof of Theorem 1.3.)

This is part of the proof of the key structure theorem for definable sets in o-minimal structures, the Cell Decomposition Theorem, due to Knight-Pillay-Steinhorn [29].

Examples. The basic example is the ordered field $\mathbb{R}_{\text{alg}} = (\mathbb{R}, <, +, \times, 0, 1)$. The o-minimality of this structure follows from quantifier elimination, due to Tarski.

A second key example is $\mathbb{R}_{\text{exp}} = (\mathbb{R}, <, +, \times, 0, 1, e^x)$, due to Wilkie [51].

Another example is

$$\mathbb{R}_{\text{an}} = (\mathbb{R}, <, +, \times, 0, 1, \{f : B \to \mathbb{R}\})$$

where B ranges over all closed bounded boxes $B \subset \mathbb{R}^n$, for all n, and f over all functions that are real analytic on an open neighbourhood of B. The o-minimality of

this structure of *restricted analytic functions* follows from Gabrielov's Theorem in real analysis.

Finally, one can add e^x to $\mathbb{R}_{\text{an}}$ to form the structure $\mathbb{R}_{\text{an exp}}$. Note that the graph of e^x is not restricted analytic. This structure seems to suffice for diophantine applications. For example, the sets required in the proof of Theorem 1.8 are definable in $\mathbb{R}_{\text{an exp}}$ as they are defined using e^x and restricted sine and cosine.

Proving the counting theorem. The key to this is a realising a definable set as an image in a suitable way i.e. a suitable parameterisation. Such a result for semi-algebraic sets was proved by Yomdin [53] and refined by Gromov [23].

2.2. Definition. Let $Z \subset P \times \mathbb{R}^n$ be a definable family of sets in an o-minimal expansion of $\mathbb{R}$, and $r \geq 1$ an integer. A *definable r-parameterisation* of Z is a finite set Φ of definable families

$$\phi : P \times (0,1)^k \to (0,1)^n$$

of maps $(0,1)^k \to (0,1)^n$ such that, for each $y \in P$, the finite set of fibres $\{\phi_y\} : \phi \in \Phi$ is an r-parameterisation of Z_y.

2.3. Theorem. (The r-Parameterisation Theorem; [43]) *Let Z be a definable family of sets in $(0,1)^n$, and $r \in \mathbb{N}$. Then there exists a definable r-parameterisation of Z.* □

Sketch proof of the counting theorem. Let $\epsilon > 0$ be given. We choose r large enough and r-parameterise Z. Now on a small sub-box of $(0,1)^k$ (but not too small! $\ll H^\epsilon$ boxes cover $(0,1)^k$) all the rational points up to height H lie on one algebraic hypersurface of some degree $d = d(\epsilon)$. The intersections of X with all such hypersurfaces form a definable family. Generally, the intersections will have lower dimension, and we can repeat.

There are a number of technicalities, and one must see how the "algebraic part" presents. Essentially, when a definable set is intersected with an algebraic variety of dimension ℓ and the intersection has the same dimension ℓ then such pieces are in the algebraic part. So in fact one intersects a k-dimensional set Z with k dimensional real algebraic sets by imposing an algebraic relation on every $k+1$ coordinates. The intersections components of dimension k are then in the algebraic part. □

Observation. In this proof, the rational points on all Z are contained in $\ll H^\epsilon$ "pieces" that are either points or positive dimensional pieces contained in Z^{alg}.

One can count algebraic points of bounded degree. For a set $Z \subset \mathbb{R}^n$, integer $k \geq 1$ and $H \geq 1$ we set (using the multiplicative Weil height, but you could also use the maximum height of the coefficients of the minimal polynomial)

$$Z(k,H) = \big\{z = (z_1, \ldots, z_n) \in Z : [\mathbb{Q}(z_i) : \mathbb{Q}] \leq k, \quad H(z_i) \leq H, i = 1, \ldots, n\big\},$$

$$N(k, Z, H) = \#Z(k,H).$$

2.4. Theorem. ([39, Theorem 1.6]) *Let $Z \subset \mathbb{R}^n$ be definable in an o-minimal structure, $k \geq 1$, and $\epsilon > 0$. Then there is a constant $c(Z,k,\epsilon)$ such that, for all $H \geq 1$,*

$$N(k, Z^{\text{trans}}, H) \leq c(Z,k,\epsilon)H^\epsilon.$$

Proving this encounters an issue that will be discussed in the next lecture.

Lecture 3

Synopsis. On unlikely intersections for a curve in $Y(1)^n$. This is a true "unlikely intersection" problem, rather than a "special point" problem. We will go through the proof (of a partial result) emphasizing the arithmetic aspects, functional transcendence, point-counting, and further issues where o-minimality plays a role.

The modular curve $Y(1)$. For background on elliptic curves and their j-invariants see e.g. [54]. Or see [41, Ch. 4].

If $\Lambda = \mathbb{Z}.\omega_1 \oplus \mathbb{Z}.\omega_2$ is a lattice then $\Lambda\backslash\mathbb{C}$ has the structure of an elliptic curve. Scaling the lattice or changing basis produces an isomorphic curve (over $\mathbb{C}$), so one can assume that the lattice has the form $\Lambda_\tau = \mathbb{Z}.1 \oplus \mathbb{Z}.\tau$ with $\tau \in \mathbb{H}$, the complex upper half-plane (positive imaginary part).

The elliptic curve corresponding to Λ_τ is determined up to isomorphism by its j-invariant, a complex number $j(\tau)$ associated to $\tau \in \mathbb{H}$. The *modular function* $j : \mathbb{H} \to \mathbb{C}$ is a holomorphic function that is invariant under the action of the *modular group* $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$ given by

$$z \mapsto \frac{az+b}{cz+d}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

I will write z rather than τ for the variable in $\mathbb{H}$. This action has the classical fundamental domain F consisting of the region in between real parts $\pm 1/2$ and outside the unit circle, and half the boundary.

The modular function j maps onto $\mathbb{C}$; hence the moduli space of complex elliptic curves, which is classically denoted $Y(1)$, is just the complex line $\mathbb{C}$. It is the simplest (positive dimensional) example of a *Shimura variety*.

Special subvarieties. For a generic elliptic curve E, the only complex numbers λ which preserve the lattice (i.e. satisfy $\lambda\Lambda \subset \Lambda$) are the integers. Equivalently, this means that the endomorphism ring of E is generically $\mathbb{Z}$. For some elliptic curves there are non-integer λ preserving Λ. Such elliptic curves are said to have *complex multiplication* and their j-invariants are algebraic numbers (even algebraic integers) known as *singular moduli*. Thirteen of them are rational numbers. Let $\Sigma \subset \mathbb{C}$ denote the set of singular moduli.

The elliptic lattice curve corresponding to Λ_τ has CM just if $\tau \in \mathbb{H}$ is quadratic over $\mathbb{Q}$. Thus singular moduli are the values of j at quadratic points, just as roots of unity are the values of $e^{2\pi iz}$ at rational points.

We also consider the following relations on pairs of elliptic curves E, E': for each positive integer N, one may consider when E, E' are related by a cyclic isogeny of degree N, meaning that $\Lambda' \subset \Lambda$ (up to scaling) and the quotient $\Lambda/\Lambda' = \mathbb{Z}/N\mathbb{Z}$. It turns out that this relation is captured by a polynomial $\Phi_N(j, j')$ on the corresponding j-invariants. Thus $\Phi_1 = X - Y$ while for $N > 1$ one proves that $\Phi_N \in \mathbb{Z}[X, Y]$ and are symmetric. These polynomials are classically called *modular polynomials.*

If all the above is unfamiliar, you may just take as given that there is a certain countably infinite subset of algebraic numbers $\Sigma \subset \mathbb{C}$ and a sequence of bivariate polynomials $\Phi_N \in \mathbb{Z}[X, Y]$ with rather remarkable properties. In particular if $\sigma \in \Sigma$ and $\Phi_N(\sigma, y)$ then $y \in \Sigma$.

Then a *special subvariety* of $Y(1)^3$ is an irreducible component of an algebraic subvariety defined by some number of equations of the form $x_i = \sigma$, where σ is "special" (i.e. a singular modulus), or $\Phi_N(x_j, x_k) = 0$ where Φ_N is a classical modular polynomial.

We define also a broader class of *weakly special subvarieties* where one allows constant coordinates $x_i = c$ where c is any complex number, not necessarily special. So a special subvariety is weakly special and a weakly special subvariety that contains a special point is a special subvariety. (These properties hold in general.)

The André-Oort conjecture concerns special points (singular moduli). The simplest case concerns a curve $V \subset Y(1)^2$. The set of special points of $Y(1)^2$ is Σ^2. The conjecture (proved in this case by André [2]) asserts that if $V \cap \Sigma^2$ is infinite then V must be the zero-set of a modular polynomial. This is the analogue of Lang's problem for a curve $V \subset (\mathbb{C}^\times)^2$ and torsion points. (And one can give a counting proof that extends to $Y(1)^n$ in analogy with the proof of Theorem 1.8 by counting quadtraic points in a suitable definable set.)

In general, for a Shimura variety X, its special points S and special subvarieties $\{T\}$, one has always that special points are dense in a special subvariety. The André-Oort conjecture is the converse statement: if special points are (Zariski-)dense in $V \subset X$ then V is a special subvariety.

Unlikely intersections for a curve in $Y(1)^3$. The André-Oort conjecture, like Multiplicative Manin-Mumford, is a "special point problem". The much broader Zilber-Pink conjecture ([56, 14, 48]) considers more generally "unlikely intersections".

Say $V \subset Y(1)^3$ is a curve and $T \subset Y(1)^3$ is a one-dimensional special subvariety (say defined by two modular conditions $\Phi_N(x, y) = 0, \Phi_M(y, z) = 0$ or one special coordinate and one modular relation, but beware that the intersection of the two modular relations is in general not irreducible: a special subvariety is one of its components. However we want to consider the union of all one-dimensional special subvarieties).

Since V, T are both one-dimensional inside $Y(1)^3$ one would expect them not to intersect, although they might. And there are countably many possibilities for T. Also, if V satisfies some modular condition (e.g. $\Phi_N(x, y) = 0$) identically, then an additional condition will in general still intersect.

The following is the simplest unlikely intersection problem in a Shimura variety.

3.1. Conjecture. (Special case of ZP) Let $V \subset X = Y(1)^3$ be a curve that is not contained in any proper special subvariety of X. Then the intersection of V with the union $X^{[2]}$ of all special subvarieties of codimension ≥ 2 is a finite set.

The multiplicative analogue of this problems was considered in [13], obtaining a partial result completed in [33], and further extended in [15]. A point-counting approach is in [18]. See also [12].

The ZP conjecture ([56, 14, 48]) concerns, more generally, a mixed Shimura variety X in place of $Y(1)^3$. These have a countably infinite collection of "special subvarieties" $\mathcal{T} = \{T\}$ including a countably infinite set of "special points". Suppose $V \subset X$. The conjecture addresses intersections $V \cap T$ for $T \in \mathcal{T}$ that are **atypical in dimension**. This includes the **unlikely intersections**, whose "expected" dimension would be negative such as (as above) two curves in a space of dimension 3 or more.

3.2. Definition. Let X be a mixed Shimura variety with its collection $\mathcal{T} = \{T\}$ of special subvarieties. A subvariety $A \subset V$ is called an *atypical component* (of V in X) if $A \subset_{\rm cpt} V \cap T$ for some $T \in \mathcal{T}$

$$\text{codim } A < \text{codim } V + \text{codim } T, \quad (\text{i.e. } \dim A > \dim V + \dim T - \dim X).$$

3.3. Zilber Pink Conjecture. Let X be a mixed Shimura variety and $V \subset X$. Then the union of atypical components of V in X is a finite union (hence a closed algebraic subset of V).

A number of partial results towards 3.1 are known [26], [40], [20]. The latter considers more generally curves in the Siegel modular varieties $\mathcal{A}_g$ of abelian varieties. In general the required counting results in the counting approach are not effective, but the Galois lower bounds can be made effective in the known cases. In [26] these use isogeny estimates to get Galois lower bounds from height upper bounds.

The points $(x_1, x_2, x_3) \in V \cap X^{[2]}$ fall into a few different types, the "generic" one being that the coordinates x_1, x_2, x_3 are non-special with the corresponding elliptic curves being pairwise isogenous. Let's call these "totally isogenous points".

Point-counting strategy for totally isogenous points. I want to describe this, concentrating on the counting aspect. I will say a little about the arithmetic aspects, which are the central focus of Project 1, at the end.

Suppose that $V \subset Y(1)^3$ is a curve and that $P = (x_1, x_2, x_3)$ is a totally isogenous point. This requires the points x_i to be non-special. Then the special curve T they lie on is unique (otherwise the intersection of two distinct special curves is a special point). So we have unique L, N, M such that

$$\Phi_L(x_1, x_2) = 0, \quad \Phi_M(x_2, x_3) = 0, \quad \Phi_N(x_3, x_1) = 0.$$

We define the *complexity* of P to be $B(P) = \max(L, M, N)$.

Let $z_1, z_2, z_3 \in F$ be the j-pre-images of the x_i. Then $z_2 = gz_2$ for some 2×2 rational matrix of determinant L, and one can show that the entries of the matrix have height at most cB^7 ([26, Lemma 5.2] or see [41, 21.9]; a better exponent is obtained in [32]).

I want first to describe: how (and where) an unlikely intersection leads to a rational point.

Let $G = \mathrm{GL}_2^+(\mathbb{R})$, the group of 2×2 real matrices with positive determinant, and let $Z = j^{-1}(V)$. For $(\alpha, \beta) \in G^2$ let

$$Y_{\alpha,\beta} = \{(z_1, z_2, z_3) \in \mathbb{H}^3 : z_2 = \alpha z_1, z_3 = \beta z_2\}.$$

Then a totally isogenous point $P \in V$ gives rise to a rational point on

$$W = \{(\alpha, \beta) \in G^2 : Y_{\alpha,\beta} \cap Z \neq \emptyset\}.$$

And this is a definable set.

Some consequences of o-minimality. Let us make some further observations. First, since each $Y_{\alpha,\beta}$ is definable, the intersection $Y_{\alpha,\beta} \cap Z$ is either finite or contains a

real analytic arc. But in the latter case, since $Y_{\alpha,\beta}$ and $j^{-1}(V)$ are complex analytic sets, the intersection is complex analytic and positive dimensional, hence $j^{-1}(V) \subset Y_{\alpha,\beta}$.

Now by an analogue of Ax's theorem for the modular function, as this amounts to the "modular logarithm" of V being not Zariski-dense in $\mathbb{H}^3$, we must actually have that V is contained in a proper weakly special subvariety.

Now we rules out V being contained in a proper special subvariety, so the only possibility is that some coordinate is constant on V. This greatly simplifies the problem, and so I want to assume that we are not in this case.

Then each intersection $Y_{\alpha,\beta} \cap Z$ is a finite set.

Now, the sets $Y_{\alpha,\beta}$ for $(\alpha,\beta) \in G^2$ form a definable family, meaning that the set

$$Y = \{(\alpha, \beta, z_1, z_2, z_3) \in G^2 \times \mathbb{H}^3 : z_2 = \alpha z_2, z_3 = \beta z_2\}$$

whose fibres over G^2 are the $Y_{\alpha,\beta}$ is a definable set. It is then an consequence of o-minimality that the finite size of $Y_{\alpha,\beta} \cap Z$ is uniformly bounded over all α, β.

A problem and a work-around. For any point $z \in (z_1, z_2, z_3) \in \mathbb{H}^3$ there is a positive dimensional set of (α,β) such that $z \in Y_{\alpha,\beta}$. This is because G is transitive on $\mathbb{H}$ and each point has a stabiliser. This implies that $W^{\mathrm{alg}} = W$ and so the counting theorem as presented in Lecture 1 says something trivial.

However, all is not lost! The proof of the counting theorem gives something more than stated. I will be a bit sketchy. It implies that the rational points of W are contained in its intersections with $\ll H^\epsilon$ algebraic sets of suitable degree and real dimension that are either points or pieces in the algebraic part (because some dimension did not drop on intersection). These pieces are called "blocks" and defined in detail [39] (with a variant in [41]).

But each individual $Y_{\alpha,\beta}$ only accounts for a finite bounded number of points on Z. If the point P has "many" conjugates, then the "few" pieces (points or subsets of W^{alg}) cannot account for so many points unless there is some one-real-dimensional semi-algebraic subset in one of the pieces such that the intersection point with Z moves. I.e. the "pieces" cannot all be contained in stabilisers.

But then by "modular Ax" we get an algebraic surface that contains $j^{-1}(V)$, which contradicts our assumptions.

Conclusion. Thus, if we can prove that a totally isogenous point has "many" Galois conjugates, then this strategy will succeed. What we need is the following.

Conjecture. For given V defined over $\overline{\mathbb{Q}}$ there are positive constants $c(V), \delta(V)$ such that if $(x_1, x_2, x_3) \in V$ is a totally isogenous point of complexity $B = B(P)$ then

$$[\mathbb{Q}(x_1, x_2, x_3) : \mathbb{Q}] \geq cB^\delta.$$

This conjecture in turn follows if such points have small height, and it can be established ([26]) when V is "asymmetric".

Conjecture. For given V defined over $\overline{\mathbb{Q}}$ and $\epsilon > 0$ there is a constant $c(V, \epsilon)$ such that if $(x_1, x_2, x_3) \in V$ is a totally isogenous point of complexity $B = B(P)$ then

$$h(x_1, x_2, x_3) \leq c(V, \epsilon)B^\epsilon.$$

This conjecture in turn follows from a conjecture on likely intersections (see [24, Appendix B] and [41, 21.23]). For other cases where it holds see [20].

Lecture 4

Synopsis. In the last lecture we will, as time permits, describe further applications and problems.

References

References for the project are included here.

1. J. Armitage, Pfaffian control of some polynomials involving the j-function and Weierstrass elliptic functions, arXiv:2011.09382.
2. Y. André, Finitude des couples d'invariants modulaire singuliers sur une courbe algébrique plane non modulaire, *J. Reine Angew. Math.* **505** (1998), 203–208.
3. J. Ax, On Schanuel's conjectures, *Annals* **93** (1971), 252–268.
4. J.-P. Bézivin, Suites d'entières et fonctions entières arithmétiques, *Ann. Fac. Sci. Toulouse Math.* **3** (1994) 313–334.
5. J.-P. Bézivin, Sur les fonctions entières q-arithmétiques, *Rend. Circulao Math. Palermo (2)* **47** (1998), 447–462.
6. Y. Bilu, D. Masser, and U. Zannier, An effective "theorem of André" for CM points on plane curves, *Math. Proc. Camb. Phil. Soc.* **154** (2013), 145–152.
7. G. Binyamini, Some effective estimates for André-Oort in $Y(1)^n$, with an appendix by E. Kowalski, *Crelle* **767** (2020), 17–35.
8. G. Binyamini, G. O. Jones, H. Schmidt, and M. E. M. Thomas, An effective Pila-Wilkie theorem for sets definable using Pfaffian functions, with some diophantine applications, arXiv:2301.09883.
9. G. Binyamini and D. Masser, Effective André-Oort for non-compact curves in Hilbert modular varieties, arXiv:2101.06412, *C. R. Acad. Sci. Paris, Ser. I*, to appear.
10. G. Binyamini, D. Novikov, and B. Zack, Wilkie's conjecture for Pfaffian structures, arXiv:2202.05305.
11. R. P. Boas, Comments on [49], in George Pólya: Collected Ppers, Volume 1, R. P. Boas, editor, MIT Press, Cambridge, 1974, 771–773.
12. E. Bombieri, P. Habegger, D. Masser, and U. Zannier, A note on Maurin's theorem, *Rend. Lincei. Mat. Appl.* **21** (2010), 251–260.
13. E. Bombieri, D. Masser, and U. Zannier, Intersecting a curve with algebraic subgroups of multiplicative groups, *IMRN* **20** (1999), 1119–1140.
14. E. Bombieri, D. Masser, and U. Zannier, Anomalous subvarieties – structure theorems and applications, *IMRN* **19** (2007), 33 pages.
15. E. Bombieri, D. Masser, and U. Zannier, On unlikely intersections of complex varieties with tori, *Acta Arithmetica* **133** (2008), 309–323.
16. E. Bombieri and J. Pila, The number of integral points on arcs and ovals, *Duke Math. J.* **59** (1989) 337–357.
17. T. D. Browning and D. R. Heath-Brown, Plane curves in boxes and equal sums of two powers, *Math. Z.* **251** (2005), 233–247.

18. L. Capuano, D. Masser, J. Pila, and U. Zannier, Rational points on Grassmannians and unlikely intersections in tori, *Bull. London Math. Soc* **48** (2016), 141–154.
19. P. Corvaja, D. Masser, and U. Zannier, Torsion hypersurfaces on abelian schemes and Betti coordinates, *Math. Ann.* **371** (2018), 1013–1045.
20. C. Daw and M. Orr, Some cases of the Zilber-Pink conjecture for curves in $\mathcal{A}_g$, arXiv:2211.06763.
21. L. van den Dries, Remarks on Tarski's problem concerning $(\mathbb{R}, +, \cdot, \exp)$, in *Logic colloquium '82,* pp. 97–121, Lolli, Longo, and Marcja, editors, *Studies in Logic and the Foundations of Mathematics* **112**, North Holland, 1984.
22. L. van den Dries, *Tame Topology and O-minimal Structures,* LMS Lecture Note Series **248**, CUP, 1998.
23. M. Gromov, Entropy, homology and semi-algebraic geometry [after Y. Yomdin], Séminaire Bourbaki, 1985-86, exposé 663, *Astérisque* **145-146** (1987), 225–240.
24. P. Habegger, Effective height upper bounds on algebraic tori, arXiv:1201.1815, *Autour de la conjecture de Zilber-Pink*, Course notes, CIRM, 2011, 167–242, Panor. Synthèses **52**, Soc. Math. France, Paris, 2017.
25. P. Habegger, G. Jones, and D. Masser, Six unlikely intersection problems in search of effectivity, *Math. Proc. Camb. Phil. Soc.* **162** (2017), 447–477.
26. P. Habegger and J. Pila, Some unlikely intersections beyond André-Oort, *Compositio* **148** (2012), 1–27.
27. G. O. Jones and S. Qiu, Integer-valued definable functions in $\mathbb{R}_{\text{an exp}}$, *IJNT* **17** (2021), 1739–1752.
28. G. O. Jones, M. E. M. Thomas, and A. J. Wilkie, Integer-valued definable functions, *Bull. LMS* **44** (2012), 1285–1291.
29. J. Knight, A. Pillay, and C. Steinhorn, Definable sets in ordered structures. II, *Trans. AMS* **295** (1986), 593–605.
30. L. Kühne, An effective result of André-Oort type, *Annals* **176** (2012), 651–671.
31. D. Masser and U. Zannier, Torsion anomalous points and families of elliptic curves, *C. R. Acad. Sci. Paris, Ser. I* **346** (2008), 491–494, and *Amer. J. Math* **132** (2010), 1677–1691.
32. D. Masser and U. Zannier, Abelian varieties isogenous to no Jacobian, *Annals* **191** (2020), 635–674.
33. G. Maurin, Courbes algébriques et équations multiplicatives, *Math. Annalen* **341** (2008), 789-824.
34. F. Pellarin, Sur une majoration explicite pour un degré disogénie liant deux courbes elliptiques, *Acta Arith.* **100** (2001), 203–243.
35. A. Perelli and U. Zannier, Su un teorema di Pólya, *Boll. Un. Mat. Ital.* A (5) **18** (1981), 305–307.
36. J. Pila, Geometric postulation of a smooth function and the number of rational points, *Duke Math. J.* **63** (1991) 449–463.
37. J. Pila, Concordant sequences and concordant entire functions, *Ren. Circ. Math. Palermo (2)* **51** (2002) 51–82.
38. J. Pila, Entire functions having a concordant value sequence, *Israel J. Math.* **134** (2003) 317–343
39. J. Pila, On the algebraic points of a definable set, *Selecta Math. N. S.* **15** (2009), 151–170.

40. J. Pila, On a modular Fermat equation, *Commentarii Mathematici Helvetici* **92** (2017), 85–103.
41. J. Pila, *Point-Counting and the Zilber–Pink Conjecture,* Cambridge Tracts in Mathematics **228**, CUP, 2022.
42. J. Pila and F. R. Villegas, Concordant sequences and integral-valued entire functions, *Acta Arithmetica* **88** (1999) 239–268.
43. J. Pila and A. J. Wilkie, The rational points of a definable set, *DMJ* **133** (2006), 591–616.
44. J. Pila and U. Zannier, Rational points in periodic analytic sets and the Manin-Mumford conjecture, *Rend. Mat. Acc. Lincei (9)* **19** (2008) 149–162.
45. A. Pillay and C. Steinhorn, Definable sets in ordered structures, *Bulletin AMS* **11** (1984), 159–162.
46. A. Pillay and C. Steinhorn, Definable sets in ordered structures I, *Trans. AMS* **295** (1986), 565–592.
47. A. Pillay and C. Steinhorn, Definable sets in ordered structures III, *Trans. AMS* **309** (1988),469–476.
48. R. Pink, A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang, manuscript dated 17 April 2005 available from the author's webpage.
49. G. Pólya, Über ganzwertige ganze Funktionen, *Rend. Circ.Mat.Palermo* **40** (1915), 1–16. Also collected papers.
50. M. Waldschmidt, Integer–valued functions, Hurwitz functions and related topics: a survey, *31st meeting of the Journées Arithmétiques,* arXiv:2002.01223
51. A. J. Wilkie, Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function, *J. Amer. M. Soc.* **9** (1996), 1051–1094.
52. A. J. Wilkie, Complex continuations of $\mathbb{R}_{\text{an exp}}$-definable unary functions with a diophantine application, *J. LMS* **93** (2016), 547–566.
53. Y. Yomdin, C^k-resolution of semi-algebraic mappings. Addendum to "Volume growth and entropy", *Israel J. Math.* **57** (1987), 301–317.
54. D. Zagier, Elliptic modular forms and their applications, *The 1-2-3 of modular forms,* 1–103, J.H. Brunier, G. van der Geer, G. Harder, and D. Zagier, Springer, Berlin, 2008.
55. U. Zannier, *Some problems of unlikely intersections in arithmetic and geometry,* with appendices by D. Masser, Annals of Mathematics Studies **181**, PUP, 2012.
56. B. Zilber, Exponential sums equations and the Schanuel conjecture, *J. London Math. Soc. (2)* **65** (2002), 27–44.

20230131