20090302

Rational points of definable sets and results of André-Oort-Manin-Mumford type

Jonathan Pila

Abstract

We prove some simple special cases, partly new, of results of André-Oort-Manin-Mumford type using an extension to algebraic points of bounded degree of a result of Pila-Wilkie on the density of rational points on sets definable in an o-minimal structure. The strategy follows that of a recent new proof of the Manin-Mumford conjecture by Pila-Zannier, and a proof of a special (but new) case of Pink's relative Manin-Mumford conjecture by Masser-Zannier.

2000 Mathematics Subject Classification: 11G18, 03C64

Keywords: Definable set, rational point, André-Oort conjecture, Manin-Mumford conjecture.

1. Introduction

This paper is devoted to giving proofs of three results in the André-Oort-Manin-Mumford circle of conjectures. Theorem 1.1 is an early case (due to André [2] and Edixhoven [21]) of the André-Oort conjecture itself. Our proof apparently yields some new features. Theorem 1.2 is a variant of a result of André [3] that is a special case of a "generalized André-Oort" conjecture described there. Theorem 1.3 is a variant of a result of Nekovar-Schappacher [34] on non-triviality of Heegner points. (The André-Oort conjecture has recently been proved in full under GRH in work of Klingler, Ullmo, and Yafaev, see [54].)

Our approach has been motivated by the following recent developments. Pila and Zannier [45], implementing a strategy proposed by Zannier, gave a new proof of the Manin-Mumford conjecture, and Masser and Zannier [31] saw that a similar strategy (though the technicalities are more delicate) could be used to prove a very special case (but of an entirely new type) of Pink's conjecture [47, 1.1], more precisely a special case of Pink's relative Manin-Mumford conjecture [47, 6.1]. In its general formulation [47, 1.1] (see also [46]), Pink's conjecture includes the conjectures of Mordell-Lang, Manin-Mumford and André-Oort, and generalizes the questions raised in the multiplicative setting by Bombieri-Masser-Zannier [7] (see also [8, 9]) and the conjectures proposed independently by Zilber [55]. One key ingredient of the proofs in [45, 31] is a result of Pila and Wilkie [44] on the height density of rational points on subsets of $\mathbb{R}^n$ that are *definable in an o-minimal structure* (see §3, or [17]), or its predecessors (on the other ingredients see further below). This result can be extended [43] to a density result for algebraic points of bounded degree on such sets, stated as Theorem 1.5 below. Now, for example, CM values of the j-invariant of elliptic curves correspond to quadratic algebraic points τ in the upper half plane, so such τ are algebraic of bounded degree. Thus this extension, in light of the previous applications to Manin-Mumford-type problems, suggested that problems of André-Oort type might be amenable to the same general strategy. The simple examples considered here show that this is indeed the case.

The strategy of proof may be sketched as follows. One has an algebraic variety A over $\overline{\mathbb{Q}}$ of suitable type (e.g. an abelian variety or a modular variety such as the affine line $\mathbb{C}$ parametrizing the j-invariants of elliptic curves, or a combination of these), its special points (e.g. torsion points of an abelian variety, or CM j-invariants, or combinations of these), and a subvariety V of A. One wishes to show that V contains only finitely many special points unless it contains a special subvariety of A of positive dimension. The variety A is uniformized by certain transcendental functions $\pi : U \to A$ on a complex domain U (e.g. $\mathbb{C}^n$ for an abelian variety or the upper-half plane $\mathcal{H}$ for moduli of elliptic curves), automorphic for some discrete group Γ (e.g. period lattice in $\mathbb{C}^n$, or $\mathrm{SL}_2(\mathbb{Z})$ action on $\mathcal{H}$).

The preimages in U of special points of A have certain rationality properties (e.g. rational points with respect to the lattice, or quadratic algebraic points $\tau \in \mathcal{H}$). Let $Z = \pi^{-1}(V) \subset U$, and let $\mathcal{Z}$ be the intersection of Z with a fundamental domain for Γ. To count special points on V we may count instead their preimages in $\mathcal{Z}$. We consider $Z \subset \mathbb{R}^{2n}$ with suitable real coordinates. In these coordinates, if $\mathcal{Z}$ is definable in some o-minimal structure (and observe that Z itself, with its invariance properties under the infinite discrete group Γ, cannot in general be so definable), then [44] (respectively [43], see 1.5 below) gives an *upper bound* on the number of rational points (respectively algebraic points of some bounded degree) up to a given height in $\mathcal{Z}$ that do not lie on some connected positive dimensional semi-algebraic subset of $\mathcal{Z}$. In the cases considered it turns out that if V contains no special subvarieties of A of positive dimension then $\mathcal{Z}$ has no such algebraic subsets containing special points and the upper bound applies to $\mathcal{Z}$ in its entirety. On the other hand, as special points are algebraic points on V, their suitable galois conjugates also lie on V. A second key ingredient is a *lower bound* for the number of conjugates of such points in terms of their "complexity" (order of torsion point or discriminant of CM-field). The upper and lower estimates contradict each other once the complexity of a special point on V gets too large, giving the finiteness.

Each of our theorems concerns a curve V in a certain surface A. The same approach should work in suitable higher-dimensional situations as well but we will pursue this and other generalizations elsewhere. In each case, the curve V (which we do not assume to be irreducible) may be taken to be defined over $\mathbb{C}$, although the finiteness of special points is immediate for any component of V that is not defined over $\overline{\mathbb{Q}}$, so in the proofs we quickly reduce to the case that V is defined over $\overline{\mathbb{Q}}$. The upper bounds from [44, 43] can probably be made effective (though not too easily; see Remark 2.4.1), but since we rely on Siegel's ineffective lower bound for class numbers (see e.g. [28]), all our theorems below are ineffective.

Theorem 1.1 concerns the variety $A = \mathbb{C}^2$ parametrizing pairs of elliptic curves, up to isomorphism, by their j-invariants. A point $(j, j') \in \mathbb{C}^2$ will be called *special* if j, j' are both j-invariants of CM elliptic curves. For a positive integer N, it is well-known that, for $\tau \in \mathcal{H}$, $j(\tau)$ and $j(N\tau)$ are algebraically related, i.e. that $F_N(j(\tau), j(N\tau)) = 0$ for some bivariate polynomial F_N where, classically, F_N is symmetric and has integer coefficients. The curve $F_N(z_1, z_2) = 0$ in $\mathbb{C}^2$ is then a planar image of the modular curve $X_0(N)$ (see e.g. Milne [32]). The *special subvarieties* of positive dimension of $\mathbb{C}^2$ are accordingly the "vertical" lines, $\{(z, z_2) : z_2 \in \mathbb{C}\}$ where z is a CM j-invariant, the "horizontal" lines $\{(z_1, z) : z_1 \in \mathbb{C}\}$ where z is a CM j-invariant, the modular curves in $\mathbb{C}^2$ defined by $F_N = 0$, and $\mathbb{C}^2$ itself. (Here and below the special subvarieties of dimension 0 are just the special points.)

1.1. Theorem. *Let V be a curve in $\mathbb{C}^2$. Then V contains only finitely many special points of $\mathbb{C}^2$ unless V contains a special subvariety of $\mathbb{C}^2$ of positive dimension.*

Suppose V contains no special subvarieties of positive dimension. The maximal possible number of special points on V (and indeed the height of the corresponding points $\tau_1, \tau_2 \in \mathcal{H}$ in a fundamental domain, i.e. $(\tau_1, \tau_2) \in \mathcal{Z}$) is then bounded uniformly (though ineffectively) for V of given bidegree defined over a number field of given degree over $\mathbb{Q}$. This unconditional uniformity in our proof seems to be new. Edixhoven's proof [21] of Theorem 1.1 is conditional on GRH for imaginary quadratic fields ("GRHIQ" in the sequel) and (under GRHIQ) uniform in the above sense (as noted in [12]: the uniformity is stated in [21] only for curves of given bidegree defined over $\mathbb{Q}$). Breuer [12] refines Edixhoven's proof to yield (under GRHIQ) an effectively computable bound on the height of a CM point on a curve (in $\mathbb{C}^n$) of given multidegree and degree of field of definition over $\mathbb{Q}$ that is not a modular curve. André's proof [2] of Theorem 1.1 is unconditional but appears not to be uniform (cf the comments in [12]). By general theory of Shimura varieties (see e.g. [22, end of §1, or 25, Prop. 2.1]) the conclusion of Theorem 1.1 extends automatically to special points on curves in the product of any two modular curves (I thank a referee for pointing this out). I sketch such an argument in 4.2.2 for products $Y_0(N) \times Y_0(M)$. The result for such products has the same unconditional uniformity.

Theorem 1.2 concerns the Legendre family of elliptic curves i.e. the quasiprojective surface A in $\mathbb{A}^1(\mathbb{C}) \times \mathbb{P}^2(\mathbb{C})$ defined by

$$Y^2Z - X(X-Z)(X-\lambda Z) = 0, \quad \lambda \neq 0, 1.$$

For a given $\lambda \neq 0, 1$, the points $P = (X, Y, Z)$ such that $(\lambda, X, Y, Z) \in A$ comprise an elliptic curve that we denote E_λ.

The *special points* of A are points (λ, P) where E_λ is a CM curve and $P \in E_\lambda$ is a torsion point. The *special subvarieties* of positive dimension of A are the "vertical" subvarieties i.e. those of the form $\{\lambda\} \times E_\lambda$ for some $\lambda \neq 0, 1$, with E_λ a CM curve, the "torsion" subvarieties i.e. curves $W \subset A$ whose points are all of the form $\{(\lambda, P_\lambda)\}$ where P_λ is a torsion point of E_λ for all λ (and therefore a torsion point of fixed order independent of λ), and A itself.

1.2. Theorem. *Let V be a curve in A. Then V contains only finitely many special points of A unless V contains a special subvariety of A of positive dimension.*

Let $f : B \to T$ be a non-isotrivial elliptic pencil, and σ a non-torsion section. André [3] (see also [36, §3.3]) proves finiteness for the number of parameters $t \in T$ for which B_t is CM and $\sigma(t)$ is a torsion point of B_t. The curve V in 1.2 need not arise from a section, and I have not found this variant in the literature. When V contains no special subvarieties of positive dimension, the number (and "complexity", i.e. height of $\tau \in \mathcal{H}$ and order of torsion point z of the corresponding point $(\tau, z) \in \mathcal{Z}$) of special points is again bounded uniformly in terms of the bidegree and degree of defining field of V.

Theorem 1.3 concerns Heegner points. Suppose that E is an elliptic curve over $\mathbb{Q}$ of conductor N, and $\Phi_E : X_0(N) \to E$ a modular parametrization. Let us call the image in E of a special point of $X_0(N)$ a *CM-point* of E, where a special point of $X_0(N)$ is simply the image under $\mathcal{H} \to X_0(N)$ of a special (i.e. quadratic) point $\tau \in \mathcal{H}$. Such a point is called a *Heegner point* of E if certain additional "Heegner" conditions on the discriminant of τ with respect to the level N are satisfied, see e.g. [16, 34]. Sometimes, the term Heegner point is also used for a suitable sum of such points having a much smaller field of definition, see e.g. [5].

Take now $A = X_0(N) \times E$. The *special points* of A are those (P, Q) with P a special point of $X_0(N)$ (as above), and Q a torsion point of E. The *special subvarieties* of positive dimension of A are the "vertical" subvarieties $\{P\} \times E$ where $P \in X_0(N)$ is special, the curves $W \subset A$ whose points (P, Q) are such that Q is always torsion, and hence are constant: i.e. "horizontal" subvarieties $X_0(N) \times \{Q\}$, $Q \in E$ a torsion point, and A itself. Let $V \subset X_0(N) \times E$ be the graph of Φ_E. A point $(P, Q) \in V$ is special just if P is special and $Q = \Phi_E(P)$ is torsion. Now V evidently contains no special subvarieties of positive dimension so it should contain only finitely many special points. More generally, one can consider any curve V in $X_0(N) \times E$, and then E need not be parametrized by $X_0(N)$ in the formulation and need only be defined over $\overline{\mathbb{Q}}$, with special subvarieties as defined above.

1.3. Theorem. *Let V be a curve in $X_0(N) \times E$, where E is an elliptic curve defined over $\overline{\mathbb{Q}}$. Then V contains only finitely many special points of $X_0(N) \times E$ unless V contains a special subvariety of $X_0(N) \times E$ of positive dimension.*

When $X_0(N)$ and E are fixed and V contains no special subvarieties of positive dimension, the number of special points on V is bounded uniformly (along with the "complexity" of the corresponding points in $\mathcal{Z}$) as in 1.1 and 1.2. For modular E, Theorem 1.3. implies the non-triviality of all but finitely many CM-points, and therefore of all but finitely many Heegner points. The latter is proved in greater generality by Nekovar-Schappacher [34]. The connection between results of André-Oort type with Heegner points is exploited in an alternative proof of a much deeper result on non-triviality of Heegner points by Cornut [15].

Note added. A referee has pointed out to me the preprint [13] of Buium and Poonen whose (much more general) main result completely contains 1.3. (They state their result with $X_1(N)$ rather than $X_0(N)$, though this immaterial in view of general theory of Shimura varieties alluded to above.) They prove a finiteness result for CM points of $X_1(N)$ mapping under a correspondence into a finitely generated subgroup Γ of E. In fact, they can replace E by an abelian variety, they can replace Γ with $\Gamma+B_\epsilon$, where B_ϵ is a set of points of small Néron-Tate height (torsion points have Néron-Tate height zero), and they prove variants where the modular curve is replaced by a Shimura curve. See also [14]. Interestingly, they observe that their theorem is apparently not a consequence of Pink's conjectures. It may be possible to recover further parts of the Buium-Poonen result by our methods, but we will pursue this elsewhere.

Since the result giving our upper estimates applies to sets definable in an o-minimal structure, a crucial third ingredient in the application of these results is a theorem of Peterzil and Starchenko [38] establishing this property for the Weierstrass $\wp$-function $\wp(z,\tau)$ when the variables are suitably restricted. We take $\mathcal{D} = \{\tau \in \mathcal{H} : -1/2 \le \mathrm{Re}(\tau) < 1/2, |\tau| \ge 1\}$ to be the standard fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$. For $\tau \in \mathcal{H}$ we define $\mathcal{L}_\tau = \{t_1 + t_2\tau \in \mathbb{C} : t_1, t_2 \in \mathbb{R}, 0 \le t_1, t_2 < 1\}$. This is a fundamental domain for the action of the lattice $\Lambda_{1,\tau} = \mathbb{Z} \oplus \mathbb{Z}\tau$ on $\mathbb{C}$. We define further

$$\mathcal{F} = \{(\tau, z) \in \mathcal{H} \times \mathbb{C} : \tau \in \mathcal{D}, z \in \mathcal{L}_\tau\}.$$

The result of Peterzil and Starchenko [38], stated with the attendant definitions in §3, is that the restriction of $\wp(z,\tau)$ to $\{(z,\tau) : (\tau, z) \in \mathcal{F}\}$ is definable in the o-minimal structure $\mathbb{R}_{\mathrm{an,exp}}$. (I follow the more usual practice of ordering (z,τ) the variables of $\wp$, but it is convenient otherwise to have τ as the first variable. I hope that this causes no confusion.) All the definability properties required in this paper flow from this result.

For sets Z definable in an o-minimal structure, [44] (respectively [43]) gives a bound for the rational (respectively algebraic of bounded degree) points of $Z - Z^{\mathrm{alg}}$ up to a given height, where Z^{alg} is defined as follows.

1.4. Definition. Let $Z \subset \mathbb{R}^m$. The *algebraic part* of Z, denoted Z^{alg} is the union of all connected positive-dimensional semi-algebraic subsets of Z.

We can now state the upper estimate result. For a set $Z \subset \mathbb{R}^m$ and an integer $k \ge 1$, denote by

$$Z(k) = \{x = (x_1, \ldots, x_m) \in Z : \max_i([\mathbb{Q}(x_i) : \mathbb{Q}]) \le k\}$$

the set of algebraic points of Z of degree $\le k$. For $k = 1$ this is just the rational points of Z. Note however that when $k > 1$ the definition allows points $x \in Z(k)$ and indeed even the various coordinates of a point $x \in Z(k)$ to be defined over different numberfields. For $T \ge 1$ set

$$Z(k,T) = \{x \in Z(k) : \max_i H(x_i) \le T\}$$

where $H(\alpha)$ is the absolute multiplicative height of an algebraic number, as defined in [6], and

$$N_k(Z,T) = \#Z(k,T).$$

The result employed here is that the estimate of [44] for rational points on a definable set Z up to height T holds for any fixed k.

1.5. Theorem. ([44], for $k = 1$, and [43]) *Let $Z \subset \mathbb{R}^m$ be definable in an o-minimal structure over $\mathbb{R}$, let $k \ge 1$ and $\epsilon > 0$. There is a constant $c(Z,k,\epsilon) > 0$ such that*

$$N_k(Z - Z^{\mathrm{alg}}, T) \le c(Z,k,\epsilon)T^\epsilon.$$

The main work in this paper is the identification of the possible constituents of $Z^{\rm alg}$ in §2. Since our $Z \subset \mathbb{C}^n$ we take $Z \subset \mathbb{R}^{2n}$ in suitable coordinates. In each case we consider it turns out that the real semi-algebraic curves in $Z \subset \mathbb{R}^{2n}$ are contained in complex algebraic curves in $Z \subset \mathbb{C}^n$ and the curves that may occur correspond almost precisely to the preimages of special subvarieties of A of positive dimension, i.e. $Z^{\rm alg}$ is almost the preimage under π of the special set of V. (The "almost" is explained in connection with our definitions in §2 below. In short, our definitions admit some additional sets, but they have no special points on them.) This phenomenon should be true more generally (see Remark 2.9) as it is very restrictive for a union of complex algebraic sets to be have the necessary invariance properties under an infinite discrete group. One might then obtain new proofs of the corresponding André-Oort-Manin-Mumford assertion provided the other requisite ingredients are available i.e. suitable rationality properties of pre-images of special points, a suitable lower bound for the size of galois orbits of special points (apparently such bounds are not presently available in general; see [24, Problem 5], cf bounds obtained in [53]), and a suitable analogue of the definability result of Peterzil-Starchenko.

Note added. In the first instance, Theorem 1.1 can be extended to the case of a curve in the product of two Shimura curves associated to indefinite quaternion algebras over $\mathbb{Q}$, affirming the André-Oort conjecture in that case. This seems to give the first *unconditional* proof of this result, which has been established under GRHIQ by Yafaev [52]. The proof will appear elsewhere, but it proceeds along the same lines. Indeed, when the Shimura curve is realized as a compact quotient of $\mathcal{H}$, definability of $\mathcal{Z}$ in $\mathbb{R}_{\rm an}$ is immediate. The special points may be taken to be suitable algebraic points of degree 4, and the analogue of 2.3 holds. I thank a referee for the suggestion to pursue this particular result. The present approach may well yield an unconditional proof of the André-Oort conjecture for Hilbert modular surfaces as well (the result has been proved under GRHIQ by Edixhoven [22]).

The proofs of the theorems are carried out in §4. They are quite straightforward, as our lower bounds (as well as our upper bounds) come from results in the literature. For the j-function at a quadratic point $\tau \in \mathcal{H}$, the number of conjugates of $j(\tau)$ over $\mathbb{Q}$ is equal to the class number of the corresponding order. Here we use the aforementioned (ineffective) lower bound due to Siegel. We need also a lower bound for the degree of a torsion point of an elliptic curve in terms of its order of torsion. In Theorem 1.2, the elliptic curves concerned are varying but have CM and so we can use a result of Silverberg [49]. The elliptic curve in Theorem 1.3 need not be CM, but is fixed and we use the result of Masser [30] that was also employed in [45]. Since the lower bounds for the number of conjugates of a special point depend on the corresponding discriminants (in 1.1) and (in 1.2, 1.3) order of torsion, a final but elementary task is to relate the height of the preimage in $\mathcal{Z}$ of a special point of V to these quantities.

2. Algebraic subsets

For each theorem we will have a suitable uniformization $\pi : U \to A$, where $U \subset \mathbb{C}^n$, and $Z = \pi^{-1}(V)$. We define the *complex algebraic part* of Z, denoted $Z^{\rm ca}$, to be the union of of connected components Y of $W \cap U$ such that $Y \subset Z$ over positive-dimensional irreducible closed complex algebraic sets $W \subset \mathbb{C}^n$. We will in each case use suitable real coordinates, giving $Z \subset \mathbb{R}^{2n}$. We must identify the possible real semialgebraic subsets of the set Z, that is, the constituents of $Z^{\rm alg}$. This was also a key step in [45] and we adapt some of the arguments there. With the real coordinates we use we will always have $Z^{\rm ca} = Z^{\rm alg}$.

Let $U \subset \mathbb{C}^n$ be an open domain. By giving real coordinates for U we mean giving functions $x_1, \ldots, x_{2n} : U \to \mathbb{R}$ such that the assignment $z \mapsto x(z) = (x_1(z), \ldots, x_{2n}(z))$ gives a bijection of U with an open domain in $\mathbb{R}^{2n}$. The inverse functions we denote $z_i = z_i(x)$. We identify subsets of U (including U itself) with their images in $\mathbb{R}^{2n}$. By a complex analytic subset of an open domain in $\mathbb{C}^n$ we mean a set defined in a neighbourhood of each point of the domain by the vanishing of a finite number of complex analytic functions. The following generalizes [45, Lemma 2.1], in which we determined real coordinates by $z_j = x_{2j-1} + iz_{2j}$.

2.1. Lemma. *Let $U \subset \mathbb{C}^n$ be an open domain. Suppose that real coordinates $x_1, \ldots, x_{2n}$ on U are given such that $U \subset \mathbb{R}^{2n}$ is semialgebraic and the functions $z_i(x)$ are given by polynomials in $\mathbb{C}[x_1, \ldots, x_{2n}]$. Suppose that $Z \subset U$ is (complex) analytic. Then*

$$Z^{\text{ca}} = Z^{\text{alg}}.$$

Proof. Suppose that W is a connected closed algebraic subset of $\mathbb{C}^n$ of positive dimension with $W \cap U \subset Z$. Since $U \subset \mathbb{R}^{2n}$ is semialgebraic, the set $W \cap U \subset \mathbb{R}^{2n}$ is semialgebraic. It is also connected and of positive dimension. Thus we have $Z^{\text{ca}} \subset Z^{\text{alg}}$.

Suppose now that $Z \subset \mathbb{R}^{2n}$ contains an arc of an irreducible real algebraic curve C. Let P be a smooth point of this arc, which for notational simplicity we may take to be the origin in $\mathbb{R}^{2n}$. So $t = x_1$, say, is a uniformizing parameter at P, and the functions x_i on C are algebraic over $\mathbb{R}(t) \subset \mathbb{C}(t)$. The functions z_j on C are then also algebraic over $\mathbb{C}(t)$. Suppose z_j is non-constant on the arc (there must be a z_j with this property otherwise the arc would reduce to a point). Then all the functions z_i are algebraic over $\mathbb{C}(z_j)$. Now the functions $z_i(t)$ are real-analytic in a real neighbourhood of $t = 0$, and hence they are complex-analytic in some complex neighbourhood of $t = 0$. Then the image of the assignment $t \mapsto (z_1(t), \ldots, z_n(t))$ in this neighbourhood is a neighbourhhod of P in some complex alagebraic curve $\mathcal{C}$. Then $\mathcal{C}$, in a neighbourhood of P, contains C, and so some irreducible component $\mathcal{C}'$ of $\mathcal{C}$ contains C in the neighbourhood. As $C \subset Z$ and Z is complex analytic, we have $\mathcal{C}' \subset Z$ (the functions defining Z vanish on the image of real t under $t \mapsto z(t)$, hence on the complex image as well).

Now any connected semialgebraic set of positive dimension can be covered by smooth arcs of irreducible semialgebraic curves except perhaps for finitely many points. At every point of these arcs we find an irreducible complex analytic curve containing it in some neighbourhood and contained in Z. The remaining finitely many points also belong to Z^{ca} by continuity (they are in the closures of the arcs), and we conclude that $Z^{\text{alg}} \subset Z^{\text{ca}}$. □

Remark. One could relax further the conditions on $z_i(x)$ in the lemma. It suffices if they are algebraic functions analytic at each point of U. Perhaps just algebraic suffices.

Consider first the situation of Theorem 1.1. Here we have $U = \mathcal{H}^2 = \mathcal{H} \times \mathcal{H}$ with the group $\mathrm{SL}_2(\mathbb{Z})^2$ acting as fractional linear transformations of each variable. The uniformization $\pi : \mathcal{H}^2 \to \mathbb{C}^2$ given by $\pi(\tau_1, \tau_2) = (j(\tau_1), j(\tau_2))$ is $\mathrm{SL}_2(\mathbb{Z})^2$-invariant, so our sets $Z \subset U$ will be $\mathrm{SL}_2(\mathbb{Z})^2$-invariant.

For any positive $N \in \mathbb{N}$ we have a polynomial relation $F_N(j(\tau), j(N\tau))$, which determines some plane curve in $\mathbb{C}^2$. We let Z_N be its preimage in $\mathcal{H}^2$. Note that Z_N does not consist in just the (intersection of $\mathcal{H}^2$ with the) line $\tau_2 = N\tau_1$, as we have $F_N(j(\tau), Y) = \prod(Y - j(N\gamma_i\tau))$ where the γ_i run over a set of representatives for the (right) cosets of the congruence subgroup $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$, see e.g. [32, Ch. V]. Indeed, for any $g \in \mathrm{SL}_2(\mathbb{R})$ whose image in $\mathrm{PSL}_2(\mathbb{R})$ is in the image of $\mathrm{GL}_2(\mathbb{Q})$, i.e. such that the ratio of any two entries g is rational so that the same fractional linear transformation of $\mathcal{H}$ is effected by some matrix with rational entries, the locus $Z_g = \{(\tau_1, g\tau_1) : \tau_1 \in \mathcal{H}\}$ is contained in Z_N for suitable N, and the union of Z_g and its conjugates under $\mathrm{SL}_2(\mathbb{Z})^2$ is Z_N. For $\tau \in \mathcal{H}$ we set $V_\tau = \{(\tau, \tau_2) : \tau_2 \in \mathcal{H}\}$ and $H_\tau = \{(\tau_1, \tau) : \tau_1 \in \mathcal{H}\}$.

The sets $V_\tau, H_\tau, Z_g, \mathcal{H}^2$ are our analogues of “special subvarieties” of positive dimension, and accordingly we make the following definition.

2.2. Definition. Let $Z \subset \mathcal{H}^2$. We define $Z^{\text{quasispecial}}$ (for Γ) to be the union of all sets of the form Z_g (for rational $g \in \mathrm{SL}_2(\mathbb{R})$), $V_\tau, H_\tau, \mathcal{H}^2$ contained in Z.

Our notion of quasispecial subset of $\mathcal{H}^2$ differs somewhat from the corresponding notion of special subvariety of positive dimension in $\mathbb{C}^2$ in that we do not insist that the τ corresponding to a quasispecial set V_τ or H_τ be a special (i.e. CM) point. Evidently, as already mentioned, the sets V_τ, H_τ that do not correspond to special subvarieties do not contain any pre-images of special points.

We write $\tau_1 = u + iv, \tau_2 = x + iy$ and use u, v, x, y as real coordinates. The hypotheses of Lemma 2.1 are satisfied. The sets Z_g, V_τ, H_τ are evidently complex algebraic, thus

$$Z^{\text{quasispecial}} \subset Z^{\text{ca}} = Z^{\text{alg}},$$

and our result is that, in analogy with [45, Theorem 2.1], these sets in fact coincide when $Z = \pi^{-1}(V)$.

2.3. Proposition. *Let $V \subset \mathbb{C}^2$ be an algebraic curve defined over $\mathbb{C}$ and let $Z = \pi^{-1}(V) \subset \mathcal{H}^2$. Then $Z^{\text{alg}} = Z^{\text{quasispecial}}$.*

Proof. Since V is a curve, Z is a one-dimensional complex analytic set and does not contain $\mathcal{H}^2$. Suppose that Z contains a non-empty $C \cap \mathcal{H}^2$ for some irreducible complex algebraic curve C. If τ_1 (or τ_2) is constant on C, then it is quasispecial and belongs to $Z^{\text{quasispecial}}$. Otherwise suppose that $\tau_2 = \phi(\tau_1)$ is an algebraic function such that $G(j(\tau_1), j(\phi(\tau_1)) = 0$, where G is a bivariate polynomial defining V. Since j has essential singularities densely on the real line, but is regular in $\mathcal{H}$, there cannot be an algebraic relation between the functions $j(\tau_1)$ and $j(\phi(\tau_1))$ in a neighbourhood of a point where τ_1 is real and $\phi(\tau_1)$ is not real, or vice versa. Therefore the algebraic function $\phi(\tau_1)$ cannot map any point of $\mathcal{H}$ to $\mathbb{R}$, and the same holds for its inverse.

I claim that ϕ is a fractional linear transformation (with real coefficients, after possible scaling, and positive determinant, as it preserves $\mathcal{H}$).

To see this, choose τ_0 such that the equation $G(j(\tau_0), y)$ has its maximum number of distinct roots y_i. Then $G(j(\tau_0), j(\phi(\tau_0)) = 0$. If $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ then $G(j(\gamma\tau_0), j(\phi(\gamma\tau_0))) = 0$ also by our assumptions. However $j(\gamma\tau_0) = j(\tau_0)$ by the properties of j, and so $\phi(\gamma\tau_0)$ must be equivalent, under $\mathrm{SL}_2(\mathbb{Z})$, to one of the roots y_i. As there are just finitely many y_i, there must exist in particular $n \neq m$ such that

$$\phi(\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \tau_0), \quad \phi(\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \tau_0)$$

are $\mathrm{SL}_2(\mathbb{Z})$ equivalent, i.e. there is $\gamma' \in \mathrm{SL}_2(\mathbb{Z})$, with

$$\phi(\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \tau_0) = \gamma'\phi(\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \tau_0).$$

Now vary τ in a small neighbourhood of τ_0 so that the roots of $G(j(\tau), y)$ remain distinct. Then the two roots

$$j(\phi \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \tau), \quad j(\phi \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \tau)$$

of $G(j(\tau), y)$ must be $\mathrm{SL}_2(\mathbb{Z})$ equivalent for all τ in this neighbourhood, and so by the discreteness of $\mathrm{SL}_2(\mathbb{Z})$ we conclude that

$$\phi(\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \tau) = \gamma'\phi(\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \tau)$$

for all τ in the neighbourhood, and hence identically.

Replacing τ by $\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \tau$ and putting $N = n - m$ we then have $\phi \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} = \gamma'\phi$ (in the first place up to a scaling, but both sides have the same determinant).

Next, by composing ϕ with a suitable element of $\mathrm{SL}_2(\mathbb{R})$ we may assume the composition ψ maps ∞ to ∞, and satisfies

$$\psi \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} = \gamma''\psi$$

where now $\gamma'' \in \mathrm{SL}_2(\mathbb{R})$. By considering large $|\tau|$ it follows that $\gamma''(\tau) = a\tau + b$ for some $a, b \in \mathbb{R}$. So we have an algebraic function ψ satisfying

$$\psi(\tau + N) = a\psi(\tau) + b.$$

Taking a Puiseux expansion one sees rapidly that ψ must itself be of the form $\psi(\tau) = c\tau + d$ for some $c, d \in \mathbb{R}$, and it follows that ϕ is given by an element of $\mathrm{SL}_2(\mathbb{R})$.

I claim that the image of ϕ in $\mathrm{PSL}_2(\mathbb{R})$ is in the image of $\mathrm{GL}_2(\mathbb{Q})$, i.e. that the fractional linear transformation can be effected with rational coefficients. From the above argument we have

$$\phi \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} = \gamma' \phi$$

for some $\gamma' \in \mathrm{SL}_2(\mathbb{Z})$. Writing

$$\phi = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

we have

$$\phi \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \phi^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{N}{\alpha\delta - \beta\gamma} \begin{pmatrix} -\alpha\gamma & \alpha^2 \\ -\gamma^2 & \alpha\gamma \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Therefore, if $\gamma \neq 0$ then $\alpha/\gamma \in \mathbb{Q}$, while if $\gamma = 0$ then ($\delta \neq 0$ and) $\alpha/\delta \in \mathbb{Q}$. Now if $j(\tau)$ and $j(\phi\tau)$ are algebraically related then so are $j(\tau)$ and $j(\phi\gamma\tau)$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ (with the same algebraic relation), and so the above now implies that the ratio $(\alpha + m\beta)/(\gamma + m\delta) \in \mathbb{Q}$ for all integers m. The ratio is not constant (as $\alpha/\gamma \neq \beta/\delta$ since ϕ is invertible) and so taking two distinct values of m we conclude that $\beta/\gamma, \delta/\gamma \in \mathbb{Q}$, if $\gamma \neq 0$. If $\gamma = 0$ we already know $\alpha/\delta \in \mathbb{Q}$ and now get $\beta/\delta \in \mathbb{Q}$. Therefore the image of ϕ in $\mathrm{PSL}_2(\mathbb{R})$ is in the image of $\mathrm{GL}_2(\mathbb{Q})$ and the locus $(\tau, \phi\tau)$ is the special subvariety Z_ϕ.

Thus Z^{alg} is contained in the union of sets V_τ, H_τ, Z_ϕ (with ϕ as above) contained in Z, which is $Z^{\mathrm{quasispecial}}$. □

2.4. Remarks.

1. To make the present proof effective one would need, in addition to GRHIQ (or an effective $D^\delta, \delta > 0$ lower bound for class numbers of imaginary quadratic fields), effectivity for the bounds implicit in the definability result (such as e.g. explicit bounds for the maximum number of zeros $\tau \in \mathcal{D}$ of a polynomial $F(\tau, j^{(k)}(\tau))$ in terms of the degree of F and k). Some such explicit bounds [4, 35, 40] have been obtained in connection with transcendence measures for the j-function.

2. I conjecture that $Z^{\mathrm{alg}} = Z^{\mathrm{quasispecial}}$ for any complex analytic $\mathrm{SL}_2(\mathbb{Z})^2$ invariant set $Z \subset \mathcal{H}^2$ such that $\mathcal{Z} = Z \cap \mathcal{D}^2$ is definable in an o-minimal structure. This would be analogous to the result established in [45], where the intersection of an analytic set in $\mathbb{C}^m$ with a fundamental domain for a period lattice is always definable in $\mathbb{R}_{\mathrm{an}}$.

3. More generally, I would conjecture the corresponding result for sets $Z = \pi^{-1}(V) \subset \mathcal{H}^n$, where $\pi : \mathcal{H}^n \to \mathbb{C}^n$ is the j-function on each coordinate and $V \subset \mathbb{C}^n$ is a closed algebraic set defined over $\mathbb{C}$. Let us call two matrices $g, h \in \mathrm{SL}_2(\mathbb{R})$ *dependent over* $\mathrm{GL}_2(\mathbb{Q})$ if the matrix gh^{-1} has the property that the ratio of any two entries is rational (so its image in $\mathrm{PSL}_2(\mathbb{R})$ is in the image of $\mathrm{GL}_2(\mathbb{Q})$). A locus $(\tau_1, \ldots, \tau_m) \in \mathcal{H}^m$ we call a *basic quasispecial locus* if $\tau_i = g_{ij}\tau_j$ for every i, j where the matrices g_{ij} are pairwise dependent over $\mathrm{GL}_2(\mathbb{Q})$. We allow $m = 1$, so that $\mathcal{H}$ is a basic quasispecial locus. Let $S = S^{\mathrm{point}} \cup S_1 \cup \ldots S_k$ be a disjoint partition of $\{1, \ldots, n\}$ allowing $k = 0$ and allowing S^{point} to be empty. For each $i \in S^{\mathrm{point}}$ take a point $\tau_i^S \in \mathcal{H}$, and for each $S_i \neq \emptyset$ we choose a basic quasispecial locus L_i on the variables $\tau^{S,i} = \{\tau_j : j \in S_i\}$. We define Z_S to be the locus of points

$$Z_S = \{(\tau_1, \ldots, \tau_n) : \tau_i = \tau_i^S \text{ all } i \in S^{\mathrm{point}}, \tau^{S,i} \in L_i \text{ all } i = 0, \ldots, k\}.$$

For a set $Z \subset \mathcal{H}^n$ let $Z^{\mathrm{quasispecial}}$ be the union of all the Z_S of positive dimension contained in Z. (Cf the definition of special subvarieties of products of modular curves set out in [23]. Our definition again differs only in allowing points corresponding to coordinates in S^{point} to be non-CM.) I conjecture then that $Z^{\mathrm{alg}} = Z^{\mathrm{quasispecial}}$ when $Z = \pi^{-1}(V)$, and indeed more generally when Z is any $\mathrm{SL}_2(\mathbb{Z})^n$-invariant analytic set with $Z \cap \mathcal{D}^n$ definable in an o-minimal structure.

Note added. Kobi Peterzil [37] has informed me that the conjecture in 2.4.2 is true in that it reduces to the case that $\pi(Z)$ is algebraic which is established in Proposition 2.3. Namely, it may be shown, using results from [39], that if $Z \subset \mathcal{H}^n$ is complex analytic and $\mathrm{SL}_2(\mathbb{Z})^n$ invariant with $Z \cap \mathcal{D}^n$ definable in some o-minimal structure, then $\pi(Z)$ is an algebraic subset of $\mathbb{C}^n$. Likewise, in 2.4.3, the more general assertion would follow from the assertion for $Z = \pi^{-1}(V)$ for algebraic V.

4. The assertion of 2.4.3 restricted to those Z arising as preimages of algebraic subsets $V \subset \mathbb{C}^n$ (even those defined over $\overline{\mathbb{Q}}$) yields a proof (unconditional but ineffective) of the corresponding case of the André-Oort conjecture, namely for subvarieties of $\mathbb{C}^n$, or, more generally, for subvarieties of products of modular curves (a result of Edixhoven [23] under GRHIQ) by a proof analogous to the proof of 1.1 given here. As already mentioned, such higher-dimensional generalizations of the present results will be pursued elsewhere.

Moving to the corresponding result required for 1.2, we need suitable functions, analogous to the Weierstarss $\wp$-function, giving the uniformization of the quasi-projective set $A \subset \mathbb{A}^1 \times \mathbb{P}^2$ defined by

$$Y^2Z - X(X-Z)(X-\lambda Z) = 0, \quad \lambda \neq 0, 1.$$

We set up some notation while recalling some parts of the classical theory of elliptic modular functions to define these functions with a view to observing the definability properties we need in §3.

Recall that, for a lattice $\Lambda_{\omega_1,\omega_2} = \mathbb{Z}\omega_1 \oplus \mathbb{Z}\omega_2 \subset \mathbb{C}$ with basis ω_1, ω_2 (ordered so that $\tau = \omega_2/\omega_1 \in \mathcal{H}$), the Weierstrass function $\wp(z;\omega_1,\omega_2)$ is $\Lambda_{\omega_1,\omega_2}$ periodic in z, while, for any $c \in \mathbb{C} - \{0\}$,

$$\wp(z/c;\omega_1/c,\omega_2/c) = c^2\wp(z;\omega_1,\omega_2).$$

We define

$$\wp(z,\tau) = \wp(z;1,\tau).$$

Then, for fixed τ, the function $\wp(z,\tau)$ is $\Lambda_{1,\tau}$ periodic in z, while, for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$,

$$\wp(\frac{z}{c\tau+d}, \gamma\tau) = \wp(\frac{z}{c\tau+d}; \frac{c\tau+d}{c\tau+d}, \frac{a\tau+b}{c\tau+d}) = (c\tau+d)^2\wp(z; c\tau+d, a\tau+b) = (c\tau+d)^2\wp(z,\tau).$$

The function $\wp(z,\tau)$ and its z-derivative

$$\wp(z,\tau)'(z,\tau) = \frac{d}{dz}\wp(z,\tau)$$

satisfy an algebraic relation

$$(\wp(z,\tau)')^2 = 4(\wp(z,\tau))^3 - g_2(\tau)\wp(z,\tau) - g_3(\tau)$$

by virtue of which the map

$$z \mapsto \begin{cases} (\wp(z,\tau), \wp'(z,\tau), 1), & z \notin \Lambda_{1,\tau} \\ \infty = (0,1,0), & z \in \Lambda_{1,\tau} \end{cases}$$

parameterizes an elliptic curve (as the discriminant $\Delta(\tau)$ of the cubic is nowhere vanishing for $\tau \in \mathcal{H}$). The functions $\wp(z,\tau), \wp'(z,\tau)$ are meromorphic on $\mathcal{H} \times \mathbb{C}$ indeed, for each τ, they are regular except for poles at the lattice points $\Lambda_{1,\tau}$.

The roots of the cubic are given by the $\wp$-function at half periods

$$e_1(\tau) = \wp(\frac{1}{2},\tau), \quad e_2(\tau) = \wp(\frac{1+\tau}{2},\tau), \quad e_3(\tau) = \wp(\frac{\tau}{2},\tau),$$

so that

$$(\wp'(z,\tau))^2 = 4\big(\wp(z,\tau) - e_1(\tau)\big)\big(\wp(z,\tau) - e_2(\tau)\big)\big(\wp(z,\tau) - e_3(\tau)\big).$$

Define

$$\xi(z,\tau) = \frac{\wp(z,\tau) - e_1(\tau)}{-(e_1(\tau) - e_3(\tau))}.$$

This function is again meromorphic in $\mathbb{C} \times \mathcal{H}$ as $e_1 - e_3$ never vanishes, and for this reason also the function $e_1 - e_3$ has a regular square-root for $\tau \in \mathcal{H}$. Putting

$$\eta(z,\tau) = 2\big(e_3(\tau) - e_1(\tau)\big)^{3/2}\wp'(z,\tau),$$

we have

$$\eta^2(z,\tau) = \xi(z,\tau)(\xi(z,\tau) - 1)(\xi(z,\tau) - \lambda(\tau))$$

where

$$\lambda(\tau) = \frac{e_2(\tau) - e_3(\tau)}{e_1(\tau) - e_3(\tau)}.$$

We observe that since $e_1(\tau), e_2(\tau), e_3(\tau)$ are distinct for any $\tau \in \mathcal{H}$ (as the discriminant of the cubic is non-zero), $\lambda \neq 0, 1$ but ([26]) takes on every other complex value.

Accordingly, the map $\pi : \mathcal{H} \times \mathbb{C} \to A$ given by

$$\pi(z,\tau) = \begin{cases} (\lambda(\tau), \xi(z,\tau), \eta(z,\tau,1), 1), & z \notin \Lambda_{1,\tau} \\ (\lambda(\tau), 0, 1, 0), & z \in \Lambda_{1,\tau} \end{cases}$$

gives a parameterization of the curves in the Legendre family (the point $(\lambda, 0, 1, 0)$ corresponds to the point at infinity on $E_{\lambda(\tau)}$).

The j-invariant of E_λ is

$$2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}$$

i.e. we have, for $\tau \in \mathcal{H}$,

$$j(\tau) = 2^8 \frac{(\lambda(\tau)^2 - \lambda(\tau) + 1)^3}{\lambda(\tau)^2(\lambda(\tau) - 1)^2}.$$

We have $\lambda(\gamma\tau) = \lambda(\tau)$ for γ in a certain finite index subgroup Γ_λ of $\mathrm{SL}_2(\mathbb{Z})$, namely those $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ for which a, b are odd and b, c are even [26].

We let $\mathcal{B} \subset \mathcal{H}$ be a fundamental domain for Γ_λ, as described in [26], which consists of a suitable choice of 6 fundamental domains for $\mathrm{SL}_2(\mathbb{Z})$. For $\gamma \in \Gamma_\lambda$,

$$\xi(\frac{z}{c\tau + d}, \gamma\tau) = \xi(z,\tau), \quad \eta(\frac{z}{c\tau + d}, \gamma\tau) = \eta(z,\tau)$$

(the factor $(c\tau + d)^2$ in the transformation of $\wp$ cancels with a similar factor from the $e_i(\tau)$). For given τ, the functions $\xi(z,\tau), \eta(z,\tau)$ are $\Lambda_{1,\tau}$-periodic. Therefore, the uniformization $\pi : \mathcal{H} \times \mathbb{C} \to A$ is automorphic for a semidirect product Γ of Γ_λ with (normalizing) $\mathbb{Z}^2$, and our sets Z will have the following property with respect to Γ:

$$(\tau, z) \in Z \to (\tau, z + \lambda) \in Z, \quad \text{all } \lambda \in \Lambda_{1,\tau},$$

$$(\tau, z) \in Z \to (\gamma\tau, \frac{z}{c\tau + d}) \in Z, \quad \text{all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\lambda.$$

For $\tau \in \mathcal{H}$ we define $V_\tau = \{(\tau, z) : z \in \mathbb{C}\}$. For $r, s \in \mathbb{Q}$ we define $H_{r,s} = \{(\tau, r\tau + s) : \tau \in \mathcal{H}\}$. These are our analogues of special subvarieties of $\mathcal{H} \times \mathbb{C}$ of dimension 1. Again, the V_τ with τ not a CM point do not correspond to special subvarieties (i.e. CM curves $E_{\lambda(\tau)}$) of A.

2.5. Definition. For a set $Z \subset \mathcal{H} \times \mathbb{C}$ we define $Z^{\text{quasispecial}}$ (for Γ) to be the union of all sets $V_\tau, H_{r,s}, \mathcal{H} \times \mathbb{C}$ contained in Z.

We write $\tau = u + iv$ but now write $z = x + \tau y$ to get real coordinates (u, v, x, y) on $\mathbb{C}^2$ in such a way that a point (u, v, x, y) with rational (x, y) corresponds to a torsion point of the corresponding lattice $\Lambda_{1,\tau}, \tau = u + iv$. We have $Z^{\text{ca}} = Z^{\text{alg}}$ for any set complex analytic Z. The sets $V_\tau, H_{r,s}$ are complex algebraic, so we have again

$$Z^{\text{quasispecial}} \subset Z^{\text{ca}} = Z^{\text{alg}}.$$

2.6. Proposition. *Let $V \subset A$ be an algebraic curve defined over $\mathbb{C}$ and $Z = \pi^{-1}(V) \subset \mathcal{H} \times \mathbb{C}$. Then $Z^{\text{alg}} = Z^{\text{quasispecial}}$.*

Proof. Let the curve $V \subset A$ be defined by some equations $G_i = 0$ where the polynomials G_i are homogeneous in the $\mathbb{P}^2$ variables. Let us first observe that Z is an analytic set. From the definition we have given this is evident only away from the poles of ξ, η, i.e. for points (τ, z) such that $z \notin \Lambda_{1,\tau}$. However, we can alternatively define Z as the zero set of a suitable expression in theta functions and the e_i that are regular everywhere on $\mathcal{H} \times \mathbb{C}$. As already observed, by Lemma 2.1, $Z^{\text{alg}} = Z^{\text{ca}}$.

Now suppose we have some complex algebraic curve contained in Z. If τ is constant on this curve, then it is $\{\tau\} \times \mathbb{C}$ and is quasispecial. Otherwise, we have some non-constant algebraic function $\phi(\tau)$ such that $\{(\tau, \phi(\tau))\} \subset Z$, i.e. $G_i(\lambda(\tau), \xi(\phi(\tau), \tau), \eta(\phi(\tau), \tau), 1) = 0$, for all i, away from poles.

Let G be one of the G_i. Generically, given λ, there are only finitely many points P on E_λ such that $G(\lambda, P) = 0$, except for any vertical components of V. Indeed, no more than the degree of the curve V as a covering of the affine line. So there exists an integer K such that for all but finitely many values of λ_0, the points $(\tau, z) \in Z$ with $\lambda(\tau) = \lambda_0$ correspond to at most K distinct points of E_{λ_0}. The function $\phi(\tau)$ has a convergent Puiseaux expansion for $|\tau|$ sufficiently large.

Now

$$\begin{pmatrix} 1 & 2k \\ 0 & 1 \end{pmatrix} \in \Gamma_\lambda$$

for any $k \in \mathbb{Z}$, so $\tau, \tau + 2, \tau + 4, \ldots$ give the same value of λ and the same lattice, so among the points $\phi(\tau + 2k), k = 0, \ldots, K$ at least two coincide mod $\Lambda_{1,\tau}$. Thus there exists $2k > 0$ such that $\phi(\tau + 2k) - \phi(\tau) \in \Lambda_\tau$ for a dense set of τ in some small neighbourhood, and so $a, b \in \mathbb{Z}$ are constant and $\phi(\tau + 2k) - \phi(\tau) = a\tau + b$ identically in τ. So the Puiseaux series has no terms with exponent > 2. Now by using binomial series to expand terms $(\tau + 2k)^\alpha$ we get a Puiseaux series for $\phi(\tau + 2k)$ and we see that $\phi(\tau) = q\tau^2 + r\tau + s$ for some $q, r, s \in \mathbb{C}$ with in fact $q, r \in \mathbb{Q}$.

We need to show $q = 0, s \in \mathbb{Q}$. We have $\lambda(\gamma\tau) = \lambda(\tau)$ for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\lambda$. For such γ the point $(\gamma\tau, \phi(\gamma\tau)$ is also in Z, and therefore the point $(\tau, (c\tau + d)\phi(\gamma\tau)) \in Z$. Considering again $\tau + 2, 4, \ldots$, the mapping $\tau \mapsto (c\tau + d)\phi(\gamma\tau)$ must have the same form established for ϕ, and this shows that $q = 0$ and $s \in \mathbb{Q}$. □

For 1.3, we have again $U = \mathcal{H} \times \mathbb{C}$, but $\mathbb{C}$ now uniformizes the fixed elliptic curve E, with the lattice points mapping to the point at infinity of E, while $\mathcal{H}$ uniformizes the modular curve (except for the finitely many images of cusps). The map $\mathbb{C} \to E$ is periodic with period lattice $\Lambda = \Lambda_{\omega_1,\omega_2}$, and the map $\mathcal{H} \to X_0(N)$ is invariant under $\Gamma_0(N)$. So here $\Gamma = \Gamma_0(N) \times \Lambda$. For $\tau \in \mathcal{H}$ we set $V_\tau = \{\tau\} \times \mathbb{C}$ and for $z \in \mathbb{C}$ we set $H_z = \mathcal{H} \times \{z\}$. These will be our analogues of special subvarieties of positive dimension, though again the sets V_τ with τ not special, and the sets H_z with z not torsion with respect to Λ are not preimages of special subvarieties of $X_0(N) \times E$.

2.7. Definition. For a set $Z \subset \mathcal{H} \times \mathbb{C}$ we define $Z^{\text{quasispecial}}$ (for Γ) to be the union of all sets $V_\tau, H_z, \mathcal{H} \times \mathbb{C}$ contained in Z.

We write $\tau = u + iv$ and $z = x\omega_1 + y\omega_2$ giving real coordinates (u, v, x, y) on $\mathcal{H} \times \mathbb{C}$. We have again

$$Z^{\text{quasispecial}} \subset Z^{\text{ca}} = Z^{\text{alg}},$$

and we prove the corresponding equality.

2.8. Proposition. *Let $V \subset X_0(N) \times E$ be an algebraic curve, $Z = \pi^{-1}(V)$. Then $Z^{\rm alg} = Z^{\rm quasispecial}$.*

Proof. The function field of $X_0(N)$ is $\mathbb{C}(j(\tau), j(N\tau))$. Using suitable elements $r_i \in \mathbb{C}(j(\tau))$ as coordinates, we can give a projective embedding of $X_0(N)$. The curve V is defined by some equations in these coordinates, and these correspond to some polynomial relations between the functions $\wp(z;\omega_1,\omega_2), \wp'(z;\omega_1,\omega_2), r_i$. Writing $\wp$ as a ratio of theta functions we can at any point of $Y_0(N) \times E$ express these equations locally using regular functions. The set $Z = \pi^{-1}(V \cap Y_0(N))$ is then a (relatively) closed analytic subset of $\mathcal{H} \times \mathbb{C}$. We have $Z^{\rm alg} = Z^{\rm ca}$ by Lemma 2.1 as previously.

Suppose now we have a complex algebraic curve contained in Z. If τ is constant on this curve, then it is $\{\tau\} \times \mathbb{C}$ and quasispecial. Otherwise we have a non-constant algebraic function $z = \phi(\tau)$ such that $\{(\tau, \phi(\tau))\} \subset Z$, and ϕ has a Puiseaux series convergent for sufficiently large $|\tau|$.

For all but finitely many pairs, given $(j(\tau), j(N\tau))$ there are only finitely many points z modulo Λ such that $(\tau, z) \in Z$. The pair $(j(\tau), j(N\tau))$ is invariant under $\Gamma_0(N)$. Since $\Gamma_0(N)$ contains translation by N, this proves, as in the proof of 2.6, that ϕ must be of the form $\phi(\tau) = \ell\tau + m$ where ℓ is a torsion point with respect to Λ. Considering general elements $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$ one finds that $(c\tau + d)\phi(\gamma\tau)$ must have the same form as above, so that m is also a torsion point with respect to Λ.

We can find a complex number K with positive imaginary part such that $K\ell$ is a lattice point. On the other hand, $M = nN\ell$ is a lattice point for a suitable integer n. Then the points $\tau + \kappa K + \mu M$ have the same image in E, but correspond to infinitely many different points in a fundamental domain for $\Gamma_0(N)$. But this is impossible, as for an algebraic subset $V \subset A$, the number of points of $X_0(N)$ corresponding to a point in E is finite unless V contains a vertical component. A map ϕ as above with $\ell \neq 0$ would mean that V has vertical fibres at every image point, i.e. $V = A$. If $V = A$ then $Z = \mathcal{H} \times \mathbb{C} = Z^{\rm quasispecial}$, otherwise $\ell = 0$ and the curve $z = \phi(\tau)$ is constant and contained in $Z^{\rm quasispecial}$. □

2.9. Remark. In each of the settings 1.1, 1.2, 1.3, considering the possible algebraic subsets of an analytic $Z \subset U$ automorphic for the appropriate group led to a definition (2.2, 2.5, 2.7) of $Z^{\rm quasispecial}$ that almost corresponds to the inverse image of the special set when $Z = \pi^{-1}(V)$, and coincides so far as special points are concerned in that the additional sets we admit to do not contain any pre-images of special points. The same phenomenon occured in [45] in considering possible algebraic subsets of analytic sets $Z \subset \mathbb{C}^n$ periodic under a lattice Λ. There we found $Z^{\rm alg} = Z^{\rm torus\ coset}$, the union of cosets of subtori contained in Z. When $\mathbb{C}^n/\Lambda = A$ is an abelian variety, torus cosets correspond to special subvarieties of A just if the coset is a torsion coset, i.e. just if it contains special points.

Note added: For the Shimura varieties A of 1.1, i.e. $\mathbb{C}^2$ or the product of two modular curves, our $Z^{\rm quasispecial}$ apparently coincides with the union of preimages of *totally geodesic subvarieties* of A contained in V, as defined by Moonen [33], which we might denote $Z^{\rm geodesic}$. One might then expect that the identification $Z^{\rm geodesic} = Z^{\rm alg}$ holds more generally. For $\mathbb{C}^n$ or the product of several modular curves this is the content of 2.4.3 when $V \subset \mathbb{C}^n$ is algebraic, and one could expect that it holds still much more generally. It is striking that the algebraic part of Z, a very coarse analogue of the special set, turns out in these circumstances to essentially coincide with its preimage.

3. Definability

An o-minimal structure is, informally, a collection of subsets of $\mathbb{R}^n, n = 1, 2, \ldots$ that is sufficiently rich and flexible to be closed under some basic operations on the one hand, while enjoying strong geometric finiteness properties on the other hand. The paradigm example is the collection of real semi-algebraic sets. Formally, the definitions are as follows (following [51]), but the reader should refer to [17] and the other references cited below for further explication.

3.1. Definition. A *pre-structure* is a sequence $\mathcal{S} = (S_n : n \geq 1)$ where each $\mathcal{S}_n$ is a collection of subsets of $\mathbb{R}^n$. A pre-structure $\mathcal{S}$ is called a *structure (over the real field)* if, for all $n, m \geq 1$, the following conditions are satisfied:

(1) $\mathcal{S}_n$ is a boolean algebra (under the usual set-theoretic operations)
(2) $\mathcal{S}_n$ contains every semi-algebraic subset of $\mathbb{R}^n$
(3) if $A \in \mathcal{S}_n$ and $B \in \mathcal{S}_m$ then $A \times B \in \mathcal{S}_{n+m}$
(4) if $m \geq n$ and $A \in \mathcal{S}_m$ then $\pi(A) \in \mathcal{S}_n$, where $\pi : \mathbb{R}^m \to \mathbb{R}^n$ is projection onto the first n coordinates.

If $\mathcal{S}$ is a structure and $X \subset \mathbb{R}^n$, we say X is *definable in* $\mathcal{S}$ if $X \in \mathcal{S}_n$. If $\mathcal{S}$ is a structure and, in addition,

(5) the boundary of every set in $\mathcal{S}_1$ is finite

then $\mathcal{S}$ is called an *o-minimal* structure (over the real field).

The fact that the semi-algebraic sets form an o-minimal struicture follows from the Tarski-Seidenberg theorem that the projection of a semi-algebraic set is semi-algebraic. Another example is afforded by the *globally subanalytic* sets, $\mathbb{R}_{\text{an}}$. (The globally subanalytic sets are those sets $X \subset \mathbb{R}^n$ that are subanalytic when considered as subsets of $\mathbb{P}^n(\mathbb{R})$, see e.g. [20].) Here the main difficulty in establishing o-minimality is the closure under complementation, which is furnished by a theorem of Gabrielov [27]. The subsets of $\mathbb{R}^n$ definable using the exponential function form an o-minimal structure, denoted $\mathbb{R}_{\text{exp}}$. Its o-minimality is due to Wilkie [50], the main difficulty again being behaviour under complementation (the requisite finiteness properties then follow from Khovanskii's theory of "Fewnomials" [29]). Neither of these structures contains the other: for example $\mathbb{R}_{\text{exp}}$ includes sets such as

$$\{(x, \exp(-1/x)), x > 0\}, \quad \{(x, x^r), x > 0\}$$

for positive irrational r that are not subanalytic (at the origin) in $\mathbb{R}^2$, and hence are not in $\mathbb{R}_{\text{an}}$ ([19]). However, the structure $\mathbb{R}_{\text{an,exp}}$ generated by $\mathbb{R}_{\text{an}}$ and $\mathbb{R}_{\text{exp}}$ is o-minimal ([19], see also [18]). For further examples, see e.g. [51, 48], the latter reference shows there is no largest o-minimal structure over $\mathbb{R}$.

Let us state the result of Peterzil-Starchenko [38]. In the theorem below and subsequent discussion, "definable" means definable in $\mathbb{R}_{\text{an,exp}}$.

3.2. Theorem. ([38, Theorem 4.1]) *The restriction of the function $\wp(z, \tau)$ to $(\tau, z) \in \mathcal{F}$ is definable.* □

Then $\wp(z, \tau)$ is definable when $\mathcal{D}$ is replaced by any other fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$, and by the union of a finite number of such domains. The fundamental domain $\mathcal{B} \subset \mathcal{H}$ (for the λ-function) consists of finitely many (in fact 6) fundamental domains for $\mathrm{SL}_2(\mathbb{Z})$. So $\wp(z, \tau)$ is also definable on $\{(z, \tau) : (\tau, z) \in \mathcal{F}_\mathcal{B}\}$ where

$$\mathcal{F}_\mathcal{B} = \{(\tau, z) \in \mathcal{H} \times \mathbb{C} : \tau \in \mathcal{B}, z \in \mathcal{L}_\tau\}.$$

The functions $e_1(\tau), e_2(\tau), e_3(\tau) : \mathcal{B} \to \mathbb{C}$ are then also definable. Then the function $\lambda(\tau) : \mathcal{B} \to \mathbb{C}$ is definable, and so are $\xi(z, \tau), \eta(z, \tau)$ for $(\tau, z) \in \mathcal{F}_\mathcal{B}$ (the requisite square-root is definable). Finally the map $j : \mathcal{D} \to \mathbb{C}$ (or $\mathcal{B} \to \mathbb{C}$) is also definable as it is a rational function (exhibited above) of $\lambda(\tau)$ (this is also observed in [38, §5.4]).

The sets $\mathcal{Z}$ will in each case be defined (except perhaps at finitely many points) by the vanishing of some polynomials in definable functions, and hence these sets will be definable as well.

4. Proofs of Theorems

The point $j(\tau)$ is special just when $\tau \in \mathcal{H}$ belongs to a quadratic (imaginary) algebraic number field. Consider such τ, the root of some irreducible polynomial $ax^2 + bx + c$ where $a, b, c \in \mathbb{Z}$ and where we may assume $a > 0$ and $(a, b, c) = 1$. Since $\tau \notin \mathbb{R}$ we must have $c > 0$ also. We will refer to the discriminant $D = b^2 - 4ac$ of the quadratic as the *discriminant* of τ. Then $\Lambda_{1,\tau}$ has CM by the order $\mathcal{O}_D = \mathbb{Z}[(D + \sqrt{D})/2]$ of discriminant D in the ring of integers of the field $\mathbb{Q}(\sqrt{D})$. The number of imaginary quadratic $\tau \in \mathcal{H}$ of discriminant D that are inequivalent under the action of $\mathrm{SL}_2(\mathbb{Z})$ (acting on τ as fractional linear transformations) is equal to the class number $h(D)$. According to the theory of complex multiplication of elliptic curves (see e.g. [11], in particular Ch. II), the corresponding $j(\tau)$ are a complete set of conjugate algebraic integers over $\mathbb{Q}(\tau)$, and also over $\mathbb{Q}$. Thus $h(D) = [\mathbb{Q}(j(\tau)) : \mathbb{Q}]$.

For the size $h(D)$ of the class group of an imaginary quadratic order we have Siegel's result that, given $\nu > 0$ there is an (ineffective) constant $c(\nu)$ such that

$$h(D) \geq c(\nu)|D|^{1/2-\nu}.$$

The upper bound

$$h(D) \leq C(\nu)|D|^{1/2+\nu}$$

holds effectively (we need this in the proof of 1.2).

By $H(\alpha)$ we denote the absolute multiplicative height of an algebraic number, i.e. $H(\alpha) = \exp h(\alpha)$ where $h(\alpha)$ is the absolute logarithmic height of as defined in [6, p16] (not to be confused with the class number $h(D)$). We extend the height H to tuples by setting $H(\alpha_1, \dots, \alpha_n) = \max_i H(\alpha_i)$.

If $\tau \in \mathcal{D}$ then $|\tau| \geq 1$ so that, by the formulae in [6, 1.6.5, 1.6.6], we have

$$H(\tau)^2 = a \prod_{\alpha=\tau,\overline{\tau}} \max(1, |\alpha|) = ac.$$

We need to relate the height and discriminant for $\tau \in \mathcal{D}$. The condition that $\tau \in \mathcal{D}$ is the condition that the integer triple (or the corresponding binary form) (a, b, c) is reduced, namely $|b| \leq a \leq c$ (and $b \geq 0$ if $b = a$ or $a = c$). Then from $b^2 - 4ac = D$ and the above inequalities we find $4ac = b^2 - D \leq ac - D$, whence

$$H(\tau)^2 = ac \leq |D|.$$

In counting quadratic algebraic points up to a given height in $\mathcal{Z}$, we consider Z as a real surface in $\mathbb{R}^4$, and thus we consider the real and complex parts of τ. When τ is quadratic, so too are its real and imaginary parts. Indeed, writing $\tau = u + iv$ we see that $u = -b/2$ is rational with

$$H(u) \leq 2|b| \leq 2|b|^2 \leq 8ac = 8H(\tau)$$

and $v = \sqrt{D}/2$ is a root of $4(x - \sqrt{D}/2)(x + \sqrt{D}/2) = 4x^2 - D$ whence

$$H(v) \leq 4|D| \leq 16ac \leq 16H(\tau).$$

4.1. Proof of Theorem 1.1. We have $\pi(\tau_1, \tau_2) = (j(\tau_1), j(\tau_2))$ mapping $\mathcal{H}^2$ onto $\mathbb{C}^2$ and $Z = \pi^{-1}(V)$. We consider Z as a two-real-dimensional subset of $\mathbb{R}^4$ in the coordinates (u, v, x, y) where $\tau_1 = u + iv$ and $\tau_2 = x + iy$. We let $\mathcal{Z} = Z \cap \mathcal{D}^2$. If V contains no special subvarieties of $\mathbb{C}^2$ of positive dimension, then $Z^{\text{alg}} = Z^{\text{quasispecial}}$ is either empty or contains some sets V_τ, H_τ that have no preimages of special points in them, so that $\mathcal{Z}^{\text{alg}}$ contains no pre-images of special points. We will show that $\mathcal{Z}$ contains only finitely many pre-images of special points. As special points are algebraic, the finiteness is clear for any components of V that are not defined over $\overline{\mathbb{Q}}$. Removing such components, we may assume then that V is defined over $\overline{\mathbb{Q}}$.

Suppose that V is defined by $G = 0$ for some bivariate polynomial G. Then

$$\mathcal{Z} = \{(\tau_1, \tau_2) \in \mathcal{D}^2 : G(j(\tau_1), j(\tau_2)) = 0\}$$

and is therefore definable in $\mathbb{R}_{\text{an,exp}}$.

The pre-images of special points are the points $(u, v, x, y) \in \mathcal{Z}$ for which $u+iv, x+iy$ are quadratic, and are a subset of the points $(u, v, x, y) \in \mathcal{Z}$ whose coordinates are all of degree ≤ 2. Let us denote by $N_2^{\text{special}}(\mathcal{Z}, T)$ the number of points in $\mathcal{Z}(2, T)$ that are preimages of special points. By Theorem 1.5 we have that

$$N_2(\mathcal{Z} - \mathcal{Z}^{\text{alg}}, T) \leq c(\mathcal{Z}, \epsilon)T^\epsilon,$$

for any positive ϵ.

The same estimate clearly holds for $N_2^{\text{special}}(\mathcal{Z}-\mathcal{Z}^{\text{alg}},T)$, and since $\mathcal{Z}^{\text{alg}}$ contains no preimages of special points we conclude that

$$N_2^{\text{special}}(\mathcal{Z},T)\le c(\mathcal{Z},\epsilon)T^{\epsilon}$$

for every positive ϵ.

Suppose that $\mathcal{Z}$ contains a point (τ_1,τ_2) such that τ_1 is quadratic with discriminant D_1, and τ_2 is quadratic with discriminant D_2. Then $H(u,v,x,y)\le 16\max(\sqrt{|D_1|},\sqrt{|D_2|})$. The number of conjugates of $j(\tau_1)$ is $h(D_1)$ and likewise for $j(\tau_2)$. The image of (τ_1,τ_2) in $\mathbb{C}^2$ has therefore, for any $\nu>0$, at least

$$c(\nu)\max(|D_1|^{1/2-\nu},|D_2|^{1/2-\nu})$$

conjugates of which at least

$$c(V)c(\nu)\max(|D_1|^{1/2-\nu},|D_2|^{1/2-\nu})$$

lie on V, where $c(V)$ is the degree over $\mathbb{Q}$ of a numberfield of definition of V. These conjugate have distinct preimages (τ_1',τ_2') in $\mathcal{D}^2$. The $\tau_1'=u'+iv'$ have discriminant D_1, and the $\tau_2'=x'+iy'$ have discriminant D_2, so that $H(u',v',x',y')\le 16\max(\sqrt{|D_1|},\sqrt{|D_2|})$. Putting $\Delta=\max(|D_1|,|D_2|)$ we find that

$$c(V)c(\nu)\Delta^{1/2-\nu}\le N_2^{\text{special}}(\mathcal{Z},16\sqrt{\Delta})\le c(\mathcal{Z},\epsilon)(16\sqrt{\Delta})^{\epsilon}.$$

Taking some ν,ϵ with $1-2\nu>\epsilon$, these inequalities are untenable once Δ is sufficiently large. So the discriminants D_1,D_2 that occur are of bounded size, and the possible $\tau\in\mathcal{D}$ are finite in number. □

4.2. Remarks.

1. The constant $c(\mathcal{Z},\epsilon)$ is uniform in definable families, hence is uniform for V of given bidegree. The constant $c(V)$ is equal to – and so depends only on – the degree of the field of definition of V. Hence the result is uniform for V of given bidegree and degree of field of definition. The constant $c(\nu)$ is, as already noted, ineffective, but effective under GRHIQ.

2. Suppose V is a curve in $X_0(N)\times X_0(M)$. The modular curves $X_0(N),X_0(M)$ are the compactifications of $Y_0(N)=\Gamma_0(N)\backslash\mathcal{H}$, $Y_0(M)=\Gamma_0(M)\backslash\mathcal{H}$ by the finite index subgroups $\Gamma_0(N),\Gamma_0(M)$ of $\mathrm{SL}_2(\mathbb{Z})$. The fundamental domains $\mathcal{D}_N,\mathcal{D}_M$ each comprise finitely many fundamental domains of $\mathrm{SL}_2(\mathbb{Z})$, and the restriction of the uniformization is definable. Let $Z=\pi^{-1}(V\cap Y_0(N)\times Y_0(M))$. Then $\mathcal{Z}=Z\cap\mathcal{D}_N\times\mathcal{D}_M$ is definable. The proof of 2.3 goes through with minor changes (we take elements of $\Gamma_0(N),\Gamma_0(M)$ rather than $\mathrm{SL}_2(\mathbb{Z})$) to show that $Z^{\text{alg}}=Z^{\text{quasispecial}}$, with the same definition of quasispecial. The proof of 1.1 goes through using H_2^{poly} (defined below) in place of the height H, with the observation made below to relate heights and discriminants on $\mathcal{D}_N,\mathcal{D}_M$. The conclusion of Theorem 1.1 therefore holds for curves V in $X_0(N)\times X_0(M)$ where we take the *special points* of $X_0(N)\times X_0(M)$ to be points (z_1,z_2) with both z_i CM points (i.e. images of quadratic $\tau\in\mathcal{H}$) and *special subvarieties* of positive dimension of $X_0(N)\times X_0(M)$ to be vertical fibres $\{z\}\times X_0(M)$ with z a CM point, horizontal fibres $X_0(N)\times\{z\}$ with z a CM point, images of Z_ϕ for ϕ rational, or $X_0(N)\times X_0(M)$ itself. This is also proved in [21, 2].

In Theorem 1.2 we will look for quadratic imaginary points in the fundamental domain $\mathcal{B}$ for Γ, which consists of finitely many (6) fundamental domains for $\mathrm{SL}_2(\mathbb{Z})$. Rather than relate the height of a quadratic point $\tau\in\mathcal{B}$ to its discriminant as we did above, it is simpler to use a different height.

For an algebraic number α of degree $\le k$ we define

$$H_k^{\text{poly}}(\alpha)=\min\{H(q):q=(q_0,\dots,q_k)\in\mathbb{Q}^{k+1}-\{(0,\dots,0)\},\sum_{j=0}^{k}q_j\alpha^j=0\}.$$

(with $H_k^{\text{poly}}(\alpha)=\infty$ if $[\mathbb{Q}(\alpha):\mathbb{Q}]>k$).

An elementary use of [6, 1.6.5, 1.6.6] shows in [43] that

$$H_k^{\rm poly}(\alpha) \le 2^k H(\alpha)^k$$

for an algebraic number α of degree $\le k$ over $\mathbb{Q}$. We extend $H_k^{\rm poly}$ coordinate-wise to n-tuples:

$$H_k^{\rm poly}(\alpha_1, \ldots, \alpha_n) = \max_i H_k^{\rm poly}(\alpha_i),$$

and for $Z \subset \mathbb{R}^n$ we let

$$N_k^{\rm poly}(Z, T) = \#\{x = (x_1, \ldots, x_n) \in Z : H_k^{\rm poly}(x) \le T\}$$

be the associated counting function. The density result of [43] may be applied using $H_k^{\rm poly}$ rather than H, in fact it is proved using the former.

If $\sigma = \gamma\tau$ for some $\gamma \in {\rm SL}_2(\mathbb{Z})$ then

$$H_k^{\rm poly}(\sigma) \le C(\gamma) H_k^{\rm poly}(\tau)$$

for some constant $C(\gamma)$, as may be seen by computing a polynomial satisfied by σ in terms of one satisfied by τ and the entries of γ^{-1}. If $\sigma \in \mathcal{B}$ then we have $\sigma = \gamma\tau$ for some $\tau \in \mathcal{D}$ and γ from a finite subset of ${\rm SL}_2(\mathbb{Z})$. Let $C(\mathcal{B}) = \max C(\gamma)$ over this set and if $\sigma \in \mathcal{B}$ is quadratic of discriminant D, then $\tau \in \mathcal{D}$ also has discriminant D and we have

$$H_2^{\rm poly}(\sigma) \le C(\mathcal{B}) H_2^{\rm poly}(\tau) \le 4C(\mathcal{B}) H(\tau)^2 \le 4C(\mathcal{B})|D|.$$

If $\sigma + u + iv$ with $u, v \in \mathbb{R}$, and σ is a root of $ax^2 + bx + c$ then $u = -b/2, v = \sqrt{D}/2$ so that

$$H_2^{\rm poly}(u) \le 2|b| \le 2H_2^{\rm poly}(\sigma) \le 8C(\mathcal{B})|D|,$$

$$H_2^{\rm poly}(v) \le 4|D|.$$

4.3. Proof of Theorem 1.2. Let $Z = \pi^{-1}(V)$. We may assume that V contains no special subvarieties of A of positive dimension and must show that V can have only finitely many special points. As in the proof of 1.1 we may assume that V is defined over $\overline{\mathbb{Q}}$.

We set $\mathcal{Z} = Z \cup \mathcal{F}_{\mathcal{B}}$ so that $\mathcal{Z}$ is definable in $\mathbb{R}_{\rm an,exp}$. The pre-images in $\mathcal{Z}$ of special points are among the points $(u, v, x, y) \in \mathcal{Z}$ with all coordinates of degree ≤ 2, and by [43] we have

$$N_2^{\rm poly}(\mathcal{Z} - \mathcal{Z}^{\rm alg}, T) \le c(\mathcal{Z}, \epsilon) T^\epsilon$$

for any positive ϵ.

Under our assumption that V contains no special subvarieties of positive dimension, $\mathcal{Z}^{\rm alg}$ contains no special points. If we denote by $N_2^{\rm poly,special}(Z, T)$ the number of special points $(\tau, z) = (u, v, x, y) \in Z$ with $H_2^{\rm poly}(u, v, x, y) \le T$ then

$$N_2^{\rm poly,special}(\mathcal{Z}, T) \le c(\mathcal{Z}, \epsilon) T^\epsilon$$

for any $\epsilon > 0$.

Suppose that V contains a special point $(\lambda, P) = \pi(\tau, z)$ where $\tau = u + iv, z = x + \tau y$. Then $\lambda(\tau)$ is algebraic with degree at least the degree of $j(\tau)$, so has at least

$$c(\nu)|D|^{1/2-\nu}$$

conjugates λ' (over $\mathbb{Q}$), for any positive ν, where D is the discriminant of τ. The point P is torsion on the elliptic curve E_λ, which is defined over $\mathbb{Q}(\lambda)$ where $[\mathbb{Q}(\lambda) : \mathbb{Q}] = h(D)$. If T is the order of P as a torsion point then, by [49], the degree $d(P)$ over $\mathbb{Q}(\lambda)$ of P satisfies

$$d(P) \ge c(\mu) T^{1-\mu} [\mathbb{Q}(\lambda) : \mathbb{Q}]^{-1} = c(\mu) T^{1-\mu} h(D)^{-1}$$

for any $\mu > 0$, where $c(\mu)$ is effective.

Putting $c(V) = [\mathbb{Q}(V) : \mathbb{Q}]^{-1}$, we find using an (effective) upper estimate

$$c'(\nu)|D|^{1/2+\nu}$$

for $h(D)$ as well as the lower estimate that (λ, P) has at least

$$\max\left(\frac{c(\mu)c'(\nu)T^{1-\mu}}{|D|^{1/2+\nu}}, c(V)c(\nu)|D|^{1/2-\nu}\right)$$

conjugates $(\lambda', P',)$ on V. The pre-images (τ', z') in $\mathcal{Z}$ of these conjugates are distinct. We have $\lambda' = \lambda(\tau')$ with τ' quadratic of discriminant D, and the P' are torsion points of order T. If $\tau' = u +' iv', z' = x' + \tau' y'$ then the height of these points satisfies

$$H_2^{\rm poly}(u, v, x, y) \leq \max(c|D|, T)$$

for any positive μ. Thus

$$\max\left(\frac{c(V)c(\mu)c'(\nu)T^{1-\mu}}{|D|^{1/2+\nu}}, c(V)c(\nu)|D|^{1/2-\nu}\right) \leq N_2^{\rm poly,special}(\mathcal{Z}, \max(c|D|, T)) \leq c(\mathcal{Z}, \epsilon) \max(c|D|, T)^\epsilon$$

for any $\epsilon, \nu, \mu > 0$. Take say $\nu = 1/4, \mu = 1/4, \epsilon = 1/8$, so that we have

$$\max\left(c' \frac{T^{3/4}}{|D|^{3/4}}, c''|D|^{1/4}\right) \leq c''' \max(c|D|, T)^{1/8}. \qquad (*)$$

Suppose that $T \geq |D|^3$. We see that $T^{3/4}/|D|^{3/4} \geq T^{1/2}$, also $c'T^{3/4}/|D|^{3/4} \geq c''|D|^{1/4}$ and $T \geq c|D|$ once $|D|$ is large enough. But then $(*)$ becomes

$$c'T^{1/2} \leq c'''T^{1/8}$$

which is impossible once T is sufficiently large. So $|D|$ is bounded in this case, and then $(*)$ shows that T is bounded as well. But if $T \leq |D|^3$ then $(*)$ implies that $|D|$ is bounded, and so T as well. □

4.4. Proof of Theorem 1.3. Let $Z = \pi^{-1}(V)$. We may assume that V contains no special subvarieties of positive dimension, so that $Z^{\rm alg} = Z^{\rm quasispecial}$ contains no special points, and that V is defined over $\overline{\mathbb{Q}}$.

We take a fundamental domain $\mathcal{D}_N$ for $\Gamma_0(N)$, and a fundamental domain $\mathcal{L}$ for the lattice Λ. The intersection $\mathcal{Z} = Z \cap \mathcal{D}_N \times \mathcal{L}$ is definable, and $\mathcal{Z}^{\rm alg} = Z^{\rm alg} \cap \mathcal{D}_N \times \mathcal{L}$ contains no special points. Therefore, by [43],

$$N_2^{\rm poly,special}(\mathcal{Z}, T) \leq c(\mathcal{Z}, \epsilon)T^\epsilon$$

for any positive ϵ.

Suppose that V contains a special point $(P, Q) = \pi(\tau, z) \in \mathcal{D}_N \times \mathcal{L}$, where $\tau = u + iv, z = x\omega_1 + y\omega_2$. As in the previous proof, putting $C(N) = C(\mathcal{D}_N)$, we have

$$H_2^{\rm poly}(u, v, x, y) \leq \max(c(N)|D|, T)$$

where D is the discriminant of τ, and T is the height (i.e. the order) of the torsion point z. The point P has at least

$$c(\nu)|D|^{1/2-\nu}$$

conjugates over $\mathbb{Q}$.

If Q is a torsion point of order T then, by a result of Masser [30] (the same result used in [45] for abelian varieties), Q has at least $c(E)T^\rho$ conjugates over $\mathbb{Q}$, for some (explicit) $\rho > 0$ (in general ρ depends on the dimension of the abelian variety concerned).

So (P,Q) has at least

$$c(V)\max(c(\nu)|D|^{1/2-\nu}, c(E)T^{\rho})$$

conjuages lying on V whose preimages in $\mathcal{Z}$ have height

$$\leq \max(c(N)|D|, T).$$

Taking $\nu = 1/4$ (say) and $\epsilon < \min(\rho, 1/4)$, the estimates

$$c(V)\max(c(\nu)|D|^{1/2-\nu}, c(E)T^{\rho}) \leq N_2^{\text{poly,special}}(\mathcal{Z}, \max(c(N)|D|, T)) \leq c(\mathcal{Z}, \epsilon)\max(c(N)|D|, T)^{\epsilon}$$

are untenable once $\max(T, |D|)$ is sufficiently large. Therefore these quantities are bounded for special points on V, and the special points (including perhaps those corresponding to the cusps of $X_0(N)$) are finite in number. □

4.5. Final Remark. In [31], it is observed that it suffices to appeal to [41] rather than [44] as the set in question is subanalytic of (real) dimension 2. While here we apply [43] only to sets $\mathcal{Z}$ of (real) dimension 2, the results of [41] (or [42]) are not sufficient. In the first place, the result for algebraic points requires implicitly results for higher dimensional sets. But a result for subanalytic sets in arbitrary dimension would still not suffice as $\mathcal{Z}$ may not be subanalytic: Peterzil-Starchenko establish definability of $\wp(z,\tau)$ on $\mathcal{F}$ in the larger structure $\mathbb{R}_{\text{an,exp}}$ and show that the "exp" is necessary. This demonstrates that the o-minimal setting of [44] is not merely a convenient one, but gives it added generality over its predecessors [41, 42, 10] that can be crucial in applications.

Acknowledgements. My thanks to Umberto Zannier for proposing our collaboration to give a new proof of the Manin-Mumford conjecture, and for telling me of his related work with Masser on Pink's conjecture. I thank Kobi Peterzil for his comments on an earlier version of this paper, in particular for his explanations and references related to 2.4.2 and 2.4.3. I am indebted to the referees for their numerous comments on the text originally submitted, including pointing out some cavalier assertions, and especially for the suggestions (to be pursued elsewhere) to consider general Shimura curves in Theorem 1.1, and to try and recover more of the stronger results of Buium-Poonen in Theorem 1.3. I thank Roger Heath-Brown and the Mathematical Institute, Oxford, for extending hospitality to me as an Academic Visitor, and I am very grateful to the Leverhulme Trust for its support of my work through a Research Fellowship.

References

1. M. Alsina and P. Bayer, *Quaternion orders, quadratic forms, and Shimura curves,* CRM Monograph Series **22**, AMS, 2004.
2. Y. André, Finititude des couples d'invariants modulaire singuliers sur une courbe algébrique plane non modulaire, *J. Reine Angew. Math.* **505** (1998), 203–208.
3. Y. André, Shimura varieties, subvarieties, and CM points, Six lectures at the Franco-Taiwan arithmetic festival, August-September 2001, notes from http://www.math.umd.edu/∼ yu/notes.shtml.
4. K. Barré, Mesure d'approximation simultanée de q et $J(q)$, *J. Number Theory* **66** (1997), 102–128.
5. B. Birch, Heegner points: the beginnings, *Heegner points and Rankin L-series,* H. Darmon and S.-W. Zhang, editors, MSRI Publications **49**, CUP, 2004.
6. E. Bombieri and W. Gubler, *Heights in diophantine geometry,* CUP 2006.
7. E. Bombieri, D. Masser and U. Zannier, Intersecting a curve with algebraic subgroups of multiplicative groups, *IMRN* **20** (1999), 1119–1140.

8. E. Bombieri, D. Masser and U. Zannier, Anomalous subvarieties–structure theorems and applications, *IMRN* **19** (2007), 33 pages.
9. E. Bombieri, D. Masser and U. Zannier, Intersecting a plane with algebraic subgroups of multiplicative groups, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **7** (2008), 51–80.
10. E. Bombieri and J. Pila, The number of integral points on arcs and ovals, *DMJ* **59** (1989), 337–357.
11. A. Borel, S. Chowla, C. S. Herz, K. Iwasawa and J.-P. Serre, *Seminar on complex multiplication,* (1957-8), Lecture notes in mathematics 21, Springer, 1966.
12. F. Breuer, Heights of CM points on complex algebraic curves, *Ramanujan J.* **5** (2001), 311-317.
13. A. Buium and B. Poonen, Independence of points on elliptic curves arising from special points on modular and Shimura curves, I: Global results, preprint dated October 5, 2008 available from http://www-math.mit.edu/∼poonen/
14. A. Buium and B. Poonen, Independence of points on elliptic curves arising from special points on modular and Shimura curves, II: Local results, preprint dated October 20, 2008 available from http://www-math.mit.edu/∼poonen/
15. C. Cornut, Non-trivialité des points de Heegner, *C. R. Acad. Sci. Paris, Se. I* **334** (2002), 1039–1042.
16. H. Darmon, *Rational points on modular elliptic curves,* CBMS Regional Conference Series in Mathematics **101**, AMS, 2004.
17. L. van den Dries, *Tame topology and o-minimal structures,* CUP 1998.
18. L. van den Dries, A. Macintyre, D. Marker, The elementary theory of restricted analytic fields with exponentiation, *Ann. Math. (2)* **140** (1994), 183–205.
19. L. van den Dries and C. Miller, On the real exponential field with restricted analytic functions, *Israel J. Math.* **85** (1994), 19–56.
20. L. van den Dries and C. Miller, Geometric categories and o-minimal structures, *DMJ* **84** (1996), 497–540.
21. S. J. Edixhoven, Special points on the product of two modular curves, *Compositio* **114** (1998), 315–328.
22. S. J. Edixhoven, The André-Oort conjecture for Hilbert modular surfaces, *Moduli of abelian varieties (Texel Island, 1999),* 135–155, Prog. Math. **195**, Birkhauser, Basel, 2001.
23. S. J. Edixhoven, Special points on products of modular curves, *DMJ* **126** (2005), 325–348.
24. S. J. Edixhoven, B. J. J. Moonen, F. Oort, Open problems in algebraic geometry, *Bull. Sci. Math.* **125** (2001), 1–22.
25. S. J. Edixhoven and A. Yafaev, Subvarieties of Shimura varieties, *Ann. Math. (2)* **157** (2003), 621–645.
26. L. Ford, *Automorphic functions,* second edition (1951), reprinted by AMS Chelsea Publishing, 2004.
27. A. Gabrielov, Projections of semi-analytic sets, *Funct. Anal. Appl.* **2** (1968), 282–291.
28. D. Goldfeld, Gauss' class number problem for imaginary quadratic fields, *Bull. Amer. Math. Soc.* **13** (1985), 23–37.
29. A. G. Khovanskii, *Fewnomials,* Translations of Math. Monographs **88**, AMS, Providence, 1991.
30. D. Masser, Small values of the quadratic part of the Néron-Tate height on an abelian variety, *Compositio* **53** (1984), 153–170.
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.
32. J. S. Milne, *Elliptic curves,* BookSurge Publishers, 2006.
33. B. J. J. Moonen, Linearity properties of Shimura varieties, I, *J. Alg. Geom.* **7** (1998), 539–567.
34. J. Nekovar and N. Schappacher, On the asymptotic behaviour of Heegner points, *Turkish J. Math.* **23** (1999), 549–556.

35. Yu. V. Nesterenko, Modular functions and transcendence questions, *Mat. Sb.* **187** (1996), 65–96 (Russian); translation in *Sb. Math.* **187** (1996), 1319–1348.
36. R. Noot, Correspondances de Hecke, Action de Galois et la conjecture d'André-Oort [d'après Edixhoven et Yafaev], *Seminaire Bourbaki*, 2004-05, No. 942.
37. Y. Peterzil, email of 30 November 2008.
38. Y. Peterzil and S. Starchenko, Uniform definability of the Weierstrass $\wp$ functions and generalized tori of dimension one, *Selecta Math. N.S.* **10** (2004), 525–550.
39. Y. Peterzil and S. Starchenko, Complex analytic geometry and analytic-geometric categories, *J. Reine Angew. Math.* **626** (2009), 39–74.
40. G. Philibert, Un lemme de zéros modulaire, *J. Number Theory* **66** (1997), 306–313.
41. J. Pila, Integer points on the dilation of a subanalytic surface, *QJM* **55** (2004), 207–223.
42. J. Pila, Rational points on a subanalytic surface, *AIF* **55** (2005), 1501–1516.
43. J. Pila, On the algebraic points of a definable set, submitted.
44. J. Pila and A. J. Wilkie, The rational points of a definable set, *DMJ* **133** (2006), 591–616.
45. J. Pila and U. Zannier, Rational points in periodic analytic sets and the Manin-Mumford conjecture, *Rend. Lincei Mat. Appl.* **19** (2008), 149–162.
46. R. Pink, A combination of the conjectures of Mordell-Lang and André-Oort, *Geometric methods in algebra and number theory,* F. Bogomolov, Y. Tschinkel, editors, pp 251–282, Prog. in Math. **253**, Birkhauser (2005).
47. R. Pink, A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang, manuscript dated 17 April 2005 available from http://www.math.ethz.ch/∼pink/.
48. J.-P. Rolin, P.Speissegger and A.J. Wilkie, Quasianalytic Denjoy-Carleman classes and o-minimality, *J. Amer. M. Soc.* **16** (2003), 751–777.
49. A. Silverberg, Torsion points on abelian varieties of CM-type, *Compositio* **68** (1988), 241–249.
50. 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.
51. A. J. Wilkie, A theorem of the complement and some new o-minimal structures, *Selecta Math. N.S.* **5** (1999), 397–421.
52. A. Yafaev, Special points on products of two Shimura curves, *Manuscripta Math.* **104** (2001), 163–171.
53. A. Yafaev, A conjecture of Yves André's, *DMJ* **132** (2006), 393–407.
54. A. Yafaev, The André-Oort conjecture – a survey, *L-functions and Galois representations,* D. Burns, K. Buzzard, editors, pp 381–406, London Math. Soc. Lecture Notes **320**, CUP 2007.
55. B. Zilber, Exponential sums equations and the Schanuel conjecture, *J. London Math. Soc. (2)* **65** (2002), 27–44.

School of Mathematics
University of Bristol
Bristol, BS8 1TW
UK

j.pila@bristol.ac.uk