The André-Oort conjecture

Edixhoven, B.; Taelman, L.

Published in:
Nieuw Archief voor Wiskunde

Citation for published version (APA):
The André–Oort conjecture is a problem in algebraic geometry from around 1990, with arithmetic, analytic and differential geometric aspects. Klingler, Ullmo and Yafaev, as well as Pila and Tsimerman have now shown that the Generalized Riemann Hypothesis implies the André–Oort conjecture. Both proofs appeared in the Annals of Mathematics in 2014. In this article Bas Edixhoven and Lenny Taelman describe the conjecture and these recent solutions.

The story of the André–Oort conjecture starts in 1988, with a question, posed by Yves André [1, X.4.3] at the end of his book on solutions of differential equations coming from algebraic varieties defined over \( \mathbb{Q} \).

Elliptic integrals and complex multiplication
The simplest such differential equation is the equation

\[
\lambda(\lambda - 1)\eta''(\lambda) + (2\lambda - 1)\eta'(\lambda) + \frac{1}{2}\eta(\lambda) = 0,
\]

which was already studied by Gauss [9]. It arises from the Legendre family of elliptic curves:

\[
y^2 = x(x - 1)(x - \lambda), \quad \lambda \in \mathbb{C} \setminus \{0, 1\}.
\]

For each \( \lambda \), the set of solutions \( (x, y) \in \mathbb{C}^2 \) is a Riemann surface that can be compactified by adding one point. The compactification \( E \) is an elliptic curve. It is homeomorphic to a torus \( S^1 \times S^1 \). If \( \lambda \) is not in \( (-\infty, 1) \), then the solutions with \( x \) in the real segments \([0, 1]\) and \([1, \lambda]\) form circles \( C_1 \) and \( C_2 \) on \( E \). These two circles intersect transversally, see Figure 1, in a unique point, and therefore generate the fundamental group of \( E \). Integrating the algebraic differential form \( \omega = y^{-1}dx \) along the two circles gives the periods (defined up to sign):

\[
\eta_1 = \int_{C_1} \omega = 2 \int_0^1 \frac{dx}{\sqrt{x(x - 1)(x - \lambda)}},
\]

and

\[
\eta_2 = \int_{C_2} \omega = 2 \int_1^\lambda \frac{dx}{\sqrt{x(x - 1)(x - \lambda)}}.
\]

For varying \( \lambda \), these form a basis of the complex vector space of solutions of (1). The integral linear combinations of \( \eta_1 \) and \( \eta_2 \) form a lattice

\[
\Lambda = \{ k_1\eta_1 + k_2\eta_2 : k_1, k_2 \in \mathbb{Z} \} \subset \mathbb{C}.
\]

The special \( \lambda \) form a countable subset of \( \mathbb{C} \).

Each morphism of elliptic curves \( f \colon E_1 \to E_2 \) corresponds to a homothety \( z \mapsto \alpha z \) such that \( \alpha \Lambda_1 \subset \Lambda_2 \).

The complex number \( \lambda \) is called special if the lattice \( \Lambda \) has complex multiplications, meaning that there are non-real complex numbers \( \alpha \) such that \( \alpha \Lambda \subset \Lambda \) (such an \( \alpha \) defines an endomorphism of \( \mathbb{C} / \Lambda \)). For example, \( \lambda = 2 \) is special as the corresponding lattice,

\[
\Lambda = \mathbb{Z}\{2.622\ldots + 2i\cdot2.622\ldots\},
\]

has multiplication by \( i \). This gives the map

\[
E \to E, \quad (x, y) \mapsto (2 - x, iy).
\]

The special \( \lambda \) form a countable subset of \( \mathbb{C} \).

Figure 1. The intersection of \( C_1 \) and \( C_2 \) near the point \((1,0)\), projected to the complex \( y \)-coordinate (for \( \lambda = i \)).
Manin–Mumford for $C^* \times C^*$
Let $C \subset C^* \times C^*$ be the set of zeros of an irreducible complex polynomial $f$ in two variables. Assume that $C$ contains infinitely many torsion points (pairs $(x, y)$ in $C^* \times C^*$ with both $x$ and $y$ a root of unity). The Manin–Mumford conjecture for $C^* \times C^*$ predicts that $f = aX^nY^m - b$ with $n$ and $m$ coprime, and $b/a$ a root of unity. Equivalently, it predicts that $C$ is the image of a complex line

$$\{(x, y) \in C^2 : ax + by + y = 0 \} \ (\alpha, \beta, y \in \mathbb{Q})$$

under the exponential map

$$C \times C \to C^* \times C^*, \ (x, y) \to (e^{2\pi ix}, e^{2\pi i y}).$$

In fact the Manin–Mumford conjecture for $C^* \times C^*$ is not hard to prove, and would be suitable for the problem section in this journal.

Statement of the conjecture
The Andrè–Oort conjecture is the following statement.

Conjecture. Let $A$ be a Shimura variety, and $Z \subset A$ an irreducible algebraic subvariety. Assume that $Z$ contains a subset $\Sigma$ of special points that is not contained in a strict subvariety of $Z$. Then $Z$ is a Shimura subvariety.

We will say more about Shimura varieties and special points below. This conjecture was formulated (as a question) by Andrè for $Z$ of dimension 1. Independently, this was also formulated by Frans Oort, for the Shimura variety $A_g$ (see below). Both Andrè and Oort were inspired by the analogy with the Manin–Mumford conjecture (see box). Oort’s motivation also came from work of Johan de Jong and Rutger Noot [10] on a conjecture of Robert Coleman on curves with complex multiplications.

Since the general theory of Shimura varieties is rather technical, we will restrict ourselves to examples.

Lattices $\Lambda_1$ and $\Lambda_2$ give isomorphic elliptic curves $C/\Lambda_1$ and $C/\Lambda_2$ if and only if they are isomorphic, that is, if there is a complex number $\alpha$ such that $\alpha\Lambda_1 = \Lambda_2$. Every lattice is homothetic to a lattice of the form

$$\Lambda_\tau := \mathbb{Z} \cdot \tau + \mathbb{Z} \cdot \tau$$

for $\tau \in \mathbb{H}^\tau := \mathbb{C} - \mathbb{R}$. The group $GL_2(\mathbb{R})$ acts transitively on $\mathbb{H}^\tau$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau := \frac{a\tau + b}{c\tau + d}$$

and homothety classes of lattices correspond to orbits under the discrete subgroup $GL_2(\mathbb{Z})$. The quotient

$$A_1 := GL_2(\mathbb{Z}) \backslash \mathbb{H}^\tau$$

is an example of a Shimura variety. It is the moduli space of elliptic curves: the points of $A_1$ are in bijection with isomorphism classes of complex elliptic curves. As before, a point $x \in A_1$ is special if its corresponding lattice $\Lambda$ has complex multiplications. These are precisely the images of the points in $\mathbb{H}^\tau$ of the form $a + b\sqrt{-d}$ with $a$, $b$ and $d$ rational.

The only Shimura subvarieties of $A_1$ are the special points and $A_1$ itself, so that the Andrè–Oort conjecture holds for trivial reasons.

The simplest non-trivial case of the conjecture is for the Shimura variety

$$A_1 \times A_1 = (GL_2(\mathbb{Z}) \times GL_2(\mathbb{Z})) \backslash (\mathbb{H}^\tau \times \mathbb{H}^\tau).$$

The special points are the pairs $(x, y)$ with $x$ and $y$ special. There are three types of one-dimensional Shimura subvarieties: $A_1 \times \{y\}$ with $y$ special, $\{x\} \times A_1$ with $x$ special, and the image of

$$\{(\alpha\tau, \beta\tau) : \tau \in \mathbb{H}^\tau \} \subset \mathbb{H}^\tau \times \mathbb{H}^\tau,$$

with $\alpha, \beta \in GL_2(\mathbb{Q})$. The Andrè–Oort conjecture for $A_1 \times A_1$ is equivalent to the statement given in the previous section.

The most interesting case of the Andrè–Oort conjecture is for the moduli space $A_g$ of (principally polarized) complex abelian varieties of dimension $g$:

$$A_g := GSp_{2g}(\mathbb{Q}) \backslash H_g^\tau,$$

where $H_g^\tau$ is the space of symmetric complex $g$ by $g$ matrices whose imaginary part is definite. The group $GSp_{2g}(\mathbb{R})$ of symplectic similitudes acts transitively on $H_g^\tau$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau := (a\tau + b(c\tau + d))^{-1}$$

where now $a, b, c, d$ are real by $g$ matrices. The special points of $A_g$ correspond to abelian varieties $A$ with many endomorphisms ($End(A) \otimes \mathbb{Q}$ is generated over $\mathbb{Q}$ by $End(A)$ by one element).

A general Shimura variety $A$ is of the form $\Gamma \backslash \mathbb{H}^\tau$, where $\Gamma$ is a discrete subgroup of $G(\mathbb{R})$ for some matrix group $G$ over $\mathbb{Q}$, and where $G(\mathbb{R})$ acts transitively on $\mathbb{H}^\tau$. The Shimura subvarieties of $A$ are images of orbits $H(\mathbb{R}) \cdot x$ for certain algebraic subgroups $H$ of $G$ over $\mathbb{Q}$ and certain $x$ in $\mathbb{H}^\tau$. The zero-dimensional Shimura subvarieties are precisely the special points.

The set of special points is dense in $A$.

Shimura and Deligne have shown that each Shimura variety has a natural structure of algebraic variety, defined over a number field $K$. It is the subspace of a complex projective space defined by a finite system of polynomial equalities and inequalities ($a_i < c_i$) with coefficients in $K$. The special points are defined over finite extensions of $K$, and all Galois conjugates of special points are special points.
Strategies and results

In his thesis Ben Moonen gave a thorough treatment of Shimura subvarieties, including two characterisations. Using one of these he proved the André–Oort conjecture for subvarieties of \( A_g \) under an additional hypothesis on the set \( \Sigma \) \{3, 3.7 and 4.5\}.

As remarked above, André proved the conjecture for \( A_1 \times A_1 \) in [2]. Let \( A = \Gamma \backslash X \) be a Shimura variety, and \( Z \subset A \) an irreducible algebraic subvariety. Assume that \( Z \) contains an infinite subset \( \Sigma \) of special points that is not contained in a strict subvariety of \( Z \). A general strategy for proving the conjecture is to first show that for almost all \( z \in \Sigma \) there is a positive-dimensional Shimura subvariety \( Y_z \) with \( z \in Y_z \subset Z \), and to deduce from this that \( Z \) itself is a Shimura subvariety. The first step is the hardest.

One method, introduced in [4], for producing such \( Y_z \) is to exploit the Galois action on the set of special points in \( A \) combined with the action of \( G(\mathbb{Q}) \) on \( X \). The idea is to intersect \( Z \) with a \( Z' \) obtained from the action by a carefully chosen \( g \) in \( G(\mathbb{Q}) \), such that \( Z \cap Z' \) contains \( \text{Gal}(k) \cdot Z \) and such that the Galois orbit has so many elements that \( Z \cap Z' \) cannot be finite. Then \( Y_z \) is obtained as an irreducible component of a repeated such intersection. This method works if one has sufficiently good lower bounds for the sizes of the Galois orbits \( \text{Gal}(k) \cdot Z \), and sufficient control on the complexity of the \( g \) that can be used. Lower bounds for Galois orbits depend on lower bounds for class numbers of number fields. For the choice of \( g \) one needs sufficiently many small primes in number fields. Both are hard problems in analytic number theory. The best known bounds depend on the Generalised Riemann Hypothesis (GRH) for number fields.

Using this strategy, the André–Oort conjecture was proved, under GRH, for \( A_1 \times A_1 \) in [4], for Hilbert modular surfaces in [6], for curves in general Shimura varieties (André’s question) by Andrei Yafaev in [21] (building on [7, 20]), for powers of \( A_1 \) in [8]. Finally, Bruno Klingler, Emmanuel Ullmo and Yafaev treated the general case (under GRH) in [12, 19]. To make the strategy work in this case is a real tour de force, which took the authors (and the referees) quite a few years (the first versions are from 2006).

Another strategy was introduced more recently by Jonathan Pila in [16], where he proved the conjecture for powers \( A = A_1^n \). The main idea is to work in \( X = (\mathbb{H}^n)^m \) instead of in the quotient \( A = \Gamma \backslash X \). Of course, everything that takes place in \( A \) can be seen in \( X \), but, because the quotient map is not algebraic, the inverse images of algebraic subvarieties of \( A \) are then genuine complex analytic subvarieties of \( X \). On the other hand, \( X \) is an open subset of \( \mathbb{C}^n = \mathbb{R}^{2n} \) in which one again has the notion of algebraic subsets (defined by polynomial equations) and even the notion of algebraic subsets defined over \( \mathbb{Q} \). This is relevant to the problem: the inverse images of special subvarieties are of that kind. Pila imported the tool of O-minimal structures [3] to deal with this mixed algebraic and analytic context. Here, Pila could apply (a generalisation of) his result with Alex Wilkie [15] (see box). This result is used to show that large Galois orbits of special points \( z \) in \( Z \subset A \) give rise to positive dimensional special subvarieties \( Y_z \subset Z \), as in the previous strategy, but now \textit{without having} to take intersections. An important intermediate result is the Ax–Lindeman theorem, which is also proved using [15]; it says that maximal irreducible algebraic subsets of the inverse image of \( Z \) are already very close to being the inverse image of a special subvariety.

Pila and Jacob Tsimerman generalised Pila’s strategy to \( A_g \) in [17]. They proved the conjecture in that case, under GRH. In the case of \( A_g \) for \( g \leq 6 \) (and products of those) they give an unconditional proof. In this case GRH is not needed, as there are sufficiently strong unconditional lower bounds for Galois orbits.

Epilogue

We have seen that the André–Oort conjecture has been proved, under GRH, but many questions remain open.

At this moment, Klingler, Ullmo, Yafaev and Chris Daw are making Pila’s strategy work for general Shimura varieties [31]. It is not inconceivable that number theorists can make the proof unconditional, by providing sufficient lower bounds for Galois orbits (see [5, Problem 14] for what is needed). Ziyang Gao has announced a proof, under GRH, of the conjecture generalised to mixed Shimura varieties [22].

Each Shimura variety \( A \) has a natural probability measure. One expects that for a sequence \( z_n \) of special points in \( A \) such that no subsequence is contained in a strict Shimura subvariety, the Galois orbits of the \( z_n \) are equidistributed. This would imply André–Oort immediately, but this is known only in very special cases.

In the wider context of ‘unlikely intersections’, Richard Pink has formulated [18] a conjecture on subvarieties of mixed Shimura varieties, simultaneously generalising the André–Oort, Manin–Mumford, Mordell–Lang and Zilber conjectures. Pink’s conjecture remains wide open.

\[ \text{The André–Oort conjecture} \]

A subset of \( \mathbb{R}^n \) is called definable in \( \mathbb{R}_{\text{an,exp}} \) if it can be defined using finitely many formulas involving the logical symbols \( \exists, \forall, \wedge, \vee, +, \cdot, \text{exp} \), and functions \( [0, 1]^m \rightarrow \mathbb{R} \) that can be extended to a real analytic function on an open neighbourhood of \( [0, 1]^m \). For example, semi-algebraic sets (defined by polynomial inequalities) are definable. The Pila–Wilkie theorem roughly states that if a definable \( X \) contains many points with rational coordinates, then these must accumulate on semi-algebraic subsets of \( X \). More precisely, for a subset \( X \) of \( \mathbb{R}^n \) we define the counting function

\[ N(X, t) := \#\left\{ \begin{pmatrix} p_1 \ldots p_n \\ q_1 \ldots q_n \end{pmatrix} \in X \mid p_i, q_i \in \mathbb{Z} \cap [-t, t] \right\}. \]

For \( X = \mathbb{R}^n \) we see that \( N(\mathbb{R}^n, t) \sim ct^{2n} \) for some \( c \). Now let \( X \) be a definable subset of \( \mathbb{R}^n \), and let \( X^{\text{alg}} \) be the union of all positive-dimensional semi-algebraic subsets of \( X \).

\[ \text{Theorem (Pila–Wilkie [15]). For every } \epsilon > 0 \text{ there is a } c \text{ such that for all } t \text{ we have } N(X - X^{\text{alg}}, t) < ct^{\epsilon}. \]

As an example, let \( X \subset \mathbb{R}_2 \) be the graph of a function \( f : [0, 1] \rightarrow \mathbb{R} \). If \( f \) is a polynomial with rational coefficients, then \( f(x) \) is rational for every rational \( x \) and \( N(X, t) \) will grow polynomially in \( t \). But if we take \( f(x) = \sin(\pi x) \) then the theorem says that this cannot happen, since \( X^{\text{alg}} = \emptyset \). In fact by Niven’s theorem \( N(X, t) \leq 5 \).
References

9. C.F. Gauss, Disquisitiones generales circa seriem infinitam \(1 + \frac{\alpha \beta}{\gamma x + \cdots}\), Comm. Soc. Regia Sci. Göttingen Rec. 2 (1812).