Annals of Mathematics **191** (2020), 635–674
https://doi.org/10.4007/annals.2020.191.2.7

Abelian varieties isogenous to no Jacobian

By David Masser and Umberto Zannier

Abstract

We prove among other things the existence of Hodge generic abelian varieties defined over the algebraic numbers and not isogenous to any Jacobian. Actually, we also show that in various interpretations these abelian varieties make up the majority, and we give certain uniform bounds on the possible degree of the fields of definition. In particular, this yields a new answer (in strong form) to a question of Katz and Oort, compared to previous work of Chai and Oort (2012, conditional on the André-Oort Conjecture) and by Tsimerman (2012 unconditionally); their constructions provided abelian varieties with complex multiplication (so not "generic"). Our methods are completely different, and they also answer a related question posed by Chai and Oort in their paper.

1. Introduction

1.1. *Preamble.* In [8, p. 589] Chai and Oort raise following the question, which they attribute to Katz (who in turn attributes it back to Oort): *Is there an abelian variety defined over the field* $\overline{\mathbf{Q}}$ *of all algebraic numbers, not isogenous to the Jacobian of any (stable) curve?*

It is classically known that the dimension g of such an abelian variety must be at least 4. This is because every abelian variety, even over $\mathbf{C}$, is isogenous to something principally polarized, and if $g = 1, 2, 3$, then the latter is even isomorphic to such a Jacobian. (Here we take the opportunity to emphasize that, as in [8], our isogenies and isomorphisms are not required to respect polarizations.)

But for $g = 4$, it is also classical that the space of all principally polarized abelian varieties has dimension 10, while the space of all Jacobians has dimension 9. As the set of isogenies is countable, this implies (for example by measure theoretic considerations) that there is a principally polarized abelian variety over $\mathbf{C}$ not isogenous to any Jacobian (and even that "almost all" are not), but it gives no further information about the field of definition. Of course

Keywords: abelian varieties, number fields, Jacobians, isogenies
AMS Classification: Primary: 14H40, 14H52, 14K02.

the field can be taken as finitely generated over $\mathbf{Q}$. This fits with Serre's theme in [39, p. 1; see also pp. 2, 3]: *Il s'agit de prouver que "tout" ce qui est réalisable sur un corps de type fini sur* $\mathbf{Q}$ *l'est aussi* (*par spécialisation*) *sur un corps de nombres.* But such a specialization seems here not straightforward, due to the lack of information about the connecting isogeny.

The following remarks may illustrate the main difficulty. Restricting even to $\overline{\mathbf{Q}}$ leaves only countably many abelian varieties, which could conceivably be covered by the infinitely many arising from isogenies acting on Jacobians or even a single one. For example, it is not hard to prove that if A is principally polarized with endomorphism ring $\mathbf{Z}$, then A/G is principally polarized as soon as the finite group G has order a gth power n^g, and that taking n as different primes leads to non-isomorphic quotients; or better, one may use the isogenies that come from elements of the symplectic group $\mathrm{Sp}_{2g}(\mathbf{Q})$, which is dense in $\mathrm{Sp}_{2g}(\mathbf{R})$ — see [37, p. 154] — and which in turn acts transitively on the Siegel space; see [18, p. 25].

Chai and Oort themselves gave an ingenious construction that supplied an affirmative answer for all $g \geq 4$, but was conditional on a then unproved special case of the André-Oort Conjecture about special sets in Shimura varieties.

Shortly afterwards, Tsimerman [43] gave a no less ingenious unconditional proof. With the help of a powerful equidistribution theorem due to Katz he constructed an infinite sequence of so-called Weyl complex multiplication CM fields with many small split primes and large Galois orbits. He also has to avoid possible Siegel zeroes. Then he used a version of André-Oort due to Klingler and Yafaev [20] but avoiding their use of the Generalized Riemann Hypothesis. He finished as in [8], which by the way uses a different form of equidistribution. (Later on the relevant André-Oort conjecture was proved by Pila and Tsimerman [34] for $g \leq 6$ and by Tsimerman [44] for all g.)

As done in [8] and [43], one can interpret the question (and answer it) for any $g \geq 2$ by considering, instead of (the Torelli locus of all) Jacobians, a general algebraic hypersurface $\mathcal{H}$ in the Siegel moduli space $\mathcal{A}_g$ of all principally polarized abelian varieties of dimension $g \geq 2$.

The abelian varieties thus constructed in [8] and [43] all have complex multiplication CM. In several senses these are well known to be "sparse." (For example, the j-invariant of any CM elliptic curve must be an algebraic integer and there are only thirteen CM elliptic curves defined over the rationals $\mathbf{Q}$.) Perhaps with this in mind, Chai and Oort [8, p. 604] also asked (in our notation): *Given an algebraic* (*special*) *hypersurface* $\mathcal{H}$ *over* $\overline{\mathbf{Q}}$ *in* $\mathcal{A}_g$, *can we find explicitly something in* $\mathcal{A}_g$ *without CM*, *or with CM but not Weyl CM*, *that is not isogenous to anything over* $\overline{\mathbf{Q}}$ *in* $\mathcal{H}$?

1.2. *Our results.* Here we show that there certainly exist such abelian varieties without CM, and even abelian varieties with endomorphism ring $\mathbf{Z}$.

In fact we can escape yet more from CM by using the concept of "Hodge generic." One can consult, for example, [36] in the more general context of Shimura varieties; but for us it is equivalent both to the Mumford-Tate group being GSp_{2g} and to not being contained in any proper special subvariety of $\mathcal{A}_g$. It is thus in some strong sense the opposite of "special" as used in the standard terminology (for example of unlikely intersections).

Back over $\mathbf{C}$, the non-Hodge generic principally polarized abelian varieties are contained in a countable union of proper subvarieties of $\mathcal{A}_g$, while the CM varieties are themselves countable so cannot even fill a semialgebraic curve. An abelian variety that is Hodge generic is not only simple but also has endomorphism ring $\mathbf{Z}$.

Our proofs even yield abelian varieties with the property (apparently stronger, but conjecturally equivalent) of being "Galois generic" (see also [36]).

Our constructions also provide a set of abelian varieties that is dense, not only in the Zariski topology but also in the euclidean topology. The abelian varieties can even be taken as pairwise non-isogenous.

In the discussion above over $\mathbf{C}$ we used the words "almost all." We can even make these more precise over $\overline{\mathbf{Q}}$ by counting the number of exceptions in suitable families.

Our first result shows that we can even take the abelian varieties defined over extensions of $\mathbf{Q}$ whose degree is bounded only as a function of g (so independent of $\mathcal{H}$). Note that the CM abelian varieties in [8], [43] have endomorphisms rings of large discriminant and then one expects their fields of definition to be large (very likely implied by [44] — in particular, the Galois result).

THEOREM 1.1. *Given an algebraic hypersurface $\mathcal{H}$ in $\mathcal{A}_g$ with $g \geq 2$, there is A in $\mathcal{A}_g$, with A defined over an extension of $\mathbf{Q}$ of degree at most 2^{16g^4} and Hodge generic, that is not isogenous to any B in $\mathcal{H}$.*

Our construction will show that even the points of order 16 are defined over the same extension.

The following consequence is clear.

COROLLARY 1.2. *For any $g \geq 4$, there is a principally polarized abelian variety of dimension g, defined over an extension of $\mathbf{Q}$ of degree at most 2^{16g^4} and Hodge generic, that is not isogenous to any Jacobian.*

Now the dimensions counting argument above (in general we have $G = g(g+1)/2$ against $3g-3$) shows that "almost all" abelian varieties over $\mathbf{C}$ will satisfy the requirements. As mentioned, such a statement over $\overline{\mathbf{Q}}$ does not come out of the proof in [43], because CM already holds for "almost no" A.

Our second result shows that almost all A over $\overline{\mathbf{Q}}$ will do, even over bounded extensions as above. To do a precise counting in $\mathcal{A}_g$ it is convenient to consider rational maps (not necessarily morphisms) down to affine $\mathbf{A}^G$ defined over $\overline{\mathbf{Q}}$, and to control more closely the fields arising it is best to allow maps also from covers $\tilde{\mathcal{A}}$, both defined over $\overline{\mathbf{Q}}$, of $\mathcal{A}_g$. All these will be finite in the sense of being generically finite-to-one as measured by degrees.

THEOREM 1.3. *For $g \geq 2$, let $\tilde{\mathcal{A}}$ be any finite cover of $\mathcal{A}_g$, Ψ a finite map from $\tilde{\mathcal{A}}$ to $\mathbf{A}^G$ and $\mathcal{H}$ in $\mathcal{A}_g$ an algebraic hypersurface. Let $\gamma < 1/2$. Then there are $C = C(\tilde{\mathcal{A}}, \Psi, \mathcal{H}, \gamma)$ and $D = D(\tilde{\mathcal{A}}, \Psi)$ with the following property. For any integer $N \geq 1$, there are at most $CN^{G-\gamma}$ elements $\mathbf{n} = (n_1, \ldots, n_G)$ in $\mathbf{Z}^G$ with $1 \leq n_1, \ldots, n_G \leq N$ such that the projection of any element of $\Psi^{-1}(\mathbf{n})$ to $\mathcal{A}_g$ is either*

(a) *not defined over an extension of $\mathbf{Q}$ of degree at most D*
or
(b) *isogenous to some B in $\mathcal{H}$.*

Further, there exist $\tilde{\mathcal{A}}, \Psi$ (defined over $\mathbf{Q}$) such that $D(\tilde{\mathcal{A}}, \Psi) = 2^{16g^4}$.

As there are N^G different $\mathbf{n}$ altogether, indeed we get almost all of them, which we will see can also be taken as Hodge generic.

It is relatively easy to show that these $\Psi^{-1}(\mathbf{n})$ represent at least $C_0^{-1}N^{G-\epsilon}$ different isogeny classes for any $\epsilon > 0$, where $C_0 = C_0(\Psi, \epsilon) > 0$ (see Section 5.3).

By comparison it is expected that the number that have CM is at most $C_1(\Psi)$ independently of N (see Section 5.4).

In fact if g is odd or $g = 2, 6$, then we can take any $\gamma < 1$.

It will turn out that

$$D(\tilde{\mathcal{A}}, \Psi) = [\tilde{F} : \mathbf{Q}][F_\Psi : \mathbf{Q}]D_\Psi, \tag{1}$$

where $\tilde{F}$ is the field of definition of the covering map from $\tilde{\mathcal{A}}$ to $\mathcal{A}_g$, F_Ψ is the field of definition of Ψ, and D_Ψ is the degree of Ψ. And for the particular $\tilde{\mathcal{A}}, \Psi$ mentioned in Theorem 1.3, the 2^{16g^4} can be halved, arising from $\tilde{F} = F_\Psi = \mathbf{Q}$ and $D_\Psi \leq 2^{16g^4-1}$.

We have the following consequence.

COROLLARY 1.4. *For any $g \geq 4$, there is a set of principally polarized abelian varieties of dimension g, dense in the euclidean topology, with each defined over an extension of $\mathbf{Q}$ of degree at most 2^{16g^4} and not isogenous to any of the others or to any Jacobian.*

Probably extending our method of proof will give some sort of ultrametric density. This would also go against CM; in that case, for example when $g = 1$ (suitably interpreted below), we have $|j|_p \leq 1$ for every p.

Finally the field degree 2^{16g^4} can be greatly improved in small dimensions. For $g = 2, 3$, we have unirationality over $\mathbf{Q}$ — that is, dominant rational maps from $\mathbf{A}^G$ to $\mathcal{A}_g$ defined over $\mathbf{Q}$. Certainly $\mathcal{A}_4, \mathcal{A}_5$ are unirational over $\mathbf{C}$ (see [13] for example), so also over some number field; and probably for $\mathcal{A}_4$ the literature implicitly yields $\mathbf{Q}$. Very recently it was shown [11] that $\mathcal{A}_6$ is not unirational. And $\mathcal{A}_g$ is of general type for $g \geq 7$, so far from unirational. Also some diophantine conjectures of Lang-Vojta type would imply that then the set of points over $\mathbf{Q}$ would not be Zariski dense.

THEOREM 1.5. *For $g = 2, 3, 4, 5$, assume there exists a dominant rational map Ξ from $\mathbf{A}^G$ to $\mathcal{A}_g$ defined over $\mathbf{Q}$. Let $\mathcal{H}$ in $\mathcal{A}_g$ be an algebraic hypersurface. Then for any $\gamma < 1/2$, there is $C = C(\Xi, \mathcal{H}, \gamma)$ with the following property. For any integer $N \geq 1$, there are at most $CN^{G-\gamma}$ elements $\mathbf{n} = (n_1, \ldots, n_G)$ in $\mathbf{Z}^G$ with $1 \leq n_1, \ldots, n_G \leq N$ such that $\Xi(\mathbf{n})$ is either*

(a) *not defined over $\mathbf{Q}$*

or

(b) *isogenous to some B in $\mathcal{H}$.*

Again the $\Xi(\mathbf{n})$ excluded from (a) and (b) can be taken as Hodge generic, and again we get at least $C_0^{-1} N^{G-\epsilon}$ isogeny classes (see also Section 5.3) and at most C_1 abelian varieties with CM (see also Section 5.4); and for $g = 2, 3, 5$, we can take any $\gamma < 1$.

Here are some examples of Theorem 1.5.

When $g = 2$, then $G = 3$ and we may take $\Xi(a, b, c)$ as the Jacobian of

$$y^2 = x^5 + x^3 + ax^2 + bx + c \tag{2}$$

or of

$$y^2 = x(x-1)(x-a)(x-b)(x-c) \tag{3}$$

in so-called "Rosenhain coordinates." An example of $\mathcal{H}$ with geometrical (or even physical) significance is defined by the vanishing of certain standard invariants I_2 or I_4 of binary sextic forms. The case $I_2 = 0$ was recently studied by Dunajski and Penrose [12] in connection with twistor theory. (We thank Igor Dolgachev for this reference.) It comes down to taking the coefficient of x in (2) as $-3/20$, and the vanishing of a quartic polynomial for (3). Thus there are many (a, b, c) in $\mathbf{Z}^3$ for which the Jacobian of (2) is not isogenous to the Jacobian of any

$$y^2 = x^5 + x^3 + Ax^2 - \frac{3}{20}x + C. \tag{4}$$

When $g = 3$, then $G = 6$ and hyperelliptic examples like (2), (3) are inadequate, as their moduli space $\mathcal{M}$ has dimension only 5. A better example is $\Xi(a, b, c, d, e, f)$ as the Jacobian of the non-hyperelliptic

$$xy^3 + y^2 + (ax^2 + bx + c)y + x^3 + dx^2 + ex + f = 0, \tag{5}$$

which can be derived from Weierstrass's Collected Works III (or see H. F. Baker [3, p. 589] after the normalizations of his $c = -1, u_3 = x^3 + dx^2 + ex + f$ and the replacement of y by xy to desingularize). Now one can take $\mathcal{H} = \mathcal{M}$ as above. Thus there are many (a, b, c, d, e, f) in $\mathbf{Z}^6$ for which the Jacobian of (5) is not isogenous to the Jacobian of any hyperelliptic curve.

When $g = 4$, we therefore obtain the following (with probably superfluous assumption).

COROLLARY 1.6. *Assume $\mathcal{A}_4$ is unirational over $\mathbf{Q}$. Then there is a principally polarized abelian fourfold, defined over $\mathbf{Q}$ and Hodge generic, that is not isogenous to any Jacobian.*

In fact the problems in [8] and [43] make sense even for $g = 1$, where they amount to showing that there are infinitely many isogeny classes of elliptic curves over $\overline{\mathbf{Q}}$. This turns out to be a much simpler problem, as we shall see in Section 3.1; but it becomes rather more interesting if we consider, instead of the complex hypersurface $\mathcal{H}$, a *real* algebraic curve in complex $\mathcal{A}_1(\mathbf{C})$ identified with $\mathbf{R}^2$. The identification can be conveniently done via the j-invariant in $\mathbf{C}$ and elliptic curves E_j defined by

$$y^2 = 4x^3 - \frac{27j}{j - 1728}x - \frac{27j}{j - 1728} \tag{6}$$

(at least for $j \neq 0, 1728$) and then taking real and imaginary parts of j. Thus we shall prove

THEOREM 1.7. *Given a real algebraic curve $\mathcal{C}$ in $\mathcal{A}_1(\mathbf{C}) = \mathbf{R}^2$, there is $C = C(\mathcal{C})$ with the following property. For any integer $N \geq 2$, there are at most $CN(\log N)^{10}$ integers n_1, n_2 with $1 \leq n_1, n_2 \leq N$ such that E_j for $j = n_1 + in_2$ either*

(a) *has complex multiplication*
or
(b) *is isogenous to some $E_{\mathbf{c}}$ with $\mathbf{c}$ in $\mathcal{C}$.*

Note that an elliptic curve is Hodge generic if and only if it has no complex multiplication.

Again from the N^2 possible (n_1, n_2) we get at least $C_0^{-1}N^2/(\log N)^4$ different isogeny classes, with $C_0 > 0$ absolute (end of Section 2). And in fact none of these E_j have CM, thanks to the solution of the class number $h = 2$ problem. (This could also be achieved by applying the same arguments to $j = \frac{1}{p} + n_1 + in_2$ for any fixed prime p.)

We could formulate and prove similar assertions for models other than (6); for example, the analogue $y^2 = x^3 + x + a$ of (2) or the Legendre analogue $y^2 = x(x-1)(x-a)$ of (3), or even the Hesse version $y^3 + 3axy + x^3 + 1 = 0$.

Thus, for example, not every elliptic curve over $\overline{\mathbf{Q}}$ is isogenous to a real elliptic curve, or to an elliptic curve with j purely imaginary. (We will see that the second is rather harder to treat than the first.)

1.3. *About our proofs.* These are necessarily quite different from those of [8] and [43]. Here is an outline of our strategy. We observe that our main results Theorems 1.1, 1.3, and 1.5 would become somewhat easier to prove if we could bound in advance the degrees of any possible connecting isogenies. Of course we cannot do this; nevertheless we start by fixing a large positive integer M, and we construct many "candidate" Hodge generic abelian varieties A not connected with anything in $\mathcal{H}$ by an isogeny of degree at most M. This construction uses Serre's version of the Hilbert Irreducibility Theorem via the Frattini subgroup to obtain the Hodge generic property (actually via p-Galois generic).

We next argue by contradiction. If there is an isogeny connecting some A as above with some $\tilde{A}$ in $\mathcal{H}$, then it has degree $\tilde{m} > M$. Now "isogeny estimates" [25] provide an upper bound for $\tilde{m}$. This bound involves among other things some power of the degree $\tilde{D}$ of the field of definition of $\tilde{A}$. If $\tilde{D}$ is too large, we can exploit this by using Galois conjugation, say by σ, and so A is connected with every $\tilde{A}^{\sigma}$. The corresponding connections in the Siegel upper half-space lead to integer points ρ_σ on a certain definable variety W in $\mathbf{R}^{4g^2}$, inviting the use of Pila-Wilkie. Often in previous applications of this result the so-called "algebraic part" was safely empty. Here it is not, and our algebraic part W^{alg} is even the whole of W, so we cannot use Pila-Wilkie as it stands. Now W comes with a projection π to another definable set Z (now in $\mathbf{R}^{g^2+g}$), but unfortunately even Z^{alg} can be non-empty. However using the blocks refinement due to Pila, we can show that the cardinality of those $\pi(\rho_\sigma)$ not in Z^{alg} is of order at most T^ϵ for any $\epsilon > 0$, where T is an upper bound for the entries of certain rational representations. Finally using the Hodge generic property and results on weakly special varieties due to Pila and Tsimerman we can show that all our $\pi(\rho_\sigma)$ indeed do not lie in Z^{alg}.

This approach works even in the special case $g = 1$. But here a special role is played by complex conjugation (and we end up in $\mathbf{R}^8$ not $\mathbf{R}^4$). Here we can avoid the general concepts of Hodge generic (which is equivalent to no CM) and of weakly special (by using the algebraic independence of j evaluated at two suitable algebraic functions). Here we can quite easily show that $T \le 2\tilde{m}^{3/2}$.

Then choosing ϵ sufficiently small gives the required contradiction.

But unfortunately for $g > 1$ it is in principle impossible to bound T in terms of the degree $\tilde{m}$ alone. Here the "endomorphism estimates" [26], which were designed to control the totality of isogenies from A to $\tilde{A}$ (by interpreting them as endomorphisms of the product $A \times \tilde{A}$ and considering certain discriminants), do provide an isogeny with T suitably bounded in terms of certain "lengths" coming from Rosati quadratic forms, and the argument can be concluded as above.

1.4. *Outline*. This paper is arranged as follows. In Section 2 we record some preliminary observations on elliptic curves. These are then used in Section 3 to prove Theorem 1.7; also here we make some supplementary remarks about this result. For example, we sketch a proof that it would become false for real transcendental curves. Then Section 4 contains preliminaries on abelian varieties, following the lines of Section 2 but now rather more technical. We can then establish Theorems 1.1, 1.3, and 1.5 in Section 5; also here we include some extra remarks about transcendental hypersurfaces, products and the Jacobian locus.

In connection with the original Jacobian question, there may be some intuition that non-simple abelian varieties are less likely to be related to Jacobians. And indeed, for example, a generic product of two elliptic curves is not the Jacobian of a smooth irreducible curve of genus 2. But it is the Jacobian of a stable curve, and thus in the closure of the Jacobian locus. We will show that this holds up to genus 4, even for arbitrary products.

We may also note that the classical Legendre construction (see [7, p. 157] for example) shows that there is an isogeny of degree 2 between the product of any two elliptic curves and a Jacobian; and it seems plausible that a similar assertion for three elliptic curves can be deduced from Cassels's construction in [6, p. 202]. We do not know if this can be done for four elliptic curves.

Chai and Oort [8, p. 605] also consider the analogous questions over $\overline{\mathbf{F}_p}$ (which apparently had especially interested Katz too). But see Shankar and Tsimerman [40] for evidence that the situation then changes.

It seems likely that our methods can detect suitable abelian varieties even inside proper algebraic subvarieties $\mathcal{K}$ of $\mathcal{A}_g$, provided $\mathcal{K}$ has a dense set of Galois generic points and is not contained in any isogeny translate of $\mathcal{H}$. Thus, for example, one might be able to prove that almost all the jacobians of (3), even with just $b = 2, c = 3$ say, are not isogenous to the Jacobian of any (4).

We are grateful to Yves André, Bas Edixhoven, Ziyang Gao and Andrei Yafaev for their valuable help with the relationship between weakly special and Hodge generic in Section 5.1; and also to Yuri Zarhin for remarks leading to the considerations about products in Section 5.4. Especially we thank Gal Binyamini for pointing out the need for blocks throughout and Gareth Jones for observing that our original arguments in $\mathbf{R}^4$ for Theorem 1.7 were inadequate.

2. Preliminaries on elliptic curves

Any isogeny between elliptic curves has an integer matrix depending on choices of representatives in the upper half-plane $\mathbf{H}$. When these are taken in the standard fundamental domain $\mathcal{F}$ the entries of the matrix can be estimated rather simply in terms of the degree of the isogeny. We use j simultaneously for the invariant of the elliptic curve and the elliptic modular function.

Lemma 2.1. *Suppose $E, \tilde{E}$ are elliptic curves related by an isogeny of degree m, and let $\tau, \tilde{\tau}$ be in $\mathcal{F}$ with $j(\tau) = j(E), j(\tilde{\tau}) = j(\tilde{E})$. Then*

$$\tilde{\tau} = \frac{a\tau + b}{c\tau + d}$$

for integers a, b, c, d with $ad - bc = m$ and

$$\max\{|a|, |b|, |c|, |d|\} \le 2m^{3/2}.$$

Proof. The relation between $\tau, \tilde{\tau}$ is classical, coming from the existence of κ with $[\mathbf{Z}\tau + \mathbf{Z} : \kappa(\mathbf{Z}\tilde{\tau} + \mathbf{Z})] = m$. We then have $\tilde{y} = my|c\tau + d|^{-2}$ for the imaginary parts $y = \Im\tau, \tilde{y} = \Im\tilde{\tau}$. We may suppose $y \le \tilde{y}$, for if not, then we can switch the curves, and the matrix becomes its adjoint. Then $|c\tau + d|^2 = my/\tilde{y} \le m$. Thus also

$$(cy)^2 \le my/\tilde{y} \le m, \tag{7}$$

and since $y \ge \sqrt{3}/2$, we get $|c| \le (4/3)m^{1/2}$. Also $(cx + d)^2 \le m$ for the real part $x = \Re\tau$ with $|x| \le 1/2$, so $|d| \le (5/3)m^{1/2}$.

Now if $\tilde{y} > 2m/\sqrt{3}$, then (7) gives $c^2 < 1$ so $c = 0$. Then

$$ad = m \tag{8}$$

so $|a| \le m$. Also $\tilde{x} = (ax + b)/d$ for $\tilde{x} = \Re\tilde{\tau}$, so

$$|b| = |d\tilde{x} - ax| \le \frac{1}{2}|d| + \frac{1}{2}|a| < \frac{5}{3}m,$$

and we are done in this case.

If $\tilde{y} \le 2m/\sqrt{3}$, then

$$|a\tau + b|^2 = |c\tau + d|^2|\tilde{\tau}|^2 \le \frac{my}{\tilde{y}}\left(\frac{1}{4} + \tilde{y}^2\right), \tag{9}$$

which is at most

$$\frac{m}{4} + my\tilde{y} \le \frac{m}{4} + \frac{2m^2y}{\sqrt{3}} \le \frac{5m^2y}{2\sqrt{3}}.$$

So as above, $|a| \le \sqrt{5/3}m$.

Finally, for the troublesome b (compare the b_{ij} in Lemma 4.1 below), we go back to (9) for

$$(ax + b)^2 \le m\left(\frac{1}{4} + \tilde{y}^2\right) \le \frac{19}{12}m^3,$$

so

$$|b| \le \sqrt{\frac{19}{12}}m^{3/2} + \frac{1}{2}\sqrt{\frac{5}{3}}m < 2m^{3/2}, \tag{10}$$

and we are done in this case too. □

We note that a result with exponent 10 was obtained as Lemma 5.2 of Habegger and Pila [17, p. 19]. It might be interesting to find the best possible exponent. In view of $ad - bc = m$ this might be supposed to be $1/2$, but there is a counterexample $\tau = i, \tilde{\tau} = ai$ with $a = m$ and exponent 1. (Since we wrote that, Orr has informed us that his work [30] does indeed give exponent 1; and we were then able to get the sharp upper bound m itself.) It is true that in Lemma 4.1 of [24, p. 10] the bound $Cm^{1/2}$ was obtained, but that was with C depending on certain heights of $E, \tilde{E}$ (when they are defined over $\overline{\mathbf{Q}}$). In fact the extra heights would make no trouble for us, as we shall soon see in the cases $g > 1$, and even the exponent of m is unimportant; still it is nice to see a result without heights. But already for $g = 2$ there can be no upper bound involving only the degree of the isogeny, even for endomorphisms of E^2. In fact for $g > 1$, we shall use not only the degree, but also a "length" coming from a Rosati quadratic form.

But heights are seemingly unavoidable in the following "isogeny estimate," and we shall use the absolute logarithmic height h.

LEMMA 2.2. *Suppose $E, \tilde{E}$ are isogenous elliptic curves defined over a number field of degree at most $D \geq 2$. Then there is an isogeny between them of degree at most*

$$cD^2(\log D)^2(1 + h(j(E)))^2$$

for c absolute.

Proof. Without specifying dependence on D but with $(1 + h(j(E)))^4$ this was proved in [24]. Gaudron and Rémond [15, p. 347] obtained the above result even with explicit c. □

From this result we can easily see that the E_j for $j = n_1 + in_2$ in Theorem 1.7 represent at least $C_0^{-1}N^2/(\log N)^4$ different isogeny classes. Namely, if some E_j is isogenous to some other fixed E also defined over $\mathbf{Q}(i)$, then there is a connecting isogeny of degree $m \leq C_1(\log N)^2$ with C_1 absolute. Thus E_j is isomorphic to E/G with G of cardinality at most m. We can identify G with a subgroup of $(\mathbf{Q}/\mathbf{Z})^2$ and, for example, Lemma 6.1 of [26, p. 469] shows that there are at most m^2 of these. The assertion follows with $C_0 = C_1^2$.

3. Proof of Theorem 1.7

3.1. *Preamble.* The relatively simple situation with elliptic curves can well be illustrated by a discussion about isogeny classes in general.

It is rather clear that not all complex elliptic curves are isogenous, for instance because there are only denumerably many of them isogenous to a given one (whereas an elliptic curve up to complex isomorphism is classified by its j-invariant, which can be any complex number).

This argument fails on replacing $\mathbf{C}$ by a denumerable subfield, like $\overline{\mathbf{Q}}$, say; still, the fact that $\overline{\mathbf{Q}}$ is dense in $\mathbf{C}$ seems to suggest our intuition that the same assertion holds. (See, however, the remarks at the end of this section.) In fact we also know that not all elliptic curves defined over $\overline{\mathbf{Q}}$ are isogenous, and there are several ways to prove this. Here are five possibilities:

(i) A curve with complex multiplication CM, e.g. $y^2 = x^3 - x$, cannot be isogenous to one without CM, e.g. $y^2 = x^3 - x + 1$ (which indeed cannot have CM because, e.g. its j-invariant $-2^8 3^3/23$ is not an algebraic integer).

(ii) If two elliptic curves are isogenous, then there is a *cyclic* isogeny between them, so the corresponding invariants j_1, j_2 satisfy some modular equation $\Phi_m(j_1, j_2) = 0$. Now, the modular polynomials Φ_m are known to be in $\mathbf{Z}[x, y]$ and monic with respect to both variables; therefore j_2 must be integral over $\mathbf{Z}[j_1]$, and it is now very easy to pick a lot of algebraic j_1, j_2 for which this does not hold.

(iii) Elliptic curves over $\mathbf{Q}$ isogenous over $\mathbf{Q}$, with good reduction at a prime p, are known to have the same number of points over $\mathbf{F}_p$. Now it is easy to pick elliptic curves over $\mathbf{F}_p$ with different numbers of points; lifting them to curves over $\mathbf{Q}$ we obtain the assertion (and likewise with finite extensions of $\mathbf{F}_p$ in case the isogeny is not over $\mathbf{Q}$).

(iv) In a similar flavor, a well-known theorem of Serre-Tate (the easier case of elliptic curves being sufficient) asserts that the set of primes of good reduction is invariant by isogeny, and this easily leads to a further possibility.

(v) Suppose as above that elliptic curves of invariants j_1, j_2 in $\overline{\mathbf{Q}}$ are isogenous. Taking say $j_1 = 0$ and $j_2 = n$ for $n = 1, 2, \ldots, N$ we get by Lemma 2.2 an isogeny between them of degree $m \leq M \ll (\log N)^2$, with absolute implied constants. Now Φ_m has degree

$$\psi(m) = m \prod_{p|m} \left(1 + \frac{1}{p}\right) \;\leq\; \sum_{d|m} d$$

in each variable, so the equation $\Phi_m(0, n) = 0$ has at most $\psi(m)$ solutions. As

$$\sum_{m \leq M} \psi(m) \;\leq\; \sum_{d \leq M} d \;\leq\; M^2,$$

we get $\ll (\log N)^4$ possible values of n, a contradiction for large enough N.

Regarding these arguments, we note that the first one is somewhat unsatisfactory, because it considers "special" curves (in fact of the sort used in [43]) whereas we would expect two "general" curves not to be isogenous. The fifth one is by far the most demanding; however, as we shall see, it will lead to much more substantial information.

In any case, all of these arguments prove indeed sharper results: first, that $\overline{\mathbf{Q}}$ could be replaced with $\mathbf{Q}$, and also that there are in fact *infinitely many* isogeny classes of elliptic curves over $\overline{\mathbf{Q}}$ (or $\mathbf{Q}$), as indicated just before the statement of Theorem 1.7.

To go one step further, let us modify the original issue by asking whether each isogeny class of elliptic curves over $\overline{\mathbf{Q}}$ is represented by curves whose invariant j lies in some "natural" restricted region $\mathcal{R}$ of $\mathbf{C}$ (a kind of "extended fundamental domain" for isogeny equivalence), as simple and as small as possible.

The above shows that such an $\mathcal{R}$ must be in any case infinite. Note also that, using the correspondence between τ in $\mathbf{H}$ and $j = j(\tau)$ in $\mathbf{C}$, and observing (as in Lemma 2.1) that isogenous curves have corresponding τ that are related by a transformation in $\mathrm{GL}_2(\mathbf{Q})$ (and conversely), it is not difficult to see that if $\mathcal{R}$ contains an open set, then indeed it represents all isogeny classes. So, let us think of an $\mathcal{R}$ that is 1-dimensional. A natural choice for $\mathcal{R}$ then seems a real curve $\mathcal{C}$, supposed to be algebraic, both for the sake of simplicity of description and because then we are sure it shall contain many algebraic points (if defined over $\overline{\mathbf{Q}}$).

We see however from Theorem 1.7 that *an algebraic curve never suffices* (even disregarding the special CM invariants). By contrast, we note that without the algebraicity assumption this fails: all isogeny classes may indeed be represented within a suitable real-analytic curve; see the remarks at the end of this section.

3.2. *Main proof.* We now proceed to prove Theorem 1.7 following the strategy outlined in Section 1. We may clearly assume that $\mathcal{C}$ is absolutely irreducible.

If $\mathcal{C}$ is in a standard sense modular, then we can finish rather quickly using complex conjugation and Lemma 2.2, so that Pila-Wilkie or Pila is not needed.

Otherwise we need Galois conjugation as well and then Pila, on a definable set W projecting down to Z as described in the introduction. But since $\mathcal{C}$ is not modular, a result of Pila implies that Z^{alg} is empty.

Let us now carry out the details of this. We start with some elimination.

LEMMA 3.1. *Given $f \neq 0$ in $\mathbf{C}[y_1, y_2]$ there is $c = c(f)$ such that for any m, there is $G_m \neq 0$ in $\mathbf{C}[x_1, x_2]$, of degree at most $c\psi(m)^2$, with the property that $G_m(\xi_1, \xi_2) = 0$ for any $\xi_1, \xi_2, \eta_1, \eta_2$ in $\mathbf{C}$ with*

$$\Phi_m(\xi_1, \eta_1) = \Phi_m(\xi_2, \eta_2) = f(\eta_1, \eta_2) = 0.$$

Proof. If f is in $\mathbf{C}$, then we can take $G_m = 1$ (vacuously). So we assume f is not in $\mathbf{C}$.

If f is in $\mathbf{C}[y_2]$, then we can take G_m as the resultant of $\Phi_m(x_2, y_2)$ and $f(y_1, y_2)$ with respect to y_2; it is well defined as $\Phi_m(x_2, y_2)$ involves y_2 and it is non-zero because $\Phi_m(x_2, y_2)$ is irreducible and f does not involve x_2. This G_m has the vanishing property that we want. So we assume f is not in $\mathbf{C}[y_2]$.

Now the resultant $R(x_1, y_2)$ of $\Phi_m(x_1, y_1)$ and $f(y_1, y_2)$ with respect to y_1 is defined and is non-zero because f does not involve x_1.

If R is in $\mathbf{C}[x_1]$, then we can take $G_m = R$. If R is not in $\mathbf{C}[x_1]$, then the resultant G_m of $\Phi_m(x_2, y_2)$ and $R(x_1, y_2)$ with respect to y_2 is defined and is non-zero because R does not involve x_2.

The degree bounds are straightforward. □

We use this with a polynomial $f \neq 0$ defining our curve $\mathcal{C}$ in the sense that

$$f(u + iv, u - iv) = 0$$

with $u = \Re j, v = \Im j$ for j in $\mathcal{C}$. We may clearly assume that f is over $\overline{\mathbf{Q}}$ and absolutely irreducible. We will shortly see that G_m above has something to do with the invariants of elliptic curves isogenous to E_j for some j in $\mathcal{C}$.

The next result provides plenty of candidates for the elliptic curve not isogenous to anything coming from $\mathcal{C}$. From now on in this section all constants in $\ll, \gg$ may depend only on f and later $\epsilon > 0$.

LEMMA 3.2. *Given integers $M \geq 2$ and $N \geq 1$, there are only*

$$\ll NM^3 \log M$$

pairs $\mathbf{n} = (n_1, n_2)$ of integers with $1 \leq n_1, n_2 \leq N$ such that $E_{\mathbf{n}}$ for $j = n_1 + in_2$ is isogenous to the complex conjugate of $E_{\mathbf{n}}$, or to any $E_{\mathbf{c}}$ for $\mathbf{c}$ in $\mathcal{C}$, via an isogeny of degree at most M.

Proof. The expression $\Phi_m(n_1 + in_2, n_1 - in_2)$ is not identically zero in n_1, n_2 and so it vanishes for $\ll \psi(m)N$ pairs. Summing over m we get the contribution NM^2.

Also from $\Phi_m(n_1 + in_2, \mathbf{c}) = 0$ it follows that the complex conjugate $\Phi_m(n_1 - in_2, \overline{\mathbf{c}}) = 0$. Thus by Lemma 3.1 we deduce $G_m(n_1 + in_2, n_1 - in_2) = 0$. This too does not vanish identically, and so the number of pairs is $\ll \psi(m)^2 N$. Now

$$\sum_{m \leq M} \psi(m)^2 \;\leq\; \sum_{m \leq M} \sum_{d|m} \sum_{d'|m} dd' \;\leq\; \sum_{d \leq M} \sum_{d' \leq M} \frac{dd'M}{[d, d']}$$

for the lowest common multiple. Converting to the highest common factor $(d, d') = e$, we get

$$M \sum_{d \leq M} \sum_{d' \leq M} (d, d') \;\leq\; M \sum_{e \leq M} e \left(\frac{M}{e}\right)^2 \ll M^3 \log M.$$

So this part contributes $NM^3 \log M$ and we are done. □

We illustrate the use of Lemma 3.2 by proving Theorem 1.7 with $\mathcal{C}$ as the real axis (so that $f(y_1, y_2) = y_1 - y_2$).

Suppose then that some $E_{\mathbf{n}}$ for $j = n_1 + in_2$ is isogenous to some $E_{\mathbf{c}}$ for $\mathbf{c}$ in $\mathcal{C}$. Taking complex conjugates we see that the conjugate of $E_{\mathbf{n}}$ is also isogenous to $E_{\mathbf{c}}$. Thus $E_{\mathbf{n}}$ and its conjugate are isogenous. By Lemma 2.2 there is an isogeny between them of degree at most $M \ll (\log N)^2$ But then by Lemma 3.2 this can happen for $\ll N(\log N)^7$ pairs $\mathbf{n}$.

Notice here that $f = \Phi_1$ already defines a modular curve. A similar argument holds for any $f = \Phi_m$.

In this case complex conjugation shows that the conjugates of $E_{\mathbf{n}}, E_{\mathbf{c}}$ are isogenous. But $E_{\mathbf{c}}$ and its conjugate are connected by an isogeny (of degree m). Thus as above $E_{\mathbf{n}}$ and its conjugate are isogenous.

Up to now all the arguments are effective.

In this connection we may note the following "hybrid" between our counting arguments and those of [8] and [43]. Namely, if we were content with just the existence of elliptic curves over $\overline{\mathbf{Q}}$ not isogenous to any $E_{\mathbf{c}}$, then we could try E_j with CM. It would follow that $E_{\mathbf{c}}$ and its complex conjugate both have CM. By André's Theorem [1] there are at most finitely many $\mathbf{c}$ unless $f = \Phi_m$ (up to constants) as above; and we jump back to j using Lemma 2.2. In view of [21] and [4] this is also effective.

Note, however, that our counting arguments to avoid CM do not work for $\mathcal{C}$ as the imaginary axis, because now $f(y_1, y_2) = y_1 + y_2$ comes from no Φ_m. Indeed for generic $\mathbf{c}$, the curves $E_{\mathbf{c}}$ and its complex conjugate are not isogenous. We will exploit this fact in the following proof of Theorem 1.7 when $f \neq \Phi_m$ up to constants.

We now choose $M = [(\log N)^3] \geq 2$ in Lemma 3.2, giving $\ll N(\log N)^{10}$ exceptional $\mathbf{n}$.

We next show that if N is large enough, then outside these exceptions, $E_{\mathbf{n}}$ for $j = n_1 + in_2$ works for Theorem 1.7.

LEMMA 3.3. *Suppose some $E_{\mathbf{n}}$ as above is isogenous to some $\tilde{E} = E_{\mathbf{c}}$ for $\mathbf{c}$ in $\mathcal{C}$. Then there is an isogeny between them of degree*

$$\tilde{m} \ll \tilde{D}^7, \tag{11}$$

where $\tilde{D} \geq 2$ is an upper bound for the degree of the field of definition $\tilde{K}$ of $\tilde{E}$. Furthermore, $\log N \ll \tilde{D}^2(\log \tilde{D})^2$.

Proof. By Lemma 2.2 we have

$$\tilde{m} \ll \tilde{D}^2(\log \tilde{D})^2(\log N)^2.$$

By construction we have $M < \tilde{m}$, and there follows $\log N \ll \tilde{D}^2(\log \tilde{D})^2$ as claimed and then (11). ∎

We are going to play this (11), informally written $\tilde{m} \ll \tilde{D}^{O(1)}$, off against a Pila-Wilkie, or better Pila, estimate

$$\tilde{D} \ll \tilde{m}^{o(1)}. \tag{12}$$

From that follows $\tilde{D} \ll 1$ and so $N \ll 1$ as we wanted.

So the rest of the argument is devoted to a precise version of (12).

There is also an isogeny of degree $\tilde{m}$ between the complex conjugates of $E_{\mathbf{n}}$ and $E_{\mathbf{c}}$. Take any embedding σ of $\tilde{K}$ in $\mathbf{C}$ fixing i and the coefficients of f. Then there is an isogeny, also of degree $\tilde{m}$, from $E_{\mathbf{n}}$ to $\tilde{E}^{\sigma} = E_{\sigma(\mathbf{c})}$.

Similarly there is an isogeny, also of degree $\tilde{m}$, from the Galois conjugate of the complex conjugate of $E_{\mathbf{n}}$ (which is just the complex conjugate of $E_{\mathbf{n}}$ itself) to the Galois conjugate of the complex conjugate of $E_{\mathbf{c}}$; we call this Galois conjugate $E_{\sigma(\mathbf{c})'}$.

Next choose $\tau_{\mathbf{n}}, \tilde{\tau}_{\sigma}, \tau'_{\mathbf{n}}, \tilde{\tau}'_{\sigma}$ in the fundamental domain $\mathcal{F}$ with

$$j(\tau_{\mathbf{n}}) = n_1 + in_2,\ j(\tilde{\tau}_{\sigma}) = \sigma(\mathbf{c}),\ j(\tau'_{\mathbf{n}}) = n_1 - in_2,\ j(\tilde{\tau}'_{\sigma}) = \sigma(\mathbf{c})'. \tag{13}$$

We get

$$\tilde{\tau}_{\sigma} = \frac{a_{\sigma}\tau_{\mathbf{n}} + b_{\sigma}}{c_{\sigma}\tau_{\mathbf{n}} + d_{\sigma}},\ \tilde{\tau}'_{\sigma} = \frac{a'_{\sigma}\tau'_{\mathbf{n}} + b'_{\sigma}}{c'_{\sigma}\tau'_{\mathbf{n}} + d'_{\sigma}} \tag{14}$$

for a point

$$\rho_{\sigma} = (a_{\sigma}, b_{\sigma}, c_{\sigma}, d_{\sigma}, a'_{\sigma}, b'_{\sigma}, c'_{\sigma}, d'_{\sigma})$$

in $\mathbf{Z}^8$ with $a_{\sigma}d_{\sigma} - b_{\sigma}c_{\sigma} = a'_{\sigma}d'_{\sigma} - b'_{\sigma}c'_{\sigma} = \tilde{m}$, and by Lemma 2.1,

$$\max\{|a_{\sigma}|, |b_{\sigma}|, |c_{\sigma}|, |d_{\sigma}|, |a'_{\sigma}|, |b'_{\sigma}|, |c'_{\sigma}|, |d'_{\sigma}|\} \ \leq\ 2\tilde{m}^{3/2}. \tag{15}$$

Write j^2 for the product map from $\mathbf{H}^2$ to $\mathbf{C}^2$. Now $f = 0$ defines a curve $\mathcal{C}_{\mathbf{C}}$ in $\mathbf{C}^2$; we define

$$Z = \mathcal{F}^2 \ \cap\ (j^2)^{-1}(\mathcal{C}_{\mathbf{C}}).$$

For (τ, τ') in $\mathbf{H}^2$, write also $W_{\tau,\tau'}$ as the set of all

$$(x_a, x_b, x_c, x_d, x'_a, x'_b, x'_c, x'_d) \tag{16}$$

in $\mathbf{R}^8$ with x_c, x_d not both zero and x'_c, x'_d not both zero and

$$\left(\frac{x_a\tau + x_b}{x_c\tau + x_d},\ \frac{x'_a\tau' + x'_b}{x'_c\tau' + x'_d}\right) \tag{17}$$

in Z.

This is a *definable* variety in $\mathbf{R}^8$. (See [35] or [46, Ch. 4] for definitions and for these properties.) In fact $W^{\mathrm{alg}}_{\tau,\tau'} = W_{\tau,\tau'}$ because, for example, a point (16) gives rise to a semi-algebraic curve parametrized by

$$(\gamma x_a, \gamma x_b, \gamma x_c, \gamma x_d, x'_a, x'_b, x'_c, x'_d)$$

for varying real γ — similarly with γ' and the last four coordinates.

These problems could probably be overcome by de-homogenizing. But even worse, the stabilizer group of τ in $\mathrm{GL}_2(\mathbf{R})$ has dimension 2 and its elements $\left(\begin{smallmatrix}\gamma_a & \gamma_b\\ \gamma_c & \gamma_d\end{smallmatrix}\right)$ multiply $\left(\begin{smallmatrix}x_a & x_b\\ x_c & x_d\end{smallmatrix}\right)$ on the right, resulting in a semi-algebraic surface.

Thus [35] cannot usefully be applied.

We can get around this problem using blocks as in Pila's [32]. As we also need uniformity in τ, τ' we regard $W_{\tau,\tau'}$ as a fibre of a set in $\mathbf{R}^8 \times \mathbf{R}^4$ corresponding, say, to the element

$$(\Re\tau, \Im\tau, \Re\tau', \Im\tau') \tag{18}$$

of $\mathbf{R}^4$.

For $T \geq 1$, let $W_{\tau,\tau'}(T)$ be the set of integral points in $W_{\tau,\tau'}$ with coordinates bounded in absolute value by T. Let $\epsilon > 0$. Then Theorem 3.5(2) of [32, p. 158] says that $W_{\tau,\tau'}(T)$ is contained in basic blocks B whose cardinality is $\ll T^\epsilon$, where the implicit constant now may depend on ϵ (and the family of varieties $W_{\tau,\tau'}$) but not on T or, crucially for us, τ and τ'. (This relies on the fact that $\mathcal{F}$, though not compact, is definable. A simple deduction of this fact from basic general results appears in [46, Ch. 4, Notes].)

We can regard Z above (also definable) as in $\mathbf{R}^4$, also as in (18), and there is an obvious map $\pi_{\tau,\tau'}$ from $W_{\tau,\tau'}$ to Z obtained by taking real and imaginary parts of (17). This is semi-algebraic, and so by Definition 3.2(2) of [32, p. 157] the various $\pi_{\tau,\tau'}(B)$ are blocks C in Z.

We now show that all these blocks have dimension zero. Indeed, if some C has positive dimension, then by Proposition 3.4(2) of [32, p. 158] it would lie in Z^{alg}.

But in fact Z^{alg} is empty. This could be seen using Pila's algebraic independence results in [33, Th. 1.6]. Or we may use a version in [46, p. 113]. Namely, if Γ is a real algebraic arc contained in Z, then we are in the situation of Lemma 4.4 of [46] (which in fact needs the absolute irreducibility of $\mathcal{C}$, not just the irreducibility over $\mathbf{Q}$ as used on p. 108). As our $\mathcal{C}$ is not modular (note that $f(y_1, y_2)$ must involve both y_1 and y_2, so vertical and horizontal lines cannot turn up), we see that Γ cannot exist.

Thus the $\pi_{\tau,\tau'}(B)$ indeed have dimension zero and so are points.

We apply this to $(\tau, \tau') = (\tau_{\mathbf{n}}, \tau'_{\mathbf{n}})$ above. Note that via (14) and (15) above, each σ gives rise to a point ρ_σ in $W_{\tau,\tau'}(T)$ with $T \ll \tilde{m}^{3/2}$. Also $\pi_{\tau,\tau'}(\rho_\sigma) = (\tilde{\tau}_\sigma, \tilde{\tau}'_\sigma)$, of which the cardinality as σ varies is $\gg \tilde{D}$ by (13). This ρ_σ lies in some basic block $B = B_\sigma$, so $\pi_{\tau,\tau'}(\rho_\sigma)$ must be the point $\pi_{\tau,\tau'}(B_\sigma)$, of which the cardinality is $\ll T^\epsilon$. It follows that $\tilde{D} \ll \tilde{m}^{3\epsilon/2}$ as in (12) above; fixing $\epsilon < 2/21$ gives, on recalling (11), the required $\tilde{D} \ll 1$ and $N \ll 1$.

Finally note that $E_{\mathbf{n}}$ cannot have complex multiplication, otherwise so would its complex conjugate, and with the same CM field, so they would after all be isogenous, also against Lemma 3.2. This completes the proof of Theorem 1.7.

We could also have used notions equivalent to block: namely, a "maximal semi-algebraic connected component", or the "class of the equivalence relation generated by the relation between two points that they are connected by a semi-algebraic set".

3.3. *Further remarks.* Theorem 1.7 can be extended from the two-parameter family $j = n_1 + in_2$ to certain one-parameter families of j. But, for example, $j = n$ for integers n is ruled out by taking $\mathcal{C}$ as the real axis $\mathbf{R}$ (even though it is not difficult to show that $j = n$ is permissible for the imaginary axis $i\mathbf{R}$). However $j = n+i$ does work for $\mathbf{R}$. So we could consider $j = n+in_0$ for some integer n_0 depending on $\mathcal{C}$. In fact given f of degree at most $d \geq 1$ in each variable, we can show that there is n_0 with $0 \leq n_0 \leq 2d^3+1$ such that $j = n + in_0$ is permissible for the corresponding $\mathcal{C}$. We proceed to sketch the proof.

The obstacle to any specific n_0 comes from the proof of Lemma 3.1 and the possibility that $G_m(x+in_0, x-in_0)$ is identically zero in x. For example, with $\mathcal{C} = \mathbf{R}$ we have $f = y_1 - y_2$ so $G_1(x_1, x_2) = x_1 - x_2$ ruling out $n_0 = 0$.

To deal with this we have to interpret G_m geometrically as defining the projection to $\mathbf{C}^2$ of the curve displayed in Lemma 3.1. Then for generic x, we would have extensions $F_1 = \mathbf{C}(x, \eta_1), F_2 = \mathbf{C}(x, \eta_2)$ defined by

$$\Phi_m(x+in_0, \eta_1) = 0, \;\; \Phi_m(x-in_0, \eta_2) = 0$$

together with $f(\eta_1, \eta_2) = 0$.

Now F_1 is ramified only above $x = -in_0, 1728-in_0, \infty$, and F_2 only above $x = in_0, 1728+in_0, \infty$. There is just one overlap (if $n_0 \neq 0$), and this implies that F_1, F_2 are linearly disjoint over $\mathbf{C}(x)$. So $[\mathbf{C}(x, \eta_1, \eta_2) : \mathbf{C}(x)] = \psi(m)^2$. On the other hand, $f = 0$ shows that this degree is at most $d\psi(m)$. Thus $\psi(m) \leq d$.

And now $n_0 \geq 1$ can be chosen to avoid the above obstacle with $n_0 - 1$ at most the total degree of G_m. By standard resultant estimates this is at most $2d\psi(m)^2 \leq 2d^3$.

Certainly the proof allows us to strengthen the theorem also by finding j in any non-real number field K. (In the proof we used $K = \mathbf{Q}(i)$ for simplicity.) We can even insist on further properties of the E_j in question, having generic Galois group of p-power-torsion for a corresponding elliptic curve; compare Lemma 5.1 below.

Next we sketch an argument proving the existence of a real-analytic curve Z in $\mathbf{C} = \mathbf{R}^2$ such that each elliptic curve over $\overline{\mathbf{Q}}$ is isogenous to one with j in Z; by Theorem 1.7 this curve will necessarily be transcendental. In fact our Z will be bounded.

Our basic tool will be a Newton series

$$F(t) = a_0 + \sum_{n=1}^{\infty} a_n \prod_{m=0}^{n-1} \frac{t - t_m}{t_n - t_m} \tag{19}$$

with different $t_0, t_1, t_2, \ldots$ that we will be able to keep in the real interval $[1, 2]$. They lead to positive $\epsilon_0, \epsilon_1, \epsilon_2, \ldots$ such that $F(t)$ converges to an entire function when all $|a_n| \leq \epsilon_n$. For given real $s_0, s_1, s_2, \ldots$, we will try to solve $F(t_n) = s_n$ $(n = 0, 1, 2, \ldots)$. Then the inequalities amount to

$$|s_0| < \epsilon_0, \; |c_{10}s_0 + s_1| < \epsilon_1, \; |c_{20}s_0 + c_{21}s_1 + s_2| < \epsilon_2, \ldots \tag{20}$$

with certain $c_{10}, c_{20}, c_{21}, \ldots$ depending only on $t_0, t_1, \ldots$.

We start our construction by enumerating as $\tau_0, \tau_1, \tau_2, \ldots$ all τ in $\mathbf{H}$ with $j(\tau)$ algebraic. Then we successively modify τ_n by multiplying by a positive rational (automatically preserving the isogeny class) such that the resulting $y_n = \Im\tau_n$ are all different with, say, $1 \leq y_n \leq 2$. Then we successively modify τ_n by adding a rational (again preserving the isogeny class as well as y_n) such that (20) are satisfied for $t_n = y_n$ and $s_n = x_n$ $(n = 0, 1, 2, \ldots)$. Then $F(y) = x$ defines a real analytic curve in $\mathbf{R}^2$ containing the modified τ_n, and we take Z as part of its image under j; namely, the set of all $j = j(F(y) + iy)$ $(1 \leq y \leq 2)$. As $y \leq 2$, we see that Z is bounded (and, in fact, $|j| \leq 2079 + e^{4\pi}$).

Can the diameter be made arbitrarily small?

Can one prove that any Z as above cannot be "too simple," for example $v = u^\pi$?

4. **Preliminaries on abelian varieties**

Any isogeny between principally polarized abelian varieties of dimension g also has an integer matrix, also depending on choices of representatives, now in the Siegel upper half-space $\mathbf{H}_g$. But even when these are taken in a fundamental domain, it may be in principle impossible to estimate the entries of the matrix in terms of the degree of the isogeny as in Lemma 2.1. This is due to the possibility of non-trivial units in an endomorphism ring, which with their powers are all isogenies of degree 1.

To overcome this problem it is convenient to use the Rosati quadratic form (on the endomorphism ring), which is positive definite. We may informally refer to its square root ℓ as the *length*. Thus, for example, given any L, there are at most finitely many endomorphisms with length at most L.

As in [27] we are going to identify an isogeny between A and $\tilde{A}$ of dimension g with an endomorphism of $A \times \tilde{A}$, so the dimension becomes $2g$. There it was a device allowing the estimation of all isogenies, not just one; see, for example, (29) below.

Generally let $\underline{A}$ be a principally polarized abelian variety of dimension $\underline{g}$, with a matrix $\underline{\tau}$ in $\mathbf{H}_{\underline{g}}$ and $\underline{x} = \Re\underline{\tau}$, $\underline{y} = \Im\underline{\tau}$. (We are underlining everything in this general situation.) An endomorphism v of $\underline{A}$ gives rise to an equation

$$\underline{\kappa}(\underline{\iota}\ \underline{\tau}) = (\underline{\iota}\ \underline{\tau})\underline{\rho} \tag{21}$$

with $\underline{\kappa} = \underline{\kappa}(v)$ in $M_{\underline{g}}(\mathbf{C})$ the complex representation and $\underline{\rho} = \underline{\rho}(v)$ in $M_{2\underline{g}}(\mathbf{Z})$ the rational representation and $\underline{\iota}$ the identity. The Rosati involution (see [27, pp. 643, 644]) associated with the principal polarization gives rise to a positive quadratic form defined by

$$\ell(v)^2 = \mathrm{tr}(\underline{\kappa}\underline{y}\overline{\underline{\kappa}}^t\underline{y}^{-1}) = \mathrm{tr}(\underline{\rho}\underline{\varepsilon}\underline{\rho}^t\underline{\varepsilon}^{-1}) \tag{22}$$

for the traces and transposes (of matrices, not for example with respect to endomorphism rings) and

$$\underline{\varepsilon} = \begin{pmatrix} \underline{o} & -\underline{\iota} \\ \underline{\iota} & \underline{o} \end{pmatrix}$$

for $\underline{o}$ the zero matrix. In particular, $\ell(v) \geq 1$ for all $v \neq 0$.

Here now is our partial substitute for Lemma 2.1. For technical reasons we have to allow a domain larger than the standard fundamental domain. The latter is defined, for example, in [18, p. 194]. If $\underline{\tau} = \underline{x} + i\underline{y}$ is in $\mathbf{H}_{\underline{g}}$, we write $\underline{y}^{(0)}$ for the diagonal matrix with the same diagonal entries $\underline{y}_1, \ldots, \underline{y}_{\underline{g}}$ as $\underline{y}$. We use $r \geq s$ or $r > s$ to indicate that the (symmetric) matrix $r - s$ is positive semi-definite or positive definite; for example, $\underline{y} > \underline{o}$. Then the existence of $\delta > 0$ depending only on $\underline{g}$ (and we may take $\delta \leq 1$) with

$$\delta^{-1}\underline{y}^{(0)} \geq \underline{y} \geq \delta\underline{y}^{(0)},\ \underline{y}^{(0)} \geq \delta\underline{\iota} \tag{23}$$

follows from Corollary 2 of [18, p. 193] together with

$$\frac{\sqrt{3}}{2} \leq \underline{y}_1 \leq \cdots \leq \underline{y}_{\underline{g}}; \tag{24}$$

see [18, p. 192 and Lemma 15, p. 195].

In fact it is specifically (24) that we wish to avoid using. Namely, if $\underline{A}$ is a product $A \times \tilde{A}$ and we form $\underline{\tau} = \left(\begin{smallmatrix} \tau & o \\ o & \tilde{\tau} \end{smallmatrix}\right)$ with $\tau, \tilde{\tau}$ satisfying the analogues of (24), then $\underline{\tau}$ might not, due to the ordering. So in the next result we drop (24).

Finally the entries of $\underline{x}$ are bounded in absolute value by $1/2$; see [18, (S.3), p. 194].

Lemma 4.1. *Given $\underline{g} \geq 1$ and δ with $0 < \delta \leq 1$ there is $C = C(\underline{g}, \delta)$ with the following property. Suppose that (23) holds for $\underline{\tau} = \underline{x} + i\underline{y}$ corresponding to $\underline{A}$ and also that the entries of $\underline{x}$ are bounded above in absolute value by δ^{-1}.*

Then for any endomorphism v of $\underline{A}$ we have for the rational representation

$$\underline{\rho} = \begin{pmatrix} \underline{a} & -\underline{b} \\ -\underline{c} & \underline{d} \end{pmatrix}$$

with entries $a_{ij}, b_{ij}, c_{ij}, d_{ij}$ the estimates

$$|a_{ij}| \leq C\sqrt{\underline{y}_i/\underline{y}_j}\,\ell(v),\ \ |b_{ij}| \leq C\sqrt{\underline{y}_i\underline{y}_j}\,\ell(v),\ \ (i,j = 1,2,\ldots,\underline{g})$$
$$|c_{ij}| \leq C\frac{1}{\sqrt{\underline{y}_i\underline{y}_j}}\ell(v),\ \ |d_{ij}| \leq C\sqrt{\underline{y}_j/\underline{y}_i}\,\ell(v),\ \ (i,j = 1,2,\ldots,\underline{g}).$$

Proof. From (21) we have $\underline{\kappa} = \underline{a} - \underline{\tau}\underline{c} = \underline{\tilde{a}} - i\underline{y}\underline{c}$ for $\underline{\tilde{a}} = \underline{a} - \underline{x}\underline{c}$ and $\underline{x} = \Re(\underline{\tau})$, so also $\overline{\underline{\kappa}}^t = \underline{\tilde{a}}^t + i\underline{c}^t\underline{y}$. Substituting these into the first equality in (22) and ignoring the imaginary parts, we get

$$\ell(v)^2 = \mathrm{tr}(\underline{\tilde{a}}\underline{y}\underline{\tilde{a}}^t\underline{y}^{-1}) + \mathrm{tr}(\underline{y}\underline{c}\underline{y}\underline{c}^t). \tag{25}$$

Now we have $\underline{y} \geq \delta\underline{y}^{(0)}$, and it follows that $\mathrm{tr}(\underline{y}\underline{c}\underline{y}\underline{c}^t) \geq \delta\mathrm{tr}(\underline{y}^{(0)}\underline{c}\underline{y}\underline{c}^t)$ using the well-known fact that $\mathrm{tr}(rs) \geq 0$ when $r \geq \underline{o}, s \geq \underline{o}$. By the same token the first term on the right of (25) is non-negative and also

$$\mathrm{tr}(\underline{y}^{(0)}\underline{c}\underline{y}\underline{c}^t) = \mathrm{tr}(\underline{c}^t\underline{y}^{(0)}\underline{c}\underline{y}) \ \geq\ \delta\mathrm{tr}(\underline{c}^t\underline{y}^{(0)}\underline{c}\underline{y}^{(0)}) = \delta\sum_{i=1}^{\underline{g}}\sum_{j=1}^{\underline{g}} c_{ij}^2\underline{y}_i\underline{y}_j.$$

From this follow the $\underline{c}$-estimates in the present lemma.

Also from $\underline{y} \leq \delta^{-1}\underline{y}^{(0)}$ it follows that $\underline{y}^{-1} \geq \delta(\underline{y}^{(0)})^{-1}$, using another well-known fact that $r \geq s > \underline{o}$ implies $r^{-1} \leq s^{-1}$ (for example by simultaneous diagonalization). So a similar argument gives

$$\mathrm{tr}(\underline{\tilde{a}}\underline{y}\underline{\tilde{a}}^t\underline{y}^{-1}) \ \geq\ \delta^2\mathrm{tr}(\underline{\tilde{a}}\underline{y}^{(0)}\underline{\tilde{a}}^t(\underline{y}^{(0)})^{-1}) = \delta^2\sum_{i=1}^{\underline{g}}\sum_{j=1}^{\underline{g}}\tilde{a}_{ij}^2\frac{\underline{y}_j}{\underline{y}_i}.$$

From this follow the $\underline{a}$-estimates, or rather the analogous estimates for $\underline{\tilde{a}}$. But from $\underline{a} = \underline{\tilde{a}} + \underline{x}\underline{c}$ and the fact that all entries of $\underline{x}$ are bounded in absolute value by δ^{-1} together with $\underline{y}_i \geq \delta$ $(i = 1,2,\ldots,\underline{g})$ and the above $\underline{c}$-estimates, we can easily check that the same indeed holds for $\underline{a}$.

What about $\underline{b}, \underline{d}$? We proceed via $\underline{\kappa} = \underline{a} - \underline{\tau}\underline{c} = \underline{a} - \underline{x}\underline{c} - i\underline{y}\underline{c}$. From the $\underline{a}$-estimates and the $\underline{c}$-estimates, together with the fact that the (i,j) entry of $\underline{y}$ is at most $\sqrt{\underline{y}_i\underline{y}_j}$ in absolute value, we can verify that the (complex) entries of $\underline{\kappa}$ satisfy the same sort of inequalities as those in the $\underline{a}$-estimates. Then we can similarly verify that the entries of $\underline{\kappa}\underline{\tau}$ satisfy the same sort of inequalities as those in the $\underline{b}$-estimates. But $\underline{\kappa}\underline{\tau} = -\underline{b} + \underline{\tau}\underline{d} = -\underline{b} + \underline{x}\underline{d} + i\underline{y}\underline{d}$ and so at once we get these inequalities for $\underline{\tilde{b}} = \underline{b} - \underline{x}\underline{d}$ and $\underline{y}\underline{d}$. Then using $\underline{d} = \underline{y}^{-1}(\underline{y}\underline{d})$ and $\underline{y}^{-1} \leq \delta^{-1}(\underline{y}^{(0)})^{-1}$ we get the required $\underline{d}$-estimates. Finally, using $\underline{b} = \underline{\tilde{b}} + \underline{x}\underline{d}$ we get the $\underline{b}$-estimates, which are the worst (as in (10) above) and most dangerous (in some applications at least). This completes the proof. □

We remark that the exponent of $\ell(v)$ cannot be improved here because when v is an integer n, then $\ell(v) = \sqrt{2g}n$. But by using a "multiplication trick" as in (8), e.g. $|a_{ij}d_{ij}| \ll \ell(v)^2$, one might hope to get bounds like $|a_{ij}| \ll \ell(v)^\nu$ for small $\nu = \nu(\underline{g})$ as in the elliptic case (when $\ell(v)^2$ is essentially the degree of v). We do not know if such bounds actually exist. The problem is of course vanishing entries. From (43) below we will see that there will be some, and probably even more. This is because in the applications $\mathrm{End}(A \times \tilde{A}) = M_2(\mathbf{Z})$ is of rank only 4 in the group $M_{4g}(\mathbf{Z})$.

Soon we will see how to deal with the possibly extraneous $\underline{y}_i$ in Lemma 4.1.

To get from a product back to its factors, we need the following observation. For $\underline{A}$ as above, we write $\mathcal{D}(\underline{A}) \geq 1$ for the discriminant of $\mathrm{End}(\underline{A})$ with respect to the Rosati form; see [27, p. 642]. Thus if $\mathrm{End}(\underline{A})$ has rank r over $\mathbf{Z}$ and we take representations $\underline{\kappa}_1, \ldots, \underline{\kappa}_r$ and $\underline{\rho}_1, \ldots, \underline{\rho}_r$ of any basis elements, as in (21), then as in (22), we have

$$\mathcal{D}(\underline{A}) = \det_{i,j} \mathrm{tr}(\underline{\kappa}_i \underline{y} \overline{\underline{\kappa}_j}^t \underline{y}^{-1}) = \det_{i,j} \mathrm{tr}(\underline{\rho}_i \underline{\varepsilon} \underline{\rho}_j^t \underline{\varepsilon}^{-1}).$$

We remark that $\mathcal{D}(\underline{A})$ is not the same discriminant as that used by Tsimerman [42]: first he considers only the centre of $\mathrm{End}(\underline{A})$ (and anyway his $\underline{A}$ has complex multiplication) and second he uses no Rosati form, just the standard field trace norm.

LEMMA 4.2. *There is $c = c(g)$ with the following property. Let $A, \tilde{A}$ be principally polarized abelian varieties of dimension g that are isogenous. Then there are isogenies f from A to $\tilde{A}$ and $\tilde{f}$ from $\tilde{A}$ to A such that the endomorphism v of $A \times \tilde{A}$ defined by*

$$v(\alpha, \tilde{\alpha}) = (\tilde{f}(\tilde{\alpha}), f(\alpha)) \tag{26}$$

has $\ell(v) \leq c\mathcal{D}(A \times \tilde{A})^{1/2}$.

Proof. This is implicit in the arguments of [27, §6]. Define $v_0, \tilde{v}_0$ in $\mathcal{E} = \mathrm{End}(A \times \tilde{A})$ by

$$v_0(\alpha, \tilde{\alpha}) = (\alpha, 0),\ \ \tilde{v}_0(\alpha, \tilde{\alpha}) = (0, \tilde{\alpha}),$$

and for any v in $\mathcal{E}$, write

$$v^\# = v_0 v \tilde{v}_0 + \tilde{v}_0 v v_0$$

also in $\mathcal{E}$. For each v in $\mathcal{E}$, there exist u in $\mathrm{End}A$, $\tilde{u}$ in $\mathrm{End}\tilde{A}$ and homomorphisms f from A to $\tilde{A}$, $\tilde{f}$ from $\tilde{A}$ to A with

$$v(\alpha, 0) = (u(\alpha), f(\alpha)),\ \ v(0, \tilde{\alpha}) = (\tilde{f}(\tilde{\alpha}), \tilde{u}(\tilde{\alpha})), \tag{27}$$

and we check

$$v^\#(\alpha, \tilde{\alpha}) = (\tilde{f}(\tilde{\alpha}), f(\alpha)). \tag{28}$$

This reflects the matrix identity (6.2) of [27, p. 652].

If r is the rank of $\mathcal{E}$ over $\mathbf{Z}$ (with $r \leq (4g)^2$ from the rational representation), then the Geometry of Numbers gives $v_1, \ldots, v_r$ in $\mathcal{E}$, linearly independent over $\mathbf{Z}$, with

$$\ell(v_1) \cdots \ell(v_r) \ll \mathcal{D}^{1/2} \tag{29}$$

with $\mathcal{D} = \mathcal{D}(A \times \tilde{A})$ and implied constants depending only on g.

It is well known that the degree

$$\deg(m_1 v_1^{\#} + \cdots + m_r v_r^{\#}) = P(m_1, \ldots, m_r) \tag{30}$$

for a polynomial P, homogeneous of degree $4g$.

Here P is not identically zero because there are isogenies $f, \tilde{f}$ as above, and then we can define v by (50) with $u = 0$ and $\tilde{u} = 0$. The resulting $v^{\#}$ in (28) has non-zero degree, leading to a non-zero value of P.

In particular, we can find $m_1, \ldots, m_r$ between 0 and $4g$ with (30) non-zero. The resulting $v = m_1 v_1 + \cdots + m_r v_r$ has

$$\ell(v) \ll \max\{\ell(v_1), \ldots, \ell(v_r)\} \ll \mathcal{D}^{1/2}$$

by (29). Now $\ell(v_0) = \ell(\tilde{v}_0) = \sqrt{2g}$ by (22). From the multiplicativity in Lemma 2.2 of [27, p. 644], we deduce $\ell(v^{\#}) \ll \ell(v)$. Now it is $v^{\#}$ that we need, and the lemma is proved. □

In fact the result contains a bound for the degree of f, because by Lemma 2.3 of [27, p. 644] we have

$$(\deg f)(\deg \tilde{f}) = \deg v^{\#} \leq \ell(v^{\#})^{4g}. \tag{31}$$

But to apply Lemma 4.1 we need the length.

To estimate the various $\underline{y}_i$ in Lemma 4.1 we need heights. For an abelian variety $\underline{A}$ defined over $\overline{\mathbf{Q}}$, we use the logarithmic (semi-stable) Faltings height $h(\underline{A})$ (not always non-negative). Here we need even fewer properties of the fundamental domain.

LEMMA 4.3. *Given $\underline{g} \geq 1$ and δ with $0 < \delta \leq 1$, there is $C = C(\underline{g}, \delta)$ with the following property. Suppose that for $\underline{\tau} = \underline{x} + i\underline{y}$ corresponding to $\underline{A}$, defined over a number field of degree at most D, we have $\underline{y} \geq \delta \underline{y}^{(0)}$ and $\underline{y}^{(0)} \geq \delta \underline{\iota}$ from (23). Then*

$$\underline{y}_i \leq CD \max\{1, h(\underline{A})\} \ (i = 1, 2, \ldots, \underline{g}).$$

Proof. We follow [25], first using the proof of Lemma 8.6 (p. 440). We start with the observation that there exist cusp forms $\varphi_1, \ldots, \varphi_r$ on the full modular group Γ, of the same weight w and with Fourier coefficients in $\mathbf{Q}$, having no common zeroes. There is therefore a $k = 1, 2, \ldots, r$ such that $u = \varphi_k(\underline{\tau}) \neq 0$. We next use Lemma 8.4 (p. 439) with $r = 1$ and γ_1 as the identity, as well as

σ as the identity, so $\tau(\sigma)$ can be chosen as $\underline{\tau}$. The existence of m is clear. The conclusion with $t = 0$, after throwing away the other conjugates, is

$$D^{-1}\log\max\{1, |u|^{-1}\} \ll \max\{1, h(O)\} \ll \max\{1, h(\underline{A})\}$$

(for the second inequality, see (7.4) of [25] p. 436) for the origin O of $\underline{A}$.

We next use the arguments of Lemma 20 of [18, p. 104]. The statement involves $\underline{y}$ in the fundamental domain for Γ, but the proof works under our weaker hypotheses on $\underline{y}$ (as Igusa himself says). It gives $|u| \le M\exp(-2\pi c\,\mathrm{tr}(\underline{y}))$ for $c > 0$ depending only on $\underline{g}$, and $M > 0$ depending only on $\underline{g}$ and φ_k. So we get

$$\underline{y}_i \;\le\; \underline{y}_1 + \cdots + \underline{y}_{\underline{g}} = \mathrm{tr}(\underline{y}) \ll 1 + \log|u|^{-1} \ll D\max\{1, h(\underline{A})\}$$

as required. □

Finally, to estimate the discriminant we use the following "endomorphism estimate" in the style of Lemma 2.2.

Lemma 4.4. *There are $c = c(\underline{g})$ and $\lambda = \lambda(\underline{g})$ with the following property. If $\underline{A}$ is defined over a number field of degree at most D, then $\mathcal{D}(\underline{A}) \le c\max\{D, h(\underline{A})\}^{\lambda}$.*

Proof. This is the main theorem of [27, p. 641], together with an inequality for C (p.650) for $\delta = 1$. □

Recently similar problems about relating rational representations to degrees have turned up in work [29] of Orr. (See also arXiv `1209.3653v4` incorporating remarks of Dill.) The solutions there do not suffice for our own problem, because neither A nor $\tilde{A}$ is "fixed". However a more recent paper [30] of Orr (using finiteness results about reductive **Q**-algebraic groups) apparently would suffice. This is because our $A, \tilde{A}$ are Hodge generic, so have trivial endomorphism rings, and therefore any connecting isogeny automatically respects polarizations as needed in [30]. We have preferred to present our own approach via [27] (which works for any $A, \tilde{A}$ and may well have other applications).

5. Proof of Theorems 1.1, 1.3 and 1.5

5.1. *Main proofs.* We prove Theorems 1.1 and 1.3 simultaneously by adjoining the extra condition

(c) not Galois generic

to (a) and (b) of Theorem 1.3, and we also include the remark about $\gamma < 1$; we call this the "strong Theorem 1.3." (See [36, p. 274] for the fact that Galois generic implies Hodge generic.) The proof of Theorem 1.5 runs along exactly the same lines, so we shall say no more about it.

We follow the strategy of the proof used for Theorem 1.7, but without any need for complex conjugation. We consider hypersurfaces obtained from $\mathcal{H}$ by isogenies of degree at most M. With Ψ the map from a finite cover of

$\mathcal{A}_g$ to $\mathbf{A}^G$ we then choose our candidates $A_{\mathbf{n}}$ from the various $\Psi^{-1}(\mathbf{n})$. If there is an isogeny between $A_{\mathbf{n}}$ and some $\tilde{A}$ in $\mathcal{H}$, then the analogues Lemma 4.4, Lemma 4.2 and (31) of Lemma 2.2 show that the degree of the field of definition of $\tilde{A}$ must be large. Now Pila-Wilkie-Pila applies again, but we need Lemmas 4.1 and 4.3 to estimate the size of the integral points.

Here is our analogue of Lemma 3.2. Write D for the expression (1). From now on all constants in $\ll, \gg$ may depend only on $\tilde{\mathcal{A}}, \Psi, \mathcal{H}$ (and later ϵ, γ).

It is especially convenient to use the concept of p-Galois generic (see [36, p. 275]) for any fixed p (even $p = 2$ will do). This is the assertion that the group of the division field of points of order p^∞ is open in $\mathrm{GSp}_{2g}(\mathbf{Z}_p)$. It is known that this implies Galois generic; a modern reference is Cadoret [5, Th. 1.2 p. 6].

LEMMA 5.1. *There is $\mu = \mu(g)$ with the following property. Given integers $M \geq 1$ and $N \geq 2$, there are only*

$$\ll N^{G-1}M^{2g} + N^{G-1}(\log N)^{\mu} + N^{G-\frac{1}{2}}\log N \tag{32}$$

elements $\mathbf{n} = (n_1, \ldots, n_G)$ of $\mathbf{Z}^G$ with $1 \leq n_1, \ldots, n_G \leq N$ such that any $A_{\mathbf{n}}$ arising from the projection of $\Psi^{-1}(\mathbf{n})$ to $\mathcal{A}_g$ is either

(a) *not defined over an extension of $\mathbf{Q}$ of degree at most D*
or

(b) *isogenous to some $\tilde{A}$ in $\mathcal{H}$ via an isogeny to $\tilde{A}$ of degree at most M,*
or

(c) *not p-Galois generic.*

Furthermore, if g is odd or $g = 2, 6$, then the last term in (32) can be omitted.

Proof. Clearly each element of $\Psi^{-1}(\mathbf{n})$ has degree bigger than $[F_\Psi : \mathbf{Q}]D_\Psi$ for $\ll N^{G-1}$ possible $\mathbf{n}$, and projecting from $\tilde{\mathcal{A}}$ to $\mathcal{A}_g$ leads at most to an extra factor $[\tilde{F} : \mathbf{Q}]$, so we can forget about (a).

As for (b), we note that the condition for (A', A'') in $\mathcal{A}_g \times \mathcal{A}_g$ that A' is isomorphic to A''/Γ for some subgroup of A'' with cardinality at most M is a correspondence of degree $\ll M^{2g}$ in the first factor, as this counts the number of Γ; for example, see again Lemma 6.1 of [26, p. 469].

Thus if $A_{\mathbf{n}}$ is isomorphic to $\tilde{A}/\Gamma$ as above, then we are intersecting this correspondence with a fixed hypersurface in the second factor, so in all we get a hypersurface in the first factor of degree $\ll M^{2g}$ to avoid. This gives $\ll N^{G-1}M^{2g}$ points, so we can forget about these too.

As for (c), it follows from Serre's famous results [39, p. 35] that any $A_{\mathbf{n}}$ is p-Galois generic when g is odd or $g = 2, 6$ provided $\mathrm{End}A_{\mathbf{n}} = \mathbf{Z}$. Now that holds for the generic point of $\mathbf{A}^G$, and so the main theorem of [23, p. 459; see also p. 474] gives an estimate $\ll N^{G-1}(\log N)^{\mu}$ for the exceptional $\mathbf{n}$.

Even for $g = 4$ it is known that endomorphism ring $\mathbf{Z}$ does not suffice. But we can obtain p-Galois genericity in our situation using the following arguments, based on those appearing in Serre's book [38] and his letter to Ribet in [39].

For generic $\mathbf{x}$ in $\mathbf{A}^G$, any $A_{\mathbf{x}}$ in the projection of $\Psi^{-1}(\mathbf{x})$ to $\mathcal{A}_g$ is defined over a finite extension $k_{\mathbf{x}}$ of $\mathbf{Q}(\mathbf{x})$. Let $K_{\mathbf{x}} = k_{\mathbf{x}}(A_{\mathbf{x}}[p^\infty])$ be the division field. By Lemma 4.4.16 of Deligne [10, p. 56] the group $G = \mathrm{Gal}(K_{\mathbf{x}}/k_{\mathbf{x}})$ contains an open subgroup of $\mathrm{Sp}_{2g}(\mathbf{Z}_p)$. So from the Weil pairing, G is open in $\mathrm{GSp}_{2g}(\mathbf{Z}_p)$.

The analogue of the assumption "Ram" in Serre [38, p. 149] holds because torsion points yield unramified extensions outside bad reduction. Thus for $\mathbf{x}$ outside some hypersurface Ω_0 in $\mathbf{A}^G$, the decomposition group $G_{\mathbf{x}}$ at $\mathbf{x}$ is the Galois group of the corresponding residue field $K_{\mathbf{x}}/k_{\mathbf{x}}$; it is a subgroup of G (defined up to conjugation).

Let N be the Frattini subgroup of G. This is also open in $\mathrm{GSp}_{2g}(\mathbf{Z}_p)$ by Proposition (iv) of [38, p. 148; see also Example 1, p. 149]. Let $F_{\mathbf{x}}$ be the fixed field of N in $K_{\mathbf{x}}$, a finite extension of $k_{\mathbf{x}}$ with $H = \mathrm{Gal}(F_{\mathbf{x}}/k_{\mathbf{x}}) = G/N$ finite. For the number field $k = \overline{\mathbf{Q}} \cap k_{\mathbf{x}}$, we can identify $F_{\mathbf{x}}$ with the function field $k(X)$ of a variety X irreducible over k. Then by Proposition 2 of [38, p. 123] there is a thin set Ω in k^G such that for all $\mathbf{y}$ in k^G outside Ω (and Ω_0), the decomposition group $H_{\mathbf{y}}$ is the same as H. For such $\mathbf{y}$, we have $G_{\mathbf{y}}.N = G$, hence $G_{\mathbf{y}} = G$ by the Frattini property.

Thus $A_{\mathbf{y}}$ is p-Galois generic.

Finally by the proposition of [38, p. 128], also $\Omega \cap \mathbf{Q}^G$ is thin in $\mathbf{Q}^G$. Thus by Cohen's Theorem [9, p. 229] — see also [38, p. 177] — this intersection contains $\ll N^{G-\frac{1}{2}} \log N$ integral points $\mathbf{n} = (n_1, \ldots, n_G)$ with $1 \le n_1, \ldots, n_G \le N$. This is the last contribution to (32). In the general context of Hilbert Irreducibility it is well known that the saving $\frac{1}{2}$ in the exponent cannot be improved (unless other methods such as those of [23] are available).

This completes the proof of Lemma 5.1. □

In connection with the original question, we note that the Jacobian locus always contains something Hodge generic. (See, for example, [19], or [2, Th. 8.1.1] even for just the hyperelliptic case.) This prevents us from answering the question affirmatively simply by picking something in $\mathcal{A}_g(\overline{\mathbf{Q}})$ (for example using Lemma 5.1 for $M = 1$) that is Hodge generic and then appealing to the isogeny invariance of this property.

As in the proof of Theorem 1.7 we will choose M in terms of N so that the number of exceptions in Lemma 5.1 is essentially that appearing in Theorem 1.3. However the argument is a bit more elaborate than that of Lemma 3.3, so the choice will be done after the following analogue of that lemma.

Fix any $\gamma < 1$ when g is odd or $g = 2, 6$; and otherwise any $\gamma < 1/2$. Thus the subsequent implied constants may depend on γ as well. Let $\lambda = \lambda(g)$ be as in Lemma 4.4.

LEMMA 5.2. *Suppose some $A = A_{\mathbf{n}}$ as above is isogenous to some $\tilde{A}$ in $\mathcal{H}$. Then there is an isogeny f from A to $\tilde{A}$ of degree*

$$\tilde{m} \ll \max\{\tilde{D}, \log N\}^{2g\lambda}, \tag{33}$$

where $\tilde{D} \geq 2$ is an upper bound for the degree of the field of definition $\tilde{K}$ of $\tilde{A}$. Further, we have

$$\mathcal{D}(A \times \tilde{A}) \ll \max\{\tilde{D}, \log N + h(\tilde{A})\}^{\lambda} \tag{34}$$

and

$$\max\{1, h(A), h(\tilde{A})\} \ll \log N + \log \tilde{m}. \tag{35}$$

Proof. In Section 3 for elliptic curves, we argued with degrees and the help of Lemma 2.1 but in view of its analogue Lemma 4.1 we now need lengths. By Lemma 4.2 there are isogenies f from A to $\tilde{A}$ and $\tilde{f}$ from $\tilde{A}$ to A such that $\ell(v) \ll \mathcal{D}(A \times \tilde{A})^{1/2}$ for v defined by (26). In particular, by (31) we have $\tilde{m} = \deg f \ll \mathcal{D}(A \times \tilde{A})^{2g}$. Here Lemma 4.4 gives

$$\mathcal{D}(A \times \tilde{A}) \ll \max\{\tilde{D}, h(A \times \tilde{A})\}^{\lambda} \ll \max\{\tilde{D}, \log N + h(\tilde{A})\}^{\lambda}$$

and so (34). Also a standard property of Faltings heights yields

$$h(\tilde{A}) \ \leq \ h(A) + \frac{1}{2}\log \tilde{m} \ll \log N + \log \tilde{m}$$

and so (35). We also get

$$\tilde{m} \ll \max\{\tilde{D}, \log N + \log \tilde{m}\}^{2g\lambda}.$$

Thus we can omit the $\log \tilde{m}$ on the right to end up with (33). □

We can now fix M. By our choice of $\mathbf{n}$ we have $M < \tilde{m}$, and so if we choose $M = [(\log N)^{\nu}]$ for any fixed $\nu > 2g\lambda$, we get

$$(\log N)^{\nu} \ll \tilde{m} \ll \tilde{D}^{2g\lambda}. \tag{36}$$

So now the number of exceptional $\mathbf{n}$ is $\ll N^{G-\gamma}(\log N)^{2g\nu}$, which is indeed essentially the upper bound appearing in Theorem 1.3.

We next show that if N is sufficiently large, then for any $\mathbf{n}$ outside the exceptional set of Lemma 5.1, the $A_{\mathbf{n}}$ works for the strong Theorem 1.3. As in the proof of Theorem 1.7, this will be via assuming that $A_{\mathbf{n}}$ is isogenous to some $\tilde{A}$ in $\mathcal{H}$, then using Pila [32] to get

$$\tilde{D} \ll 1 \tag{37}$$

and then the contradiction $N \ll 1$. So the rest of the argument is devoted to (37).

We take Galois conjugates as in the proof of Theorem 1.7. But the arguments are complicated by the lack of a simple analogue of Lemma 2.1; also at the end we will have to play off Hodge generic against weakly special.

Let σ be any embedding of $\tilde{K}$ in $\mathbf{C}$ fixing the fields of definition of A and $\mathcal{H}$. Then $\tilde{A}^\sigma$ is in $\mathcal{H}$, and $A, \tilde{A}^\sigma$ are isogenous. By Lemma 4.2 there are isogenies f_σ from A to $\tilde{A}^\sigma$ and $\tilde{f}_\sigma$ from $\tilde{A}^\sigma$ to A such that

$$\ell_\sigma(v_\sigma) \ll \mathcal{D}(A \times \tilde{A}^\sigma)^{1/2} = \mathcal{D}(A \times \tilde{A})^{1/2} \tag{38}$$

for v_σ defined by (26) with $f_\sigma, \tilde{f}_\sigma$ and a length ℓ_σ coming from any Rosati form on $A \times \tilde{A}^\sigma$. (Recall that all the estimates are independent of choice of principal polarization.)

Next choose $\tau_{\mathbf{n}}$ and $\tilde{\tau}_\sigma$ in a Siegel fundamental domain $\mathcal{F}_g$ in $\mathbf{H}_g$ corresponding to $A = A_{\mathbf{n}}$ and $\tilde{A}^\sigma$ respectively. Just as in (21), the isogenies $f_\sigma, \tilde{f}_\sigma$ lead to matrix equations

$$\kappa_\sigma(\iota\ \tau_{\mathbf{n}}) = (\iota\ \tilde{\tau}_\sigma)\rho_\sigma,\ \tilde{\kappa}_\sigma(\iota\ \tilde{\tau}_\sigma) = (\iota\ \tau_{\mathbf{n}})\tilde{\rho}_\sigma \tag{39}$$

for integral

$$\rho_\sigma = \begin{pmatrix} a_\sigma & -b_\sigma \\ -c_\sigma & d_\sigma \end{pmatrix},\ \tilde{\rho}_\sigma = \begin{pmatrix} \tilde{a}_\sigma & -\tilde{b}_\sigma \\ -\tilde{c}_\sigma & \tilde{d}_\sigma \end{pmatrix}.$$

We pause to show that

$$\det(c_\sigma \tau_{\mathbf{n}} + d_\sigma) \neq 0 \tag{40}$$

and

$$\tilde{\tau}_\sigma = (a_\sigma \tau_{\mathbf{n}} + b_\sigma)(c_\sigma \tau_{\mathbf{n}} + d_\sigma)^{-1}. \tag{41}$$

Namely, from the first of (39) we get

$$\tilde{\tau}_\sigma(c_\sigma \tau_{\mathbf{n}} + d_\sigma) = a_\sigma \tau_{\mathbf{n}} + b_\sigma \tag{42}$$

and then

$$\begin{pmatrix} a_\sigma & -b_\sigma \\ -c_\sigma & d_\sigma \end{pmatrix} \begin{pmatrix} \iota & -\tau_{\mathbf{n}} \\ o & \iota \end{pmatrix} = \begin{pmatrix} a_\sigma & \tilde{\tau}_\sigma(c_\sigma \tau_{\mathbf{n}} + d_\sigma) \\ -c_\sigma & c_\sigma \tau_{\mathbf{n}} + d_\sigma \end{pmatrix}.$$

If (40) fails, then there is a non-zero column $\mathbf{p}$ with $(c_\sigma \tau_{\mathbf{n}} + d_\sigma)\mathbf{p} = \mathbf{0}$. But then multiplying the above on the right by $\binom{\mathbf{0}}{\mathbf{p}}$ gives $\rho_\sigma \binom{-\tau_{\mathbf{n}}\mathbf{p}}{\mathbf{p}} = 0$ showing $\det \rho_\sigma = 0$. However this determinant is none other than the degree of f_σ.

Thus (40) holds, and (23) follows from (42).

We have a natural matrix $\underline{\tau} = \begin{pmatrix} \tau_{\mathbf{n}} & o \\ o & \tilde{\tau}_\sigma \end{pmatrix}$ for $A_{\mathbf{n}} \times \tilde{A}^\sigma$ in $\mathbf{H}_{2g}$. But as explained this might not be in the standard fundamental domain $\mathcal{F}_{2g}$, due to the stringent condition of Minkowski-reduced on the imaginary part. Possibly this problem can be solved just by permuting the successive minima. However we have

$$\begin{pmatrix} o & \tilde{\kappa}_\sigma \\ \kappa_\sigma & o \end{pmatrix} \begin{pmatrix} \iota & o & \tau_{\mathbf{n}} & o \\ o & \iota & o & \tilde{\tau}_\sigma \end{pmatrix} = \begin{pmatrix} \iota & o & \tau_{\mathbf{n}} & o \\ o & \iota & o & \tilde{\tau}_\sigma \end{pmatrix} \begin{pmatrix} o & \tilde{a}_\sigma & o & -\tilde{b}_\sigma \\ a_\sigma & o & -b_\sigma & o \\ o & -\tilde{c}_\sigma & o & \tilde{d}_\sigma \\ -c_\sigma & o & d_\sigma & o \end{pmatrix}. \tag{43}$$

(Note the "skew-diagonal" and the position of the block matrices.)

We can apply Lemma 4.1 because for

$$\underline{y} = \Im \underline{\tau} = \begin{pmatrix} \Im\tau_{\mathbf{n}} & o \\ o & \Im\tilde{\tau}_\sigma \end{pmatrix} = \begin{pmatrix} y_{\mathbf{n}} & o \\ o & \tilde{y}_\sigma \end{pmatrix}$$

(say) and the corresponding diagonal matrices $y_{\mathbf{n}}^{(0)}, \tilde{y}_\sigma^{(0)}$, we certainly have

$$y_{\mathbf{n}}^{(0)} \gg y_{\mathbf{n}} \gg y_{\mathbf{n}}^{(0)} \gg \iota, \ \tilde{y}_\sigma^{(0)} \gg \tilde{y}_\sigma \gg \tilde{y}_\sigma^{(0)} \gg \iota,$$

and these easily imply the same for $\underline{y}, \underline{y}^{(0)}$ as in (23). We conclude by cherry-picking (43) that

$$M_\sigma = \max\{||a_\sigma||, ||b_\sigma||, ||c_\sigma||, ||d_\sigma||\} \ll ||\underline{y}||\ell_\sigma(v_\sigma) \tag{44}$$

for the supremum norms.

Here by Lemma 4.3 we have

$$||\underline{y}|| \ll \tilde{D}\max\{1, h(A) + h(\tilde{A}^\sigma)\} = \tilde{D}\max\{1, h(A) + h(\tilde{A})\},$$

which by (35) and (36) is

$$\ll \tilde{D}(1 + \log N + \log \tilde{m}) \ll \tilde{D}^2.$$

Similarly, by (38) and (34) we get

$$\ell_\sigma(v_\sigma) \ll \max\{\tilde{D}, \log N + h(\tilde{A})\}^{\lambda/2} \ll \tilde{D}^{\lambda/2}.$$

Thus from (44) we get

$$M_\sigma \ll \tilde{D}^{2+\lambda/2} \ll \tilde{D}^\lambda \tag{45}$$

assuming $\lambda \geq 4$.

With J from $\mathbf{H}_g$ to $\mathcal{A}_g$ an analogue of j, write

$$Z = \mathcal{F}_g \cap J^{-1}(\mathcal{H})$$

containing the $\tilde{\tau}_\sigma$.

For τ in $\mathbf{H}_g$, also write W_τ as the set of all $X = \left(\begin{smallmatrix} x_a & -x_b \\ -x_c & x_d \end{smallmatrix}\right)$ in the matrix ring $M_2(M_g(\mathbf{R})) = \mathbf{R}^{4g^2}$ with

$$\det(x_c\tau + x_d) \neq 0$$

corresponding to (40) and

$$\pi_\tau(X) = (x_a\tau + x_b)(x_c\tau + x_d)^{-1}$$

in Z. As in Section 3 this is definable. (See, in particular, the work [31] of Peterzil and Starchenko.) Here too the algebraic part is full, so we use blocks as in Section 3, with the projection π_τ to Z.

All goes through as before with $\tau = \tau_{\mathbf{n}}$ and the integral point $\rho_\sigma = \left(\begin{smallmatrix} a_\sigma & -b_\sigma \\ -c_\sigma & d_\sigma \end{smallmatrix}\right)$. Using (45) we see that ρ_σ lies in the analogous set $W_\tau(T)$ with $T \ll \tilde{D}^\lambda$, provided the associated block $\pi_\tau(B_\sigma)$ has zero dimension. But the latter is no longer automatic, as Z^{alg} might not be empty. (Indeed it might not be if $\mathcal{H}$ contains a positive dimensional special variety, as was the case for modular $\mathcal{C}$ in Section 3.)

We claim anyway that $\pi_\tau(B_\sigma)$ has zero dimension. It would then follow by fixing any $\epsilon < 1/\lambda$ in the T^ϵ from Pila's [32], and noting that the number of different $\pi_\tau(\rho_\sigma) = \tilde{\tau}_\sigma$ is $\gg \tilde{D}$, that $\tilde{D} \ll 1$ as in (37) above. Thus by (36) we conclude $N \ll 1$, and we would be done.

If, on the contrary, $\pi_\tau(B_\sigma)$ had positive dimension, it would lie in Z^{alg}, and so $\pi_\tau(\rho_\sigma) = \tilde{\tau}_\sigma$ also. Thus $\tilde{\tau}_\sigma$ lies also in some semi-algebraic curve Γ in Z. Then Γ lies in $J^{-1}(\mathcal{H})$. By Theorem 6.1 of [34, p. 670] there is some weakly special $\mathcal{K}$ in $\mathcal{H}$ with Γ in $J^{-1}(\mathcal{K})$, and so $J(\tilde{\tau}_\sigma)$ is in $\mathcal{K}$. But $J(\tilde{\tau}_\sigma) = \tilde{A}^\sigma$ of course.

Now $\tilde{A}$ is isogenous to our p-Galois generic A, and so $\tilde{A}$ is also p-Galois generic (as isogenies change the Galois groups only up to finite index). Thus also the Galois conjugate $\tilde{A}^\sigma$ is p-Galois generic. So $\tilde{A}^\sigma$ is Hodge generic.

In fact it is known that the only weakly special $\mathcal{K}$ containing something Hodge generic are points or the whole $\mathcal{A}_g$. A precise reference can be found in Gao's appendix to a paper [2] by André, Corvaja and the second author of the present work. Namely, $\mathcal{K}$ satisfies (ii) of Lemma 10.2.6 of [2]. Thus it also satisfies (iv), which says that it is not contained in any proper bi-algebraic variety of positive dimension. By Ullmo and Yafaev [45, Th. 1.2, p. 264] the latter are precisely the weakly special; thus we see that indeed $\mathcal{K}$ is either a point or the whole $\mathcal{A}_g$. Both are excluded by the fact that $J(\Gamma)$ lies in $\mathcal{K}$, which lies in $\mathcal{H}$. This justifies our claim about $\pi_\tau(B_\sigma)$.

Note that the (closure of the) set of Jacobians (when $g \geq 4$) cannot be weakly special because we already remarked that the Jacobian of a generic hyperelliptic curve is Hodge generic. However it does contain many weakly special varieties of positive dimension, for example products. Or less trivially for $g = 4$ there are the Jacobians of the well-known family $y^5 = x(x-1)(x-a)$ with CM by $\mathbf{Q}(\sqrt{-5})$. For much more, see Moonen and Oort [28].

5.2. *On degrees.* To finish the proofs it remains to exhibit $\tilde{\mathcal{A}}, \Psi$ with sharpened $D(\tilde{\mathcal{A}}, \Psi) = 2^{16g^4-1}$ in (1). For $\gamma = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ in $\Gamma = \mathrm{Sp}_{2g}(\mathbf{Z})$, write

$$\Delta(\gamma, \tau) = \det(r\tau + s)$$

as in the standard automorphy factor. Let Λ be normal of index n in Γ. We say that a function φ analytic on $\mathbf{H}_g$ is a Λ-form of weight k if $\varphi(\lambda(\tau)) = \Delta(\lambda,\tau)^k\varphi(\tau)$ for every λ in Λ.

LEMMA 5.3. *Let φ be a Λ-form of weight k. Then if*

$$\Gamma = \bigcup_{i=1}^{n} \gamma_i\Lambda = \bigcup_{i=1}^{n} \Lambda\gamma_i$$

with γ_1 the identity, the product

$$\Phi_1(\tau) = \prod_{i=2}^{n} \frac{\varphi(\gamma_i(\tau))}{\Delta(\gamma_i,\tau)^k}$$

is a Λ-form of weight $(n-1)k$ and $\Phi = \varphi\Phi_1$ is a Γ-form of weight nk.

Proof. We consider Φ first. Write $\varphi_i(\tau) = \varphi(\gamma_i(\tau))$ $(i = 1,\ldots,n)$. For any γ in Γ, there are $\lambda = \lambda(i,\gamma)$ in Λ and $j = j(i,\gamma)$ with $\gamma_i\gamma = \lambda\gamma_j$; for each γ, we get a permutation of $\{1,\ldots,n\}$. It follows that

$$\begin{aligned}\varphi_i(\gamma(\tau)) &= \varphi((\gamma_i\gamma)(\tau)) = \varphi((\lambda\gamma_j)(\tau))\\ &= \Delta(\lambda,\gamma_j(\tau))^k\varphi(\gamma_j(\tau)) = \Delta(\lambda,\gamma_j(\tau))^k\varphi_j(\tau).\end{aligned}$$

Thus γ permutes $\{\varphi_1,\ldots,\varphi_n\}$ modulo automorphy.

It follows that

$$\Phi(\gamma(\tau)) = \prod_{i=1}^{n} \frac{\varphi_i(\gamma(\tau))}{\Delta(\gamma_i,\gamma(\tau))^k} = \prod_{i=1}^{n} \frac{\Delta(\lambda,\gamma_j(\tau))^k}{\Delta(\gamma_i,\gamma(\tau))^k}\varphi_j(\tau), \tag{46}$$

which is

$$\Phi(\tau)\prod_{i=1}^{n} \frac{\Delta(\lambda,\gamma_j(\tau))^k}{\Delta(\gamma_i,\gamma(\tau))^k}\Delta(\gamma_j,\tau)^k = \Phi(\tau)\Delta(\gamma,\tau)^{nk}, \tag{47}$$

where we have used the well-known composition rule

$$\Delta(gh,\tau) = \Delta(g,h(\tau))\Delta(h,\tau)$$

(an identity in g,h,τ).

Now (46) and (47) give the functional equation for Γ-forms of weight nk, and it is clear that Φ is analytic on the Siegel space. So Φ is indeed a Γ-form.

Next Φ,φ satisfy the equations for Λ-forms of weights nk, k respectively, so the same holds for $\Phi_1 = \Phi/\varphi$ with weight $(n-1)k$. Again Φ_1 is analytic, so it is a Λ-form of weight $(n-1)k$. □

Presumably Φ is independent of the choice of $\gamma_1,\ldots,\gamma_n$.

Given a Λ-form φ there is an integer $d \geq 1$ and a Fourier expansion (recall $g \geq 2$, so no Koecher needed)

$$\varphi(\tau) = \sum_M a(M) \exp(\pi i \operatorname{tr}(M\tau))$$

taken over all positive semi-definite symmetric matrices M with dM half-integral in the usual sense. For $\varphi \neq 0$, we define

$$\operatorname{ord}(\varphi) = \min_{a(M)\neq 0} \operatorname{tr}(M) \geq 0$$

(not necessarily an integer). It is easy to see that $\operatorname{ord}(\varphi_1\varphi_2) \geq \operatorname{ord}(\varphi_1) + \operatorname{ord}(\varphi_2)$. (Perhaps with a lexicographic argument one could prove equality.)

Define κ_g as in Igusa [18, p. 197].

LEMMA 5.4. *Let $\varphi \neq 0$ be a Λ-form of weight k. Then*

$$\operatorname{ord}(\varphi) \leq \frac{\kappa_g n k}{4\pi}.$$

Proof. With Φ, Φ_1 as in Lemma 5.3 we have $\Phi \neq 0$ too. Then

$$\operatorname{ord}(\Phi) \geq \operatorname{ord}(\Phi_1) + \operatorname{ord}(\varphi) \geq \operatorname{ord}(\varphi).$$

By Theorem 7 of [18, p. 206] we have $\operatorname{ord}(\Phi) \leq \kappa_g nk/(4\pi)$. □

We now specialize to $\Lambda = \Gamma(e, 2e)$ as in [25, p. 422], with e an (even) integer (not a matrix) so that normality follows from [18, pp. 177, 178]. For row vectors m, m^* in $\mathbf{R}^g$, we use the standard

$$\theta_{mm^*}(\tau) = \sum_h \exp\{\pi i (h+m)\tau(h+m)^t + 2\pi i (h+m) m^{*t}\}$$

with the sum over all row vectors h in $\mathbf{Z}^g$, where t denotes the transpose. Here we may regard m in the quotient $(\mathbf{R}/\mathbf{Z})^g$. Then the various $\theta_{m0}(e\tau)^2$ (m in $e^{-1}\mathbf{Z}^g/\mathbf{Z}^g$) of weight 1 (see [18, p. 185] or [25, p. 423]) are Λ-forms. It is easy to check that we may take the d above as e.

Recall that $G = g(g+1)/2$.

LEMMA 5.5. *Given any $\mathbf{C}$-linear combinations $\chi_1, \ldots, \chi_{G+2}$ of the $\theta_{m0}(e\tau)$, any real $W \geq 0$ and any integer $D \geq 0$ with*

$$(D+1)^{G+1} > (G+1)!(4eW+1)^G, \tag{48}$$

there is a non-zero polynomial P in $\mathbf{C}[X_1, \ldots, X_{G+2}]$, homogeneous of degree D, with $\operatorname{ord}(\varphi) \geq W$ for $\varphi = P(\chi_1^2, \ldots, \chi_{G+2}^2)$ provided $\varphi \neq 0$.

Proof. We have

$$\frac{(D+1)\cdots(D+G+1)}{(G+1)!} \geq \frac{(D+1)^{G+1}}{(G+1)!}$$

coefficients at our disposal. The number of conditions is the number of M with eM half-integral and $\mathrm{tr}(M) \le W$. Thus $\mathrm{tr}(eM) \le eW$, and we get an upper bound $(4eW+1)^G$ (see [18, p. 208]). This completes the proof. □

Thus if we choose $W = W_0 = Nk$ with $N = \kappa_g n/(4\pi)$ and $k = D$ in Lemma 5.5, we must get $\varphi = 0$ by Lemma 5.4. Therefore the degree of the Zariski closure of the variety V_e parametrized by the $\theta_{m0}(e\tau)$ (which is known to be quasi-projective when 8 divides e; see [18, p. 415], where we also take e as a square) is at most $2D$.

By (48) this holds for any D with $(D+1)^{G+1} > (G+1)!(4eN(D+1)^G$, which we can secure with $D \le (G+1)!(4eN)^G$. Now $\kappa_g \le (2g/\sqrt{3})c_g$ (see [18, p. 197]) for the Minkowski constant

$$c_g \;\le\; \left(\frac{4}{\pi}\right)^g \Gamma\left(\frac{g+1}{2}\right)^2 \left(\frac{3}{2}\right)^{(g-1)(g-2)}$$

(see, for example, [22, p. 63]). Also $n \le e^{2g^2}(2e)^{2g^2}$ can be seen by noting that $\Gamma(e, 2e)$ contains the group of all $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ in Γ with p, s congruent to 1 mod e and q, r congruent to 0 mod $2e$ (also between the principal congruence subgroups mod e and mod $2e$). We conclude with $e = 16$ a projective degree at most

$$2(G+1)!\left(\frac{32g}{\pi\sqrt{3}}512^{2g^2}c_g\right)^G \;\le\; 2^{16g^4-1}.$$

Now we take the cover $\tilde{\mathcal{A}}$, which is algebraically V_{16} and analytically the quotient of $\mathbf{H}_g$ by $\Gamma(16, 32)$, and the map Ψ defined, for example, by a suitable subset of the $\theta_{m0}(e\tau)/\theta_{00}(e\tau)$ $(m \ne 0)$. This makes it clear that $D_\Psi \le 2^{16g^4-1}$ and $F_\Psi = \mathbf{Q}$ in (1); also $\tilde{F} = \mathbf{Q}$ because $\tilde{\mathcal{A}}, \mathcal{A}_g$ are both defined over $\mathbf{Q}$. (For the former, see [25, p. 415]; the latter is well known.)

This completes the proof of Theorems 1.1 and 1.3 (and Theorem 1.5).

Note that by the definition of $\Gamma(e, 2e)$, the extension of $\mathbf{Q}$ in the above construction is a field of definition not only for the abelian variety but also its points of order 16.

5.3. *Isogeny classes and proof of Corollary* 1.4. We shall need standard isogeny estimates from [26] generalizing Lemma 2.2. These suffice to prove the assertions made in connection with Theorems 1.3 and 1.5 about many different isogeny classes. We omit the details, as they follow closely the remarks at the end of Section 2, now using $(\mathbf{Q}/\mathbf{Z})^{2g}$.

To prove Corollary 1.4 we start by finding a dense set $\mathcal{A}_g^*$ of principally polarized abelian varieties of dimension g, with each defined over an extension of $\mathbf{Q}$ of degree at most 2^{16g^4} and not isogenous to any Jacobian. (This cannot be done by applying elements of $\mathrm{Sp}_{2g}(\mathbf{Q})$ to a single abelian variety because the field degrees would blow up.) After that we will thin it out to separate the isogeny classes.

We take the $\tilde{\mathcal{A}}, \Psi$ in Theorem 1.3 with $D(\tilde{\mathcal{A}}, \Psi) = 2^{16g^4-1}$. To get something in $\mathcal{A}_g$ close to any given A_0 in $\mathcal{A}_g(\mathbf{C})$, we lift A_0 to $\tilde{A}_0$ in $\tilde{\mathcal{A}}$. Then we choose $(\xi_1, \ldots, \xi_G)$ in $\mathbf{Q}(i)^G$ close to $\Psi(\tilde{A}_0)$ in $\mathbf{C}^G$. We borrow a trick sometimes used in Hilbert Irreducibility. (See, for example, Fried and Jarden [14, p. 264].) Let d be a large positive integer, and apply Theorem 1.3 instead to $\Lambda_d \circ \Psi$, where

$$\Lambda_d(x_1, \ldots, x_G) = \left(\frac{1}{d(x_1 - \xi_1)}, \ldots, \frac{1}{d(x_G - \xi_G)}\right). \tag{49}$$

The inverse images

$$\mathbf{x} = \Lambda_d^{-1}(\mathbf{n}) = \left(\xi_1 + \frac{1}{dn_1}, \ldots, \xi_G + \frac{1}{dn_G}\right)$$

are also in $\mathbf{Q}(i)^G$. Now for almost all $\mathbf{n}$ (we will need only a single one) it is easy to see (as in the proof of (a) of Lemma 5.1) that anything in $\Psi^{-1}(\mathbf{x})$ is defined over an extension of $\mathbf{Q}$ of degree at most $2D_\Psi \leq 2^{16g^4}$; and so we end up with this degree in $\mathcal{A}_g$. The required density follows by making d tend to infinity.

For the thinning out, we can certainly find non-empty open subsets $U_1, U_2, \ldots$ of $\mathcal{A}_g$, whose diameters tend to zero, such that for any $n \geq 1$, the union of $U_n, U_{n+1}, \ldots$ is $\mathcal{A}_g$. We pick any A_1 in $\mathcal{A}_g^*$ also in U_1. By density there are infinitely many A in $\mathcal{A}_g^*$ also in U_2, but again by isogeny estimates they cannot all be isogenous to A_1. So we can pick A_2 in $\mathcal{A}_g^*$ also in U_2 not isogenous to A_1. Then with A_3 in U_3 not isogenous to A_1 or A_2, and so on. The resulting subset $\{A_1, A_2, \ldots\}$ of $\mathcal{A}_g^*$ remains dense and so completes the proof of Corollary 1.4.

Probably extending our method of proof and using instead of (49)

$$\left(\frac{x_1 - \xi_1}{p^e}, \ldots, \frac{x_G - \xi_G}{p^e}\right)$$

for a large prime-power p^e will give some sort of ultrametric analogue of Corollary 1.4.

5.4. *Further remarks.* To verify the counting assertions for $\Psi^{-1}(\mathbf{n})$ and $\Xi(\mathbf{n})$ with complex multiplication made in connection with Theorems 1.3 and 1.5, we recall that the degrees of their fields of definition are $\ll 1$, that is, bounded independently of N. Thus by Conjecture 7.1 of Pila and Tsimerman [34, p. 673], which has been proved by Tsimerman in [42] for $g = 1, 2, 3, 4, 5, 6$ unconditionally and $g \geq 7$ under the Generalized Riemann Hypothesis, the relevant discriminants are similarly $\ll 1$. Thus by Lemma 7.4 of [34, p. 675] the number of our CM abelian varieties in $\mathcal{A}_g$ is also $\ll 1$.

Next we prove that there is an analytic hypersurface W in $\mathcal{A}_g$ such that each element of $\mathcal{A}_g(\overline{\mathbf{Q}})$ is isogenous to something in W. By Theorems 1.1, 1.3 and 1.5 this W is necessarily transcendental.

Before, in scalars $\tau = x + iy$ we could express x analytically in terms of y. Now, for matrices we try instead to express the entries of x, y in terms of a single real "ghost parameter" t, so again we end up with something like a real analytic curve. The positivity $y > 0$ makes problems, which we solve by writing $y = y'y'^t$ and using y' instead of y. Actually it is convenient to stay away from $y = 0$ by demanding $y > \iota$ so that $y = \iota + ww^t$; we then parametrize w (and x).

We again use the Newton series (19), but now we are able to choose $t_0, t_1, \ldots$ in $[1, 2]$ in advance, and we assume this is done.

Any real symmetric $\tilde{z} > 0$ can be written as $\tilde{z} = \tilde{w}\tilde{w}^t$ for real $\tilde{w}$ that is lower triangular. (This is closely related to the "Cholesky factorization.") It is known that $\tilde{w}$ can be chosen continuously in $\tilde{z}$; in fact from Theorem 4.1 of Stewart [41, p. 518], one can easily deduce that there are $c > 0$ and C, depending only on $\tilde{z}$ and $\tilde{w}$, such that for any $z > 0$ with supremum norm $||z - \tilde{z}|| < c$, there is lower triangular w with $z = ww^t$ and

$$||w - \tilde{w}|| \leq C||z - \tilde{z}||. \tag{50}$$

Much as before, we enumerate as $\tau_0, \tau_1, \tau_2, \ldots$ all τ in $\mathbf{H}_g$ corresponding to abelian varieties in $\mathcal{A}_g$ defined over $\overline{\mathbf{Q}}$.

We start the construction by picking any real $\tilde{x}_0$ satisfying

$$||\tilde{x}_0|| < \epsilon_0 \tag{51}$$

as in (20). At the same time we pick any real non-singular lower triangular $\tilde{w}_0$ with

$$||\tilde{w}_0|| < \epsilon_0. \tag{52}$$

Then

$$\tilde{\tau}_0 = \tilde{x}_0 + i(\iota + \tilde{w}_0\tilde{w}_0^t)$$

lies in $\mathbf{H}_g$. Using $\mathrm{Sp}_{2g}(\mathbf{Q})$ (automatically preserving the isogeny class) we modify the original τ_0 to lie sufficiently near $\tilde{\tau}_0$. Then

$$\tau_0 = x_0 + i(\iota + z_0)$$

for x_0 near $\tilde{x}_0$. So, in particular, we can secure

$$||x_0|| < \epsilon_0 \tag{53}$$

from (51).

And z_0 is near $\tilde{w}_0\tilde{w}_0^t$, so also $z_0 > 0$. Thus by (50) we can find w_0 near $\tilde{w}_0$ with $z_0 = w_0w_0^t$. So, in particular, we can secure

$$||w_0|| < \epsilon_0 \tag{54}$$

from (52); and we record

$$\tau_0 = x_0 + i(\iota + w_0 w_0^t).$$

We continue the construction by picking any real $\tilde{x}_1$ with

$$||c_{10}x_0 + \tilde{x}_1|| < \epsilon_1 \tag{55}$$

as in (20) and then any real non-singular lower triangular $\tilde{w}_1$ with

$$||c_{10}w_0 + \tilde{w}_1|| < \epsilon_1. \tag{56}$$

We define

$$\tilde{\tau}_0 = \tilde{x}_0 + i(\iota + \tilde{w}_0 \tilde{w}_0^t)$$

in $\mathbf{H}_g$. Again using $\mathrm{Sp}_{2g}(\mathbf{Q})$ we modify the original τ_1 to lie sufficiently near $\tilde{\tau}_1$, so

$$\tau_1 = x_1 + i(\iota + z_1),$$

say. In particular, we can secure

$$||c_{10}x_0 + x_1|| < \epsilon_1 \tag{57}$$

from (55). Also (50) shows that $z_1 = w_1 w_1^t$ for w_1 near $\tilde{w}_1$. Now we can secure

$$||c_{10}w_0 + w_1|| < \epsilon_1 \tag{58}$$

from (56), and we record

$$\tau_1 = x_1 + i(\iota + w_1 w_1^t).$$

And so on; for example, we next choose $\tilde{x}_2, \tilde{w}_2$ in accordance with the inequality in (20) involving ϵ_2, and similar arguments enable us to take

$$\tau_2 = x_2 + i(\iota + w_2 w_2^t)$$

with

$$||c_{20}x_0 + c_{21}x_1 + x_2|| < \epsilon_2,\ ||c_{20}w_0 + c_{21}w_1 + w_2|| < \epsilon_2. \tag{59}$$

It is now clear from (53), (54), (57), (58), (59) and so on, that we can construct matrices F_x, F_w of entire functions such that

$$x_n = F_x(t_n),\ w_n = F_w(t_n)\ (n = 0, 1, 2, \ldots)$$

for

$$\tau_n = x_n + i(\iota + w_n w_n^t)\ (n = 0, 1, 2, \ldots).$$

So $\tau_0, \tau_1, \ldots$ all lie on the image of $[1, 2]$ under

$$F = F_x + i(\iota + F_w F_w^t); \tag{60}$$

this image stays in $\mathbf{H}_g$ and has all the appearance of being a real analytic curve. The same is true of its image K in $\mathcal{A}_g$ under J. At any rate we may without difficulty find a complex analytic hypersurface W in $\mathcal{A}_g$ containing K. For example, we pick any two affine coordinates and note that the corresponding analytic functions of t are analytically dependent locally at each point of $[1,2]$. We then use compactness to reduce to a finite number of analytic equations, and finally we multiply them together (or we can appeal to well-known results of Remmert).

Can one take W bounded, as we could in the elliptic analogue? Even though the $\Im F \geq \iota$ in (60), this does not imply the boundedness of $J(F)$.

Finally, to check the assertions about products being in the closure of the Jacobian locus, we use the "Igusa modular form"

$$F_g(\tau) = 2^g U_g(\tau) - V_g(\tau)^2$$

with

$$U_g(\tau) = \sum_{mm^*} \theta_{mm^*}(\tau)^{16}, \ V_g(\tau) = \sum_{mm^*} \theta_{mm^*}(\tau)^8$$

and the sums over all row vectors m, m^* in $2^{-1}\mathbf{Z}^g/\mathbf{Z}^g$. This is actually a Γ-form (of weight 8). For $g = 1, 2, 3$ it vanishes identically on $\mathcal{A}_g$. For $g = 4$, its vanishing defines the closure of the Jacobian locus. See, for example, Grushevsky [16, Th. 3.8].

When

$$\tau = \begin{pmatrix} \dot{\tau} & o \\ o & \ddot{\tau} \end{pmatrix}$$

for blocks of order $\dot{g}, \ddot{g}$ respectively, it is easily seen that

$$U_g(\tau) = U_{\dot{g}}(\dot{\tau}) U_{\ddot{g}}(\ddot{\tau}), \ V_g(\tau) = V_{\dot{g}}(\dot{\tau}) V_{\ddot{g}}(\ddot{\tau}).$$

Thus, for example, if $g = 4$ and $\dot{g} = \ddot{g} = 2$, then we have

$$0 = F_2(\dot{\tau}) = 4U_2(\dot{\tau}) - V_2(\dot{\tau})^2, \ 0 = F_2(\ddot{\tau}) = 4U_2(\ddot{\tau}) - V_2(\ddot{\tau})^2$$

and so $F_4(\tau) = 0$. Thus products of two principally polarized abelian surfaces are in the closure (and, in particular, the product of four elliptic curves). A similar proof works for the product of an elliptic curve and a principally polarized abelian threefold.

References

[1] Y. André, Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire, *J. Reine Angew. Math.* **505** (1998), 203–208. MR 1662256. Zbl 0918.14010. https://doi.org/10.1515/crll.1998.118.

[2] Y. André, P. Corvaja, and U. Zannier, The Betti map associated to a section of an abelian scheme. arXiv 1802.03204.

[3] H. F. Baker, Note on the foregoing paper "Commutative ordinary differential operators," by J. L. Burchnall and J. W. Chaundy, *Proc. Royal Soc. London A* **118** (1928), 584–593. JFM 54.0439.02. https://doi.org/https://doi.org/10.1098/rspa.1928.0070.

[4] Y. Bilu, D. Masser, and U. Zannier, An effective "Theorem of André" for CM-points on a plane curve, *Math. Proc. Cambridge Philos. Soc.* **154** no. 1 (2013), 145–152. MR 3002589. Zbl 1263.14028. https://doi.org/10.1017/S0305004112000461.

[5] A. Cadoret, An open adelic image theorem for abelian schemes, *Int. Math. Res. Not. IMRN* no. 20 (2015), 10208–10242. MR 3455865. Zbl 1352.14030. https://doi.org/10.1093/imrn/rnu259.

[6] J. W. S. Cassels, The arithmetic of certain quartic curves, *Proc. Roy. Soc. Edinburgh Sect. A* **100** no. 3-4 (1985), 201–218. MR 0807702. Zbl 0589.14029. https://doi.org/10.1017/S0308210500013779.

[7] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a Middlebrow Arithmetic of Curves of Genus* 2, *London Math. Soc. Lecture Note Ser.* **230**, Cambridge Univ. Press, Cambridge, 1996. MR 1406090. Zbl 0857.14018. https://doi.org/10.1017/CBO9780511526084.

[8] C.-L. Chai and F. Oort, Abelian varieties isogenous to a Jacobian, *Ann. of Math.* (2) **176** no. 1 (2012), 589–635. MR 2925391. Zbl 1263.14032. https://doi.org/10.4007/annals.2012.176.1.11.

[9] S. D. Cohen, The distribution of Galois groups and Hilbert's irreducibility theorem, *Proc. London Math. Soc.* (3) **43** no. 2 (1981), 227–250. MR 0628276. Zbl 0484.12002. https://doi.org/10.1112/plms/s3-43.2.227.

[10] P. Deligne, Théorie de Hodge. II, *Inst. Hautes Études Sci. Publ. Math.* no. 40 (1971), 5–57. MR 0498551. Zbl 0219.14007. Available at http://www.numdam.org/item/PMIHES_1971__40__5_0/.

[11] M. Dittmann, R. Salvati Manni, and N. Scheithauer, Harmonic theta series and the Kodaira dimension of $\mathcal{A}_6$, 2019. arXiv 1909.07062v1.

[12] M. Dunajski and R. Penrose, On the quadratic invariant of binary sextics, *Math. Proc. Cambridge Philos. Soc.* **162** no. 3 (2017), 435–445. MR 3628200. Zbl 1388.13018. https://doi.org/10.1017/S0305004116000542.

[13] G. Farkas and A. Verra, The universal abelian variety over $\mathcal{A}_5$, *Ann. Sci. Éc. Norm. Supér.* (4) **49** no. 3 (2016), 521–542. MR 3503825. Zbl 1357.14058. https://doi.org/10.24033/asens.2289.

[14] M. D. Fried and M. Jarden, *Field Arithmetic*, second ed., *Ergeb. Math. Grenzgeb.* **11**, Springer-Verlag, Berlin, 2005. MR 2102046. Zbl 1055.12003.

[15] E. Gaudron and G. Rémond, Théorème des périodes et degrés minimaux d'isogénies, *Comment. Math. Helv.* **89** no. 2 (2014), 343–403. MR 3225452. Zbl 1297.11058. https://doi.org/10.4171/CMH/322.

[16] S. Grushevsky, The Schottky problem, in *Current Developments in Algebraic Geometry*, *Math. Sci. Res. Inst. Publ.* **59**, Cambridge Univ. Press, Cambridge, 2012, pp. 129–164. MR 2931868. Zbl 1254.14054.

[17] P. Habegger and J. Pila, Some unlikely intersections beyond André-Oort, *Compos. Math.* **148** no. 1 (2012), 1–27. MR 2881307. Zbl 1288.11062. https://doi.org/10.1112/S0010437X11005604.

[18] J.-i. Igusa, *Theta Functions*, **194**, Springer-Verlag, New York, 1972, Grundlehren Math. Wiss. MR 0325625. Zbl 0251.14016.

[19] N. M. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues, and Monodromy*, *Amer. Math. Soc. Colloq. Publ.* **45**, Amer. Math. Soc., Providence, RI, 1999. MR 1659828. Zbl 0958.11004.

[20] B. Klingler and A. Yafaev, The André-Oort conjecture, *Ann. of Math.* (2) **180** no. 3 (2014), 867–925. MR 3245009. Zbl 1377.11073. https://doi.org/10.4007/annals.2014.180.3.2.

[21] L. Kühne, An effective result of André-Oort type, *Ann. of Math.* (2) **176** no. 1 (2012), 651–671. MR 2925393. Zbl 1341.11035. https://doi.org/10.4007/annals.2012.176.1.13.

[22] C. G. Lekkerkerker, *Geometry of Numbers*, *Bibliotheca Math.* **VIII**, Wolters-Noordhoff Publishing, Groningen; North-Holland Publishing Co., Amsterdam-London, 1969. MR 0271032. Zbl 0198.38002.

[23] D. W. Masser, Specializations of endomorphism rings of abelian varieties, *Bull. Soc. Math. France* **124** no. 3 (1996), 457–476. MR 1415735. Zbl 0866.11040. https://doi.org/10.24033/bsmf.2288.

[24] D. W. Masser and G. Wüstholz, Estimating isogenies on elliptic curves, *Invent. Math.* **100** no. 1 (1990), 1–24. MR 1037140. Zbl 0722.14027. https://doi.org/10.1007/BF01231178.

[25] D. W. Masser and G. Wüstholz, Periods and minimal abelian subvarieties, *Ann. of Math.* (2) **137** no. 2 (1993), 407–458. MR 1207211. Zbl 0796.11023. https://doi.org/10.2307/2946542.

[26] D. W. Masser and G. Wüstholz, Isogeny estimates for abelian varieties, and finiteness theorems, *Ann. of Math.* (2) **137** no. 3 (1993), 459–472. MR 1217345. Zbl 0804.14019. https://doi.org/10.2307/2946529.

[27] D. W. Masser and G. Wüstholz, Endomorphism estimates for abelian varieties, *Math. Z.* **215** no. 4 (1994), 641–653. MR 1269495. Zbl 0826.14025. https://doi.org/10.1007/BF02571735.

[28] B. Moonen and F. Oort, The Torelli locus and special subvarieties, in *Handbook of Moduli. Vol. II*, *Adv. Lect. Math.* (*ALM*) **25**, Int. Press, Somerville, MA, 2013, pp. 549–594. MR 3184184. Zbl 1322.14065.

[29] M. Orr, Families of abelian varieties with many isogenous fibres, *J. Reine Angew. Math.* **705** (2015), 211–231. MR 3377393. Zbl 1349.14143. https://doi.org/10.1515/crelle-2013-0058.

[30] M. Orr, Height bounds and the Siegel property, *Algebra Number Theory* **12** no. 2 (2018), 455–478. MR 3803710. Zbl 06880895. https://doi.org/10.2140/ant.2018.12.455.

[31] Y. Peterzil and S. Starchenko, Definability of restricted theta functions and families of abelian varieties, *Duke Math. J.* **162** no. 4 (2013), 731–765. MR 3039679. Zbl 1284.03215. https://doi.org/10.1215/00127094-2080018.

[32] J. Pila, On the algebraic points of a definable set, *Selecta Math.* (*N.S.*) **15** no. 1 (2009), 151–170. MR 2511202. Zbl 1218.11068. https://doi.org/10.1007/s00029-009-0527-8.

[33] J. Pila, O-minimality and the André-Oort conjecture for $\mathbb{C}^n$, *Ann. of Math.* (2) **173** no. 3 (2011), 1779–1840. MR 2800724. Zbl 1243.14022. https://doi.org/10.4007/annals.2011.173.3.11.

[34] J. Pila and J. Tsimerman, Ax-Lindemann for $\mathcal{A}_g$, *Ann. of Math.* (2) **179** no. 2 (2014), 659–681. MR 3152943. Zbl 1305.14020. https://doi.org/10.4007/annals.2014.179.2.5.

[35] J. Pila and A. J. Wilkie, The rational points of a definable set, *Duke Math. J.* **133** no. 3 (2006), 591–616. MR 2228464. Zbl 1217.11066. https://doi.org/10.1215/S0012-7094-06-13336-7.

[36] R. Pink, A combination of the conjectures of Mordell-Lang and André-Oort, in *Geometric Methods in Algebra and Number Theory*, *Progr. Math.* **235**, Birkhäuser Boston, Boston, MA, 2005, pp. 251–282. MR 2166087. Zbl 1200.11041. https://doi.org/10.1007/0-8176-4417-2_11.

[37] B. Poonen, *Rational Points on Varieties*, *Grad. Stud. Math.* **186**, Amer. Math. Soc., Providence, RI, 2017. MR 3729254. Zbl 1387.14004.

[38] J-P. Serre, *Lectures on the Mordell-Weil Theorem*, third ed., *Aspects of Math.*, Friedr. Vieweg & Sohn, Braunschweig, 1997, translated from the French and edited by Martin Brown from notes by Michel Waldschmidt, With a foreword by Brown and Serre. MR 1757192. Zbl 0863.14013. https://doi.org/10.1007/978-3-663-10632-6.

[39] J-P. Serre, *Œuvres. Collected Papers. IV*, Springer-Verlag, Berlin, 2000, 1985–1998. MR 1730973. Zbl 0933.01034.

[40] A. N. Shankar and J. Tsimerman, Unlikely intersections in finite characteristic, *Forum Math. Sigma* **6** (2018), e13, 17. MR 3841493. Zbl 1403.11050. https://doi.org/10.1017/fms.2018.15.

[41] G. W. Stewart, Perturbation bounds for the QR factorization of a matrix, *SIAM J. Numer. Anal.* **14** no. 3 (1977), 509–518. MR 0436566. Zbl 0358.65038. https://doi.org/10.1137/0714030.

[42] J. Tsimerman, Brauer-Siegel for arithmetic tori and lower bounds for Galois orbits of special points, *J. Amer. Math. Soc.* **25** no. 4 (2012), 1091–1117. MR 2947946. Zbl 1362.11057. https://doi.org/10.1090/S0894-0347-2012-00739-5.

[43] J. Tsimerman, The existence of an abelian variety over $\overline{\mathbb{Q}}$ isogenous to no Jacobian, *Ann. of Math.* (2) **176** no. 1 (2012), 637–650. MR 2925392. Zbl 1250.14032. https://doi.org/10.4007/annals.2012.176.1.12.

[44] J. Tsimerman, The André-Oort conjecture for $\mathcal{A}_g$, *Ann. of Math.* (2) **187** no. 2 (2018), 379–390. MR 3744855. Zbl 1415.11086. https://doi.org/10.4007/annals.2018.187.2.2.

[45] E. Ullmo and A. Yafaev, A characterization of special subvarieties, *Mathematika* **57** no. 2 (2011), 263–273. MR 2825237. Zbl 1236.14029. https://doi.org/10.1112/S0025579311001628.

[46] U. Zannier, *Some Problems of Unlikely Intersections in Arithmetic and Geometry*, *Ann. of Math. Stud.* **181**, Princeton University Press, Princeton, NJ,

2012, With appendixes by David Masser. MR 2918151. Zbl 1246.14003. https://doi.org/10.23943/princeton/9780691153704.001.0001.

(Received: January 14, 2019)
(Revised: August 16, 2019)

Universität Basel, Basel, Switzerland
E-mail: David.Masser@unibas.ch

Scuola Normale Superiore, Pisa, Italy
E-mail: u.zannier@sns.it