Integral points of Shimura varieties: an “all or nothing" principle

Haohao LIU
email: kyung@mail.ustc.edu.cn

April 6, 2024

Abstract

We define the integral Lang locus for algebraic varieties over number fields. It measures the failure of finiteness of integral points of the algebraic variety. For Shimura varieties, Lang conjectures that the locus is empty when the level structure is high, and we prove that the locus is either full or empty.

Contents

1 Introduction

A complex analytic space X is called *Brody hyperbolic*, if every morphism $\mathbb{C} \to X$ is constant. For example, by [Cos05, p.78], a genus g compact Riemann surface is Brody hyperbolic if and only if $g \geq 2$. Conjecture 1.1 predicts that hyperbolicity (geometric property) restricts the behavior of rational points (arithmetic result).

Conjecture 1.1 (Lang, [Lan74, (1.3)], [Lan86, p.160])**.** *Let X be an integral projective variety over a number filed $F(\subset \mathbb{C})$. If the complex analytification $X(\mathbb{C})$ is Brody hyperbolic, then the set of rational points $X(F)$ is finite.*

Ullmo and Yafaev [UY10] study Conjecture 1.1 in the case of Shimura varieties. Let (G, X) be a Shimura datum (in the sense of [Mil17, Def. 5.5]).

Let $K \leq G(\mathbb{A}_f)$ be a compact open subgroup. For every connected component $S \subset \mathrm{Sh}_K(G, X)$, denote the Baily-Borel compactification of S be S^*. Fact 1.2 is derived from [Nad89, Thm. 0.2] in the paragraph following [UY10, Thm. 2.1].

Fact 1.2 (Nadel). *There is an open subgroup $K' \leq K$, such that for every induced finite étale cover $S' \to S$, the Baily-Borel compactification S'^* is Brody hyperbolic.*

For one thing, by Fact 1.2, shrinking K to a sufficiently small open subgroup, one may and will assume that the Shimura variety S is Brody hyperbolic. For another, S^* has a natural structure of projective variety over a number field $F(\subset \mathbb{C})$. Then Conjecture 1.1 predicts $S(F')$ to be finite for every finite extension F'/F. Ullmo and Yafaev [UY10] introduce "Lang locus" (Example 2.2) for algebraic varieties over $\bar{\mathbb{Q}}$ to measure the failure of Conjecture 1.1. In particular, the Lang locus of an algebraic variety over $\bar{\mathbb{Q}}$ is empty if and only if it has only finitely many rational points over every number field where it can be defined. The Lang locus of Shimura varieties satisfies an alternative principle.

Fact 1.3 ([UY10, Thm. 1.1]). *Let S be a Shimura variety of sufficiently high level. Then its Lang locus is either empty or full S.*

As Ullmo and Yafaev put it, Fact 1.3 means that for Shimura varieties, Conjecture 1.1 is either true or very false.

As Shimura varieties are not proper in general, it is natural to consider integral points. Conjecture 1.1 predicts that S has only finitely many integral points. We derive an analogue of Fact 1.3 for integral points. We define a notion of "integral Lang locus" (Definition 5.1) for algebraic varieties over $\bar{\mathbb{Q}}$ that measures the failure of finiteness of integral points.

Theorem (Theorem 5.12). *The integral Lang locus of a Shimura variety S is either empty or full S.*

Notation and conventions

Let $\bar{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Let $\mathbb{A}_f$ be the ring of finite adèles of $\mathbb{Q}$. Unless otherwise specified, an algebraic variety means a finite type, separated, geometrically reduced scheme over a field. The closure of a subset of an algebraic variety is taken in the Zariski topology. A subvariety is assumed to be Zariski closed. A Zariski-closed subset of a variety is endowed with the reduced induced closed subscheme structure, hence a subvariety.

By an *étale cover* $X \to Y$, we mean a finite étale morphism between integral algebraic varieties. In particular, it is surjective. If $\mathrm{Aut}(X/Y)$ acts transitively on each fiber, then $X \to Y$ is called a *Galois cover*, of Galois group $\mathrm{Aut}(X/Y)$. For a topological space X, we write $X^{>0}$ for the union of irreducible components of positive Krull dimension. Then for every subspace $Y \subset X$, one has $Y^{>0} \subset X^{>0}$.

Acknowledgment

I would like to express my thanks to my advisor Anna Cadoret, who suggested me to write these things down. My friends Long Liu, Yachen Liu, Hui Zhang and Ruishen Zhao have kindly answered me quite a few questions.

2 Locus formation

We shall show that an alternative principle (Corollary 4.3) for an abstract locus is a consequence of some axioms.

Suppose that for every integral algebraic variety X over $\bar{\mathbb{Q}}$, we define a subvariety $X^L \subset X$. For a reducible algebraic variety Z over $\bar{\mathbb{Q}}$, let $Z = \cup_{i=1}^n Z_i$ be the decomposition into irreducible components. Set $Z^L := \cup_{i=1}^n Z_i^L$. Suppose that the formation $(\cdot)^L$ satisfies Assumption 2.1.

Assumption 2.1. For any integral algebraic varieties X, Y over $\bar{\mathbb{Q}}$:

1. (Dimension) If $X^L \neq \emptyset$, then every irreducible component of X^L has positive dimension;
2. (Inheritance) For every closed immersion $i : X \to Y$ over $\bar{\mathbb{Q}}$, one has $i(X^L) \subset Y^L$;
3. (Cover) For every étale cover $f : X \to Y$ over $\bar{\mathbb{Q}}$, one has $f(X^L) \subset Y^L$;
4. (Iteration) One has $X^L \subset (X^L)^L$;
5. (Birational) For every finite birational morphism $f : X \to Y$ over $\bar{\mathbb{Q}}$, one has $f(X^L) \subset Y^L$.

For every integral algebraic variety X over $\bar{\mathbb{Q}}$, by Assumption 2.1 2, one has $X^L \supset (X^L)^L$. From Assumption 2.1 4, one has $X^L = (X^L)^L$.

Example 2.2. The Lang locus defined in [UY10, Sec. 2.2] satisfies Assumption 2.1. For every integral algebraic variety X over $\bar{\mathbb{Q}}$, by [Sta24, Tag 01ZM (1), Tag 01ZQ], there exist a number field F, an algebraic variety X_F over F and an isomorphism $X_F \otimes_F \bar{\mathbb{Q}} \to X$ over $\bar{\mathbb{Q}}$. For each finite subextension M/F, let $X(X_F, M)$ be the image of the natural injection[1] $X_F(M) \to X(\bar{\mathbb{Q}})$. The *Lang locus* of X relative to X_F is defined to be the Zariski closure of

$$\cup_M \overline{X(X_F, M)}^{>0}$$

in X, where M runs through all finite subextensions of F. By Lemma 2.3, the Lang locus depends only on X. From [UY10, Lemmes 2.3, 2.5], the Lang locus satisfies Assumption 2.1. It measures the failure of finiteness of rational points, since $X^L = \emptyset$ if and only if $X_F(M)$ is finite for every finite subextension M/F.

Lemma 2.3. *The Lang locus of X is independent of the choice of X_F.*

[1] The natural map $X_F(M) \to X_F$ is not injective in general.

Proof. Take another model $X_{F'}$ over a number field F'. There is a $\bar{\mathbb{Q}}$-isomorphism $X_F \otimes_F \bar{\mathbb{Q}} \to X_{F'} \otimes_{F'} \bar{\mathbb{Q}}$. Because $X_{F'}$ is separated, by [Gro65, Prop. 4.8.13], the morphism is defined over a number field F'' containing both F and F'. For every finite extension M/F, there is a number field M' containing M and F''. Then $X(X_F, M) \subset X(X_{F'}, M')$, so $\overline{X(X_F, M)}^{>0} \subset \overline{X(X_{F'}, M')}^{>0}$ and hence the Lang locus relative to X_F is contained in that relative to $X_{F'}$. The reverse inclusion follows by symmetry. □

Remark 2.4.

1. The Lang locus X^L in Example 2.2 is slightly different from the "lieu de Lang" X_F^L (a Zariski closed subset of X_F) defined by [UY10, (1)]. Let $\phi : X \to X_F$ be the natural morphism of schemes. For every finite extension M/F, let $X_F[M]$ be the image of the natural map $X_F(M) \to X_F$. Then $\phi(X(X_F, M)) = X_F[M]$. Because ϕ is integral, and surjective integral morphisms preserve the dimension, one has $\phi(\overline{X(X_F, M)}) = \overline{X_F[M]}$ and $\phi(\overline{X(X_F, M)}^{>0}) = \overline{X_F[M]}^{>0}$. Hence $\phi(X^L) = X_F^L$.

2. For a finite birational morphism $f : X \to Y$ of integral algebraic varieties over $\bar{\mathbb{Q}}$, it is not clear whether the Lang locus of Y is the image of the Lang locus of X (even if this is stated in [UY10, p.697]). That is why we require only inclusion but not equality in Assumption 2.1 5.

We gather some consequences of Assumption 2.1.

Lemma 2.5. *Let X be an algebraic variety over $\bar{\mathbb{Q}}$.*

1. *If $X = \cup_{i=1}^r Z_i$, where each Z_i is a subvariety of X, then $X^L = \cup_{i=1}^r Z_i^L$.*

2. *If Z is an irreducible component of X^L, then $Z^L = Z$.*

3. *If $f : X \to Y$ is Galois cover over $\bar{\mathbb{Q}}$, then $f^{-1}(f(X^L)) = X^L$. If $Z \subset Y$ is an irreducible subvariety, and Z' is an irreducible component of $f^{-1}(Z)$, then $f(f^{-1}(Z)^L) = f(Z'^L)$.*

Proof.

1. By Assumption 2.1 2, one has $\cup_{i=1}^r Z_i^L \subset X^L$. If Y is an irreducible component of X, then there exists an index i such that $Y \subset Z_i$. From Assumption 2.1 2, one has $Y^L \subset Z_i^L$ and hence $X^L \subset \cup_{i=1}^r Z_i^L$.

2. Write $X^L = \cup_{i=1}^n Z_i$ for the decomposition into irreducible components with $Z_1 = Z$. By Assumption 2.1 4, one has
$$Z \subset X^L = (X^L)^L = \cup_{i=1}^n Z_i^L.$$
As Z is irreducible, there is an index i such that $Z \subset Z_i^L \subset Z_i$. As $Z = Z_1$ is an irreducible component of X^L, one has $i = 1$ and $Z = Z^L$.

3. For every $x \in f^{-1}(f(X^L))$, there is $x' \in X^L$ with $f(x') = f(x)$. Let Θ be the Galois group of $f : X \to Y$. There is $\theta \in \Theta$ with $\theta(x') = x$, so $x \in X^L$ by Assumption 2.1 2. Therefore, $f^{-1}(f(X^L)) = X^L$. Since

Θ permutes transitively the irreducible components of $f^{-1}(Z)$, one has $f^{-1}(Z) = \Theta \cdot Z'$. By Part 1, one has $f^{-1}(Z)^L = \Theta \cdot Z'^L$ and hence $f(f^{-1}(Z)^L) = f(Z'^L)$.

□

Given an étale cover $f : X \to Y$ over $\bar{\mathbb{Q}}$, the induced morphism $X^L \to Y^L$ may not be surjective. We introduce a sublocus that lifts along all étale covers.

For an integral algebraic variety X over $\bar{\mathbb{Q}}$, define its *locus at infinite level* by

$$X^{L_\infty} := \cap_{f:T\to X} f(T^L),$$

where $f : T \to X$ runs through all étale covers of X. By Assumption 2.1 3, the sublocus X^{L_∞} is a subvariety of X^L. As X is topologically Noetherian, and $f(T^L) \subset X$ is closed for every such $f : T \to X$, there exists a particular cover $f_1 : X_1 \to X$ with $f_1(X_1^L) = X^{L_\infty}$. For every étale cover $X_2 \to X_1$, the composition $X_2^L \to X_1^L \xrightarrow{f_1} X^{L_\infty}$ is still surjective.

Remark 2.6. By Assumption 2.1 1, if $X^{L_\infty} \neq \emptyset$, then its irreducible components are *positive dimensional.*

For a reducible algebraic variety Y over $\bar{\mathbb{Q}}$, let $Y = \cup_{i=1}^n Y_i$ be the decomposition into irreducible components. Define $Y^{L_\infty} = \cup_{i=1}^n Y_i^{L_\infty}$, which is a subvariety of Y^L.

Lemma 2.7. *Let $f : T \to S$ be an étale cover over $\bar{\mathbb{Q}}$. Then $f^{-1}(S^{L_\infty}) = T^{L_\infty}$. In particular, $T^{L_\infty} = T$ is equivalent to $S^{L_\infty} = S$, and $S^{L_\infty} = S^L$ implies $T^{L_\infty} = T^L$.*

Proof. • We show $T^{L_\infty} \subset f^{-1}(S^{L_\infty})$.

Fix $t \in T^{L_\infty}$, and set $s = f(t)$. For every étale cover $g : S' \to S$, there is a commutative diagram

$$\begin{tikzcd} & T' \arrow[dl, "g'"'] \arrow[dr, "f'"] & \\ T \arrow[dr, "f"'] & & S' \arrow[dl, "g"] \\ & S & \end{tikzcd}$$

where each arrow is an étale cover. There is $t' \in T'^L$ with $g'(t') = t$. Then by Assumption 2.1 3, one has $s' := f'(t') \in S'^L$ and $s = g(s') \in g(S'^L)$. Hence $s \in S^{L_\infty}$.

- We show $T^{L_\infty} \supset f^{-1}(S^{L_\infty})$.

Take $t \in f^{-1}(S^{L_\infty})$. Then $s := f(t) \in S^{L_\infty}$. For every étale cover $u : Z \to T$, there is an étale cover $v : N \to Z$ such that the composition $N \xrightarrow{v} Z \xrightarrow{u} T \xrightarrow{f} S$

is a Galois cover. One has

$$\begin{aligned}(u\circ v)^{-1}(t) \subset (f\circ u\circ v)^{-1}(s) \subset (f\circ u\circ v)^{-1}(S^{L_\infty})\\ \subset (f\circ u\circ v)^{-1}((f\circ u\circ v)(N^L)) \overset{\text{(a)}}{=} N^L,\end{aligned}$$

where (a) uses Lemma 2.5 3. Thus, one has $u^{-1}(t) \subset v(N^L) \subset Z^L$ and $t \in u(Z^L)$. Hence $t \in T^{L_\infty}$.

- The equality $T^{L_\infty} = T$ is equivalent to $S^{L_\infty} = S$.

If $T^{L_\infty} = T$, then $S^{L_\infty} = f(f^{-1}(S^{L_\infty})) = f(T^{L_\infty}) = f(T) = S$. Conversely, if $S^{L_\infty} = S$, then $T^{L_\infty} = f^{-1}(S^{L_\infty}) = f^{-1}(S) = T$.

- The equality $S^{L_\infty} = S^L$ implies $T^{L_\infty} = f^{-1}(S^{L_\infty}) = f^{-1}(S^L) \overset{\text{(b)}}{\supset} T^L$, where (b) uses Assumption 2.1 3. Hence $T^{L_\infty} = T^L$.

□

3 Shimura varieties

We review some basic facts about Shimura varieties, the main objects of interest in this note. We use essentially results on the geometry of Hecke correspondences and special subvarieties from [UY10, UY14].

Basics

Let G be an affine algebraic group over $\mathbb{Q}$.

Definition 3.1 ([Pin90, Sec. 0.1, p.13])**.** For every prime number p, choose an embedding $\bar{\mathbb{Q}} \to \bar{\mathbb{Q}}_p$.

1. For an element $g = (g_p)_p \in \mathrm{GL}_n(\mathbb{A}_f)$, let $\Gamma_p \leq \bar{\mathbb{Q}}_p^\times$ be the subgroup generated by all eigenvalues of $g_p \in \mathrm{GL}_n(\mathbb{Q}_p)$. If the intersection of the torsion subgroups
$$\cap_p(\bar{\mathbb{Q}}^\times \cap \Gamma_p)_{\mathrm{tor}} = \{1\}$$
for p running through all primes, then g is called *neat*.

2. An element of $G(\mathbb{A}_f)$ is called neat if its image under some faithful algebraic representation of $G \to \mathrm{GL}_{n/\mathbb{Q}}$ is neat.

3. A subgroup of $G(\mathbb{A}_f)$ is called neat if all its elements are neat.

Fact 3.2 ([Pin90, p.13])**.**

1. *Let $K \leq G(\mathbb{A}_f)$ be a compact open subgroup. Then there is an open normal subgroup $K' \leq K$ that is neat.*

2. *Let $K \leq G(\mathbb{A}_f)$ be a neat subgroup. Then $K \cap G(\mathbb{Q})$ is a neat subgroup of* $G(\mathbb{Q})$ (in the sense of [Mil17, p.34]).

Let (G, X) be a Shimura datum. The set $G(\mathbb{R})$ is naturally a (real) Lie group. For a Lie group L, let L^+ be its identity component. Let G^{ad} be the quotient of G by its center. Set $G(\mathbb{R})_+$ to be the preimage of $G^{\mathrm{ad}}(\mathbb{R})^+$ under the natural morphism $G(\mathbb{R}) \to G^{\mathrm{ad}}(\mathbb{R})$ of Lie groups. Then $G(\mathbb{Q})^+ \subset G(\mathbb{Q})_+ \subset G(\mathbb{Q})$. By [Noo06, p.168] and [Mil17, Prop. 5.9], X is naturally a finite disjoint union of isomorphic hermitian symmetric domains. Let X^+ be a connected component of X. By [Mil17, Prop. 5.7 (b)], the stabilizer of X^+ in $G(\mathbb{Q})$ is $G(\mathbb{Q})_+ := G(\mathbb{Q}) \cap G(\mathbb{R})_+$.

Let $K \leq G(\mathbb{A}_f)$ be a compact open subgroup. From Fact 3.2 1, by passing to an open subgroup of K, we may and *always* assume that K is neat. Then by [Pin90, Prop. 3.3 (b)], $\mathrm{Sh}_K(G, X) := G(\mathbb{Q}) \backslash X \times G(\mathbb{A}_f)/K$ is naturally a complex manifold. For every $g \in G(\mathbb{A}_f)$, put $\Gamma_g := gKg^{-1} \cap G(\mathbb{Q})_+$ and $S_g := \Gamma_g \backslash X^+$. By Fact 3.2 2 and [Mil17, Prop. 4.1], Γ_g is a neat (hence torsion-free) arithmetic subgroup of $G(\mathbb{Q})$ (in the sense of [Mil17, p.33]). From [Mil17, Prop. 3.1], $S_g = [X^+, g]_K$ is naturally a connected complex manifold. Let C be a set of representatives for the double coset space $G(\mathbb{Q})_+ \backslash G(\mathbb{A}_f)/K$. From [Mil17, Lemmas 5.12 and 5.13], the set C is finite, and as complex manifold $\mathrm{Sh}_K(G, X) = \sqcup_{g \in C} S_g$.

By [Mil17, Thm. 3.12; Cor. 3.16], the complex manifold S_g has a canonical structure of a complex algebraic variety. The algebraic variety S_g is an irreducible, *smooth* arithmetic locally symmetric variety ([Mil17, p.58]). It is Zariski-open in its Baily-Borel compactification S_g^* ([Mil17, p.40]), which is a projective variety. Thus, $\mathrm{Sh}_K(G, X)$ is also a smooth quasi-projective (reducible) complex algebraic variety.

Let $E(G, X) \subset \bar{\mathbb{Q}}$ be the reflex field of the Shimura datum (G, X) (in the sense of [Mil17, Def. 12.2]). By [Mil17, Rk. 12.3 (a)], $E(G, X)$ is a number field. From [Mil99, Rk. 2.4 (b)] and [Mil17, p.128], $\mathrm{Sh}_K(G, X)$ admits a unique (up to a unique isomorphism) *canonical model* over $E(G, X)$ (in the sense of [Mil17, Def. 12.8]). Hence a *smooth* quasi-projective variety $\mathrm{Sh}_K(G, X)$ over $E(G, X)$. By [Del71, Cor. 5.4], for every morphism of Shimura data $f : (G', X') \to (G, X)$ and every compact open subgroup $K \leq G(\mathbb{A}_f)$ containing $f(K')$, the induced morphism $\mathrm{Sh}_{K'}(G', X') \to \mathrm{Sh}_K(G, X)$ is defined over a number field. Assume that f is the identity. Then the induced morphism (denoted by $p_{K',K}$) is finite étale and defined over $E(G, X)$. For every irreducible component $S' \subset \mathrm{Sh}_{K'}(G', X')$, its image $S \subset \mathrm{Sh}_K(G, X)$ is an irreducible component, and the restriction $S' \to S$ is an étale cover defined over a finite extension of $E(G, X)$. From [CK16, p.1901], when K' is normal in K, this étale cover is Galois.

By [Moo98b, p.282] and [GN20, Remark (3), p.56], every connected component $S \subset \mathrm{Sh}_K(G, X)$ and its inclusion $S \to S^*$ to the Baily-Borel compactification are defined over a finite abelian extension of $E(G, X)$. Such an S is called a Shimura variety[2] associated with (G, X, K).

[2]By [Moo98b, Prop. 2.9] and [Mil17, Thm. 5.17], when K is sufficiently small, S is a

Hecke correspondences

By [Mil17, Thm. 13.6], for every $g \in G(\mathbb{A}_f)$, there is an isomorphism $T(g) : \mathrm{Sh}_K(G,X) \to \mathrm{Sh}_{g^{-1}Kg}(G,X)$ of algebraic varieties over $E(G,X)$. For every $h \in G(\mathbb{A}_f)$, the morphism $T(g)$ sends the connected component $[X^+, h]_K \subset \mathrm{Sh}_K(G,X)$ isomorphically to $[X^+, hg]_{g^{-1}Kg} \subset \mathrm{Sh}_{g^{-1}Kg}(G,X)$. The algebraic correspondence

$$\mathrm{Sh}_K(G,X) \overset{p_{K\cap gKg^{-1},K}}{\leftarrow} \mathrm{Sh}_{K\cap gKg^{-1}}(G,X) \overset{p_{K\cap gKg^{-1},gKg^{-1}}}{\rightarrow} \mathrm{Sh}_{gKg^{-1}}(G,X) \overset{T(g)}{\rightarrow} \mathrm{Sh}_K(G,X)$$

over $E(G,X)$ is denoted by $T_g^{\mathbb{A}}$, and called the *adelic Hecke correspondence* induced by g.

Let $S = (K \cap G(\mathbb{Q})_+)\backslash X^+$. For every $q \in G(\mathbb{Q})_+$, let $S_q = (K \cap q^{-1}Kq \cap G(\mathbb{Q})_+)\backslash X^+$. Then S_q is the connected component $[X^+, 1]$ of $\mathrm{Sh}_{K\cap q^{-1}Kq}(G,X)$ (resp. $\mathrm{Sh}_K(G,X)$). The map Id_{X^+} (resp. $X^+ \to X^+, \quad x \mapsto q\cdot x$) induces an étale cover $\alpha_q : S_q \to S$ (resp. $\beta_q : S_q \to S$). There is a commutative diagram

$$\begin{tikzcd}
S & S_q \arrow[l, "\alpha_q"'] \arrow[r, "\beta_q"] & S \\
\mathrm{Sh}_K(G,X) \arrow[from=u, hook] & \mathrm{Sh}_{K\cap q^{-1}Kq}(G,X) \arrow[from=u, hook] \arrow[l, "p_{K\cap q^{-1}Kq,K}"] \arrow[r, "T(q^{-1})p_{K\cap q^{-1}Kq,q^{-1}Kq}"'] & \mathrm{Sh}_K(G,X) \arrow[from=u, hook]
\end{tikzcd}$$

of complex manifolds. Therefore, the correspondence

$$S \overset{\alpha_q}{\leftarrow} S_q \overset{\beta_q}{\rightarrow} S$$

is algebraic and defined over a number field. It is called the *(rational) Hecke correspondence* induced by q, and denoted by T_q.

Let $\{q_i\}_{i=1}^n$ be elements of $G(\mathbb{Q})_+ \cap KgK$ satisfying

$$G(\mathbb{Q})_+ \cap KgK = \sqcup_{i=1}^n \Gamma q_i^{-1}\Gamma, \Gamma := K \cap G(\mathbb{Q})_+.$$

By [KY14, p.881], the correspondence on $[X^+, 1] \subset \mathrm{Sh}_K(G,X)$ induced by $T_g^{\mathbb{A}}$ decomposes as $\sum_{i=1}^n T_{q_i}$. For instance, the correspondences $T_1^{\mathbb{A}}$ and T_1 are the identity.

Special subvarieties

Definition 3.3. [Moo98a, Def. 2.5] An irreducible subvariety $Z \subset \mathrm{Sh}_K(G,X)$ over $\mathbb{C}$ is called special, if there exists a connected, reductive algebraic subgroup $H \leq G$ defined over $\mathbb{Q}$, an element $g \in G(\mathbb{A}_f)$ and a connected component D_H^+ of

$$D_H := \{x \in X | h_x : \mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_m) \to G_{\mathbb{R}} \text{ factors through } H_{\mathbb{R}}\},$$

such that $Z(\mathbb{C})$ is the image of $D_H^+ \times gK$ in $\mathrm{Sh}_K(G,X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K$.

connected Shimura variety (in the sense of [Mil17, Def. 4.10]).

By [Moo98a, 2.4], D_H is a finite union of $H(\mathbb{R})$-conjugacy classes. Let C be the $H(\mathbb{R})$-conjugacy class containing D_H^+. Then (H, C) is a Shimura subdatum[3] of (G, X). Then from [Del71, Cor. 5.4] and [Moo98a, Rk. 2.6], every special subvariety of $\mathrm{Sh}_K(G, X)$ is defined over a number field.

Example 3.4. 1. A complex point $s \in \mathrm{Sh}_K(G, X)$ is a special subvariety, if and only if there is a special point $x \in X$ (in the sense of [Mil17, Def. 12.5]) and $g \in G(\mathbb{A}_f)$ with $s = [x, g]_K$.

2. When $H = G$, the corresponding special subvarieties of $\mathrm{Sh}_K(G, X)$ are precisely the connected components.

For every $g \in G(\mathbb{A}_f)$ and every irreducible subvariety $Z \subset \mathrm{Sh}_K(G, X)$ over $\mathbb{C}$, Z is special if and only if $T(g)(Z)$ is special in $\mathrm{Sh}_{g^{-1}Kg}(G, X)$. By [Moo98a, Sec. 2.9], an irreducible component of the intersection of a family of special subvarieties of $\mathrm{Sh}_K(G, X)$ over $\mathbb{C}$ is again special. Therefore, for a complex, irreducible subvariety $Y \subset \mathrm{Sh}_K(G, X)$, there is a smallest special subvariety $Z_Y \subset \mathrm{Sh}_K(G, X)$ containing Y. We say that Y is *Hodge generic* in Z_Y. Let $\mathbb{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m$ be the Deligne torus.

Definition 3.5. The generic Mumford-Tate group (denoted by $\mathrm{MT}(X)$) of the Shimura datum (G, X) is the smallest closed subgroup H of G over $\mathbb{Q}$, such that every $h : \mathbb{S} \to G_{\mathbb{R}}$ in X factors through $H_{\mathbb{R}}$. If $\mathrm{MT}(X) = G$, then the Shimura datum (G, X) is called irreducible.

The subgroup $\mathrm{MT}(X) \leq G$ is normal, connected and reductive. By [Che09, Def. 1.3.3], $(\mathrm{MT}(X), X)$ is a Shimura subdatum of (G, X). Fact 3.6 characterizes special subvarieties as Hecke image of irreducible components of a Shimura subvariety. Recall that $K \leq G(\mathbb{A}_f)$ is a *neat*, compact open subgroup. For $g \in G(\mathbb{A}_f)$, the quotient $S_g = \Gamma_g \backslash X^+$ is an irreducible component of $\mathrm{Sh}_K(G, X)$.

Fact 3.6 ([UY10, Lem. 2.7])**.** *Let $(H, X_H) \subset (G, X)$ be an* irreducible *Shimura subdatum. Let X_H^+ be a connected component of X_H contained in X^+. Set $\Gamma_{H,g} = gKg^{-1} \cap H(\mathbb{Q})_+$ and $\tilde{Z}_g := \Gamma_{H,g} \backslash X_H^+$ (an irreducible component of $\mathrm{Sh}_{gKg^{-1} \cap H(\mathbb{A}_f)}(H, X_H)$). Then the image Z_g of $\tilde{Z}_g$ under the $\mathbb{C}$-morphism*

$$\mathrm{Sh}_{gKg^{-1} \cap H(\mathbb{A}_f)}(H, X_H) \to \mathrm{Sh}_K(G, X), \quad [x, h] \mapsto [x, hg]$$

is a special subvariety of S_g. The induced morphism $\pi : \tilde{Z}_g \to Z_g$ is finite and birational. Conversely, every special subvariety of S_g arises in this way.

Remark 3.7. If the special subvariety Z_g in Fact 3.6 is normal, then by Zariski's main theorem (see, *e.g.*, [Liu06, Cor. 4.6]), $\pi : \tilde{Z}_g \to Z_g$ is an isomorphism.

Let $S = S_g$ be a Shimura variety associated with (G, X, K).

Lemma 3.8. *Let $Z \subset S$ be a special subvariety over $\bar{\mathbb{Q}}$. Let $\pi : \tilde{Z} \to Z$ be a finite birational morphism given by Fact 3.6. Then there is a Galois cover*

[3] in the sense of [CLZ16, p.894]

$f : S' \to S$ over $\bar{\mathbb{Q}}$ with $S'^L = S'^{L_\infty}$, such that for every irreducible component $Z' \subset f^{-1}(Z)$, one has $Z'^L = Z'^{L_\infty}$ and $f : Z' \to Z$ factors through an étale cover $Z' \to \tilde{Z}$.

Proof. The Hecke isomorphism $T(g) : \mathrm{Sh}_{gKg^{-1}}(G,X) \to \mathrm{Sh}_K(G,X)$ sends $[X^+,1]_{gKg^{-1}}$ to S_g. It keeps the special subvarieties. Thus, one may assume $g = 1$ (by replacing K with gKg^{-1}). Let $(H, X_H) \subset (G,X)$ be an irreducible Shimura subdatum inducing Z *via* Fact 3.6. Then the restriction

$$\pi : \tilde{Z}(:= [X_H^+,1]_{K\cap H(\mathbb{A}_f)}) \to Z$$

of $\mathrm{Sh}_{K\cap H(\mathbb{A}_f)}(H,X_H) \to \mathrm{Sh}_K(G,X)$ is finite and birational.

The system $\{[X_H^+,1]_U\}$ (U running through all open subgroups of $K\cap H(\mathbb{A}_f)$) is cofinal in all the étale covers of $\tilde{Z}$. So there is an open subgroup $K_{0,H} \leq K$ such that the étale cover $g_0 : \tilde{Z}_0(:= [X_H^+,1]_{K_{0,H}}) \to \tilde{Z}$ satisfies $g_0(\tilde{Z}_0^L) = \tilde{Z}^{L_\infty}$. Similarly, there is an open subgroup $K_1 \leq K$ such that $K_1 \cap H(\mathbb{A}_f) \subset K_{0,H}$ and the étale cover $f_1 : S_1 := ([X^+,1]_{K_1}) \to S$ satisfies $f_1(S_1^L) = S^{L_\infty}$. By Lemma 2.7, one has $S_1^L = S_1^{L_\infty}$.

- There is an open subgroup $K_2 \leq K_1$, such that $K_2 \cap H(\mathbb{A}_f) = K_1 \cap H(\mathbb{A}_f)$ and the natural morphism $i_1 : \mathrm{Sh}_{K_2\cap H(\mathbb{A}_f)}(H,X_H) \to \mathrm{Sh}_{K_2}(G,X)$ is a closed immersion.

Indeed, by [Del71, Prop. 1.15], there is a compact open subgroup $K_m \leq G(\mathbb{A}_f)$ containing $K_1 \cap H(\mathbb{A}_f)$, such that the morphism

$$i_2 : \mathrm{Sh}_{K\cap H(\mathbb{A}_f)}(H,X_H) \to \mathrm{Sh}_{K_m}(G,X)$$

is a closed immersion. Let $K_2 = K_1 \cap K_m$. Then $K_2 \cap H(\mathbb{A}_f) = K_1 \cap H(\mathbb{A}_f)$ and $i_2 = p_{K_2,K_m} i_1$. Since $p_{K_2,K_m} : \mathrm{Sh}_{K_2}(G,X) \to \mathrm{Sh}_{K_m}(G,X)$ is separated, by magic square, i_1 is a closed immersion.

Then the morphism $\tilde{Z}_2(:= [X_H^+,1]_{K_2\cap H}) \to S_2(:= [X^+,1]_{K_2})$ is a closed immersion. The induced morphism $\pi_2 : \tilde{Z}_2 \to (f_1 f_2)^{-1}(Z)$ is a closed immersion.

- The closed immersion π_2 identifies $\tilde{Z}_2$ with an irreducible component of $(f_1 f_2)^{-1}(Z)$.

Since $\tilde{Z}_2$ is irreducible, it is contained in an irreducible component $C \subset (f_1 f_2)^{-1}(Z)$. As π is birational, by [Sta24, Tag 0BAC], there is a nonempty open subset $\tilde{U} \subset \tilde{Z}$, such that $U := \pi(\tilde{U})$ is open in Z and $\pi|_{\tilde{U}} : \tilde{U} \to U$ is an isomorphism. Consider the commutative square

$$\begin{array}{ccc} g_2^{-1}(U) & \longrightarrow & (f_1 f_2)^{-1}(U) \\ \downarrow & & \downarrow \\ \tilde{U} & \xrightarrow{\pi|_{\tilde{U}}} & U. \end{array}$$

The morphism $g_2^{-1}(\tilde{U}) \to (f_1f_2)^{-1}(U) \xrightarrow{f_1f_2} U$ (resp. $f_1f_2 : (f_1f_2)^{-1}(U) \to U$) is a base change of the étale morphism $g_2 : \tilde{Z}_2 \to Z_2$ (resp. $f_1f_2 : S_2 \to S$), so it is étale. By [Sta24, Tag 03PC (10)], the morphism $\pi_2|_{g_2^{-1}(\tilde{U})} : g_2^{-1}(\tilde{U}) \to (f_1f_2)^{-1}(U)$ is étale. From [Sta24, Tag 03PC (9)], $g_2^{-1}(\tilde{U})$ is an open subset of $(f_1f_2)^{-1}(U)$, hence a nonempty open subset of C. Since C is irreducible, $g_2^{-1}(\tilde{U})$ is dense in C. Therefore, $C \subset \tilde{Z}_2$.

There is a normal, open subgroup $K' \leq K$ with $K' \subset K_2$. Let $f_3 : S'(:= [X^+, 1]_{K'}) \to S_2$ be the induced étale cover. Since K' is normal in K, the composition $f(= f_1f_2f_3) : S' \to S$ is a Galois cover. Since $S_1^L = S_1^{L_\infty}$, by Lemma 2.7, one has $S'^L = S'^{L_\infty}$.

Let $\tilde{Z}_3$ be an irreducible component of $f_3^{-1}(\tilde{Z}_2)$. The morphism $f_3 : f_3^{-1}(\tilde{Z}_2) \to \tilde{Z}_2$ is a base change of the étale cover $f_3 : S_3 \to S_2$, so it is finite and étale. The algebraic variety $\tilde{Z}_2$ is smooth, so is $f_3^{-1}(\tilde{Z}_2)$. Therefore, $\tilde{Z}_3$ is open in $f_3^{-1}(\tilde{Z}_2)$. The morphism $g_3 : \tilde{Z}_3 \to \tilde{Z}_2$ is finite étale, and $\tilde{Z}_2$ is connected, so g_3 is surjective. The situation is depicted as a commutative diagram

$$\begin{tikzcd}
\tilde{Z}_3 \arrow[r, hook] \arrow[dr, "g_3"'] & f_3^{-1}(\tilde{Z}_2) \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\square"] & f^{-1}(Z) \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\square"] & S' \arrow[d, "f_3"] \arrow[ddd, bend left, "f"] \\
& \tilde{Z}_2 \arrow[r, hook, "\pi_2"'] \arrow[d] \arrow[dd, bend right, "g_2"'] & (f_1f_2)^{-1}(Z) \arrow[r, hook] \arrow[dd] \arrow[ddr, phantom, "\square"] & S_2 \arrow[d, "f_2"] \\
& \tilde{Z}_0 \arrow[d, "g_0"] & & S_1 \arrow[d, "f_1"] \\
& \tilde{Z} \arrow[r, "\pi"'] & Z \arrow[r, hook] & S.
\end{tikzcd}$$

Then $\tilde{Z}_3$ is an étale cover of $\tilde{Z}$ and an irreducible component of $f^{-1}(Z)$. The Galois group of the Galois cover $f : S' \to S$ permutes the irreducible components of $f^{-1}(Z)$, so they have similar properties. $\square$

Lemma 3.9 is used in the proof of Theorem 4.1.

Lemma 3.9. *If $S^{L_\infty} \neq \emptyset$ is a finite union of special subvarieties of S, then $S^L = S$.*

Proof. Write $S^{L_\infty} = \cup_{i=1}^n Z_i$ for the decomposition into irreducible components. By assumption, for every $1 \leq i \leq n$, the subvariety $Z_i \subset S$ is special. Let $\pi_i : \tilde{Z}_i \to Z_i$ be a finite birational morphism given by Fact 3.6. Let $f_i : S_i \to S$ be a Galois cover corresponding to π_i given by Lemma 3.8. There is a Galois cover $f : S' \to S$, such that for every $1 \leq i \leq n$, there is an étale cover $g_i : S' \to S_i$ with $f_ig_i = f$. Then $S'^L = S'^{L_\infty}$. Hence $S'^L = (S'^L)^L = (S'^{L_\infty})^L$.

1. One has $S^{L_\infty} \subset \cup_{i=1}^n \pi_i(\tilde{Z}_i^{L_\infty})$.

Indeed, one has $f(S'^L) = S^{L_\infty}$. For every irreducible component $C \subset S'^L$, the subset $f(C)$ of S^{L_∞} is irreducible. Then there is $1 \leq i \leq n$ with $f(C) \subset$

Z_i. Thus, $g_i(C)$ is an irreducible subset of $f_i^{-1}(Z_i)$. There is an irreducible component $Z' \subset f_i^{-1}(Z_i)$ containing $g_i(C)$. By Lemma 3.8, the morphism $f_i : Z' \to Z_i$ factors through an étale cover $Z' \to \tilde{Z}_i$. Therefore, Z' and $g_i^{-1}(Z')$ are smooth. One has

$$g_i^{-1}(Z') \subset f^{-1}(Z_i) \subset f^{-1}(S^{L_\infty}) \overset{\text{(a)}}{=} S'^L,$$

where (a) uses Lemma 2.7. Then C is an irreducible component of $g_i^{-1}(Z')$, hence an étale cover of Z'. One has $f(C^L) \subset \pi_i(\tilde{Z}_i^{L_\infty})$. Thus, 1 is proved.

2. One has $\tilde{Z}_1^{L_\infty} = \tilde{Z}_1$.

From 1, one has $Z_1 \subset \cup_{i=1}^n \pi_i(\tilde{Z}_i^{L_\infty})$. Since Z_1 is irreducible, there is $1 \le j \le n$ with $Z_1 \subset \pi_j(\tilde{Z}_j^{L_\infty}) \subset Z_j$. As Z_1 is an irreducible component of S^{L_∞}, one has $j = 1$. Then $\dim \tilde{Z}_1^{L_\infty} \ge \dim Z_1 = \dim \tilde{Z}_1$. The irreducibility of $\tilde{Z}_1$ proves 2.

3. For every $q \in G(\mathbb{Q})_+$, one has $T_q Z_1 \subset S^L$.

Let $(H, X_H) \subset (G, X)$ be an irreducible Shimura subdatum inducing Z_1 *via* Fact 3.6. Then $\tilde{Z}_1 = [X_H^+, 1]_{K \cap H(\mathbb{A}_f)}$. For every irreducible component $Z_q \subset \alpha_q^{-1}(Z_1)$, there is an irreducible component $\tilde{Z}_q$ of $\mathrm{Sh}_{K \cap q^{-1}Kq \cap H(\mathbb{A}_f)}(H, X_H)$ with the following properties:

- The morphism $\mathrm{Sh}_{K\cap q^{-1}Kq\cap H(\mathbb{A}_f)}(H, X_H) \to \mathrm{Sh}_{K\cap H(\mathbb{A}_f)}(H, X_H)$ restricts to an étale cover $\alpha'_q : \tilde{Z}_q \to \tilde{Z}_1$.
- The image of $\tilde{Z}_q$ under the morphism $\mathrm{Sh}_{K\cap q^{-1}Kq\cap H(\mathbb{A}_f)}(H, X_H) \to \mathrm{Sh}_{K\cap q^{-1}Kq}(G, X)$ is Z_q.

Conjugating by q gives another irreducible Shimura subdatum $(qHq^{-1}, q\cdot X_H) \subset (G, X)$, and a morphism of Shimura data $(H, X_H) \to (qHq^{-1}, q \cdot X_H)$. Let $\tilde{Z}'_q$ be the image of $\tilde{Z}_q$ under the induced morphism $\mathrm{Sh}_{K\cap q^{-1}Kq\cap H(\mathbb{A}_f)}(H, X_H) \to \mathrm{Sh}_{K\cap qH(\mathbb{A}_f)q^{-1}}(qHq^{-1}, q \cdot X_H)$. Then $\tilde{Z}'_q$ is an irreducible component of

$$\mathrm{Sh}_{K\cap qH(\mathbb{A}_f)q^{-1}}(qHq^{-1}, q \cdot X_H),$$

and the restriction $\beta'_q : \tilde{Z}_q \to \tilde{Z}'_q$ is an étale cover. By Fact 3.6, the morphism $\mathrm{Sh}_{K\cap qH(\mathbb{A}_f)q^{-1}}(qHq^{-1}, q \cdot X_H) \to \mathrm{Sh}_K(G, X)$ restricts to a finite birational morphism $\pi'_q : \tilde{Z}'_q \to \beta_q(Z_q)$. Consider the commutative diagram

$$\begin{tikzcd}
 & \tilde{Z}_q \arrow[dl, "\alpha'_q"'] \arrow[d] \arrow[dr, "\beta'_q"] & \\
\tilde{Z}_q \arrow[d, "\pi_1"] & Z_q \arrow[dl, "\alpha_q"] \arrow[dr, "\beta_q"] & \tilde{Z}'_q \arrow[d, "\pi'_q"] \\
Z_1 & & \beta_q(Z_q).
\end{tikzcd}$$

From 2 and Lemma 2.7, one has $\tilde{Z}'_q = \tilde{Z}'^{L_\infty}_q = \tilde{Z}'^{L}_q$. Then $\beta_q(Z_q) = \pi'_q(\tilde{Z}'_q) = \pi'_q(\tilde{Z}'^{L}_q) \overset{(a)}{\subset} \beta_q(Z_q)^L$, where (a) uses Assumption 2.1 5. By Assumption 2.1 2, one has $\beta_q(\tilde{Z}_q) \subset S^L$. Thus, 3 is proved.

Since Z_1 is a special subvariety of S, from [KUY18, Lem. 2.5], Z_1 contains a special point z. By [LZ19, Rk. 2.7], $\{T_q z\}_{q\in G(\mathbb{Q})_+}$ is dense in the complex manifold $S(\mathbb{C})$. By 3, the Zariski closed subset $S^L \subset S$ contains $\{T_q z\}_{q\in G(\mathbb{Q})_+}$. Hence $S^L = S$. □

4 Ullmo-Yafaev alternative principle

In Theorem 4.1, we show that an alternative principle results from Assumption 2.1. Let $S = S_g = (gKg^{-1} \cap G(\mathbb{Q})_+)\backslash X^+$ be a Shimura variety associated with (G, X, K).

Theorem 4.1 (Ullmo-Yafaev alternative)**.** *Either $S^{L_\infty} = \emptyset$ or $S^{L_\infty} = S$.*

Proof. By Hecke isomorphisms, one may assume $g = 1$ and $S = [X^+, 1]_K$. By Lemma 2.7, one may replace S by an étale cover induced by an open subgroup of K. One may thereby assume $S^L = S^{L_\infty} \neq \emptyset$. For every irreducible component $Z \subset S^L$, by Assumption 2.1 1 (resp. Lemma 2.5 2), one has $\dim(Z) > 0$ (resp. $Z^L = Z$).

1. The subvariety $Z \subset S$ is special.

Let $S_M \subset S$ be the smallest special subvariety containing Z. From Fact 3.6, there is a Shimura subdatum $(H, X_H) \subset (G, X)$, such that the restriction $\pi : \tilde{S}_M := [X^+_H, 1]_{K\cap H(\mathbb{A}_f)} \to S_M$ of $\mathrm{Sh}_{K\cap H(\mathbb{A}_f)}(H, X_H) \to \mathrm{Sh}_K(G, X)$ is finite birational.

Take a Galois cover $f : S' \to S$ given by Lemma 3.8 for the special subvariety $S_M \subset S$. Since f is finite surjective, there is an irreducible component $T \subset f^{-1}(Z)$ with $f(T) = Z$.

Since $Z \subset S^L$ is an irreducible component, T is an irreducible component of

$$f^{-1}(S^L) = f^{-1}(S^{L_\infty}) \overset{(a)}{=} S'^{L_\infty} \overset{(b)}{=} S'^L.$$

Here (a) and (b) use Lemma 2.7. Then by Lemma 2.5 2, one has $T^L = T$. There is an irreducible component $S'_M \subset f^{-1}(S_M)$ containing T.

By Lemma 3.8, one has $S'^L_M \overset{(c)}{=} S'^{L_\infty}_M$, and $f : S'_M \to S_M$ factors through an étale cover

$$g : S'_M \to \tilde{S}_M.$$

2. One has $g(T) \subset \tilde{S}^{L_\infty}_M$.

Consider the commutative diagram

$$
\begin{tikzcd}
T \arrow[r, hook] & S'_M \arrow[r, hook] \arrow[d, dashed, "g"] & f^{-1}(S_M) \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\square"] & S' \arrow[d, "f"] \\
& \tilde{S}_M \arrow[r, "\pi"] & S_M \arrow[r, hook] & S.
\end{tikzcd}
$$

One has

$$T = T^L \subset S'^L_M = S'^{L_\infty}_M.$$

Hence $g(T) \subset g(S'^{L_\infty}_M) = \tilde{S}^{L_\infty}_M$. Thus, 2 is proved.

3. The nonempty, irreducible, closed subset $g(T) \subset \tilde{S}^{L_\infty}_M$ is Hodge generic in $\tilde{S}_M$.

Since π is finite surjective, there is an irreducible component $\tilde{Z} \subset \pi^{-1}(Z)$ with $\pi(\tilde{Z}) = Z$. For every special subvariety $V \subset \tilde{S}_M$ containing $g(T)$, by [KY14, p.879], $\pi(V) \subset S$ is a special subvariety containing $\pi g(T) = f(T) = Z$. Hence $\pi(V) = S_M$. Therefore, $\dim V \geq \dim S_M = \dim \tilde{S}_M$. Since $\tilde{S}_M$ is irreducible, one has $V = \tilde{S}_M$. Thus, 3 is proved.

By 2, 3 and Lemma 4.2, one has $\tilde{S}_M = \tilde{S}^{L_\infty}_M = \tilde{S}^L_M$. One has $S_M = \pi(\tilde{S}_M) = \pi(\tilde{S}^L_M) \overset{(a)}{\subset} S^L_M \subset S^L$, where (a) uses Assumption 2.1 5. Since Z is an irreducible component of S^L and S_M is irreducible, one has $Z = S_M$. Thus, 1 is proved.

By 1, the locus S^L is a finite union of special subvarieties. From Lemma 3.9, one has $S^{L_\infty} = S$. $\square$

Lemma 4.2 (Ullmo-Yafaev)**.** *Let $S = [X^+, 1]_K \subset \mathrm{Sh}_K(G, X)$. If S^{L_∞} contains a nonempty, irreducible closed subset that is Hodge generic in S, then $S^{L_\infty} = S$.*

Proof. For every $q \in G(\mathbb{Q})^+$, by Lemma 2.7, one has $T_q S^{L_\infty} = \beta_q(\alpha_q^{-1}(S^{L_\infty})) = \beta_q(S_q^{L_\infty}) = S^{L_\infty}$. Write $S^{L_\infty} = U_1 \cup U_2$, where U_1 is the union of irreducible components of S^{L_∞} that are Hodge generic in S, and U_2 is the union of the remaining irreducible components. By Remark 2.6 and assumption, one has $\dim U_1 > 0$.

Let C be an irreducible component of $T_q U_2$. Then there is an irreducible subvariety $C_q \subset S_q$ with $\beta_q(C_q) = C$ and $\alpha_q(C_q) \subset U_2$. Then $\alpha_q(C_q)$ is not Hodge generic in S. Thus, there is a strict, special subvariety $V \subset S$ containing $\alpha_q(C_q)$. Then $C \subset T_q(\alpha_q(C_q)) \subset T_q V$. There is an irreducible component $W \subset T_q V$ containing C. By [LZ19, Remark 2.7], W is a special subvariety of S. Since $\dim W \leq \dim V < \dim S$, the subvariety $C \subset S$ is not Hodge generic. As every irreducible component of $T_q U_2$ is not Hodge generic in S, and $U_1 \subset T_q S^{L_\infty} = T_q U_1 \cup T_q U_2$, one has $U_1 \subset T_q U_1$. By $\dim U_1 > 0$ and [UY10, Thm. 1.2], one has $U_1 = S$ and $S^{L_\infty} = S$. $\square$

Corollary 4.3 ([UY10, Thm. 1.1])**.** *If a Shimura variety S over $\bar{\mathbb{Q}}$ is of sufficiently high level, then either $S^L = \emptyset$ or $S^L = S$.*

Proof. As the level is high, one has $S^L = S^{L_\infty}$. The result follows from Theorem 4.1. $\square$

5 “All or nothing" principle for integral points

We define an locus concerning integral points, analogous to the Lang locus concerning rational points. We verify Assumption 2.1 for this locus. Then an alternative principle follows.

Let X be an integral algebraic variety over $\bar{\mathbb{Q}}$. As in Example 2.2, there is a number field $F \subset \bar{\mathbb{Q}}$, an integral algebraic variety X_F over F and an isomorphism $X_F \otimes_F \bar{\mathbb{Q}} \to X$ over $\bar{\mathbb{Q}}$. For every finite set Σ of places of F including all archimedean ones, let $O_{F,\Sigma}$ be the ring of Σ-integers. When Σ is sufficiently large, there exists an integral scheme $\mathcal{X}$ that is finite type and separated over $O_{F,\Sigma}$, whose generic fiber is X_F. (From [Har77, III, Prop. 9.7], $\mathcal{X}$ is flat over $O_{F,\Sigma}$.) We call $\mathcal{X}$ an integral model for X relative to (F, Σ). By a finite extension $(M, \Omega)/(F, \Sigma)$, we mean a finite extension M/F together with a finite set Ω of places of M containing all the places above Σ.

For every $(M, \Omega)/(F, \Sigma)$, let $X(\mathcal{X}, M, \Omega)$ be the image of the injection

$$\mathcal{X}(O_{M,\Omega}) \to X(\bar{\mathbb{Q}}), \quad x \mapsto x|_{\operatorname{Spec} \bar{\mathbb{Q}}}.$$

Definition 5.1. Let $\mathcal{X}^I$ be the Zariski closure of

$$\cup_{(M,\Omega)/(F,\Sigma)} \overline{X(\mathcal{X}, M, \Omega)}^{>0}$$

inside X, where (M, Ω) runs though all finite extensions of (F, Σ). We call $\mathcal{X}^I$ the integral Lang locus of X relative to $(\mathcal{X}, F, \Sigma)$.

The integral Lang locus $\mathcal{X}^I$ is a subvariety of the Lang locus of X.

Lemma 5.2. *Given models $\mathcal{X}_i$ over O_{F_i,Σ_i} $(i = 1, 2)$ for X, one has $\mathcal{X}_1^I = \mathcal{X}_2^I$.*

Proof. By [Gro66, Cor. 8.8.2.5], there is a common finite extension (F_3, Σ_3) of (F_i, Σ_i) $(i = 1, 2)$, such that there is an O_{F_3,Σ_3}-isomorphism

$$\mathcal{X}_1 \otimes_{O_{F_1,\Sigma_1}} O_{F_3,\Sigma_3} \to \mathcal{X}_2 \otimes_{O_{F_2,\Sigma_2}} O_{F_3,\Sigma_3}$$

extending the isomorphism between the generic fibers. For every finite extension $(M_1, \Omega_1)/(F_1, \Sigma_1)$, there is a common finite extension (M_2, Ω_2) of (F_3, Σ_3) and (M_1, Ω_1). Then

$$\mathcal{X}_1(O_{M_1,\Omega_1}) \subset \mathcal{X}_1(O_{M_2,\Omega_2}) = \mathcal{X}_2(O_{M_2,\Omega_2}),$$

so $X(\mathcal{X}_1, M_1, \Omega_1) \subset X(\mathcal{X}_2, M_2, \Omega_2)$. Therefore,

$$\overline{X(\mathcal{X}_1, M_1, \Omega_1)}^{>0} \subset \overline{X(\mathcal{X}_2, M_2, \Omega_2)}^{>0} \subset \mathcal{X}_2^I.$$

Hence $\mathcal{X}_1^I \subset \mathcal{X}_2^I$. The other inclusion follows by symmetry. $\square$

By Lemma 5.2, one may use the notation X^I for $\mathcal{X}^I$ and call it *integral Lang locus* of X. We extend the definition to reducible algebraic varieties as in Section 2.

Remark 5.3. Assume that X is proper over $\bar{\mathbb{Q}}$. Then there is an integral model $(\mathcal{X}, F, \Sigma)$ for X, such that $\mathcal{X}$ is proper over $O_{F,\Sigma}$. By [Poo17, Thm. 3.2.13 (ii)], X^I coincides with the Lang locus of X.

Definition 5.4. [Ull04, Déf. 2.3] An integral algebraic variety X over $\bar{\mathbb{Q}}$ is called arithmetically hyperbolic if $X^I = \emptyset$.

An integral algebraic variety X over $\bar{\mathbb{Q}}$ is arithmetically hyperbolic if and only if for one (hence for every by Lemma 5.2) model $(\mathcal{X}, F, \Sigma)$, the set of integral points $\mathcal{X}(O_{M,\Omega})$ is finite for every finite extension $(M, \Omega)/(F, \Sigma)$ (so [Ull04, Lem. 2.4] follows from Lemma 5.2).

Example 5.5. Let $X = \mathbf{P}^1 \setminus \{0, 1, \infty\} = Y(2)$ be a modular curve over $\bar{\mathbb{Q}}$. Its Baily-Borel compactification is $X^* = \mathbf{P}^1$, and the Lang locus of X is full. By the Siegel-Mahler theorem (see, *e.g.*, [HS00, Thm. D.8.1]), X is arithmetically hyperbolic.

A complex analytic space is called *Kobayashi hyperbolic*, if its Kobayashi pseudo-distance (in the sense of [Kob98, p.50]) is a metric. Every Kobayashi hyperbolic, complex analytic space is Brody hyperbolic. Conversely, Brody [Bro78, p.213] proves that every compact, Brody hyperbolic complex analytic space is Kobayashi hyperbolic. In view of Remark 5.3, Conjecture 5.6 implies Conjecture 1.1.

Conjecture 5.6 ([Lan91, IX, Conjecture 5.1], [Ull04, Conjecture 2.5])**.** *Let X be a quasi-projective, integral algebraic variety over $\bar{\mathbb{Q}}$. If the complex analytic space $X(\mathbb{C})$ is Kobayashi hyperbolic, then X is arithmetically hyperbolic.*

Fact 5.7 is an evidence of Conjecture 5.6. It relies on Faltings's solution [Fal83, Satz 6] to Shafarevich's conjecture.

Fact 5.7 ([Ull04, Thm. 3.2 (a)])**.** *Let (G, X) be an adjoint Shimura datum of abelian type* (in the sense of [Ull04, p.4118])*. Let $K \leq G(\mathbb{A}_f)$ be a neat compact open subgroup. Then every irreducible component of* $\mathrm{Sh}_K(G, X)_{\bar{\mathbb{Q}}}$ *is arithmetically hyperbolic.*[4]

We prove that an alternative principle holds for integral points on Shimura varieties, by checking Assumption 2.1. Since an irreducible component of X^I with dimension 0 is an isolated point, Assumption 2.1 1 holds. Lemma 5.8 verifies Assumptions 2.1 2, 3 and 5.

Lemma 5.8. *Let $f : Z_1 \to Z_2$ be a morphism of integral algebraic varieties over $\bar{\mathbb{Q}}$. If f has finite geometric fibers, then $f(Z_1^I) \subset Z_2^I$.*

Proof. One may choose a number field F, a finite set Σ of places of F containing all the archimedean ones, a model $\mathcal{Z}_i$ over $O_{F,\Sigma}$ for Z_i $(i = 1, 2)$ and an $O_{F,\Sigma}$-morphism $f' : \mathcal{Z}_1 \to \mathcal{Z}_2$ whose base change to F is f. For every finite extension

[4]By [Moo98b, 2.17], the model over $\bar{\mathbb{Q}}$ defined by Faltings [Fal82] (used in [Ull04, Thm. 3.2 (a)]) is the scalar extension of the canonical model along $E(G, X) \to \bar{\mathbb{Q}}$.

$(M,\Omega)/(F,\Sigma)$, one has $f'(\mathcal{Z}_1(O_{M,\Omega})) \subset \mathcal{Z}_2(O_{M,\Omega})$, so $f(Z_1(\mathcal{Z}_1, M, \Omega)) \subset Z_2(\mathcal{Z}_2, M, \Omega)$. Hence

$$f(\overline{Z_1(\mathcal{Z}_1, M, \Omega)}) \subset \overline{Z_2(\mathcal{Z}_2, M, \Omega)}.$$

Let $C \subset \overline{Z_1(\mathcal{Z}_1, M, \Omega)}$ be an irreducible component of positive dimension. Then $f(C)$ is irreducible but not a singleton. (For otherwise, C is a finite set by assumption, which is a contradiction). Hence

$$f(C) \subset \overline{Z_2(\mathcal{Z}_2, M, \Omega)}^{>0} \subset Z_2^I.$$

Therefore, $f(\overline{Z_1(\mathcal{Z}_1, M, \Omega)}^{>0}) \subset Z_2^I$ and $f(Z_1^I) \subset Z_2^I$. □

Corollary 5.9 ([Ull04, Prop. 2.6]). *A locally closed subvariety of an arithmetically hyperbolic variety is also arithmetically hyperbolic.*

Proof. It follows from Lemma 5.8. □

Lemma 5.10 verifies Assumption 2.1 4 for integral Lang loci.

Lemma 5.10. *Let X be an integral algebraic variety over $\bar{\mathbb{Q}}$. Then $X^I \subset (X^I)^I$.*

Proof. Write $X^I = \cup_{i=1}^n Y_i$ as the union of irreducible components. Take a model $(\mathcal{X}, F, \Sigma)$ for X. Let $\mathcal{Y}_i$ be the scheme-theoretic image of the composition $Y_i \to X \to \mathcal{X}$, which is model of Y_i relative to (F, Σ). For every finite extension $(M,\Omega)/(F,\Sigma)$, the Zariski closed subset $\overline{X(\mathcal{X}, M, \Omega)} \subset X$ is the disjoint union of $\overline{X(\mathcal{X}, M, \Omega)}^{>0}$ with a finite set $\{p_1, \ldots, p_t\} \subset X(\bar{\mathbb{Q}})$.

Consider $x \in \mathcal{X}(O_{M,\Omega})$, *i.e.*, a section $x : \mathrm{Spec}(O_{M,\Omega}) \to \mathcal{X}$ to the structure morphism $\mathcal{X} \to \mathrm{Spec}(O_{M,\Omega})$. If $x|_{\mathrm{Spec}\,\bar{\mathbb{Q}}} \notin \{p_1, \ldots, p_t\}$, then

$$x|_{\mathrm{Spec}\,\bar{\mathbb{Q}}} \in \overline{X(\mathcal{X}, M, \Omega)}^{>0} \subset X^I.$$

Thus, there exists an index $1 \leq i \leq n$ with $x|_{\mathrm{Spec}\,\bar{\mathbb{Q}}} \in Y_i$. Since $\mathcal{Y}_i$ is Zariski closed in $\mathcal{X}$, the section x factors through $\mathcal{Y}_i$, *i.e.*, $x \in \mathcal{Y}_i(O_{M,\Omega})$. Therefore,

$$X(\mathcal{X}, M, \Omega) \subset \cup_{i=1}^n Y_i(\mathcal{Y}_i, M, \Omega) \cup \{p_1, \ldots, p_t\}.$$

Then

$$\overline{X(\mathcal{X}, M, \Omega)}^{>0} \subset \cup_{i=1}^n \overline{Y_i(\mathcal{Y}_i, M, \Omega)}^{>0} \subset \cup_{i=1}^n Y_i^I = (X^I)^I,$$

so $X^I \subset (X^I)^I$. □

Lemma 5.11 implies [Ull04, Prop. 2.8].

Lemma 5.11 (Chevalley-Weil). *If $f : X \to Y$ is an étale cover over $\bar{\mathbb{Q}}$, then $f(X^I) = Y^I$. In particular, $X^{I^\infty} = X^I$. Moreover, $X^I = X$ (resp. $X^I = \emptyset$) is equivalent to $Y^I = Y$ (resp. $Y^I = \emptyset$).*

Proof. By Lemma 5.8, one has $f(X^I) \subset Y^I$. There is a number field F, a finite set Σ of places of F containing all the archimedean ones, and a finite étale $O_{F,\Sigma}$-morphism $f' : \mathcal{X} \to \mathcal{Y}$ between models whose base change to the generic fiber recovers f.

For every finite extension $(M, \Omega)/(F, \Sigma)$, by the Chevalley-Weil theorem (see, *e.g.*, [Ser97, p.50]), there is a finite extension $(M', \Omega')/(M, \Omega)$ with $Y(\mathcal{Y}, M, \Omega) \subset f(X(\mathcal{X}, M', \Omega'))$. Since zero dimensional schemes are discrete,

$$\overline{Y(\mathcal{Y}, M, \Omega)}^{>0} \subset f(\overline{X(\mathcal{X}, M', \Omega')}^{>0}) \subset f(X^I).$$

Hence $Y^I \subset f(X^I)$. □

Theorem 5.12. *The integral Lang locus of a Shimura variety S is either empty or whole S.*

Proof. By Lemmas 5.8 and 5.10, the formation of the integral Lang locus $(\cdot)^I$ satisfies Assumption 2.1. The result is a combination of Theorem 4.1 and Lemma 5.11. □

References

[Bro78] Robert Brody. Compact manifolds in hyperbolicity. *Transactions of the American Mathematical Society*, 235:213–219, 1978.

[Che09] Ke Chen. *Special subvarieties of mixed Shimura varieties*. PhD thesis, Université Paris-Sud, 2009. `http://www.numdam.org/item/BJHTUP11_2009__0773__A1_0/`.

[CK16] Anna Cadoret and Arno Kret. Galois-generic points on Shimura varieties. *Algebra & Number Theory*, 10(9):1893–1934, 2016.

[CLZ16] Ke Chen, Xin Lu, and Kang Zuo. On the Oort conjecture for Shimura varieties of unitary and orthogonal types. *Compositio Mathematica*, 152(5):889–917, 2016.

[Cos05] Izzet Coskun. The arithmetic and geometry of Kobayashi hyperbolicity. In *Snowbird Lectures in Algebraic Geometry*, volume 388 of *Contemporary Mathematics*, pages 77–88. American Mathematical Society, 2005.

[Del71] Pierre Deligne. Travaux de Shimura. In *Séminaire Bourbaki vol. 1970/71 Exposés 382–399*, pages 123–165. Springer, 1971.

[Fal82] G Faltings. Arithmetic varieties and rigidity. In *Seminar on number theory, Paris*, pages 63–77, 1982.

[Fal83] Gerd Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Inventiones mathematicae*, 73(3):349–366, 1983.

[GN20] A Genestier and BC Ngô. Lectures on Shimura varieties. In *Shimura varieties*. Cambridge University Press, 2020.

[Gro65] Alexander Grothendieck. Éléments de géométrie algébrique : IV. Étude locale des schémas et des morphismes de schémas, Seconde partie. *Publications Mathématiques de l'IHÉS*, 24:5–231, 1965.

[Gro66] Alexander Grothendieck. Éléments de géométrie algébrique : IV. Étude locale des schémas et des morphismes de schémas, Troisième partie. *Publications Mathématiques de l'IHÉS*, 28:5–255, 1966.

[Har77] Robin Hartshorne. *Algebraic geometry*, volume 52. Springer Science & Business Media, 1977.

[HS00] Marc Hindry and Joseph H Silverman. *Diophantine geometry: an introduction*, volume 201. Springer Science & Business Media, 2000.

[Kob98] Shoshichi Kobayashi. *Hyperbolic complex spaces*, volume 318. Springer Science & Business Media, 1998.

[KUY18] Bruno Klingler, Emmanuel Ullmo, and Andrei Yafaev. Bi-algebraic geometry and the André-Oort conjecture. In *Algebraic geometry: Salt Lake City 2015*, volume 97, pages 319–359. American Mathematical Society Providence, 2018.

[KY14] Bruno Klingler and Andrei Yafaev. The André-Oort conjecture. *Annals of mathematics*, 180(3):867–925, 2014.

[Lan74] Serge Lang. Higher dimensional Diophantine problems. *Bulletin of the American Mathematical Society*, 80(5):779 – 787, 1974.

[Lan86] Serge Lang. Hyperbolic and Diophantine analysis. *Bulletin of the American Mathematical Society*, 14(2):159–205, 1986.

[Lan91] Serge Lang. *Number theory III: Diophantine geometry*, volume 60 of *Encyclopaedia of Mathematical Sciences*. Springer Science & Business Media, 1991.

[Liu06] Qing Liu. *Algebraic geometry and arithmetic curves*. Oxford University Press, 2nd edition, 2006.

[LZ19] Xin Lu and Kang Zuo. The Oort conjecture on Shimura curves in the Torelli locus of curves. *Journal de Mathématiques Pures et Appliquées*, 123:41–77, 2019.

[Mil99] JS Milne. Descent for Shimura varieties. *Michigan Mathematical Journal*, 46(1):203–208, 1999.

[Mil17] James S Milne. Introduction to Shimura varieties (revised version). `https://www.jmilne.org/math/`, 2017.

[Moo98a] Ben Moonen. Linearity properties of Shimura varieties, I. *Journal of algebraic geometry*, 7(3):539–568, 1998.

[Moo98b] Ben Moonen. Models of Shimura varieties in mixed characteristics. In *Galois representations in arithmetic algebraic geometry. Proceedings of the symposium, Durham, UK, July 9–18, 1996*, pages 267–350. Cambridge University Press, 1998.

[Nad89] Alan Michael Nadel. The nonexistence of certain level structures on abelian varieties over complex function fields. *Annals of Mathematics*, 129(1):161–178, 1989.

[Noo06] Rutger Noot. Correspondances de Hecke, action de Galois et la conjecture d'André-Oort. In *Séminaire Bourbaki : volume 2004/2005, exposés 938-951*, number 307 in Astérisque, pages 165–197. Société mathématique de France, 2006. talk:942.

[Pin90] Richard Pink. *Arithmetical compactification of mixed Shimura varieties*. PhD thesis, Rheinischen Friedrich-Wilhelms-Universitat Bonn, 1990. `https://people.math.ethz.ch/~pink/ftp/phd/PinkDissertation.pdf`.

[Poo17] Bjorn Poonen. *Rational points on varieties*, volume 186. American Mathematical Soc., 2017.

[Ser97] Jean-Pierre Serre. *Lectures on the Mordell-Weil theorem*. Springer, 3rd edition, 1997.

[Sta24] The Stacks project authors. The stacks project. `https://stacks.math.columbia.edu`, 2024.

[Ull04] Emmanuel Ullmo. Points rationnels des variétés de Shimura. *International Mathematics Research Notices*, 2004(76):4109–4125, 2004.

[UY10] Emmanuel Ullmo and Andrei Yafaev. Points rationnels des variétés de Shimura: un principe du "tout ou rien". *Mathematische Annalen*, 348(3):689–705, 2010.

[UY14] Emmanuel Ullmo and Andrei Yafaev. Galois orbits and equidistribution of special subvarieties: towards the André-Oort conjecture. *Annals of mathematics*, 180(3):823–865, 2014.