Hecke orbits as Shimura varieties in positive characteristic

Ching-Li Chai*

Abstract. Let p be a prime number, and let $\mathcal{M}$ be a modular variety of PEL type over $\overline{\mathbb{F}}_p$ which classifies abelian varieties in characteristic p with extra symmetries of a fixed PEL type. Consider the p-divisible group with extra symmetries consisting of all p-power torsions of the universal abelian scheme over $\mathcal{M}$. The locus in $\mathcal{M}$ corresponding to a fixed isomorphism type of a p-divisible group with extra symmetry is called a *leaf* by F. Oort. Each leaf is a smooth locally closed subvariety of the modular variety $\mathcal{M}$ which is stable under all prime-to-p Hecke correspondences on $\mathcal{M}$. Oort conjectured that every Hecke orbit is dense in the leaf containing it. Tools fashioned for this conjecture include (a) rigidity, (b) global monodromy, and (c) canonical coordinates. The theory of canonical coordinates generalizes the classical Serre–Tate coordinates; it asserts that locally at the level of jet-spaces, every leaf is built up from p-divisible formal groups through a finite family of fibrations in a canonical way. The Hecke orbit conjecture is affirmed when $\mathcal{M}$ is a Siegel modular variety classifying principally polarized abelian varieties of a fixed dimension, and also when $\mathcal{M}$ is a Hilbert modular variety classifying abelian varieties with real multiplications. The proof of the Siegel case, joint with F. Oort, uses the irreducibility of non-supersingular leaves in Hilbert modular varieties due to C.-F. Yu. That proof relies heavily on a special property of Siegel modular varieties: The set of $\overline{\mathbb{F}}_p$-rational points of a Siegel modular variety $\mathcal{A}_{g,n}$ is filled up by $\overline{\mathbb{F}}_p$-rational points of Hilbert modular varieties contained in $\mathcal{A}_{g,n}$. Possible directions for further progress include Tate-linear subvarieties and p-adic monodromy. The title of this article suggests that each leaf deserves to be viewed as a Shimura variety in characteristic p in its own right.

Mathematics Subject Classification (2000). Primary 14K10; Secondary 11G10.

Keywords. Shimura variety, Hecke correspondence, moduli, leaf, Barsotti–Tate group, abelian variety, monodromy.

1. Introduction

Let p be prime number and let $k \supset \mathbb{F}_p$ be an algebraically closed field fixed throughout this article; the field k will serve as the base field of modular varieties. The reader may want to take $k = \overline{\mathbb{F}}_p$.

We are interested in Hecke symmetries on the reduction to k of a Shimura variety. Because the theory of integral models of Shimura varieties are not fully developed yet, we will restrict to a small class of Shimura varieties, call modular varieties of PEL

*Partially supported by NSF grant DMS04-00482.

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006

type. Such a modular variety classifies abelian varieties with prescribed polarization and endomorphisms of a fixed type.

We will further restrict our attention to the prime-to-p Hecke symmetries. Since these symmetries come from finite étale isogeny correspondences for the universal abelian scheme over a modular variety $\mathcal{M}$, they preserve all p-adic invariants of geometric fibers of the universal Barsotti–Tate group on $\mathcal{M}$. Familiar examples of p-adic invariants include the p-rank and the Newton polygon. In 1999 F. Oort had the insight that if one uses "the mother of all p-adic invariants", namely the isomorphism class of the geometric fibers of the universal Barsotti–Tate group, then instead of a stratification of $\mathcal{M}$ by a finite number of subvarieties, one gets a decomposition of $\mathcal{M}$ into an infinite number of smooth locally closed subvarieties. In [30] these subvarieties are called *central leaves*, which we simplify to *leaves* here. In general the leaves in a given modular variety have moduli: They are parametrized by a scheme of finite type over k. Oort's Hecke orbit conjecture asserts that the leaves are determined by the Hecke symmetries: Every prime-to-p Hecke orbit is dense in the leaf containing it.

In this article we explain techniques motivated by the Hecke orbit problem: global ℓ-adic monodromy (Proposition 3.3), canonical coordinates on leaves (§4), hypersymmetric points (§5), local stabilizer principle (Proposition 6.1) and local rigidity (Theorem 6.2). The first is group-theoretic, while the rest four constitutes an effective "linearization method" of the Hecke orbit problem, which is illustrated in 6.3. The techniques developed so far are strong enough to affirm the Hecke orbit conjecture for Siegel modular varieties and Hilbert modular varieties. Advances in p-adic monodromy may lead to further progress; see §7.

In many ways a leaf in a modular variety $\mathcal{M}$ over a field of characteristic p resembles a Shimura variety in characteristic zero:

- The action of the Hecke symmetries on a leaf is topologically transitive according to the Hecke orbit Conjecture 3.2.
- By Proposition 3.3, the ℓ-adic monodromy of a leaf of positive dimension in $\mathcal{M}$ is $G(\mathbb{Q}_\ell)$, where G is a semisimple group attached to $\mathcal{M}$.
- A leaf is homogeneous in the sense that the local structure of a leaf are the same throughout the leaf, in view of the theory of canonical coordinates in §4.
- It seems plausible that a suitable parabolic subgroup P of an inner twist G' of G attached to a leaf $\mathcal{C}$ is closely related to the p-adic monodromy of $\mathcal{C}$; see §7.2. Hints of such a connection already appeared in [21].
- The theory of canonical coordinates suggests that one considers a leaf $\mathcal{C}$ as above to be "uniformized" by G'/P in a weak sense.

The above analogy depicts a scene in which a Shimura variety spawns an infinitude of morphed characteristic p replicas while reducing itself modulo p, an image that resonates with the mantra of *Indra's Pearls*[1]. We hope that the readers find this analogy somewhat sound, or perhaps even pleasing.

2. Hecke symmetry on modular varieties

A modular varieties of PEL type classifies abelian varieties with a prescribed type of polarization, endomorphisms and level structure. To a given PEL type is associated a tower of modular varieties. A locally compact group, consisting of prime-to-p finite adelic points of a reductive algebraic group over $\mathbb{Q}$, operates on this tower; this action induces Hecke correspondences on a fixed modular variety in the tower.

2.1. PEL data. Let B be a finite dimensional semisimple algebra over $\mathbb{Q}$, let $\mathcal{O}_B$ be a maximal order of B maximal at p, and let $*$ be a positive involution on B preserving $\mathcal{O}_B$. Let V be a B-module of finite dimension over $\mathbb{Q}$, let $\langle \cdot, \cdot \rangle$ be a $\mathbb{Q}$-valued nondegenerate alternating form on V compatible with $(B, *)$, and let $h\colon \mathbb{C} \to \mathrm{End}_{B\otimes_{\mathbb{Q}}\mathbb{R}}(V \otimes_{\mathbb{Q}} \mathbb{R})$ be a $*$-homomorphism such that

$$(v, w) \ \mapsto \ \langle v, h(\sqrt{-1})w \rangle$$

defines a positive definite real-valued symmetric form on $V \otimes_{\mathbb{Q}} \mathbb{R}$. The 6-tuple $(B, *, \mathcal{O}_B, V, \langle \cdot, \cdot \rangle, h)$ is called a *PEL datum unramified at* p if B is unramified at p and there exists a self-dual $\mathbb{Z}_p$-lattice in $V \otimes_{\mathbb{Q}} \mathbb{Q}_p$ stable under $\mathcal{O}_B$.

2.2. Modular varieties of PEL type. Suppose that we are given a PEL datum $(B, *, \mathcal{O}_B, V, \langle \cdot, \cdot \rangle, h)$ unramified at p, one associates a tower of modular varieties $(\mathcal{M}_{K^p})$ indexed by the set of all compact open subgroups K^p of $G(\mathbb{A}_f^p)$, where G is the unitary group attached to the pair $(\mathrm{End}_B(V), *)$, and $\mathbb{A}_f^p = \prod'_{\ell \neq p} \mathbb{Q}_\ell$ is the ring of prime-to-p finite adeles attached to $\mathbb{Q}$. The modular variety $\mathcal{M}_{K^p}$ classifies abelian varieties A with endomorphisms by $\mathcal{O}_B$ plus prime-to-p polarization and level-structure, whose H_1 is modeled on the given PEL datum; see [20, §5] for details.

2.3. Hecke symmetries. The group $G(\mathbb{A}_f^p)$ operates on the whole projective system $(\mathcal{M}_{K^p})$. If a level subgroup K_0^p is fixed, then on the corresponding modular variety $\mathcal{M}_{K_0^p}$ the remnant from the action of $G(\mathbb{A}_f^p)$ takes the form of a family of algebraic finite étale algebraic correspondences on $\mathcal{M}_{K_0^p}$; they are known as *Hecke correspondences*.

For a given point $x \in \mathcal{M}_{K_0^p}(k)$, denote by $\mathcal{H}^p \cdot x$ the subset of $\mathcal{M}_{K_0^p}(k)$ consisting of all elements which belongs to the image of x under some prime-to-p Hecke correspondence on $\mathcal{M}_{K_0^p}$. The countable set $\mathcal{H}^p \cdot x$ is called the prime-to-p Hecke orbit of x; it is equal to the image of $G(\mathbb{A}_f^p) \cdot \tilde{x}$ in $\mathcal{M}_{K_0^p}(k)$, where $\tilde{x} \in \varprojlim_{K^p} \mathcal{M}_{K^p}(k)$ is a pre-image of x.

[1] Cf. the preface of [25].

2.4. Examples

2.4.1. Siegel modular varieties. Let $g, n \in \mathbb{N}$, $(n, p) = 1$, and $n \geq 3$. Denote by $\mathcal{A}_{g,n}$ the modular variety over k which classifies all g-dimensional principally polarized abelian varieties (A, λ) over k with a symplectic level-n structure η. Two k-points $[(A_1, \lambda_1, \eta_1)]$, $[(A_2, \lambda_2, \eta_2)]$ in $\mathcal{A}_{g,n}$ are in the same prime-to-p Hecke orbit if and only if there exists a prime-to-p quasi-isogeny β (="$\beta_2 \circ \beta_1^{-1}$")

$$\beta\colon A_1 \xleftarrow{\beta_1} A_3 \xrightarrow{\beta_2} A_2$$

defined by prime-to-p isogenies β_1 and β_2 such that β respects the principal polarizations λ_1 and λ_2 in the sense that $\beta_1^*(\lambda_1) = \beta_2^*(\lambda_2)$. The semisimple algebra B in the PEL datum is equal to $\mathbb{Q}$. The reductive group G attached to the PEL datum is the symplectic group Sp_{2g}. The modular variety attached to the principal congruence subgroup of level-n in $\mathrm{Sp}_{2g}(\mathbb{A}_f^p)$ is $\mathcal{A}_{g,n}$.

2.4.2. Hilbert modular varieties. Let $E = F_1 \times \cdots \times F_r$, where $F_1, \ldots, F_r$ are totally real number fields. Consider the PEL datum where $B = E$, $* = \mathrm{Id}_E$, and V is a free E-module of rank two. Then the reductive group attached to the PEL datum is the kernel of the composition

$$\prod_{E/\mathbb{Q}} \mathrm{GL}_2 \xrightarrow{\det} \prod_{E/\mathbb{Q}} \mathbb{G}_m \xrightarrow{\mathrm{Nm}_{E/\mathbb{Q}}} \mathbb{G}_m,$$

where $\prod_{E/\mathbb{Q}}$ denotes Weil's restriction of scalars functor from E to $\mathbb{Q}$. A typical member of the associated tower of modular varieties is $\mathcal{M}_{E,n}$, with $n \geq 3$ and $(n, p) = 1$, which classifies $[E : \mathbb{Q}]$-dimensional abelian varieties A over k, together with a ring homomorphism $\iota\colon \mathcal{O}_E = \mathcal{O}_{F_1} \times \cdots \times \mathcal{O}_{F_r} \to \mathrm{End}(A)$, an $\mathcal{O}_E$-linear level-n structure η, and an $\mathcal{O}_E$-linear polarization of A.

There are different versions of polarizations; the one in [14] is as follows. It is a positivity-preserving $\mathcal{O}_E$-linear homomorphism $\lambda\colon \mathcal{L} \to \mathrm{Hom}^{\mathrm{sym}}_{\mathcal{O}_E}(A, A^t)$ from an projective rank-one $\mathcal{O}_E$-module $\mathcal{L}$ with a notion of positivity, which induces an $\mathcal{O}_E$-linear isomorphism $\lambda\colon A \otimes_{\mathcal{O}_E} \mathcal{L} \xrightarrow{\sim} A^t$, where A^t is the dual abelian variety of A. The modular variety $\mathcal{M}_{E,n}$ is not smooth over k if any one of the totally real fields F_i is ramified above p; if so then $\mathcal{M}_{E,n}$ has moderate singularities – it is a local complete intersection.

The prime-to-p Hecke orbit of a point $[(A, \iota_A, \lambda_A, \eta_A)] \in \mathcal{M}_{E,n}(k)$ consists of all points $[(B, \iota_B, \lambda_B, \eta_B)] \in \mathcal{M}_{E,n}(k)$ such that there exists a prime-to-p $\mathcal{O}_E$-linear quasi-isogeny from A to B which preserves the polarizations.

2.4.3. Picard modular varieties. Let L be an imaginary quadratic field contained in $\mathbb{C}$ such that p is unramified in L. Let a, b be positive integers, and let $g = a + b$. For the PEL datum, we take $B = L$ with the involution induced by the complex conjugation, a g-dimensional vector space V over L, and an L-linear complex structure $h\colon \mathbb{C} \to \mathrm{End}_{L\otimes_{\mathbb{Q}}\mathbb{R}}(V \otimes_{\mathbb{Q}} \mathbb{R})$ on $V \otimes_{\mathbb{Q}} \mathbb{R}$ satisfying the following condition: For every element $u \in L \subset \mathbb{C}$, the trace of the action of u on V_1 is equal to $au + b\bar{u}$, where $V_1 = \{v \in V \otimes_{\mathbb{Q}} \mathbb{C} : h(z)(v) = z \cdot v \text{ for all } z \in \mathbb{C}\}$. We also fix a ring homomorphism $\varepsilon\colon \mathcal{O}_L \to k$, i.e. a $\mathcal{O}_L$-algebra structure on k.

Let $n \geq 3$ be a positive integer, $(n, p) = 1$. The Picard modular variety $\mathcal{M}_{L,a,b,n}$ over k classifies $\mathcal{O}_L$-linear$g$-dimensional abelian varieties (A, ι) of signature (a, b), together with a principal polarization $\lambda\colon A \to A^t$ such that $\lambda \circ \iota(\bar{u}) = \iota(u)^t \circ \lambda$ for all $u \in \mathcal{O}_L$, and a symplectic $\mathcal{O}_L$-linear level-n structure. The signature condition above is that

$$\det_{\mathcal{O}_L\otimes_{\mathbb{Z}}k} (T \cdot \mathrm{Id} - \iota(u)|\mathrm{Lie}(A, \iota)) = (T-\varepsilon(u))^a \cdot (T-\varepsilon(u))^b \in k[T] \quad \text{for all } u \in \mathcal{O}_L.$$

As before the prime-to-p Hecke orbit of a point $[(A, \iota_A, \lambda_A, \eta_A)] \in \mathcal{M}_{L,a,b,n}(k)$ consists of all points $[(B, \iota_B, \lambda_B, \eta_B)] \in \mathcal{M}_{L,a,b,n}(k)$ such that there exists a prime-to-p $\mathcal{O}_L$-linear quasi-isogeny from A to B which preserves the polarizations.

3. Leaves and the Hecke orbit conjecture

3.1. Leaves in modular varieties in characteristic *p*

Definition 3.1. Let $\mathcal{M} = \mathcal{M}_{K_0^p}$ be a modular variety of PEL type over k as in 2.2. Let x_0 be a point in $\mathcal{M}(k)$. The *leaf* $\mathcal{C}_{\mathcal{M}}(x_0)$ in $\mathcal{M}$ passing through x_0 is the reduced locally closed subvariety of $\mathcal{M}$ over k such that $\mathcal{C}_{\mathcal{M}}(x_0)(k)$ consists of all points $x = [(A, \lambda, \iota, \eta)] \in \mathcal{C}_{\mathcal{M}}(x_0)(k)$ such that the Barsotti–Tate group (or p-divisible group) $(A[p^\infty], \lambda[p^\infty], \iota[p^\infty])$ with prescribed polarization and endomorphisms attached to x is isomorphic to that attached to x_0.

Remark. (i) The notion of leaves was introduced in [30]; it was studied later by Vasiu in [38].

(ii) In addition to being locally closed, every leaf in $\mathcal{M}$ is a smooth subvariety of $\mathcal{M}$ and is closed under all prime-to-p Hecke correspondences.

3.2. The Hecke orbit conjectures. The Hecke orbit conjecture HO, due to Oort, asserts that the decomposition of a modular variety $\mathcal{M}$ of PEL type into leaves is determined by the Hecke symmetries on $\mathcal{M}$. It is equivalent to the conjunction of the continuous version $\mathrm{HO}_{\mathrm{ct}}$ and the discrete version $\mathrm{HO}_{\mathrm{dc}}$ below.

Conjecture 3.2 (HO). Every prime-to-p Hecke orbit in a modular variety of PEL type $\mathcal{M}$ over k is dense in the leaf in $\mathcal{M}$ containing it.

Conjecture ($\mathrm{HO}_{\mathrm{ct}}$). The closure of any prime-to-p Hecke orbit in the leaf $\mathcal{C}$ containing it is an open-and-closed subset of $\mathcal{C}$, i.e. it is a union of irreducible components of the smooth variety $\mathcal{C}$.

Conjecture ($\mathrm{HO}_{\mathrm{dc}}$). Every prime-to-$p$ Hecke orbit in a leaf $\mathcal{C}$ meets every irreducible component of $\mathcal{C}$.

3.3. Global ℓ-adic monodromy. The discrete Hecke orbit conjecture is essentially an irreducibility statement, in view of the following result on global monodromy.

Proposition 3.3. *Let $\mathcal{M}$ be a modular variety of PEL type attached to a reductive group G over $\mathbb{Q}$ as in 2.2. Let $G_{\mathrm{der}}^{\mathrm{sc}}$ be the simply connected cover of the derived group of G. Let $x_0 \in \mathcal{M}(k)$ be a point of $\mathcal{M}$ such that the prime-to-p Hecke orbit of x_0 with respect to every simple factor of $G_{\mathrm{der}}^{\mathrm{sc}}$ is infinite. Let $Z(x_0)$ be the Zariski closure of the prime-to-p Hecke orbit of x_0 for the group $G_{\mathrm{der}}^{\mathrm{sc}}$ in the leaf $\mathcal{C}(x_0)$ in $\mathcal{M}$ containing x_0. Then $Z(x_0)$ is irreducible, and the Zariski closure of the ℓ-adic monodromy group of $Z(x_0)$ is $G_{\mathrm{der}}(\mathbb{Q}_\ell)$ for every prime number $\ell \neq p$.*

Remark. (i) A stronger version of 3.3 for Siegel modular varieties is proved in [5]. The argument in [5] works for all modular varieties of PEL type.

(ii) The irreducibility statement in Proposition 3.3 is a useful tool for proving irreducibility of a given subvariety Z of modular varieties of PEL type which are stable under the prime-to-p Hecke correspondences: It reduces the task to proving Hecke transitivity on $\pi_0(Z)$.

3.4. Some known cases of the Hecke orbit conjecture

Theorem 3.4. *The Hecke orbit conjecture* HO *holds for Siegel modular varieties.*

Theorem 3.5. *The Hecke orbit conjecture* HO *holds for Hilbert modular varieties attached to a finite product $F_1 \times \cdots \times F_r$ of totally real fields. Here the prime p may be ramified in any or all of the totally real fields $F_1, \ldots, F_r$.*

Remark. (i) Theorem 3.4 is joint work with F. Oort. Details of the proof of Theorem 3.4 will appear in a monograph with F. Oort. The proof of the continuous version $\mathrm{HO}_{\mathrm{ct}}$ in the Siegel case uses Theorem 3.5.

(ii) The proof of Theorem 3.5 is the result of joint work with C.-F. Yu; the proof of the discrete version, i.e. the irreducibility of non-supersingular leaves in Hilbert modular varieties, is the work of C.-F. Yu.

(iii) Among the methods used in the proof of Theorem 3.4, the action of the local stabilizer subgroup and the trick of using Hilbert modular subvarieties first appeared in [2], where the case of Theorem 3.4 for ordinary principally polarized abelian varieties was proved.

(iv) A detailed sketch of the proof of Theorem 3.4 can be found in [4].

4. Canonical coordinates on leaves

4.1. Classical Serre–Tate coordinates. Recall that an abelian variety A over k is *ordinary* if the Barsotti–Tate group $A[p^\infty]$ is the extension of an étale Barsotti–Tate group by a toric Barsotti–Tate group over k. It has been more than forty years when Serre and Tate discovered that the local deformation space of an ordinary abelian variety A over k has a natural structure as a formal torus over $W(k)$ of relative dimension $\dim(A)^2$. For Siegel modular varieties, their result says that if $x = [(A, \lambda)] \in \mathcal{A}_{g,n}(k)$ is a closed point of $\mathcal{A}_{g,n}$ such that A is an ordinary abelian variety, then the formal completion $\mathcal{A}_{g,n}^{/x}$ of $\mathcal{A}_{g,n} \to \mathrm{Spec}(\mathbb{Z})$ has a natural structure as a formal torus over $W(k)$ of relative dimension $\frac{g(g+1)}{2}$, where $g = \dim(A)$. Notice that the ordinary locus of $\mathcal{A}_{g,n}$ over k is equal to the dense open stratum in the Newton polygon stratification of $\mathcal{A}_{g,n}$.

The above approach generalizes to modular variety of PEL type, to the effect that the mixed-characteristic local deformation space for a point in the dense open Newton polygon stratum of a modular variety $\mathcal{M}$ of PEL type can be built up from Barsotti–Tate groups over $W(k)$ by a system of fibrations; see [23].

There is a long-standing question as to whether one can find a reasonable theory of canonical coordinates for points outside the generic Newton polygon stratum of a modular variety of PEL type. It turns out that the answer is *yes* if one restricts to a leaf in a modular variety.

4.2. The slope filtration. The starting point is the observation that there exists a natural *slope filtration* on the restriction of the universal Barsotti–Tate group to a leaf; moreover the slope filtration gives the local moduli of a leaf.

Proposition 4.1. *Let $\mathcal{M}$ be a modular variety of PEL type over k attached to a PEL datum $(B, *, \mathcal{O}_B, V, \langle\cdot,\cdot\rangle, h)$ as in 2.2. Let $\mathcal{C}$ be a leaf in $\mathcal{M}$. Let $\tilde{X} \to \mathcal{C}$ be the restriction to $\mathcal{C}$ of the Barsotti–Tate group attached to the universal abelian variety.*

(i) *There exist rational numbers $\mu_1, \dots \mu_m$ with $1 \geq \mu_1 > \cdots > \mu_m \geq 0$ and Barsotti–Tate groups*

$$0 = \tilde{X}_0 \subset \tilde{X}_1 \subset \tilde{X}_2 \subset \cdots \subset \tilde{X}_m = \tilde{X}$$

over $\mathcal{C}$ such that $\tilde{Y}_i := \tilde{X}_i/\tilde{X}_{i-1}$ is a non-trivial isoclinic Barsotti–Tate group over $\mathcal{C}$ of slope μ_i for each $i = 1, \dots, m$.

(ii) *The filtration $0 = \tilde{X}_0 \subset \tilde{X}_1 \subset \tilde{X}_2 \subset \cdots \subset \tilde{X}_m = \tilde{X}$ is uniquely determined by $\tilde{X} \to \mathcal{C}$. Each subgroup $\tilde{X}_i \subseteq \tilde{X}$ is stable under the natural action of $\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p$.*

(ii) *For each $i = 1, \dots, r$ the Barsotti–Tate group $\tilde{Y}_i \to \mathcal{C}$ is geometrically constant, hence $\tilde{Y}_i$ is isomorphic to the twist of a constant Barsotti–Tate group by a smooth étale $\mathbb{Z}_p$-sheaf over $\mathcal{C}$.*

Remark. (i) That $\tilde{Y}_i$ is *isoclinic* of slope μ_i means that the kernel of the N-th iterate of the relative Frobenius for $\tilde{Y}_i$ is comparable to the kernel of $[p^{N\mu_i}]_{\tilde{X}_i/\tilde{X}_{i-1}}$ for all sufficiently large multiples of the denominator of μ_i.

(ii) The proof of Proposition 4.1 depends on [42] and [34].

4.3. The cascade structure. Combining Proposition 4.1 with the Serre–Tate theorem, one sees that the local moduli of a leaf $\mathcal{C}$ comes from the deformation of the slope filtration.

Let $x = [(A, \iota, \lambda, \eta)] \in \mathcal{M}(k)$ be a closed point of the modular variety $\mathcal{M}$, and let $\mathrm{Fil}_{A[p^\infty]} = (0 = X_0 \subset X_1 \subset \cdots \subset X_m = A[p^\infty])$ be the slope filtration of $A[p^\infty]$. Let $Y_i = X_i/X_{i-1}$ for $i = 1, \ldots, m$. For each pair (i, j) with $1 \le i \le j \le m$, let $\mathfrak{Def}([i, j], \iota) = \mathfrak{Def}([i, j], \iota[p^\infty])$ be the deformation functor over k of the filtered Barsotti–Tate group

$$0 \subset X_i/X_{i-1} \subset X_{i+1}/X_{i-1} \subset \cdots \subset X_j/X_{i-1}$$

with action by $\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p$. Each $\mathfrak{Def}([i, j], \iota)$ is a smooth formal scheme over k, and $\mathfrak{Def}([i, i], \iota) = \mathrm{Spec}(k)$ for each i. For $1 \le i < j \le m$, let $\mathfrak{DE}(i, j; \iota) = \mathfrak{DE}(i, j; \iota[p^\infty])$ be the deformation functor of the filtered Barsotti–Tate group $0 \subset Y_i \subset Y_i \times_{\mathrm{Spec}(k)} Y_j$ with action by $\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p$; it is a smooth formal scheme over k. Each $\mathfrak{DE}(i, j; \iota)$ has a natural structure as a smooth commutative formal group over k; the group structure comes from via Baer sum. Notice that $\mathfrak{Def}([i, i+1], \iota)$ is a torsor over $\mathfrak{DE}(i, i+1; \iota])$ for $i = 1, \ldots, m-1$.

We have a family of forgetful morphisms

$$\pi_{[i+1,j],[i,j]} : \mathfrak{Def}([i, j], \iota]) \to \mathfrak{Def}([i+1, j], \iota), \quad 1 \le i < j \le m,$$

and

$$\pi_{[i,j-1],[i,j]} : \mathfrak{Def}([i, j], \iota) \to \mathfrak{Def}([i, j-1], \iota), \quad 1 \le i < j \le m$$

such that

$$\pi_{[i+1,j-1],[i+1,j]} \circ \pi_{[i+1,j],[i,j]} = \pi_{[i+1,j-1],[i,j-1]} \circ \pi_{[i,j-1],[i,j]} \quad \text{if } i \le j-2.$$

Each morphism $\pi_{[i+1,j],[i,j]}$ is smooth, same for each $\pi_{[i,j-1],[i,j]}$.

For each pair (i, j) with $1 \le i < j \le m$, define a commutative smooth formal group

$$\pi'_{[i+1,j],[i,j]} \colon \mathfrak{DE}(i, [i+1, j]; \iota) \to \mathfrak{Def}([i+1, j], \iota)$$

as follows. For each Artinian local ring R over k and for each R-valued point $f \colon \mathrm{Spec}(R) \to \mathfrak{Def}([i+1, j], \iota)$ corresponding to a deformation

$$0 \subset \tilde{X}_{[i+1,i+1]} \subset \cdots \subset \tilde{X}_{[i+1,j]}$$

of the filtration $(0 \subset X_{i+1}/X_i \subset \cdots \subset X_j/X_i)$ over R, define the set of R-valued points of $\mathfrak{DE}(i, [i+1, j], \iota)$ over f to be the set of all isomorphism classes of extensions of $\tilde{X}_{[i+1,j]}$ by $Y_i \times_{\mathrm{Spec}(k)} \mathfrak{Def}([i+1, j], \iota)$. It is easy to see that $\pi_{[i+1,j],[i,j]}$ has a natural structure as a torsor for $\pi'_{[i+1,j],[i,j]}$. Similarly one can define a commutative formal group

$$\pi'_{[i,j-1],[i,j]} \colon \mathfrak{DE}([i, j-1], j; \iota) \to \mathfrak{Def}([i, j-1]$$

for $1 \leq i < j \leq m$ so that $\pi_{[i,j-1],[i,j]}$ has a natural structure as a torsor for $\pi'_{[i,j-1],[i,j]}$.

Consider the natural map

$$\pi_{[i,j]}\colon \mathfrak{Def}([i,j],\iota) \to \mathfrak{Def}([i+1,j],\iota) \times_{\mathfrak{Def}([i+1,j-1],\iota)} \mathfrak{Def}([i,j-1],\iota)$$

defined by the maps $\pi_{[i+1,j],[i,j]}$ and $\pi_{[i,j-1],[i,j]}$. It turns out that in a suitable sense the map $\pi_{[i,j]}$ has a natural structure as a *torsor for a biextension* of

$$\big(\mathfrak{DE}([i+1,j-1],j;\iota),\ \mathfrak{DE}(i,[i+1,j-1];\iota)\big)$$

by (the base extension to $\mathfrak{Def}([i+1,j-1],\iota)$ of) the commutative smooth formal group $\mathfrak{DE}(i,j;\iota)$ if $i \leq j-2$. Notice that for the two factors of the target of the map $\pi_{[i,j]}$, the first factor $\mathfrak{Def}([i+1,j],\iota) \to \mathfrak{Def}([i+1,j-1],\iota)$ is a torsor for the group $\mathfrak{DE}([i+1,j-1],j;\iota) \to \mathfrak{Def}([i+1,j-1],\iota)$, while the second factor $\mathfrak{Def}([i,j-1],\iota) \to \mathfrak{Def}([i+1,j-1],\iota)$ is a torsor for the group $\mathfrak{DE}(i,[i+1,j-1];\iota) \to \mathfrak{Def}([i+1,j-1],\iota)$.

The formal structure of a family such as

$$\begin{aligned}\mathfrak{MDE} = \big(&\mathfrak{Def}([i,j],\iota]),\ \mathfrak{Def}([i,j],\iota),\pi_{[i+1,j],[i,j]},\\ &\pi_{[i,j-1],[i,j]},\pi'_{[i+1,j],[i,j]},\pi'_{[i,j-1],[i,j]},\pi_{[i,j]}\big)\end{aligned}$$

will be called a *cascade*, following the terminology in [23], although the situation here is somewhat more complicated than [23].

When x is a point of the generic Newton polygon stratum of $\mathcal{M}$, the maximal subcascade of $\mathfrak{MDE}$ fixed by the involution induced by the polarization λ coincides with the formal completion $\mathcal{M}^{/x}$ of $\mathcal{M}$ at x. So $\mathcal{M}^{/x}$ is built up from suitable subgroups of the commutative formal groups $\mathfrak{DE}(i,j;\iota[p^\infty])$ over k through a family of fibrations; see [23].

4.4. Maximal p-divisible subcascade. Suppose that x lies outside the generic Newton polygon stratum. Then when one deforms the slope filtration, the resulting Barsotti–Tate group may fail to remain geometrically constant. It turns out that the maximal reduced closed formal subscheme of $\mathfrak{Def}([0,m];\iota)$ is in some sense the *maximal p-divisible subcascade* of the cascade $\mathfrak{MDE}$ of formal groups attached to x; the latter is built up from the maximal p-divisible formal subgroups $\mathfrak{DE}(i,j;\iota)_{\mathrm{pdiv}}$ of $\mathfrak{DE}(i,j;\iota)$, with (i,j) running through all pairs with $1 \leq i < j \leq m$.

The polarization λ of the abelian variety A induces an involution on the formal scheme $\mathfrak{Def}([0,m];\iota])$. and also an involution of the maximal p-divisible subcascade $\mathfrak{MDE}_{\mathrm{pdiv}}$ of the cascade of formal groups $\mathfrak{MDE}$ attached to $x \in \mathcal{M}(k)$. The maximal closed formal subscheme of the formal scheme underlying $\mathfrak{MDE}_{\mathrm{pdiv}}$ is equal to the formal completion $\mathcal{C}^{/x}$ of the leaf $\mathcal{C}$ in $\mathcal{M}$ containing x. In particular $\mathcal{C}^{/x}$ is built up from p-divisible formal groups over k through a family of fibrations.

4.5. The two slope case. Let X, Y be isoclinic Barsotti–Tate groups over k with slopes $\mu_X < \mu_Y$. Let h_X and h_Y be the height of X and Y respectively. Let $\mathfrak{DE}(X, Y)$ be the deformation functor over k of the filtration $0 = Y \subset X \times_{\mathrm{Spec}(k)} Y$; it is a commutative smooth formal group over k. Let $\mathfrak{DE}(X, Y)_{\mathrm{pdiv}}$ be the maximal p-divisible formal subgroup of the commutative smooth formal group $\mathfrak{DE}(X, Y)$ over k.

Let $\mathrm{M}(X)$ and $\mathrm{M}(Y)$ be the Cartier module of X and Y respectively. We refer to [43] for the Cartier theory. On $H_{\mathbb{Q}} := \mathrm{Hom}_{W(k)}(\mathrm{M}(X), \mathrm{M}(Y)) \otimes_{\mathbb{Z}} \mathbb{Q}$ we have a σ-linear operator F and a σ^{-1}-linear operator V on $H_{\mathbb{Q}}$ given by

$$(V \cdot h)(u) = V(h(V^{-1}u)), \quad (F \cdot h)(u) = F(h(V(u))) \quad \text{for all } h \in H_{\mathbb{Q}}, u \in \mathrm{M}(X).$$

Clearly $\mathrm{Hom}_{W(k)}(\mathrm{M}(X), \mathrm{M}(Y))$ is stable under the action of F.

Theorem 4.2. *Notation as above.*

(i) *The p-divisible formal group $\mathfrak{DE}(X, Y)_{\mathrm{pdiv}}$ is isoclinic of slope $\mu_Y - \mu_X$; its height is equal to $h_X \cdot h_Y$.*

(ii) *The Cartier module of $\mathfrak{DE}(X, Y)_{\mathrm{pdiv}}$ is naturally isomorphic to the maximal $W(k)$-submodule of* $\mathrm{Hom}_{W(k)}(\mathrm{M}(X), \mathrm{M}(Y))$ *which is stable under the actions of F and V.*

(iii) *Suppose that $Y = X^t$ is the Serre dual of X. Then we have a natural involution $*$ on $\mathfrak{DE}(X, X^t)_{\mathrm{pdiv}}$, and the Cartier module of the maximal formal subgroup of $\mathfrak{DE}(X, X^t)_{\mathrm{pdiv}}$ fixed under $*$ is the maximal $W(k)$-submodule of* $\mathrm{Hom}_{W(k)}(\mathrm{S}^2(\mathrm{M}(X)), W(k))$ *which is stable under the actions of F and V.*

Remark. (i) See [8] for a proof of Theorem 4.2.

(ii) The set of all p-typical curves in the reduced Cartier ring functor, with three compatible actions by the reduced Cartier ring over k, plays a major role in the proof.

(iii) The case when we have a maximal order $\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p$ in an unramified semisimple algebra $B \otimes_{\mathbb{Q}} \mathbb{Q}_p$ over $\mathbb{Q}_p$ operating on X and Y is easily deducible from Theorem 4.2.

5. Hypersymmetric points

Over a field of characteristic zero one has the notion of *special points* in modular varieties of PEL type, corresponding to abelian varieties of CM-type (or, with sufficiently many complex multiplications). On the other hand, it is well-known that every abelian variety over $\overline{\mathbb{F}}_p$ has sufficiently many complex multiplications, so one can say that every $\overline{\mathbb{F}}_p$-point of a modular variety $\mathcal{M}$ of PEL type is “special”. But there are points in $\mathcal{M}$ that are more distinguished than others – they correspond to abelian varieties whose $\mathcal{O}_B$-endomorphism ring is “as big as allowed by the slope constraint”.

Definition 5.1. (i) Let B be a simple algebra over $\mathbb{Q}$, and let $\mathcal{O}_B$ be an order of B. Let A be an abelian variety over k, and let $\iota\colon \mathcal{O}_B \to \mathrm{End}(A)$ be a ring homomorphism. We say that (A, ι) is a *hypersymmetric $\mathcal{O}_B$-linear abelian variety* if the canonical map $\mathrm{End}_{\mathcal{O}_B}(A) \otimes_{\mathbb{Z}} \mathbb{Z}_p \to \mathrm{End}_{\mathcal{O}_B}(A[p^\infty])$ is an isomorphism.

(ii) Let $\mathcal{M}$ be a modular variety of PEL type as in 2.2. A point $x \in \mathcal{M}(k)$ is *hypersymmetric* if the underlying $\mathcal{O}_B$-linear abelian variety (A_x, ι_x) is hypersymmetric.

Remark 5.2. (i) When $B = \mathbb{Q}$, it is easy to see that an abelian variety A over k is hypersymmetric if and only if it is isogenous to a finite product of abelian varieties $B_1 \times \cdots \times B_r$ defined over a finite field $\mathbb{F}_q$ such that the action of the Frobenius element $\mathrm{Fr}_{B_i,\mathbb{F}_q}$ on the first ℓ-adic cohomology group of B_i has at most two eigenvalues for each $i = 1, \ldots, r$, and B_i and B_j share no common slope if $i \neq j$. See [9, §2, §3].

(ii) One can use the method in [9, §5] to show that there exist hypersymmetric points on any leaf of a modular variety of PEL type over k. However one has difficulty showing the existence of hypersymmetric points on every irreducible component of a given leaf of a modular variety, without knowing or assuming the irreducibility of the leaf.

(iii) If the semisimple $\mathbb{Q}$-rank of the reductive group G attached to the PEL datum for the modular variety $\mathcal{M}$ is equal to one, then every $\overline{\mathbb{F}}_p$-point of $\mathcal{M}$ is hypersymmetric. For instance, every $\overline{\mathbb{F}}_p$-point of the modular curve is hypersymmetric. One consequence of this phenomenon is that one cannot simply substitute "special points" by "hypersymmetric points" and expect to get a reasonable formulation of the André-Oort conjecture in characteristic p; see [9, §7].

6. Action of stabilizer subgroups and rigidity

6.1. Stabilizer subgroups. Let $\mathcal{M}$ be a modular variety over k of PEL type as in 2.2. Attached to a point $x \in \mathcal{M}(k)$ corresponding to a quadruple $(A_x, \iota_x, \lambda_x, \eta_x)$ are two compact p adic groups:

- Let $\mathrm{G_x}(\mathbb{Z}_p) = \mathrm{Aut}_{\mathcal{O}_B}(A_x[p^\infty], \lambda_x[p^\infty])$. We call $G_x(\mathbb{Z}_p)$ the *local p-adic automorphism group* at x, and

- Let $\mathrm{H_x}$ be the unitary group attached to the semisimple algebra with involution $(\mathrm{End}_{\mathcal{O}_B}(A_x) \otimes_{\mathbb{Z}} \mathbb{Q}, *_x)$, where $*_x$ is the Rosati involution attached to λ_x. Let $\mathrm{H}_x(\mathbb{Z}_p)$ be the group of $\mathbb{Z}_p$-points of $\mathrm{H_x}$ with respect to the integral structure given by $\mathrm{End}(A_x)$. We call $\mathrm{H}_x(\mathbb{Z}_p)$ the *local stabilizer subgroup* at x. We have a natural embedding $\mathrm{H}_x(\mathbb{Z}_p) \hookrightarrow G_x(\mathbb{Z}_p)$.

6.2. Action on deformation space. We have a natural action of $G_x(\mathbb{Z}_p)$ on $\mathcal{M}^{/x}$, the formal completion of $\mathcal{M}^{/x}$. This action comes from the combination of

(a) a classical theorem of Serre and Tate, which states that the deformation functor for the abelian variety A_x is canonically isomorphic to the deformation functor attached to the Barsotti–Tate group $A_x[p^\infty]$, and

(b) the action of $G_x(\mathbb{Z}_p)$ on the deformation functor for the $\mathcal{O}_B$-linear Barsotti–Tate group $(A_x[p^\infty], \iota_x[p^\infty])$ by "change of marking", or "transport of structure".

The local stabilizer subgroup $\mathrm{H}_x(\mathbb{Z}_p)$ can be regarded as the p-adic completion of the stabilizer subgroup at x in the set of all prime-to-p Hecke correspondences. Consequently the *local stabilizer principle* holds:

Proposition 6.1 (Local stabilizer principle). *If Z is a closed subvariety of $\mathcal{M}$ stable under all prime-to-p Hecke correspondence and $x \in Z(k)$ is a closed point of Z, then the formal completion $Z^{/x} \subset \mathcal{M}^{/x}$ of Z at x is stable under* $\mathrm{H}_x(\mathbb{Z}_p)$ *for the action described in* 6.2*.*

The local stabilizer principle can be effectively deployed for studying Hecke-invariant subvarieties when combined with the rigidity result below.

Theorem 6.2 (Local rigidity). *Let X be a p-divisible formal group over k. Let H be a connected reductive linear algebraic subgroup over $\mathbb{Q}_p$ and let $\rho\colon H(\mathbb{Q}_p) \to (\mathrm{End}_k(X) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^\times$ be a rational linear representation of $H(\mathbb{Q}_p)$ such that the composition of ρ with the left regular representation of* $\mathrm{End}_k(X) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ *does not contain the trivial representation of $H(\mathbb{Q}_p)$ as a subquotient. Let Z be an irreducible closed formal subscheme of X. Assume that Z is stable under the natural action of an open subgroup U of $H(\mathbb{Q}_p)$ on X. Then Z is a p-divisible formal subgroup of X.*

The proof of 6.2 is elementary; see [6]. An instructive special case of 6.2 asserts that any irreducible closed formal subvariety of a formal torus over k which is stable under multiplication by $1 + p^n$ for some $n \geq 1$ is a formal subtorus.

6.3. Linearization of the Hecke orbit problem. The combination of the local stabilizer principle and local rigidity leads to an effective linearization of the Hecke orbit problem: Consider the case when $\mathcal{M}$ is a Siegel modular variety $\mathcal{A}_{g,n}$ and $x \in \mathcal{A}_{g,n}(\overline{\mathbb{F}}_p)$ corresponds to a g-dimensional principally polarized abelian variety A_x defined over $\overline{\mathbb{F}}_p$ with two slopes $\lambda < 1-\lambda$. Then the formal completion $\mathcal{C}(x)^{/x}$ at x of the leaf $\mathcal{C}(x)$ containing x has a natural structure as an isoclinic p-divisible formal group of hight $\frac{g(g+1)}{2}$ and slope $1-2\lambda$. The local stabilizer principle and Theorem 6.2 imply that the formal completion at x of the Zariski closure of the prime-to-p Hecke orbit of x is a p-divisible formal subgroup of $C(x)^{/x}$; moreover this p-divisible formal subgroup is stable under the natural action of the local stabilizer subgroup $\mathrm{H}_x(\mathbb{Z}_p)$.

Continuing the situation above, and assume that A_x is hypersymmetric in the sense of 5.1. Then the Zariski closure of the prime-to-p Hecke orbit of x coincides with the irreducible component of $\mathcal{C}(x)$ in an open neighborhood of x, because the action on the local stabilizer subgroup $\mathrm{H}_x(\mathbb{Z}_p)$ on the Cartier module of the p-divisible formal group $C(x)^{/x}$ underlies an absolutely irreducible representation of $\mathrm{H}_x(\mathbb{Q}_p)$.

6.4. Hypersymmetric points and the Hecke orbit conjecture. Let $x \in \mathcal{M}(\overline{\mathbb{F}}_p)$ be an $\overline{\mathbb{F}}_p$ point of a modular variety $\mathcal{M}$ of PEL type, and let $Z(x)$ be the Zariski closure in the leaf $\mathcal{C}(x)$ of the prime-to-p Hecke orbit of x. The argument in 6.3 shows that the continuous Hecke orbit conjecture $\mathrm{HO}_{\mathrm{ct}}$ for $Z(x)$ would follow if one can show that there exists a hypersymmetric point y in $Z(x)$.

The Hecke orbit conjecture HO for Hilbert modular varieties comes into the proof of the continuous Hecke orbit conjecture $\mathrm{HO}_{\mathrm{ct}}$ for Siegel modular varieties at this juncture. After a possibly inseparable isogeny correspondence, one can assume that the given point $x \in \mathcal{A}_{g,n}$ lies in a Hilbert modular subvariety in $\mathcal{A}_{g,n}$. After another application of the local stabilizer principle, one is reduced to the case when the abelian variety A_x has only two slopes. With the help of Theorem 3.5, one sees that $Z(x)$ contains the leaf through x in a Hilbert modular subvariety containing x, hence $Z(x)$ contains a hypersymmetric point in $\mathcal{A}_{g,n}$. See [4] for a more detailed outline of the argument.

Remark 6.3. (i) The proof of the Hecke orbit conjecture for Siegel modular varieties outlined above relies on a special property of Siegel modular varieties: For every point $x \in \mathcal{A}_{g,n}(\overline{\mathbb{F}}_p)$, there exists a Hilbert modular variety $\mathcal{M}_{E,n}$, a finite-to-one correspondence $f : \mathcal{M}_{E,n} \to \mathcal{A}_{g,n}$ equivariant with respect to the prime-to-p Hecke correspondences, and a point $y \in \mathcal{M}_{E,n}(\overline{\mathbb{F}}_p)$ above x. See [4, §9], labeled as the "Hilbert trick". The point is that, every $\overline{\mathbb{F}}_p$-point of $\mathcal{A}_{g,n}$ lies in a subvariety which is essentially the reduction of a "small" Shimura subvariety of positive dimension, namely a Hilbert modular variety attached to a product E of totally real fields such that $\dim_{\mathbb{Q}}(E) = g$. Here "small" means that every factor of the reductive group attached to the Shimura subvariety has semisimple $\mathbb{Q}$-rank one.

(ii) The property that every rational point over a finite field lies in the image of a small Shimura variety of positive dimension holds for modular varieties of PEL type C, but fails for modular varieties of type A and D. Consequently, the Hecke orbit conjecture for modular varieties of PEL type C is within reach by available methods, while new ideas are needed for PEL types A and D, or the reduction of general Shimura varieties.

7. Open questions and outlook

We discuss two approaches toward a proof of the Hecke orbit conjecture for Siegel modular varieties without resorting to the Hilbert trick.

7.1. Tate-linear subvarieties in leaves. For simplicity, we consider the special case of a leaf $\mathcal{C}$ in a Siegel modular variety $\mathcal{A}_{g,n}$, such that every point of $\mathcal{C}$ corresponds to a g-dimensional principally polarized abelian variety with two slopes $\lambda < 1 - \lambda$. This assumption implies that the formal completion $\mathcal{C}^{/x}$ of $\mathcal{C}$ has a natural structure as an isoclinic p-divisible formal group with slope $1-2\lambda$ and height $\frac{g(g+1)}{2}$, for any closed point $x \in \mathcal{C}$.

An irreducible closed subvariety $Z \subset \mathcal{C}$ is said to be *Tate linear* at a closed point $x \in Z$ if the formal completion $Z^{/x}$ is a p-divisible formal subgroup of $\mathcal{C}^{/x}$. It can be shown that if Z is Tate-linear at one closed point of $\mathcal{C}$, then it is Tate-linear at every closed point of the smooth locus of Z

Remark. (i) The proof that the property of being Tate-linear propagates from one point of Z to every point of Z depends on a global version of canonical coordinates. The case when $\mathcal{C}$ is the ordinary locus of $\mathcal{A}_{g,n}$ has been documented in [7], where several issues related to the notion of Tate-linear subvarieties are addressed.

(ii) The notion of Tate-linear subvarieties is inspired by the Hecke orbit problem: Suppose that $\mathcal{M}$ is a modular variety of PEL type contained in a Siegel modular variety $\mathcal{A}_{g,n}$, and $x \in \mathcal{M}(\overline{\mathbb{F}}_p) \cap \mathcal{C}(\overline{\mathbb{F}}_p)$ is a point of $\mathcal{M}$ such that the abelian variety A_x attached to x has two slopes $\lambda < 1-\lambda$. Then the Zariski closure of the Hecke orbit $\mathcal{H}^p \cdot x$ in the leaf $\mathcal{C}_{\mathcal{M}}(x)$ is a Tate linear subvariety of $\mathcal{C}$.

Question 7.1. The most intriguing question about the notion of Tate-linear subvarieties is whether every Tate-linear subvariety of a leaf $\mathcal{C}$ in $\mathcal{A}_{g,n}$ is (an irreducible component of) the intersection of $\mathcal{C}$ with the reduction of a Shimura subvariety of $\mathcal{A}_{g,n}$ in characteristic 0.

It seems plausible that the answer is a *qualified yes*. This naive expectation will be termed the *global rigidity conjecture*.

Remark. (i) If the global rigidity conjecture is true, then the notion of Tate-linear subvarieties provides a geometric characterization for subvarieties of $\mathcal{C}$ which are equal to (an irreducible component of) the intersection of $\mathcal{C}$ with the reduction of a Shimura subvariety of $\mathcal{A}_{g,n}$.

(ii) The global rigidity conjecture should be considered as being stronger than the continuous Hecke orbit conjecture $\mathrm{HO}_{\mathrm{ct}}$: Continuing the set-up as in §7.1. Let Z be the Zariski closure in the leaf $\mathcal{C}_{\mathcal{M}}(x)$ in $\mathcal{M}$ containing x of the prime-to-p Hecke orbit in $\mathcal{M}$. Then Z is a Tate-linear subvariety, by Theorem 6.2. Moreover if the global rigidity conjecture is true, then one can deduce without difficulty that Z is a union of irreducible components of $\mathcal{C}(x)$.

7.2. *p*-adic monodromy. As we saw in 6.3, the combination of the local stabilizer principle, canonical coordinates and the local rigidity theory achieves a certain level of localization for the Hecke orbit problem. This linearization allows one to approach the Hecke orbit problem through the p-adic monodromy: Let Z be the Zariski closure of a given prime-to-p Hecke orbit $\mathcal{H}(x)$ in the leaf $\mathcal{C}$ containing $\mathcal{H}(x)$. Consider the restriction to Z of the universal Barsotti–Tate group $A[p^\infty] \to \mathcal{M}$ over the modular variety $\mathcal{M}$. Over the leaf $\mathcal{C}$, the Barsotti–Tate group $A[p^\infty] \to \mathcal{C}$ admits a *slope filtration*; the p-adic monodromy attached to the associated graded of the slope filtration of $A[p^\infty] \to \mathcal{Z}$ will be called the *naive p-adic monodromy* of $A[p^\infty] \to Z$.

Conjecture 7.2. The naive p-adic monodromy of the family $A[p^\infty] \to \mathcal{Z}$ is "as large as possible", in the sense that the image of the naive p-adic monodromy representation is an open subgroup of the group of $\mathbb{Q}_p$-points of a suitable Levi subgroup L of an inner twist G' of G attached to Z, where G is the reductive group attached to the PEL data for $\mathcal{M}$.

Remark. (i) Conjecture 7.2 for the Zariski closure Z of an Hecke orbit implies the continuous Hecke orbit conjecture $\mathrm{HO}_{\mathrm{cont}}$.

(ii) Conjecture 7.2 is a p-adic analogue of Proposition 3.3.

(iii) As a weak converse to Conjecture 7.2, the method of the proof of Proposition 7.4 below should enable one to show that the Hecke orbit conjecture HO implies Conjecture7.2, using a hypersymmetric point as the base point.

Remark 7.3. Given an abelian scheme $A \to S$, where S a scheme over $\mathbb{F}_p$, we would like to show that the naive p-adic monodromy for $A \to S$ is "as large as possible", subject to obvious constraints, such as cycles on the family $A \to S$. As an intermediate step toward this goal, one would like to show that, when S is the spectrum of a Noetherian local integral domain and the Newton polygon of the closed fiber of $A \to S$ is different from the Newton polygon of the generic fiber, the naive p-adic monodromy for the generic fiber of $A \to S$ is *large* in a suitable sense.

When $\dim(A/S) = 1$, the above *wish* is a classical theorem of Igusa. The argument of Igusa was generalized in [3] to the case of a one-dimensional p-divisible formal group with ordinary generic fiber. The same argument applies to the case of a p-divisible formal group with ordinary generic fiber such that the dimension and the codimension are coprime; details will appear in an article with D. U. Lee.

Proposition 7.4. *Let $\mathcal{A}^{\mathrm{or}}_{g,n}$ be the ordinary locus of a Siegel modular variety $\mathcal{A}_{g,n}$ over k, where $g \geq 1$, $n \geq 3$, $(n, p) = 1$, and the base field $k \supseteq \mathbb{F}_p$ is algebraically closed. Let $A \to \mathcal{A}^{\mathrm{or}}_{g,n}$ be the universal abelian scheme over $\mathcal{A}^{\mathrm{or}}_{g,n}$. Let $A[p^\infty]_{\mathrm{et}} \to \mathcal{A}^{\mathrm{or}}_{g,n}$ be the maximal étale quotient of $A[p^\infty] \to \mathcal{A}^{\mathrm{or}}_{g,n}$; it is an étale Barsotti–Tate group of height g. Let E_0 be an ordinary elliptic curve defined over $\overline{\mathbb{F}}_p$, and let $x_0 = (A_0, \lambda_0)$, where A_0 is the product of g copies of E_0, and λ_0 is the product principal polarization on A_0. Let $\mathrm{T}_p = \mathrm{T}_p(A_0[p^\infty])_{\mathrm{et}})$ be the p-adic Tate module of the étale p-divisible group $A_0[p^\infty]_{\mathrm{et}}$; it is naturally isomorphic to the direct sum of g copies of $\mathrm{T}_p(E_0[p^\infty]_{\mathrm{et}}) \cong \mathbb{Z}_p$, so $\mathrm{GL}(\mathrm{T}_p)$ is naturally isomorphic to $\mathrm{GL}_g(\mathbb{Z}_p)$. Let $\rho\colon \pi_1(\mathcal{A}^{\mathrm{or}}_{g,n}, x_0) \to \mathrm{GL}(\mathrm{T}_p)$ be the naive p-adic monodromy representation of $A[p^\infty] \to \mathcal{A}^{\mathrm{or}}_{g,n}$. Then the image of ρ is equal to $\mathrm{GL}(\mathrm{T}_p) \cong \mathrm{GL}_g(\mathbb{Z}_p)$.*

Proof. Let $\mathcal{B}$ be the product of g copies of $\mathcal{A}_{1,n}$, diagonally embedded in $\mathcal{A}_{g,n}$. Let E_0 be an ordinary elliptic curve defined over a $\overline{\mathbb{F}}_p$, and let $x_0 = (A_0, \lambda_0)$, where A_0 is the product of g copies of E_0, and λ_0 is the product principal polarization on A_0. Let $\mathcal{O} = \mathrm{End}(E_0)$. Then $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong \mathbb{Z}_p \times \mathbb{Z}_p$, corresponding to the natural splitting of $E_0[p^\infty]$ into the product of its toric part $E_0[p^\infty]_{\mathrm{tor}}$ and its étale part $E_0[p^\infty]_{\mathrm{et}}$. So we have an isomorphism $\mathrm{End}(A_0) \cong \mathrm{M}_g(\mathcal{O})$, and a splitting $\mathrm{End}(A_0) \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong$

$\mathrm{M}_g(\mathcal{O}) \times \mathrm{M}_g(\mathcal{O})$ corresponding to the splitting of $A_0[p^\infty]$ into the product its toric and étale parts. Denote by $\mathrm{pr}\colon (\mathrm{End}(A_0) \otimes_{\mathbb{Z}} \mathbb{Z}_p)^\times \to \mathrm{GL}(\mathrm{T}_p) \cong \mathrm{GL}_g(\mathbb{Z}_p)$ the projection corresponding to the action of $\mathrm{End}(A_0) \otimes_{\mathbb{Z}} \mathbb{Z}_p$ on the étale factor $A_0[p^\infty]_{\mathrm{et}}$ of $A_0[p^\infty]$. The Rosati involution $*$ on $\mathrm{End}(A_0)$ interchanges the two factors of $\mathrm{End}(A_0) \otimes_{\mathbb{Z}} \mathbb{Z}_p$. It follows that $\mathrm{U}(\mathcal{O}_{(p)} \otimes_{\mathbb{Z}} \mathbb{Z}_p, *)$ is isomorphic to $\mathrm{GL}(\mathrm{T}_p)$ under the projection map pr, therefore the image of $\mathrm{U}(\mathcal{O}_{(p)}, *)$ in $\mathrm{GL}(\mathrm{T}_p)$ is dense in $\mathrm{GL}(\mathrm{T}_p)$. Here $\mathcal{O}_{(p)} = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$, and $\mathbb{Z}_{(p)} = \mathbb{Q} \cap \mathbb{Z}_p$ is the localization of $\mathbb{Z}$ at the prime ideal $(p) = p\mathbb{Z}$.

By a classical theorem of Igusa, the p-adic monodromy group of the restriction to $\mathcal{B}$, i.e. $\rho(\mathrm{Im}(\pi_1(\mathcal{B}, x_0) \to \pi_1(\mathcal{A}_{g,n}, x_0)))$, is naturally identified with the product of g copies of $\mathbb{Z}_p^\times$ diagonally embedded in $\mathrm{GL}(\mathrm{T}_p) \cong \mathrm{GL}_g(\mathbb{Z}_p)$. Denote by D this subgroup of $\mathrm{GL}(\mathrm{T}_p)$.

Let $R_{(p)} = \mathrm{End}(A_0) \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)} \cong \mathrm{M}_g(\mathcal{O}) \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$. Every element $u \in R_{(p)}$ such that $u^*u = uu^* = 1$ gives rise to a prime-to-p isogeny from A_0 to itself respecting the polarization λ_0. Such an element $u \in R_{(p)}$ gives rise to

- a prime-to-p Hecke correspondence h on $\mathcal{A}_{g,n}$ having x_0 as a fixed point, and
- an irreducible component $\mathcal{B}'$ of the image of $\mathcal{B}$ under h such that $\mathcal{B}' \ni x_0$.

By the functoriality of the fundamental group, the image of the fundamental group $\pi_1(\mathcal{B}', x_0)$ of $\mathcal{B}'$ in $\pi_1(\mathcal{A}_{g,n}^{\mathrm{or}}, x_0)$ is mapped under the p-adic monodromy representation ρ to the conjugation of D by the element $\mathrm{pr}(h) \in \mathrm{GL}(\mathrm{T}_p)$. In particular, $\rho(\pi_1(\mathcal{A}_{g,n}^{\mathrm{or}}, x_0))$ is a closed subgroup of $\mathrm{GL}(\mathrm{T}_p)$ which contains all conjugates of D by elements in the image of $\mathrm{pr}\colon \mathrm{U}(E_{(p)}, *) \to \mathrm{GL}(\mathrm{T}_p)$.

Recall that the image of $\mathrm{U}(E_{(p)}, *)$ in $\mathrm{GL}(\mathrm{T}_p)$ is a dense subgroup. So the monodromy group $\rho(\pi_1(\mathcal{A}_{g,n}, x_0))$ is a closed normal subgroup of $\mathrm{GL}(\mathrm{T}_p) \cong \mathrm{GL}_g(\mathbb{Z}_p)$ which contains the subgroup D of all diagonal elements. An easy exercise in group theory shows that the only such closed normal subgroup is $\mathrm{GL}_g(\mathbb{Z}_p)$ itself. □

Remark. (i) There are at least two published proofs of Proposition 7.4 in the literature, in [15] and [16, chap. V §7] respectively.

(ii) In the proof of 7.4, one can use as the base point any element $[(A_1, \lambda)]$ of $\mathcal{A}_{g,n}^{\mathrm{or}}$ such that A_1 is separably isogenous to a product of g copies of an ordinary elliptic curve E_1 over $\overline{\mathbb{F}}_p$.

(iii) As already mentioned before, the argument of Proposition 7.4 applies to leaves in modular varieties of PEL type. Since we used a hypersymmetric point as the base point, a priori this argument applies only to those irreducible components of a given leaf which contain hypersymmetric points.

References

[1] Andreatta, F., Goren, E. Z., Hilbert modular varieties of low dimension. In *Geometric Aspects of Dwork's Theory* (ed. by A. Adolphson, F. Baldassarri, P. Berthelot, N. Katz, and F. Loeser), Volume I, Walter de Gruyter, Berlin 2004, 113–175.

[2] Chai, C.-L., Every ordinary symplectic isogeny class in positive characteristic is dense in the moduli. *Invent. Math.* **121** (1995), 439–479.

[3] Chai, C.-L., Local monodromy for deformations of one-dimensional formal groups. *J. Reine Angew. Math.* **524** (2000), 227–238.

[4] Chai, C.-L., Hecke orbits on Siegel modular varieties. In *Geometric Methods in Algebra and Number Theory*, Progr. Math. 235, Birkhäuser, Boston, MA, 2004, 71–107.

[5] Chai, C.-L., Monodromy of Hecke-invariant subvarieties. *Quaterly J. Pure Applied Math.* **1** (Borel Special Issues, part I) (2005), 291–303.

[6] Chai, C.-L., A rigidity result for p-divisible formal groups. Preprint, 2003. Available from http://www.math.upenn.edu/~chai/.

[7] Chai, C.-L., Families of ordinary abelian varieties: canonical coordinates, p-adic monodromy, Tate-linear subvarieties and Hecke orbits. Preprint, 2003.

[8] Chai, C.-L., Canonical coordinates on leaves of p-divisible groups: The two-slope case. Preprint, 2004.

[9] Chai, C.-L., Oort, F., Hypersymmetric abelian varieties. *Quaterly J. Pure Applied Math.* **2** (1) (2006) (Coates Special Issue), 1–27.

[10] Chai, C.-L., Oort, F., *Hecke Orbits*. Monograph in preparation.

[11] Chai, C.-L., Yu, C.-F., Fine structures and Hecke orbits on Hilbert modular varieties. Preliminary, 2004; revision in progress.

[12] de Jong, A. J., Homomorphisms of Barsotti-Tate groups and crystals in positive characteristic. *Invent. Math.* **134** (1998), 301–333.

[13] de Jong, A. J., Oort F., Purity of the stratification by Newton polygons, *J. Amer. Math. Soc.* **13** (2000), 209–241.

[14] P. Deligne and G. Pappas, Singularités des espaces de modules de Hilbert, en les caractéristiques divisant le discriminant. *Composito Math.* **90** (1994), 59–79.

[15] Ekedahl, T., The action of monodromy on torsion points of jacobian. In *Arithmetic Algebraic Geometry*, Progr. Math. 89, Birkhäuser, Boston, MA, 1991, 41–49.

[16] Faltings, G. and Chai, C.-L., *Degeneration of Abelian Varieties*. Ergeb. Math. Grenzgeb. (3) 22, Springer-Verlag, Berlin 1990.

[17] Goren, E. Z., Oort, F., Stratification of Hilbert modular varieties. *J. Algebraic Geom.* **9** (2000), 111–154.

[18] Grothendieck, I., *Groupes de Monodromie en Géométrie Algébrique* (SGA 7) I. Lecture Notes in Math. 288, Springer-Verlag, Berlin1972.

[19] Katz, N. M., Slope filtration of F-crystals. *Astérisque* **63** (1979), 113–164.

[20] Kottwitz, R. E., Points on some Shimura varieties over finite fields. *J. Amer. Math. Soc.* **2** (1992), 373–444.

[21] Kottwitz, R. E., Isocrystals with additional structures. *Compositio Math.* **56** (1985), 201–220.

[22] Manin, Y. I., The theory of commutative formal groups over fields of finite characteristic. *Uspehi Mat. Nauk* **1**8 (1963), 3–90; English transl. *Russian Math. Surveys* **18** (1963), 1–80.

[23] Moonen, B., Serre-Tate theory for moduli spaces of PEL type. *Ann. Sci. École Norm. Sup.* **37** (2004), 223–269.

[24] Mumford, D., Bi-extensions of formal groups. In *Algebraic Geometry* (Bombay, 1968), Oxford University Press, London 1969, 307–322.

[25] Mumford, D., Series, C., Wright, D., *Indra's Pearls. The Vision of Felix Klein*. Cambridge University Press, Cambridge 2002.

[26] Oda, T., Oort, F., Supersingular abelian varieties. In *Proceedings of the International Symposium on Algebraic Geometry* (Kyoto, 1977), Kinokuniya Book Store, Tokyo 1978, 595–621.

[27] Oort, F., Some questions in algebraic geometry. Preliminary version, 1995; available from http://www.math.uu.nl/people/oort/.

[28] Oort, F., Newton polygons and formal groups: conjectures by Manin and Grothendieck. *Ann. of Math.* **152** (2000), 183–206.

[29] Oort, F., A stratification of a moduli space of abelian varieties. In *Moduli of Abelian Varieties*, Progr. Math. 195, Birkhäuser, Basel 2001, 345–416.

[30] Oort, F., Foliations in moduli spaces of abelian varieties. *J. Amer. Math. Soc.* **17** (2004), 267–296.

[31] Oort, F., Minimal p-divisible groups. *Ann. of Math.* **161** (2005), 1021–1036

[32] Oort, F., Monodromy, Hecke orbits and Newton polygon strata. Seminar at Max Planck Institut, Bonn, Feb. 14, 2004; available from http://www.math.uu.nl/people/oort/.

[33] Oort, F., Hecke orbits in moduli spaces. Notes for the 2005 AMS Summer Institute on Algebraic Geometry, 22 pp.

[34] Oort, F., Zink, T., Families of p-divisible groups with constant Newton polygon. *Doc. Math.*, to appear.

[35] Ribet, K., p-adic interpolation via Hilbert modular forms. *Algebraic Geometry* (Arcata 1974), Proc. Symp. Pure Math. 29, Amer. Math. Soc., Providence, R.I., 1975, 581–592.

[36] Tate, J., Endomorphisms of abelian varieties over finite fields. *Invent. Math.* **2** (1966), 134–144.

[37] Tate, J., Classes d'isogenie de variétés abéliennes sur un corps fini (d'après T. Honda). In *Séminaire Bourbaki 1968/69*, Exposé 352, Lecture Notes in Math. 179, Springer-Verlag, Berlin 1971, 95–110.

[38] Vasiu, A., Crystalline boundedness principle. Preprint, 2005; available from http://math.arizona.edu/~vasiu/.

[39] Yu, C.-F., On the supersingular locus in the Hilbert-Blumenthal 4-folds. *J. Algebraic Geom.* **12** (2003), 653–698.

[40] Yu, C.-F., On reduction of Hilbert-Blumenthal varieties. *Ann. Inst. Fourier* **53** (2003), 2105–2154.

[41] Yu, C.-F., Discrete Hecke orbit problem for Hilbert-Blumenthal varieties. Preprint, 2004.

[42] Zink, T., On the slope filtration. *Duke Math. J.* **109** (2001), 79–95.

[43] Zink, T., Cartiertheorie kommutativer formaler Gruppen. Teubner-Texte Math. 68, B. G. Teubner Verlagsgesellschaft, Leipzig 1984.

Department of Mathematics, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, PA 19104-6395, U.S.A.

E-mail: chai@math.upenn.edu