Annals of Mathematics, **161** (2005), 1111–1142

Serre's conjecture over $\mathbb{F}_9$

By Jordan S. Ellenberg*

Abstract

In this paper we show that an odd Galois representation $\bar{\rho} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_9)$ having nonsolvable image and satisfying certain local conditions at 3 and 5 is modular. Our main tools are ideas of Taylor [21] and Khare [10], which reduce the problem to that of exhibiting points on a Hilbert modular surface which are defined over a solvable extension of $\mathbb{Q}$, and which satisfy certain reduction properties. As a corollary, we show that Hilbert-Blumenthal abelian surfaces with ordinary reduction at 3 and 5 are modular.

Introduction

In 1986, J-P. Serre proposed the following conjecture [16]:

Conjecture. *Let $\mathbb{F}$ be a finite field of characteristic p, and*

$$\bar{\rho} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F})$$

an absolutely irreducible representation such that $\det \bar{\rho}$ *applied to complex conjugation yields* -1. *Then* $\bar{\rho}$ *is the mod* p *representation attached to a modular form on* $\mathrm{GL}_2(\mathbb{Q})$.

Serre's conjecture, if true, would provide the first serious glimpse into the nonabelian structure of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. The work of Langlands and Tunnell shows that Serre's conjecture is true when $\mathrm{GL}_2(\mathbb{F})$ is solvable; that is, when $\mathbb{F}$ is $\mathbb{F}_2$ or $\mathbb{F}_3$. Work of Shepherd-Barron and Taylor [17] and Taylor [21] have shown that the conjecture is also true, under some local and global conditions on $\bar{\rho}$, when $\mathbb{F}$ is $\mathbb{F}_4$ or $\mathbb{F}_5$; the work of Breuil, Conrad, Diamond, and Taylor [2] proves the conjecture when $\mathbb{F}$ is $\mathbb{F}_5$ and $\det \bar{\rho}$ is cyclotomic. More recently, Manoharmayum [12] has proved Serre's conjecture when $\mathbb{F} = \mathbb{F}_7$, again subject

*Partially supported by NSA Young Investigator Grant MDA905-02-1-0097 and NSF Grant DMS-0401616.

to local conditions. His argument, like ours, uses the ideas of [21] and [10], together with a construction of solvable points on a certain modular variety.

In the present work, we show that Serre's conjecture is true, again subject to certain local and global conditions, when $\mathbb{F} = \mathbb{F}_9$. To be precise, we prove the following theorem.

THEOREM. *Let*

$$\bar{\rho} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_9)$$

be an odd Galois representation such that

- $\bar{\rho}$ *has nonsolvable image*;
- *The restriction of* $\bar{\rho}$ *to* D_3 *can be written as*

$$\bar{\rho}|D_3 \cong \begin{bmatrix} \psi_1 & * \\ 0 & \psi_2 \end{bmatrix},$$

  where $\psi_1|I_3$ *is the* mod 3 *cyclotomic character*, *and* ψ_2 *is unramified*;
- *The image of the inertia group* I_5 *lies in* $\mathrm{SL}_2(\mathbb{F}_9)$, *and has odd order.*

Then $\bar{\rho}$ *is modular.*

As a corollary, we get the following result towards a generalized Shimura-Taniyama-Weil conjecture for Hilbert-Blumenthal abelian surfaces:

COROLLARY. *Let* $A/\mathbb{Q}$ *be a Hilbert-Blumenthal abelian surface which has good ordinary or multiplicative reduction at* 3 *and* 5*. Then* A *is a quotient of* $J_0(N)$ *for some integer* N*.*

The corresponding theorem when A is an elliptic curve has now been proved without any hypotheses, thanks to the results of [24], [20], and [2]. The case where A is a Hilbert-Blumenthal abelian variety with real multiplication by a field with an ideal of norm 5 is treated in [17]. Our method follows theirs; one starts with a case of Serre's conjecture that one knows, and uses lifting theorems to prove modularity of a Hilbert-Blumenthal abelian variety.

We prove the theorem above by exhibiting $\bar{\rho}$ as the Galois representation on the 3-torsion subscheme of a certain Hilbert-Blumenthal abelian surface defined over a totally real extension $F/\mathbb{Q}$ with solvable Galois group. We then use an idea of Taylor, together with a theorem of Skinner and Wiles [19], to prove the modularity of the abelian surface, and consequently of $\bar{\rho}$.

The key algebro-geometric point is that a certain twisted Hilbert modular variety has many points defined over solvable extensions of $\mathbb{Q}$. This suggests that we consider the class of varieties X such that, if K is a number field, and

Σ is the set of all solvable Galois extensions L/K, then

$$\bigcup_{L\in\Sigma} X(L)$$

is Zariski-dense in X. We say X has "property **S**" in this case. Certainly if X has a Zariski-dense set of points over a *single* number field—for example, if X is unirational—it has property **S**. The Hilbert modular surfaces we consider, on the other hand, are varieties of general type with property **S**.

To indicate our lack of knowledge about solvable points on varieties, note that at present there does not exist a variety which we can prove does *not* have property **S**! Nonetheless, it seems reasonable to guess that "sufficiently complicated" varieties do not have property **S**.

One might consider the present result evidence for the truth of Serre's conjecture. On the other hand, it should be pointed out that the theorems here and in [17], [21] rely crucially on the facts that

- the GL_2 of small finite fields is solvable, and
- certain Hilbert modular varieties for number fields of small discriminant have property **S**.

These happy circumstances may not persist very far. In particular, it is reasonable to guess that only finitely many Hilbert modular varieties have property **S**. If so, one might say that we have much philosophical but little numerical evidence for the truth of Serre's conjecture in general. Our ability to compute has progressed mightily since Serre's conjecture was first announced. It would be interesting, given the present status of the conjecture, to carry out numerical experiments for $\mathbb{F}$ a "reasonably large" finite field—whatever that might mean.

The author gratefully acknowledges several helpful conversations with Brian Conrad, Eyal Goren, and Richard Taylor, and the careful reading and suggestions of the referee.

Added in Proof. Since the original submission of this paper, substantial progress has been made towards a resolution of Serre's conjecture. The recently announced work of Khare and Khare-Wintenberger proves Serre's conjecture in level 1 for an arbitrary coefficient field; this result, unlike ours, avoids the use of special geometric properties of low-degree Hilbert modular varieties, and thus presents a very promising direction for further progress. Recent work of Kisin generalizes the results we cite on lifting of modularity to handle many potentially supersingular cases; it seems likely that his methods could substantially simplify the argument of the present paper, by eliminating the necessity of showing that the abelian varieties we construct in Section 2 have ordinary reduction in characteristics 3 and 5.

Notation. If v is a prime of a number field F, we write G_F for the absolute Galois group of F and $D_v \subset G_F$ for the decomposition group associated to v, and I_v for the corresponding inertia group. The p-adic cyclotomic character of Galois is denoted by χ_p, and its mod p reduction by $\bar{\chi}_p$.

If $V \subset \mathbb{P}^N$ is a projective variety, write $F_1(V)$ for the Fano variety of lines contained in V.

If $\mathcal{O}$ is a ring, an *$\mathcal{O}$-module scheme* is an $\mathcal{O}$-module in the category of schemes.

All Hilbert modular forms are understood to have all weights equal.

We denote by ω a primitive cube root of unity.

1. Realizations of Galois representations on HBAV's

Recall that a *Hilbert-Blumenthal abelian variety* (HBAV) over a number field is an abelian d-fold endowed with an injection $\mathcal{O} \hookrightarrow \mathrm{End}(A)$, where $\mathcal{O}$ is the ring of integers of a totally real number field of degree d over $\mathbb{Q}$. Many Hilbert-Blumenthal abelian varieties can be shown to be modular; for example, see [17]. It is therefore sometimes possible to show that a certain mod p Galois representation $\bar{\rho}$ is modular by realizing it on the p-torsion subscheme of some HBAV.

We will show that, given a Galois representation $\bar{\rho} : \mathrm{Gal}(\bar{K}/K) \to \mathrm{GL}_2(\mathbb{F}_9)$ satisfying some local conditions at $3, 5$ and ∞, we can find $\bar{\rho}$ in the 3-torsion of an abelian surface over a solvable extension of K, satisfying some local conditions at 3 and 5. One of these conditions—that certain representations be "D_p-distinguished"—requires further comment.

Definition 1.1. Let $\bar{\rho} : \mathrm{Gal}(\bar{K}/K) \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ be a Galois representation, and let $\mathfrak{p}|p$ be a prime of K. We say that $\bar{\rho}$ is *$D_{\mathfrak{p}}$-distinguished* if the semisimplification of the restriction $\bar{\rho}|D_{\mathfrak{p}}$ is isomorphic to $\theta_1 \oplus \theta_2$, with θ_1 and θ_2 distinct characters from $D_{\mathfrak{p}}$ to $\bar{\mathbb{F}}_p^*$.

This condition is useful in deformation theory, and is required, in particular, in the main theorem of [19]. A natural source of $D_{\mathfrak{p}}$-distinguished Galois representations is provided by abelian varieties with ordinary reduction at $\mathfrak{p}$.

PROPOSITION 1.2. *Let p be an odd prime. Let K_v be a finite extension of $\mathbb{Q}_p$ with odd ramification degree, and let A/K_v be a principally polarized* HBAV *with good ordinary or multiplicative reduction and real multiplication by $\mathcal{O}$, and let $\mathfrak{p}$ be a prime of $\mathcal{O}$ dividing p.*

Then the semisimplification of the $\mathrm{Gal}(\bar{K}_v/K_v)$*-module $A[\mathfrak{p}]$ is isomorphic to $\theta_1 \oplus \theta_2$, with θ_1 and θ_2 distinct characters of* $\mathrm{Gal}(\bar{K}_v/K_v)$.

Proof. If A has multiplicative reduction, the theory of the Tate abelian variety yields an exact sequence

$$0 \to (\mu_p)^g \to A[p] \to (\mathbb{Z}/p\mathbb{Z})^g \to 0$$

over some unramified extension of K_v. If, on the other hand, A has good ordinary reduction, then A extends to an abelian scheme $\mathcal{A}$ over the ring of integers R_v of K_v. The finite flat group scheme $\mathcal{A}[p]/R_v$ then fits into the connected-étale exact sequence

$$0 \to \mathcal{A}[p]^0 \to \mathcal{A}[p] \to \mathcal{A}[p]^{\mathrm{et}} \to 0$$

and we denote by $A[p]^0/K_v$ and $A[p]^{\mathrm{et}}/K_v$ the generic fibers of the corresponding group schemes over R_v. Note that $A[p]^{\mathrm{et}}$ is unramified as a Galois representation, and has dimension g.

So in either case $A[p]$ has an unramified g-dimensional quotient A''. The Weil pairing yields an isomorphism of group schemes $A[p] \cong \mathrm{Hom}(A[p], \mu_p)$; the unramified quotient A'' thus gives rise to a g-dimensional submodule of $A[p]$ on which I_v acts cyclotomically.

Since the ramification degree of $K_v/\mathbb{Q}_p$ is odd, the cyclotomic character of I_v is nontrivial. It follows that $A[p]$ fits into an exact sequence of Galois representations

$$0 \to A' \to A[p] \to A'' \to 0$$

in which A'' is the I_v-coinvariant quotient of $A[p]$, and $\dim A' = \dim A'' = g$. Since the endomorphisms in $\mathcal{O}$ are defined over K_v, they respect this quotient; we conclude that the above exact sequence can be interpreted as a sequence of $\mathcal{O}$-modules. We know by [15, 2.2.1] that $A[\mathfrak{p}]$ is a two-dimensional vector space over $\mathcal{O}/\mathfrak{p}$. Since the action of $\mathcal{O}$ is compatible with Weil pairing, we have $\wedge^2 A[\mathfrak{p}] \cong \mu_p \otimes_{\mathbb{F}_p} \mathcal{O}/\mathfrak{p}$ as $\mathcal{O}$-modules. In particular, inertia acts cyclotomically on $\wedge^2 A[\mathfrak{p}]$, which means that $A[\mathfrak{p}] \cap A'$ must have dimension 1 over $\mathcal{O}/\mathfrak{p}$. We conclude that $A[\mathfrak{p}]$ fits into an exact sequence of $\mathcal{O}$-modules

$$0 \to A[\mathfrak{p}] \cap A' \to A[\mathfrak{p}] \to B \to 0$$

which shows that the semisimplification of $A[\mathfrak{p}]$ is indeed isomorphic to the sum of two characters θ_1 and θ_2. Since $\theta_1|I_v$ is cyclotomic and $\theta_2|I_v$ is trivial, the two characters are distinct. □

We are now ready to state the main theorem of this section.

Proposition 1.3. *Let K be a totally real number field, and let*

$$\bar{\rho} : \mathrm{Gal}(\bar{K}/K) \to \mathrm{GL}_2(\mathbb{F}_9)$$

be a Galois representation such that $\det \bar{\rho} = \bar{\chi}_3$. *Suppose that*

- *The absolute ramification degree of K is odd at every prime of K above* 3 *and* 5.

- *For any prime w of K over* 3, *the restriction of $\bar{\rho}$ to the decomposition group D_w is*
$$\bar{\rho}|D_w \cong \begin{bmatrix} \bar{\chi}_3 & * \\ 0 & 1 \end{bmatrix}.$$

- *The image of the inertia group I_v in* $\mathrm{GL}_2(\mathbb{F}_9)$ *has odd order for every prime v of K over* 5.

Then there exists a totally real number field F with F/K a solvable Galois extension, *and a Hilbert-Blumenthal abelian variety A/F with real multiplication by $\mathcal{O} = \mathcal{O}_{\mathbb{Q}[\sqrt{5}]}$, such that*

- *The absolute ramification degree of F is odd at every prime of F over* 3 *and* 5;

- *A has multiplicative reduction at all primes of F above* 3, *and good ordinary or multiplicative reduction at all primes of F above* 5;

- *The* mod $\sqrt{5}$ *representation*
$$\bar{\rho}_{A,\sqrt{5}} : \mathrm{Gal}(\bar{F}/F) \to \mathrm{GL}_2(\mathbb{F}_5)$$
is surjective;

- *There exists a symplectic isomorphism of* $\mathrm{Gal}(\bar{F}/F)$*-modules*
$$\iota : A[3] \cong \bar{\rho}|\,\mathrm{Gal}(\bar{F}/F).$$

2. Proof of Proposition 1.3

In order to produce Hilbert-Blumenthal abelian varieties, we will produce rational points on certain moduli spaces. Our main tool is an explicit description of the complex moduli space of HBAV's with real multiplication by $\mathcal{O} = \mathcal{O}_{\mathbb{Q}[\sqrt{5}]}$ and full 3-level structure, worked out by Hirzebruch and van der Geer. For the rest of this paper, an HBAV over a base S will be understood to mean a triple (A, m, λ), where

- A/S is an abelian surface;

- $m : \mathcal{O} \hookrightarrow \mathrm{End}(A)$ is an injection such that $\mathrm{Lie}(A/S)$ is, locally on S, a free $\mathcal{O} \otimes_{\mathbb{Z}} \mathcal{O}_S$ module (*the Rapoport condition*);

- λ is a principal polarization.

See [14] for basic properties of this definition.

2.1. *Twisted Hilbert modular varieties.* We first describe some twisted versions of the moduli space of HBAV's with full level 3 structure.

Suppose $\bar{\rho} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_9)$ is a Galois representation with cyclotomic determinant. Let N be the product of the ramified primes of $\bar{\rho}$. We also denote by $\bar{\rho}$ the $\mathcal{O}$-module scheme over $\mathbb{Z}[1/N]$ associated to the Galois representation.

Choose for all time an isomorphism $\eta : \wedge^2 \bar{\rho} \cong \mu_3 \otimes_{\mathbb{Z}} \mathcal{O}$. Now suppose A is an HBAV with real multiplication by $\mathcal{O}$ over a scheme T, and suppose A is endowed with an isomorphism $\phi : A[3] \cong \bar{\rho}$. Then Weil pairing gives an isomorphism $\wedge^2 A[3] \cong \mu_3 \otimes_{\mathbb{Z}} \mathcal{O}$. Now composing $\wedge^2 \phi$ with Weil pairing and with η yields an automorphism of $\mu_3 \otimes_{\mathbb{Z}} \mathcal{O}$. If this automorphism is the identity, we say ϕ has determinant 1. If this automorphism is obtained by tensoring an automorphism of μ_3 with $\mathcal{O}$, we say ϕ has integral determinant.

We define functors $\tilde{F}^{\bar{\rho}}$ and $F^{\bar{\rho}}$ from $\mathbf{Sch}/\mathbb{Z}[1/N]$ to $\mathbf{Sets}$ as follows:

> $\tilde{F}^{\bar{\rho}}(T)$ = isomorphism classes of pairs (A, ϕ), where A/T is a principally polarized Hilbert-Blumenthal abelian variety with RM by $\mathcal{O}$ and $\phi : A[3] \xrightarrow{\sim} \bar{\rho}$ is an isomorphism of $\mathcal{O}$-module schemes over T, with integral determinant.

and

> $F^{\bar{\rho}}(T)$ = isomorphism classes of pairs (A, ϕ), where A/T is a principally polarized Hilbert-Blumenthal abelian variety with RM by $\mathcal{O}$ and $\phi : A[3] \xrightarrow{\sim} \bar{\rho}$ is an isomorphism of $\mathcal{O}$-module schemes over T, with determinant 1.

Proposition 2.1. *The functor $\tilde{F}^{\bar{\rho}}$ is represented by a smooth scheme $\tilde{X}^{\bar{\rho}}$ over* $\mathrm{Spec}\,\mathbb{Z}[1/N]$. *The functor $F^{\bar{\rho}}$ is represented by a smooth geometrically connected scheme $X^{\bar{\rho}}$ over* $\mathrm{Spec}\,\mathbb{Z}[1/N]$.

Proof. We begin by observing that $\tilde{F}^{\bar{\rho}}$ is an étale sheaf on $\mathbf{Sch}/\mathbb{Z}[1/N]$. This follows exactly as in [4, Th. 2]; the key points are, first, that level 3 structure on HBAV's is rigid, and, second, that HBAV's are projective varieties and thus have effective descent.

For the first statement of the proposition, it now suffices to show that $\tilde{F}^{\bar{\rho}} \times_{\mathrm{Spec}\,\mathbb{Z}[1/N]} \mathcal{O}_L[1/N]$ is represented by a scheme, where L is a finite extension of $\mathbb{Q}$ unramified away from N. In particular, we may take L to be the fixed field of $\ker \bar{\rho}$. Then $\tilde{F}^{\bar{\rho}} \times_{\mathrm{Spec}\,\mathbb{Z}[1/N]} \mathcal{O}_L[1/N]$ is isomorphic to the functor $\tilde{F}$ parametrizing principally polarized HBAV's A together with isomorphisms $A[3] \cong (\mathcal{O}/3\mathcal{O})^2$ with integral determinant. This functor is representable by a smooth quasi-projective scheme $\tilde{X}$ over $\mathrm{Spec}\,\mathbb{Z}[1/3]$ (cf. [14, Th. 1.22], [3, Th. 4.3.ix]).

Now the functor $\tilde{F}^{\bar{\rho}}$ admits a map to $\mathrm{Aut}(\mu_3) \cong \underline{(\mathbb{Z}/3\mathbb{Z})^*}$, by the rule $(A, \phi) \mapsto (\eta \circ \wedge^2 \phi)$. It is clear that $F^{\bar{\rho}}$ is the preimage under this map of $1 \in (\mathbb{Z}/3\mathbb{Z})^*$. By changing base to L and invoking Theorem 1.28 ii) and the discussion below Theorem 1.22 in [14], we see that $X^{\bar{\rho}}$ is geometrically connected. □

We will sometimes refer to $X_L^{\bar{\rho}}$ simply as X. The group $\mathrm{PSL}_2(\mathbb{F}_9)$ acts on X by means of its action on $(\mathcal{O}/3\mathcal{O})^2$. (Note that (A, ϕ) and $(A, -\phi)$ are identified in X.) One can define exactly as in [14, §6.3] a line bundle $\underline{\omega}$ on $X^{\bar{\rho}}$ which is invariant under the action of $\mathrm{PSL}_2(\mathbb{F}_9)$.

When R is a ring containing $\mathcal{O}_L[1/N]$, the sections of $\underline{\omega}^{\otimes k}$ on X_R are called Hilbert modular forms of weight k and level 3 over R; the space of Hilbert modular forms over $\mathbb{C}$ is in natural isomorphism with the analytically defined space of Hilbert modular forms of the same weight and level [14, Lemma 6.12].

Within the space $H^0(X_{\bar{\mathbb{Q}}}, \underline{\omega}^{\otimes 2})$ of weight 2 modular forms of level 3 over $\bar{\mathbb{Q}}$ there is a 5-dimensional space of cuspforms, which we call C. The automorphism group $\mathrm{PSL}_2(F_9)$ acts on C through one of its irreducible 5-dimensional representations. It is shown by Hirzebruch and van der Geer that this space of modular forms provides a birational embedding of X into $\mathbb{P}^5$. To be precise: fix for all time an isomorphism $\mathrm{PSL}_2(\mathbb{F}_9) \cong A_6$ such that

- A_6 acts on C through the 5-dimensional quotient of its permutation representation;
- $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ is sent to the double flip $(01)(23)$.
- The subgroup of upper triangular unipotent matrices is sent to the group generated by (014) and (235).

Let s_0 be a generator of the 1-dimensional subspace of C fixed by the stabilizer of a letter in A_6, and let $s_0, \ldots, s_5$ be the A_6-orbit of s_0. Note that $s_0 + \cdots + s_5 = 0$.

Proposition 2.2. *Let $S_{\mathbb{Z}}$ be the surface in $\mathbb{P}^5/\mathbb{Z}$ defined by the equations*

$$\sigma_1(s_0, \ldots, s_5) = \sigma_2(s_0, \ldots, s_5) = \sigma_4(s_0, \ldots, s_5) = 0,$$

where σ_i is the i^{th} symmetric polynomial. Note that $A_6 \cong \mathrm{PSL}_2(\mathbb{F}_9)$ acts on $S_{\mathbb{C}}$ by permutation of coordinates.

Then the map $X_{\mathbb{C}} \to \mathbb{P}^5_{\mathbb{C}}$ given by $[s_0 : s_1 : s_2 : s_3 : s_4 : s_5]$ factors through a birational isomorphism $X_{\mathbb{C}} \to S_{\mathbb{C}}$.

Proof. [22, VIII.(2.6)] □

Note that the map $X_{\mathbb{C}} \to S_{\mathbb{C}}$ is equivariant for the action of $\mathrm{PSL}_2(\mathbb{F}_9)$ on the left and A_6 on the right. The form $\sigma_k(s_0, \ldots, s_5)$ is invariant under $\mathrm{PSL}_2(\mathbb{F}_9)$, and is therefore a cusp form of level 1 and weight $2k$. Let τ be the involution of X induced from the Galois involution of $\mathcal{O}$ over $\mathbb{Z}$. We say a Hilbert modular form is *symmetric* if it is fixed by τ. By a result of Nagaoka [13, Th. 5.2], the ring $M_{2*}(\mathrm{SL}_2(\mathcal{O}), \mathbb{Z}[1/2])$ of even-weight level 1 symmetric modular forms over $\mathbb{Z}[1/2]$ is generated by forms ϕ_2, χ_6, and χ_{10} of weights $2, 6$, and 10. The form ϕ_2 is the weight 2 Eisenstein series, while χ_6 and χ_{10} are cuspforms. It follows that the ideal of cuspforms in $M_{2*}(\mathrm{SL}_2(\mathcal{O}), \mathbb{Z}[1/2])$ is generated by χ_6 and χ_{10}. One has from [22, VIII.2.4] that there is no nonsymmetric modular form of even weight less than 20. It follows that $\sigma_k(s_0, \ldots, s_5)$ can be expressed in terms of ϕ_2, χ_6, and χ_{10}. For simplicity, write σ_k for $\sigma_k(s_0, \ldots, s_5)$. Then by a series of computations on q-expansions, one has

$$
\begin{aligned}
\phi_2 &= -3\sigma_5^{-1}(\sigma_3^2 - 4\sigma_6), \\
\chi_6 &= \sigma_3, \\
\chi_{10} &= (-1/3)\sigma_5.
\end{aligned} \tag{2.1.1}
$$

The details can be found in the appendix.

(Note that the constants here depend on our original choice of the weight 2 forms s_i. Modifying that choice by a constant c would modify each formula above by $c^{k/2}$, where k is the weight of the modular form in the expression.)

We now show that the theorem of Hirzebruch and van der Geer above allows us to compute equations for birational models of $X^{\bar{\rho}}$ over $\mathbb{Q}$. Recall that $\mathrm{PSL}_2(\mathbb{F}_9)$ acts on $X_{\bar{\mathbb{Q}}}$; the action of $\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on $\mathrm{PSL}_2(\mathbb{F}_9) \subset \mathrm{Aut}(X_{\bar{\mathbb{Q}}})$ is conjugation by $\bar{\rho}(\sigma)$. Note that the image of $\bar{\rho}(\sigma)$ in $\mathrm{PGL}_2(\mathbb{F}_9)$ is actually contained in $\mathrm{PSL}_2(\mathbb{F}_9)$, since $\bar{\rho}$ has cyclotomic determinant.

In particular, the action of Galois on $\mathrm{PSL}_2(\mathbb{F}_9) \cong A_6$ permutes the six letter-stabilizing subgroups; thus it permutes the six lines $\bar{\mathbb{Q}}s_0, \ldots, \bar{\mathbb{Q}}s_5$ in $H^0(X_{\bar{\mathbb{Q}}}, \underline{\omega}^{\otimes 2})$, since each of these lines is the fixed space of a letter-stabilizing subgroup. The fact that $s_0 + \cdots + s_5 = 0$ implies that the action of Galois on the set $s_0, \ldots, s_5$ is the composition of a permutation with a scalar multiplication in $\bar{\mathbb{Q}}^*$. By Hilbert 90, we can multiply $s_0, \ldots, s_5$ by a scalar to ensure that σ permutes the six variables by means of the permutation in A_6 attached to the projectivization of $\bar{\rho}(\sigma)$.

Write $C_{\bar{\rho}}$ to denote the cuspidal subspace of $H^0(X^{\bar{\rho}}, \underline{\omega}^{\otimes 2})$. Then our determination of the action of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on the forms $s_0, \ldots, s_5$ suffices to determine the 5-dimensional $\mathbb{Q}$-vector space $C_{\bar{\rho}}$ as a subspace of $\bar{\mathbb{Q}}s_0 + \cdots + \bar{\mathbb{Q}}s_5$. Any basis $s'_0, \ldots, s'_4$ of $C_{\bar{\rho}}$ induces a birational embedding of $X^{\bar{\rho}}$ in $\mathbb{P}^4$, by Proposition 2.2; the image of this embedding is the intersection of a quadratic hypersurface $Q_2^{\bar{\rho}}$ and a quartic hypersurface $Q_4^{\bar{\rho}}$; here $Q_i^{\bar{\rho}}$ is the variety in the $\mathbb{P}^4$ with coordinates $s'_0, \ldots, s'_4$ defined by the vanishing of the degree-i form $\sigma_i(s_0, \ldots, s_5)$.

We will often make use of the following important example. Let $\bar{\rho}_0$ be the representation

$$\bar{\rho}_0 = \begin{bmatrix} \bar{\chi}_3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then the modular forms

$$\begin{aligned} x_0 &= \omega s_0 + \omega^2 s_1 + s_4, \\ x_1 &= \omega^2 s_0 + \omega s_1 + s_4, \\ x_4 &= s_0 + s_1 + s_4, \\ x_2 &= \omega s_2 + \omega^2 s_3 + s_5, \\ x_3 &= \omega^2 s_2 + \omega s_3 + s_5, \\ x_5 &= s_2 + s_3 + s_5 \end{aligned}$$

lie in $H^0(X^{\bar{\rho}_0}, \underline{\omega}^{\otimes 2})$. This coordinate system yields a map $X^{\bar{\rho}_0} \to \mathbb{P}^5$, which is a birational isomorphism between $X^{\bar{\rho}_0}$ and the intersection of the three hypersurfaces

$$\begin{aligned} Q_1^{\bar{\rho}_0} &= V(x_4 + x_5), \\ Q_2^{\bar{\rho}_0} &= V(x_4^2 + x_5^2 - x_0x_1 - x_2x_3 + 3x_4x_5), \\ Q_4^{\bar{\rho}_0} &= V(-3x_0x_1x_4x_5 - 3x_2x_3x_4x_5 + 3x_0x_1x_2x_3 + x_4x_5(x_4^2 + 3x_4x_5 + x_5^2) \\ &\qquad -3x_0x_1x_5^2 - 3x_2x_3x_4^2 + x_0^3x_5 + x_1^3x_5 + x_2^3x_4 + x_3^3x_4). \end{aligned}$$

In this case, symmetry considerations lead us to think of $S^{\bar{\rho}}$ as contained in a $\mathbb{Q}$-rational hyperplane in $\mathbb{P}^5$, as opposed to placing $S^{\bar{\rho}}$ directly into $\mathbb{P}^4$.

Our overall strategy is as follows. To prove Proposition 1.3, we will need to find a point on a twisted Hilbert modular variety $X^{\bar{\rho}}$ defined over a solvable extension of K. The geometric observation allowing us to produce such points is the following.

Let L/K be a line contained in the variety $Q_2^{\bar{\rho}}$. Then $L \cap Q_4^{\bar{\rho}}$ is a 0-dimensional subscheme Σ of degree 4 in $S^{\bar{\rho}}$. Generically, Σ will split into four distinct points over a degree 4 (whence solvable!) extension of K. Now $Q_2^{\bar{\rho}}$ is a quadric hypersurface in $\mathbb{P}^4$, so its Fano variety is rational. This means we have plenty of lines in $Q_2^{\bar{\rho}}$, whence plenty of points in $S^{\bar{\rho}}$ defined over solvable extensions of K. What remains is to make sure we can find such points which satisfy the local conditions at $3, 5$, and ∞ required in the proposition. Our strategy will be to define suitable lines over completions of K at the relevant primes, and finally to use strong approximation on the Fano variety $F_1(Q_2^{\bar{\rho}})$ to find a global line which is adelically close to the specified local ones.

2.2. *Archimedean primes.* Let c be a complex conjugation in $\mathrm{Gal}(\bar{K}/K)$, and let u be the corresponding real place of K.

The fact that $\bar{\rho}$ is odd implies that $\bar{\rho}(c)$ is conjugate to

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

In particular, we have

$$\bar{\rho}_0 | \mathrm{Gal}(\mathbb{C}/K_u) \cong \bar{\rho} | \mathrm{Gal}(\mathbb{C}/K_u),$$

whence

$$S^{\bar{\rho}} \times_K K_u \cong S^{\bar{\rho}_0} \times_{\mathbb{Q}} K_u = S^{\bar{\rho}_0} \times_{\mathbb{Q}} \mathbb{R}.$$

If $s_0, \ldots, s_5$ are our standard coordinates on S, we may take $x_0, \ldots, x_5$ as coordinates on $S^{\bar{\rho}_0}_{\mathbb{Q}}$ as in the previous section. Now choose a real line $L_{\mathbb{R}}$ in $F_1(Q_1^{\bar{\rho}_0} \cap Q_2^{\bar{\rho}_0})(\mathbb{R})$ with the property that $L_{\mathbb{R}} \cap S^{\bar{\rho}_0}$ consists of four distinct real points. For instance, we may choose $L_{\mathbb{R}}$ to be the line

$$\begin{aligned}&(x_0, x_1, x_2, x_3, x_4, x_5)\\ &= \left(\frac{8}{15}u - \frac{4}{3}t, -\frac{82}{15}u - \frac{4}{3}t, -\frac{4}{3}u - \frac{8}{3}t, -\frac{4}{3}u + \frac{10}{3}t, -\frac{16}{15}u + \frac{8}{3}t, \frac{16}{15}u - \frac{8}{3}t\right).\end{aligned}$$

Let L_u be the corresponding line in $F_1(Q_1^{\bar{\rho}} \cap Q_2^{\bar{\rho}})(K_u)$.

2.3. *Primes above* 5. Let K_v be the completion of K at a prime v dividing 5, and let E_v^0 be the splitting field of $\bar{\rho}|G_{K_v}$. Note that, by hypothesis, E_v^0 has odd absolute ramification degree.

As above, our aim is to find a suitable line in $Q_2^{\bar{\rho}}$ over some unramified extension of E_v^0. Since $\bar{\rho}$ is trivial on $\mathrm{Gal}(\bar{\mathbb{Q}}_5/E_v^0)$, the morphism $X^{\bar{\rho}} \to S$ is defined over E_v^0. Write Q_i for the hypersurface $\sigma_i(s_0, \ldots, s_5) = 0$, where $i = 1, 2, 4$. So $S = Q_1 \cap Q_2 \cap Q_4$, and we are looking for lines on $Q_1 \cap Q_2$. Denote by U an open dense subvariety of S which is isomorphic to an open dense subvariety of $X^{\bar{\rho}}$. Write Z for the complement of U in S.

LEMMA 2.3. *There exists a finite unramified extension E_v of E_v^0 and a line L_v/E_v contained in $Q_1 \cap Q_2/E_v$ such that*

- *L_v is disjoint from Z;*
- *$(L_v \cap Q_4)(E_v)$ consists of 4 distinct E_v-points;*
- *For each $x \in (L_v \cap Q_4)(E_v)$, the functions*

  $$\sigma_5^{-6}(\sigma_3^2 - 4\sigma_6)^5$$

  and

  $$\sigma_5^{-3}\sigma_3^{-1}(\sigma_3^2 - 4\sigma_6)^3$$

  have nonpositive valuation when evaluated at x.

Proof. One checks that $Q_1 \cap Q_2$ is isomorphic over $\mathbb{Z}_5^{unr}$ to the Plücker quadric threefold

$$T := V(y_0y_1 + y_2y_3 + y_4^2) \subset \mathbb{P}^4.$$

We also know (see [8, §6, Ex. 22.6]) an explicit 3-parameter family of lines on T, which is to say a map

$$\lambda : \mathbb{P}^3/\operatorname{Spec}\mathbb{Z}_5 \to F_1(T);$$

moreover, λ is an isomorphism over any algebraically closed field. Composing λ with an isomorphism between T and $Q_1 \cap Q_2$ yields a map

$$L : \mathbb{P}^3/\operatorname{Spec}\mathbb{Z}_5^{unr} \to F_1(Q_1 \cap Q_2)$$

which is an isomorphism over any algebraically closed field.

The set of $\bar{\mathbf{p}} \in \mathbb{P}^3(\bar{\mathbb{F}}_5)$ such that $L(\bar{\mathbf{p}}) \cap Q_4/\bar{\mathbb{F}}_5$ consists of four distinct $\bar{\mathbb{F}}_5$-points is Zariski-open. To check that it is not empty, we need only exhibit a single such line L in $(Q_1 \cap Q_2)/\bar{\mathbb{F}}_5$. One such line is

$$\begin{aligned}&(s_0, s_1, s_2, s_3, s_4, s_5)\\ &= ((1-\sqrt{-3})t, (1+\sqrt{-3})t, -t+(1+\sqrt{-3})u, -t+(1-\sqrt{-3})u, t, -t-2u).\end{aligned}$$

One checks that the restriction of Q_4 to L is $-3t(8u^3 - t^3)$, which indeed has four distinct roots over $\bar{\mathbb{F}}_5$.

Let V be the closed subscheme of $S/\bar{\mathbb{F}}_5$ where the form $\sigma_3^2 - 4\sigma_6$ vanishes. Then V is a curve. Moreover, if x is a point in $S/\bar{\mathbb{F}}_5$, the subscheme of $\mathbb{P}^3/\bar{\mathbb{F}}_5$ parametrizing lines passing through x is one-dimensional. So the subscheme of $\mathbb{P}^3/\bar{\mathbb{F}}_5$ parametrizing lines intersecting V is at most two-dimensional. We may thus choose a point $\bar{\mathbf{p}} \in \mathbb{P}^3(\bar{\mathbb{F}}_5)$ such that $L(\bar{\mathbf{p}}) \cap Q_4/\bar{\mathbb{F}}_5$ consists of four distinct $\bar{\mathbb{F}}_5$-points, none contained in V.

Now let $\mathbf{p}$ be a lift of $\bar{\mathbf{p}}$ to $\mathbb{P}^3(\mathbb{Q}_5^{nr})$. Then $L(\mathbf{p})$ is a line contained in $Q_1 \cap Q_2$ whose intersection with Q_4 consists of four distinct points defined over some unramified extension of $\mathbb{Q}_5$. Let E_v be the compositum of this extension with E_v^0. Since Z is at most one-dimensional, we may choose $\mathbf{p}$ such that $L(\mathbf{p}) \cap Q_4$ is disjoint from Z, by the same argument as above.

Let x be a point in $L(\mathbf{p}) \cap Q_4(E_v)$, and choose integral coordinates for x so that at least one coordinate has nonpositive valuation. Then $(\sigma_3^2 - 4\sigma_6)(x)$ has nonpositive valuation, so that the third desired condition on $L(\mathbf{p})$ is satisfied. This completes the proof. $\square$

Now take L_v and E_v as in the lemma. Let x_1, x_2, x_3, x_4 be the four E_v-points making up $(L_v \cap S)(E_v)$. Then each x_i corresponds to an abelian variety A_i/E_v with real multiplication by $\mathcal{O}$ admitting an isomorphism $A[3] \cong \bar{\rho} \cong \mathbb{F}_9^{\oplus 2}$ of $\mathcal{O}$-module schemes over E. It follows that A_i has semistable reduction over $\mathcal{O}_E$, since no nontrivial finite-order element of $\mathrm{GL}_2(\mathbb{Z}_3(\sqrt{5}))$ is congruent to 1 mod 3.

We now want to show that each A_i has good ordinary or multiplicative reduction. We have computed above that the weight 2 modular form ϕ_2 can be written as $-3\sigma_5^{-1}(\sigma_3^2 - 4\sigma_6)$. Therefore, our choice of L_v guarantees that the modular functions ϕ_2^3/χ_6 and ϕ_2^5/χ_{10} have nonpositive valuation when evaluated on A_i. The desired ordinarity now follows from the next lemma.

LEMMA 2.4. *Let A be a semi-*HBAV *over a finite extension $\mathcal{O}_E/\mathbb{Z}_5$. Suppose that the modular functions ϕ_2^3/χ_6 and ϕ_2^5/χ_{10} evaluated at A have nonpositive valuation. Then A has good ordinary or multiplicative reduction.*

Proof. Let Ω be the determinant of the pullback via the identity section of the relative cotangent sheaf of $A/\mathcal{O}_E$. Then Ω is a free rank 1 $\mathcal{O}_E$-module. Let η be a section generating Ω. Then every modular form f with coefficients in $\mathcal{O}_E$ has a well-defined value $f(A,\eta)$. Suppose $\phi_2(A,\eta) \in \mathfrak{m}_E$. Then by the hypothesis of the theorem, we have also that $\chi_6(A,\eta)$ and $\chi_{10}(A,\eta) \in \mathfrak{m}_E$.

The involution τ preserves integrality, by the q-expansion principle. It follows that every modular form f over $\mathcal{O}_E$ is integral over the ring of symmetric even-weight modular forms studied by Nagaoka. In particular, since ϕ_2, χ_6, and χ_{10} generate this ring, we have that $f(A,\eta) \in \mathfrak{m}_E$ for all symmetric modular forms f of positive even weight. But this is impossible, since for any sufficiently large k the sheaf $\underline{\omega}^{\otimes k}$ on X is generated by its global sections [3, 4.3(iii)].

We conclude that $\phi_2(A,\eta) \notin \mathfrak{m}_E$. So the mod 5 reduction $\phi_2(\bar{A},\bar{\eta})$ is not equal to 0.

The q-expansion of ϕ_2^2 reduces to 1 (mod 5) [13, (5.12)]. By [1, 7.12,7.14], the Hasse invariant h is a weight 4 modular form which also has q-expansion equal to 1; it follows that h is the reduction mod 5 of ϕ_2^2. So $h(A,\eta) \neq 0$. But this implies that A has good ordinary or multiplicative reduction by [1, 7.14.2]. $\square$

2.4. *Primes above* 3. This section will be the most technically complicated part of the paper, owing to the fact that we do not have at our disposal a good model for $X^{\bar{\rho}}$ in characteristic 3.

Let w be a prime of K dividing 3, and let K_w be the completion of K at w. We have by hypothesis that

$$\bar{\rho}|D_w \cong \begin{bmatrix} \bar{\chi}_3 & * \\ 0 & 1 \end{bmatrix}. \tag{2.4.2}$$

Now the $*$ in (2.4.2) is a cocycle corresponding to an element $\lambda \in K_w^* \otimes_{\mathbb{Z}} \mathbb{F}_9$. Write $\bar{\rho}_\lambda$ for the representation of D_w on the right-hand side of (2.4.2), which is isomorphic to $\bar{\rho}|D_w$. As in the 5-adic case, let Z be a proper closed subscheme of $S^{\bar{\rho}_\lambda}/K_w$ such that the complement of Z is isomorphic to an open dense subset of $X^{\bar{\rho}_\lambda}/K_w$.

LEMMA 2.5. *There exists a line L_w in $\mathbb{P}^5_{K_w}$ satisfying the following conditions*:

- *L_w is disjoint from Z.*
- *L_w is contained in $Q_1^{\bar\rho_\lambda} \cap Q_2^{\bar\rho_\lambda}$.*
- *The intersection $L_w \cap Q_4^{\bar\rho_\lambda}$ splits into four distinct points over an unramified extension E_w of K_w.*
- *The four* HBAV'*s A_1, A_2, A_3, A_4 corresponding to the four points of $L_w \cap Q_4^{\bar\rho_\lambda}(E_w)$ have multiplicative reduction.*

Proof. We first remark that the truth of the lemma depends only on the isomorphism class of $\bar\rho_\lambda$; in particular, the conclusion of the lemma also holds for $\bar\rho|D_w$, whose image might lie in a Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_9)$ other than the upper triangular one discussed here.

By means of our chosen isomorphism between $\mathrm{PSL}_2(\mathbb{F}_9)$ and A_6, we interpret $*$ as a cocycle from D_w to the group G generated by the 3-cycles (014) and (235). Each 3-cycle generates a cyclic factor of G, and the projection of $*$ onto the cyclic factor yields a cocycle in $H^1(D_w, \mu_3)$. Kummer theory attaches to each of the resulting cocycles an element of $K_w^*/(K_w^*)^3$; we call these elements λ_1 and λ_2. It is easy to check that the forms

$$\begin{aligned}
y_0 &= (\lambda_1)^{-1/3}(\omega s_0 + \omega^2 s_1 + s_4),\\
y_1 &= (\lambda_1)^{1/3}(\omega^2 s_0 + \omega s_1 + s_4),\\
y_4 &= s_0 + s_1 + s_4,\\
y_2 &= (\lambda_2)^{-1/3}(\omega s_2 + \omega^2 s_3 + s_5),\\
y_3 &= (\lambda_2)^{1/3}(\omega^2 s_2 + \omega s_3 + s_5)\\
y_5 &= s_2 + s_3 + s_5
\end{aligned}$$

lie in $H^0(X^{\bar\rho_\lambda}, \underline{\omega}^{\otimes 2})$.

With these coordinates, one checks that $Q_1^{\bar\rho_\lambda}$ is defined by $y_4 + y_5$ and $Q_2^{\bar\rho_\lambda}$ by

$$y_4^2 + y_5^2 - y_0y_1 - y_2y_3 + 3y_4y_5.$$

So a family of lines in $Q_1^{\bar\rho_\lambda} \cap Q_2^{\bar\rho_\lambda}$ is given by

$$L_{a,b,c} : y_0 = ay_2 + by_4, y_3 = -ay_1 + cy_4, y_4 = -(by_1 + cy_2), y_5 = -y_4.$$

One checks that the equation for $Q_4^{\bar\rho_\lambda}$ is given by

$$\begin{aligned}
-3y_0y_1y_4y_5 - 3y_2y_3y_4y_5 + 3y_0y_1y_2y_3 + y_4y_5(y_4^2 + 3y_4y_5 + y_5^2)&\\
- 3y_0y_1y_5^2 - 3y_2y_3y_4^2 + \lambda_1 y_0^3 y_5 + \lambda_1^{-1} y_1^3 y_5 + \lambda_2 y_2^3 y_4 + \lambda_2^{-1} y_3^3 y_4.&
\end{aligned}$$

Since λ_1 and λ_2 are defined only up to cubes, we may assume that both have even valuation.

The equation for $Q_4^{\bar\rho_\lambda}$ restricted to $L_{a,b,c}$ is of the form

$$P = \sum_{i=0}^{4} P_i(a,b,c) y_1^i y_2^{4-i}.$$

Suppose that $\mathrm{ord}_w(b)$ and $\mathrm{ord}_w(c)$ are approximately equal and that both are much greater than $\mathrm{ord}_w(a)$, which is in turn much greater than 0. Then one checks that

$$\begin{aligned}
P_4(a,b,c) &= \lambda_1^{-1} b + \text{higher order terms},\\
P_3(a,b,c) &= \lambda_1^{-1} c + \text{higher order terms},\\
P_2(a,b,c) &= -3a^2 + \text{higher order terms},\\
P_1(a,b,c) &= -\lambda_2 b + \text{higher order terms},\\
P_0(a,b,c) &= -\lambda_2 c + \text{higher order terms}.
\end{aligned}$$

It follows that the vanishing locus of P in the projective line with coordinates y_1 and y_2 consists of two points reducing to $[0:1]$ and two reducing to $[1:0]$. So P factors over K_w^{unr} into a constant and two quadratics:

$$P = -3a^2(e_1 y_1^2 + e_2 y_1 y_2 + e_3 y_2^2)(f_1 y_1^2 + f_2 y_1 y_2 + f_3 y_2^2)$$

where e_3 and f_1 are units. One checks that $\mathrm{ord}_w(e_1) = \mathrm{ord}_w(b) + 1 \pmod 2$ and $\mathrm{ord}_w(f_3) = \mathrm{ord}_w(c) + 1 \pmod 2$, and that

$$\begin{aligned}
\mathrm{ord}_w(e_2) &\geq \min(\mathrm{ord}_w(b), \mathrm{ord}_w(c)) + \mathrm{ord}_w(\lambda_1^{-1}/3a^2),\\
\mathrm{ord}_w(f_2) &\geq \min(\mathrm{ord}_w(b), \mathrm{ord}_w(c)) + \mathrm{ord}_w(\lambda_2/3a^2).
\end{aligned}$$

So when b and c have odd valuation, the two quadratic factors of P split over K_w^{nr}. In other words, the four points of $L_{a,b,c} \cap Q_4^{\bar\rho_\lambda}$ are distinct and defined over an unramified extension E_w of K_w. Since Z is at most 1-dimensional, we may choose a, b, c such that $L_{a,b,c}$ is disjoint from Z, as in the previous section.

We now show that the HBAV's parametrized by $L_{a,b,c} \cap Q_4^{\bar\rho_\lambda}$ have potentially multiplicative reduction.

The points of $L_{a,b,c} \cap Q_4^{\bar\rho_\lambda}$ are w-adically close to $[0:1:0:0:0:0]$ and $[0:0:1:0:0:0]$. In coordinates $[s_0 : \cdots : s_5]$, these points are $[\omega : \omega^2 : 0 : 0 : 1 : 0]$ and $[0 : 0 : \omega^2 : \omega : 0 : 1]$. At each point, the symmetric functions σ_k in $s_0, \dots, s_5$ are w-adically close to 0 for $k = 5, 6$, while σ_3 is close to 1.

A technical complication arises here: we would like to say that if a point is w-adically close to a cusp of X, the corresponding HBAV has potentially multiplicative reduction. The right way to proceed would be to make use of a modular interpretation of a formal neighborhood of a cusp in some good model of X over $\mathcal{O}_{K_w}$. Since such an interpretation is not yet in the published literature, we resort to a less pleasant argument on level 1 modular forms.

Let A/E_w be an HBAV associated to a point of $L_{a,b,c} \cap Q_4^{\bar{\rho}_\lambda}$. Then from (2.1.1) we see that the weight 0 modular functions χ_6/ϕ_2^3 and χ_{10}/ϕ_2^5 take values at A which are w-adically close to 0.

Let q be some large prime inert in $\mathcal{O}$, let k_w be the residue field of K_w, and let $X(q)/\bar{k}_w$ be the proper Hilbert modular variety parametrizing generalized HBAV's with full level q structure. (The auxiliary prime q is introduced only so that we can speak of schemes instead of stacks.) Passing to a ramified extension if necessary, take $A_0/\bar{k}_w$ to be the reduction of a semistable model of A. We want to show that A_0 is not smooth. Let $\phi : A_0[q] \cong (\mathcal{O}/q\mathcal{O})^2$ be an arbitrary choice of level structure.

Define Ω, η as in the proof of Lemma 2.4. As in that proof, we know that $\phi_2(A_0, \eta), \chi_6(A_0, \eta)$, and $\chi_{10}(A_0, \eta)$ do not all vanish. It follows from (2.1.1) that $\chi_6(A_0, \eta) = \chi_{10}(A_0, \eta) = 0$, while $\phi_2(A_0, \eta) \neq 0$.

By [3, 4.3(x)], the ideal of cusp forms of level k (defined holomorphically, in terms of q-expansions) is the same as the algebraically defined ideal of forms vanishing at the cusps. So if the point of $X(q)$ parametrized by (A_0, ϕ) were not a cusp, there would be a cusp form f of level q such that $f(A_0, \phi, \eta) \neq 0$. By squaring if necessary, we may assume f has even weight. Let $f_1, \ldots, f_r$ be the set of images of f under the action of the involution τ and the group $\mathrm{PGL}_2(\mathcal{O}/q\mathcal{O})$. Every symmetric function in the $f_i(A_0, \phi, \eta)$ is a symmetric level 1 cusp form of even weight. Since the ideal of cusp forms in the ring of symmetric modular forms of even weight is generated by χ_6 and χ_{10}, we have shown that $f(A_0, \phi, \eta) = 0$.

We conclude that the point parametrizing (A_0, ϕ) is a cusp, which is to say that A has potentially multiplicative reduction, as desired. $\square$

It remains to prove that A has semistable reduction.

LEMMA 2.6. *Let p_1 and p_2 be the points of $S^{\bar{\rho}_\lambda}(E_w)$ whose coordinates $[s_0 : \ldots s_5]$ are $[\omega : \omega^2 : 0 : 0 : 1 : 0]$ and $[0 : 0 : \omega^2 : \omega : 0 : 1]$ respectively. Then there exist w-adic neighborhoods U_1 and U_2 of p_1 and p_2 with the following property.*

Let (A, ϕ) be an HBAV over E_w endowed with a 3-level structure $\phi : A[3] \xrightarrow{\sim} \bar{\rho}_\lambda$, such that (A, ϕ) is parametrized by a point of U_1 or U_2. Then A has multiplicative reduction.

Proof. We have already shown A has potentially multiplicative reduction. It follows that either A or its twist by a quadratic ramified character is semistable. So $A[3]$ has a well-defined "canonical subgroup" G which, over $\bar{\mathbb{Q}}_3$, is the subgroup obtained by pulling back $\mu_3 \otimes_{\mathbb{Z}} \mathfrak{d}^{-1}$ from the Tate uniformization

$$A \cong (\mathbb{G}_m \otimes \mathfrak{d}^{-1}/q^{\mathcal{O}}).$$

The reduction of A is multiplicative if and only if G is the subgroup of $\bar{\rho}_\lambda$ on which Galois acts cyclotomically. (Note that this condition is automatic unless λ is trivial.)

Denote by **Tate** a semi-abelian variety attached to a compactification of the level-1 Hilbert modular surface [3, 3.5]. We call a 3-level structure

$$\mathbf{Tate}[3] \cong (\mathcal{O}/3\mathcal{O})^2$$

canonical if it attaches the canonical subgroup of **Tate** to the first coordinate of $(\mathcal{O}/3\mathcal{O})^2$. We say a cusp of X is canonical if the associated level structure on **Tate** is canonical. Note that the canonical cusp is the unique one which is fixed by the action of the upper triangular matrices in $\mathrm{PSL}_2(\mathbb{F}_9)$. Moreover, A has multiplicative reduction precisely when (A, ϕ) is a pullback of **Tate**, supplied with a canonical 3-level structure.

The condition above is geometric, so from now on we consider X to be defined over $\bar{\mathbb{Q}}_3$.

Choose a distinct value α_C in $\bar{\mathbb{Q}}_3$ for each cusp C of X. Since $\underline{\omega}^{\otimes 2k}$ is very ample on the Hilbert modular surface X for k large enough, there exists a modular form $E \in H^0(X, \underline{\omega}^{\otimes 2k})$ which takes the value α_C at each C [3, 4.5.1b]. In other words, the q-expansion of E at C has constant term α_C.

Let $E = E_0, \ldots, E_r$ be the images of E under $\mathrm{PSL}_2(\mathbb{F}_9)$ and the involution τ. Let t be an indeterminate and define

$$F = \prod_j (t - E_j).$$

Then the coefficient of t^i in F is a symmetric modular form f_i of level 1 and weight $2ki$, and is therefore a polynomial of weight $2ki$ in ϕ_2, χ_6, and χ_{10}. From (2.1.1) we now see that f_i/ϕ_2^{ki} lies in the ring $R = \bar{\mathbb{Q}}_3[\sigma_3, \sigma_5, \sigma_6, (\sigma_3^2 - 4\sigma_6)^{-1}]$. So E/ϕ_2^k is integral over R.

The surface S has only the 30 points in the S_6-orbit of $[1 : \omega : \omega^2 : 1 : \omega : \omega^2]$ as singularities, so that S is normal, whence projectively normal as a subscheme of $\mathbb{P}^5$. We can therefore write

$$\tilde{E} = E/\phi_2^k = (\sigma_3^2 - 4\sigma_6)^{-n} P(s_0, \ldots, s_5)$$

for some homogeneous polynomial P.

Now $\tilde{E}$ has weight 0, so for each γ in $\mathrm{PSL}_2(\mathbb{F}_9)$ the value $\tilde{E}(A, \gamma \circ \phi)$ is well-defined. Let U be a small w-adic neighborhood of p_1 or p_2. We may assume U is preserved by the action of the upper triangular matrices in $\mathrm{PSL}_2(\mathbb{F}_9)$, since p_1 and p_2 are fixed points for this action.

The function $\tilde{E}$ is continuous on U, since $\sigma_3^2 - 4\sigma_6$ has no zeroes on U. In particular, the values $\tilde{E}(A, \beta \circ \phi)$ all lie in a small w-adic neighborhood, where β ranges over upper triangular matrices. On the other hand, if (A, ϕ') is pulled back from some cusp C of X, the value of $\tilde{E}(A, \phi')$ can be computed by substituting a value of q into the q-expansions at C of E and ϕ_2. Moreover, we can ensure that $\mathrm{ord}_w(q)$ is large by making U sufficiently small; for instance, by choosing U very small we can ensure that the values of

$$\chi_6/\phi_2^3 = -3^{-3}\sigma_5^3\sigma_3(\sigma_3^2 - 4\sigma_6)^{-3}$$

and

$$\chi_{10}/\phi_2^5 = 3^{-6}\sigma_5^6(\sigma_3^2 - 4\sigma_6)^{-5}$$

lie in a small w-adic neighborhood whenever u is in U, simply by virtue of the fact that we can force σ_5 to be as small as we like by shrinking U. Then the value of $\mathrm{ord}_w(q)$ can be computed by means of [7, Prop. 2.22]; in particular, when χ_6/ϕ_2^3 and χ_{10}/ϕ_2^5 are both w-adically close to 0, so is q. It follows that $\tilde{E}(A, \phi')$ lies in a small w-adic neighborhood of α_C, and this determines C.

So the points $(A, \beta \circ \phi)$ are all pulled back from the same cusp C of X. Thus, C is fixed by upper triangular matrices, and must be the canonical cusp. $\square$

2.5. *The global construction.* We now combine the local arguments above into the global statement we desire, thereby completing the proof of Proposition 1.3.

Choose a finite Galois extension K'/K such that;

- K' is totally real;
- K'/K is solvable;
- The completion of K' at any prime v above 5 is isomorphic to an unramified extension of E_v;
- The completion of K' at any prime w above 3 is isomorphic to an unramified extension of K_w;
- $Y = F_1(Q_1^{\bar\rho} \cap Q_2^{\bar\rho})$ is rational over K'. (Since $F_1(Q_1^{\bar\rho} \cap Q_2^{\bar\rho})$ is geometrically rational, this amounts to trivializing an element of the Brauer group; the existence of L_u tells us that this element is already trivial at every real place, so it can be killed by a totally real solvable extension.)

(See [21, Lemma 2.2] for the existence of K'.)

From now on, write Y for $F_1(Q_1^{\bar\rho} \cap Q_2^{\bar\rho})$.

Since Y is a rational variety, we can choose $L \in Y(K')$ such that the image of L under the map

$$Y(K') \to \bigoplus_{v_i|5} Y(K'_{v_i}) \oplus \bigoplus_{w_i|3} Y(K'_{w_i}) \oplus \bigoplus_{u|\infty} Y(K'_u)$$

is arbitrarily adelically close to $(L_{v_1}, \dots, L_{w_1}, \dots, L_{u_1}, \dots)$.

The intersection $L \cap S^{\bar\rho}$ is a zero-dimensional scheme of degree 4 over K'. Modifying our choice of L if necessary, we can arrange for $L \cap S^{\bar\rho}$ to be in the image of the rational map from $X^{\bar\rho}$. Let F be a splitting field for $L \cap S^{\bar\rho}$. Note that F is solvable over K', whence also over K. Then we can think of $L \cap S^{\bar\rho}$ as specifying four HBAV's A_i/F, with $A_i[3] \cong \bar\rho|F$.

Let T be the subvariety of $S^{\bar\rho} \times Y$ consisting of pairs (x, L) such that x is contained in L. Then the projection map $\pi_2 : T \to Y$ is generically a 4-fold cover. Let π_1 be the projection $T \to S^{\bar\rho}$. Let $\mathcal{A} \xrightarrow{p} V$ be the universal object over a suitable dense open subscheme V of $S^{\bar\rho}$, and let $\mathcal{F}_\ell$ be the l-adic sheaf $R^1 p_* \mathbb{Z}_\ell$ on V. Finally, define $\mathcal{G}_\ell = \pi_{2;*}\pi_1^* \mathcal{F}_\ell$. Then $\mathcal{G}_\ell$ is an l-adic sheaf which is lisse on a dense open subset of Y. To be concrete, the stalk of $\mathcal{G}_\ell$ at a point L of Y is dual to the direct sum $\oplus_i T_\ell A_i$, where the A_i are the four abelian varieties parametrized by $L \cap S^{\bar\rho}$.

By our choices of L_u, the field F is totally real. Similarly, our choices of L_v and L_w guarantee that A_i and F satisfy the local conditions at 3 and 5 stated in the theorem. The latter fact follows from a theorem of Kisin [11, Thm. 5.1] on ℓ-adic local constancy of Galois representations, applied to the sheaf $\mathcal{G}_\ell$. For instance, when $w|3$ and $\ell \neq 3$, the fact that L is very w-adically close to L_w implies that the stalk of $\mathcal{G}_\ell$ at L is isomorphic, as representation of D_w, to the stalk of $\mathcal{G}_\ell$ at L_w. Since the reduction type of an abelian variety at w is determined by its ℓ-adic Galois representation, the abelian varieties parametrized by $L \cap S^{\bar\rho}$, like those parametrized by $L_w \cap S^{\bar\rho}$, have multiplicative reduction at w. The local conditions at 5 are established similarly.

It remains only to check that L can be chosen so that $A_i[\sqrt{5}]$ is an absolutely irreducible $\mathrm{Gal}(\bar F/F)$-module, for some i. Let N be an integer whose divisors include 15 and all ramification primes of $\bar\rho$. By the arguments of section 2.1, the functor parametrizing HBAV's over $\mathbb{Z}[1/N]$ together with determinant-1 isomorphisms $A[3] \cong \bar\rho$ and $A[\sqrt{5}] \cong \mu_5 \oplus (\mathbb{Z}/5\mathbb{Z})$ is representable by an irreducible scheme $(X^{\bar\rho})'/\operatorname{Spec}\mathbb{Z}[1/N]$, which is an étale Galois cover of $X^{\bar\rho}$ with Galois group $\mathrm{SL}_2(\mathbb{F}_5)$. A dense open U of $X^{\bar\rho}$ is isomorphic to a dense open of $S^{\bar\rho}$; we now define U' to be $U \times_{X^{\bar\rho}} (X^{\bar\rho})'$.

Finally, define T' to be the pullback

$$\begin{array}{ccc} T' & \longrightarrow & T \times_{S^{\bar\rho}} U \\ \big\downarrow & & \big\downarrow \\ U' & \longrightarrow & U. \end{array}$$

We may think of T' as the variety parametrizing HBAV's A endowed with level 3 structure, level $\sqrt{5}$-structure, and a choice of a line in Y passing through the point of $S^{\bar\rho}$ parametrizing A.

We claim that $T'_{\bar{\mathbb{Q}}}$ is irreducible. The map $T'_{\bar{\mathbb{Q}}} \to U'_{\bar{\mathbb{Q}}}$ is proper and has irreducible closed fibers, since it is a base change from the map $T_{\bar{\mathbb{Q}}} \to S^{\bar\rho}_{\bar{\mathbb{Q}}}$, which has the same properties. (A fiber of $T_{\bar{\mathbb{Q}}} \to S^{\bar{\rho}}_{\bar{\mathbb{Q}}}$ is just the family of lines contained in a smooth quadric 3-fold and passing through a fixed point–in fact, such a family is isomorphic to $\mathbb{P}^1$ [8, 22.5].) Now suppose $T'_{\bar{\mathbb{Q}}}$ were the union of two closed subvarieties T'_1 and T'_2; then by properness the images of T'_1 and T'_2 in $U'_{\bar{\mathbb{Q}}}$ are closed, and by irreducibility of $U'_{\bar{\mathbb{Q}}}$, this means that T'_1 and T'_2 both map surjectively onto $U'_{\bar{\mathbb{Q}}}$. But this contradicts the irreducibility of the closed fibers in the map $T'_{\bar{\mathbb{Q}}} \to U'_{\bar{\mathbb{Q}}}$.

We now apply Ekedahl's version of the Hilbert Irreducibility Theorem [6, Th. 1.3] to the composition

$$\pi : T' \to T \times_{S^{\bar{\rho}}} U \to Y$$

replacing Y with a dense open, if necessary, to make sure π is an étale cover. Y, being a rational variety, has weak approximation over K'. It follows that we can choose an L in $Y(K')$ which is adelically close to $(L_{v_1}, \dots, L_{w_1}, \dots, L_{u_1}, \dots)$, and such that the fiber of T' over L is connected. This implies that the A_i/F are all Galois-conjugate over K', and that for each i the image of the representation $\mathrm{Gal}(\bar{\mathbb{Q}}/F_i) \to \mathrm{Aut}_{\mathcal{O}}(A_i[\sqrt{5}])$ surjects onto the determinant-1 subgroup, where $F_i \supset F$ is the (non-Galois) field over which A_i is defined. It follows that $\mathrm{Gal}(\bar{\mathbb{Q}}/F)$ also surjects onto the determinant-1 subgroup of $\mathrm{Aut}_{\mathcal{O}}(A_i[\sqrt{5}])$, since $\mathrm{SL}_2(\mathbb{F}_5)$ has no proper normal subgroup with solvable quotient. The fact that F has odd absolute ramification degree over 5 now implies that $\mathrm{Gal}(\bar{\mathbb{Q}}/F)$ surjects onto the whole of $\mathrm{Aut}_{\mathcal{O}}(A[\sqrt{5}])$. This completes the proof of Proposition 1.3. Moreover, if K'' is a finite extension of K', we can choose L in such a way that F/K' and K''/K' are linearly disjoint; this follows easily from the argument of [6], applied to the cover $T' \times_{K'} K'' \to Y$.

3. Modularity

Now that we have exhibited $\bar{\rho}$ as a representation appearing on the torsion points of an abelian variety, we can prove that $\bar{\rho}$ is modular. Our argument proceeds along the lines of [10] and [21], utilizing several different Galois representations; the reader may find it helpful to refer to the "chutes and ladders" diagrams at the end of this section, which give a schematic picture of the proof.

We begin by recording a theorem of Skinner and Wiles.

THEOREM 3.1. *Let K be a totally real number field, let $p > 2$ be a rational prime, let L be a finite extension of $\mathbb{Q}_p$, and let*

$$\rho : \mathrm{Gal}(\bar{K}/K) \to \mathrm{GL}_2(L)$$

be a continuous odd absolutely irreducible representation ramified at only finitely many primes. Suppose

- $\det \rho = \psi\chi_p^{k-1}$ *for some finite-order character* ψ *and some integer* $k > 1$, *called the* weight *of* ρ;

- *For each prime* v *of* K *dividing* p,

$$\rho|I_v \cong \begin{bmatrix} \psi\chi_p^{k-1} & * \\ 0 & 1 \end{bmatrix};$$

  (*A* p*-adic representation satisfying the first two conditions will be called* ordinary.)

- *The semisimplification of* $\bar{\rho}$ *is absolutely irreducible and* D_v*-distinguished for all primes* v *of* K *dividing* p;

- *There exist an ordinary modular Galois representation* ρ' *and an isomorphism between the* mod p *representations* $\bar{\rho}$ *and* $\bar{\rho}'$.

Then ρ *is modular.*

Proof. This is a special case of [19, Th. 5.1]. Note that the ordinariness of ρ' implies that ρ' is a χ_2-good lift of $\bar{\rho}$ in the sense of [19, §5]. □

We are now ready to prove the first part of our main result.

THEOREM 3.2. *Let* K *be a totally real number field whose absolute ramification indices over* 3 *and* 5 *are both odd. Let*

$$\bar{\rho} : G_K \to \mathrm{GL}_2(\mathbb{F}_9)$$

be an odd, absolutely irreducible Galois representation such that

- *For each prime* w *of* K *dividing* 3, *the restriction of* $\bar{\rho}$ *to* D_w *is*

$$\bar{\rho}|D_w \cong \begin{bmatrix} \psi_1 & * \\ 0 & \psi_2 \end{bmatrix},$$

  where $\psi_1|I_w = \bar{\chi}_3$ *and* ψ_2 *is unramified*;

- *For each prime* v *of* K *dividing* 5, *the image of* I_v *under* $\bar{\rho}$ *lies in* $\mathrm{SL}_2(\mathbb{F}_9)$, *and has odd order.*

Then there exist

- *a totally real, solvable Galois extension* F/K *such that* $\bar{\rho}|G_F$ *is* D_w*-distinguished for all primes* w *of* F *dividing* 3; *and*

- *an ordinary modular representation* $\rho : G_F \to \mathrm{GL}_2(\bar{\mathbb{Q}}_3)$ *reducing to* $\bar{\rho}|G_F$.

Proof. For each prime w of K dividing 3, let

$$\psi_1^w, \psi_2^w : D_w \to \mathbb{F}_9^*$$

be the characters given by the hypotheses of the theorem. Let F_0 be a totally real abelian extension of K, unramified at 3 and 5, such that $\psi_1^w \bar{\chi}_3^{-1}$ and ψ_2^w vanish when restricted to any decomposition group of F_0 over w. (Such an extension exists by class field theory, as in [21, Lemma 2.2].)

Now $\theta = (\det \bar{\rho})^{-1} \bar{\chi}_3$ is a character of $\mathrm{Gal}(\bar{\mathbb{Q}}/F_0)$ which annihilates all complex conjugations, since $\bar{\rho}$ is odd. We thus have a totally real abelian extension F_1/F_0 defined by $\mathrm{Gal}(\bar{\mathbb{Q}}/F_1) = \ker \theta$. Since $\det(\bar{\rho})(I_5)$ is trivial, I_5 lies in the kernel of θ, and F_1/F_0 is unramified at 5. Likewise, $\theta(I_3)$ is trivial, so F_1/F_0 is unramified at 3. Now the local conditions on $\bar{\rho}$ at primes dividing 3 and 5 imply the corresponding local conditions in Proposition 1.3, and the determinant of ρ is cyclotomic when restricted to F_1. We may now choose an extension F_2/K and an abelian variety A/F_2 satisfying the four hypotheses given in Proposition 1.3.

From here, we proceed along the lines of [21]. We will first prove that the irreducible representation

$$\bar{\rho}_{A,\sqrt{5}} : \mathrm{Gal}(\bar{\mathbb{Q}}/F_2) \to \mathrm{GL}_2(\mathbb{F}_5)$$

induced by the torsion subscheme $A[\sqrt{5}]$ is modular. The main tool is the following lemma.

Lemma 3.3. *There exists a solvable totally real extension F/F_2, unramified at 3 and 5, and an elliptic curve E/F, such that*

- *E has multiplicative reduction at all primes over 3 and 5;*
- *$E[5] \cong A[\sqrt{5}]$ as Galois modules;*
- *For each prime $w|3$ of F, we have $E[3] \cong \mu_3 \oplus (\mathbb{Z}/3\mathbb{Z})$ as I_w-modules.*
- *$\bar{\rho}_{E,3}$ is absolutely irreducible.*

Proof. For every characteristic 0 field L and every $\alpha \in L^*$, let $\bar{\rho}_\alpha$ be the extension of $\mathbb{Z}/5\mathbb{Z}$ by μ_5 in Kummer correspondence with the class of α in $L^*/(L^*)^5$.

We have arranged for A to have multiplicative reduction at all primes of F_2 over 3. So the subgroup

$$\bar{\rho}_{A,\sqrt{5}}(I_w) \subset \mathrm{SL}_2(\mathbb{F}_5)$$

is unipotent, for every prime $w|3$. In particular, after passing to an unramified extension F_w of $F_{2;w}$, we have $\bar{\rho}_{A,\sqrt{5}} \cong \bar{\rho}_\alpha$ for some α in F_w^*. We can choose α to be contained in the maximal ideal of the ring of integers of F_w. Then we set E_w/F_w to be the elliptic curve $\mathbb{G}_m/\alpha^{\mathbb{Z}}$. So $E_w[5] \cong A[\sqrt{5}]$ as Galois modules. Since α is defined only up to 5^{th} powers, we may assume further that $\alpha \in (F_w^*)^3$. This implies that $E_w[3] \cong \mu_3 \oplus (\mathbb{Z}/3\mathbb{Z})$ as I_w-modules.

Now suppose v is a prime of F_2 dividing 5. Since A is good ordinary or multiplicative at v, we have

$$\bar{\rho}_{A,\sqrt{5}}|I_v \cong \begin{bmatrix} \bar{\chi}_5 & * \\ 0 & 1 \end{bmatrix}.$$

Once again, over some unramified extension F_v of $F_{2;v}$, we have $\bar{\rho}_{A,\sqrt{5}} \cong \bar{\rho}_\beta$ for some β in the maximal ideal of the ring of integers of F_v^*. Let E_v/F_v be the elliptic curve $\mathbb{G}_m/\beta^{\mathbb{Z}}$.

By [21, Lemma 2.2], there is a solvable totally real extension F/F_2 (necessarily unramified at 3 and 5) such that all completions of F over 3 and 5 are isomorphic to some $F_{2;w}$ or $F_{2;v}$. Let C/F be the modular curve parametrizing elliptic curves E over F with $E[5] \cong A[\sqrt{5}]$. Then C is a rational curve over F [17, Lemma 1.1], and in particular C has weak approximation. So there is a point P in $C(F)$ which is arbitrarily close to the points parametrizing E_v and E_w for all $v|5$ and all $w|3$. Let E/F be the elliptic curve parametrized by P. Using the result of Kisin [11] as in Section 2.5, we conclude that E has multiplicative reduction at all v, w. Ekedahl's version of Hilbert irreducibility [6] guarantees that E can be chosen such that $\bar{\rho}_{E,3}$ is surjective, just as in the proof of [21, Lemma 2.3]. □

Let ρ_0 be the composition of the mod 3 representation $G_F \to \mathrm{End}(E[3])$ with an injection $\mathrm{GL}_2(\mathbb{F}_3) \hookrightarrow \mathrm{GL}_2(\mathbb{C})$.

The Langlands-Tunnell theorem implies that there exists an automorphic form π_1 of weight 1 on $\mathrm{GL}_2(F)$ such that $L(\pi_1, s) = L(\rho_0, s)$. In order to use the Skinner-Wiles theorem, we need to lift ρ_0 to an ordinary automorphic representation of weight at least 2. For this, we use Wiles's theorem on Hida families of ordinary Hilbert modular forms. Let w be a prime of F dividing 3. The hypothesis that $E[3]$ is semisimple as I_w-module implies that the local factor of $L(\rho_0, s)$ at w is

$$L_w(\rho_0, s) = (1 - a_w(\rho_0)(\mathbf{N}w)^{-s})^{-1}$$

where $a_w(\rho_0) = \pm 1$.

Let f_1 be a Hilbert modular newform of weight 1 associated to π_1. Let $c(w, f_1)$ be the eigenvalue of the Hecke operator $T(w)$ acting on f_1, as in [23]. Then

$$L_w(f_1, s) = (1 - c(w, f_1)(\mathbf{N}w)^{-s})^{-1}$$

and it follows that $c(w, f_1) = \pm 1$. In particular, $c(w, f_1)$ is a unit mod 3, and so Theorem 1.4.1 of [23] shows the existence of a Λ-adic modular form (i.e. a Hida family) $\mathcal{F}$ which specializes to f_1 in weight 1. Let f be the specialization of $\mathcal{F}$ to weight 2. Then Theorem 2.1.4 of [23] associates to f an ordinary Galois representation

$$\rho' : G_F \to \mathrm{GL}_2(\bar{\mathbb{Q}}_3)$$

of weight 2, which reduces to $\bar\rho_{E,3}$.

Evidently, $\bar\rho_{E,3}$ is distinguished at all $w|3$. We now know, by Theorem 3.1, that T_3E is modular. It follows that T_5E is modular, so $E[5]$, whence also $A[\sqrt{5}]/F$, is modular.

By hypothesis, A has good ordinary or multiplicative reduction at 5, so that $T_{\sqrt{5}}A$ is an ordinary representation. Because $F/\mathbb{Q}$ has odd ramification degree over 5, $A[\sqrt{5}]$ is D_v-distinguished for all primes v dividing 5. Now $T_{\sqrt{5}}A/F$ is modular by another application of Theorem 3.1, with $\rho' = T_5E$. This implies that T_3A/F is also modular.

Theorem 3.2 is now proved, with T_3A as ρ. We note that F has odd absolute ramification degree at every prime over 3, which guarantees that $\bar\rho|G_F$ is D_w-distinguished for all $w|3$. Note also that this is the point where we use Lemma 2.6; without that fact, we would not necessarily be able to find an F with odd ramification degree over 3 such that $\bar\rho|G_F$ admits an *ordinary* modular lift. □

The following two propositions, essentially due to Khare, Ramakrishna, and Taylor, allow us to use Theorem 3.2 to prove Serre's conjecture over $\mathbb{F}_9$ under some local hypotheses.

PROPOSITION 3.4. *Let K be a totally real number field whose absolute ramification indices over* 3 *and* 5 *are both odd. Let*

$$\bar\rho : G_K \to \mathrm{GL}_2(\mathbb{F}_9)$$

be an odd, absolutely irreducible Galois representation, and let F be a totally real solvable Galois extension of K. Suppose that

- *$\bar\rho|G_F$ is absolutely irreducible and D_w-distinguished for all primes w of F dividing* 3;
- *$\bar\rho|G_F$ is the reduction of an ordinary modular representation $\rho' : G_F \to \mathrm{GL}(L')$, for some finite extension $L'/\mathbb{Q}_3$.*
- *$\bar\rho$ is the reduction of an ordinary representation $\rho : G_K \to \mathrm{GL}_2(L)$, for some finite extension $L/\mathbb{Q}_3$.*

Then $\bar\rho$ is modular.

Proof. We apply Khare's idea of using cyclic descent ([10], [21]). Let F^1 be a subfield of F such that F/F^1 is a cyclic Galois extension. It follows from Theorem 3.1 that $\rho|G_F$ is modular. The automorphic form π on $\mathrm{GL}_2(F)$

corresponding to ρ is preserved by $\mathrm{Gal}(F/F^1)$. Therefore, π descends to an automorphic form π^1 on $\mathrm{GL}_2(F^1)$. The Galois representation ρ^1 of G_{F^1} associated to π^1 restricts to $\rho|G_F$; thus after applying a twist we have $\rho^1 = \rho|G_{F^1}$. Continuing inductively, one finds that ρ itself is associated to a modular form on $\mathrm{GL}_2(K)$; therefore, its mod 3 reduction $\bar{\rho}$ is modular. □

The following proposition tells us that the conditions of Proposition 3.4 hold in many cases of interest when $K = \mathbb{Q}$.

PROPOSITION 3.5. *Let*

$$\bar{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_9)$$

be a representation with nonsolvable image, such that

$$\bar{\rho}|I_3 \cong \begin{bmatrix} \bar{\chi}_3 & * \\ 0 & 1 \end{bmatrix}.$$

Then there exists an ordinary representation

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(W(\mathbb{F}_9))$$

reducing to $\bar{\rho}$.

Proof. This is a special case of a theorem of Taylor [21, Thm. 1.3], which refines results of Ramakrishna. For completeness's sake, one might add to the proof of [21, Thm. 1.3] the fact, checkable by direct computation, that $H^1(A_5, ad^0\tau) = 0$, where τ is the embedding of A_5 into $\mathrm{PGL}_2(\mathbb{F}_9)$. This is needed to ensure that the statement of the theorem is correct in case the projective image of $\bar{\rho}$ is A_5 and the characteristic of the coefficient field is 3. □

The combination of Theorem 3.2 and Propositions 3.4 and 3.5 gives:

THEOREM 3.6. *Let*

$$\bar{\rho} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_9)$$

be an odd Galois representation such that

- *$\bar{\rho}$ has nonsolvable image;*
- *The restriction of $\bar{\rho}$ to D_3 can be written as*

  $$\bar{\rho}|D_3 \cong \begin{bmatrix} \psi_1 & * \\ 0 & \psi_2 \end{bmatrix},$$

  where $\psi_1|I_3$ is the mod 3 *cyclotomic character, and ψ_2 is unramified;*
- *The image of the inertia group I_5 lies in $\mathrm{SL}_2(\mathbb{F}_9)$, and has odd order.*

Then $\bar{\rho}$ is modular.

Moreover, we can prove the modularity of Hilbert-Blumenthal abelian surfaces over $\mathbb{Q}$, under some conditions on reduction at 3 and 5.

Corollary 3.7. *Let $A/\mathbb{Q}$ be a Hilbert-Blumenthal abelian surface which has good ordinary or multiplicative reduction at* 3 *and* 5*. Then A is a quotient of $J_0(N)$ for some integer N.*

Proof. Let λ be a prime of the field of real multiplication dividing 3. If $A[\lambda]$ is absolutely reducible, then the corollary follows from Theorem A of [18]. So we may assume that $A[\lambda]$ is absolutely irreducible.

Suppose λ has norm 3. Then by the ordinariness of A, there exists an elliptic curve $E_3/\mathbb{Q}_3$ with multiplicative reduction such that $E_3[3] \cong A[\lambda]$ as D_3-modules. The variety parametrizing elliptic curves $E/\mathbb{Q}$ such that $E[3] \cong A[\lambda]$ as $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$-modules is isomorphic to $\mathbb{P}^1$; it follows from weak approximation on $\mathbb{P}^1(\mathbb{Q})$ that there is an elliptic curve $E/\mathbb{Q}$ with multiplicative reduction at 3 and $E[3] \cong A[\lambda]$. Now E is modular by [2] (or, if we choose E semistable at 5, by the earlier result of [5];) so T_3E is an ordinary modular lifting of $A[\lambda]$, and the modularity of $T_\lambda A$ follows from Theorem 3.1.

Suppose on the other hand that λ has norm 9. Then Theorem 3.2 produces a totally real solvable extension $F/\mathbb{Q}$ and an ordinary modular representation ρ of G_F reducing to $A[\lambda]|G_F$. Then modularity of $T_\lambda A$ follows from Proposition 3.4 as soon as we verify that $A[\lambda]|G_F$ is absolutely irreducible. This is immediate from the solvability of $F/\mathbb{Q}$ unless the image of Galois on $A[\lambda]$ is solvable. In that case, we must be a bit more careful. It suffices to show that the field F constructed by Theorem 3.2 can be chosen to be linearly disjoint from the finite extension $L/\mathbb{Q}$ whose Galois group is $\ker\bar{\rho}$, so that $\bar{\rho}(G_F) = \bar{\rho}(G_{\mathbb{Q}})$.

It is easy to see that the solvable extensions satisfying specified local conditions, as constructed by [21, Lemma 2.2], can be chosen to be linearly disjoint from L. Thus the extension F_1/K in the first paragraph of the proof of Theorem 3.2 can be chosen to be linearly disjoint from L, since the character θ is trivial and the other conditions on F_1 are local. Now the extension F_2/F_1 arises from the global construction in section 2.5. In that construction, the conditions on K' are local, so K' can be chosen linearly disjoint from L; the remark at the end of section 2.5 explains why the degree-4 extension F_2/K' can be chosen to be linearly disjoint from L. Finally, the solvable extension F/F_2 is required only to satisfy local conditions, so F can be chosen linearly disjoint from L as claimed. □

Chutes and ladders

Arguments along the lines of the present paper, [12], and [21] can be diagrammed in a way that may clarify the relationships between the various Galois representations involved.

In each of the diagrams below, the Galois representations in the top row are ℓ-adic, and those in the bottom row are mod ℓ. The theorem of Langlands and Tunnell introduces modularity into the picture at a $\mathrm{GL}_2(\mathbb{F}_3)$ representation. Modularity then proceeds up ladders by means of the Skinner-Wiles theorem, across horizontal segments by compatibility of ℓ-adic Galois representations, and down chutes by definition. In each case, F is a totally real number field solvable over $\mathbb{Q}$ and ρ is a representation produced by a lifting theorem à la Ramakrishna; the modularity of $\rho|G_F$ implies that of ρ by base change as above.

As the terminology suggests, sliding down is easier than climbing up.

Icosahedral representations [21]:

$$\begin{tikzcd}
& T_3E \arrow[r] & T_5E \arrow[dr] & & \rho|G_F \\
E[3] \arrow[ur] & & & E[5] = \bar{\rho}|G_F \arrow[ur] &
\end{tikzcd}$$

$\mathrm{GL}_2(\mathbb{F}_7)$ *representations* [12]:

$$\begin{tikzcd}
& T_3E \arrow[r] & T_7E \arrow[dr] & & \rho|G_F \\
E[3] \arrow[ur] & & & E[7] = \bar{\rho}|G_F \arrow[ur] &
\end{tikzcd}$$

$\mathrm{GL}_2(\mathbb{F}_9)$ *representations*:

$$\begin{tikzcd}
& T_3E \arrow[r] & T_5E \arrow[dr] & & T_{\sqrt{5}}A \arrow[r] & T_3A \arrow[dr] & & \rho|G_F \\
E[3] \arrow[ur] & & & E[5] = A[\sqrt{5}] \arrow[ur] & & & A[3] = \bar{\rho}|G_F \arrow[ur] &
\end{tikzcd}$$

4. Appendix: computation of Fourier coefficents of Hilbert modular forms

In this appendix we compute the first few terms of the q-expansions of the Hilbert modular forms $s_0, \ldots, s_6$ defined in section 2.1, and as a result derive the expressions given there for the symmetric functions $\sigma_k(s_0, \ldots, s_6)$ in terms of level 1 modular forms. We include these computations for two reasons: because they are necessary to the proof of the present theorem, and because the ideas apply in general to the computation of q-expansions of Hilbert modular forms of small level, given the q-expansions of the forms in level 1.

We recall that the q-expansion of a cusp form f of level n for $\mathrm{GL}_2(\mathcal{O})$ is a power series

$$\sum_{\alpha \in (1/n)\mathfrak{d}^{-1}} a_\alpha(f) q^\alpha$$

where $\mathfrak{d}^{-1}$ is the inverse different of $\mathcal{O}$, and $a_\alpha = 0$ unless α is totally positive. We write $\mathbb{Z}[[q^{(1/n)\mathfrak{d}^{-1}}]]$ for the ring of such power series.

In the present case, we have $n = 3$ and $\mathcal{O} = \mathbb{Z}[\frac{1+\sqrt{5}}{2}]$. We can thus write the q-expansion of a modular form f as $\sum a_{\alpha/3}(f) q^{\alpha/3}$, where α ranges over

$$\frac{1}{2} - \frac{\sqrt{5}}{10},\ \frac{1}{2} + \frac{\sqrt{5}}{10},\ 1 - \frac{2\sqrt{5}}{5},\ 1 - \frac{\sqrt{5}}{5},\ 1,\ 1 + \frac{\sqrt{5}}{5},\ 1 + \frac{2\sqrt{5}}{5}, \dots .$$

The values of α not listed above are precisely those with trace at least 3.

Since the coefficients of these modular forms vary in two directions, it is convenient to write the q-expansion in a grid: for instance, a q-expansion beginning

$$q^{(1/3)(1/2-\sqrt{5}/10)} - q^{(1/3)(1-\sqrt{5}/5)} + q^{(1/3)(1+2\sqrt{5}/5)}$$

can be written

$$\begin{array}{ccccccc} 0 & -1 & & 0 & & 0 & 1 \\ & & 1 & & 0 & & \end{array}$$

Here, a row in the grid contains the coefficients of all exponents α with a specified $\mathrm{Tr}(\alpha)$, and a column contains the coefficients of all exponents with a specified $\mathrm{Tr}(\sqrt{5}\alpha)$.

By the "trace n term" of a modular form we mean the sum of all terms $a_\alpha q^\alpha$ of the Fourier expansion with $\mathrm{Tr}\alpha = n$.

We now compute the q-expansions of $s_0, \dots, s_5$. Write ω for a cube root of unity, to be fixed through the whole discussion. The symbol Tr always denotes trace from $\mathfrak{d}^{-1}$ to $\mathbb{Z}$.

It follows from [22, VIII.2.5] that an even-weight nonsymmetric form of level 1 must have weight at least 20. The even-weight level 1 forms σ_3, σ_5, and σ_6 are thus automatically symmetric. Therefore, they lie in the ring generated by ϕ_2, χ_6, and χ_{10}. The q-expansions of ϕ_2, χ_6, and χ_{10} can be found in [13, (5.12)].

The following useful lemma forces many Fourier coefficients to be 0.

LEMMA 4.1. *Choose i in $\{0, \dots, 5\}$, and elements $\alpha, \beta \in \mathfrak{d}^{-1}$ which generate the abelian group $\mathfrak{d}^{-1}/3\mathfrak{d}^{-1}$. Then either $a_{\alpha/3}(s_i) = 0$ or $a_{\beta/3}(s_i) = 0$.*

Proof. For each $\gamma \in \mathcal{O}$, let t_γ be the automorphism of $\mathbb{Z}[[q^{(1/3)\mathfrak{d}^{-1}}]]$ obtained by replacing $q^{\alpha/3}$ by $q^{\alpha/3}\omega^{\mathrm{Tr}(\alpha\gamma)}$ for all α in $\mathfrak{d}^{-1}$. Note that t_γ fixes the subring of $\mathbb{Z}[[q^{(1/3)\mathfrak{d}^{-1}}]]$ with exponents in $\mathfrak{d}^{-1}$; this means that t_γ fixes all level

1 modular forms. In particular t_γ fixes all symmetric functions $\sigma_k(s_i)$, which means that t_γ permutes the s_i. Let $t_{\mathcal{O}/3\mathcal{O}}$ be the group generated by the t_γ. Then $t_{\mathcal{O}/3\mathcal{O}}$ is isomorphic to $(\mathbb{Z}/3\mathbb{Z})^2$ and embeds in S_6 via its action on the s_i.

Suppose $a_{\alpha/3}(s_i)$ and $a_{\beta/3}(s_i)$ are both nonzero. Then, by the hypothesis on α and β, the automorphism t_γ fixes s_i only if $\gamma \in 3\mathcal{O}$. But the group generated by commuting 3-cycles in S_6 must contain a nontrivial permutation fixing s_i, which is a contradiction. $\square$

We now proceed with the computation of Fourier expansions of $s_0, \dots, s_5$. The form $\sigma_3(s_0, \dots, s_5)$ is a weight 6 cusp form and is thus a multiple of χ_6, whose Fourier expansion begins

$$q^{(1/2-\sqrt{5}/10)} + q^{(1/2+\sqrt{5}/10)} + \dots .$$

This is enough information to compute the Fourier coefficients of the s_i with trace 1/3. By the lemma above, each s_i has either $a_{(1/3)(1/2-\sqrt{5}/10)} = 0$ or $a_{(1/3)(1/2+\sqrt{5}/10)} = 0$. It is easy to see that, for σ_3 to have the right trace-1 term, the trace-1/3 terms of $s_0, \dots, s_5$ must be

$$\begin{aligned}
s_4 &= q^{(1/3)(1/2-\sqrt{5}/10)} + \dots \\
s_0 &= \omega q^{(1/3)(1/2-\sqrt{5}/10)} + \dots \\
s_1 &= \omega^2 q^{(1/3)(1/2-\sqrt{5}/10)} + \dots \\
s_5 &= q^{(1/3)(1/2+\sqrt{5}/10)} + \dots \\
s_2 &= \omega q^{(1/3)(1/2+\sqrt{5}/10)} + \dots \\
s_3 &= \omega^2 q^{(1/3)(1/2+\sqrt{5}/10)} + \dots .
\end{aligned}$$

Note that these expansions are determined only up to renumbering of the s_i and multiplication by a common constant. In fact, changing the numbering of the forms by an even permutation amounts to applying an automorphism in $\mathrm{PSL}_2(\mathbb{F}_9)$ to the level structure, which in turn amounts to computing the q-expansions at a different cusp of X. We will speak at the end of the appendix about the description of the particular cusp at which the q-expansions here are taken.

Using the lemma again, the q-expansion of s_4 up to the trace-2/3 part is

$$\begin{array}{ccccccccc}
0 & & a & & 0 & & 0 & & b \\
& & & 1 & & 0 & & &
\end{array}$$

and that of s_5 is

$$\begin{array}{ccccccccc}
c & & 0 & & 0 & & d & & 0 \\
& & & 0 & & 1. & & &
\end{array}$$

Note that $s_0 = t_1 s_4, s_1 = t_1^2 s_4, s_2 = t_1 s_5, s_3 = t_1^2 s_5$, so it suffices for our purposes to compute the q-expansions of s_4 and s_5.

We now compute that $\sigma_2(s_0, \ldots, s_5)$ has trace 1 term

$$-3(a+c)q^{1/2-\sqrt{5}/10} - 3(b+d)q^{1/2+\sqrt{5}/10}$$

which implies that $c = -a$ and $d = -b$. We make this substitution, and compute that $\sigma_5(s_0, \ldots, s_6)$ has trace 2 term

$$3aq^{1-\sqrt{5}/5} + (3b-3a)q - 3bq^{1-\sqrt{5}/5}$$

and zero trace 1 term. A weight 10 level 1 cusp form with zero trace 1 term must be a multiple of χ_{10}, whose trace 2 term is

$$q^{1-\sqrt{5}/5} - 2q + q^{1-\sqrt{5}/5}$$

from which we conclude that $b = -a$ and $\sigma_5 = 3a\chi_{10}$.

It remains to compute a. The coefficient of $q^{1-2\sqrt{5}/5}$ in σ_3 is $c^3 = -a^3$, because the only triple of totally positive elements of $(1/3)\mathfrak{d}^{-1}$ summing to $1-2\sqrt{5}/5$ is three copies of $(1/3)(1-2\sqrt{5}/5)$. Since the coefficient of $q^{1-2\sqrt{5}/5}$ in χ_6 is 1, we have $a^3 = -1$. Making different choices of cube root amounts to multiplying all s_i by a cube root of unity and permuting; since we are only working up to constants, we may take $a = -1$.

We conclude that, up to terms of trace higher than $2/3$, the q-expansions of s_4 and s_5 are

$$s_4 = \begin{array}{ccccccc} 0 & -1 & & 0 & & 0 & 1 \\ & & 1 & & 0 & & \end{array}$$

$$s_5 = \begin{array}{ccccccc} 1 & 0 & & 0 & & -1 & 0 \\ & & 0 & & 1 & & \end{array}$$

It follows, as computed above, that $\sigma_3(s_0, \ldots, s_5) = \chi_6$ and $\sigma_5(s_0, \ldots, s_5) = -3\chi_{10}$. The trace 2 term of σ_6 consists of the single monomial q. Now σ_6 is a weight 12 cuspform, and is thus a linear combination of $\phi_2\chi_{10}$ and χ_6^2. The trace 2 term of $\phi_2\chi_{10}$ is

$$q^{1-\sqrt{5}/5} - 2q + q^{1-\sqrt{5}/5}$$

and the trace 2 term of χ_6^2 is

$$q^{1-\sqrt{5}/5} + 2q + q^{1-\sqrt{5}/5}.$$

It follows that $\sigma_6(s_0, \ldots, s_5) = (1/4)(\chi_6^2 - \phi_2\chi_{10})$. We now have the desired relations

$$\begin{aligned} \phi_2 &= -3\sigma_5^{-1}(\sigma_3^2 - 4\sigma_6) \\ \chi_6 &= \sigma_3 \\ \chi_{10} &= (-1/3)\sigma_5. \end{aligned}$$

We now return to the question of describing the cusp at which the q-expansions given here are computed. We need to recall some basic facts about

Tate HBAV's ([3],[7]). Let S be a set of d linearly independent elements of $\mathcal{O}$, and say an element α of $\mathfrak{d}^{-1}$ is *S-semipositive* if $\mathrm{Tr}(x\alpha) \geq 0$ for all $x \in S$. Let $\mathbb{Z}[[\mathfrak{d}^{-1}, S]]$ be the ring of power series of the form

$$\sum_{\alpha \in (1/n)\mathfrak{d}^{-1}} a_\alpha(f) q^\alpha$$

where $a_\alpha = 0$ unless α is S-semipositive. Then Mumford's construction yields a semi-HBAV G over $\mathrm{Spec}\,\mathbb{Z}[[\mathfrak{d}^{-1}, S]]$ which can be thought of as the "quotient" $(\mathbb{G}_m \otimes_{\mathbb{Z}} \mathfrak{d}^{-1})/q^{\mathcal{O}}$. The 3-torsion subscheme $G[3]$ fits into an exact sequence

$$0 \to \mu_3 \otimes_{\mathbb{Z}} \mathfrak{d}^{-1} \xrightarrow{\iota} G[3] \to \mathcal{O}/3\mathcal{O} \to 0.$$

Let $\phi : G[3]/T \cong (\mathcal{O}/3\mathcal{O})^2/T$ be a 3-level structure for G over some $\mathbb{Z}[[\mathfrak{d}^{-1}, S]]$-scheme T. We say ϕ is *canonical* if the image of $\phi \circ \iota$ is the first factor of $(\mathcal{O}/3\mathcal{O})^2$. We also refer to the image of ι as the canonical subgroup of $G[3]$.

Let η_{can} be the canonical generator for $\mathrm{Lie}(G)$. Then we may define the q-expansion of a form f at a cusp ϕ over $T = \mathrm{Spec}\,R$ to be $f(G, \phi, \eta_{\mathrm{can}}) \in R$.

The six q-expansions computed above were apparently numbered arbitrarily, which is to say that we have not specified the cusp at which the q-expansions are being computed. We now wish to argue that, with the conventions used here, the q-expansions above are actually being computed at a canonical cusp.

Denote by $t_{\mathcal{O}/3\mathcal{O}}$ the group of automorphisms of the ring of q-expansions described in the proof of Lemma 4.1. When ϕ is canonical, it is easy to check that the action of the subgroup

$$U = \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \subset \mathrm{PSL}_2(\mathbb{F}_9)$$

on the space of modular forms acts on q-expansions via $t_{\mathcal{O}/3\mathcal{O}}$. Recall that we have chosen our isomorphism between $\mathrm{PSL}_2(\mathbb{F}_9)$ and A_6 to send U to the group generated by the 3-cycles (014) and (235). It is then easy to see from the computations above that, indeed, U acts on q-expansions via $t_{\mathcal{O}/3\mathcal{O}}$.

Now suppose ϕ is not canonical; then $\phi = g\phi_{\mathrm{can}}$, where ϕ_{can} is canonical and g is an element of $\mathrm{PGL}_2(\mathbb{F}_9)$. In this case, it is the group $g^{-1}Ug$ which acts on the q-expansions of $s_0, \ldots, s_5$ via $t_{\mathcal{O}/3\mathcal{O}}$. In our case, the only elements of $\mathrm{PSL}_2(\mathbb{F}_9)$ whose action on modular forms is via $t_{\mathcal{O}/3\mathcal{O}}$ are the elements of U, which implies g is in the normalizer of U; but this means that g fixes the first factor of $(\mathcal{O}/3\mathcal{O})^2$, whence ϕ is in fact canonical.

Princeton University, Princeton, NJ
E-mail address: ellenber@math.princeton.edu

References

[1] F. Andreatta and E. Z. Goren, *Hilbert Modular Forms*: mod p *and* p*-adic Aspects*, *Mem. Amer. Math. Soc* **173** (2005), no. 819; Available from arXiV as `math.NT/0308040`.

[2] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc.* **14** (2001), 843–939.

[3] C.-L. Chai, Arithmetic minimal compactification of the Hilbert-Blumenthal moduli spaces, *Ann. of Math.* **131** (1990), 541–554.

[4] B. Conrad and S. Wong, Remarks on mod-ℓ^n representations, $\ell = 3, 5$, *J. Number Theory* **78** (1999), 253–270.

[5] F. Diamond, On deformation rings and Hecke rings, *Ann. of Math.* **144** (1996), 137–166.

[6] T. Ekedahl, An effective version of Hilbert's irreducibility theorem, in *Séminaire de Théorie des Nombres* (Paris, 1988–1989) *Progr. Math.* **91**, 241–249, Birkaüser Boston, Boston, MA, 1990.

[7] J. Ellenberg, Finite flatness of torsion subschemes of Hilbert-Blumenthal abelian varieties, *J. Reine Angew. Math.* **532** (2001), 1–32.

[8] J. Harris, *Algebraic Geometry*, Springer-Verlag, New York, 1992.

[9] F. Hirzebruch, The ring of Hilbert modular forms for real quadratic fields of small discriminant, in *Modular Functions of One Variable VI*, *Lecture Notes in Math.* **627**, 288–323, Springer-Verlag, New York, 1977.

[10] C. Khare, Mod p descent for Hilbert modular forms, *Math. Res. Lett.* **7** (2000), 455–462.

[11] M. Kisin, Local constancy in p-adic families of Galois representations, *Math. Z.* **230** (1999), 569–593.

[12] J. Manoharmayum, On the modularity of certain $\mathrm{GL}_2(\mathbb{F}_7)$ Galois representations, *Math. Res. Lett.* **8** (2001), 703–712.

[13] S. Nagaoka, On the ring of Hilbert modular forms over $\mathbf{Z}$, *J. Math. Soc. Japan* **35** (1983), 589–608.

[14] M. Rapoport, Compactifications de l'espace de modules de Hilbert-Blumenthal, *Compositio Math.* **37** (1978), 255–335.

[15] K. Ribet, Galois actions on division points of abelian varieties with real multiplications, *Amer. J. Math.* **98** (1976), 751–804.

[16] J-P. Serre, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, *Duke Math. J.* **54** (1987), 179–230.

[17] N. Shepherd-Barron and R. Taylor, Mod 2 and mod 5 icosahedral representations, *J. Amer. Math. Soc.* **10** (1997), 283–298.

[18] C. Skinner and A. Wiles, Residually reducible representations and modular forms, *Inst. Hautes Études Sci. Publ. Math.* **89** (1999), 5–126.

[19] C. Skinner and A. Wiles, Nearly ordinary deformations of irreducible residual representations, *Ann. Fac. Sci. Toulouse Math.* **10** (2001), 185–215.

[20] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math.* **141** (1995), 553–572.

[21] R. Taylor, On icosahedral Artin representations II, *Amer. J. Math.* **125** (2003), 549–566.

[22] G. van der Geer, *Hilbert Modular Surfaces*, Springer-Verlag, New York, 1988.

[23] A. Wiles, On ordinary λ-adic representations attached to modular forms, *Invent. Math.* **94** (1988), 529–573.

[24] ———, Modular elliptic curves and Fermat's last theorem, *Ann. of Math.* **141** (1995), 443–551.

(Received July 26, 2001)