JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 10, Number 2, April 1997, Pages 283–298
S 0894-0347(97)00226-9

MOD 2 AND MOD 5 ICOSAHEDRAL REPRESENTATIONS

N. I. SHEPHERD–BARRON AND R. TAYLOR

Introduction

We shall call a simple abelian variety $A/\mathbb{Q}$ modular if it is isogenous (over $\mathbb{Q}$) to a factor of the Jacobian of a modular curve. In this paper we shall call a representation $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\bar{\mathbb{F}}_l)$ modular if there is a modular abelian variety $A/\mathbb{Q}$, a number field $F/\mathbb{Q}$ of degree equal to $\dim A$, an embedding $\mathcal{O}_F \hookrightarrow \mathrm{End}(A/\mathbb{Q})$ and a homomorphism $\theta : \mathcal{O}_F \to \bar{\mathbb{F}}_l$ such that $\bar{\rho}$ is equivalent to the action of $G_{\mathbb{Q}}$ on the $\ker\theta$ torsion points of A. If $\bar{\rho}$ is modular in our sense, then there exists a Hecke eigenform f and a homomorphism from the ring generated by its Fourier coefficients to $\bar{\mathbb{F}}_l$ such that $\mathrm{Tr}\,\bar{\rho}(\mathrm{Frob}_p)$ is congruent to the eigenvalue of T_p on f modulo $\ker\theta$ for almost all primes p. (One may further assume either that f is of weight 2 and cuspidal, or that it has level coprime to l. This follows from the results of [AS].) If $\bar{\rho}$ is irreducible, then $\bar{\rho}$ being modular is equivalent to the existence of a Hecke eigenform f related to $\bar{\rho}$ in this way. However if $\bar{\rho}$ is reducible, then our definition appears to be stronger than requiring the existence of such an f, even if we insist that f be cuspidal of weight 2. (This is because the existence of such an f depends only on the semi-simplification of $\bar{\rho}$ and not on $\bar{\rho}$ itself.)

Serre [S2] has conjectured that any continuous irreducible homomorphism $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\bar{\mathbb{F}}_l)$ with odd determinant is modular. Very little is known about this conjecture. It has been proved for representations $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_2)$ and $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_3)$. In both cases this can be achieved by lifting $\bar{\rho}$ to a continuous odd irreducible representation $G_{\mathbb{Q}} \to GL_2(\mathcal{O})$, where $\mathcal{O}$ is the ring of integers of some number field, and then using the Langlands-Tunnell theorem which asserts that this latter representation is modular of weight 1 (see [L] and [T]). This makes essential use of the fact that $GL_2(\mathbb{F}_2)$ and $GL_2(\mathbb{F}_3)$ are soluble.

The purpose of this paper is to treat Serre's conjecture in two further cases where the image of $\bar{\rho}$ is no longer soluble. We show that if $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_5)$ is unramified at 3 and has determinant the cyclotomic character, then $\bar{\rho}$ is modular. We also show that if $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_4)$ is unramified at 3 and 5, then $\bar{\rho}$ is modular. We do this by proving first that $\bar{\rho}$ (up to twist) can be realized on respectively the 5-division points on an elliptic curve over $\mathbb{Q}$ and the 2-division points of an abelian surface with multiplication by $\mathcal{O}_{\mathbb{Q}(\sqrt{5})}$ over $\mathbb{Q}$. (In fact we prove this over any given field of characteristic zero. Since $SL_2(\mathbb{F}_4)$ is isomorphic to A_5, this can be stated as follows: given any quintic polynomial over a field K of characteristic zero, its splitting field can be obtained by adjoining first the square

Received by the editors June 10, 1996.
1991 *Mathematics Subject Classification.* Primary 11F41, 11G10, 14G05, 14G35.
The second author was partially supported by the EPSRC.

root of the discriminant d and then the 2-division points of an abelian surface with multiplication by $\mathcal{O}_{\mathbb{Q}(\sqrt{5})}$ defined over $K(\sqrt{d})$.) We then use a recent result of Wiles (see [W]), as completed in [TW] and extended in [D], to show that these abelian varieties are modular (using respectively their 3-adic and the $\sqrt{5}$-adic Tate modules). We remark that the ramification conditions arise for technical reasons and do not seem to be essential to the method. On the other hand the restriction on the determinant for representations to $GL_2(\mathbb{F}_5)$ does seem to be essential to the method.

Wiles has asked (private communication) whether in fact any continuous indecomposable homomorphism $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\bar{\mathbb{F}}_l)$ with odd determinant is modular. He has (unpublished) answered this question affirmatively for representations $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_3)$, by exploiting the fact that in this case any odd, reducible but indecomposable representation must have dihedral image. Our methods give an affirmative answer to Wiles' question for representations $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_5)$ which have determinant the cyclotomic character and which are unramified at 3, as well as to representations $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_4)$ which are unramified at 3 and 5.

The question of realizing Galois groups of number fields as the action of Galois on division points of abelian varieties, especially elliptic curves, has a long history, going back at least to Hermite [H]; see also Klein [K] and for a modern explanation Serre [S1]. Our results seem to be new for the group $GL_2(\mathbb{F}_4)$, while one of us (RT) explained to Wiles in 1992 that the Shimura-Taniyama conjecture implied the result for $GL_2(\mathbb{F}_5)$ (see [W]). (Serre has informed us that he was already aware of 1.1 below.)

§1. Elliptic curves and mod 5 representations

If $N \geq 3$ is an integer we will let $X_N/\mathbb{Q}$ denote the fine moduli scheme of elliptic curves E with an isomorphism $E[N] \cong \mu_N \times \mathbb{Z}/N\mathbb{Z}$, where μ_N denotes the group scheme of N^{th} roots of unity and where the isomorphism takes the Weil pairing on $E[N]$ to the standard pairing on $\mu_N \times \mathbb{Z}/N\mathbb{Z}$:

$$(\zeta_1, a_1) \times (\zeta_2, a_2) \mapsto \zeta_2^{a_1}/\zeta_1^{a_2}.$$

We will let $\mathcal{E} \to X_N$ denote the universal elliptic curve. It is known that $X_N/\mathbb{Q}$ is smooth, geometrically irreducible and of dimension 1. We will let $\overline{X}_N$ denote the unique smooth projective model of its function field. Then $X_N \subset \overline{X}_N$. (In fact $\overline{X}_N$ is the fine moduli space of generalised elliptic curves all of whose geometric fibres are either smooth or N-gons together with a level N structure in the above sense (see [DR]).)

If we fix a primitive N^{th} root of unity ζ_N, then the action of $SL_2(\mathbb{Z}/N\mathbb{Z})$ on $(\mu_N \times \mathbb{Z}/N\mathbb{Z})/\mathbb{Q}(\zeta_N)$ induces a homomorphism

$$\theta_N : SL_2(\mathbb{Z}/N\mathbb{Z}) \longrightarrow \mathrm{Aut}(\mathcal{E} \to X_N)(\mathbb{Q}(\zeta_N)).$$

Moreover if $\chi_N : G_{\mathbb{Q}} \to (\mathbb{Z}/N\mathbb{Z})^{\times}$ denotes the cyclotomic character, then we have that

$$\sigma(\theta_N(\alpha)) = \theta_N\left(\begin{pmatrix}\chi_N(\sigma) & 0\\ 0 & 1\end{pmatrix}\alpha\begin{pmatrix}\chi_N(\sigma)^{-1} & 0\\ 0 & 1\end{pmatrix}\right).$$

Because of the uniqueness of the smooth projective model we see that this action extends uniquely to one on $\overline{X}_N$.

Let K be a field of characteristic zero. Let $\bar{\rho} : G_K \to GL_2(\mathbb{Z}/N\mathbb{Z})$ be a continuous representation with $\det \bar{\rho} = \chi_N$. We define

$$\phi_{\bar{\rho}} : G_K \longrightarrow \mathrm{Aut}(\mathcal{E} \to \overline{X}_N)(\bar{K})$$

by

$$\sigma \longmapsto \theta_N \left(\bar{\rho}(\sigma) \begin{pmatrix} \chi_N(\sigma)^{-1} & 0 \\ 0 & 1 \end{pmatrix} \right).$$

This is a 1-cocycle; that is, $\phi_{\bar{\rho}}(\sigma\tau) = \phi_{\bar{\rho}}(\sigma)\sigma(\phi_{\bar{\rho}}(\tau))$. We define $\mathcal{E}_{\bar{\rho}} \to \overline{X}_{\bar{\rho}}$ to be the twist of $\mathcal{E} \to X_N$ by $\phi_{\bar{\rho}}$, and we will denote by $X_{\bar{\rho}} \subset \overline{X}_{\bar{\rho}}$ the twist of $X_N \subset \overline{X}_N$. If $x \in X_{\bar{\rho}}(K)$ and if E_x denotes the fibre of $\mathcal{E}_{\bar{\rho}}$ above x, then the action of G_K on $E_x[N]$ is via the representation $\bar{\rho}$, i.e.

$$i(\sigma a) = \bar{\rho}(\sigma) i(a).$$

Lemma 1.1. *If $\bar{\rho} : G_K \to GL_2(\mathbb{F}_5)$ and $\det \bar{\rho} = \chi_5$, then $\overline{X}_{\bar{\rho}} \cong \mathbb{P}_1/\mathbb{Q}$.*

Proof. It is well known that $\overline{X}_5 \cong \mathbb{P}_1/\mathbb{Q}$. (See for instance [RS]. Alternatively note that $\overline{X}_5$ is geometrically isomorphic to $\mathbb{P}_1$ and has a point defined over $\mathbb{Q}$. One such is provided by the elliptic curve $y^2 - 10xy - 11y = x^3 - 11x^2$ and another by the generalised elliptic curve with smooth locus $\mathbb{G}_m \times \mathbb{Z}/5\mathbb{Z}$.) We can think of $\phi_{\bar{\rho}}$ as giving an element of $H^1(G_K, PSL_2(\bar{K}))$. We must show that this element is trivial. It suffices to show that it comes from an element of $H^1(G_K, SL_2(\bar{K})) = (0)$. For this it suffices to check that $\theta_5 : SL_2(\mathbb{F}_5) \to PSL_2(\bar{K})$ lifts to a homomorphism $\tilde{\theta}_5 : SL_2(\mathbb{F}_5) \to SL_2(\bar{K})$ also satisfying

$$\tilde{\theta}_5 \left(\begin{pmatrix} \chi_5(\sigma) & 0 \\ 0 & 1 \end{pmatrix} \alpha \begin{pmatrix} \chi_5(\sigma)^{-1} & 0 \\ 0 & 1 \end{pmatrix} \right) = \sigma(\tilde{\theta}_5(\alpha)),$$

for all $\sigma \in G_K$ and $\alpha \in SL_2(\mathbb{F}_5)$.

Note that $\theta_5 : PSL_2(\mathbb{F}_5) \hookrightarrow PSL_2(\bar{K})$. As $PSL_2(\mathbb{F}_5)$ has no two-dimensional non-trivial representation and a unique non-trivial double cover (see [A], page 2) we see that the preimage of $\theta_5(SL_2(\mathbb{F}_5))$ in $SL_2(\bar{K})$ is isomorphic to $SL_2(\mathbb{F}_5)$. Thus θ_5 lifts to a homomorphism $\tilde{\theta}_5 : SL_5(\mathbb{F}_5) \to SL_2(\bar{K})$, and moreover as $SL_2(\mathbb{F}_5)$ is perfect this lift is unique. Because of this uniqueness we see that

$$\alpha \longmapsto \sigma(\tilde{\theta}_5) \left(\begin{pmatrix} \chi_5(\sigma)^{-1} & 0 \\ 0 & 1 \end{pmatrix} \alpha \begin{pmatrix} \chi_5(\sigma) & 0 \\ 0 & 1 \end{pmatrix} \right)$$

must equal $\tilde{\theta}_5$, as desired. □

Theorem 1.2. *Let K be a field of characteristic zero and let $\bar{\rho} : G_K \to GL_2(\mathbb{F}_5)$ be a homomorphism with $\det \bar{\rho} = \chi_5$. Then there exists an elliptic curve E/K such that $\bar{\rho}$ is equivalent to the representation of G_K on $E[5]$, and if K is a number field we may further suppose that the image of G_K in $\mathrm{Aut}(E[3])$ contains $SL_2(\mathbb{F}_3)$.*

Proof. For the first part we need only note that because $X_{\bar{\rho}}$ is a non-empty Zariski open subset of $\mathbb{P}_1/K$, it has a k-rational point. For the second part let $Y_{\bar{\rho}}/X_{\bar{\rho}}$ denote the quotient of

$$\{(a, b) \in (\mathcal{E}_{\bar{\rho}} \times_{X_{\bar{\rho}}} \mathcal{E}_{\bar{\rho}})[3] | \langle a, b \rangle \neq 1\}$$

by the equivalence relation $(a, b) \sim (\pm a, b)$. Then geometrically $Y_{\bar{\rho}}$ is isomorphic to X_{15} and so is geometrically irreducible. Moreover $Y_{\bar{\rho}} \to X_{\bar{\rho}}$ is finite of degree 24. If K is a number field we can find (by Hilbert irreducibility) a point $x \in X_{\bar{\rho}}(K)$

such that if y is a point of $\overline{Y}_{\bar\rho}$ above x, then $[K(y):K]=24$. Then if E denotes the fibre of $\mathcal{E}_{\bar\rho}$ at x we see that the action of G_K on $E[5]$ is via $\bar\rho$. Moreover $[K(E[3]):K]\geq 24$, and so, as $GL_2(\mathbb{F}_3)$ has a unique subgroup of index 2, we see that the image of G_K in $\mathrm{Aut}(E[3])$ contains $SL_2(\mathbb{F}_3)$. □

We remark that Lemma 1.1 remains valid with the prime 3 replacing 5. Essentially the same argument applies, since $PSL_2(\mathbb{F}_3)$ has a unique non-trivial double cover and although $SL_2(\mathbb{F}_3)$ is not perfect it has no character of order two. However, there are representations $\bar\rho: G_K\to GL_2(\mathbb{Z}/4\mathbb{Z})$ with cyclotomic determinant that cannot be realized on the 4-division points of an elliptic curve. For example, suppose that $K=\mathbb{Q}$, c is complex conjugation and $\bar\rho(c)=\begin{pmatrix}-1&2\\2&1\end{pmatrix}$. Then $\bar\rho$ cannot be lifted to a λ-adic representation for any λ dividing 2.

§2. Abelian surfaces with trivial level 2 structure

Let $W=\mathbb{F}_4^2$ and fix an alternating pairing $\langle\ ,\ \rangle: W\times W\to\mathbb{F}_2$ by

$$(a_1,a_2)\times(b_1,b_2)\mapsto \mathrm{tr}_{\mathbb{F}_4/\mathbb{F}_2}(a_1b_2-a_2b_1).$$

Let $\eta=(1+\sqrt{5})/2$, let $\mathcal{O}=\mathbb{Z}[\eta]$, let τ denote the non-trivial involution of $\mathcal{O}/\mathbb{Z}$ and identify $\mathcal{O}/2\cong\mathbb{F}_4$. We will denote the group of automorphisms of W as an $\mathbb{F}_2$-vector space together with its alternating pairing by $Sp_4(\mathbb{F}_2)$. We will denote by $SL_2(\mathbb{F}_4)$ the subgroup preserving the $\mathcal{O}$ structure. We will let $\widetilde{SL}_2(\mathbb{F}_4)$ denote the subgroup generated by $SL_2(\mathbb{F}_4)$ and the element τ. There is a bijection $\{i,j\}\mapsto w_{\{i,j\}}$ from cardinality 2 subsets of $\{1,...,6\}$ to $W-\{0\}$ such that: $w_{\{i,j\}}+w_{\{j,k\}}=w_{\{i,k\}}$ if i,j,k are all distinct; $w_{\{i,j\}}+w_{\{k,l\}}=w_{\{m,n\}}$ if $\{i,j,k,l,m,n\}=\{1,...,6\}$; and $\langle w_S,w_T\rangle\equiv \#S\cap T \bmod 2$ (see [DO], page 142). Fix such an isomorphism. It gives rise to an isomorphism $S_6\tilde{\to}Sp_4(\mathbb{F}_2)$, and hence to an embedding $\tilde\jmath:\widetilde{SL}_2(\mathbb{F}_4)\hookrightarrow S_6$. This embedding arises from the action of $\widetilde{SL}_2(\mathbb{F}_4)$ on its Sylow 5-subgroups, suitably numbered.

By a principally polarised abelian surface with full level 2 structure we shall mean an abelian surface A together with a principal polarisation $\lambda: A\tilde{\to}A^\vee$ and an isomorphism $\beta: W\tilde{\to}A[2]$ which takes the Weil pairing onto our standard pairing on W. By an HB principally polarised abelian surface with full level 2 structure we shall mean the same data but in addition a homomorphism $i:\mathcal{O}\to\mathrm{End}(A)$ such that for all $x\in\mathcal{O}$ we have $i(x)^\vee\circ\lambda=\lambda\circ i(x)$ and such that the action of $\mathcal{O}$ on $A[2]$ matches (via β) that on W. We will let $\mathcal{A}_2/\mathbb{Q}$ denote the coarse moduli scheme of principally polarised abelian surfaces with full level 2 structure and $\mathcal{H}_2/\mathbb{Q}$ the coarse moduli scheme of HB principally polarised abelian surfaces with full level 2 structure. Note that $\mathcal{A}_2$ has a natural action of $Sp_4(\mathbb{F}_2)$ and that $\mathcal{H}_2$ has a natural action of $\widetilde{SL}_2(\mathbb{F}_4)$ (where τ acts by sending (A,λ,i,β) to $(A,\lambda,i\circ\tau,\beta\circ\tau)$). We have a natural morphism $\mathcal{H}_2\to\mathcal{A}_2$, compatible with the inclusion $\widetilde{SL}_2(\mathbb{F}_4)\hookrightarrow Sp_4(\mathbb{F}_2)$.

We will also let $\mathcal{H}_{2\sqrt{5}}$ denote the fine moduli space of quintuples $(A,\lambda,i,\beta,\alpha)$, where (A,λ,i,β) is a principally polarised HB abelian surface with full level 2 structure and where $\alpha:\mu_5\times\mathbb{Z}/5\mathbb{Z}\tilde{\to}A[\sqrt{5}]$ takes the standard pairing on $\mu_5\times\mathbb{Z}/5\mathbb{Z}$ to the Weil pairing on $A[\sqrt{5}]$. (We remark that one can define a Weil pairing on $A[\sqrt{5}]$ by pulling back the standard pairing between $A[\sqrt{5}]$ and $A^\vee[\sqrt{5}^\vee]$ and using the fact that $i(\sqrt{5})^\vee\circ\lambda=\lambda\circ i(\sqrt{5})$.) Equivalently $\mathcal{H}_{2\sqrt{5}}$ is the moduli space of sextuples (A,λ,i,β,C,P), where (A,λ,i,β) is a principally polarised HB

abelian surface with full level 2 structure, $C \subset A[\sqrt{5}]$ is a subgroup of order 5 and $P \in A[\sqrt{5}] - C$. (Associate $(A, \lambda, i, \beta, \alpha)$ to $(A, \lambda, i, \beta, \alpha(\mu_5), \alpha(1,1))$. Conversely given $(A, \lambda, i, \beta, C, P)$ define α by $\alpha(1,1) = P$ and $\alpha(\zeta, 0) = Q_\zeta$, where $Q_\zeta \in C$ and the Weil pairing of Q_ζ and P equals ζ.)

We have the usual analytic description of $\mathcal{H}_{2\sqrt{5}}$ as

$$\Gamma_{2\sqrt{5}} \backslash GL_2^+(\mathbb{R})^2 / K_\infty,$$

where $\Gamma_{2\sqrt{5}}$ is the subgroup of $SL_2(\mathcal{O})$ consisting of matrices congruent to 1 modulo $2\sqrt{5}$, $GL_2^+(\mathbb{R})$ denotes the elements of $GL_2(\mathbb{R})$ of positive determinant and K_∞ denotes the commutator in $GL_2^+(\mathbb{R})^2$ of an element I_0 satisfying $I_0^2 = -1$ (I_0 is unique up to conjugation and multiplication by $(\pm 1, \pm 1)$). In particular $\mathcal{H}_{2\sqrt{5}}$ is geometrically irreducible. (We briefly recall how this analytic description of $\mathcal{H}_{2\sqrt{5}}$ is obtained. If we fix a primitive 5^{th} root of unity in $\mathbb{C}$, then to give a quintuple $(A, \lambda, i, \beta, \alpha)/\mathbb{C}$ is the same as giving the following data. A finitely generated torsion free $\mathcal{O}$-module Λ of rank 2; a perfect alternating pairing $e : \Lambda \times \Lambda \to \mathbb{Z}$ such that $e(ax, y) = e(x, ay)$ if $a \in \mathcal{O}$ and $x, y \in \Lambda$; an isomorphism $\beta : W \tilde{\to} (1/2\Lambda)/\Lambda$ as $\mathbb{F}_4$-vector spaces with an alternating pairing; an isomorphism $\alpha : \mathbb{F}_5^2 \tilde{\to} (1/\sqrt{5}\Lambda)/\Lambda$ taking the standard alternating pairing on $\mathbb{F}_5^2$ to the one induced by e on $(1/\sqrt{5}\Lambda)/\Lambda$; and an element $I \in \mathrm{Aut}_{\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{R}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R})$ such that $I^2 = -1$ and $(x, y) \mapsto e(x, Iy)$ is a symmetric positive definite bilinear form on $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$. Because $\mathcal{O}$ has class number 1 we may take $\Lambda = \mathcal{O}^2$, e to be the pairing $(x_1, x_2) \times (y_1, y_2) \mapsto \mathrm{tr}_{\mathbb{Q}(\sqrt{5})/\mathbb{Q}}(x_1 y_2 - x_2 y_1)/\sqrt{5})$, and α and β to be the obvious maps. The automorphisms of this structure are $\Gamma_{2\sqrt{5}}$. It remains to describe the possible choices for I. As $\mathrm{Aut}_{\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{R}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R}) \cong GL_2(\mathbb{R})^2$, we need a pair $I = (I_1, I_2) \in GL_2(\mathbb{R})^2$ such that $I^2 = -1$. Then we have automatically that $(x, y) \mapsto e(x, Iy)$ is symmetric. There is a unique $GL_2(\mathbb{R})^2$-conjugacy class of possible I, which splits into four $GL_2^+(\mathbb{R})^2$-conjugacy classes. A unique one of these makes $(x, y) \mapsto e(x, Iy)$ positive definite.)

We now proceed to describe some concrete models for $\mathcal{H}_2$ and $\mathcal{A}_2$. Our aim is simply to introduce a certain model Y for $\mathcal{H}_2$ and to prove Proposition 2.5.

Let $\mathcal{A}_2^*$ (resp. $\mathcal{H}_2^*$) denote the Satake compactification of $\mathcal{A}_2$ (resp. $\mathcal{H}_2$). Then there is an embedding $\mathcal{A}_2^* \hookrightarrow \mathbb{P}_4$ given by the modular forms of weight 2 (which are spanned by the 4^{th} powers of the 10 even theta-nulls). Moreover the image is a quartic hypersurface and the action of $Sp_4(\mathbb{F}_2)$ is by linear transformations of $\mathbb{P}_4$ (see [vG], pages 217–218). Consider the map $\mathcal{H}_2^* - \to \mathcal{A}_2^* \hookrightarrow \mathbb{P}_4$. It is given by some space of weight 2 modular forms on $\mathcal{H}_2$. This space must be $\widetilde{SL}_2(\mathbb{F}_4)$-invariant. The full space of weight 2 modular forms on $\mathcal{H}_2$ gives an embedding $\mathcal{H}_2^* \hookrightarrow \mathbb{P}^4$ and the image is the intersection of a quadric and a quartic (see [vG], pages 190–191). Moreover the action of $\widetilde{SL}_2(\mathbb{F}_4) \cong S_5$ is by its permutation representation on $\mathbb{P}_4$ (loc. cit.). Thus the map $\mathcal{H}_2^* - \to \mathcal{A}_2^* \subset \mathbb{P}_4$ is given by projection onto an $\widetilde{SL}_2(\mathbb{F}_4)$-invariant subspace of $\mathbb{P}^4$. Using the explicit equations (over $\mathbb{C}$) for the image of $\mathcal{H}_2^* \hookrightarrow \mathbb{P}_4$ it is not hard to see that none of the (three) projections to proper invariant subspaces is injective on $\mathcal{H}_2$. Thus the weight 2 modular forms on $\mathcal{A}_2$ must restrict to the whole space of weight 2 modular forms on $\mathcal{H}_2$. Moreover $\mathcal{H}_2^* \subset \mathcal{A}_2^* \subset \mathbb{P}_4$, with $\mathcal{A}_2^*$ a quartic hypersurface and $\mathcal{H}_2^*$ the intersection of a quadric and a quartic. Degree considerations show that $\mathcal{H}_2^*$ is a quadric section of $\mathcal{A}_2^*$. The intersection of the singular locus of $\mathcal{A}_2^*$ (i.e. $\mathcal{A}_2^* - \mathcal{A}_2$) with $\mathcal{H}_2^*$ is contained in the singular locus of $\mathcal{H}_2^*$ (i.e. $\mathcal{H}_2^* - \mathcal{H}_2$) (because $\mathcal{H}_2 \to \mathcal{A}_2$).

Consider the space $P_1^n = PGL_2\backslash(\mathbb{P}_1^n)^{ss}$ (see for example [DO]). There is a natural map $\pi : P_1^{n+1} \to P_1^n$, which is easily seen to be a $\mathbb{P}_1$-bundle in the Zariski topology on the stable locus $(P_1^n)^s$ (see Chapter II of [DO]). It also comes equipped with n-sections $s_1, ..., s_n$ which are induced by the maps

$$s_i : \mathbb{P}_1^n \longrightarrow \mathbb{P}_1^{n+1},$$

$$a_1 \times ... \times a_n \mapsto a_1 \times ... \times a_n \times a_i.$$

The symmetric group S_n acts on P_1^n and P_1^{n+1} so that these actions commute with π and so that

$$\sigma \circ s_i = s_{\sigma i} \circ \sigma.$$

The variety P_1^6 can be identified with the closed subset of $\mathbb{P}_5$ defined by $\sum_1^6 z_i = \sum_1^6 z_i^3 = 0$ (see page 16 of [DO]). Note that $\mathrm{Out}(S_6) \cong C_2$ (see [A], page 4). If we denote the embedding $P_1^6 \hookrightarrow \mathbb{P}_5$ by α, then we have that for all $z \in P_1^6$ and all $\sigma \in S_6$:

$$(\alpha \circ \sigma(z))_{\sigma^\dagger(i)} = (\alpha(z))_i$$

where $\dagger : S_6 \to S_6$ is some non-inner automorphism. (If it were inner we could assume that $(\alpha \circ \sigma(z))_{\sigma i} = (\alpha(z))_i$. Then the -1 eigenspace of the involution $\sigma = (12)$ on $(R_1^6)_1$ (see [DO], page 9) would have dimension 1, but in the notation of [DO], page 15, both t_0 and t_2 are -1 eigenvectors, a contradiction.) The stable locus $(P_1^6)^s$ coincides with the smooth locus of P_1^6 (see [DO], page 31) which can be identified as P_1^6 less the 15 S_6-conjugates of $(1 : 1 : 1 : -1 : -1 : -1)$. (Although the arguments of [DO] work over $\mathbb{C}$ the arguments for these facts work equally well over $\mathbb{Q}$.)

Over the open subset U of P_1^6 where the sections $s_1, ..., s_6$ are distinct, we have a principally polarised abelian surface with full level 2 structure obtained as the Jacobian of a hyperelliptic curve over U which maps $2:1$ to $P_1^7|_U$ ramified exactly at $s_1, ..., s_6$. This gives a morphism $U \to \mathcal{A}_2$, which is injective on geometric points by Torelli's theorem. As P_1^6 and $\mathcal{A}_2$ are both three-dimensional we obtain a birational equivalence $P_1^6 - \to \mathcal{A}_2^*$ which gives an isomorphism between U and the open subset of $\mathcal{A}_2$ corresponding to Jacobians. Under this equivalence the $Sp_4(\mathbb{F}_2)$ action on $\mathcal{A}_2$ is compatible with the S_6 action on P_1^6 under our isomorphism $Sp_4(\mathbb{F}_2) \cong S_6$ (see [DO], page 142).

Lemma 2.1. *The birational equivalence between $P_1^6 \subset \mathbb{P}^4$ and $\mathcal{A}_2^* \subset \mathbb{P}^4$ is given by projective duality (up to an automorphism of $\mathbb{P}_4$).*

Proof. By Thomae's formula (see [M], page 120), the fact that $\mathcal{A}_2^* \hookrightarrow \mathbb{P}_4$ is given by the 4^{th} powers of the 10 even theta-nulls and the formulae in [DO], pages 15–17, we see that it is given by quadratic expressions in the coordinates of $P_1^6 \subset \mathbb{P}_4$. As a representation of S_6 these span a space isomorphic to $\mathrm{Symm}^2 W_5 \cong W_1 \oplus W_5 \oplus W_9$ where W_i is an irreducible representation of dimension i. As the rational map $P_1^6 - \to \mathcal{A}_2^* \subset \mathbb{P}_4$ is equivariant for a linear S_6 action on $\mathbb{P}_4$ we see that it is given by the unique irreducible five-dimensional space of quadratic functions on $P_1^6 \subset \mathbb{P}_4$. Such a space is provided by the space spanned by the partial derivatives of the defining equation of $P_1^6 \subset \mathbb{P}_4$. The lemma follows. □

Let Y denote the smooth cubic surface in $\mathbb{P}_4$ defined by $\sum_1^5 y_i = \sum_1^5 y_i^3 = 0$. Note that the permutation action of S_5 on $\mathbb{P}_4$ preserves Y and so we get an action of

S_5 on Y (so that $(\sigma(y))_{\sigma i} = (y)_i$). Let Y^1 denote Y less the 10 S_5-conjugates of $(0 : 0 : 0 : 1 : -1)$, let $\tilde{Y}$ denote Y blown up in the 10 S_5-conjugates of $(0 : 0 : 0 : 1 : -1)$ and let Y^0 denote Y less the 15 lines which are S_5 conjugate to $(s : -s : t : -t : 0)$. According to [vG], pages 190–191, we have the following facts. The surface Y is birationally equivalent to $\mathcal{H}_2^*$ in such a way that the $\widetilde{SL}_2(\mathbb{F}_4)$ action on $\mathcal{H}_2^*$ goes over to the S_5 action on Y under an isomorphism $\widetilde{SL}_2(\mathbb{F}_4) \cong S_5$ (which is unique up to inner automorphism: this can be easily deduced from $\mathrm{Out}(A_5) \cong C_2$, see [A], page 2). We fix such a birational equivalence and the corresponding isomorphism $\widetilde{SL}_2(\mathbb{F}_4) \cong S_5$. Moreover Y^0 becomes isomorphic to an open subset of $\mathcal{H}_2$. The birational map $Y - \to \mathcal{H}_2^* \subset \mathbb{P}_4$ is given (up to automorphism of $\mathbb{P}_4$) by

$$(y_1 : \ldots : y_5) \mapsto (1/y_1 : \ldots : 1/y_5).$$

It can be factored $Y \leftarrow \tilde{Y} \to \mathcal{H}_2^*$, where $\tilde{Y} \to \mathcal{H}_2^*$ contracts the strict transforms of the five hyperplane sections $Y \cap \{y_i = 0\}$ to the five singular points of $\mathcal{H}_2^*$ but is otherwise an isomorphism. (The arguments of [vG] are over $\mathbb{C}$ but work equally well over $\mathbb{Q}$. To explain this we will use for the rest of this paragraph the notation of pages 190–191 of [vG]. Note that the sheaf of weight 2 modular forms on X_Γ, the set of cusps and the action of $\widetilde{SL}_2(\mathbb{F}_4)$ are all defined over $\mathbb{Q}$. As each cusp is stabilised by a different subgroup of $\widetilde{SL}_2(\mathbb{F}_4)$ we see that each cusp is individually defined over $\mathbb{Q}$. The Eisenstein series E_i span the space of weight 2 modular forms. The only ones that vanish at all cusps except cusp 0 are the scalar multiples of E_0. Thus replacing E_0 by some scalar multiple we may assume that it is defined over $\mathbb{Q}$. As a subgroup of $\widetilde{SL}_2(\mathbb{F}_4)$ of index 5 fixes the line generated by E_0 we see that E_0 has 5 $\widetilde{SL}_2(\mathbb{F}_4)$-conjugates all defined over $\mathbb{Q}$. Replacing the E_i by these new conjugates, it is now clear that the arguments of [vG] work over $\mathbb{Q}$.)

We will let E denote the sum of the exceptional divisors of $\tilde{Y} \to Y$ and $\tilde{H}$ the pull back of the hyperplane section H of Y. Note that the sum of the exceptional divisors of $\tilde{Y} \to \mathcal{H}_2^*$ is the sum of the strict transforms of the five hyperplane sections $Y \cap \{y_i = 0\}$ of Y. Hence it is linearly equivalent to $5\tilde{H} - 3E$.

Lemma 2.2. *Y contains six disjoint lines each defined over $\mathbb{Q}(\sqrt{5})$ which are permuted amongst themselves by the action of A_5.*

Proof. The A_5 orbit of the line

$$\{t_1(1 : -1 : \eta : -\eta : 0) + t_2(0 : -\eta : \eta : -1 : 1)\}$$

will do. □

Corollary 2.3. *The pull back under $\tilde{Y} \to \mathcal{H}_2^* \subset \mathbb{P}_4$ of a hyperplane section is linearly equivalent to $4\tilde{H} - 2E$.*

Proof. The rational map $Y - \to \mathcal{H}_2^* \subset \mathbb{P}_4$ is given by a system of quartics whose only base points are the 10 S_5-conjugates of $(0 : 0 : 0 : 1 : -1)$. Thus the pull back to $\tilde{Y}$ of a hyperplane section is $4\tilde{H}$ plus a divisor supported on the exceptional divisors of $\tilde{Y} \to Y$. Because the S_5 action on $\tilde{Y}$ is compatible with a linear action on $\mathcal{H}_2^* \subset \mathbb{P}_4$ we see that the pull back of a hyperplane section of $\mathcal{H}_2^*$ defines an S_5-invariant class in $\mathrm{Pic}(\tilde{Y})$. Thus it is of the form $4\tilde{H} + mE$. On the other hand if E_1 denotes the preimage of $(0 : 0 : 0 : 1 : -1)$ in $\tilde{Y}$, then the image of E_1 in $\mathcal{H}_2^*$ is the conic consisting of points $(a : b : c : 0 : 0)$ such that $ab + bc + ca = 0$. Thus $-m = (E_1, 4\tilde{H} + mE) = 2$ and the corollary follows. □

The image of $\mathcal{H}_2$ in $\mathcal{A}_2$ contains a point corresponding to a Jacobian (e.g. the Jacobian of $y^2 = x(x^5 - 1)$). Thus the composite of rational maps

$$j : Y - \to \mathcal{H}_2^* \longrightarrow \mathcal{A}_2^* - \to P_1^6$$

makes sense. The actions of $\widetilde{SL}_2(\mathbb{F}_4) \cong S_5$ and of S_6 are compatible with $\tilde{\jmath} : \widetilde{SL}_2(\mathbb{F}_4) \hookrightarrow S_6$.

Lemma 2.4. *The rational map j is in fact a morphism and, up to conjugation by an element of S_6, is given by*

$$j : Y \longrightarrow P_1^6 \cong Z,$$

$$(y) \mapsto (4y_1^2 - s : 4y_2^2 - s : 4y_3^2 - s : 4y_4^2 - s : 4y_5^2 - s : s),$$

where $s = y_1^2 + y_2^2 + y_3^2 + y_4^2 + y_5^2$.

Proof. The action of S_5 on $\tilde{Y}$ is compatible with a linear action on $P_1^6 \subset \mathbb{P}_4$ and so the pull back under j of a hyperplane section of P_1^6 defines an S_5-invariant class in $\mathrm{Pic}(\tilde{Y})$. The map $\mathcal{H}_2^* \hookrightarrow \mathcal{A}_2^* - \to P_1^6 \subset \mathbb{P}_4$ is given by a linear system of cubics whose only base points are the intersection of $\mathcal{H}_2^*$ with the singular locus of $\mathcal{A}_2^*$ and so are contained in the singular locus of $\mathcal{H}_2^*$. Thus the pull back of a hyperplane section of P_1^6 to $\tilde{Y}$ is of the form $3(4\tilde{H} - 2E)$ plus a divisor supported on the exceptional divisors of $\tilde{Y} \to \mathcal{H}_2^*$. Using the S_5-invariance we see that it must in fact be

$$3(4\tilde{H} - 2E) - r(5\tilde{H} - 3E) = (12 - 5r)\tilde{H} + 3(r - 2)E,$$

where $r \in \mathbb{Z}_{\geq 0}$.

We will show that under $\tilde{Y} - \to P_1^6$ an irreducible component E_i of E gets contracted to a point. According to [vG], pages 190–191, the image of E_i in $\mathcal{H}^*$ parametrises principally polarised abelian surfaces which split as a product of two elliptic curves. Hence by [vG], page 217, we see that the image of E_i in $\mathcal{A}^*$ is contained in the locus of vanishing of some theta-null. However it is not difficult to see that the hyperplane defined by the vanishing of a theta-null is tangent to $\mathcal{A}$ at every point on its intersection with $\mathcal{A}$. It follows that E_i is contracted to a point in P_1^6. (To verify the fact that the hyperplane defined by the vanishing of a theta-null is tangent to $\mathcal{A}$ at every point of contact with $\mathcal{A}$ we use the model of $\mathcal{A}^* \subset \mathbb{P}^5$ given in the notation of page 217 of [vG] by $s_1 = s_2^2 - 4s_4 = 0$. We see that the singular locus of $\mathcal{A}^*$ consists of the S_6-conjugates of the line $(r : r : s : s : t : t)$ where $r + s + t = 0$. Finding the six sets of five disjoint lines of singularity we see that with a suitable numbering the triple $\{\lambda_{12}, \lambda_{13}, \lambda_{23}\}$ (see [vG], page 216) becomes the set of lines $\{(r : s : t : r : s : t), (r : s : t : s : t : r), (r : s : t : t : r : s)\}$. The unique hyperplane containing these three lines is $x_1 + x_2 + x_3 = x_4 + x_5 + x_6 = 0$ and so the 10 S_6-conjugates of this hyperplane are the 10 hyperplanes defined by the vanishing of an even theta-null. These hyperplanes each meet $\mathcal{A}^*$ in a quadric taken with multiplicity 2.)

Hence either we have $(E, (12-5r)\tilde{H}+3(r-2)E) = 0$ or $\tilde{Y} - \to P_1^6$ has base points on E. We deduce that either $r = 2$ or that all points in the intersection of E with an exceptional divisor of $\tilde{Y} - \to \mathcal{H}_2^*$ are in the base locus of $\tilde{Y} - \to P_1^6$ (because the base locus is S_5-invariant). In the latter case $((12-5r)\tilde{H}+3(r-2)E, 5\tilde{H}-3E) \geq 30$ and so $r \geq 2$. On the other hand $(12 - 5r)\tilde{H} + 3(r - 2)E$ must be effective so that $r \leq 2$. We deduce that $r = 2$ and that the pull back under j of a hyperplane section is linearly equivalent to $2H$ on Y.

Let $V \subset |2H|$ be the linear system defining j. As a representation of S_5 we have $V \subset \mathrm{Symm}^2 V_4 \cong V_1 \oplus V_4 \oplus V_5$, where V_1 is the trivial representation, $V_1 \oplus V_4$ is the five-dimensional permutation representation and V_5 is an irreducible five-dimensional representation. Thus $V = V_1 \oplus V_4$ and is spanned by $y_1^2, ..., y_5^2$. Using the S_5-symmetry and letting $s = y_1^2 + ... + y_5^2$ we see that up to composition with an element of S_6

$$j : (y_1 : ... : y_5) \longmapsto (Ay_1^2 + Bs : ... : Ay_5^2 + Bs : -(A + 5B)s).$$

In order that $(A + 5B)^3 s^3 - \sum_{i=1}^5 (Ay_i^2 + Bs)^3$ vanish on Y it is easy to check that we must have $A/B = -4$. Thus j has the form described in the lemma as is easily checked to be everywhere defined. □

We remark that the map $Y \hookrightarrow P_1^6$ described in [vG], pages 191, 218 and 220, is not the one induced by their moduli space interpretation. The correct map is the one given in the above lemma.

We see that $j(Y^1) \subset (P_1^6)^s$. We will let C denote the pull back to Y^1 of P_1^7/P_1^6. We will now summarise what we will need from the proceeding calculations.

Proposition 2.5. *The surface $\mathcal{H}_2$ is birationally equivalent to Y. Under this equivalence Y^0 embeds as an open subset of $\mathcal{H}_2$ and the action of S_5 on Y matches with the action of $\widetilde{SL}_2(\mathbb{F}_4)$ on $\mathcal{H}_2$ under an isomorphism $S_5 \cong \widetilde{SL}_2(\mathbb{F}_4)$ (which is unique up to inner automorphism). Moreover there is a $\mathbb{P}_1$-bundle for the Zariski topology C/Y^1 which has an action of $\widetilde{SL}_2(\mathbb{F}_4)$ compatible with the action on Y^1 and six sections $s_1, ..., s_6$ such that if $\sigma \in \widetilde{SL}_2(\mathbb{F}_4)$, then*

$$\sigma \circ s_i = s_{\tilde{j}(\sigma) i} \circ \sigma.$$

If y is a geometric point of Y^0, then the principally polarised abelian surface which is parametrised by y can be realised as the Jacobian of the genus 2 curve which maps $2:1$ to C_y ramified exactly at $s_1(y), ..., s_6(y)$. Over the $\widetilde{SL}_2(\mathbb{F}_4)$-conjugates of $(1 : -1 : 1 : -1 : 0)$ the six sections $s_1, ..., s_6$ coalesce in pairs.

Proof. The only part which needs further explanation is the final sentence. Note that $j(1 : -1 : 1 : -1 : 0) = (0 : 0 : 0 : 0 : -1 : 1)$. Also from the formulae in [DO], pages 15–17, we see that $(0 : 0 : 0 : 0 : -1 : 1)$ parametrises 6-tuples (P, Q, R, R, Q, P) with P, Q, R distinct points in $\mathbb{P}_1$. □

As $\mathcal{H}_2$ is a coarse moduli space we cannot hope for a universal family of abelian surfaces over $\mathcal{H}_2$. However one might hope for a universal family of Kummer surfaces, or, what comes to much the same thing, a universal family of conics with six marked points. Our family C is probably such a family. At least it has a natural action of $\widetilde{SL}_2(\mathbb{F}_4)$ and from a geometric fibre of this family we can reconstruct the abelian surface associated to that point in Y^0. We cannot use C to construct a universal family of abelian surfaces as there is no canonical choice of genus 2 curve mapping $2:1$ to C ramified at the six marked points. Also note that our calculations show that C extends to a larger open set (namely Y^1) than one might have expected. This will be crucial in the sequel.

The surface Y is not rational over $\mathbb{Q}$. It will be useful to replace it by a related variety X which is rational. In fact X will be a certain blow up of the restriction of scalars from $\mathbb{Q}(\sqrt{5})$ to $\mathbb{Q}$ of $Y \times_{\mathbb{Q}} \mathbb{Q}(\sqrt{5})$. More precisely consider the blow up X' of $Y \times Y$ along its diagonal. It has an involution t which switches the two factors

and an action of $\widetilde{SL}_2(\mathbb{F}_4)$ which commutes with t and is compatible with the two projections $X' \to Y$. We also have a rational map

$$\theta : X' - \to Y$$

which takes a pair $(y^{(1)}, y^{(2)})$ to the third point of intersection of the line through $y^{(1)}$ and $y^{(2)}$ with Y. The rational map θ is well defined outside the strict transforms of the subvarieties $L \times L \subset Y \times Y$ as L runs over lines lying in Y. Also θ is compatible with the $\widetilde{SL}_2(\mathbb{F}_4)$ actions on X' and Y. We will let X denote the twist of X' by the homomorphism

$$G_{\mathbb{Q}} \to \mathrm{Gal}(\mathbb{Q}(\sqrt{5})/\mathbb{Q}) \tilde{\to} \{1, t\} \subset \mathrm{Aut}(X'/\mathbb{Q}).$$

Then X inherits an action of $\widetilde{SL}_2(\mathbb{F}_4)$ and a rational map $\theta : X - \to Y$ compatible with the $\widetilde{SL}_2(\mathbb{F}_4)$ actions, all still defined over $\mathbb{Q}$. We note that X is birationally equivalent over $\mathbb{Q}$ to the restriction of scalars from $\mathbb{Q}(\sqrt{5})$ to $\mathbb{Q}$ of $Y \times_{\mathbb{Q}} \mathbb{Q}(\sqrt{5})$. We let $X^1 \subset X$ denote the locus where θ is defined and maps to Y^1. We note that the complement of X^1 in X has codimension at least 2. We will also let D denote the pull back via θ of C to X^1. It inherits an action of $\widetilde{SL}_2(\mathbb{F}_4)$ and six sections $s_1, ..., s_6$.

Over some non-empty open subset $X^0 \subset X$, $s_1, ..., s_6$ are distinct and we can form an abelian surface A/X^0 as the Jacobian of a genus 2 curve over X^0 which maps 2 : 1 to D ramified exactly at $s_1, ..., s_6$. Then A comes with a principal polarisation $\lambda : A \to A^\vee$ and an isomorphism $\beta : A[2] \tilde{\to} W$ which takes the Weil pairing to the standard pairing on W. Note however that the construction of A is not canonical and that A does not have an action of $SL_2(\mathbb{F}_4)$. Let x (resp. $\overline{x}$) denote the generic (resp. geometric generic) point of X. From the interpretation of Y^0 as a coarse moduli space we see that there is an embedding $i : \mathcal{O} \hookrightarrow \mathrm{End}(A_{\overline{x}})$ which is compatible via θ with the moduli theoretic interpretation of Y and such that the action of $\mathcal{O}$ (via i) on $A_{\overline{x}}[2]$ matches that on W. In fact there is a unique such map i. For if i' were a second such map, then $i' = f \circ i \circ f^{-1}$ for some automorphism f of $A_{\overline{x}}$ which is the identity on $A_{\overline{x}}[2]$. We see that the representation $T_2(f)$ on the 2-adic Tate module of A can only have eigenvalues ± 1. Suppose that $f \neq \pm 1$. Then $(1+f)A$ must be an elliptic curve, A must be isogenous to $(1+f)A \oplus (1-f)A$ and we have a non-zero isogeny

$$(1 + f)A \longrightarrow A \longrightarrow (1 - f)A,$$

where the first map is multiplication by $i(\sqrt{5})$ and the second $(1 - f)$. This will all be defined over a non-empty open subset U of some finite cover of $\mathcal{H}_2$. Then we can map U to the j-line by taking z to the j-invariant of $(1 + f)A_z$. It is easy to see that this morphism has countable fibres and so must in fact have finite fibres, a contradiction. Thus $f = \pm 1$ and $i' = i$. In particular we see that i is stable under the action of $\mathrm{Gal}(k(\overline{x})/k(x))$ and so defined (non-canonically) over $k(x)$. Thus replacing X^0 by a smaller non-empty open subset we see that there is a homomorphism $i : \mathcal{O} \hookrightarrow \mathrm{End}(A)$ such that $i(a)^\vee \circ \lambda = \lambda \circ i(a)$ for all $a \in \mathcal{O}$ and such that the actions of $\mathcal{O}$ on $A[2]$ and W match under β.

We will let

$$\tilde{X}^0 = \{(a, b) \in (A \times_{X^0} A)[\sqrt{5}] | \langle a, b \rangle \neq 1\} / \sim,$$

where $(a, b) \sim (a', b')$ if and only if $(a, b) = \pm(\mu a', b')$ for some $\mu \in \mathbb{F}_5^\times$. (Although the construction of A is non-canonical, $A_{\overline{x}}$ is canonical and the only automorphisms

of $(A_{\overline{x}}, \lambda, i, \beta)$ are ± 1. Thus at least generically the construction of $\tilde{X}^0$ is canonical (i.e. if we start with a different hyperelliptic curve mapping $2:1$ to F we get in some neighbourhood of the generic point of X^0 the same variety $\tilde{X}^0 \to X^0$).)

Lemma 2.6. *After perhaps again shrinking X^0, $\tilde{X}^0$ is geometrically irreducible.*

Proof. We shall work over $\mathbb{C}$. We can perform the same construction over some non-empty open subset $Y^{00} \subset Y^0$, except that we will use the equivalence relation $(a,b) \sim (a',b')$ if and only if $(a,b) = (\mu a', b')$ for some $\mu \in \mathbb{F}_5^\times$. We obtain a variety Z/Y^{00} which has an action of $\{\pm 1\}$ which associates $(y, [(a,b)])$ with $(y, [(-a,-b)])$. We will denote by $\tilde{Y}^0$ the quotient $Z/\{\pm 1\}$. Using the fact that the construction of $\tilde{X}^0$ is generically canonical we see that (after perhaps shrinking X^0) $\tilde{X}^0 = \tilde{Y}^0 \times_{Y^{00}} X^0$. After perhaps again shrinking X^0 we may suppose that $\tilde{X}^0$ dominates X^0. Assume for now that $\tilde{Y}^0$ is irreducible. The general fibres of $X^0 \to Y^{00}$ and hence of $\tilde{X}^0 \to \tilde{Y}^0$ are irreducible (in fact birationally equivalent to Y). If y is the generic point of Y^{00}, then $\tilde{Y}^0_y$ is a field and $\tilde{X}^0_y$ is irreducible. Thus the generic fibre of $\tilde{X}^0 \to X^0$ is irreducible and hence $\tilde{X}^0$ is irreducible.

It remains to show that $\tilde{Y}^0$ is irreducible. We have a family $\tilde{A}/Z$ of principally polarised HB abelian surfaces with full level 2 structure and with a subgroup $C \subset A[\sqrt{5}]$ of order 5 and a point $P \in \tilde{A}[\sqrt{5}] - C$. (Let $\tilde{A} = Z \times_{Y^{00}} A$, let C be the subgroup generated by a over a point $(y, [(a,b)])$ and let $P = b$.) Thus we get a map $Z \to \mathcal{H}_{2\sqrt{5}}$ over $Y^{00} \to \mathcal{H}_2$. Looking at points one sees that $Z \to \mathcal{H}_{2\sqrt{5}}$ factors through a map $\tilde{Y}^0 \to \mathcal{H}_{2\sqrt{5}}$. Because $[\tilde{Y}^0 : Y^{00}] \leq 60$ and $[\mathcal{H}_{2\sqrt{5}} : \mathcal{H}_2] = 60$ we see that $\tilde{Y}^0$ is birationally equivalent to $\mathcal{H}_{2\sqrt{5}}$ and hence irreducible. □

§3. Abelian surfaces with twisted level 2 structure

We will now fix both a field K of characteristic zero and a homomorphism $\bar{\rho} : G_K \to SL_2(\mathbb{F}_4)$. If V/K is a variety with an action of $SL_2(\mathbb{F}_4)$ defined over K we will let $V_{\bar{\rho}}$ denote the twist of V by $\bar{\rho} : G_K \to SL_2(\mathbb{F}_4) \to \mathrm{Aut}(V)(K)$. Thus we have

$$\begin{array}{ccccc} D_{\bar{\rho}} & \longrightarrow & X^1_{\bar{\rho}} & \subset & X_{\bar{\rho}} \\ \downarrow & & \downarrow & & \\ C_{\bar{\rho}} & \longrightarrow & Y^1_{\bar{\rho}} & \subset & Y_{\bar{\rho}}. \end{array}$$

$D_{\bar{\rho}}/X^1_{\bar{\rho}}$ is now a $\mathbb{P}_1$-bundle in the étale topology, but not necessarily in the Zariski topology. We have sections $s_1, ..., s_6 : X^1_{\bar{\rho}} \times \bar{K} \to D_{\bar{\rho}} \times \bar{K}$ such that for all $\sigma \in G_K$

$$\sigma(s_i) = s_{\tilde{j} \circ \bar{\rho}(\sigma) i}.$$

In particular the collection $\{s_1, ..., s_6\}$ is defined over K.

Lemma 3.1. *The surface $Y_{\bar{\rho}}$ is birationally equivalent to $\mathbb{P}_2$ over $K(\sqrt{5})$. $X_{\bar{\rho}}$ is birationally equivalent to $\mathbb{P}_4$ over K.*

Proof. By Lemma 2.2 there is a system of six disjoint lines on $Y_{\bar{\rho}}$ collectively defined over $K(\sqrt{5})$. Blowing them down we obtain a variety V which is geometrically isomorphic to $\mathbb{P}_2$ and has a zero cycle of degree 10 rationally defined over $K(\sqrt{5})$, namely the $SL_2(\mathbb{F}_4)$ orbit of $(1:-1:0:0:0) \in Y(\bar{K})$. As 10 is coprime to 3 we see that $V \tilde{\to} \mathbb{P}_2$ over $K(\sqrt{5})$ (because on $V^\vee$ the line bundle $\mathcal{O}(1) \cong \mathcal{O}(10) \otimes (\Omega^2_{V^\vee})^{\otimes 3}$ is defined over $K(\sqrt{5})$). For the second part note that $X_{\bar{\rho}}$ is birationally equivalent over K to the restriction of scalars from $K(\sqrt{5})$ to K of $Y_{\bar{\rho}} \otimes K(\sqrt{5})$. □

Proposition 3.2. *$D_{\bar\rho}/X^1_{\bar\rho}$ is a $\mathbb{P}_1$-bundle in the Zariski topology.*

Proof. We will use $\mathrm{Br}_2(V)$ to denote $H^2_{et}(V,\mathbb{G}_m)[2]$, the 2-torsion in the cohomological Brauer group of V. We must check that $[D_{\bar\rho}] \in \mathrm{Br}_2(X^1_{\bar\rho})$ is trivial. If $[L:K]$ is odd, then $\mathrm{Br}_2(X^1_{\bar\rho}) \hookrightarrow \mathrm{Br}_2(X^1_{\bar\rho}\otimes L)$, and so we may reduce to the case that the image of $\bar\rho$ is contained in a Sylow 2-subgroup of $SL_2(\mathbb{F}_4)$. (We thank Bogomolov for pointing this out to us.)

The natural inclusion

$$\mathrm{Br}_2(K) \hookrightarrow \mathrm{Br}_2(X^1_{\bar\rho})$$

is an isomorphism because

$$\mathrm{Br}_2(X^1_\rho) \cong \mathrm{Br}_2(X_\rho) \cong \mathrm{Br}_2(\mathbb{P}_4/K) \cong \mathrm{Br}_2(K)$$

(using respectively: purity for the cohomological Brauer group ([G3], Corollary 6.2), birational invariance of the cohomological Brauer group ([G3], Theorem 7.4) and the calculation of the cohomological Brauer group of projective space). Thus if x is any K-rational point of X^1_ρ, then the map $\mathrm{Br}_2(X^1_{\bar\rho})\to\mathrm{Br}_2(K)$ obtained by taking the fibre at x is an isomorphism. We see that it suffices to find a K-rational point on $D_{\bar\rho}$.

Suppose as we may that $\mathrm{im}\,\bar\rho$ acts on Y so as to fix the first coordinate. Then $(0:1:1:-1:-1)$, $(0:1:-1:1:-1)$ and $(0:1:-1:-1:1) \in Y(\bar K)$ all define K-rational points on $Y^1_{\bar\rho}$. Above these points the fibre of $C_{\bar\rho}$ has a rationally defined cycle of degree 3 (coming from the sections $s_1,\dots,s_6$; see Proposition 2.5) and so splits. We can find a Severi-Brauer variety $C'_{\bar\rho}\to Y_{\bar\rho}$ such that $C'_{\bar\rho}|_{Y^1_{\bar\rho}} \sim C_{\bar\rho}$ (purity of the cohomological Brauer group: [G3], Corollary 6.2, and the cohomological description of the Brauer group of a regular surface: [G2], Theorem 2.1). The fibre of $C'_{\bar\rho}$ at these three points will also be split and so the restriction of $C'_{\bar\rho}$ to each of the three copies of $\mathbb{P}_1$ that make up $(Y\cap\{y_1=0\})_{\bar\rho}$ is split (as $\mathrm{Br}_2(K)\tilde{\to}\mathrm{Br}_2(\mathbb{P}_1/K)$). We deduce that above any K-rational point of $(Y^1\cap\{y_1=0\})_{\bar\rho}$ there is a K-rational point of $C_{\bar\rho}$. Thus it suffices to find a K-rational point of $X^1_{\bar\rho}$ which maps under θ to a point of $(Y\cap\{y_1=0\})_{\bar\rho}$.

Let L denote the line $\{(0:s:-s:t:-t)\}\subset Y$. Let $y\in L_{\bar\rho}(K)$ and suppose that, considered as a point of $Y(\bar K)$, $y\neq(0:1:-1:1:-1)$, $(0:1:-1:-1:1)$, $(0:1:-1:0:0)$ or $(0:0:0:1:-1)$. Consider the blow up of $Y_{\bar\rho}$ at y less the strict transform of $L_{\bar\rho}$, which we will denote by V. It has an involution w which maps a point y' to the third point of intersection of the line yy' with Y. Let V' denote the twist of V by the homomorphism

$$G_K \longrightarrow \mathrm{Gal}(K(\sqrt{5})/K) \hookrightarrow \{1,w\}\subset \mathrm{Aut}(V/K).$$

Then V' is isomorphic over K to the fibre of θ at y. On the other hand V' has a K-rational point, namely the intersection of the exceptional divisor with the strict transform of the conic $(Y\cap T_yY - L)_{\bar\rho}$. This concludes the proof of the proposition. □

The same arguments used before Lemma 2.6 allow one to deduce the following corollary.

Corollary 3.3. *There is a non-empty Zariski open subset $X^0(\bar\rho)\subset X^1_{\bar\rho}$, a principally polarised HB abelian surface $(A(\bar\rho),\lambda,i)/X^0(\bar\rho)$ and an isomorphism β :*

$(A(\bar\rho)\times\bar K)[2]\tilde\to W$ *such that for all* $\sigma\in G_K$

$$\beta\circ\sigma=\bar\rho(\sigma)\circ\beta.$$

Note that as $SL_2(\mathbb{F}_4)$ does not act on A/X^0 we cannot define $A_{\bar\rho}$ directly. The abelian surface $A(\bar\rho)$ may be thought of as a substitute.

We will let

$$\tilde X^0(\bar\rho)=\{(a,b)\in(A(\bar\rho)\times_{X^0(\bar\rho)}A(\bar\rho))[\sqrt{5}]|\langle a,b\rangle\neq 1\}/\sim,$$

where $(a,b)\sim(a',b')$ if and only if $(a,b)=\pm(\mu a',b')$ for some $\mu\in\mathbb{F}_5^\times$. Again, although the construction of $A(\bar\rho)$ is non-canonical, generically the construction of $\tilde X^0(\bar\rho)$ is canonical. Thus over $\bar K$ we have that $\tilde X^0(\bar\rho)$ is birationally equivalent to $\tilde X^0$ and in particular $\tilde X^0(\bar\rho)$ is absolutely irreducible (by Lemma 2.6).

Theorem 3.4. *Suppose that* K *is a field of characteristic* 0 *and suppose that* $\rho: G_K\to SL_2(\mathbb{F}_4)$ *is a continuous homomorphism. Then there is an abelian surface* A/K *together with a principal polarisation* $\lambda: A\to A^\vee$ *and an embedding* $i:\mathcal{O}\hookrightarrow\mathrm{End}(A)$ (*both defined over* K) *such that*

(1) $\lambda\circ i(a)=i(a)^\vee\circ\lambda$ *for all* $a\in\mathcal{O}$*;*

(2) *the action of* G_K *on* $A[2]\cong\mathbb{F}_4^2$ *is equivalent to* $\bar\rho$*; and*

(3) *if* K *is a number field, the action of* G_K *on* $A[\sqrt{5}]\cong\mathbb{F}_5^2$ *is via a homomorphism* $G_K\twoheadrightarrow GL_2(\mathbb{F}_5)$ *whose image contains* $SL_2(\mathbb{F}_5)$.

Proof. Choose a point $x\in X^0(\bar\rho)(K)$. If K is a number field we may, by Hilbert irreducibility, assume that if $\tilde x\in\tilde X^0(\bar\rho)$ is a point above x, then $[K(\tilde x):K]=60$. Let $A=A(\bar\rho)_x$. The only part of the theorem that is not obvious is part (3). So suppose that K is a number field. If G denotes the image of G_K in $GL_2(\mathbb{F}_5)\cong\mathrm{Aut}(A[\sqrt{5}])$, then we have that

$$\#G\Big/G\cap\left\{\begin{pmatrix}\mu&0\\0&\nu\end{pmatrix}\middle|\,\nu=\pm1,\ \ \mu\in\mathbb{F}_5^\times\right\}=60.$$

An easy bit of group theory now shows that $G\supset SL_2(\mathbb{F}_5)$. □

Corollary 3.5. *Suppose* K *is a field of characteristic zero,* $f\in K[X]$ *is a quintic polynomial with discriminant* d *and* L/K *is the splitting field for* f*. Then there is a principally polarised HB abelian surface* $A/K(\sqrt{d})$ *such that* $L=K(\sqrt{d})(A[2])$.

Hermite [H] showed that in the situation of this corollary there is a solvable extension K_1/K and an elliptic curve E/K_1 such that $K_1(E[5])\supset L$. However we do not in general have equality, nor even an equality between $K_1(E[5])$ and $L(\zeta_5)$. Apparently Kronecker found this lack of equality unsatisfactory (see [K]).

We remark, in the context of Theorem 3.4, that if F is a real quadratic field in which 2 is inert such that every $\rho: G_K\to SL_2(\mathbb{F}_4)$ can be realized on $A[2]$ for some abelian surface A/K with multiplication by $\mathcal{O}_F$, then necessarily $F=\mathbb{Q}(\sqrt{5})$. The reason is that then we could, for any prime $p\equiv1 \bmod 5$, find an abelian surface with multiplication by F such that 5 divides the image of inertia at p in $\mathrm{Aut}(T_2A)$ and hence in $\mathrm{Aut}(T_\lambda A)$ for all λ not dividing p. Thus $[F_\lambda(\zeta_5):F_\lambda]\le2$ for all but finitely many primes λ and so $[F(\zeta_5):F]\le2$. Thus $F=\mathbb{Q}(\sqrt{5})$.

§4. Applications

Theorem 4.1. *Suppose that $\bar\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_5)$ has determinant the cyclotomic character and that $\#\bar\rho(I_3)|10$. Then $\bar\rho$ is modular.*

Proof. After twisting by a character we may replace the third condition by the assumption that $\bar\rho(I_3)$ is unipotent. Then by Theorem 1.2 we can find an elliptic curve $E/\mathbb{Q}$ such that $\bar\rho_{E,5} \cong \rho$ and $\bar\rho_{E,3}$ is surjective onto $GL_2(\mathbb{F}_3)$. Then we see by considering the action of I_3 on the 5-adic Tate module of E that E is semi-stable at 3. Hence by [D], E is modular and so $\bar\rho$ is modular. □

Theorem 4.2. *Suppose that $A/\mathbb{Q}$ is an abelian surface, $\lambda : A \to A^\vee$ is a principal polarisation and $i : \mathcal{O} \to \mathrm{End}(A)$ is an embedding all defined over $\mathbb{Q}$ and such that $\lambda \circ i(a) = i(a)^\vee \circ \lambda$ for all $a \in \mathcal{O}$. Suppose moreover that A has semi-stable reduction at 3 and 5 and that the representation of $G_{\mathbb{Q}(\sqrt{5})}$ on the $\sqrt{5}$-division of A is irreducible. Then A is modular (i.e. a factor of the Jacobian of a modular curve).*

Proof. Let $\bar\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_5)$ denote the representation of $G_{\mathbb{Q}}$ on $A[\sqrt{5}]$. The Weil pairing on $A[\sqrt{5}]$ shows that $\det\bar\rho$ is the cyclotomic character. Also $\bar\rho(I_3)$ has order 1 or 5 because A has semi-stable reduction at 3. Thus by Theorem 4.1 $\bar\rho$ is modular.

Let $\rho : G_{\mathbb{Q}} \to GL_2(\mathcal{O}_{\sqrt{5}})$ denote the representation of $G_{\mathbb{Q}}$ on the $\sqrt{5}$-adic Tate module of A. Then ρ is either flat or Selmer (in the sense of Wiles [W]), $\det\rho$ is the cyclotomic character and the reduction of ρ modulo $\sqrt{5}$ is modular and irreducible when restricted to $G_{\mathbb{Q}(\sqrt{5})}$. By [D], ρ is modular and hence by Faltings isogeny theorem A is modular. □

We remark that we have in fact proved the following more general result, which may be convenient to record here.

Theorem 4.2′. *Let F be a totally real number field of degree d and let μ be a prime of $\mathcal{O}_F$ with norm 5 (resp. 3). Let $A/\mathbb{Q}$ be an abelian d-fold with semi-stable reduction at 3 and 5 (resp. at 3). Suppose that there is an isomorphism $i : F \xrightarrow{\sim} \mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. This gives rise to a representation (well defined up to semi-simplification) $\bar\rho_{A,\mu} : G_{\mathbb{Q}} \to GL_2(\mathcal{O}_F/\mu)$. We further suppose that this representation is absolutely irreducible when restricted to $\mathbb{Q}(\sqrt{5})$ (resp. $\mathbb{Q}(\sqrt{-3})$). Then A is modular.*

We should first explain what we mean by $\bar\rho_{A,\mu}$. Choose an abelian variety A' isogenous to A with an embedding $\mathcal{O}_F \hookrightarrow \mathrm{End}(A')$. Then let $\bar\rho_{A,\mu}$ denote the action of $G_{\mathbb{Q}}$ on $A'[\mu]$. If for one choice of A' this is absolutely irreducible it is for all choices of A' and is independent of the choice of A'.

Thus to prove the proposition we may assume that $R = \mathcal{O}_F$. We may choose a polarisation λ on A and it is known that the Rosati involution it induces on $i(\mathcal{O}_F)$ must be positive and hence the identity. We see that the action of $G_{\mathbb{Q}}$ on $T_\mu A \cong \mathcal{O}_{F,\mu}^2$ has determinant the cyclotomic character and so $\det\bar\rho_{A,\mu}$ is the cyclotomic character. The result now follows exactly as in the previous proposition. We remark that the case μ has norm 3 is due to Wiles et al. (see [W], [TW], [D]).

Theorem 4.3. *Suppose that $\bar\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_4)$ is a representation which is unramified at 3 and 5. Then $\bar\rho$ is modular.*

Proof. Twisting, we may replace $GL_2(\mathbb{F}_4)$ by $SL_2(\mathbb{F}_4)$. By Theorem 3.4 we can find an abelian surface $A/\mathbb{Q}$, a principal polarisation $\lambda : A \to A^\vee$ and an embedding $i : \mathcal{O} \hookrightarrow G_{\mathbb{Q}}$ all defined over $\mathbb{Q}$ and such that

(1) $\lambda \circ i(a) = i(a)^\vee \circ \lambda$ for all $a \in \mathcal{O}$;
(2) $\bar{\rho}$ is equivalent to the representation of $G_{\mathbb{Q}}$ on $A[2]$;
(3) the representation of $G_{\mathbb{Q}(\sqrt{5})}$ on $A[\sqrt{5}]$ is irreducible.

It is not hard to see that the action of the inertia groups I_3 and I_5 on the 2-adic Tate module of A can be conjugated (separately) to lie in the group of matrices of the form

$$\pm \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} q.$$

Thus, possibly after replacing A by a quadratic twist, we may also assume that A is semi-stable at 3 and 5. Thus, by Theorem 4.2, A is modular; and hence $\bar{\rho}$ is modular. □

Acknowledgements

We are grateful to Fedya Bogomolov, for explaining to us how to improve our original argument for Proposition 3.2 so that it would work over any field of characteristic 0, rather than just over $\mathbb{Q}$. The first author is grateful to Columbia University for its hospitality, and especially to Bob Friedman and Troels Jorgenson. Both authors are grateful to John Coates for his support.

References

[A] J.Conway, R.Curtis, S.Norton, R.Parker and R.Wilson, *ATLAS of finite groups*, Oxford, 1985. MR **88g:**20025

[AS] A.Ash and G.Stevens, *Modular forms in characteristic l and special values of their L-functions*, Duke Math. J. **53** (1986), 849-868. MR **88h:**11036

[D] F. Diamond, *On deformation rings and Hecke rings*, Annals of Math. (2) **144** (1996), 137-166. CMP 96:17

[DO] I.Dolgachev and D.Ortland, *Point sets in projective space and theta functions*, Astérisque **165** (1988). MR **90i:**14009

[DR] P.Deligne and M.Rapoport, *Les schémas de modules de courbes elliptiques*, Modular functions of one variable II, LNM 349 (1973). MR **52:**3177

[G2] A.Grothendieck, *Le groupe de Brauer II: théorie cohomologique*, in Dix exposés sur la cohomologie des schémas, Advanced Studies in Pure Mathematics **3** (1968). MR **39:**5586b

[G3] A.Grothendieck, *Le groupe de Brauer III: exemples et complements*, in Dix exposés sur la cohomologie des schémas, Advanced Studies in Pure Mathematics **3** (1968). MR **39:**5586c

[H] C.Hermite, *Sur la résolution de l'équation du cinqième degré*, Comptes Rendus **46** (1858).

[K] F. Klein, *Lectures on the icosahedron (transl. G.G. Morrice)*, Trübner, 1888.

[L] R.Langlands, *Base Change for $GL(2)$*, Princeton, 1980. MR **82a:**10032

[M] D.Mumford, *Tata lectures on theta II*, Birkauser, 1984. MR **86b:**14017

[RS] K.Rubin and A.Silverberg, *Families of elliptic curves with constant mod p representations*, Ser. Number Theory, vol. 1, Internat. Press, Cambridge, MA, 1995, pp. 148-161. MR **96j:**11078

[S1] J-P. Serre, *Extensions icosaédriques*, Oeuvres III (no. 123 (1980)), Springer, 1986. MR **89h:**01109c

[S2] J-P. Serre, *Sur les représentations modulaires de degré* 2 *de* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Duke Math. J. **54** (1987), 179-230. MR **88g:**11022

[T] J.Tunnell, *Artin's conjecture for representations of octahedral type*, Bull. AMS **5** (1981), 173-175. MR **82j:**12015

[TW] R. Taylor and A. Wiles, *Ring theoretic properties of certain Hecke algebras*, Annals of Math. **141(3)** (1995), 553-572. MR **96d:**11072

[vG] G. van der Geer, *Hilbert modular surfaces*, Springer, 1988. MR **89c:**11073
[W] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Annals of Math. **141(3)** (1995), 443-551. MR **96d:**11071

Department of Pure Mathematics and Mathematical Statistics, Cambridge University, 16 Mill Lane, Cambridge CB2 1SB, United Kingdom
E-mail address, N. Shepherd-Barron: nist@pmms.cam.ac.uk

Current address, R. Taylor: Department of Mathematics, Harvard University, 1 Oxford St., Cambridge, Massachusetts 02138
E-mail address, R. Taylor: rtaylor@math.harvard.edu