manuscripta math. 105, 201 – 234 (2001)

David Carlton

Moduli for pairs of elliptic curves with isomorphic N-torsion

Received: 20 September 2000 / Revised version: 7 February 2001

Abstract. We study the moduli surface for pairs of elliptic curves together with an isomorphism between their N-torsion groups. The Weil pairing gives a "determinant" map from this moduli surface to $(\mathbf{Z}/N\mathbf{Z})^*$; its fibers are the components of the surface. We define spaces of modular forms on these components and Hecke correspondences between them, and study how those spaces of modular forms behave as modules for the Hecke algebra. We discover that the component with determinant -1 is somehow the "dominant" one; we characterize the difference between its spaces of modular forms and the spaces of modular forms on the other components using forms with complex multiplication. In addition, we prove Atkin–Lehner-style results about these spaces of modular forms. Finally, we show some simplifications that arise when N is prime, including a complete determination of such CM-forms, and give numerical examples.

1. Introduction

If R is the ring of integers in a totally real number field, one can consider the Hilbert modular variety associated to R, which parameterizes abelian varieties of dimension $[R : \mathbf{Z}]$ together with a map from R into their endomorphism ring. This modular variety is disconnected; its components correspond to polarization types, and are indexed by elements of the narrow class group of R. One can define spaces of modular forms associated to the modular variety and to its components; the former are adelic in nature, while the latter are more classical.

In this paper, we consider a variant of the above situation, where we replace R by the order $(\mathbf{Z} \times \mathbf{Z})_{\equiv(N)}$ that consists of pairs of integers that are congruent mod N. Thus, we replace our totally real number field by the totally real "number algebra" $\mathbf{Q} \times \mathbf{Q}$, and in addition consider a non-maximal order rather than the full ring of integers. As in the traditional situation, one can associate a modular variety to this situation, and study its components, which are indexed by $(\mathbf{Z}/N\mathbf{Z})^*$; this has been done in Hermann [5] and Kani and Schanz [7]. One can also define spaces of classical and adelic modular forms, which we do in this paper.

These "degenerate" Hilbert modular varieties and modular forms should have properties very similar to those of traditional Hilbert modular varieties and modular

D. Carlton: Department of Mathematics, Stanford University, Stanford, CA 94305-2125, USA. e-mail: carlton@math.stanford.edu

Mathematics Subject Classification (2000): 11F41, 11F30, 11F12

forms. However, they can also be related to modular curves and elliptic modular forms, which have been the subject of extensive study. For example, these surfaces have an interpretation as moduli spaces for pairs of elliptic curves with isomorphic N-torsion, and can be constructed as a quotient of $X(N) \times X(N)$. Thus, we expect them to be a particularly suitable test ground for exploring properties of Hilbert modular surfaces and modular forms. We expect the generalization to the case where R is an order in a product of ring of integers of totally real number fields to be of interest as well: for example, R might be the Hecke algebra $\mathbf{T}_0(N)$ associated to the modular curve $X_0(N)$.

One such new property, which is the main goal of this paper, involves studying how these components of the degenerate Hilbert modular variety vary. It is easy to see that two components whose index differs by a square are isomorphic, but there is no reason why other components should be isomorphic. Indeed, Hermann has shown that, for example, if $N = 7$ then the component indexed by 1 is a rational surface and the component indexed by -1 is a K3 surface; similarly, if $N = 11$, the component indexed by 1 is an elliptic surface and the component indexed by -1 is of general type. As Kani and Schanz noted, this change in geometric complexity is reflected by the geometric genera of the components.

These geometric genera can be studied via modular forms. Thus, in Sect. 2, we define spaces of modular forms associated to these surfaces; we also use new techniques to prove an Atkin–Lehner-style result about how these spaces are related under change of level. In Sects. 3 and 4, we define Hecke algebras associated to these spaces; we prove multiplicity one theorems for the action of these Hecke algebras in Sects. 4 and 5. We also show in Sect. 5 that, for N fixed, the component indexed by -1 always has the largest geometric genus of any of the components. The geometric genus of a component is the dimension of a suitable space of cusp forms of weight $(2, 2)$; we exhibit this difference in genera as the dimension of a certain special subspace of the space of cusp forms on the -1 surface, which we call the *Hecke kernel* since it can be seen as the intersection of the kernels of certain Hecke operators. In Sect. 6, we give an alternative characterization of the Hecke kernel as forms with complex multiplication. Finally, we give an explicit formula for the difference of geometric genera when N is prime in Sect. 8, and give an explicit construction of the forms in the Hecke kernel when the weight is $(2, 2)$ and N is prime in Sect. 9.

2. Basic definitions

Let $X_w(N)$ be the curve over $\mathbf{C}$ parameterizing elliptic curves together with a basis for their N-torsion that maps to some specified N-th root of unity under the Weil pairing.[1] It is Galois over the curve $X_w(1)$ with group $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})/\{\pm 1\}$. Let $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$ act on the product surface $X_w(N) \times X_w(N)$ via the diagonal action; we can then form the quotient surface, which we shall denote by $X_{\simeq,1}(N)$. More

[1] This curve is traditionally denoted by $X(N)$; however, we have chosen to use the notation $X(N)$ to denote the (geometrically reducible) curve coming from the adelic mod N principal congruence subgroup, and have changed all notation accordingly.

generally, if ϵ is an element of $(\mathbf{Z}/N\mathbf{Z})^*$ and if $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$ acts on the first factor via the natural action but on the second factor via the automorphism

$$\theta_\epsilon : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & \epsilon^{-1}b \\ \epsilon c & d \end{pmatrix}$$

then we denote the quotient surface by $X_{\simeq,\epsilon}(N)$. And we set

$$X_{\simeq}(N) = \coprod_{\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*} X_{\simeq,\epsilon}(N).$$

These surfaces can also be constructed in another fashion, as degenerate Hilbert modular surfaces: let $\mathfrak{H}$ be the upper half plane, with $\Gamma(1) = \mathrm{SL}_2(\mathbf{Z})$ acting on it via fractional linear transformations. Then $\Gamma(1) \times \Gamma(1)$ acts on $\mathfrak{H} \times \mathfrak{H}$; if we denote by $\Gamma_{\simeq,\epsilon}(N)$ the subgroup of $\Gamma(1) \times \Gamma(1)$ given by

$$\left\{ \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \right) \;\middle|\; \begin{array}{rl} a_1 \equiv a_2 & \pmod{N}, \\ b_1 \equiv \epsilon b_2 & \pmod{N}, \\ \epsilon c_1 \equiv c_2 & \pmod{N}, \\ d_1 \equiv d_2 & \pmod{N} \end{array} \right\}$$

then the quotient $\Gamma_{\simeq,\epsilon}(N)\backslash\mathfrak{H} \times \mathfrak{H}$ is an open subset of $X_{\simeq,\epsilon}(N)$, and if we denote by $\mathfrak{H}^*$ the space $\mathfrak{H} \coprod \mathbf{P}^1(\mathbf{Q})$ then $\Gamma_{\simeq,\epsilon}(N)\backslash\mathfrak{H}^* \times \mathfrak{H}^*$ is all of $X_{\simeq,\epsilon}(N)$.

The surface $X_{\simeq,\epsilon}(N)$ (or, more properly, the open subset given by using $\mathfrak{H} \times \mathfrak{H}$ instead of $\mathfrak{H}^* \times \mathfrak{H}^*$) is a coarse moduli space for isomorphism classes of triples (E_1, E_2, ϕ) where the E_i's are elliptic curves and ϕ is an isomorphism from $E_1[N]$ to $E_2[N]$ such that $\wedge^2\phi$ raises the Weil pairing to the ϵ-th power; cf. Kani and Schanz [7], p. 339. The modular parameterization is given on points as follows: let $(\tau_1, \tau_2) \in \mathfrak{H} \times \mathfrak{H}$ and let E_i be the elliptic curve given by the lattice with basis $\{1, \tau_i\}$. Also, let e be an integer that reduces to ϵ mod N. We then have the map ϕ from $E_1[N]$ to $E_2[N]$ that sends τ_1/N to $e\tau_2/N$ and $1/N$ to $1/N$; it raises the Weil pairing to the ϵ-th power, the group of elements of $\Gamma(1) \times \Gamma(1)$ that preserve ϕ is the subgroup $\Gamma_{\simeq,\epsilon}(N)$ defined above, and every triple (E_1, E_2, ϕ) arises in this fashion.

The structure of the $X_{\simeq,\epsilon}(N)$'s as complex surfaces has been studied by Hermann in [5] and by Kani and Schanz in [7]; our $X_{\simeq,\epsilon}(N)$ is Hermann's $Y_{N,\epsilon^{-1}}$ and Kani and Schanz's $Z_{N,\epsilon^{-1}}$.[2] In particular, Kani and Schanz give explicit formulas and tables computing various invariants of the $X_{\simeq,\epsilon}(N)$'s, such as the dimensions of various cohomology groups. They also give explicit minimal desingularizations of the surfaces.

We now define spaces of modular forms on these surfaces. Thus, let f be a holomorphic function on $\mathfrak{H} \times \mathfrak{H}$; let $\gamma = (\gamma_1, \gamma_2)$ be an element of $\mathrm{GL}_2^+(\mathbf{R}) \times \mathrm{GL}_2^+(\mathbf{R})$, where $\mathrm{GL}_2^+(\mathbf{R})$ is the set of elements of $\mathrm{GL}_2(\mathbf{R})$ with positive determinant; and let $k = (k_1, k_2)$ be a pair of natural numbers. We define the function $f|_{k,\gamma} : \mathfrak{H} \times \mathfrak{H} \to \mathbf{C}$ by

$$f|_{k,\gamma}(z_1, z_2) = f(\gamma_1(z_1), \gamma_2(z_2)) j(\gamma_1, z_1)^{-k_1} j(\gamma_2, z_2)^{-k_2}$$

[2] We replaced their ϵ by ϵ^{-1} to simplify the normalizations given in Sect. 7; since $X_{\simeq,\epsilon}(N)$ and $X_{\simeq,\epsilon^{-1}}(N)$ are isomorphic, this is an unimportant change.

where, if $\sigma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is an element of $\mathrm{GL}_2^+(\mathbf{R})$, then $\sigma(z) = (az+b)/(cz+d)$ and $j(\sigma, z) = (ad-bc)^{-1/2}(cz+d)$. We write $f|_\gamma$ instead of $f|_{k,\gamma}$ if k is clear from context.

Defining $\Gamma(1)$ to be $\mathrm{SL}_2(\mathbf{Z})$, we say that a subgroup Γ of $\Gamma(1) \times \Gamma(1)$ is a *congruence subgroup* if it contains the group $\Gamma_w(N) \times \Gamma_w(N)$ for some N, where $\Gamma_w(N)$ is defined to be the set of matrices in $\mathrm{SL}_2(\mathbf{Z})$ that are congruent to the identity mod N. A function $f : \mathfrak{H} \times \mathfrak{H} \to \mathbf{C}$ is a *modular form for* Γ *of weight* k if $f|_{k,\gamma} = f$ for all $\gamma \in \Gamma$ and if f is holomorphic at the cusps. To explain this latter condition, assume that $\Gamma_w(N) \times \Gamma_w(N) \subset \Gamma$. Then $f(z_1 + N, z_2) = f(z_1, z_2)$ for all $(z_1, z_2) \in \mathfrak{H} \times \mathfrak{H}$; so setting $q_1 = e^{2\pi\sqrt{-1}z_1/N}$, we can write

$$f(z_1, z_2) = \sum_{m \in \mathbf{Z}} c_m(f)(z_2) q_1^m$$

for some functions $c_m(f)$. If $c_m(f)$ is zero for all $m < 0$ and if a similar condition holds if we do a Fourier expansion in z_2, we say that f is *holomorphic at infinity*. And f is *holomorphic at all of the cusps* if, for all $\gamma \in \Gamma(1) \times \Gamma(1)$, $f|_{k,\gamma}$ is holomorphic at infinity.

A modular form is a *cusp form* if it vanishes at all of the cusps; that is to say, if whenever we take a Fourier expansion of $f|_{k,\gamma}$ in either variable as above, $c_0(f)$ is zero. We denote the space of all modular forms of weight k for Γ by $M_k(\Gamma)$; we denote the space of all cusp forms by $S_k(\Gamma)$.

If $\Gamma = \Gamma_1 \times \Gamma_2$, with each Γ_i a congruence subgroup of $\Gamma(1)$, then there is a natural map from $M_{k_1}(\Gamma_1) \otimes M_{k_2}(\Gamma_2)$ to $M_{(k_1,k_2)}(\Gamma_1 \times \Gamma_2)$ which sends $f_1 \otimes f_2$ to the function

$$(z_1, z_2) \mapsto f_1(z_1) f_2(z_2).$$

Furthermore, this map sends cusp forms to cusp forms. It is in fact an isomorphism in either the modular form or cusp form case:

Proposition 2.1. *If S is a subset of $\mathfrak{H}^*$ or $\mathfrak{H}^* \times \mathfrak{H}^*$ and Γ is a congruence subgroup of $\Gamma(1)$ or $\Gamma(1) \times \Gamma(1)$, let $M_k(\Gamma, S)$ be the set of forms in $M_k(\Gamma)$ that vanish on the points in S. Then for any congruence subgroups Γ_1 and Γ_2 of $\Gamma(1)$ and subsets S_1 and S_2 of $\mathfrak{H}^*$, the natural map*

$$M_{k_1}(\Gamma_1, S_1) \otimes M_{k_2}(\Gamma_2, S_2) \to M_{(k_1,k_2)}(\Gamma_1 \times \Gamma_2, (S_1 \times \mathfrak{H}^*) \cup (\mathfrak{H}^* \times S_2))$$

is an isomorphism.

Proof. We prove the Proposition by induction on the dim $M_{k_1}(\Gamma_1, S_1)$. Set $S_{12} = (S_1 \times \mathfrak{H}^*) \cup (\mathfrak{H}^* \times S_2)$, and assume that dim $M_{k_1}(\Gamma_1, S_1)$ is zero. Let f be an element of $M_{(k_1,k_2)}(\Gamma_1 \times \Gamma_2, S_{12})$. For any $z_2 \in \mathfrak{H}^*$, the function $z \mapsto f(z, z_2)$ is an element of $M_{k_1}(\Gamma_1, S_1)$, which is therefore zero, so f is the zero function.

Now assume that dim $M_{k_1}(\Gamma_1, S_1)$ is positive, and let z_1' be an element of $\mathfrak{H}^*$ such that, setting $S_1' = S_1 \cup \{z_1'\}$, dim $M_{k_1}(\Gamma_1, S_1') =$ dim $M_{k_1}(\Gamma_1, S_1) - 1$. Let $S_{12}' = (S_1' \times \mathfrak{H}^*) \cup (\mathfrak{H}^* \times S_2)$; writing M_i for $M_{k_i}(\Gamma_i, S_i)$, M_1' for $M_1(\Gamma_1, S_1')$, and

$M_{1,2}$ and $M'_{1,2}$ for the similar spaces of weight (k_1, k_2) and level $\Gamma_1 \times \Gamma_2$, we will construct a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M'_1 \otimes M_2 & \longrightarrow & M_1 \otimes M_2 & \longrightarrow & M_2 & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \| & & \\ 0 & \longrightarrow & M'_{1,2} & \longrightarrow & M_{1,2} & \longrightarrow & M_2 & \longrightarrow & 0 \end{array}$$

with exact rows. This will prove our Proposition: the left vertical arrow is an isomorphism, by induction, so the middle one is as well.

The left horizontal arrows are the obvious injections. The right arrow on the top row sends $f_1 \otimes f_2$ to $f_1(z'_1) f_2$; the definition of S'_1 and the choice of z'_1 shows that this makes the top row exact.

Similarly, we define the right arrow on the bottom row by having it send f to the function sending z to $f(z'_1, z)$, which is in $M_{k_2}(\Gamma_2, S_2)$. This map is surjective: if we pick a function $f'_1 \in M_{k_1}(\Gamma_1, S_1)$ such that $f'_1(z) = 1$ then we can get a splitting for this map by sending f_2 to the image of $f'_1 \otimes f_2$ under the middle vertical arrow. The exactness of the bottom row then follows immediately from the definitions. □

Corollary 2.1. *Given any natural numbers k_1, k_2, and N, we have isomorphisms*

$$M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N)) \simeq (M_{k_1}(\Gamma_w(N)) \otimes M_{k_2}(\Gamma_w(N)))^{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})}$$

and

$$S_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N)) \simeq (S_{k_1}(\Gamma_w(N)) \otimes S_{k_2}(\Gamma_w(N)))^{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})},$$

where $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$ *acts on the first member of the tensor product in the natural fashion and on the second member via the automorphism θ_ϵ.*

Proof. By Proposition 2.1,

$$M_{(k_1,k_2)}(\Gamma_w(N) \times \Gamma_w(N)) \simeq (M_{k_1}(\Gamma_w(N)) \otimes M_{k_2}(\Gamma_w(N)));$$

that $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$-invariants correspond to forms in $M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N))$ follows from the definitions. The cusp form case is similar, setting S_1 and S_2 in the Proposition to be equal to $\mathbf{P}^1(\mathbf{Q})$. □

This allows us to express the dimension of the space $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(N))$ in terms of data given in Kani and Schanz [7]:

Corollary 2.2. *The dimensions of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(N))$ and $H^2(X_{\simeq,\epsilon}(N), \mathcal{O}_{X_{\simeq,\epsilon}(N)})$ are equal, and they are also equal to the geometric genus of a desingularization of $X_{\simeq,\epsilon}(N)$.*

Proof. Writing X for $X_w(N)$, we have the equalities

$$\begin{aligned}\dim S_{(2,2)}(\Gamma_{\simeq,\epsilon}(N)) &= \dim(S_2(\Gamma_w(N)) \otimes S_2(\Gamma_w(N)))^{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})} \\ &= \dim(H^1(X, \mathcal{O}_X) \otimes H^1(X, \mathcal{O}_X))^{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})} \\ &= \dim H^2(X \times X, \mathcal{O}_{X\times X})^{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})} \\ &= \dim H^2(\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})\backslash(X \times X), \mathcal{O}_{\mathrm{SL}_2\backslash X\times X}).\end{aligned}$$

The last equality follows from Kani and Schanz [6], Proposition 2.7; that the right hand side equals the geometric genus is Kani and Schanz [6], Proposition 3.1. □

Of course, this isn't too surprising: weight 2 cusp forms should correspond to holomorphic 2-forms.

If f is a modular form on $\Gamma_{\simeq,\epsilon}(N)$, it has a Fourier expansion

$$f(z_1, z_2) = \sum_{m_1,m_2\geq 0} c_{m_1,m_2}(f) q_1^{m_1} q_2^{m_2}$$

where $q_i = e^{2\pi\sqrt{-1}z_i/N}$. There is one thing that we can say immediately about the Fourier coefficients $c_{m_1,m_2}(f)$:

Proposition 2.2. *For all* $f \in M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N))$, $c_{m_1,m_2}(f) = 0$ *is zero unless* $\epsilon m_1 + m_2 \equiv 0 \pmod{N}$.

Proof. This follows from the fact that $f = f|_{\left(\left(\begin{smallmatrix}1 & e\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)\right)}$, where e is an integer congruent to ϵ mod N. □

Thus, most of the Fourier coefficients are "missing". This turns out to make it natural to also study modular forms on the surface $X_{\simeq}(N)$, even when we are only interested in one of the individual $X_{\simeq,\epsilon}(N)$'s; we shall elaborate on this theme in Sect. 5.

One way to produce forms on $X_{\simeq,\epsilon}(N)$ is to consider forms on $X_{\simeq,\epsilon}(N/d)$ to be forms on $X_{\simeq,\epsilon}(N)$, for d a divisor of N. Such forms have Fourier coefficients c_{m_1,m_2} equal to zero unless d divides m_1 (and hence m_2, by Proposition 2.2). The converse is also true:

Proposition 2.3. *Let f be a modular form of weight k on $\Gamma_{\simeq,\epsilon}(N)$, and assume that, for some $d|N$, we have $c_{m_1,m_2}(f) = 0$ unless $d|m_1$. Then f is an element of $M_k(\Gamma_{\simeq,\epsilon}(N/d))$.*

Proof. The fact that $c_{m_1,m_2}(f) = 0$ unless $d|m_1$ is equivalent to having f be invariant under $\left(\left(\begin{smallmatrix}1 & N/d\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\right)$. Thus, we have to show that the smallest subgroup Γ containing both $\left(\left(\begin{smallmatrix}1 & N/d\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\right)$ and $\Gamma_{\simeq,\epsilon}(N)$ is $\Gamma_{\simeq,\epsilon}(N/d)$. Furthermore, we can take the quotient by $\Gamma_w(N) \times \Gamma_w(N)$, and thus consider all matrices to be elements of $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$. Letting $G = \{\gamma \in \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}) \mid (\gamma, 1) \in \Gamma\}$, we see that $\Gamma = G\times\{1\}\cdot\Gamma_{\simeq,\epsilon}(N)$ and one checks that Γ is a subgroup if and only if G is normal. Thus, we have to show that the smallest normal subgroup of $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$ containing the matrix $\tau_{N/d} = \left(\begin{smallmatrix} 1 & N/d \\ 0 & 1 \end{smallmatrix}\right)$ is the kernel of the natural map from $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$ to $\mathrm{SL}_2(\mathbf{Z}/(N/d)\mathbf{Z})$. Furthermore, we can assume that d is a prime p, and by the Chinese remainder theorem we can assume that $N = p^l$ for some l.

First, assume that $l = 1$, so we want to show that the smallest normal subgroup G of $\mathrm{SL}_2(\mathbf{Z}/p\mathbf{Z})$ containing $\tau_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is the entire group. Conjugating by $\sigma = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, we see that $\tau_1' = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ is also in G. But $(\tau_1')^2\tau_1^{-1}\tau_1'\tau_1 = \sigma$, and σ and τ_1 generate $\mathrm{SL}_2(\mathbf{Z})$, hence $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$.

Finally, assume that $l > 1$, and that we have a normal subgroup G containing τ_q, where $q = p^{l-1}$. (Note that q^2 is zero in $\mathbf{Z}/p^l\mathbf{Z}$, which greatly simplifies calculations.) We then have to show that G contains all matrices of the form $\left(\begin{smallmatrix} 1+aq & bq \\ cq & 1+dq \end{smallmatrix}\right)$ with determinant 1; this condition on the determinant is equivalent to having a equal to $-d$ in $\mathbf{Z}/p\mathbf{Z}$. But it is easy to produce all such matrices by taking suitable multiples of τ_q, its conjugate by $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, and its conjugate by $\left(\begin{smallmatrix} a & a-1 \\ 1 & 1 \end{smallmatrix}\right)$. □

In fact, the following analogue of Atkin–Lehner theory is true:

Theorem 2.1. *Let f be a modular form on $\Gamma_{\simeq,\epsilon}(N)$ such that*

$$c_{m_1,m_2}(f) = 0 \text{ unless } (m_i, N) > 1. \tag{2.1}$$

Then f can be written as a sum $f = \sum_{p|N} f_p$ where the p's are the prime divisors of N and $f_p \in M_k(\Gamma_{\simeq,\epsilon}(N/p))$. Furthermore, if f is a cusp form then the f_p's can be chosen to be cusp forms.

The proof rests on two elementary linear algebra lemmas:

Lemma 2.1. *Let $V_1, \ldots, V_n$ be vector spaces and, for each i, let f_i be an endomorphism of V_i. Then*

$$\ker(f_1 \otimes \cdots \otimes f_n) = \sum_{i=1}^{n} V_1 \otimes \cdots \otimes (\ker f_i) \otimes \cdots \otimes V_n.$$

Proof. We can easily reduce to the case $n = 2$. If we write $V_i = (\ker f_i) \oplus V_i'$ then $f_i|_{V_i'}$ is an isomorphism onto its image, and

$$V_1 \otimes V_2 = ((\ker f_1) \otimes (\ker f_2)) \oplus ((\ker f_1) \otimes V_2') \oplus (V_1' \otimes (\ker f_2)) \oplus (V_1' \otimes V_2').$$

We see that $f_1 \otimes f_2$ is zero on the first three factors, and is an isomorphism from the fourth factor onto its image; $\ker(f_1 \otimes f_2)$ is therefore the sum of the first three factors, which is what we wanted to show. □

Lemma 2.2. *Let $V_1, \ldots, V_n$ be vector spaces and, for each i, let V_i' and V_i'' be subspaces of V_i. Then*

$$\left(\sum_{i=1}^{n} V_1 \otimes \cdots \otimes V_i' \otimes \cdots \otimes V_n\right) \cap (V_1'' \otimes \cdots \otimes V_n'') = \sum_{i=1}^{n} V_1'' \otimes \cdots \otimes (V_i' \cap V_i'') \otimes \cdots \otimes V_n''.$$

Proof. Again, we can assume that $n = 2$. Write $V_i = V_{i1} \oplus V_{i2} \oplus V_{i3} \oplus V_{i4}$ where $V_{i1} = V_i' \cap V_i''$, $V_i' = V_{i1} \oplus V_{i2}$, and $V_i'' = V_{i1} \oplus V_{i3}$. Then $V_1' \otimes V_2 + V_1 \otimes V_2'$ is the direct sum of those $V_{1j} \otimes V_{2k}$'s where at least one of j or k is in the set $\{1, 2\}$. Also, $V_1'' \otimes V_2''$ is the direct sum of the $V_{1j} \otimes V_{2k}$'s where j and k are both in the set $\{1, 3\}$. Thus, their intersection is $(V_{11} \otimes V_{21}) \oplus (V_{11} \otimes V_{23}) \oplus (V_{13} \otimes V_{21})$, as claimed. □

Proof of Theorem 2.1. Let $M = M_k(\Gamma(N)) \otimes M_k(\Gamma(N))$; it comes with an action of $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$. If $f \in M$ and $d|N$, define $\pi_d(f)$ to be $\sum_{d|m_1,m_2} c_{m_1,m_2}(f) q_1^{m_1} q_2^{m_2}$. Then $\pi_d(f) \in M$: in fact,

$$\pi_d(f) = \frac{1}{d^2} \sum_{b_1,b_2=0}^{d-1} f|_{\left(\left(\begin{smallmatrix} 1 & b_1N/d \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & b_2N/d \\ 0 & 1 \end{smallmatrix}\right)\right)}.$$

The principle of inclusion and exclusion implies that f satisfies (2.1) if and only if $f = \sum_{p|N} \pi_p(f) - \sum_{\substack{p_1,p_2|N \\ p_1<p_2}} \pi_{p_1p_2}(f) + \cdots$. Thus, if V is an irreducible $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$-representation contained in M, it suffices to prove our Theorem for a form in V, since the conditions of our Theorem can be expressed in terms of the action of $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$.

Let $N = \prod_{i=1}^n p_i^{n_i}$ be the prime factorization of N. Then $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}) \simeq \prod_i \mathrm{SL}_2(\mathbf{Z}/p^i\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z}/p^i\mathbf{Z})$, so $V \simeq \bigotimes_i V_i$ where V_i is a representation of $\mathrm{SL}_2(\mathbf{Z}/p^i\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z}/p^i\mathbf{Z})$. Also, π_{p_i} acts as the identity on the V_j for $j \neq i$. So if we define

$$\pi(f) = f - \sum_{p|N} \pi_p(f) + \sum_{\substack{p_1,p_2|N \\ p_1<p_2}} \pi_{p_1p_2}(f) - \cdots$$

then $\pi = (1 - \pi_{p_1}) \otimes \cdots \otimes (1 - \pi_{p_n})$ and $\ker(\pi)$ is the space of forms satisfying (2.1). Thus, Lemma 2.1 implies that

$$\ker(\pi) = \sum_{i=1}^n V_1 \otimes \cdots \otimes (\ker(1 - \pi_{p_i})) \otimes \cdots \otimes V_n.$$

Turning now to the question of a form's being in $M_k(\Gamma_{\simeq,\epsilon}(N))$, that is the case if and only if the form is both in $M_k(\Gamma(N)) \otimes M_k(\Gamma(N))$ and is invariant under the image $G(N)$ of $\Gamma_{\simeq,\epsilon}(N)$ in $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$. Also, $G(N) \simeq \prod_i G(p_i)$. Thus, setting V_i' to be $\ker(1 - \pi_{p_i})$ and V_i'' to be the space of $G(p_i)$-invariant elements of V_i, Lemma 2.2 implies that an element of V is both in $\ker \pi$ and invariant under $G(N)$ if and only if it is in

$$\sum_{i=1}^n V_1'' \otimes \cdots \otimes (V_i' \cap V_i'') \otimes \cdots \otimes V_n''. \tag{2.2}$$

But if $v_i \in V_i$ is in $V_i' \cap V_i''$ then it is invariant both under $G(p_i)$ and under π_{p_i}. Thus, each element of the summand in (2.2) corresponds to forms satisfying the conditions of Proposition 2.3, so that Proposition proves our Theorem in the modular form case. The cusp form case is exactly the same, replacing M by the space of cusp forms. □

We define $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ to be the quotient of $S_k(\Gamma_{\simeq,\epsilon}(N))$ by the subgroup of forms f whose Fourier coefficients satisfy the condition of Theorem 2.1. In the $\Gamma_1(N)$ case, this would have the effect of replacing $S_k(\Gamma_1(N))$ by a space with the same Hecke eigenspaces but where each eigenspace is one-dimensional, generated by the newform in that eigenspace; we shall see in Theorem 5.2 that Hecke eigenspaces in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ are also one-dimensional. We let

$$S_{k,\simeq}(N) = \bigoplus_{\epsilon\in(\mathbf{Z}/N\mathbf{Z})^*} S_k(\Gamma_{\simeq,\epsilon}(N)),$$

and we let

$$\overline{S}_{k,\simeq}(N) = \bigoplus_{\epsilon\in(\mathbf{Z}/N\mathbf{Z})^*} \overline{S}_k(\Gamma_{\simeq,\epsilon}(N)).$$

Note that in the definitions of $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ and $\overline{S}_{k,\simeq}(N)$ it's enough to assume that the Fourier coefficients are zero unless $(m_1, N) > 1$ (or unless $(m_2, N) > 1$), by Proposition 2.2.

Proposition 2.4. *If p is prime then the spaces $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ and $\overline{S}_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ are equal, as are the spaces $S_{(2,2),\simeq}(p)$ and $\overline{S}_{(2,2),\simeq}(p)$.*

Proof. We have to show that if f is an element of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ such that the Fourier coefficients $c_{m_1,m_2}(f)$ are zero unless $p|m_1$ then f is zero. Proposition 2.3 implies that such an f is in fact a form on $\Gamma_{\simeq,\epsilon}(1)$. By Corollary 2.1, f can be considered to be an element of $S_2(\Gamma(1)) \otimes S_2(\Gamma(1))$. But $S_2(\Gamma(1))$ is zero, so f is zero. □

In fact, this holds for all weights (k_1, k_2) where each k_i is less than 12.

Proposition 2.5. *For all $N > 0$,*

$$\dim S_k(\Gamma_{\simeq,\epsilon}(N)) = \sum_{d|N} \dim \overline{S}_k(\Gamma_{\simeq,\epsilon}(d)).$$

Proof. If, for each $d|N$, we choose a map $s_d : \overline{S}_k(\Gamma_{\simeq,\epsilon}(d)) \to S_k(\Gamma_{\simeq,\epsilon}(d))$ splitting the projection from $S_k(\Gamma_{\simeq,\epsilon}(d))$ to $\overline{S}_k(\Gamma_{\simeq,\epsilon}(d))$, and if we let t_d be the natural inclusion of $S_k(\Gamma_{\simeq,\epsilon}(d))$ into $S_k(\Gamma_{\simeq,\epsilon}(N))$, then Theorem 2.1 shows that the map

$$\Big(\sum_{d|N} t_d\Big) \circ \Big(\bigoplus_{d|N} s_d\Big) : \bigoplus_{d|N} \overline{S}_k(\Gamma_{\simeq,\epsilon}(d)) \to S_k(\Gamma_{\simeq,\epsilon}(N))$$

is an isomorphism of vector spaces. □

3. Hecke operators on $X_{\simeq,\epsilon}(N)$

Set

$$\Delta^*_{\simeq,\epsilon}(N) = \left\{ \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \right) \middle| \begin{array}{ll} a_i, b_i, c_i, d_i \in \mathbf{Z}, & \\ a_i d_i - b_i c_i > 0, & \\ (a_i d_i - b_i c_i, N) = 1, & \\ a_1 \equiv a_2 & \pmod N, \\ b_1 \equiv \epsilon b_2 & \pmod N, \\ \epsilon c_1 \equiv c_2 & \pmod N, \\ d_1 \equiv d_2 & \pmod N \end{array} \right\}.$$

This is a semigroup: we prove a more general property at the beginning of Sect. 4. We can partition $\Delta^*_{\simeq,\epsilon}(N)$ into double $\Gamma_{\simeq,\epsilon}(N)$-cosets; each double coset is called a *Hecke operator*. They act on the spaces of modular forms as follows:

Let $\gamma = (\gamma_1, \gamma_2)$ be an element of $\Delta^*_{\simeq,\epsilon}(N)$, and let

$$\Gamma_{\simeq,\epsilon}(N)\gamma\Gamma_{\simeq,\epsilon}(N) = \coprod_j \Gamma_{\simeq,\epsilon}(N)\gamma_j$$

be a decomposition of the double coset generated by γ into left cosets. As we shall see in Proposition 3.1, this decomposition is finite. Then, if f is a form in $M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N))$, we define

$$f|_{(k_1,k_2),\Gamma_{\simeq,\epsilon}(N)\gamma\Gamma_{\simeq,\epsilon}(N)} = \det(\gamma_1)^{(k_1/2)-1}\det(\gamma_2)^{(k_2/2)-1}\sum_j f|_{(k_1,k_2),\gamma_j}.$$

We see as in Shimura [12], Chapter 3, that $f|_{(k_1,k_2),\Gamma_{\simeq,\epsilon}(N)\gamma\Gamma_{\simeq,\epsilon}(N)}$ is an element of the space $M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N))$, that cusp forms are transformed into cusp forms, and that the product of two Hecke operators is a sum of Hecke operators.

Let T_{n_1,n_2} be the operator given by the sum of the double cosets containing elements (γ_1, γ_2) where $\det(\gamma_i) = n_i$. This is zero unless $n_1 \equiv n_2 \pmod N$ and $(n_i, N) = 1$. Left coset representatives for it are given as follows:

Proposition 3.1. *Let (n_1, n_2) be a pair of positive integers that are congruent mod N and that are relatively prime to N. The set of elements of $\Delta^*_{\simeq,\epsilon}(N)$ that have determinant (n_1, n_2) then has the following left coset decomposition:*

$$\coprod_{\substack{a_1,a_2>0 \\ a_i d_i = n_i \\ 0 \le b_i < d_i}} \Gamma_{\simeq,\epsilon}(N)\left(\sigma_{a_1}\begin{pmatrix} a_1 & b_1 N \\ 0 & d_1 \end{pmatrix}, \sigma_{a_2}\begin{pmatrix} a_2 & b_2 N \\ 0 & d_2 \end{pmatrix}\right)$$

where, for $a \in (\mathbf{Z}/N\mathbf{Z})^*$, σ_a is any matrix in $\Gamma(1)$ that is congruent to $\left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & a \end{smallmatrix}\right)$ mod N.

Proof. First, note that the above cosets do indeed occur in T_{n_1,n_2}. Also, it is easy to see that the above cosets are disjoint. Thus, we have to show that the cosets cover all of T_{n_1,n_2}.

Let (δ_1, δ_2) be an element of $\Delta^*_{\simeq,\epsilon}(N)$ whose determinant is (n_1, n_2). By Shimura [12], Proposition 3.36, we can multiply δ_1 on the left by an element of $\Gamma(1)$ to get it into the form $\left(\begin{smallmatrix} a_1 & b_1 \\ 0 & d_1 \end{smallmatrix}\right)$, with $a_1 > 0$, $a_1 d_1 = n_1$, and $0 \le b_1 < d_1$. Subsequently multiplying it on the left by an element of the form $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ will put it into the form $\left(\begin{smallmatrix} a_1 & b_1 N \\ 0 & d_1 \end{smallmatrix}\right)$, but possibly with a different b_1. (We can still force b_1 to be in the range $0 \le b_1 < d_1$, however.) And since σ_{a_1} is an element of $\Gamma(1)$, we have shown that there is an element γ_1 of $\Gamma(1)$ such that $\gamma_1 \delta_1$ is of the form $\sigma_{a_1} \left(\begin{smallmatrix} a_1 & b_1 N \\ 0 & d_1 \end{smallmatrix}\right)$.

We can choose an element γ_2 of $\Gamma(1)$ such that (γ_1, γ_2) is in $\Gamma_{\simeq,\epsilon}(N)$: reduce γ_1 mod N, apply θ_ϵ to it, and lift it back to $\Gamma(1)$. Multiplying (δ_1, δ_2) on the left by (γ_1, γ_2), we can thus assume that δ_1 is of the form $\sigma_{a_1} \left(\begin{smallmatrix} a_1 & b_1 N \\ 0 & d_1 \end{smallmatrix}\right)$. But then the congruence relations force δ_2 to be congruent to the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & n_1 \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & n_2 \end{smallmatrix}\right)$ $\pmod N$.

Now that we have fixed δ_1 to be of the correct form, we still have to force δ_2 to be of the correct form, and we are only allowed to multiply δ_2 on the left by elements of $\Gamma_w(N)$. Thus, we need to find an element γ_2' of $\Gamma_w(N)$ such that $\gamma_2' \delta_2$ is of the form $\sigma_{a_2} \left(\begin{smallmatrix} a_2 & b_2 N \\ 0 & d_2 \end{smallmatrix}\right)$. However, δ_2 is in what Shimura calls Δ' in [12], p. 68, so we can indeed find such a γ_2' by Proposition 3.36 of Shimura [12]. □

The action of the Hecke operators T_{n_1,n_2} descends to the spaces $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$:

Proposition 3.2. *If f is a form in $S_k(\Gamma_{\simeq,\epsilon}(N))$ such that $c_{m_1,m_2}(f) = 0$ unless $(N, m_i) > 1$ then $T_{n_1,n_2} f$ has the same property for all $n_1 \equiv n_2 \pmod N$.*

Proof. This follows from Theorem 2.1. □

Proposition 3.3. *For all $(\delta_1, \delta_2) \in \Delta^*_{\simeq,\epsilon}(N)$, the $\Gamma_{\simeq,\epsilon}(N)$-double cosets generated by (δ_1, δ_2) and $(\delta_1^\iota, \delta_2^\iota)$ are equal, where*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\iota = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Proof. We need to find matrices (γ_1, γ_2) and (γ_1', γ_2') in $\Gamma_{\simeq,\epsilon}(N)$ such that, for $i \in \{1, 2\}$, $\gamma_i \delta_i = \delta_i^\iota \gamma_i'$. Since δ_1 and δ_1^ι have the same elementary divisors, we can choose a γ_1 and γ_1' that give us equality on the first coordinate. Now pick γ_2 and γ_2' such that (γ_1, γ_2) and (γ_1', γ_2') are in $\Gamma_{\simeq,\epsilon}(N)$. Then $\gamma_2 \delta_2 \equiv \delta_2^\iota \gamma_2' \pmod N$. But by Shimura [12], Lemma 3.29(1), we can then change γ_2 and γ_2' by elements of $\Gamma_w(N)$ so that $\gamma_2 \delta_2 = \delta_2^\iota \gamma_2'$, as desired. □

We can define a Petersson inner product on the space of weight (k_1, k_2) cusp forms just as in the one-variable case:

$$\langle f, g \rangle = \int_{\Gamma_{\simeq,\epsilon}(N) \backslash \mathfrak{H} \times \mathfrak{H}} f(z_i) \overline{g(z_i)} y_1^{k_1-2} y_2^{k_2-2} \, dx_1 \, dx_2 \, dy_1 \, dy_2$$

(where $z_i = x_i + \sqrt{-1}y_i$); then just as in Shimura [12], Formula (3.4.5), we see that the Hecke operators $\Gamma_{\simeq,\epsilon}(N)(\delta_1, \delta_2)\Gamma_{\simeq,\epsilon}(N)$ and $\Gamma_{\simeq,\epsilon}(N)(\delta_1^\iota, \delta_2^\iota)\Gamma_{\simeq,\epsilon}(N)$ are adjoint with respect to that inner product. (The argument in [12] depends on having a fundamental domain for $\Gamma_{\simeq,\epsilon}(N)$ in $\mathfrak{H} \times \mathfrak{H}$; we can construct such a fundamental domain as a finite union of fundamental domains for $\mathrm{SL}_2(\mathbf{Z}) \times \mathrm{SL}_2(\mathbf{Z})$, and those are in turn products of fundamental domains for $\mathrm{SL}_2(\mathbf{Z})$ in $\mathfrak{H}$.) Thus:

Corollary 3.1. *The* $\mathbf{Z}$*-algebra generated by the Hecke operators is a commutative algebra; the Hecke operators are self-adjoint with respect to the Petersson inner product on* $S_k(\Gamma_{\simeq,\epsilon}(N))$ *and simultaneously diagonalizable.*

Proof. The self-adjointness follows from Proposition 3.3 by the above discussion; the commutativity follows from Proposition 3.3 and Shimura [12], Proposition 3.8, and the simultaneous diagonalizability follows from the self-adjointness. □

The effect of Hecke operators on Fourier expansions is given as follows:

Proposition 3.4. *Let* f *be an element of* $M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N))$*; if* a *is an element of* $(\mathbf{Z}/N\mathbf{Z})^*$*, let* $f|_{\left(\sigma_a, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)}$ *have the Fourier expansion*

$$f|_{\left(\sigma_a, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)}(z_1, z_2) = \sum_{m_1,m_2\geq 0} c_{a,m_1,m_2} q_1^{m_1} q_2^{m_2}.$$

If we set

$$T_{n_1,n_2} f(z_1, z_2) = \sum_{m_1,m_2\geq 0} d_{m_1,m_2} q_1^{m_1} q_2^{m_2}$$

then the d_{m_1,m_2}*'s are given by*

$$d_{m_1,m_2} = \sum_{\substack{a_1,a_2>0\\ a_i|(m_i,n_i)}} a_1^{k_1-1} a_2^{k_2-1} c_{(a_1/a_2),m_1n_1/a_1^2,m_2n_2/a_2^2}.$$

Proof. The proof is entirely parallel to the proof of the analogous fact in the one-variable case; cf. Shimura [12], (3.5.12). □

Note that the matrices $\left(\sigma_a, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)$ don't normalize $\Gamma_{\simeq,\epsilon}(N)$. This is why we have to introduce the functions $f|_{\left(\sigma_a, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)}$ instead of simply diagonalizing $M_k(\Gamma_{\simeq,\epsilon}(N))$.

Corollary 3.2. *Let* $f \in M_k(\Gamma_{\simeq,\epsilon}(N))$ *be a simultaneous eigenform for all of the Hecke operators. Then if* $\lambda_{m_1,m_2}(f)$ *is the eigenvalue for* T_{m_1,m_2}*, we have*

$$c_{m_1,m_2}(f) = \lambda_{m_1,m_2}(f) c_{1,1}(f).$$

Unfortunately, this Corollary isn't quite as useful as one might hope, since the above coefficients are all zero by Proposition 2.2 unless $\epsilon = -1$! However, in that situation, we do get the following result:

Corollary 3.3. *If f and g are elements of $S_k(\Gamma_{\simeq,-1}(N))$ that are eigenfunctions for all T_{n_1,n_2}'s with the same eigenvalues then their images in $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ differ by a multiplicative constant.*

Proof. By Proposition 2.2 and Corollary 3.2, if we let $c = c_{1,1}(f)/c_{1,1}(g)$ then $c_{m_1,m_2}(f - cg) = 0$ unless $(m_i, N) > 1$. □

This can be restated as follows: let $\overline{\mathbf{T}}_{k,\epsilon}(N)$ be the $\mathbf{C}$-algebra of endomorphisms of $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ generated by the Hecke operators T_{n_1,n_2} for $n_1 \equiv n_2 \pmod N$. Then:

Proposition 3.5. *The space $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ is a free module of rank one over $\overline{\mathbf{T}}_{k,-1}(N)$.*

Proof. By Corollary 3.1, we can find a basis for $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ consisting of simultaneous eigenforms for all of the elements of $\overline{\mathbf{T}}_{k,-1}(N)$. Furthermore, by Corollary 3.3, no two of those eigenforms have the same eigenvalues. This implies our Proposition. □

We define $\mathbf{T}^*_{k,\epsilon}(N)$ to be the $\mathbf{C}$-algebra of endomorphisms of $S_k(\Gamma_{\simeq,\epsilon}(N))$ generated by the Hecke operators T_{n_1,n_2} for $n_1 \equiv n_2 \pmod N$. Proposition 2.4 tells us that the spaces $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ and $\overline{S}_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ are equal; thus, the above Proposition has the following Corollary:

Corollary 3.4. *The space $S_{(2,2)}(\Gamma_{\simeq,-1}(p))$ is a free module of rank one over $\mathbf{T}^*_{(2,2),-1}(p)$.*

With a little bit more care, we can use the above techniques to prove similar facts for $\epsilon = -k^2$ instead of just $\epsilon = -1$. (This isn't too surprising, since $X_{\simeq,-1}(N)$ and $X_{\simeq,-k^2}(N)$ are isomorphic.) They are in fact true for arbitrary ϵ; the proof demands different techniques, and will be given as Theorem 5.2. It does seem that $X_{\simeq,-1}(N)$ is the "dominant" $X_{\simeq,\epsilon}(N)$; see Sects. 5 and 6 for further discussion of this matter.

Finally, we let $\mathbf{T}^*_{\equiv}(N)$ denote the free polynomial algebra over $\mathbf{C}$ with variables T_{n_1,n_2} for every pair n_1,n_2 of positive integers that are relatively prime to N and congruent mod N. This algebra acts on the spaces $S_k(\Gamma_{\simeq,\epsilon}(N))$ and $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ for all k and ϵ; its image in the endomorphism rings of those spaces gives us the algebras $\mathbf{T}^*_{k,\epsilon}(N)$ and $\overline{\mathbf{T}}_{k,\epsilon}(N)$ that we defined above.

4. Hecke operators on $X_{\simeq}(N)$

In analogy with the elliptic curve case, we'd like to think of the Hecke operators T_{n_1,n_2} defined above as arising from correspondences that have the following modular interpretation: let (E_1, E_2, ϕ) be a point of $X_{\simeq,\epsilon}(N)$, and let $\pi_i : E_i \to E_i'$ be maps of elliptic curves of degree n_i, where $(n_i, N) = 1$. Then ϕ induces a map from $E_1'[N]$ to $E_2'[N]$ which is an isomorphism of group schemes; T_{n_1,n_2} should send our point to the sum of all points (E_1', E_2', ϕ) that arise in such a fashion. Why, then, do we impose the restriction that n_1 be congruent to n_2 mod N? The answer is that, if $\pi : E \to E'$ is a map of degree n (with $(n, N) = 1$) then π doesn't preserve the Weil pairing:

$$(\pi x, \pi y) = (x, \pi^\vee \pi y) = (x, [n]y) = (x, y)^n.$$

So if ϕ raises the Weil pairing to the ϵ'th power then, if we push it forward via maps of order n_i as above, the resulting map raises the Weil pairing to the $\epsilon n_2/n_1$ power. This explains why we had to assume that $n_1 \equiv n_2 \pmod N$ for the Hecke operators to act on the surfaces $X_{\simeq,\epsilon}(N)$. However, we should have Hecke operators T_{n_1,n_2} for arbitrary n_i with $(n_i, N) = 1$ which act on the surface $X_{\simeq}(N)$.

The above considerations, when translated into matrices, lead us to the following definition: for any ϵ, ϵ' in $(\mathbf{Z}/N\mathbf{Z})^*$, set

$$\Delta^*_{\simeq,\epsilon,\epsilon'}(N) = \left\{ \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \right) \middle| \begin{array}{ll} a_i, b_i, c_i, d_i \in \mathbf{Z}, & \\ a_i d_i - b_i c_i > 0, & \\ (a_i d_i - b_i c_i, N) = 1, & \\ a_1 \equiv a_2 & \pmod N, \\ b_1 \equiv \epsilon' b_2 & \pmod N, \\ \epsilon c_1 \equiv c_2 & \pmod N, \\ \epsilon d_1 \equiv \epsilon' d_2 & \pmod N \end{array} \right\}.$$

It is obvious from the definitions that $\Delta^*_{\simeq,\epsilon,\epsilon} = \Delta^*_{\simeq,\epsilon}$ and one easily checks that

$$\Delta^*_{\simeq,\epsilon,\epsilon'} \cdot \Delta^*_{\simeq,\epsilon',\epsilon''} \subset \Delta^*_{\simeq,\epsilon,\epsilon''}.$$

These facts imply in particular that $\Delta^*_{\simeq,\epsilon,\epsilon'}$ is invariant under multiplication by $\Gamma_{\simeq,\epsilon}(N)$ on the left and by $\Gamma_{\simeq,\epsilon'}(N)$ on the right; thus, $\Delta^*_{\simeq,\epsilon,\epsilon'}$ can be partitioned into Hecke operators that send forms on $X_{\simeq,\epsilon}(N)$ to forms on $X_{\simeq,\epsilon'}(N)$. For any n_1 and n_2 with $(n_i, N) = 1$ and with $\epsilon n_1 \equiv \epsilon' n_2 \pmod N$, we define the Hecke operator $T^\epsilon_{n_1,n_2}$ to be the sum of the double cosets $\Gamma_{\simeq,\epsilon}(N)(\gamma_1, \gamma_2)\Gamma_{\simeq,\epsilon'}(N)$ occurring in $\Delta^*_{\simeq,\epsilon,\epsilon'}$ for which $\det(\gamma_i) = n_i$. While this depends on ϵ, it has a set of left coset representatives that is independent of ϵ:

Proposition 4.1. *Let n_1 and n_2 be positive integers that are relatively prime to N, and let ϵ and ϵ' be elements of $(\mathbf{Z}/N\mathbf{Z})^*$ such that $\epsilon n_1 \equiv \epsilon' n_2 \pmod N$. Then the set of elements of $\Delta^*_{\simeq,\epsilon,\epsilon'}(N)$ that have determinant (n_1, n_2) has the following left coset decomposition:*

$$\coprod_{\substack{a_1,a_2>0 \\ a_i d_i = n_i \\ 0 \le b_i < d_i}} \Gamma_{\simeq,\epsilon}(N) \left(\sigma_{a_1} \begin{pmatrix} a_1 & b_1 N \\ 0 & d_1 \end{pmatrix}, \sigma_{a_2} \begin{pmatrix} a_2 & b_2 N \\ 0 & d_2 \end{pmatrix} \right)$$

where, for $a \in (\mathbf{Z}/N\mathbf{Z})^*$, σ_a is any matrix that is congruent to $\left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & a \end{smallmatrix}\right)$ mod N. Furthermore, the above left cosets are also disjoint as $\Gamma(1) \times \Gamma(1)$ cosets.

Proof. The proof is the same as the proof of Proposition 3.1. □

Recall that we defined

$$S_{k,\simeq}(N) = \bigoplus_{\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*} S_k(\Gamma_{\simeq,\epsilon}(N))$$

and made a similar definition for $\overline{S}_{k,\simeq}(N)$. Also, if $\mathbf{f}$ is an element of $S_{k,\simeq}(N)$, we write $\mathbf{f}_\epsilon$ for its ϵ-th component. We then define Hecke operators T_{n_1,n_2} mapping $S_{k,\simeq}(N)$ to $S_{k,\simeq}(N)$ by setting

$$(T_{n_1,n_2}\mathbf{f})_\epsilon = T_{n_1,n_2}^{\epsilon n_2/n_1}(\mathbf{f}_{\epsilon n_2/n_1}).$$

Proposition 4.1 shows that that action of $T^\epsilon_{n_1,n_2}$ "looks the same" for all ϵ, so we will write T_{n_1,n_2} in place of $T^\epsilon_{n_1,n_2}$ from now on. The following Proposition shows that the action of these Hecke operators descends to the spaces $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$, and hence allows us to similarly define an action of them on the space $\overline{S}_{k,\simeq}(N)$:

Proposition 4.2. *If f is a form in $S_k(\Gamma_{\simeq,\epsilon}(N))$ such that $c_{m_1,m_2}(f) = 0$ unless $(N, m_i) > 1$ then $T_{n_1,n_2} f$ has the same property for all n_i relatively prime to N.*

Proof. The proof is the same as the proof of Proposition 3.2. □

The action on Fourier expansions is also as expected from Proposition 3.4, with the same proof:

Proposition 4.3. *Let f be an element of $M_{(k_1,k_2)}(\Gamma_{\simeq,\epsilon}(N))$; if a is an element of $(\mathbf{Z}/N\mathbf{Z})^*$, let $f|_{\left(\sigma_a, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)}$ have the Fourier expansion*

$$f|_{\left(\sigma_a, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)}(z_1, z_2) = \sum_{m_1,m_2 \geq 0} c_{a,m_1,m_2} q_1^{m_1} q_2^{m_2}.$$

If we set

$$T_{n_1,n_2} f(z_1, z_2) = \sum_{m_1,m_2 \geq 0} d_{m_1,m_2} q_1^{m_1} q_2^{m_2}$$

then the d_{m_1,m_2}'s are given by

$$d_{m_1,m_2} = \sum_{\substack{a_1,a_2>0 \\ a_i | (m_i,n_i)}} a_1^{k_1-1} a_2^{k_2-1} c_{(a_1/a_2),m_1n_1/a_1^2,m_2n_2/a_2^2}.$$

This Proposition (or Proposition 4.1, which it is a corollary of) allows us to translate theorems about forms on $X_w(N)$ into theorems about forms on $X_\simeq(N)$: if f is a form on some $X_{\simeq,\epsilon}(N)$ and we have a Hecke operator T_{n_1,n_2}, we can consider f to be form on $X_w(N) \times X_w(N)$ and apply $T_{n_1} \times T_{n_2}$ to it there. This gives us a form on $X_w(N) \times X_w(N)$; but by Proposition 4.1, that has the same effect as directly applying the T_{n_1,n_2} that we have defined above to f considered as a form on $X_{\simeq,\epsilon}(N)$, so our resulting form, which is a priori only a form on $X_w(N) \times X_w(N)$, is really a form on $X_{\simeq,\epsilon n_1/n_2}(N)$. Thus, the fact that the Hecke operators

T_n (with $(n, N) = 1$) on $X_w(N)$ commute implies that our Hecke operators T_{n_1,n_2} commute. Similarly, we can define a Petersson inner product on $S_{k,\simeq}(N)$ by taking the orthogonal direct sum of the inner products on the $S_k(\Gamma_{\simeq,\epsilon}(N))$'s; our Hecke operators are then normal with respect to that inner product because the Hecke operators on $X_w(N)$ are.

It is frequently useful to encapsulate this relation between forms on $X_{\simeq}(N)$ and forms on $X_w(N)$ via the map $\overline{\Sigma}\colon \overline{S}_{k,\simeq}(N) \to \overline{S}_{k_1}(\Gamma_w(N)) \otimes \overline{S}_{k_2}(\Gamma_w(N))$ which sends $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ to $\sum_{\epsilon\in(\mathbf{Z}/N\mathbf{Z})^*} \mathbf{f}_\epsilon$. (We are using the identification given by Proposition 2.1 here.) By $\overline{S}_{k_i}(\Gamma_w(N))$ we mean $S_{k_i}(\Gamma_w(N))/V$ where V is the space of forms $f \in S_{k_i}(\Gamma_w(N))$ such that $c_m(f) = 0$ unless $(m, k_i) > 1$; it is a module over the Hecke algebra generated by the operators T_n with $(n, N) = 1$, and its eigenspaces for that algebra are one-dimensional. The following two Propositions then sum up the discussion of the previous paragraph:

Proposition 4.4. *The map from $S_{k,\simeq}(N)$ to $S_{k_1}(\Gamma_w(N)) \otimes S_{k_2}(\Gamma_w(N))$ sending* $\mathbf{f}$ *to $\sum_{\epsilon\in(\mathbf{Z}/N\mathbf{Z})^*} \mathbf{f}_\epsilon$ commutes with the action of Hecke operators. It descends to an injection $\overline{\Sigma}\colon \overline{S}_{k,\simeq}(N) \to \overline{S}_{k_1}(\Gamma_w(N)) \otimes \overline{S}_{k_2}(\Gamma_w(N))$; if $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ then*

$$\mathbf{f}_\epsilon = \sum_{\substack{m_1,m_2>0\\ \epsilon m_1+m_2\equiv 0 \pmod N\\ (m_i,N)=1}} c_{m_1,m_2}(\overline{\Sigma}\mathbf{f}) q_1^{m_1} q_2^{m_2}.$$

Proof. The only parts that remain to be proved are that $\overline{\Sigma}$ is an injection and that $\mathbf{f}_\epsilon$ can be recovered in the given manner. First, we note that, for all m_1, m_2 with $(m_i, N) = 1$, $c_{m_1,m_2}(\overline{\Sigma}\mathbf{f}) = \sum_{\epsilon\in(\mathbf{Z}/N\mathbf{Z})^*} c_{m_1,m_2}(\mathbf{f}_\epsilon)$. But Proposition 2.2 says that $c_{m_1,m_2}(\mathbf{f}_\epsilon) = 0$ unless $\epsilon \equiv -m_2/m_1 \pmod N$; $c_{m_1,m_2}(\overline{\Sigma}\mathbf{f})$ therefore equals $c_{m_1,m_2}(\mathbf{f}_{-m_2/m_1})$. This together with Proposition 2.2 immediately implies our formula for $\mathbf{f}_\epsilon$. And if $\overline{\Sigma}\mathbf{f} = 0$ then this implies that, for all ϵ and for all m_i such that $\epsilon \equiv -m_2/m_1 \pmod N$, $c_{m_1,m_2}(\mathbf{f}_\epsilon)$ is zero. But that implies that $\mathbf{f}_\epsilon = 0$ by using Proposition 2.2 again. □

Proposition 4.5. *The* $\mathbf{Z}$*-algebra generated by the Hecke operators T_{n_1,n_2} acting on $S_{k,\simeq}(N)$ is a commutative algebra; the Hecke operators are normal with respect to the Petersson inner product on $S_{k,\simeq}(N)$ and simultaneously diagonalizable.*

Proof. This follows from the above reduction of these facts to facts about forms on $X_w(N)$ and from Shimura [12], Theorem 3.41. □

Let $\mathbf{f}$ be an element of $S_{k,\simeq}(N)$, and let m_1 and m_2 be integers relatively prime to N. We define $c_{m_1,m_2}(\mathbf{f})$ to be equal to $c_{m_1,m_2}(\mathbf{f}_{-m_2/m_1})$. We also make the same definition for $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$. If we set $f = \sum_{\epsilon\in(\mathbf{Z}/N\mathbf{Z})^*} \mathbf{f}_\epsilon$ then f is a form on $X_w(N) \times X_w(N)$, and $c_{m_1,m_2}(\mathbf{f}) = c_{m_1,m_2}(f)$, by Proposition 2.2, as noted in the proof of Proposition 4.4.

Proposition 4.6. *Let* $\mathbf{f}$ *be an element of $S_{k,\simeq}(N)$; for $a \in (\mathbf{Z}/N\mathbf{Z})^*$, let $\mathbf{f}_a \in S_{k,\simeq}(N)$ be defined by*

$$(\mathbf{f}_a)_\epsilon = \mathbf{f}_{(a^{-2}\epsilon)}|_{\left(\sigma_a,\, \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)}.$$

Then for all n_1, n_2 with $(n_i, N) = 1$ and for all m_1, m_2 with $(m_i, N) = 1$, we have

$$c_{m_1,m_2}(T_{n_1,n_2}\mathbf{f}) = \sum_{\substack{a_1,a_2>0 \\ a_i|(m_i,n_i)}} a_1^{k_1-1} a_2^{k_2-1} c_{m_1n_1/a_1^2, m_2n_2/a_2^2}(\mathbf{f}_{a_1/a_2}).$$

Proof. This is a corollary of Proposition 4.3. □

We define $\mathbf{T}^*(N)$ to be the free polynomial algebra over $\mathbf{C}$ with generators T_{n_1,n_2} for each pair n_1, n_2 of positive integers that are relatively prime to N. We define $\mathbf{T}^*_{k,\simeq}(N)$ to be its image in the endomorphism ring of $S_{k,\simeq}(N)$; we define $\overline{\mathbf{T}}_{k,\simeq}(N)$ to be its image in the endomorphism ring of $\overline{S}_{k,\simeq}(N)$.

Corollary 4.1. *If $\mathbf{f} \in S_{k,\simeq}(N)$ is a simultaneous eigenform for all Hecke operators T_{n_1,n_2} in $\mathbf{T}^*_{k,\simeq}(N)$ with eigenvalues $\lambda_{n_1,n_2}(\mathbf{f})$ then, for all m_1 and m_2 with $(m_i, N) = 1$, we have*

$$c_{m_1,m_2}(\mathbf{f}) = \lambda_{m_1,m_2}(\mathbf{f})c_{1,1}(\mathbf{f}).$$

Thus, if $\mathbf{f}$ is a non-zero element of $\overline{S}_{k,\simeq}(N)$ that is an eigenform for all the T_{n_1,n_2}'s then $c_{1,1}(\mathbf{f})$ is also non-zero; we call such an $\mathbf{f}$ a *normalized eigenform* if $c_{1,1}(\mathbf{f}) = 1$.

Corollary 4.2. *The space $\overline{S}_{k,\simeq}(N)$ is a free module of rank one over the algebra $\overline{\mathbf{T}}_{k,\simeq}(N)$.*

Proof. By Proposition 4.5, we can find a basis for $\overline{S}_{k,\simeq}(N)$ consisting of simultaneous eigenforms for all elements of $\overline{\mathbf{T}}_{k,\simeq}(N)$; the previous Corollary shows that the eigenspaces are one-dimensional, implying this Corollary. □

Corollary 4.3. *The space $S_{(2,2),\simeq}(p)$ is a free module of rank one over the algebra $\mathbf{T}^*_{(2,2),\simeq}(p)$.*

Proof. This follows from Corollary 4.2 and Proposition 2.4. □

There is a special class of operators contained in our Hecke algebras $\mathbf{T}^*_{k,\simeq}(N)$. Given elements ϵ and a of $(\mathbf{Z}/N\mathbf{Z})^*$, we have

$$(1, \sigma_a)^{-1}\Gamma_{\simeq,\epsilon}(N)(1, \sigma_a) = \Gamma_{\simeq,a^{-2}\epsilon}(N).$$

The action of $(1, \sigma_a)$ therefore gives an isomorphism between $S_k(\Gamma_{\simeq,\epsilon}(N))$ and $S_k(\Gamma_{\simeq,a^{-2}\epsilon}(N))$, denoted by $\langle a\rangle$; as with the operators T_{n_1,n_2}, $\langle a\rangle$ extends to the spaces $S_{k,\simeq}(N)$ and $\overline{S}_{k,\simeq}(N)$ via the definition $(\langle a\rangle\mathbf{f})_\epsilon = \langle a\rangle(\mathbf{f}_{a^2\epsilon})$. Furthermore, the action is the same up to a constant if we multiply $(1, \sigma_a)$ by $\left(\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}a&0\\0&a\end{smallmatrix}\right)\right)$; but if we consider it as an operator on $X_w(N) \times X_w(N)$, as in the discussion before Proposition 4.4, then this, up to a constant, is the product of the identity with the Hecke operator $T(a, a)$. By Shimura [12], Theorem 3.24(4), $T(a, a)$ is in the $\mathbf{Q}$-algebra generated by the $T(n)$'s, so $\langle a\rangle$ is in $\mathbf{T}^*_{k,\simeq}(N)$. Thus:

Proposition 4.7. *For all $a \in (\mathbf{Z}/N\mathbf{Z})^*$, the operator $\langle a\rangle$ given by the action of $(1, \sigma_a)$ is an isomorphism from $S_k(\Gamma_{\simeq,\epsilon}(N))$ to $S_k(\Gamma_{\simeq,a^{-2}\epsilon}(N))$; furthermore, it is contained in $\mathbf{T}^*_{k,\simeq}(N)$.*

5. Relationships between the spaces $\overline{S}_{k,\simeq}(N)$, $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$, and $\overline{S}_k(\Gamma_{\simeq,-1}(N))$

When we were trying to prove that the Hecke eigenspaces in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ are one-dimensional, we ran into problems because forms are "missing" Fourier coefficients: in particular, they don't have a (1, 1) Fourier coefficient unless $\epsilon \equiv -1 \pmod{N}$, so we couldn't simply use Corollary 3.2. However, the space $\overline{S}_{k,\simeq}(N)$ doesn't have that problem, and there is a natural projection map from $\overline{S}_{k,\simeq}(N)$ to $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$. This gives us a replacement for the missing Fourier coefficients; it also gives us a framework for seeing how the spaces $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ differ (as $\mathbf{T}^*_{\equiv}(N)$-modules) as ϵ varies.

The key Lemma here is the following:

Lemma 5.1. *The space $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ has a basis consisting of $\overline{\mathbf{T}}_{k,\epsilon}(N)$-eigenforms f that are of the form $\mathbf{f}_\epsilon$ for $\overline{\mathbf{T}}_{k,\simeq}(N)$-eigenforms $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$.*

Proof. If $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ is a $\overline{\mathbf{T}}_{k,\simeq}(N)$-eigenform then it is certainly an eigenform for those Hecke operators T_{n_1,n_2} where $n_1 \equiv n_2 \pmod{N}$; its ϵ-component $\mathbf{f}_\epsilon$ is therefore an eigenform for those operators as well. The Lemma then follows from the fact that $\overline{S}_{k,\simeq}(N)$ has a basis of eigenforms, by Proposition 4.5. □

It is possible for two different $\overline{\mathbf{T}}_{k,\simeq}(N)$-eigenforms in $\overline{S}_{k,\simeq}(N)$ to project to the same $\overline{\mathbf{T}}_{k,\epsilon}(N)$-eigenform in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$; we shall discuss this in Theorem 5.1. Also, some eigenforms in $\overline{S}_{k,\simeq}(N)$ project to zero for some choices of ϵ: see the comments after the proof of the following Proposition and Sect. 6. We shall state a slightly stronger version of this Lemma as Corollary 5.2.

Proposition 5.1. *If $f \in \overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ is a $\overline{\mathbf{T}}_{k,\epsilon}(N)$-eigenform then there is an $\overline{\mathbf{T}}_{k,-1}(N)$-eigenform $g \in \overline{S}_k(\Gamma_{\simeq,-1}(N))$ such that $c_{m_1,m_2}(g) = \lambda_{m_1,m_2}(f)$ for all $m_1 \equiv m_2 \pmod{N}$.*

Proof. By Lemma 5.1, there is an eigenform $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ such that $\lambda_{m_1,m_2}(\mathbf{f}) = \lambda_{m_1,m_2}(f)$ for all $m_1 \equiv m_2 \pmod{N}$. (We might a priori not be able to assume that $\mathbf{f}_\epsilon = f$; however, f is a linear combination of eigenforms projecting from $\overline{S}_{k,\simeq}(N)$, so those eigenforms must have the same eigenvalues as f.) We can assume that $\mathbf{f}$ is normalized. We then set $g = \mathbf{f}_{-1}$; it is a normalized eigenform contained in $\overline{S}_k(\Gamma_{\simeq,-1}(N))$, and $\lambda_{m_1,m_2}(g) = \lambda_{m_1,m_2}(\mathbf{f}) = \lambda_{m_1,m_2}(f)$. But Corollary 3.2 then tells us that $c_{m_1,m_2}(g) = \lambda_{m_1,m_2}(f)$. □

Define $\overline{K}'_{k,\epsilon}(N)$ to be the subspace of $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ generated by eigenforms whose eigenvalues are those of an eigenform in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$; define $\overline{K}_{k,\epsilon}(N)$ to be the subspace of $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ generated by eigenforms which do *not* arise in such a fashion. The Hecke algebra $\overline{\mathbf{T}}_{k,\epsilon}(N)$ is isomorphic to the image of $\overline{\mathbf{T}}_{k,-1}(N)$ in the endomorphism ring of $\overline{K}'_{k,\epsilon}(N)$: both actions are diagonalizable, so the rings are isomorphic iff the same eigenvalues occur, which is the case by the definition of $\overline{K}'_{k,\epsilon}(N)$ and by Proposition 5.1. In fact, the spaces $\overline{K}'_{k,\epsilon}(N)$ and $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ are isomorphic as $\mathbf{T}^*_{\equiv}(N)$-modules, because the eigenspaces in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ are one dimensional; we shall prove this fact later as Theorem 5.2. Thus, the space $\overline{K}_{k,\epsilon}(N)$ measures the difference between $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ and $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$; we shall study this space in Sect. 6.

Since the proof of Proposition 5.1 involved lifting eigenforms in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ to eigenforms in $\overline{S}_{k,\simeq}(N)$, we'd like to see how ambiguous the choice of such a lifting is. The following Theorem answers that question:

Theorem 5.1. *Let* $\mathbf{f}$ *be an eigenform in* $\overline{S}_{k,\simeq}(N)$*, and let* $H \subset (\mathbf{Z}/N\mathbf{Z})^*$ *be the set of* ϵ *such that* $\mathbf{f}_{-\epsilon} \neq 0$*. Then:*

1. H *is a subgroup of* $(\mathbf{Z}/N\mathbf{Z})^*$.
2. H *depends only on* $\mathbf{f}_{-1}$.
3. *Every element of* $(\mathbf{Z}/N\mathbf{Z})^*/H$ *has order one or two.*
4. *If* $\mathbf{g}$ *is another eigenform in* $\overline{S}_{k,\simeq}(N)$ *then* $\mathbf{g}_{-1} = \mathbf{f}_{-1}$ *if and only if there is a character* χ *on* H *such that* $\mathbf{g}_{-\epsilon} = \chi(\epsilon)\mathbf{f}_{-\epsilon}$ *for all* $\epsilon \in H$.

As we shall see in Sect. 6, the ambiguity in the fourth part of this theorem has to do with forms with complex multiplication. First, we prove a lemma that we shall need during the proof of the theorem.

Lemma 5.2. *Let* $\mathbf{f}$ *be an eigenform in* $\overline{S}_{k,\simeq}(N)$ *and* ϵ *an element of* $(\mathbf{Z}/N\mathbf{Z})^*$ *such that* $\mathbf{f}_\epsilon \neq 0$*. For any positive integers* m_1 *and* m_2 *there exist positive integers* n_1 *and* n_2 *such that* $(n_i, m_i) = 1$ *for* $i \in \{1, 2\}$ *and* $c_{n_1,n_2}(\mathbf{f}_\epsilon) \neq 0$.

Proof. By Proposition 4.4, $\overline{\Sigma}\mathbf{f}$ is an eigenform in $\overline{S}_{k_1}(\Gamma_w(N)) \otimes \overline{S}_{k_2}(\Gamma_w(N))$. Since the eigenspaces in $\overline{S}_{k_i}(\Gamma_w(N))$ are one-dimensional, there must exist $f_i \in \overline{S}_{k_i}(\Gamma_w(N))$ such that $\overline{\Sigma}\mathbf{f} = f_1 \otimes f_2$.

For any $\epsilon' \in (\mathbf{Z}/N\mathbf{Z})^*$, set $f_{i,\epsilon'} = \sum_{\substack{n>0 \\ n\equiv\epsilon' \pmod N}} c_n(f_i)q^n$. It is also an element of $\overline{S}_{k_i}(\Gamma_w(N))$. (This follows easily from Shimura [12], Proposition 3.64.) Then $\mathbf{f}_\epsilon = \sum_{\epsilon'\in(\mathbf{Z}/N\mathbf{Z})^*} f_{1,\epsilon'} \otimes f_{2,-\epsilon\epsilon'}$, by Proposition 4.4.

Since $\mathbf{f}_\epsilon \neq 0$, there exists $\epsilon' \in (\mathbf{Z}/N\mathbf{Z})^*$ such that $f_{1,\epsilon'}$ and $f_{2,-\epsilon\epsilon'}$ are both nonzero. By Lang [8], Theorem VIII.3.1, there exist n_i such that $(n_i, Nm_i) = 1$ and that $c_{n_1}(f_{1,\epsilon'})$ and $c_{n_2}(f_{2,-\epsilon\epsilon'})$ are both non-zero. But Proposition 4.4 then implies that $c_{n_1,n_2}(\mathbf{f}_\epsilon) \neq 0$, as desired. □

Proof of Theorem 5.1. We can assume that $\mathbf{f}$ is a normalized eigenform. To show that H is a subgroup, let ϵ_1 and ϵ_2 be elements of H. Thus, there exist $n_{1,i}$ and $n_{2,i}$ (for $i = 1, 2$) such that $c_{n_{1,i},n_{2,i}}(\mathbf{f}_{-\epsilon_i})$ is non-zero; by Lemma 5.2, we can assume that $(n_{1,1}, n_{1,2}) = (n_{2,1}, n_{2,2}) = 1$, and by Proposition 2.2, $\epsilon_i n_{1,i} \equiv n_{2,i} \pmod N$.

By Corollary 4.1, $c_{n_{1,i},n_{2,i}}(\mathbf{f}) = \lambda_{n_{1,i},n_{2,i}}(\mathbf{f})$. But

$$\lambda_{n_{1,1}n_{1,2},n_{2,1}n_{2,2}}(\mathbf{f}) = \lambda_{n_{1,1},n_{2,1}}(\mathbf{f})\lambda_{n_{1,2},n_{2,2}}(\mathbf{f}),$$

by our assumption that $(n_{i,1}, n_{i,2}) = 1$, and is therefore non-zero, as is the corresponding Fourier coefficient of $\mathbf{f}$. This is a Fourier coefficient of $\mathbf{f}_\epsilon$ for

$$\epsilon \equiv -(n_{2,1}n_{2,2}/n_{1,1}n_{1,2}) \equiv -(n_{2,1}/n_{1,1})(n_{2,2}/n_{1,2}) \equiv -\epsilon_1\epsilon_2.$$

Thus, $\epsilon_1\epsilon_2 \in H$, so H is a subgroup of $(\mathbf{Z}/N\mathbf{Z})^*$.

To see that every element of $(\mathbf{Z}/N\mathbf{Z})^*/H$ has order one or two, pick $a \in (\mathbf{Z}/N\mathbf{Z})^*$ and let $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ be an eigenform. Then $(\langle a\rangle\mathbf{f})_{-1} = \langle a\rangle(\mathbf{f}_{-a^2})$. Since $\langle a\rangle$ is an invertible operator contained in $\overline{\mathbf{T}}_{k,\simeq}(N)$, by Proposition 4.7, the fact that $\mathbf{f}_{-1} \neq 0$ implies that $(\langle a\rangle\mathbf{f})_{-1} \neq 0$ as well, so so $\mathbf{f}_{-a^2} \neq 0$ and $a^2 \in H$.

To show that H depends only on $\mathbf{f}_{-1}$, it's enough to prove the last part of the Theorem. We shall prove that if $\mathbf{g}$ is an eigenform such that $\mathbf{g}_{-1} = \mathbf{f}_{-1}$ then there is a character χ on H such that $\mathbf{g}_{-\epsilon} = \chi(\epsilon)\mathbf{f}_{-\epsilon}$; the converse (i.e. that $\mathbf{g}$'s constructed in that fashion are eigenforms) follows easily from the definitions.

Thus, assume that we have normalized eigenforms $\mathbf{f}$ and $\mathbf{g}$ such that $\mathbf{f}_{-1} = \mathbf{g}_{-1}$; let ϵ be an element of H, so $\mathbf{f}_{-\epsilon} \neq 0$. Since H is a subgroup, $\mathbf{f}_{-(1/\epsilon)}$ is also non-zero. There then exist m_1 and m_2 relatively prime to N such that $m_1 \equiv \epsilon m_2 \pmod N$ and $c_{m_1,m_2}(\mathbf{f}) \neq 0$. Therefore, $\lambda_{m_1,m_2}(\mathbf{f})$ is also non-zero. And

$$\begin{aligned}\lambda_{m_1,m_2}(\mathbf{f})\mathbf{f}_{-\epsilon} &= (T_{m_1,m_2}\mathbf{f})_{-\epsilon} = T_{m_1,m_2}(\mathbf{f}_{-\epsilon m_2/m_1}) = T_{m_1,m_2}(\mathbf{f}_{-1})\\ &= T_{m_1,m_2}(\mathbf{g}_{-1}) = \lambda_{m_1,m_2}(\mathbf{g})\mathbf{g}_{-\epsilon}.\end{aligned}$$

Since $\lambda_{m_1,m_2}(\mathbf{f})$ and $\mathbf{f}_{-\epsilon}$ are both non-zero, this implies that $\lambda_{m_1,m_2}(\mathbf{g})$ and $\mathbf{g}_{-\epsilon}$ are also both non-zero, and that if we define $\chi(\epsilon) = \lambda_{m_1,m_2}(\mathbf{f})/\lambda_{m_1,m_2}(\mathbf{g})$ (for any choice of m_i such that $m_1 \equiv \epsilon m_2 \pmod N$ and such that $c_{m_1,m_2}(\mathbf{f}_{-1/\epsilon}) \neq 0$) then $\mathbf{g}_{-\epsilon} = \chi(\epsilon)\mathbf{f}_{-\epsilon}$, as desired. We then only have to show that χ is a character, not just a function; that follows by using the same arguments that we used to show that H was a subgroup. □

We now have all the tools necessary to prove that the spaces $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ are free of rank one over $\overline{\mathbf{T}}_{k,\epsilon}(N)$ for all $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$.

Theorem 5.2. *For all $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$, every $\overline{\mathbf{T}}_{k,\epsilon}(N)$-eigenspace in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ is one-dimensional, and the space $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ is a free module of rank one over $\overline{\mathbf{T}}_{k,\epsilon}(N)$.*

Proof. Pick a $\overline{\mathbf{T}}_{k,\epsilon}(N)$-eigenspace in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$. By Lemma 5.1, it has a basis consisting of eigenforms of the form $\mathbf{f}_\epsilon$ where $\mathbf{f}$ is a normalized eigenform in $\overline{S}_{k,\simeq}(N)$. Thus, we need to show that if $\mathbf{f}$ and $\mathbf{g}$ are normalized eigenforms in $\overline{S}_{k,\simeq}(N)$ such that $\mathbf{f}_\epsilon$ and $\mathbf{g}_\epsilon$ are in the same eigenspace then $\mathbf{f}_\epsilon$ and $\mathbf{g}_\epsilon$ are in fact constant multiples of each other. However, $\lambda_{n_1,n_2}(\mathbf{f}_\epsilon) = \lambda_{n_1,n_2}(\mathbf{f}) = c_{n_1,n_2}(\mathbf{f})$, for all $n_1 \equiv n_2 \pmod N$, so the fact that $\mathbf{f}_\epsilon$ and $\mathbf{g}_\epsilon$ have the same eigenvalues simply means that $\mathbf{f}_{-1}$ and $\mathbf{g}_{-1}$ are equal. Theorem 5.1 then implies that $\mathbf{f}_\epsilon$ and $\mathbf{g}_\epsilon$ are multiples of each other. Thus, the eigenspaces are one-dimensional, and $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ is indeed a free $\overline{\mathbf{T}}_{k,\epsilon}(N)$-module of rank one. □

The basic idea behind the proof of Theorem 5.2 is that, if we have a form in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$, we can use Lemma 5.1 to fill in the Fourier coefficients that are forced to vanish by Proposition 2.2. Of course, it's often easiest just to work with $\overline{S}_{k,\simeq}(N)$ and $X_{\simeq}(N)$ directly. As usual, we have the following Corollary:

Corollary 5.1. *For all $\epsilon \in (\mathbf{Z}/p\mathbf{Z})^*$, the space $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ is a free module of rank one over $\mathbf{T}^*_{(2,2),\epsilon}(p)$.*

Proof. This follows from Theorem 5.2 and Proposition 2.4. □

We also have the following slight strengthening of Lemma 5.1:

Corollary 5.2. *For every eigenform* $f \in \overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ *there exists an eigenform* $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ *such that* $\mathbf{f}_\epsilon = f$.

Proof. By Lemma 5.1, $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ has a basis consisting of such eigenforms. Since the eigenspaces are one-dimensional, however, every eigenform must be a multiple of one of those basis elements. □

And, finally, we have the facts that $\overline{K}'_{k,\epsilon}(N)$ and $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ are isomorphic as $\mathbf{T}^*_{\equiv}(N)$-modules and a geometric consequence of that fact:

Corollary 5.3. *For all* $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$, $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ *is isomorphic to* $\overline{K}_{k,\epsilon}(N) \oplus \overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ *as a module over* $\mathbf{T}^*_{\equiv}(N)$.

Proof. By definition, $\overline{S}_k(\Gamma_{\simeq,-1}(N)) = \overline{K}_{k,\epsilon}(N) \oplus \overline{K}'_{k,\epsilon}(N)$. But $\overline{K}'_{k,\epsilon}(N)$ is a $\mathbf{T}^*_{\equiv}(N)$-module that is a direct sum of one-dimensional spaces corresponding to the Hecke eigenvalues occurring in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$; the Corollary then follows from Theorem 5.2. □

Corollary 5.4. *For all* $N > 0$*, the geometric genus of (a desingularization of)* $X_{\simeq,\epsilon}(N)$ *is maximized when* $\epsilon = -1$.

Proof. Corollary 2.2 and Proposition 2.5 allow us to reduce this Corollary to showing that, for all ϵ and for all $M|N$, the dimension of $\overline{S}_{(2,2)}(\Gamma_{\simeq,-1}(M))$ is at least as large as the dimension of $\overline{S}_{(2,2)}(\Gamma_{\simeq,\epsilon}(M))$. This in turn follows directly from the above Corollary. □

6. The Hecke kernel

In Section 5, we saw that, for all $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$, we can write $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ as $\overline{K}_{k,\epsilon}(N) \oplus \overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$. Thus, the key to understanding modular forms in all of the $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$'s is to understand the space $\overline{S}_k(\Gamma_{\simeq,-1}(N))$; once we have that, we then need to understand its subspaces $\overline{K}_{k,\epsilon}(N)$. The goal of the present section is to study those subspaces, which we call "Hecke kernels". Note that Corollary 5.4 gives us a geometric interpretation of these spaces in some situations.

We first give the alternate following characterizations of forms in $\overline{K}_{k,\epsilon}(N)$:

Proposition 6.1. *Let* f *be a nonzero eigenform in* $\overline{S}_k(\Gamma_{\simeq,-1}(N))$ *and let* ϵ *be an element of* $(\mathbf{Z}/N\mathbf{Z})^*$*. The following are equivalent:*

1. f *is in* $\overline{K}_{k,\epsilon}(N)$.
2. *For any or all eigenforms* $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ *such that* $\mathbf{f}_{-1} = f$, $\mathbf{f}_\epsilon = 0$.
3. *For all* n_1, n_2 *such that* $\epsilon n_1 + n_2 \equiv 0 \pmod N$, $T_{n_1,n_2} f = 0$.
4. *For all* m_1, m_2, n_1, *and* n_2 *with* $n_1 m_1 \equiv n_2 m_2 \pmod N$, $\epsilon n_1 + n_2 \equiv 0 \pmod N$*, and* $(n_i, m_i) = 1$ *for* $i \in \{1, 2\}$*, we have* $c_{n_1 m_1, n_2 m_2}(f) = 0$.

Proof. We can assume f is a normalized eigenform. First we, show the equivalence between 1 and 2: let $\mathbf{f}$ be an eigenform in $S_{k,\simeq}(N)$ such that $\mathbf{f}_{-1} = f$, which we can find by Corollary 5.2. By Theorem 5.1, $\mathbf{f}_\epsilon$ only depends on the choice of $\mathbf{f}$ up to a non-zero constant multiple. If $\mathbf{f}_\epsilon \neq 0$ then $\mathbf{f}_\epsilon$ is an eigenform in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ whose eigenvalues are the same as those of f, hence are the same as the Fourier coefficients of f, so f isn't in $\overline{K}_{k,\epsilon}(N)$. Conversely, if f isn't in $\overline{K}_{k,\epsilon}(N)$ then there exists an eigenform $g \in \overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ whose eigenvalues are the Fourier coefficients of f. Corollary 5.2 allows us to pick an eigenform $\mathbf{g} \in \overline{S}_{k,\simeq}(N)$ such that $\mathbf{g}_\epsilon = g$; multiplying it (and g) by a constant factor, we can assume that $\mathbf{g}$ is a normalized eigenform. Then $\mathbf{g}_\epsilon$ and $\mathbf{g}_{-1}$ have the same eigenvalues, so $\mathbf{g}_{-1}$ is a multiple of f, by our assumption on g; $\mathbf{g}$ therefore gives us an eigenform in $\overline{S}_{k,\simeq}(N)$ such that $\mathbf{g}_{-1} = f$ and $\mathbf{g}_\epsilon \neq 0$, as desired. By Theorem 5.1, this is independent of the choice of $\mathbf{g}$, justifying our use of the phrase "any or all".

Next we show that 2 and 3 are equivalent. Thus, we are given normalized eigenforms $f \in \overline{S}_k(\Gamma_{\simeq,-1}(N))$ and $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ such that $f = \mathbf{f}_{-1}$ and we want to show that $\mathbf{f}_\epsilon = 0$ iff, for all n_1 and n_2 such that $\epsilon n_1 + n_2 \equiv 0 \pmod N$, $T_{n_1,n_2} f = 0$. First assume that $\mathbf{f}_\epsilon = 0$. By part 1 of Theorem 5.1, $\mathbf{f}_{1/\epsilon} = 0$. Then for all n_i as above,

$$\begin{aligned} T_{n_1,n_2} f &= T_{n_1,n_2}(\mathbf{f}_{-1}) = (T_{n_1,n_2}\mathbf{f})_{-n_1/n_2} = (T_{n_1,n_2}\mathbf{f})_{1/\epsilon} \\ &= \lambda_{n_1,n_2}(\mathbf{f})\mathbf{f}_{1/\epsilon} = 0. \end{aligned}$$

Conversely, if $T_{n_1,n_2} f = 0$ for all n_i with $\epsilon n_1 + n_2 \equiv 0 \pmod N$ then the above series of equalities shows that $\lambda_{n_1,n_2}(\mathbf{f})\mathbf{f}_{1/\epsilon}$ is always zero, or equivalently (by Corollary 4.1), $c_{n_1,n_2}(\mathbf{f})\mathbf{f}_{1/\epsilon} = 0$. If $\mathbf{f}_\epsilon \neq 0$ then there exist such n_i such that $c_{n_1,n_2}(\mathbf{f}) \neq 0$; thus, $\mathbf{f}_{1/\epsilon} = 0$, so $\mathbf{f}_\epsilon$ is zero after all, by part 1 of Theorem 5.1.

Next we show that 3 implies 4. Assume that, for all n_1 and n_2 with $\epsilon n_1 + n_2 \equiv 0 \pmod N$, $T_{n_1,n_2} f = 0$. Then, for all m_1 and m_2 with $(m_i, n_i) = 1$, we have $T_{m_1 n_1, m_2 n_2}(f) = T_{m_1,m_2}(T_{n_1,n_2}(f))$
$= 0$, so in particular that is true for m_i with $(m_i, n_i) = 1$ and with $m_1 n_1 \equiv m_2 n_2 \pmod N$. But Corollary 3.2 then implies that $c_{m_1 n_1, m_2 n_2}(f) = 0$.

Finally, we show that 4 implies 2, so let f be a normalized eigenform such that all such coefficients $c_{m_1 n_1, m_2 n_2}(f)$ are zero, and let $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ be a lift of f. Assume that $\mathbf{f}_\epsilon \neq 0$. Thus, there exist n_1 and n_2 with $c_{n_1,n_2}(\mathbf{f}) \neq 0$, or, equivalently, $\lambda_{n_1,n_2}(\mathbf{f}) \neq 0$. Then for all m_1 and m_2 with $(m_i, n_i) = 1$ and with $m_1 n_1 \equiv m_2 n_2 \pmod N$, or equivalently $(1/\epsilon)m_1 + m_2 \equiv 0 \pmod N$,

$$0 = \lambda_{m_1 n_1, m_2 n_2}(\mathbf{f}) = \lambda_{m_1,m_2}(\mathbf{f})\lambda_{n_1,n_2}(\mathbf{f}),$$

so $\lambda_{m_1,m_2}(\mathbf{f}) = 0$ for all m_i with $(m_i, n_i) = 1$ and $(1/\epsilon)m_1 + m_2 \equiv 0 \pmod N$. By Lemma 5.2, $\mathbf{f}_{1/\epsilon} = 0$; by part 1 of Theorem 5.1, $\mathbf{f}_\epsilon = 0$, a contradiction. Thus 4 implies 2. □

For an arbitrary form in $\overline{K}_{k,\epsilon}(N)$, it is necessary for those coefficients specified in part 4 of Proposition 6.1 to vanish. The following Proposition shows that even more coefficients of elements of $\overline{K}_{k,\epsilon}(N)$ vanish:

Proposition 6.2. *For all a and ϵ in $(\mathbf{Z}/N\mathbf{Z})^*$, the spaces $\overline{K}_{k,\epsilon}(N)$ and $\overline{K}_{k,a^2\epsilon}(N)$ are equal.*

Proof. Let f be an eigenform in $\overline{K}_{k,\epsilon}(N)$, and let $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ be an eigenform such that $\mathbf{f}_{-1} = f$. Let H be the set of ϵ such that $\mathbf{f}_{-\epsilon} \neq 0$; by assumption, $-\epsilon \notin H$. By the third part of Theorem 5.1, $a^2 \in H$. Since H is a group (by the first part of Theorem 5.1), $-a^2\epsilon \notin H$, so $\mathbf{f}_{-a^2\epsilon} \neq 0$ and $f \in \overline{K}_{k,a^2\epsilon}(N)$. □

Thus, if $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ is a normalized eigenform such that $\mathbf{f}_\epsilon$ is zero for some ϵ, or equivalently that $\mathbf{f}_{-1}$ is in $\overline{K}_{k,\epsilon}(N)$, then $\mathbf{f}_{a^2\epsilon}$ is also zero for all $a \in (\mathbf{Z}/N\mathbf{Z})^*$. So if we let $f = \overline{\Sigma}\mathbf{f}$ then lots of the Fourier coefficients of f are zero. This leads one to suspect that f might be related to forms with complex multiplication, where we define an eigenform g on $X_w(N)$ to have *complex multiplication* if there exists a non-trivial character ϕ such that $\phi(p)\lambda_p(g) = \lambda_p(g)$ (or, equivalently, $\lambda_p(g) = 0$ unless $\phi(p) = 1$) for all but finitely many primes p, where $\lambda_p(g)$ is the T_p-eigenvalue for g. (This is as in Ribet [10], §3, except that we don't require g to be a newform.) This implies that, for some M, $\phi(n)\lambda_n(g) = \lambda(n)(g)$ for all n with $(M, n) = 1$. We also call such a g a *CM-form*. It is indeed the case that such forms are linked to elements of the Hecke kernel:

Theorem 6.1. *An eigenform f is in $\overline{K}_{(k_1,k_2),\epsilon}(N)$ if and only if there exist eigenforms $f_i \in S_{k_i}(\Gamma_w(N))$ such that, for all $n_1 \equiv n_2 \pmod N$ with $(n_i, N) = 1$,*

$$c_{n_1,n_2}(f) = c_{n_1}(f_1)c_{n_2}(f_2)$$

and such that the f_i have complex multiplication by some character ϕ such that $\phi(-\epsilon) = -1$. Furthermore, $\overline{K}_{(k_1,k_2),\epsilon}(N)$ is spanned by such forms.

Proof. Let $k = (k_1, k_2)$, and let $f \in \overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$ be an eigenform. Pick an eigenform $\mathbf{f} \in \overline{S}_{k,\simeq}(N)$ such that $\mathbf{f}_{-1} = f$ and let H be the subgroup of $\epsilon' \in (\mathbf{Z}/N\mathbf{Z})^*$ such that $\mathbf{f}_{-\epsilon'} \neq 0$, as in Theorem 5.1. By Proposition 4.4, $\overline{\Sigma}\mathbf{f}$ is an eigenform in $\overline{S}_{k_1}(\Gamma_w(N)) \otimes \overline{S}_{k_2}(\Gamma_w(N))$; but eigenspaces in that latter space are one-dimensional, so $\overline{\Sigma}\mathbf{f} = f_1 \otimes f_2$, where $f_i \in \overline{S}_{k_1}(\Gamma_w(N))$ is an eigenform. We wish to relate f's being an element of $\overline{K}_{k,\epsilon}(N)$, i.e. having $\mathbf{f}_\epsilon = 0$, to the f_i's being CM-forms.

For all m_1 and m_2 with $(m_i, N) = 1$, $c_{m_1,m_2}(\mathbf{f}) = c_{m_1}(f_1)c_{m_2}(f_2)$. If $\epsilon' \notin H$, i.e. $\mathbf{f}_{-\epsilon'} = 0$, then, for all m_i such that $\epsilon' m_1 \equiv m_2 \pmod N$, $c_{m_1,m_2}(\mathbf{f}) = 0$, so $c_{m_1}(f_1) = 0$ or $c_{m_2}(f_2) = 0$. Since the f_i are eigenforms, their first Fourier coefficients are non-zero; thus, setting $m_2 = 1$, $c_{m_1}(f_1) = 0$ for $m_1 \equiv 1/\epsilon' \pmod N$ where $\epsilon' \notin H$. Since H is a subgroup, this means that $c_{m_1}(f_1) = 0$ for $m_1 \notin H$ (identifying m_1 with its projection to an element of $(\mathbf{Z}/N\mathbf{Z})^*$). Similarly, $c_{m_2}(f_2) = 0$ for $m_2 \notin H$.

First, assume that $f \in \overline{K}_{k,\epsilon}(N)$, i.e. that $\mathbf{f}_\epsilon = 0$, or that $-\epsilon \notin H$. Pick a non-trivial character ϕ of $(\mathbf{Z}/N\mathbf{Z})^*$ that is trivial on H and such that $\phi(-\epsilon) \neq 1$. The previous paragraph shows that f_1 and f_2 both have complex multiplication by ϕ. By part 3 of Theorem 5.1, ϕ has order two; thus, $\phi(-\epsilon) = -1$, as desired.

Conversely, assume that there exists a character ϕ such that the forms f_i have complex multiplication by ϕ and such that $\phi(-\epsilon) = -1$. Pick m_1 and m_2 such that $\epsilon m_1 + m_2 \equiv 0 \pmod N$. Then $-\epsilon \equiv m_2/m_1 \pmod N$; since $\phi(-\epsilon) = -1$, either $\phi(m_1)$ or $\phi(m_2)$ is not equal to one. Thus, either $c_{m_1}(f_1)$ or $c_{m_2}(f_2)$ is zero, so $c_{m_1,m_2}(\mathbf{f}) = 0$. This is true for all such m_i, so $\mathbf{f}_\epsilon = 0$, i.e. $f \in \overline{K}_{k,\epsilon}(N)$.

Finally, the fact that $\overline{K}_{k,\epsilon}(N)$ is spanned by such forms follows from the fact that it has a basis of eigenforms, which is obvious from the definition of $\overline{K}_{k,\epsilon}(N)$. □

For p prime we define $K_{\simeq}(p)$ to be the subspace $\overline{K}_{(2,2),\epsilon}(p)$ of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ for any $\epsilon \in (\mathbf{Z}/p\mathbf{Z})^*$ such that $-\epsilon$ is non-square; here we identify $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ with $\overline{S}_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ by Proposition 2.4. (For this to make sense, we should assume that $p \neq 2$; since $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(2))$ is zero for all ϵ, this isn't very important.) This is independent of the choice of ϵ by Proposition 6.2; its dimension is the difference between the geometric genera of $X_{\simeq,-1}(p)$ and $X_{\simeq,\epsilon}(p)$, by Corollary 5.4. We shall give an explicit basis for this space in Sects. 8 and 9.

7. The adelic point of view

As we have seen in Sect. 4, to get a satisfactory theory of Hecke operators, we had to consider the surface $X_{\simeq}(N)$, not just the surfaces $X_{\simeq,\epsilon}(N)$. To explain this, it helps to look at $X_{\simeq}(N)$ from the adelic point of view. Thus, we review some of definitions from that theory and explain their relevance to our context. For references, see Diamond and Im [1], Sect. 11.

Let $\mathbf{A}^\infty$ denote the finite adeles, i.e. the restricted direct product of the fields $\mathbf{Q}_p$ with respect to the rings $\mathbf{Z}_p$. Let U be an open compact subgroup of $\mathrm{GL}_2(\mathbf{A}^\infty)$. We define the curve Y_U to be $\mathrm{GL}_2^+(\mathbf{Q})\backslash(\mathfrak{H} \times \mathrm{GL}_2(\mathbf{A}^\infty))/U$. Here, $\mathrm{GL}_2^+(\mathbf{Q})$ is the set of matrices in $\mathrm{GL}_2(\mathbf{Q})$ with positive determinant, acting on $\mathfrak{H}$ via fractional linear translations and on $\mathrm{GL}_2(\mathbf{A}^\infty)$ via the injection $\mathbf{Q} \hookrightarrow \mathbf{A}$; U acts trivially on $\mathfrak{H}$ and acts on $\mathrm{GL}_2(\mathbf{A}^\infty)$ via multiplication on the right. This defines Y_U as a non-compact curve over the complex numbers; it has a canonical compactification X_U given by adding a finite number of cusps. The curves X_U and Y_U in fact have canonical models over $\mathbf{Q}$ which are irreducible; over $\mathbf{C}$, however, the number of their components is given by the index of $\det U$ in $\hat{\mathbf{Z}}^\times$. If U and U' are open compact subgroups of $\mathrm{GL}_2(\mathbf{A}^\infty)$ and if g is an element of $\mathrm{GL}_2(\mathbf{A}^\infty)$ such that $g^{-1}Ug \subset U'$ then multiplication by g on the right gives a map $g^*\colon X_U \to X'_U$; it descends to the models over $\mathbf{Q}$.

We say that a function $\mathbf{f}\colon \mathfrak{H} \times \mathrm{GL}_2(\mathbf{A}^\infty) \to \mathbf{C}$ is a *cusp form of weight k on X_U* if

1. $\mathbf{f}(z, g)$ is a holomorphic function in z for fixed g.
2. $\mathbf{f}(\gamma z, \gamma g) = j(\gamma, z)^k \mathbf{f}(z, g)$ for all $\gamma \in \mathrm{GL}_2^+(\mathbf{Q})$.
3. $\mathbf{f}(z, gu) = \mathbf{f}(z, g)$ for all $u \in U$.
4. $\mathbf{f}(z, g)$, considered as a function in z, vanishes at infinity for all g.

We denote by $S_k(U)$ the space of all such forms. If $g^{-1}Ug \subset U'$ then we get a map $g_*\colon S_k(U') \to S_k(U)$ by defining $(g_*\mathbf{f})(z, h)$ to be $\mathbf{f}(z, hg)$.

Each U-double coset in $\mathrm{GL}_2(\mathbf{A}^\infty)$ gives a Hecke operator, which acts on $S_k(U)$. If $U = \mathrm{GL}_2(\mathbf{Z}_p) \times U^p$ then the Hecke operator T_p is generated by the set of inverses of those elements of $M_2(\mathbf{Z}_p)$ whose determinant is in $p\mathbf{Z}_p^\times$; defining the Hecke operator S_p to be the double coset generated by $\begin{pmatrix} p^{-1} & 0 \\ 0 & p^{-1} \end{pmatrix}$ in the $GL_2(\mathbf{Q}_p)$ component, the ring of Hecke operators consisting of those double cosets generated by elements in $GL_2(\mathbf{Q}_p)$ is generated by T_p and $S_p^{\pm 1}$.

If we define $S_k(\mathbf{C})$ to be the direct limit of the $S_k(U)$'s as U gets arbitrarily small then the above maps g_* make this into an admissible representation of $GL_2(\mathbf{A}^\infty)$; the original spaces $S_k(U)$ can be recovered from that representation by taking its U-invariants. The main fact that we need is the following adelic analogue of parts of Atkin-Lehner theory:

Theorem 7.1 (Strong Multiplicity One). *If π and π' are two irreducible constituents of $S_k(\mathbf{C})$ such that π_p and π'_p are isomorphic for almost all p then π and π' are equal. (Not just isomorphic.) Furthermore, if $\mathbf{f}$ and $\mathbf{f}'$ are elements of π and π' then this is the case iff $\mathbf{f}$ and $\mathbf{f}'$ have the same eigenvalues for almost all T_p and S_p; in this case, they have the same eigenvalues for all p such that $\mathbf{f} \in S_k(U)$ for some U of the form $GL_2(\mathbf{Z}_p) \times U^p$.*

The subgroups that we shall be concerned with are

$$U_w(N) = \left\{ g \in GL_2(\hat{\mathbf{Z}}) \,\middle|\, g \equiv \begin{pmatrix} * & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\}$$

and

$$U(N) = \left\{ g \in GL_2(\hat{\mathbf{Z}}) \,\middle|\, g \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\}.$$

These define the modular curves $X_w(N)$ and $X(N)$, respectively. The modular interpretation of $X(N)$ is given as follows: for each $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$, choose a matrix $g_\epsilon \in GL_2(\hat{\mathbf{Z}})$ congruent to $\begin{pmatrix} \epsilon^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ mod N. The strong approximation theorem for GL_2 implies that every point in $Y(N)$ has a representative of the form (z, g_ϵ) for some unique choice of ϵ; we let this point correspond to the elliptic curve $\mathbf{C}/\langle z, 1\rangle$ together with the basis for its N-torsion given by $(\epsilon z/N, 1/N)$. We then have an action of $GL_2(\mathbf{Z}/N\mathbf{Z})$ on $X(N)$ that sends a matrix $\overline{g} \in GL_2(\mathbf{Z}/N\mathbf{Z})$ to the map $(g^{-1})^* \colon X(N) \to X(N)$, where g is any lifting of $\overline{g}$ to $GL_2(\hat{\mathbf{Z}})$; it has the modular interpretation of preserving the elliptic curve and having $\overline{g}$ act on the basis for its N-torsion on the left.

Note that, in contrast, the action of $SL_2(\mathbf{Z}/N\mathbf{Z})$ on $X_w(N)$ can't easily be defined adelically; this is one reason why one can't define such an action over $\mathbf{Q}$, and thus why we find it convenient to use the curves $X(N)$ rather than $X_w(N)$ at times. However, with a bit of care it is possible to use the action of $GL_2(\mathbf{Z}/N\mathbf{Z})$ on $X(N)$ to extract information about the action of $SL_2(\mathbf{Z}/N\mathbf{Z})$ on $X_w(N)$; we shall do this in Section 8.

Now we turn to the surfaces $X_\simeq(N)$. Definitions similar to the above go through, where we replace $\mathfrak{H} \times GL_2(\mathbf{A}^\infty)$ by $\mathfrak{H} \times \mathfrak{H} \times GL_2(\mathbf{A}^\infty) \times GL_2(\mathbf{A}^\infty)$ and put in two copies of everything else. We then recover our surfaces $X_\simeq(N)$ and spaces $S_{k,\simeq}(N)$ of cusp forms by using the following subgroup:

$$U_\simeq(N) = \left\{ (g_1, g_2) \in GL_2(\hat{\mathbf{Z}}) \times GL_2(\hat{\mathbf{Z}}) \,\middle|\, g_1 \equiv g_2 \pmod N \right\}.$$

Using these definitions, we easily see that that, as claimed,

$$X_{\simeq}(N) = \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})\backslash(X(N) \times X(N)),$$

where $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$ acts diagonally with the action given above.

In contrast with this situation, there does *not* exist a subgroup $U_{\simeq,\epsilon}(N)$ that would allow us to define $X_{\simeq,\epsilon}(N)$ in the same way; this explains why we couldn't naturally define a Hecke operator T_{n_1,n_2} acting on $X_{\simeq,\epsilon}(N)$ unless $n_1 \equiv n_2 \pmod{N}$. Of course, it isn't hard to see which points on $X_{\simeq}(N)$ are on $X_{\simeq,\epsilon}(N)$ for some ϵ: they are the points that have a representative of the form (z_1, z_2, g_1, g_2) with $g_i \in \mathrm{GL}_2(\hat{\mathbf{Z}})$ and with $\det g_1 \equiv \epsilon \det g_2 \pmod{N}$. And if we are given $\mathbf{f} \in S_k(U_{\simeq}(N)) = S_{k,\simeq}(N)$, we can recover $\mathbf{f}_\epsilon$ from it by letting $\mathbf{f}_\epsilon(z_1, z_2) = \mathbf{f}(z_1, z_2, 1, g_\epsilon)$.

The above definitions of Hecke operators also pass over immediately to our situation; in particular, one can check that the operators T_{p_1,p_2} defined in Section 4 (for $(p_i, N) = 1$) can be thought of $T_{p_1} \times T_{p_2}$, where each T_{p_i} is an operator on $X(N)$ and the product descends to $X_{\simeq}(N)$. Similarly, $\langle p\rangle$ is $1 \times S_p$ (again for $(p, N) = 1$; note that $S_p \times 1$ is $\langle p^{-1}\rangle$). We shall sketch a proof of a similar statement in Lemma 8.1 below.

8. The case of prime level

In this section, we discuss facts that are special to the case of weight $(2, 2)$ forms on prime level. The main fact here is that we can ignore Fourier coefficients that are multiples of p, as stated in Proposition 2.4; this in turn implies that certain spaces of cusp forms are free of rank one over their Hecke algebras, as stated in Corollaries 4.3 and 5.1. In the rest of this section, we shall present some general calculations that lead us towards methods for calculating the spaces $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$; in the next Section, we shall give some explicit constructions of forms.

Since

$$S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p)) \simeq (S_2(\Gamma_w(p)) \otimes S_2(\Gamma_w(p)))^{\mathrm{SL}_2(\mathbf{Z}/p\mathbf{Z})},$$

to understand $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ we should understand the representation theory of $\mathrm{SL}_2(\mathbf{Z}/p\mathbf{Z})$ on $S_2(\Gamma_w(p))$. Since $\left(\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right)$ acts trivially on $S_2(\Gamma_w(p))$, we can look at the representation theory of $\mathrm{PSL}_2(\mathbf{F}_p)$ instead. We shall start by considering arbitrary weights and levels, and adding the assumptions of weight 2 and level p as it becomes convenient.

The basic fact about representations of groups on spaces of cusp forms is the Strong Multiplicity One Theorem. This tells us how to pick out the irreducible representations of $\mathrm{GL}_2(\mathbf{A}^\infty)$ that are contained in $S_k(\mathbf{C})$: they are just the Hecke eigenspaces. Taking $U(N)$-invariants, this breaks up $S_k(U(N))$ into a sum of representations of $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$, one per eigenspace. (Of course, these smaller representations may not be irreducible as representations of $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$.) Thus, we wish to understand the eigenspaces of $S_k(U(N))$.

First we recall that $\begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}^{-1} \Gamma_w(N) \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \supset \Gamma_1(N^2)$. This allows us to pass from forms on $X_w(N)$ to forms on $X_1(N^2)$: the image of $S_k(\Gamma_w(N))$ under the action of $\left(\begin{smallmatrix} N & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is the direct sum of the spaces $S_k(\Gamma_0(N^2), \chi)$ where χ is a character on $(\mathbf{Z}/N\mathbf{Z})^*$. A form $f = \sum c_m q^m$, where $q = e^{2\pi\sqrt{-1}z/N}$, gets sent to a form with the same Fourier expansion except that q is now equal to $e^{2\pi\sqrt{-1}z}$. Furthermore, if ψ is a character on $(\mathbf{Z}/N\mathbf{Z})^*$ then the form f_ψ, which is defined to have Fourier expansion $\sum c_m \psi(m) q^m$, is still a form in $S_k(\Gamma_w(N))$, by Shimura [12], Proposition 3.64.

We now try to produce forms contained in $S_k(U(N))$. A form $\mathbf{f} \in S_k(U(N))$ is a function from $\mathfrak{H} \times \mathrm{GL}_2(\mathbf{A}^\infty)$ to $\mathbf{C}$ with those properties listed in Section 7; it then follows easily that if, for $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$, we define $\mathbf{f}_\epsilon$ by setting $\mathbf{f}_\epsilon(z) = \mathbf{f}(z, g_\epsilon)$ (where g_ϵ is a matrix in $\mathrm{GL}_2(\hat{\mathbf{Z}})$ that is congruent to $\left(\begin{smallmatrix} \epsilon^{-1} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ mod N) then each of the $\mathbf{f}_\epsilon$'s is a form in $S_k(\Gamma_w(N))$. By the Strong Approximation Theorem, a choice of such $\mathbf{f}_\epsilon$'s determines $\mathbf{f}$ uniquely. Thus, we can think of forms on $S_k(U(N))$ as $\phi(N)$-tuples of forms on $S_k(\Gamma_w(N))$.

This allows us to determine the Hecke eigenspaces in $S_k(U(N))$. The dimension of $S_k(U(N))$ is $\phi(N)$ times the dimension of $S_k(\Gamma_w(N))$, so the hope is that each eigenform on $S_k(\Gamma_w(N))$ will somehow give us $\phi(N)$ different eigenforms on $S_k(U(N))$. This is indeed what happens, as we shall see in Proposition 8.1:

Lemma 8.1. *Let* $\mathbf{f}$ *be an element of* $S_k(U(N))$ *and let* q *be a prime not dividing* N*. Then, for all* $\epsilon \in (\mathbf{Z}/N\mathbf{Z})^*$, $(T_q\mathbf{f})_\epsilon = T_q(\mathbf{f}_{\epsilon q})$ *and* $(S_q\mathbf{f})_\epsilon = S_q(\mathbf{f}_{\epsilon q^2})$*. (In both equalities, the operators on the left hand side of the equations are the adelic operators mentioned in Sect. 7 while the operators on the right hand side of the equation are the classical operators.)*

Proof. Let Δ_q be the $U(N)$-double coset defining the adelic operator T_q, and let $\Delta_q = \coprod_i \delta_i U(N)$. Then, up to an appropriate normalizing constant, $(T_q\mathbf{f})_\epsilon(z) = \sum_i \mathbf{f}(z, g_\epsilon \delta_i)$. But $\det(g_\epsilon \delta_i) = q^{-1}\epsilon^{-1}$; by the Strong Approximation Theorem, we can therefore write $g_\epsilon \delta_i$ as $\delta_i' g_{q\epsilon} u_i$ for some $\delta_i' \in \mathrm{GL}_2^+(\mathbf{Q})$ and $u_i \in U(N)$. But $\mathbf{f}$ is invariant by right multiplication by $U(N)$, and multiplying $g_{q\epsilon}$ on the left by an element δ_i' of $\mathrm{GL}_2^+(\mathbf{Q})$ has the effect of replacing $\mathbf{f}_{q\epsilon}$ by $\mathbf{f}_{q\epsilon}|_{(\delta_i')^{-1}}$. One then checks that $(\delta_i')^{-1}$ is contained in $g_{q\epsilon}\Delta_q^{-1}g_\epsilon^{-1}$, and that this is the $\Gamma(N)$-double coset defining the classical Hecke operator T_q. The proof for the S_q's is similar. □

Corollary 8.1. *Let* $g \in S_k(\Gamma_w(N))$ *be an eigenform, with eigenvalues* $\{a_q, \chi(q)\}$ *(for* T_q *and* S_q *respectively, as* q *varies over primes not dividing* N*). Let* ψ *be a character of* $(\mathbf{Z}/N\mathbf{Z})^*$*. Then the form* $\mathbf{f}(g, \psi) \in S_k(U(N))$ *defined by* $\mathbf{f}(g, \psi)_\epsilon = \psi(\epsilon) g$ *is an eigenform with eigenvalues* $\{\psi(q)a_q, \psi^2(q)\chi(q)\}$.

Proof. Write $\mathbf{f}$ for $\mathbf{f}(g, \psi)$. By the Lemma,

$$\begin{aligned}(T_q\mathbf{f})_\epsilon &= T_q(\mathbf{f}_{\epsilon q}) = T_q(\psi(\epsilon q) g) \\ &= \psi(q)\psi(\epsilon) a_q g = \psi(q) a_q \mathbf{f}_\epsilon.\end{aligned}$$

The calculation for S_q proceeds in exactly the same manner. □

This allows us to produce a basis of eigenforms for $S_k(U(N))$ in terms of a basis of eigenforms for $S_k(\Gamma_w(N))$:

Proposition 8.1. *Let $\{g_j\}$ be a basis of eigenforms for $S_k(\Gamma_w(N))$. Then the set of forms $\{\mathbf{f}(g_j, \psi)\}$, as g_j varies over elements of the basis and ψ varies over characters of $(\mathbf{Z}/N\mathbf{Z})^*$, give a basis of eigenforms for $S_k(U(N))$. Every set $\{a_q, \chi(q)\}$ of eigenvalues for T_q and S_q (as q runs over primes not dividing N) that occurs in $S_k(U(N))$ occurs in $S_k(\Gamma_w(N))$. A basis for the set of eigenforms in $S_k(U(N))$ with eigenvalues $\{a_q, \chi(q)\}$ is given by taking the forms $\mathbf{f}(g, \psi)$ where ψ varies over the characters of $(\mathbf{Z}/N\mathbf{Z})^*$ and where, once ψ is fixed, g varies over a basis for those eigenforms in $S_k(\Gamma_w(N))$ which have eigenvalues $\{a_q\psi^{-1}(q), \chi(q)\psi^{-2}(q)\}$.*

Proof. Assume that we have an expression of linear dependence involving the forms $\mathbf{f}(g_j, \psi)$. Looking at the first coordinate, the fact that the forms $\{g_j\}$ form a basis for $S_k(\Gamma_w(N))$ implies that we can assume that our relation involves only forms $\mathbf{f}(g, \psi)$ for some fixed form g. But those forms are linearly independent since characters are linearly independent. This gives us $\phi(N) \cdot \dim S_k(\Gamma_w(N))$ forms; but that's the dimension of $S_k(U(N))$, so those forms give a basis for $S_k(U(N))$ that consists of eigenforms.

Every set of eigenvalues on $S_k(U(N))$ can therefore be written in the form $\{\psi(q)a_q, \psi^2(q)\chi(q)\}$, where $\{a_q, \chi(q)\}$ is the set of eigenvalues of a form $g \in S_k(\Gamma_w(N))$, by Corollary 8.1. But those are the eigenvalues of g_ψ, which is also an eigenform in $S_k(\Gamma_w(N))$. The last statement of the Proposition follows in a similarly direct manner from the first paragraph of the proof and Corollary 8.1. □

To restate the last sentence of the above Proposition: assume that g is a newform in $S_k(\Gamma_w(N))$ with eigenvalues $\{a_p, \chi(p)\}$, where by "newform" we mean that $g|_{\left(\begin{smallmatrix} N & 0 \\ 0 & 1 \end{smallmatrix}\right)} \in S_k(\Gamma_1(N^2))$ is a newform. A basis for the space of eigenforms in $S_k(U(N))$ with those eigenvalues is given by the forms $\mathbf{f}(g_{\psi^{-1}}, \psi)$ together with the forms $\mathbf{f}(h, \psi)$ where h runs over oldforms with the same eigenvalues as $g_{\psi^{-1}}$.

Let us now fix $k = 2$ and $N = p$ prime. We may assume that $p > 5$, since $S_2(\Gamma_w(p))$ is zero otherwise. Pick a set $A = \{a_q, \chi(q)\}$ of eigenvalues. Let $g \in S_2(\Gamma_w(p))$ be a newform with those eigenvalues; we wish to calculate the dimension of the space S_A of forms in $S_2(U(p))$ with eigenvalues A. For each character ψ, we can produce an element of S_A all of whose components are multiples of $g_{\psi^{-1}}$; this gives us $(p-1)$ forms. Furthermore, when $g_{\psi^{-1}}$ is an oldform, we can produce extra forms. Since $S_2(\Gamma(1))$ is zero, we can produce at most one extra form for each ψ this way: this happens when the eigenvalues $\{a_q\psi^{-1}(q), \chi(q)\psi^{-2}(q)\}$ occur in $S_2(\Gamma_1(p))$.

For how many ψ does an extra form arise in this way? By the Strong Multiplicity One Theorem, studying S_A reduces to the study of irreducible representations of $\mathrm{GL}_2(\mathbf{A}^\infty)$ and their $U(p)$-invariants. Factoring those representations, we have to study irreducible representations of $\mathrm{GL}_2(\mathbf{Q}_q)$ and their $U(p)_q$-invariants. If $q \neq p$ then $U(p)_q = \mathrm{GL}_2(\mathbf{Z}_q)$; since the space of $\mathrm{GL}_2(\mathbf{Z}_q)$ invariants of an irreducible representation of $\mathrm{GL}_2(\mathbf{Q}_q)$ is either zero- or one-dimensional, we can therefore concentrate on the irreducible representations of $\mathrm{GL}_2(\mathbf{Q}_p)$, and in particular calculating the dimension of their $U(p)_p$-invariants, where

$$U(p)_p = \left\{ g \in \mathrm{GL}_2(\mathbf{Z}_p) \;\middle|\; g \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{p} \right\}.$$

Irreducible representations of $\mathrm{GL}_2(\mathbf{Q}_p)$ can be classified as *principal series, special,* or *supercuspidal*. If the space of $U(p)_p$-invariants is nonzero then it is $(p+1)$-, p-, or $(p-1)$-dimensional, depending on which classification it falls into; thus, we have two, one, or no extra dimensions of oldforms arising in the principal series, special, and supercuspidal cases, respectively.

Let us now turn towards the space $S_2(\Gamma_w(p))$. The group $\mathrm{PSL}_2(\mathbf{F}_p)$ acts on this space; we wish to determine its irreducible representations. Since this action is not given adelically, we can't just apply the theory of irreducible $\mathrm{GL}_2(\mathbf{A}^\infty)$-representations and the Strong Multiplicity One Theorem to get the answer. However, we can use the adelic action to get information about this representation as follows: let g be an element of $S_2(\Gamma_w(p))$ and let $\mathbf{f}$ be an element of $S_2(U(p))$ such that $\mathbf{f}_1 = g$. Let $\overline{\gamma}$ be an element of $\mathrm{PSL}_2(\mathbf{F}_p)$ and let γ be an element of $\mathrm{SL}_2(\mathbf{Z}_p)$ projecting to it. Then $\overline{\gamma}$ sends g to $(\gamma_*^{-1}\mathbf{f})_1$, as can be seen by tracing through the definitions. In particular, we get representations of $\mathrm{PSL}_2(\mathbf{F}_p)$ on $S_2(\Gamma_w(p))$ by projecting the representations given in the previous paragraphs down to their first coordinate.

The map from $S_2(U(p))$ to $S_2(\Gamma_w(p))$ sending $\mathbf{f}$ to $\mathbf{f}_1$ is injective unless there is a ψ such that $g = g_\psi$, by Proposition 8.1, i.e. unless g is a CM-form, in which case all of the forms in the representation are CM-forms, and the dimension of the representation in $S_2(\Gamma_w(p))$ is half of the dimension of the representation in $S_2(U(p))$. Thus, we have decomposed $S_2(\Gamma_w(p))$ as a direct sum of representations that are either of dimension $p-1$, p, $p+1$, $(p-1)/2$, or $(p+1)/2$.

These representations may not be irreducible, however. Most of the time, they do turn out to be irreducible; we can see this by looking at the character table of $\mathrm{PSL}_2(\mathbf{F}_p)$. (See Fulton and Harris [2], §5.2, for example.) The dimensions of the irreducible representations of $\mathrm{PSL}_2(\mathbf{F}_p)$ are 1, $p-1$, p, $p+1$, and either $(p-1)/2$ (if $p \equiv 3 \pmod 4$) or $(p+1)/2$ (if $p \equiv 1 \pmod 4$). Furthermore, the only one-dimensional representation of $\mathrm{PSL}_2(\mathbf{F}_p)$ is the trivial one, which doesn't occur in $S_2(\Gamma_w(p))$ (since that would be equivalent to having a form that is invariant under $\mathrm{PSL}_2(\mathbf{F}_p)$, i.e. a form in $S_2(\Gamma(1))$). There are no 2-dimensional representations, either, so by comparing dimensions, we see that the representations that we have constructed above are either irreducible or the direct sum of two representations of dimension $(p-1)/2$ or $(p+1)/2$.

We wish to see how $\dim S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ varies as a function of ϵ. Write $\chi_w(p)$ for the character of $S_2(\Gamma_w(p))$, considered as a representation of $\mathrm{PSL}_2(\mathbf{F}_p)$. Then

$$\begin{aligned}
\dim S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p)) &= \dim(S_2(\Gamma_w(p)) \otimes S_2(\Gamma_w(p)) \circ \theta_\epsilon)^{\mathrm{PSL}_2(\mathbf{F}_p)} \\
&= \langle \chi_w(p) \otimes \chi_w(p) \circ \theta_\epsilon, 1 \rangle \\
&= \left\langle \chi_w(p), \overline{\chi_w(p)} \circ \theta_\epsilon \right\rangle.
\end{aligned}$$

(Here, θ_ϵ is the automorphism of $\mathrm{PSL}_2(\mathbf{F}_p)$ given at the beginning of Sect. 2. Assume that $S_2(\Gamma_w(p))$ has $\bigoplus_i R_i^{\oplus n_i}$ as its decomposition into a sum of irreducible representations. Then, by the above,

$$\dim S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p)) = \sum_{\substack{i,j \\ R_i \simeq \overline{R_j} \circ \theta_\epsilon}} n_i n_j.$$

Now assume that $p \equiv 1 \pmod 4$. Examining the character table of $\mathrm{PSL}_2(\mathbf{F}_p)$, we see that $R_i \simeq \overline{R_i}$ for all R_i and that $R_i \simeq R_i \circ \theta_\epsilon$ for all ϵ unless $R_i \simeq W'$ or W'', where W' and W'' are the irreducible representations of dimension $(p+1)/2$. In this latter case, composing with θ_ϵ switches W' and W'' if ϵ is not a square. Now assume that W' occurs n' times in the decomposition of $S_2(\Gamma_w(p))$ and W'' occurs n'' times. Then, if ϵ_1 is a square and ϵ_2 isn't, the above discussion shows that

$$\begin{aligned} \dim S_{(2,2)}(\Gamma_{\simeq,\epsilon_1}(p)) - \dim S_{(2,2)}(\Gamma_{\simeq,\epsilon_2}(p)) &= n'^2 + n''^2 - 2n'n'' \\ &= (n' - n'')^2. \end{aligned}$$

This is a bit misleading, however, because in this case n' and n'' are equal, so the dimension of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ is the same for all ϵ. We can see this by calculating n' and n'' using Ligozat [9], Proposition III.1.3.2.1: the characters of W' and W'' only differ in matrices that are conjugate to $\left(\begin{smallmatrix}1 & * \\ 0 & 1\end{smallmatrix}\right)$, and the only place that such matrices occur in the formula given there is in the term $\sum_{a \bmod p} \chi\left(\left(\begin{smallmatrix}1 & a \\ 0 & 1\end{smallmatrix}\right)\right)$, which equals p both for $\chi = \chi_{W'}$ and $\chi = \chi_{W''}$.

As a corollary, this implies that there are no CM-forms in $S_2(\Gamma_w(p))$ for $p \equiv 1 \pmod 4$. For if there were such a form g, it would generate an irreducible representation $R_g \subset S_2(\Gamma_w(p))$, all of whose elements would be CM-forms; there would then be a form in $R_g \otimes (R_g \circ \theta_{-1})$ that is invariant under $\mathrm{PSL}_2(\mathbf{F}_p)$. But such a form would be a CM-form in $S_{(2,2)}(\Gamma_{\simeq,-1}(p))$, so Theorem 6.1 would then imply that the dimension of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ for ϵ a non-square is strictly smaller than the dimension of $S_{(2,2)}(\Gamma_{\simeq,-1}(p))$, contradicting our calculations above.

Let us now turn to the case where $p \equiv 3 \pmod 4$. This time, $R_i \simeq \overline{R_i}$ unless $R_i \simeq X'$ or X'', where X' and X'' are the irreducible representations of dimension $(p-1)/2$; $\overline{X'} \simeq X''$ and vice-versa. Similarly, $R_i \circ \theta_\epsilon \simeq R_i$ unless $R_i \simeq X'$ or X'' and ϵ is not a square mod p; if it is, $X' \circ \theta_\epsilon \simeq X''$ and vice-versa. Thus, if X' occurs n' times and X'' occurs n'' times in the decomposition of $S_2(\Gamma_w(p))$,

$$\begin{aligned} \dim S_{(2,2)}(\Gamma_{\simeq,\epsilon_1}(p)) - \dim S_{(2,2)}(\Gamma_{\simeq,\epsilon_2}(p)) &= 2n'n'' - (n'^2 + n''^2) \\ &= -(n' - n'')^2, \end{aligned}$$

where ϵ_1 is a square mod p and ϵ_2 isn't. Since -1 is not a square, the dimension is maximized when $\epsilon = -1$, agreeing with Corollary 5.4.

This time, however, $n' - n''$ is non-zero. We can't calculate it as easily as we calculated it in the previous case, because the method used there calculates the number of times a representation occurs plus the number of times that its complex conjugate occurs, and here the character is no longer totally real. Instead, we refer to Hecke [4], where he proves that the difference is equal to the class number $h(-p)$ of $\mathbf{Q}(\sqrt{-p})$. Thus,

$$\dim S_{(2,2)}(\Gamma_{\simeq,-1}(p)) - \dim S_{(2,2)}(\Gamma_{\simeq,1}(p)) = h(-p)^2.$$

This implies that there are exactly $h(-p) \cdot (p-1)/2$ CM-forms contained in $S_2(\Gamma_w(p))$; they have been constructed by Hecke in [3]. We shall review his construction in Sect. 9, and use them to write down the Hecke kernel $K_{\simeq}(p)$ explicitly. We shall also show how to use the theory outlined in this Section to perform explicit calculations of spaces $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ for small primes.

To recap:

Theorem 8.1. *If p is a prime congruent to 1 mod 4 then there are no CM-forms contained in $S_2(\Gamma_w(p))$ and the Hecke kernel $K_{\simeq}(p)$ is zero. If $p > 3$ is congruent to 3 mod 4 then there are $h(-p)\cdot(p-1)/2$ CM-forms contained in $S_2(\Gamma_w(p))$ and $K_{\simeq}(p)$ has dimension $(h(-p))^2$, where $h(-p)$ is the class number of $\mathbf{Q}(\sqrt{-p})$.*

9. Examples

$X_{\simeq,-1}(7)$

The first $X_{\simeq,\epsilon}(p)$ to have a non-zero (2, 2)-cusp form is $X_{\simeq,-1}(7)$, as can be seen by looking at Table 1 in Kani and Schanz [7] (and using Corollary 2.2 above); in fact, we see that $\dim S_{(2,2)}(\Gamma_{\simeq,-1}(7)) = 1$. We can explicitly determine a non-zero form in this space as follows:

Conjugating $\Gamma_w(7)$ by $\left(\begin{smallmatrix} 7 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, we can think of $X_w(7)$ as lying between the curves $X_0(49)$ and $X_1(49)$. The former is an elliptic curve (after choosing a base point); its L-series gives rise to a weight two cusp form

$$f(z) = \sum_{m>0} c_m q^m$$

on $X_0(49)$ and $X_w(7)$. (Here, $q = e^{2\pi\sqrt{-1}z}$ if we are thinking of f as a form on $X_0(49)$ and $q = e^{2\pi\sqrt{-1}z/7}$ if we are thinking of f as a form on $X_w(7)$.) If χ is a non-trivial character on $(\mathbf{Z}/7\mathbf{Z})^*$ such that $\chi(-1) = 1$ then the functions

$$f_\chi(z) = \sum_{m>0} c_m \chi(m) q^m$$

and

$$f_{\chi^2}(z) = \sum_{m>0} c_m \chi^2(m) q^m$$

are also modular forms in $S_2(\Gamma_w(7))$, by Shimura [12], Proposition 3.64; since the latter space is three-dimensional, $\{f, f_\chi, f_{\chi^2}\}$ forms a basis for it. For $n \in (\mathbf{Z}/7\mathbf{Z})^*$, we have $f_\chi|_{\sigma_a} = \chi^2(a) f_\chi$ and $f_{\chi^2}|_{\sigma_a} = \chi(a) f_{\chi^2}$.

To produce an element of $S_{(2,2)}(\Gamma_{\simeq,-1}(7))$, we have to find a form contained in $S_2(\Gamma_w(7)) \otimes S_2(\Gamma_w(7))$ that is fixed by $\mathrm{PSL}_2(\mathbf{F}_7)$ (acting on the second factor via θ_{-1}). For our form to be fixed by the matrices (σ_a, σ_a), it has to be of the form

$$a_0 \cdot f \otimes f + a_1 \cdot f_\chi \otimes f_{\chi^2} + a_2 \cdot f_{\chi^2} \otimes f_\chi .$$

And for our form to be fixed by the matrix $\left(\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\right)$, we must have $a_0 = a_1 = a_2$. However, those constraints leave us with only a one-dimensional space of possible cusp forms, and since $S_{(2,2)}(\Gamma_{\simeq,-1}(7))$ is non-empty, we see that it must be generated by the form

$$g = \frac{1}{3}(f \otimes f + f_\chi \otimes f_{\chi^2} + f_{\chi^2} \otimes f_\chi) = \sum_{m_1 \equiv m_2 \pmod 7} c_{m_1} c_{m_2} q_1^{m_1} q_2^{m_2},$$

where the c_i's are the coefficients of f as above.

Now that we've got our form g in hand, we'd like to relate it to some of our general theorems about forms in $\overline{S}_k(\Gamma_{\simeq,\epsilon}(N))$. Note that g has lots of Fourier coefficients that are zero: not only is $c_{m_1,m_2}(g)$ zero unless $m_1 \equiv m_2 \pmod 7$, but it's also zero unless the m_i's are squares mod 7. (This follows from the fact that the elliptic curve $X_0(49)$ has complex multiplication by $\mathbf{Q}(\sqrt{-7})$.) By Proposition 6.1, our form is therefore in $K_{\simeq}(7)$; indeed, $S_{(2,2)}(\Gamma_{\simeq,1}(7))$ is trivial.

$X_{\simeq,-1}(p)$ *for* $p \equiv 3 \pmod 4$

The above may look like a general recipe for producing forms on $X_{\simeq,\epsilon}(p)$ out of forms on $X_0(p^2)$, but it isn't. To see why, note that the transition involved two steps: matching up characters, which involved checking invariance under the matrices (σ_a, σ_a), and making sure that certain Fourier coefficients were zero, which involved checking invariance under the matrices $\left(\begin{pmatrix}1 & \epsilon\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\right)$. Thus, we checked that our putative form is invariant under the subgroup $B(p)$ of upper-triangular matrices, not all of $\mathrm{PSL}_2(\mathbf{F}_p)$. The reason why we could get away with that above was that we knew a lot about $S_2(\Gamma_w(7))$ and that the dimension of $S_{(2,2)}(\Gamma_{\simeq,-1}(7))$ was 1.

Fortunately, all is not lost for more general p. Let R_1 and R_2 be irreducible representations occurring in $S_2(\Gamma_w(p))$. As the discussion in Sect. 8 showed, $R_1 \otimes R_2$ contributes 1 to the dimension of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$ iff $R_1 = \overline{R_2 \circ \theta_\epsilon}$. Now, assume that that is indeed the case, and that, furthermore, R_1 is irreducible as a representation of $B(p)$. Writing χ_i for the character of R_i, it will then also be the case that

$$\left\langle \chi_1 \cdot (\chi_2 \circ \theta_\epsilon), 1_{B(p)} \right\rangle_{B(p)} = \left\langle \chi_1, \overline{\chi_2 \circ \theta_\epsilon} \right\rangle_{B(p)} = 1.$$

But this says that there's only a one-dimensional space of vectors in $R_1 \otimes R_2$ that is fixed by $B(p)$, and since there is also a one-dimensional space of vectors in $R_1 \otimes R_2$ that is fixed by $\mathrm{PSL}_2(\mathbf{F}_p)$, they must be the same space. Thus, under the hypothesis that our representation is irreducible when considered as a representation of $B(p)$, we can test to see whether an element of $R_1 \otimes R_2$ is a cusp form on $X_{\simeq,\epsilon}(p)$ simply by making sure that it is invariant under (σ_n, σ_n) and $\left(\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & \epsilon\\ 0 & 1\end{pmatrix}\right)$.

To make this concrete, assume that p is congruent to $3 \pmod 4$ but not equal to 3 and that $\epsilon = -1$. In this case, the representations X' and X'' of $\mathrm{PSL}_2(\mathbf{F}_p)$ remain irreducible when restricted to $B(p)$. Thus, if we can produce representations isomorphic to X' or X'' in $S_2(\Gamma_w(p))$, we'll be able to explicitly write down forms in $S_{(2,2)}(\Gamma_{\simeq,-1}(p))$. We saw that there should be $h(-p)$ such representations coming from CM-forms; they would be good ones to look for.

Fortunately, those representations are produced in Hecke [3]. They are defined as follows: let I be an integral ideal in $\mathbf{Q}(\sqrt{-p})$ with norm A and let ρ be an element of I. We define a theta series as follows:

$$\theta_H(z;\rho,I,\sqrt{-p}) = \sum_{\substack{\mu\in I\\ \mu\equiv\rho(\operatorname{mod} I\sqrt{-p})}} \mu e^{2\pi\sqrt{-1}z\frac{\mu\bar{\mu}}{pA}},$$

where $\overline{\mu}$ is the complex conjugate of μ. Letting V_I be the vector space generated by the functions $\theta_H(z;\rho,I,\sqrt{-p})$ for $\rho\in I$, the results of Hecke [3], §7 show that V_I only depends on the ideal class of I, that these θ_H's are elements of $S_2(\Gamma_w(p))$, and that V_I is a representation of $\mathrm{PSL}_2(\mathbf{F}_p)$ isomorphic to X'. This gives us our desired $h(-p)$ different copies of X'. (These eigenforms are also discussed in Ribet [10], §3 as being eigenforms associated to Grössencharacters.)

Now that we have our representations, we follow the same program as in the $X_{\simeq,-1}(7)$ case:

Theorem 9.1. *Let p be a prime congruent to 3* mod *4. For each ideal class of $\mathbf{Q}(\sqrt{-p})$, fix an integral ideal I in that class and an element α_I of I that's not contained in $I\sqrt{-p}$. Let*

$$f_I = \sum_{a\in(\mathbf{Z}/p\mathbf{Z})^*} \theta_H\left(z; a\left(\frac{a}{p}\right)\alpha_I, I, \sqrt{-p}\right)$$

have the Fourier expansion

$$f_I(z) = \sum_{m>0} c_{I,m}q^m,$$

where $q=e^{2\pi\sqrt{-1}z/p}$. If I_1 and I_2 are (not necessarily distinct) ideal classes then the function

$$f_{I_1,I_2}(z_1,z_2) = \sum_{m_1\equiv m_2(\operatorname{mod} p)} c_{I_1,m_1}c_{I_2,m_2}q_1^{m_1}q_2^{m_2}.$$

is an element of $S_{(2,2)}(\Gamma_{\simeq,-1}(p))$ contained in $K_{\simeq}(p)$; furthermore, the f_{I_1,I_2}'s give a basis for $K_{\simeq}(p)$ as I_1 and I_2 vary over the ideal classes of $\mathbf{Q}(\sqrt{-p})$.

Proof. The same argument as in the $p=7$ case shows that multiples of f_{I_1,I_2} are the only elements of $V_{I_1}\otimes V_{I_2}$ invariant under $B(p)$, so they are indeed elements of $S_{(2,2)}(\Gamma_{\simeq,\epsilon}(p))$. Assuming that we can show that they are in $K_{\simeq}(p)$, Theorem 8.1 shows that they give us a basis. Thus, by Theorem 6.1, we just have to verify that the forms f_I are CM-forms.

This can be seen as follows: by definition,

$$c_m(\theta_H(z;\rho,I,\sqrt{-p})) = \sum_{\substack{\mu\in I \\ \mu\equiv\rho \pmod{I\sqrt{-p}} \\ \mu\overline{\mu}=mA}} \mu,$$

where A is the norm of I But $\mu\overline{\mu}$ is a square mod p for all μ in the ring of integers of $\mathbf{Q}(\sqrt{-p})$, as is A, so c_m is zero unless m is a square mod p. Thus, f_I is invariant under twisting by the quadratic character of $(\mathbf{Z}/p\mathbf{Z})^*$, hence a CM-form. $\square$

Acknowledgements. I wish to thank Fred Diamond, Jordan Ellenberg, Steven Kleiman, and Barry Mazur for the help that they have given me while writing this paper; and the N.D.S.E.G. Fellowship Program for the support that it has provided.

References

[1] F. Diamond and J. Im: Modular forms and modular curves. In V. K. Murty, editor, *Seminar on Fermat's Last Theorem*, Providence, AMS, 1995, pp. 39–133
[2] W. Fulton and J. Harris: *Representation Theory*. Springer-Verlag, New York, 1991
[3] E. Hecke: Zur Theorie der elliptischen Modulfunktionen. *Math. Annalen* **97**, 210–242 (1926) Werke, pp. 428–460
[4] E. Hecke: Über das Verhalten der Integrale 1. Gattung bei Abbildungen, insbesondere in der Theorie der elliptischen Modulfunktionen. *Abh. Math. Sem. Hamburg*, **8**, 271–281 (1930) Werke, pp. 548–558
[5] C. F. Hermann: Modulflächen quadratischer Diskriminante. *manuscripta math.* **72**, 95–110 (1991)
[6] E. Kani and W. Schanz: Diagonal quotient surfaces. *manuscripta math.* **93**, 67–108 (1997)
[7] E. Kani and W. Schanz: Modular diagonal quotient surfaces. *Math. Zeitschrift* **227**, 337–366 (1998)
[8] S. Lang: *Introduction to Modular Forms*. Springer-Verlag, New York, 1976.
[9] G. Ligozat: Courbes modulaires de niveau 11. In Serre and Zagier [11], pp. 149–237.
[10] K. A. Ribet: Galois representations attached to eigenforms with nebentypus. In Serre and Zagier [11], pp. 17–52
[11] J.-P. Serre and D. Zagier (eds.): *Modular Functions of One Variable V*, Lecture Notes in Mathematics 601, Springer-Verlag, Berlin, 1977
[12] G. Shimura: *Introduction to the Arithmetic Theory of Automorphic Functions*. Princeton University Press, Princeton, 1971