In this final lecture we survey some generalizations of the (classical) modular forms we have been studying in this course (some of this material can be found in [4, Chapter 15] but much of it is taken from the literature).

## 23.1 Half-integral weight modular forms

Recall that for integers $k$ and lattices $\Gamma \le \mathrm{SL}_2(\mathbb{R})$ we defined weight-$k$ modular forms for $\Gamma$ as holomorphic functions $f\colon \mathbf{H} \to \mathbb{C}$ that satisfy

$$f|_k\gamma = f$$

for every $\gamma \in \Gamma$ and are holomorphic at the cusps of $\Gamma$, where the weight-$k$ slash operator is

$$f|_k\gamma := j(\gamma,z)^{-k} f(\gamma z)$$

with $j(\gamma,z) := cz+d$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$. We now want to consider half integral values of $k$.

To extend this to $k \in \frac{1}{2} + \mathbb{Z}$ we need to define the square root of $j(\gamma,z)$ for $z \in \mathbf{H}$ and $\gamma \in \Gamma$ in a way that makes $f|_k\gamma$ holomorphic and which is compatible with the action of $\Gamma$ on $\mathbf{H}$. In particular, we want an analog of the cocycle condition $j(\gamma_1\gamma_2, z) = j(\gamma_1, \gamma_2 z) j(\gamma_2, z)$ to hold. One might be tempted to take $(cz+d)^{1/2} := \sqrt{cz+d}$, where $\sqrt{\cdot}$ denotes the principal branch (so $\mathrm{Arg}(\sqrt{z}) \in (-\frac{\pi}{2}, \frac{\pi}{2}])$ and then use this to define $(cz+d)^{-k}$ for $k = n/2$. Note that either $c = 0$ and $cz+d$ is constant or $cz+d\colon \mathbf{H} \to \mathbb{C}$ lies in a halfplane on which $\sqrt{z}$ is holomorphic, so we don't need to worry about what happens on the branch locus $\mathbb{R}_{<0}$. But note that $(\sqrt{z})^{-n} = \pm\sqrt{z^{-n}}$, and the sign varies with $z$, so it is not completely clear what $(cz+d)^{-k}$ means, and even if we fix a choice, neither choice satisfies the cocycle condition consistently.

Instead we extend $\mathrm{SL}_2(\mathbb{R})$ to a connected double cover of $\mathrm{SL}_2(\mathbb{R})$ known as the metaplectic group $\mathrm{Mp}_2(\mathbb{R})$; The fundamental group of $\pi_1(\mathrm{SL}_2(\mathbb{R})) \simeq \mathbb{Z}$ has a unique quotient of order 2, so $\mathrm{Mp}_2(\mathbb{R})$ is uniquely determined as a topological space, and the fact that $\mathrm{SL}_2(\mathbb{R})$ is connected and locally path connected implies that $\mathrm{Mp}_2(\mathbb{R})$ is also a topological group; it is locally compact, since $\mathrm{SL}_2(\mathbb{R})$ is, so we can consider lattices is $\mathrm{Mp}_2(\mathbb{R})$ (discrete cofinite subgroups).

We can describe $\mathrm{Mp}_2(\mathbb{R})$ more explicitly as follows. It consists of pairs $(\gamma,\varepsilon)$ with $\gamma \in \mathrm{SL}_2(\mathbb{R})$ and $\varepsilon\colon \mathbf{H} \to \mathbb{C}$ a holomorphic function that satisfies $\varepsilon(z)^2 = j(\gamma,z)$; note that without the holomorphic condition we would have uncountably many choices for $\varepsilon(z)$, but with it there are exactly two: $\pm\sqrt{j(\gamma,z)}$. The group operation is

$$(\gamma_1,\varepsilon_1)\cdot(\gamma_2,\varepsilon_2) = (\gamma_1\gamma_2, z \mapsto \varepsilon_1(\gamma_2 z)\varepsilon_2(z)),$$

so the cocycle conduction holds. We have a surjective homomorphism $\mathrm{Mp}_2(\mathbb{R}) \to \mathrm{SL}_2(\mathbb{R})$ defined by $(\gamma,\varepsilon) \mapsto \gamma$ whose kernel is $\{(1, \pm\sqrt{z})\} \simeq \{\pm 1\}$. This leads to the exact sequence

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathrm{Mp}_2(\mathbb{R}) \longrightarrow \mathrm{SL}_2(\mathbb{R}) \longrightarrow 1$$

This sequence does not split in general, but it splits over $\Gamma(4)$; indeed, for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \Gamma(4)$, if we define $\varepsilon_\gamma(z) = \left(\frac{c}{d}\right)\sqrt{j(\gamma,z)}$, then $\widetilde{\Gamma}(4) := \{(\gamma,\varepsilon_\gamma) : \gamma \in \Gamma(4)\}$ is a lattice in $\mathrm{Mp}_2(\mathbb{R})$ and the surjection $\mathrm{Mp}_2(\mathbb{R}) \to \mathrm{SL}_2(\mathbb{R})$ restricts to an isomorphism $\widetilde{\Gamma}(4) \xrightarrow{\sim} \Gamma(4)$.[1] This leads many authors to restrict their attention to half-integral weight modular forms on congruence subgroups whose level is a multiple of 4, but one can work with any lattice in $\mathrm{Mp}_2(\mathbb{R})$. We define the weight-$k$ slash operator for elements of $\mathrm{Mp}_2(\mathbb{R})$ via

$$f|_k(\gamma,\varepsilon) = \varepsilon(z)^{-2k} f(\gamma z),$$

[1] If one enlarges the metaplectic group to allow $\varepsilon(z)^2 = \pm j(\gamma,z)$ one can extend this to $\Gamma_0(4)$.

and one can then consider modular forms of weight $k \in \frac{1}{2}\mathbb{Z}$ for lattices $\Gamma \leq \mathrm{Mp}_2(\mathbb{R})$. The metaplectic group $\mathrm{Mp}_2(\mathbb{R})$ is a compact Lie group (but not a matrix group), so in particular it is a locally compact topological group in which a lattice is any discrete cofinite subgroup.

Every lattice in $\mathrm{SL}_2(\mathbb{R})$ is also a lattice in $\mathrm{Mp}_2(\mathbb{R})$, and for $k \in \mathbb{Z}$ the definition of the weight-$k$ slash operator above is the same as our original definition. The set of half-integral weight modular forms of weight $k \in \frac{1}{2}\mathbb{Z}$ for a lattice $\Gamma \in \mathrm{Mp}_2(\mathbb{R})$ is a finite dimensional $\mathbb{C}$-vector space. The canonical example of a half-integral weight modular form is the Jacobi theta function

$$\theta(\tau) = \sum_{n \in \mathbb{Z}} e^{2\pi i n^2 \tau},$$

which is a weight-1/2 modular form for $\widetilde{\Gamma}(4)$.

One can use the Jacobi theta function to define the space $M_k(4N, \chi)$ of modular forms of weight $k \in \frac{1}{2} + \mathbb{Z}$ and level $4N$ with character $\chi$ as holomorphic functions $f : \mathbf{H} \to \mathbb{C}$ for which there exists a Dirichlet character $\chi$ of modulus $4N$ for which

$$f(\gamma z) = \chi(d) \left( \frac{\theta(\gamma z)}{\theta(z)} \right)^{2k} f(z)$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(4N)$, and which are holomorphic at all the cusps of $\Gamma_0(4N)$. Most of the theory we have developed for integral weight modular forms extends to half-integral weight, but there are subtleties that arise, and the theory of Hecke operators and eigenforms is more complicated. However, as shown by Shimura [10], there is a correspondence between half-integral weight Hecke eigenforms for $M_k(4N, \chi)$ and integral weight hecke eigenforms for $M_{2k-1}(2N, \chi^2)$. This is known as the Shimura correspondence, which is not 1-to-1 in general, but if one restricts to the Kohnen plus space it is; see [8] for details.

## 23.2 Hilbert modular forms

Let $K$ be a totally real number field with embeddings $\sigma_1, \ldots, \sigma_r : K \to \mathbb{R}$. Each $\sigma_i$ induces an embedding $\sigma_i : \mathrm{GL}_2(K) \hookrightarrow \mathrm{GL}_2(\mathbb{R})$, and we define

$$\mathrm{GL}_2^+(K) := \{\alpha \in \mathrm{GL}_2(K) : \sigma_i(\alpha) \in \mathrm{GL}_2^+(\mathbb{R}) \text{ for } 1 \leq i \leq r\},$$

as the subgroup of matrices in $\mathrm{GL}_2(K)$ with totally positive determinant. The group $\mathrm{GL}_2^+(K)$ acts on $\mathbf{H}^r$ via the isometric action of $\mathrm{GL}_2^+(\mathbb{R})$ on $\mathbf{H}$. For functions $f : \mathbf{H}^r \to \mathbb{C}$, vectors of integers $k \in \mathbb{Z}^r$, and $\alpha \in \mathrm{GL}_2^+(K)$ we define the weight-$k$ slash operator

$$(f|_k\alpha)(z) := \left( \prod_{i=1}^{r} \det(\sigma_i(\alpha))^{k_i/2} j(\sigma_i(\alpha), z)^{k_i} \right) f(\alpha z),$$

which we note coincides with the usual weight-$k$ slash operator when $r = 1$.

Let $\mathbb{Z}_K$ denote the ring of integers of $K$ (the integral closure of $\mathbb{Z}$ in $K$). The group $\mathrm{GL}_2^+(\mathbb{Z}_K)$ is an arithmetic lattice in the locally compact group $\mathrm{GL}_2^+(\mathbb{R})^r$, as is any $\Gamma \leq \mathrm{GL}_2^+(K)$ commensurable with $\mathrm{GL}_2^+(\mathbb{Z}_K)$; we note that unlike the case $r = 1$, for $r > 1$ lattices in $\mathrm{GL}_2^+(\mathbb{R})$ are automatically arithmetic.

**Definition 23.1.** Let $K$ be a totally real number field of degree $r > 1$. For $k \in \mathbb{Z}^r$ and lattices $\Gamma \leq \mathrm{GL}_2^+(K)$ commensurable with $\mathrm{GL}_2^+(\mathbb{Z}_K)$, a Hilbert modular form $f : \mathbf{H}^r \to \mathbb{C}$ of weight $k$ for $\Gamma$ is a holomorphic function that satisfies $f|_k\gamma = f$ for all $\gamma \in \Gamma$.

For $r = 1$ (in which case $K = \mathbb{Q}$) this coincides with the usual definition of a modular form without the requirement that $f$ be holomorphic at the cusps. Hilbert modular forms are automatically holomorphic at the cusps (this follows from what is known as the Koecher principle).

We call weights $k = (c, \ldots, c) \in \mathbb{Z}^r$ parallel weights (or scalar weights), and these are the most commonly studied. There is a naturally analog of the congruence subgroups $\Gamma_0(N)$ that is commonly used for Hilbert modular forms: for any nonzero ideal $\mathfrak{n}$ of $\mathbb{Z}_K$ we define

$$\Gamma_0(\mathfrak{n}) := \left\{ \left( \begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in \mathrm{GL}_2^+(\mathbb{Z}_K) : c \in \mathfrak{n} \right\},$$

and say that Hilbert modular forms for $\Gamma_0(\mathfrak{n})$ have level $\mathfrak{n}$. More generally, if $\mathfrak{c}$ is any fractional $\mathbb{Z}_F$-ideal, we can define

$$\Gamma_0(\mathfrak{c}, \mathfrak{n}) := \left\{ \left( \begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in \mathrm{GL}_2^+(K) : a, d \in \mathbb{Z}_K, c \in \mathfrak{c}\mathfrak{n}, b \in \mathfrak{c}^{-1} \right\}$$

and consider the space $M_k(\mathfrak{c}, \mathfrak{n})$ of Hilbert modular forms for $\Gamma_0(\mathfrak{c}, \mathfrak{n})$ of level $(\mathfrak{c}, \mathfrak{n})$. This is a finite dimensional $\mathbb{C}$-vector space, as is its subspace $S_k(\mathfrak{c}, \mathfrak{n})$ of cusp forms (Hilbert modular forms that vanish at the cusps). These spaces have dimension zero unless $k \in \mathbb{Z}^r_{\geq 0}$.

For suitable $k$ (including parallel even weight $k$) there are dimension formulas and traces formulas, as well as a nice theory of Hecke operators and newforms; we now have Hecke operators $T_\mathfrak{p}$ associated to each prime ideal $\mathfrak{p}$ of $\mathbb{Z}_K$ and newforms $f$ have Fourier coefficients $a_\mathfrak{p}$ equal to the eigenvalue of $T_\mathfrak{p}$ on $f$ (all of which are algebraic integers). We should note that the field $\mathbb{Q}(f)$ generated by these eigenvalues need not contain $K$; indeed, we may have $\mathbb{Q}(f) = \mathbb{Q}$.

For parallel weight-2 newforms $f$ there is a conjectural analog of the Eichler-Shimura construction that yields modular abelian varieties over $K$ that is known to hold in many cases (including all cases where $K$ has odd degree) due to a theorem of Hida [7]. When $f$ has rational Hecke eigenvalues (so $\mathbb{Q}(f) = \mathbb{Q}$), one obtains an elliptic curve $E/K$.

Recall that for $K = \mathbb{Q}$ the modularity theorem gives a 1-to-1 correspondence between isogeny classes of elliptic curves $E/\mathbb{Q}$ of conductor $N$ and newforms $f$ of weight 2 and level $N$ with $\mathbb{Q}(f) = \mathbb{Q}$ in which $E$ and $f$ have the same $L$-function. It is similarly conjectured that there is a 1-to-1 correspondence between elliptic curves over totally real fields $K$ of conductor $\mathfrak{n}$ and Hilbert modular newforms of parallel weight 2 and level $\mathfrak{n}$ with $\mathbb{Q}(f) = \mathbb{Q}$ in which $E$ and $f$ have the same $L$-function.

There is much recent progress toward this conjecture, which is now known to hold for all real quadratic fields [6] all totally real cubic fields [5], and in many other cases. Unlike $E/\mathbb{Q}$, it is possible for $E/K$ to have everywhere good reduction (meaning the its conductor is the trivial ideal $\mathbb{Z}_K$), corresponding to a Hilbert modular form for the full Hilbert modular group $\mathrm{GL}_2^+(\mathbb{Z}_K)$.

**Remark 23.2.** Some authors restrict to subgroups of $\mathrm{SL}_2(\mathbb{Z}_K) \leq \mathrm{GL}_2^+(\mathbb{Z}_K)$ when discussing Hilbert modular forms. This is a nontrivial restriction: the determinant of an element of $\mathrm{GL}_2^+(\mathbb{Z}_K)$ is a unit that is positive in every real embedding (a totally positive unit), but this condition is satisfied by the square of any unit, and there are infinitely many when $K = \mathbb{Q}$.

## 23.3 Bianchi modular forms

We began this course by considering the action of $\mathrm{GL}_2(\mathbb{R})$ on $\mathbf{H} \simeq \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$, restricting to $\mathrm{GL}_2^+(\mathbb{R})$ so that the action is isometric, and then to lattices in $\mathrm{SL}_2(\mathbb{R})$, which acts on $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$ via left multiplication. Here $\mathrm{SO}_2(\mathbb{R})$ is the maximal compact subgroup of $\mathrm{SL}_2(\mathbb{R})$, and we note that $\mathrm{SU}_2(\mathbb{C})$ is the maximal compact subgroup of $\mathrm{SL}_2(\mathbb{C})$. Replacing $\mathbb{R}$ with $\mathbb{C}$, we get an action of $\mathrm{SL}_2(\mathbb{C})$ on $\mathrm{SL}_2(\mathbb{C})/\mathrm{SU}_2(\mathbb{C})$.

What is the analog of the complex upper halfplane $\mathbf{H}$ is in this setting? Well $\mathbf{H} = \mathbf{H}_2$ is a 2-dimensional real manifold equipped with the hyperbolic metric induced by the Haar measure of $\mathrm{SL}_2(\mathbb{R})$, and is often called hyperbolic 2-space. Now $\mathrm{SL}_2(\mathbb{C})/\mathrm{SU}_2(\mathbb{C})$ is a 3-dimensional real manifold isomorphic to hyperbolic 3-space

$$\mathbf{H}_3 := \mathbb{C} \times \mathbb{R}_{>0} = \{(x_1 + x_2 i, y) : x_1, x_2 \in \mathbb{R}, y \in \mathbb{R}_{>0}\}.$$

The Haar measure on $\mathrm{SL}_2(\mathbb{C})$ induces the hyperbolic metric $\frac{dx_i dx_2 dy}{y^3}$ on $\mathbf{H}_3$ and $\mathrm{SL}_2(\mathbb{C})$ then acts on $\mathbf{H}_3$ via isometries.

For any number field $K$ of degree $r$, we have $r$ embeddings of $K$ into $\mathbb{C}$, each of which induces an embedding of $\mathrm{SL}_2(K)$ into $\mathrm{SL}_2(\mathbb{C})$, allowing us to view $\mathrm{SL}_2(K)$ as a subgroup of $\mathrm{SL}_2(\mathbb{C})^r$ that acts on $\mathbf{H}_3^r$ via isometries.

In the simplest case $r = 1$ and $K$ is an imaginary quadratic field, which we assume henceforth. A complication arises because $\mathbf{H}_3$ is not a complex manifold, so we cannot define modular forms as "holomorphic" functions on $\mathbf{H}_3$. One instead works with vector-valued real analytic functions $F\colon \mathbf{H}_3 \to \mathbb{C}^{k+1}$ that are invariant under the action of certain differential operators (Casimir operators associated to the Lie algebra of $\mathrm{SL}_2(\mathbb{C})$). The weight-$k$ slash operator associated to $\alpha \in \mathrm{SL}_2(\mathbb{C})$ is then defined by

$$(F|_k\alpha) := \mathrm{Sym}^k(J(\alpha, z)^{-1})F(\alpha z),$$

where $J(\alpha, z) := \begin{pmatrix} cx+d & -cy \\ \bar{c}y & \overline{cx+d} \end{pmatrix}$ for $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $z = (x, y) \in \mathbb{C} \times \mathbb{R}_{>0} = \mathbf{H}_3$ and $\mathrm{Sym}^k$ denotes the linear operator corresponding to the $k$th symmetric power of the standard representation of $\mathrm{SL}_2(\mathbb{C})$ on $\mathbb{C}^2$. For integers $k \geq 0$ and congruence subgroups $\Gamma \leq \mathrm{SL}_2(\mathbb{Z}_K)$ one can then consider the space of Bianchi modular forms of weight $k$ for $\Gamma$, which turns out to be a finite dimension $\mathbb{C}$-vector space that contains a subspace of cusp forms and Hecke operators $T_{\mathfrak{p}}$ associated to prime ideals of $\mathbb{Z}_K$. There is a notion of newforms that are simultaneous eigenforms for all the Hecke operators, and these form a basis for the new subspace of cusp forms.

As with Hilbert modular forms, one expects to be able to associate modular abelian varieties over $K$ to weight-2 Bianchi newforms $f$, and when $\mathbb{Q}(f) = \mathbb{Q}$, to obtain elliptic curves over $K$. But unlike the case of Hilbert modular forms, this is not quite true; there are weight-2 Bianchi newforms with $\mathbb{Q}(f) = \mathbb{Q}$ that do not correspond to any elliptic curve $E/K$, but instead correspond to a simple abelian surface whose endomorphism ring is an order in a quaternion algebra (something that cannot hold for an elliptic curve over a number field).

It is however true that every elliptic curve over an imaginary quadratic field $K$ is modular, meaning it has the same $L$-function as a weight-2 Bianchi newform $f$ with $\mathbb{Q}(f) = \mathbb{Q}$; this was recently proved in [12].

## 23.4 Siegel modular forms

For each positive integer $g$ we define the Siegel upper half space

$$\mathbf{H}_g := \{z \in \mathsf{M}_g(\mathbb{C}) : z^{\mathrm{T}} = z \text{ and } \mathrm{Im}(z) > 0\}$$

of symmetric $g \times g$ matrices with positive definite imaginary part. Let $J = \begin{pmatrix} 0 & -I_g \\ I_g & 0 \end{pmatrix} \in \mathsf{M}_{2g}(\mathbb{Z})$. The symplectic group

$$\mathrm{Sp}_{2g}(\mathbb{R}) := \{\gamma \in \mathrm{SL}_{2g}(\mathbb{R}) : \gamma^{\mathrm{T}} J \gamma = J\}$$

acts on $\mathbf{H}_g$ via "linear fractional transformations" $\tau \mapsto (a\tau+b)(c\tau+d)^{-1}$ for $\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{Sp}_{2g}(\mathbb{Z})$; note that for $g > 1$ this expression involves matrices and the order of multiplication matters. This extends to an isometric action of $\mathrm{GSp}_{2g}^+(\mathbb{R})$ on $\mathbf{H}_g$, where

$$\mathrm{GSp}_{2g}(\mathbb{R}) = \{\gamma \in \mathrm{SL}_{2g}(\mathbb{R}) : \gamma^{\mathrm{T}} J\gamma = \lambda J \text{ with } \lambda \in \mathbb{R}^\times\}$$

is the group of symplectic similitudes and $\mathrm{GSp}_{2g}^+(\mathbb{R})$ is the subgroup with positive determinant.

Each $\tau \in \mathbf{H}_g$ has an associated lattice $\Lambda_\tau := \tau\mathbb{Z}^g \oplus \mathbb{Z}^g \subseteq \mathbb{C}^g$ for which the torus $A_\tau = \mathbb{C}^g/\Lambda_\tau$ is an abelian variety $A/\mathbb{C}$ equipped with the principal polarization

$$E_\tau \colon (\tau x_1 + x_2, \tau y_1 + y_2) = (x_1^{\mathrm{T}}, x_2^{\mathrm{T}}) J \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

For $V \simeq \mathbb{C}^g$, a polarization on a torus $V/\Lambda$ is a Hermitian form $H \colon V \times V \to \mathbb{C}$ with $H(\Lambda,\Lambda) = \mathbb{Z}$. Every polarization induces a homomorphism $v \mapsto H(v,\cdot)$ to the dual torus $V^*/\Lambda^*$, where

$$V^* := \left\{ f : V \to \mathbb{C} : f(\alpha v) = \bar\alpha f(v) \text{ and } f(v_1+v_2) = f(v_1)+f(v_2)\right\}$$

and $\Lambda^* := \{f \in V^* : \operatorname{Im} f(\Lambda) \subseteq \mathbb{Z}\}$. This is defines an isogeny from the abelian variety $A = V/\Lambda$ and the dual abelian variety $A^* = V^*/\Lambda^*$; for principal polarizations this is an isomorphism.

Two elements $\tau, \tau' \in \mathbf{H}_g$ define isomorphic principally polarized abelian varieties $A_\tau \simeq A_{\tau'}$ if and only if they lie in the same $\mathrm{Sp}_{2g}(\mathbb{Z})$ orbit. We thus have a bijection between isomorphism classes of principally polarized abelian varieties $A/\mathbb{C}$ of dimension $g$ and the quotient $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbf{H}_g$, which is a variety $\mathscr{A}_g/\mathbb{Q}$ of dimension $g(g+1)/2$, the moduli space of principally polarized abelian varieties of dimension $g$. For $g=1$ the moduli space $\mathscr{A}_1$ corresponds to the modular curve $Y(1) = \Gamma(1)\backslash\mathbf{H}$; note that $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z}) = \mathrm{Sp}_2(\mathbb{Z})$ and $\mathbf{H} = \mathbf{H}_1$. As with $Y(1) = \mathscr{A}_1$, the moduli space $\mathscr{A}_g$ is not compact, but it can be compactified and embedded into projective space (analogous to the construction of $X(1)$ from $Y(1)$, but a bit more complicated).

We note that for $g \le 3$ the dimensions of $\mathscr{A}_g$ and $\mathscr{M}_g$ coincide. Here $\mathscr{M}_g$ is the moduli space of nice (smooth projective geometrically connected) curves; it has dimension 1 for $g=1$ and dimension $3g-3$ for $g>1$. Thus almost all principally polarized abelian varieties of dimension $g \le 3$ are Jacobians of nice curves of genus $g$, but this is not true for any $g > 3$.

We now define

$$\Gamma(N) := \{\gamma \in \mathrm{Sp}_{2g}(\mathbb{Z}) : \gamma \equiv 1 \bmod N\},$$

and call subgroups $\Gamma \le \mathrm{Sp}_{2g}(\mathbb{Z})$ that contain $\Gamma(N)$ for some $N$ congruence subgroups. For $g>1$ a holomorphic function $f : \mathbf{H}_g \to \mathbb{C}$ is a (classical) Siegel modular form of (parallel even) weight $k \in 2\mathbb{Z}$ for a congruence subgroup $\Gamma$ if it is invariant under the weight-$k$ slash operator

$$(f|_k\gamma)(\tau) := (\det J(\gamma,\tau)^{-k} f(\gamma\tau),$$

for all $\gamma = \left(\begin{smallmatrix} A & B\\ C & D\end{smallmatrix}\right) \in \Gamma$, where $J(\gamma,\tau) = C\tau + D$. This definition coincides with the definition of a modular form for a congruence subgroup when $g=1$ except the holomorphicity condition at the cusps is omitted because it is automatically satisfied (by Koecher's principle).

This definition can be generalized in many ways, including to non-parallel weights and vector-valued functions associated to representations of $\mathrm{GL}_g(\mathbb{C})$, but for the purpose of this introduction there is one particular generalization we want to consider, and to simplify matters we will restrict to $g = 2$. Rather than a congruence subgroup $\Gamma \le \mathrm{Sp}_{2g}(\mathbb{Z})$ we instead use a paramodular group

$$K(N) := \begin{pmatrix} \mathbb{Z} & N\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ \mathbb{Z} & \mathbb{Z} & \mathbb{Z} & \mathbb{Z}/N \\ \mathbb{Z} & N\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ N\mathbb{Z} & N\mathbb{Z} & N\mathbb{Z} & \mathbb{Z} \end{pmatrix} \cap \mathrm{Sp}_4(\mathbb{Q}).$$

Points on the quotient $K(N)\backslash\mathbf{H}_2$ now correspond to $(1,N)$-polarized abelian surfaces, but this distinction won't concern us here; the connection between modular forms and abelian varieties is via their $L$-functions and operates at the level of isogeny classes (recall that polarizations are isogenies); the $L$-function of an abelian variety is independent of its polarization.

Modular forms for paramodular groups are called paramodular forms. There is a nice theory of Hecke operators, eigenforms, and newforms for paramodular groups that is conjectured to have all the properties one would like; in particular the new subspace of cuspidal paramodular forms of weight $k$ and level $N$ has a basis consisting of newforms that are normalized eigenforms for all of the Hecke operators $T_p$ for $p \nmid N$. The theory of newforms is more complicated than in the classical case (there are three different ways in which a paramodular form of level $M|N$ can arise as a paramodular form of level $N$, as well as Gritsenko lifts of Jacobi forms that need to be excluded), but this theory is now well understood [9].

## 23.5 The paramodular conjecture

For abelian surfaces there is a conjectured analog of the modularity of elliptic curves known as the paramodular conejecture, due to Brumer and Kramer [2].

**Conjecture 23.3** (Paramodular conjecture). *Let $A/\mathbb{Q}$ be an abelian surface of conductor $N$ with* $\mathrm{End}(A) = \mathbb{Z}$. *Then $L(A,s) = L(f,s)$ for some weight-2 paramodular newform of level $N$ with rational Hecke eigenvalues. Conversely, for every weight-2 paramodular newform of level $N$ with rational Hecke eigenvalues we have $L(f,s) = L(A,s)$ for an abelian variety $A/\mathbb{Q}$ that is either an abelian surface of conductor $N$ with* $\mathrm{End}(A) = \mathbb{Z}$, *or an abelian fourfold of conductor $N^2$ with* $\mathrm{End}(A)$ *an order in a non-split quaternion algebra.*

Both directions of this conjecture are open and highly non-trivial; in particular, there is no analog of the Eichler–Shimura construction that allows one to associate an abelian variety to a weight-2 paramodular newform. In fact there has been more progress toward proving the first part of the conjecture (the modularity of abelian surfaces) than the second (which for $g = 1$ was known long before the modularity of elliptic curves, due to the Eichler–Shimura construction). The necessity of allowing for the case of QM abelian fourfolds noted by Frank Calegari highlights the difficulty and is analogous to the situation with weight-2 Bianchi newforms with rational Hecke eigenvalues that correspond to QM abelian surfaces rather than elliptic curves.

The paramodular conjecture has now been verified for many abelian surfaces over $\mathbb{Q}$ with small conductor $N$, including for $N = 277$, which is the first level where one finds paramodular newforms with rational Hecke eigenvalues and believed to be the smallest conductor of an abelian surface $A/\mathbb{Q}$ with $\mathrm{End}(A) = \mathbb{Z}$. There are two abelian surfaces $A$ of conductor 277, both of which are isogenous to the Jacobian of

$$y^2 + (x^3 + x^2 + x + 1)y = -x^2 - x,$$

which is the genus 2 curve with label `277.a.277.1` in the LMFDB. Infinite families of abelian surfaces over $\mathbb{Q}$ for which the paramodular conjectures holds are now known [3], building on the potential modular results of Boxer, Calegari, Gee, and Pilloni [1]. Very recently the same authors have announced a proof that the paramodular conjecture holds for a positive proportion of abelian surfaces over $\mathbb{Q}$ (but the new result uses different techniques).

One can define paramodular groups $K(N) \le \mathrm{Sp}_{2g}(\mathbb{Q})$ for even $g > 2$, but in general the quotient $K(N)\backslash\mathbf{H}_g$ parameterizes abelian varieties whose endomorphism ring is an order in a totally real field of degree $g/2$. For $g > 2$ these abelian varieties are not fully general and do not account for a positive proportion of abelian varieties of dimension $g$.

### 23.6 Shimura varieties

We have seen that modular curves that parameterize elliptic curves equipped with a level structure, Siegel modular varieties that parameterize abelian varieties equipped with a polarization, and paramodular varieties that parameterize abelian varieties whose endomorphism rings are orders in a totally real field. These are all special cases of Shimura varieties, and in particular, PEL type Shimura varieties, that parameterize abelian varieties with a particular polarization (P), endomorphism ring (E), and level structure (L).

We won't attempt to give a formal definition of Shimura varieties here, but they generally start with a complex manifold $M$ (such as $M = \mathbf{H}$) endowed with a Hermitian metric (such as the hyperbolic metric) that is induced by the Haar measure of a locally compact topological group (such as $\mathrm{SL}_2(\mathbb{R})$) acting via isometries. Moreover, this topological group corresponds to the real points $G(\mathbb{R})$ of a $\mathbb{Q}$-algebraic group $G$ (such as $G = \mathrm{SL}_2$). This allows us to then consider arithmetic lattices $\Gamma \leq G(\mathbb{R})$ (such as $\mathrm{SL}_2(\mathbb{Z})$) and Shimura varieties $X = \Gamma\backslash M$ formed by the quotient that are defined over a finite extension of $\mathbb{Q}$ and which parameterize abelian varieties equipped with certain structures.

## References


[1] George Boxer, Frank Calegari, Toby Gee, and Vincent Pilloni, *Abelian surfaces over totally real fields are potentially modular*, Publ. Math. IHES **134** (2021), 153–501.

[2] Armand Brumer and Ken Kramer, *Paramodular abelian varieties of odd conductor*, Trans. Amer. Math. Soc. **366** (2013), 2463–2516 (but see 2018 correction in arXiv:1004.4699).

[3] Frank Calegari, Shiva Chidambaram, and David Roberts, *Abelian surfaces with fixed 3-torsion*, Proceedings of the Fourteenth Algorithmic Number Theory Symposium (ANTS XIV), Open Book Series **4**, Mathematical Sciences Publishers, 2020.

[4] Henri Cohen and Fredrik Strömberg, *Modular forms: A classical approach*, Graduate Studies in Mathematics **179**, American Mathematical Society, 2017.

[5] Maarten Derickx, Filip Najman, and Samir Siksek, *Elliptic curves over totally real cubic fields are modular*, Algebra Number Theory **14** (2020), 1791–1800.

[6] Nuno Freitas, Bao V. Le Hung, and Samir Siksek, *Elliptic curves over real quadratic fields are modular*, Invent. Math. **201** (205), 159–206.

[7] Haruzo Hida, Hilbert modular forms and Iwasawa theory, Oxford University Press, 2006.

[8] Winfried Kohnen, *Modular forms of half–integral weight on* $\Gamma_0(4)$, Math. Ann. **248** (1980), 249–266.

[9] Brooke Roberts and Ralf Schmidt, *Local newforms for* $\mathrm{GSp}(4)$, Lecture Notes in Mathematics **1918**, Springer, 2007.

[10] Goro Shimura, *On modular forms of half integral weight*, Ann. of Math. **97** (1973), 440-481.

[11] Nils–Peter Skoruppa and Don Zagier, *Jacobi forms and a certain space of modular forms*, Invent. Math. **94** (1988), 113–146.

[12] Patrick B. Allen, Frank Calegari, Ana Caraiani, Toby Gee, David Helm, Bao V. Le Hung, James Newton, Peter Scholze, Richard Taylor, and Jack A. Thorne, *Potential automorphy over CM fields*, Ann. of Math. **197** (2023), 897–1113.