MODULAR SYMBOLS

PAUL E. GUNNELLS

Abstract. Expanded notes from lectures given by Paul E. Gunnells at the 2014/2017 UNCG Summer Schools in Computational Number Theory.

http://www.uncg.edu/mat/numbertheory/summerschool/2014.html
http://www.uncg.edu/mat/numbertheory/summerschool/2017.html

Lecture 1. Introduction to modular forms

1.1. The goal of these lectures is to explain how to compute effectively with classical holomorphic modular forms. The main approach is the *modular symbol method*, due to work of Birch, Manin, Mazur, Merel, and Cremona. In this first lecture, we give an overview of classical holomorphic modular forms of weights $k \geq 2$ and give some of their applications. For further reading, you might consult [8, 12–14, 16–18, 20].

1.2. Definitions and notation. Let

$$\mathfrak{H} = \text{upper halfplane} = \{z \in \mathbb{C} \colon \operatorname{Im}(z) > 0\}$$
$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z},\ \det(\gamma) = 1 \right\}.$$

Then $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathfrak{H}$ by fractional-linear transformations:

$$z \longmapsto \frac{az+b}{cz+d}.$$

For each *weight* $k \geq 2$, we get an action on functions $f \colon \mathfrak{H} \to \mathbb{C}$ called the *slash operator*:

$$(f\big|_k\gamma)(z) = f\left(\frac{az+b}{cz+d}\right)(cz+d)^{-k}, \quad \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

1.3. Definition. A function $f \colon \mathfrak{H} \to \mathbb{C}$ is a *modular form of weight k* if

(a) f is holomorphic
(b) $(f\big|_k\gamma) = f$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$
(c) f is holomorphic "at infinity", which means as $\operatorname{Im}(z) \to \infty$, $|f(z)|$ is majorized by a polynomial in $\max\{1, \operatorname{Im}(z)^{-1}\}$

In particular, (c) implies that $|f(z)|$ cannot grow too rapidly as $\operatorname{Im}(z) \to \infty$. Let M_k denote the $\mathbb{C}$-vector space of weight k modular forms.

Date: May 22, 2017.
2010 *Mathematics Subject Classification.* Primary 11F75; Secondary 11F11, 11F67.
Key words and phrases. Modular symbols, modular forms, Hecke operators.
The author thanks the organizers for the invitation to speak. He also thanks Dan Yasaki for providing a preliminary LaTeXed version of his notes.

We get the notion of a *cuspform* by imposing stronger growth conditions, namely f decays very rapidly as $\mathrm{Im}(z) \to \infty$. More precisely, replace condition (c) by (c)′: $|f|$ is majorized by $\mathrm{Im}(z)^{-k/2}$ as $\mathrm{Im}(z) \to \infty$. Let $S_k \subset M_k$ be the subspace of cuspforms.

1.4. Fact. *The space of weight k modular forms M_k is finite-dimensional.*

1.5. Fourier expansion of f. Let $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ Observe $f\big|_k\gamma = f$ means f is invariant under $z \mapsto z+1$. Thus f has a Fourier expansion

$$f(z) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n z}, \quad a_n \in \mathbb{C}.$$

Usually we put $q = e^{2\pi i n z}$ and write this as a *q-expansion* of f:

$$f(q) = \sum_{n\in\mathbb{Z}} a_n q^n.$$

One can show that the growth conditions (c) and (c)′ are equivalent to

$$\begin{aligned} a_n = 0 \text{ for all } n < 0 \quad &\iff \quad f \in M_k, \\ a_n = 0 \text{ for all } n \leq 0 \quad &\iff \quad f \in S_k. \end{aligned}$$

Under the change of coordinates $z \mapsto q = e^{2\pi i z}$, the upper halfplane maps to the unit disk $\{q \in \mathbb{C} : |q| < 1\}$. The point at $i\infty$ gets taken to the origin in the disk. In these new coordinates, saying $f \in M_k$ means f is bounded as $q \to 0$ in the disk, and can thus be extended to a function defined on the disk. Similarly, saying $f \in S_k$ means that f extends to a function vanishing at 0 on the disk.

1.6. Why do we study modular forms? As we shall see, sometimes we have a sequence

$$\{\alpha_n \colon n \in \mathbb{Z}_{\geq 0}\} \subset \mathbb{C}$$

arising naturally. For instance, we might have $\alpha_n \in \mathbb{Z}$, and they may count something.

Combinatoricists use a generating function $\sum \alpha_n x^n$ to organize these numbers. Number theorists, on the other hand, replace x by q and make a q-series. Replacing x by q is trivial, but nevertheless suggestive. One can ask: Is the resulting series the q-expansion of a weight k modular form?

If this is true, then $f \in M_k$, and the latter is a vector space of rather small dimension (roughly $k/12$). We can then take a basis of M_k and can express the function f in terms of this basis; this typically already leads to nontrivial information about the coefficients of f. Another typical phenomenon is that we may have other sequences $g_1, g_2, \ldots$ giving rise to modular forms in M_k coming from quite different settings. Since M_k has small dimension, as soon as we have enough modular forms they can't all be linearly independent, and thus we obtain nontrivial relations among f and the g_i, relations that are usually not at all obvious from the sources of these series.

This is best understood through examples, as we now illustrate. This also gives us the chance to introduce some key players in the theory.

1.7. Example (Eisenstein series)**.** The simplest way to try to make a modular form is by averaging: we can average over $\mathrm{SL}_2(\mathbb{Z})$ to force invariance under the slash action. Put $k \geq 4$,

and define

$$E_k(z) \coloneqq \frac{(k-1)!}{2(2\pi i)^k} \sum_{m,n\in\mathbb{Z}}{}' (mz+n)^{-k}$$

(the normalizing factor is used for convenience). This sum is absolutely convergent if $k \geq 4$, and we get a modular form $E_k \in M_k$, called the *holomorphic weight k Eisenstein series*. Note E_k vanishes identically for odd k. When $k = 2$ the series doesn't converge absolutely, but there is a standard way to sum the series conditionally (*Hecke's trick*). In this case the result is not a modular form, but it's almost one.

The Eisenstein series E_k has Fourier expansion

$$E_k(q) = \frac{1}{2}\zeta(1-k) + \sum_{n\geq 1} \sigma_{k-1}(n) q^n,$$

where σ_r is the r^{th} power divisor sum

$$\sigma_r(n) \coloneqq \sum_{d|n} d^r.$$

Note

$$\frac{1}{2}\zeta(1-k) = -\frac{B_k}{2k},$$

where B_k is the k^{th} Bernoulli number. The first few q-expansions are

$$E_4 = \frac{1}{240} + q + 9q^2 + 28q^3 + \dots, \tag{1}$$

$$E_6 = -\frac{1}{504} + q + 33q^2 + 244q^3 + \dots, \tag{2}$$

$$E_8 = \frac{1}{480} + q + 129q^2 + 2188q^3 + \dots. \tag{3}$$

Now the direct sum of all the spaces of modular forms

$$M_* = \bigoplus_k M_k$$

forms a graded ring, where the weight gives the grading: if f has weight k and g has weight l, then fg is a modular form of weight $k + l$. One can prove

$$M_* \simeq \mathbb{C}[E_4, E_6]. \tag{4}$$

Thus any weight k modular form can be written as a (weighted) homogeneous polynomial in the Eisenstein series E_4, E_6, which allows one to easily compute the dimension of M_k. Immediately we can get a nontrivial identity: one can check $\dim(M_4) = \dim(M_8) = 1$, which means E_4^2 must be a multiple of E_8. Checking constant terms of Fourier expansions, we see

$$120E_4^2 = E_8.$$

Now look at the Fourier coefficients. We get

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{m=1}^{n-1} \sigma_3(m)\sigma_3(n-m), \tag{5}$$

which is not obvious.

There is a Hermitian inner product defined on (most of) M_k, called the *Petersson product*:

$$\langle f, g\rangle = \int_D y^k f\bar{g}\, dA$$

where D is a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ in $\mathfrak{H}$ and $dA = dx\, dy/y^2$ is hyperbolic measure. What do we mean by *most of*? Unfortunately we can't compute the inner product of two Eisenstein series (the integral doesn't converge), but we can compute the inner product of an Eisenstein series and any cuspform, or of two cuspforms. Using the Petersson product it's possible to prove

$$M_k \simeq \mathbb{C}E_k \oplus S_k$$

is an orthogonal decomposition. Thus given any modular form f, we can always subtract off a suitable multiple of an Eisenstein series to obtain a cuspform.

1.8. Exercise. Verify some cases of the divisor sum identity (5). Can you prove it without using modular forms?

1.9. Example (Delta function)**.** The first weight with $S_k \neq 0$ is $k = 12$: M_{12} is spanned by E_4^3, E_6^2, and these are not equal. The difference

$$\Delta(q) := 8000E_4^3 - 147E_6^2 = q - 24q^2 + 252q^3 + \dots \tag{6}$$

has no constant term and is thus a cuspform.[1] The coefficients of the q-expansion give the values of Ramanujan's τ-function:

$$\Delta(q) = \sum \tau(n)q^n.$$

Thus our expression in terms of Eisenstein series gives a way to compute $\tau(n)$ using sums of powers of divisors of n. But Δ has even more structure. One can prove that Δ satisfies an infinite product formula

$$\Delta(q) = q\prod_{n\geq 1}(1-q^n)^{24};$$

most modular forms, of course, have no such product structure. This shows that $\Delta = \eta(q)^{24}$, where η is Dedekind's eta-function.

1.10. Example (Theta series of even, unimodular lattices)**.** Now we have an arithmetic application. Let L be an even, unimodular lattice in $\mathbb{R}^n$. This means

(1) $L \subset \mathbb{R}^n$ is a discrete, cocompact subgroup,
(2) the inner product in $\mathbb{R}^n$ is $\mathbb{Z}$-valued when restricted to L,
(3) L has a $\mathbb{Z}$-basis $\{v_1, \dots, v_n\}$ such that the Gram matrix $(v_i \cdot v_j)$ has determinant 1 (*unimodular*), and
(4) $v \cdot v \in 2\mathbb{Z}$ for all $v \in L$ (*even*).

It is known that even, unimodular lattices exist in $\mathbb{R}^n$ if and only if $n \equiv 0 \bmod 8$. There are finitely many up to rotation. In general, the number of such lattices is unknown except for small values of n (cf. Table 1).

Define

$$r_L(m) = \#\Big\{x \in L\colon \frac{x\cdot x}{2} = m\Big\},$$

[1]This is correct although it looks quite ugly. Another typical normalization of the Eisenstein series puts the constant terms to be 1, i.e. $\tilde{E}_4 = 240E_4$, $\tilde{E}_6 = -504E_6$, ... With this convention, the expression (6) becomes $\Delta = (\tilde{E}_4^3 - \tilde{E}_6^2)/1728$, which is much more attractive.

TABLE 1. Even unimodular lattices in $\mathbb{R}^n$.

n	$\#L$	Name
8	1	The root lattice E_8
16	2	$E_8 \oplus E_8$ and the root lattice D_{16}
24	24	The 24 Niemeier lattices (includes the Leech lattice)
32	over 1000000000	

and form the q-expansion

$$f_L(q) = \sum_{m \geq 0} r_L(m) q^m.$$

Then one can prove the following:

1.11. Fact. *Let $L \subset \mathbb{R}^n$ be an even unimodular lattice. Then $f_L(q)$ is a modular form of weight $n/2$.*

Here are two applications of this fact. First, consider the root lattice of type E_8 (don't confuse this notation with the Eisenstein series!). By definition this is the set of points $(x_1, \ldots, x_8) \in \mathbb{R}^8$ such that

- all the coordinates are integers or all the coordinates are half-integers (a mixture of integers and half-integers is not allowed), and
- the sum of the eight coordinates is an even integer.

This is an even unimodular lattice. Thus $f_{E_8}(q) \in M_4$, which we know is spanned by the Eisenstein series E_4. Comparing constant terms, we find $f_{E_8} = 240E_4$. This implies

$$r_{E_8}(m) = 240\sigma_3(m). \tag{7}$$

1.12. Exercise. Check (7) for as many values of m as you can using a computer. Can you *prove* (7) without using modular forms?

Next consider $n = 16$. There are two even unimodular lattices in this dimension, $L_1 = E_8 \oplus E_8$ and a new one L_2, which is the root lattice D_{16}. Now $f_{L_1}(q)$ and $f_{L_2}(q)$ are both weight 8 modular forms with constant coefficient 1. Since the space of weight 8 modular forms is one-dimensional and is spanned by the Eisenstein series $E_8(q)$ (don't mix this up with the root lattice E_8!), both these modular forms must be equal. (In fact by (3) they equal $480E_8(q)$).

Thus these two lattices have the property the number of vectors of a given length is the same for both. This is relevant to a famous problem in differential geometry, which asks *Can you hear the shape of a manifold?* Precisely, the question means *Does the spectrum of the Laplacian on a Riemannian manifold uniquely determine it, up to isometry?*. The answer, as observed by Milnor, is no. The lattices determine two 16-dimensional flat tori $T_1 = \mathbb{R}^{16}/L_1$ and $T_2 = \mathbb{R}^{16}/L_2$. If $\Lambda \subset \mathbb{R}^n$ is a lattice with associated flat torus $T = \mathbb{R}^n/\Lambda$, then the eigenfunctions for the Laplacian have the form

$$f_{\lambda^*}(x) := e^{2\pi\sqrt{-1}(\lambda^* \cdot x)},$$

where λ^* is any point in the dual of Λ (by definition the dual of Λ is all λ^* such that $\lambda^* \cdot \lambda \in \mathbb{Z}$ for all $\lambda \in \Lambda$). Furthermore, the eigenvalue of $f_{\lambda^*}(x)$ is $4\pi^2|\lambda^*|^2$. The lattices L_i are self-dual, so the sequence of Laplacian eigenvalues is essentially what's encoded by the q-expansions

$f_{L_i}(q)$. Thus $f_{L_1}(q) = f_{L_2}(q)$ implies that T_1 and T_2 are isospectral. On the other hand, T_1 and T_2 are non-isometric (there is no isometry of $\mathbb{R}^{16}$ taking L_1 into L_2).

1.13. Exercise. Find how to express the theta series for the Leech lattice in terms of modular forms, and for some of the Niemeier lattices. You will only need the Eisenstein series E_{12} and the cuspform Δ to do this. (The answer is known and can be found online, say at the OEIS [19]. But don't cheat! This way you can learn how to compute with these lattices.)

1.14. Level structure. For arithmetic applications, one needs modular forms with level. To define these we need congruence subgroups.

1.15. Definition. Fix $N \in \mathbb{Z}_{>0}$. The *principal congruence subgroup* $\Gamma(N)$ is defined by

$$\Gamma(N) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \colon \gamma \equiv I \bmod N\}.$$

A subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is called a *congruence subgroup* if Γ contains $\Gamma(N)$ for some N. The minimal such N is called the *level*.

The principal congruence subgroup $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$; indeed, one can show that $\Gamma(N)$ fits into an exact sequence

$$1 \longrightarrow \Gamma(N) \longrightarrow \mathrm{SL}_2(\mathbb{Z}) \longrightarrow \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) \longrightarrow 1$$

(the tricky part is the surjectivity onto $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$). Thus every congruence subgroup also has finite index. The converse, however, is not true: not every finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup. The groups $\Gamma(N)$ are torsion-free when $N \geq 3$.

The most important congruence subgroups besides $\Gamma(N)$ are the *Hecke congruence subgroups*:

$$\Gamma_0(N) = \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \colon \gamma \equiv \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \bmod N\right\},$$
$$\Gamma_1(N) = \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \colon \gamma \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \bmod N\right\}.$$

1.16. Definition. Suppose Γ is a congruence subgroup. We say $f \colon \mathfrak{H} \to \mathbb{C}$ is a *weight k modular form on Γ* if

(a) f is holomorphic,
(b) $f\big|_k\gamma = f$ for all $\gamma \in \Gamma$, and
(c) the previous growth condition now holds for $f\big|_k\gamma$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Let $M_k(\Gamma)$ denote the $\mathbb{C}$-vector space of weight k modular forms on Γ. If $\Gamma = \Gamma_0(N)$, we usually just write $M_k(N)$ etc.

The last condition is a generalization of holomorphic at ∞. It is more complicated because there is more than one way to go to infinity, and by requiring the growth condition to hold for $f\big|_k\gamma$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, we are holomorphic at infinity for all possible cases. We will say more about this about this later.

Let $M_k(\Gamma)$ be the space of modular forms on Γ. As before this is a finite-dimensional complex vector space, and there is a distinguished subspace $S_k(\Gamma)$ of cuspforms. Just like the case of full level, f is a cuspform if $f\big|_k\gamma$ decays rapidly to zero as $\operatorname{Im} z$ goes to infinity, where γ varies over all of $\mathrm{SL}_2(\mathbb{Z})$. The Petersson product makes sense (just use the same definition

but integrate over a fundamental domain for Γ), and the complement of the cuspforms in $M_k(\Gamma)$ is the subspace of Eisenstein series $\mathrm{Eis}_k(\Gamma)$. We have

$$M_k(\Gamma) = S_k(\Gamma) \oplus \mathrm{Eis}_k(\Gamma),$$

an orthogonal decomposition with respect to the Petersson inner product.

So far everything looks the same, but there are some differences. First of all, in general there are Eisenstein series of weight 2. Secone, unlike the case of full level, it is not true in general that $M_*(\Gamma)$ is a polynomial ring over a fixed set of Eisenstein series. In fact, the Eisenstein series usually aren't sufficient to generate $M_*(\Gamma)$ as a graded ring; some cuspforms must be taken too. And once one has a set of generators, there are usually nontrivial relations among them. However, it is still true that the ring of modular forms is always finitely presented, just like the case of full level.

There is a close connection between the groups $\Gamma_0(N)$ and $\Gamma_1(N)$, and in fact one can investigate modular forms on $\Gamma_1(N)$ by enlarging the scope of objects considered on $\Gamma_0(N)$. Let $\chi\colon \mathbb{Z} \to \mathbb{C}$ be a Dirichlet character of level N. This means $\chi(n+N) = \chi(n)$; $\chi(n) = 0$ if and only if $(n, N) > 1$; and $\chi(mn) = \chi(m)\chi(n)$. Thus χ induces a map

$$\chi\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$$

that is nonzero exactly on $(\mathbb{Z}/N\mathbb{Z})^\times$, and when nonzero takes values in the roots of unity. We have

$$\Gamma_0(N)/\Gamma_1(N) \simeq (\mathbb{Z}/N\mathbb{Z})^\times \tag{8}$$

by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \longmapsto d \bmod N.$$

Hence we can understand modularity with respect to $\Gamma_1(N)$ by incorporating a character χ into the action of $\Gamma_0(N)$. More precisely, for $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(N)$, put

$$(f\big|_{k,\chi}\gamma)(z) = \overline{\chi(d)}(cz+d)^{-k} f(\gamma z).$$

We can define the space $M_k(N,\chi)$ by replacing the condition $f\big|_k\gamma = f$ with $f\big|_{k,\chi} = f$. This leads to the vector space $M_k(N,\chi)$, which is called the space of weight k modular forms of level N and *nebentype* χ. By (8) we have

$$M_k(\Gamma_1(N)) \simeq \bigoplus_\chi M_k(N,\chi).$$

1.17. Hecke operators. The space of modular forms M_k admits a huge collection of commuting linear operators, the *Hecke operators*. Moreover, they are Hermitian with respect to the natural inner product on M_k. Thus we can look for simultaneous eigenclasses. It is these eigenclasses and their eigenvalues that reveal the hidden arithmetic information in the modular forms. They are crucial for arithmetic applications, and motivate the main goal of our lectures: how to effectively compute spaces of modular forms and the Hecke action on them.

For now, we just define the Hecke operators; later we will see how to compute them. Let n be a fixed positive integer. Define a subset $\mathcal{X}_n \subset M_2(\mathbb{Z})$ by

$$\mathcal{X}_n = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : a \geq 1,\ ad = n,\ 0 \leq b < d \right\}.$$

Extend the slash action on functions $f\colon \mathfrak{H} \to \mathbb{C}$ from matrices in $\mathrm{SL}_2(\mathbb{Z})$ to $\mathrm{GL}_2(\mathbb{Q})$ via

$$(f\big|_k\gamma)(z) = (\det(\gamma))^{k-1}(cz+d)^{-k} f(\gamma z).$$

Now we can apply the elements of $\mathcal{X}_n$ to modular forms. Suppose f is a weight k modular form of full level. Then the action of the Hecke operator T_n on f is defined by

$$(T_n f)(z) \coloneqq \sum_{\gamma \in \mathcal{X}_n} (f\big|_k\gamma)(z).$$

Note that to be pedantic, we really should write fT_n (i.e. the Hecke operator should act on modular forms on the right, since the matrices in $\mathcal{X}_n$ are acting by the slash operator, which is a right action). But as we said, one knows that the Hecke operators commute with each other. Thus it doesn't matter whether we write the operators acting on the right or left.

Why is this an action, and why are these interesting operators? Certainly, if you've never seen it before, it's not clear why this is an action. The main thing to check is that if f is modular, so is $T_n f$. The point is that the set $\mathcal{X}_n$ is in bijection with a certain subset of lattices. Namely, we have

$$\mathcal{X}_n \Longleftrightarrow \{L \subset \mathbb{Z}^2 \colon [\mathbb{Z}^2 : L] = n\},$$

i.e., $\mathcal{X}_n$ is in bijection the set of sublattices of $\mathbb{Z}^2$ of index n. The bijection itself is easy to describe: any such lattice has a basis of the form $ae_1 + be_2, de_2$, where $\mathbb{Z}^2 = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2$.

1.18. Exercise. Draw examples of these sublattices for some small values of n.

Now an alternative perspective on modular forms describes them as certain functions on lattices in $\mathbb{C}$: to any weight k modular form f we can attach a function $F = F_f$, where

$$F\colon \{\text{lattices in } \mathbb{C}\} \longrightarrow \mathbb{C}$$

satisfies the homogenity condition

$$F(\lambda L) = \lambda^{-k} F(L), \quad \text{for all } \lambda \in \mathbb{C}^\times.$$

For more discussion, see [17, VII.2.2].[2] So from this perspective, the effect of the Hecke operator T_n is to define a new function $T_n F$ that averages F over the index n sublattices of its input [17, VII.5.1]. This is certainly a very natural operation on functions defined on lattices, although why this reveals the arithmetic information hidden in M_k is less obvious.

The operators satisfy

$$T_n T_m = T_{nm} \quad \text{if } (n,m) = 1, \text{ and} \tag{9}$$

$$T_{p^n} = T_{p^{n-1}} T_p - p^{k-1} T_{p^{n-2}} \quad \text{for } p \text{ prime}. \tag{10}$$

[2] You should read this book anyway, if you're interested in number theory. It's one of the greats.

These identities follow from the description of $\mathcal{X}_n$ in terms of sublattices. We can compute the operators directly on q-expansions. If $f(q) = \sum a_n q^n$, then

$$(T_n f)(q) = \sum_{m \in \mathbb{Z}} \Bigg(\sum_{\substack{d \geq 1 \\ d | (m,n)}} d^{k-1} a_{mn/d^2} \Bigg) q^m. \tag{11}$$

In particular, for p prime (11) becomes

$$(T_p f)(q) = \sum_{m \geq 0} (a_{mp} + p^{k-1} a_{m/p}) q^m. \tag{12}$$

These formulas give an algorithm to compute Hecke operators, although not a very good one: simply compute q-expansions of a basis of M_k as far as one needs, using (4) and the q-expansions of the Eisenstein series, then apply (11) and find the action of T_p in terms of the basis. (What makes this algorithm not great is that computing the coefficient of q^m in $T_p f$ needs the coefficient a_{mp}.) In any case, we see that if f is an eigenform, and if we normalize so that $a_1 = 1$, then the Fourier coefficient a_n is the eigenvalue of T_n, and from (9)–(10) the Fourier coefficients satisfy

$$a_n a_m = a_{nm} \quad \text{if } (n, m) = 1, \text{ and} \tag{13}$$

$$a_{p^n} = a_{p^{n-1}} a_p - p^{k-1} a_{p^{n-2}} \quad \text{for } p \text{ prime}. \tag{14}$$

We can also define Hecke operators for modular forms with level structure N, but we must be careful if $(n, N) \neq 1$. For T_p, if $p \mid N$ then we only use the elements $\begin{bmatrix} 1 & a \\ 0 & p \end{bmatrix} \in \mathcal{X}_p$, in other words we omit $\begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix}$. The resulting operator is usually denoted U_p.

Now that we have level structure and Hecke operators, we can give an example to show how the operators reveal the arithmetic information hidden in the modular forms. Let $E/\mathbb{Q}$ be an elliptic curve. Concretely, we can consider E to be a nonsingular plane curve defined by the equation

$$y^2 + a_1 xy + a_3 = x^3 + a_2 x^2 + a_4 + a_6, \quad a_i \in \mathbb{Z}, \tag{15}$$

although in doing so we are missing one point (the point at infinity, which serves as the identity element for the group law on E). The equation (15) can be reduced modulo any prime p since the a_i are integral, and one knows that for almost all p the resulting curve $E(\mathbb{F}_p)$ is nonsingular. Using the finitely many p for which $E(\mathbb{F}_p)$ is singular one can define the *conductor* of E; it is an integer N_E such that $E/\mathbb{F}_p$ is nonsingular if and only if $p \nmid N_E$. In general N_E is not squarefree, but there is an explicit algorithm to determine it.

Now we want to attach a Dirichlet series to E. Define a sequence $\{a_n\} \subset \mathbb{Z}$ as follows. If $p \nmid N_E$, put $a_p(E) = p + 1 - \#E(\mathbb{F}_p)$ (this enumeration of points on E mod p also includes the point at infinity). If $p \mid N_E$, then one puts $a_p(E) \in \{0, \pm 1\}$ depending on the singularity E acquires mod p. There are two possibilities: E has a cusp or a node mod p. If E has a cusp mod p we put $a_p(E) = 0$. If $E(\mathbb{F}_p)$ has a node mod p, then there are two further possibilities, depending on the structure of the tangents to the node. We put $a_p(E) = 1$ (respectively, -1) if the slopes of the two tangents to the node lie in $\mathbb{F}_p$ (respectively, lie in $\mathbb{F}_{p^2} \smallsetminus \mathbb{F}_p$.)

Thus we have a collection of integers $a_p(E)$ indexed by the primes p. We extend the definition from the $a_p(E)$ to $a_n(E)$ with n composite via an *Euler product*:

$$\sum a_n(E)n^{-s} := \prod_{p\nmid N}(1 - a_p(E)p^{-s} + p^{1-2s})^{-1} \cdot \prod_{p|N}(1 - a_p(E)p^{-s}). \tag{16}$$

In particular, the numbers $a_n(E)$ are computed by expanding the factors on the right of (16) into geometric series, just as one does to go between the the Euler product for the Riemann ζ-function and its Dirichlet series:

$$\zeta(s) = \sum n^{-s} = \prod(1 - p^{-s})^{-1}.$$

The Dirichlet series

$$L(E, s) = \sum_{n>0} a_n(E)/n^s$$

is called the *L-function of the elliptic curve* E. For instance, if E is defined by the equation $y^2 + y = x^3 - x$, then $N_E = 37$. A small table of the $a_p(E)$ is given in Table 2. We have

$$\begin{aligned} L(E, s) = 1 - 2/2^s 2 - 3/3^s + 2/4^s - 2/5^s + 6/6^s - 1/7^s + 6/9^s + 4/10^s - 5/11^s \\ - 6/12^s - 2/13^s + 2/14^s + \cdots + 32/1024^s + \cdots - 1024/1048576^s + \cdots \end{aligned} \tag{17}$$

The coefficients a_2, a_3, and a_5 are determined by counting points mod p, whereas a_1 and a_4 are determined using the Euler product (16).

p	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47
$a_p(E)$	-2	-3	-2	-1	-5	-2	0	0	2	6	-4	-1	-9	2	-9

TABLE 2. Some $a_p(E)$ for the curve $E : y^2 + y = x^3 - x$.

1.19. Exercise. Verify Table 2. Do at least some by hand, including $p = 37$. Verify (17).

Now we have the following amazing theorem. I personally consider myself extremely lucky to have been mathematically aware when this theorem was proved:

1.20. Theorem. *Let $E/\mathbb{Q}$ be an elliptic curve of conductor N_E and let $f_E = \sum a_n(E)q^n$, where the $a_n(E)$ are defined as above. Then f_E is the q-expansion of a Hecke eigenform in $S_2(N_E)$.*

Why is this so amazing? First, it is a key part of the proof of Fermat's last theorem. Indeed, the motivating idea of the proof (due to G. Frey) was to use a purported solution to the Fermat equation to make an elliptic curve over $\mathbb{Q}$ that couldn't possibly satisfy the conclusion of Theorem 1.20.[3] Second, since the space $S_2(N_E)$ is finite dimensional, it means that the numbers a_p must satisfy a lot of relations. Third, it means that the L-function $L(E, s)$ of the elliptic curve equals the L-function $L(f, s)$ of the modular form.[4] In particular $L(E, s)$ inherits all the nice analytic properties of $L(f, s)$, such as analytic continuation to

[3]Run, don't walk, and read [4] for a complete presentation of this amazing story (including an excellent paper by D. Rohrlich about the basics of modular forms).

[4]This L-function is defined by the Dirichlet series $\sum_n a_n(f)n^{-s}$, where the $a_n(f)$ are the Fourier coefficients from the q-expansion. If f is a Hecke eigenform, then $L(f, s)$ also has an expression as an Euler product a la (16), although in this case the $a_p(f)$ for $p \mid N$ are the eigenvalues of the U_p operators.

all of $\mathbb{C}$ and a functional equation. Finally, although this may not be clear the first time one comes across these objects, the L-functions $L(E,s)$ and $L(f,s)$ come from completely different sources. The L-function of f is built from taking the Fourier coefficients of f and using them in a Dirichlet series, so is fundamentally an analytic object. The L-function of E is built from arithmetic data, namely counting the solutions to a equation mod p for different primes p and then forming an Euler product. There is really no reason to expect that they should have anything to do with each other. In this sense, the identity $L(E,s) = L(f,s)$ is an example of broad theme in modern number theory, which says roughly that "motivic L-functions are automorphic L-functions." There is certainly not the time to go into this huge business here. For excellent survey articles we recommend [3, 10, 11]

Lecture 2. Modular symbols

2.1. Modular curves. Our ultimate goal is to explain how to compute with modular forms. Now that we have defined modular forms, the first step is to learn more about the geometry of *modular curves*, which are quotients of $\mathfrak{H}$ by congruence subgroups. This will also help us understand the statement of the growth conditions for modular forms on congruence subgroups.

Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, that is a group containing $\Gamma(N)$ for some N. Then $\Gamma\backslash\mathfrak{H}$ is an open Riemann surface, in other words topologically is an orientable surface of some genus with some punctures.

We can canonically compactify $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ by adding *cusps*. First define

$$\mathfrak{H}^* = \mathfrak{H} \cup \mathbb{Q} \cup \{\infty\} = \mathfrak{H} \cup \mathbb{P}^1(\mathbb{Q}),$$

where $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ and $\{\infty\}$ is considered to be a single point infinitely far up the imaginary axis.

We need to put a topology on $\mathfrak{H}^*$. We do this by first extending the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathfrak{H}$ to an action on $\mathfrak{H}^*$. There are two ways to think about this:

(1) We can act directly on $\mathbb{P}^1(\mathbb{Q})$. For $z \in \mathbb{Q} \subset \mathbb{P}^1(\mathbb{Q})$, we put

$$z \longmapsto \frac{az+b}{cz+d},$$

where we use the convention that $z \mapsto \infty$ if $z = -d/c$. In other words, we act directly on fractions where the "fraction" $1/0$ is considered to be the point at infinity in $\mathbb{P}^1(\mathbb{Q})$.

(2) We can convert to integral vectors and then act: the fraction $\dfrac{m}{n}$, written in reduced terms, is converted to the vector $\begin{bmatrix} m \\ n \end{bmatrix} \in \mathbb{Z}^2$, with ∞ corresponding to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then the action of $\mathrm{SL}_2(\mathbb{Z})$ is just by matrix multiplication:

$$\begin{bmatrix} m \\ n \end{bmatrix} \longmapsto \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix}.$$

Now we define a topology on $\mathfrak{H}^*$. For a basis of open sets of ∞ we take the sets

$$B_c \coloneqq \{z \in \mathfrak{H} \colon \operatorname{Im}(z) > c\}.$$

The $\mathrm{SL}_2(\mathbb{Z})$-translates are open disks tangent to the rational points of the real axis (cf. Figure 1). This gives a system of neighborhoods of $\partial\mathfrak{H}^* = \mathfrak{H}^* \setminus \mathfrak{H}$. This induces a topology

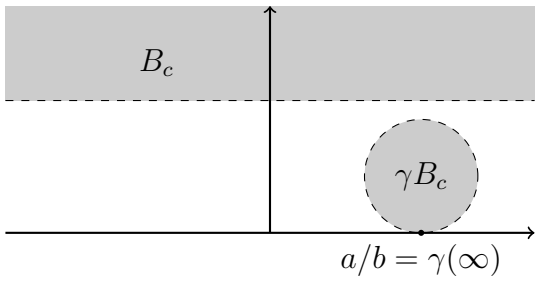

FIGURE 1. Open neighborhoods in the Satake topology on $\mathfrak{H}$.

called the *Satake topology* on $\Gamma\backslash\mathfrak{H}^*$. With this topology, the quotient $\Gamma\backslash\mathfrak{H}^*$ is now a compact Riemann surface.

2.2. Definition. The Γ-orbits in $\mathbb{P}^1(\mathbb{Q})$, and their images in the quotient $\Gamma\backslash\mathfrak{H}^*$, are called *cusps*.

Why are these points called cusps? The quotient $\Gamma\backslash\mathfrak{H}$ is more than just a topological surface. It has an induced metric, since the standard hyperbolic metric on $\mathfrak{H}$ is Γ-invariant. The metric on $\Gamma\backslash\mathfrak{H}^*$ degenerates to 0 as one approaches a cusp, and in fact the surface appears metrically to be a sharp horn in a neighborhood of a cusp. In other words, a cusp looks metrically like a cusp!

Now is also a good time to explain the connection between the cusps and the growth condition for modular forms with level. As we said before, for a general finite-index subgroup Γ of $SL_2(\mathbb{Z})$ there is more than one way to go to infinity on the quotient $\Gamma\backslash\mathfrak{H}$. The different ways correspond exactly (surprise) to the cusps, and our growth condition is effectively ensuring that the image of f doesn't blow up as one approaches a cusp on $\Gamma\backslash\mathfrak{H}^*$. However, there is a subtlety lurking here: since f is not *invariant* under the left action of Γ on $\mathfrak{H}$, f does not induce a function on the quotient $\Gamma\backslash\mathfrak{H}$. However, it is the section of a certain line bundle on $\Gamma\backslash\mathfrak{H}$, so the growth condition guarantees that this section extends over the cusps.

When Γ is one of our special congruence subgroups, we will use the following notation for its quotients:

Γ	$\Gamma\backslash\mathfrak{H}$	$\Gamma\backslash\mathfrak{H}^*$
$\Gamma(N)$	$Y(N)$	$X(N)$
$\Gamma_0(N)$	$Y_0(N)$	$X_0(N)$
$\Gamma_1(N)$	$Y_1(N)$	$X_1(N)$

Here are some examples.

2.3. Example.

$$Y(1) \simeq \mathbb{P}^1 \setminus \{\text{pt}\}, \quad X(1) \simeq \mathbb{P}^1,$$

so there is only one cusp. This is not hard to show directly: one checks that the group $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$.

2.4. Example.

$$Y(3) \simeq \mathbb{P}^1 \setminus \{4 \text{ pts}\}, \quad X(3) \simeq \mathbb{P}^1.$$

See Figure 2. Note that the four cusps correspond exactly to the four points of $\mathbb{P}^1(\mathbb{F}_3)$.

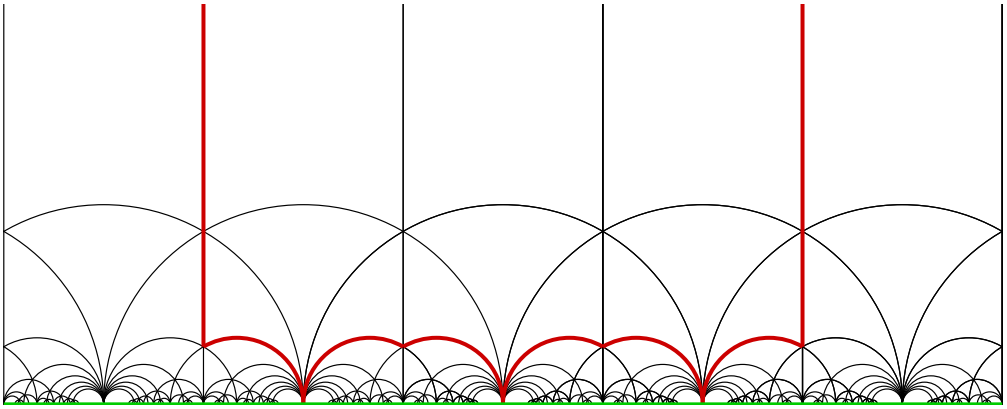

FIGURE 2. A fundamental domain for $\Gamma(3)$ is outlined in red. The four cusps are the three shown on the real axis and ∞. The edge identifications are the obvious ones that yield $X(3) \simeq \mathbb{P}^1$.

2.5. Example.

$$Y(7) \simeq C_3 \setminus \{24 \text{ pts}\}, \quad X(7) \simeq C_3, \text{ (a surface of genus 3).}$$

This time there are more cusps than points in $\mathbb{P}^1(\mathbb{F}_7)$, which has order 6.

2.6. Example.

$$Y_0(11) = C_1 \setminus \{2 \text{ pts}\}, \quad X_0(11) \simeq C_1, \text{ (surface of genus 1).}$$

2.7. Exercise. Verify all the above examples. Hint: For $\Gamma(N)$, which is torsion-free, one can take a fundamental domain in $\mathfrak{H}$ that is a finite union of the basic ideal hyperbolic triangle with vertices at 0, 1, and ∞, as in Figure 2. For $\Gamma_0(N)$, which can have torsion, one must use copies of the standard fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$, which is the geodesic (partially ideal) triangle with vertices ∞, ρ, ρ^2, where $\rho = e^{\pi i/3}$. Figure 3 shows a fundamental domain for $\Gamma_0(11)$ in $\mathfrak{H}$. The light green triangle in the middle has vertices ρ, ρ^2, and 0.

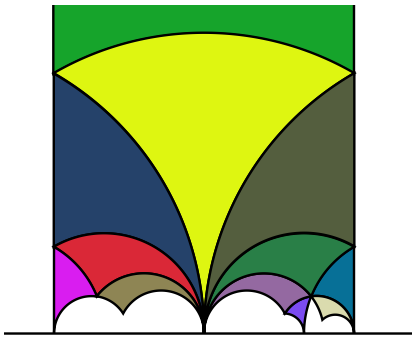

FIGURE 3. A fundamental domain for $\Gamma_0(11)$ (created using a C program by H. Verrill)

2.8. Weight 2 modular symbols. Finally, we can start talking about modular symbols. Let's focus on weight 2 for now. Suppose $f \in S_2(\Gamma)$. Then f is not a function on X_Γ, as

we said before, but $f\,dz$ is a holomorphic 1-form on X_Γ. To see why, first look at how the product $f\,dz$ transforms under Γ:

$$f\left(\frac{az+b}{cz+d}\right)\,d\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 f(z)\frac{ad-bc}{(cz+d)^2}\,dz = f\,dz.$$

Thus the modularity of f implies that the differential form $f\,dz$ is invariant under Γ. One then needs to see that $f\,dz$ is actually holomorphic on X_Γ. This is a straightforward computation in local coordinates; the only tricky parts are checking what happens at ramified points of the map $\mathfrak{H}^* \to X_\Gamma$. In particular, one can check that $f\,dz$ is actually holomorphic at the cusps, which means that any weight two cuspform determines a holomorphic 1-form on X_Γ.[5] Conversely, any holomorphic 1-form on X_Γ can be written as $f(z)dz$ for $f \in S_2(\Gamma)$. The theory of Riemann surfaces shows that $\dim(S_2(\Gamma)) = g(X_\Gamma)$, where g the genus of X_Γ as an orientable topological surface.

Now suppose that α and β are cusps that are equivalent mod Γ. We can use them to construct a homology class: we take any reasonable oriented path between α and β on $\mathfrak{H}$, say the geodesic directed from α to β, and then take the image mod Γ. Since α and β are equivalent mod Γ, the image becomes a closed oriented 1-curve on X_Γ, i.e. a 1-cycle. Thus we get a class in $H_1(X_\Gamma;\mathbb{Z})$. Let us denote this class by $\{\alpha,\beta\}$. Note that this notation looks a lot like the set $\{\alpha,\beta\}$, but it's not: it really represents an *ordered* pair, since if we change the roles of α and β we reverse the orientation on the cycle and thus get the opposite class: $\{\beta,\alpha\} = -\{\alpha,\beta\}$. This can be confusing, but the notation is traditional.

Now consider the pairing $S_2(\Gamma) \times H_1(X_\Gamma;\mathbb{Z}) \to \mathbb{C}$ given by integration

$$(f,\{\alpha,\beta\}) \longmapsto 2\pi i \int_\alpha^\beta f(z)\,dz := \langle\{\alpha,\beta\}, f\rangle. \tag{18}$$

This is independent of the path between α and β since f is holomorphic (essentially this boils down to Cauchy's theorem from complex analysis). Note also that f has to be a cuspform for the integral make sense; if f is nonvanishing at the cusp, say when f is an Eisenstein series, the integral diverges.

We can extend (18) from integral homology to real homology to get a pairing

$$S_2(\Gamma) \times H_1(X_\Gamma;\mathbb{R}) \to \mathbb{C}.$$

This is done in the obvious way. First choose an integral basis of $H_1(X_\Gamma;\mathbb{Z})$. Any class in $H_1(X_\Gamma;\mathbb{R})$ can be written as a linear combination of this basis with real coefficients, so we can extend the pairing using linearity.

Now recall that

$$\dim_{\mathbb{C}}(S_2(\Gamma)) = g,$$

from our discussion about weight 2 cuspforms and holomorphic 1-forms. Thus as a real vector space, we have

$$\dim_{\mathbb{R}}(S_2(\Gamma)) = 2g,$$

which is the same as $\dim_{\mathbb{R}}(H_1(X_\Gamma;\mathbb{R}))$. This is not a coincidence:

2.9. Claim. *The pairing $S_2(\Gamma) \times H_1(X_\Gamma;\mathbb{R}) \to \mathbb{C}$ is perfect, and identifies the dual $S_2(\Gamma)^\vee$ of $S_2(\Gamma)$ with $H_1(X_\Gamma;\mathbb{R})$.*

[5] It is important to take f to be a cuspform here. In fact, the (omitted) computation in local coordinates shows that if f is nonzero at a cusp, then the differential form $f\,dz$ will have a pole of order 1 there. This is caused by the effect of the nontrivial stabilizer of a cusp in Γ on the local coordinates. See Milne for details.

In fact, the truth of this claim has nothing to do with modular forms. It's really a combination of Poincaré duality and the Hodge theorem. There is a slight subtlety in that the differentiable structure of X_Γ is more complicated at some points, namely those whose preimages in $\mathfrak{H}^*$ have nontrivial stabilizers, but nevertheless everything works out.

Now we want to extend the notation $\{\alpha, \beta\}$ to include cusps that aren't necessarily equivalent mod Γ. This is done by integration: we can still integrate f along the geodesic from α to β, which produces a number. Thus these two cusps determine a linear form on $S_2(\Gamma)$, and so define an element of $S_2(\Gamma)^\vee = H_1(X_\Gamma; \mathbb{R})$. Thus again $\{\alpha, \beta\}$ gives a class in $H_1(X_\Gamma; \mathbb{R})$.

2.10. Definition. The *modular symbol* attached to the pair of cusps α, β is the real homology class $\{\alpha, \beta\} \in H_1(X_\Gamma; \mathbb{R})$.

Here are some basic properties of modular symbols:

(1) $\{\alpha, \beta\} = -\{\beta, \alpha\}$ (2-term relation)
(2) $\{\alpha, \beta\} = \{\alpha, \gamma\} + \{\gamma, \beta\}$ (3-term relation)
(3) $\{g\alpha, g\beta\} = \{\alpha, \beta\}$ for all $g \in \Gamma$ (Γ-action)
(4) $\{\alpha, g\alpha\} \in H_1(X_\Gamma; \mathbb{Z})$
(5) $\{\alpha, g\alpha\} = \{\beta, g\beta\}$

These are all easy to verify. The 2-term relation just says that reversing the limits of integration introduces a minus sign. The 3-term relation says that we can divide an integral into two integrals by introducing a common new endpoint. Perhaps the last is the most complicated. It can be proved by considering the square in Figure 4.

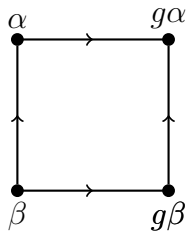

FIGURE 4. $\{\alpha, g\alpha\} = \{\beta, g\beta\}$

Properties (4) and (5) imply that we have constructed a map

$$\Gamma \longrightarrow H_1(X_\Gamma; \mathbb{Z})$$
$$g \longmapsto \{\alpha, g\alpha\}$$

that is independent of α.

By the way, our construction of modular symbols means that all we can say a priori is that $\{\alpha, \beta\} \in H_1(X_\Gamma; \mathbb{R})$, i.e. $\{\alpha, \beta\}$ is a real homology class. However, the theorem of Manin–Drinfeld tells us that this class often lies in the rational homology $H_1(X_\Gamma; \mathbb{Q}) = H_1(X_\Gamma; \mathbb{Z}) \otimes \mathbb{Q}$:

2.11. Theorem (Manin–Drinfeld)**.** *If Γ is a congruence subgroup, and α, β are cusps of Γ, then $\{\alpha, \beta\} \in H_1(X_\Gamma; \mathbb{Q})$.*

Why is this important? I.e., why should it matter whether a homology class lives in the real homology or the rational homology? The point is, this theorem has arithmetic

consequences. For instance, suppose f has q-expansion $\sum a_n q^n$. We can make an L-function from f as in the previous section using the Dirichlet series built from the a_n:

$$L(f,s) = \sum a_n/n^s.$$

(The discussion before might lead one to believe that the a_n need to be Hecke eigenvalues, since there we were connecting modular forms to elliptic curves. But this is not true.) We can make a more direct connection between f and its L-function using the Mellin transform of f. We have

$$L(f,s) = \frac{(2\pi)^s}{\Gamma(s)} \int_0^{i\infty} (-iz)^s f(z) \frac{dz}{z}. \tag{19}$$

Now the L-function of a modular form satisfies many properties, the most important of which is the existence of a functional equation taking s into $2-s$ (the 2 comes from f having weight 2). The central point $s = 1$ is especially important for many applications; for example, the Birch–Swinnerton-Dyer conjecture predicts that all the key arithmetic invariants of an elliptic curve $E/\mathbb{Q}$ are contained in the special value $L(E,1) = L(f_E,1)$, where f_E is the modular form attached to E in Theorem 1.20. Evaluating (19) at $s = 1$, we have

$$L(f,1) = -2\pi i \int_0^{i\infty} f(z)\, dz = -\langle \{0,\infty\}, f \rangle.$$

Thus the fact that the modular symbol $\{0,\infty\}$ is a rational homology class means that the special value $L(f,1)$ is a rational multiple of a period of f. In other words, the Manin–Drinfeld theorem implies that the quantity $L(f,1)$, which is a priori an extremely complicated transcendental number, actually lives in the rational span of certain other numbers, still transcendental to be sure, but nevertheless more tractable.

Let's go back to the general discussion. At this point we've written almost all the relations needed to reconstruct H_1 from our symbols. Specifically, let $\mathcal{M}_2(\Gamma)$ denote the $\mathbb{Q}$-vector space generated by the $\{\alpha,\beta\}$, modulo the 2-term and 3-term relations and Γ-action. Then we have the following result of Manin relating modular symbols to a *relative* homology group. Such groups can be unfamiliar to some, so we take a moment to recall them. Suppose X is a space with a nice subspace Y. We have the chain complexes $C_*(X)$, $C_*(Y)$ that can be used to compute their homology. We have an inclusion $C_*(Y) \to C^*(X)$ and can form the quotient chain complex $C_*(X)/C_*(Y)$. Then the relative homology of the pair (X,Y) is the homology of this complex. We denote relative homology by $H_*(X,Y;\mathbb{Z})$. Intuitively, the difference between $H_*(X)$ and $H_*(X,Y)$ is that in the latter, we consider a chain to be a cycle not only if its boundary vanishes, but also if its boundary lies in $C_*(Y)$. Now we can state Manin's key theorem:

2.12. Theorem (Manin)**.** *We have*

$$\mathcal{M}_2(\Gamma) \xrightarrow{\sim} H_1(X_\Gamma, \partial X_\Gamma; \mathbb{Q}).$$

For example, recall that the modular curve $X_0(11)$ has genus 1 and has 2 cusps. Thus $X_0(11)$ is topologically a torus, and as one learns in topology class the usual homology group $H_1(X_0(11);\mathbb{Q})$ has dimension 2. We claim the relative homology $H_1(X_0(11), \partial X_0(11);\mathbb{Q})$ is 3-dimensional. Indeed, we still have the two closed 1-cycles giving our 2 dimensions from before, and now there is an additional class, which can be represented by a path from one cusp to the other.

The space $\mathcal{M}_2(\Gamma)$ is a good start, but we are primarily interested in part of the homology relevant for studying the cusp forms, in other words $H_1(X_\Gamma;\mathbb{Q})$. But it is easy to identify the subspace of $\mathcal{M}_2(\Gamma)$ mapping onto the usual homology. From our example above, it's clear that we don't want relative classes that have boundary in the cusps. Instead we want those relative classes with vanishing boundary at the cusps. Formally, let $\mathcal{B}_2(\Gamma)$ be the $\mathbb{Q}$-vector space generated by the cusps of X_Γ, equipped with the obvious Γ-action. Define

$$\partial\colon \mathcal{M}_2(\Gamma) \longrightarrow \mathcal{B}_2(\Gamma),$$

by

$$\{\alpha,\beta\} \longmapsto \beta - \alpha.$$

A moments thought shows that this definition makes sense (the point is one has to think about the relations defining $\mathcal{M}_2(\Gamma)$ and make sure that the map is well-defined modulo them.) Put $\mathcal{S}_2(\Gamma) = \ker(\partial)$. It is clear that this is the subspace we want. Classes in $\mathcal{S}_2(\Gamma)$ are called *cuspidal modular symbols*. Manin proved that cuspidal modular symbols exactly capture the homology of X_Γ:

2.13. Theorem (Manin)**.** *We have an isomorphism*

$$\mathcal{S}_2(\Gamma) \xrightarrow{\sim} H_1(X_\Gamma;\mathbb{Q}). \tag{20}$$

After tensoring with $\mathbb{R}$, it follows from (20) that we have an isomorphism

$$\mathcal{S}_2(\Gamma)\otimes\mathbb{R} \xrightarrow{\sim} S_2(\Gamma)^\vee, \tag{21}$$

and thus have a topological model of the vector space of cuspforms.

2.14. Hecke operators and unimodular symbols. At this point we have found a way to connect the topology of the modular curve X_Γ to weight 2 modular forms on Γ. This is great but isn't good enough for number theory. The point is, we have Hecke operators acting on modular forms, and unless we can incorporate them into our model, it doesn't do us much good. But amazingly, the pairing between cuspforms and cycles, and the identification (21), are compatible with the Hecke action. Namely, there exists an action of the Hecke operators directly on the modular symbols: given a symbol $\{\alpha,\beta\}$ and an n, we can define a new (sum of) symbol(s) $T_n\{\alpha,\beta\}$, and we have the fundamental relation

$$\langle T_n\{\alpha,\beta\}, f\rangle = \langle\{\alpha,\beta\}, T_n f\rangle. \tag{22}$$

Furthermore, the action on symbols is simple to describe. We can use the matrices $\mathcal{X}_n$ from before that we used to define the Hecke action on modular forms. Let's take the set $\mathcal{X}_p$, where p is a prime not dividing the level.[6] Then we define

$$T_p\{\alpha,\beta\} = \sum_{g\in\mathcal{X}_p}\{g\alpha, g\beta\}. \tag{23}$$

The same conditions on $\mathcal{X}_p$ that guarantee that the Hecke image of a modular form is modular also guarantee that the right of (23) is a well-defined modular symbol. The relation (22) imples that if we can find eigenclasses and eigenvalues in $\mathcal{S}_2(\Gamma)$, then we can recover eigenclasses and eigenvalues in $S_2(\Gamma)$. This is great news, but unfortunately there's a catch: in its present form, our model for $\mathcal{S}_2(\Gamma)$ is not computable. The problem is that the current

[6]If $\Gamma = \Gamma_0(N)$ or $\Gamma_1(N)$, then we just mean $p\nmid N$. For a general congruence subgroup Γ, we can just fix N minimal such that $\Gamma(N)\subset\Gamma$, and then our discussion applies to $p\nmid N$.

definitions give infinite presentations of $\mathcal{M}_2(\Gamma)$ and $\mathcal{S}_2(\Gamma)$ (as spaces spanned by infinitely many symbols divided by infinitely many relations).

To address this, we want to identify a *finite* generating set of $\mathcal{M}_2(\Gamma)$. To this end, we introduce *unimodular symbols*. These are the symbols given by the pairs of cusps corresponding to the edges of the *Farey tessellation* of $\mathfrak{H}$ (Figure 5). To make this picture, take the ideal triangle in $\mathfrak{H}$ with vertices at the cusps $\{0, 1, \infty\}$. Then the $\mathrm{SL}_2(\mathbb{Z})$-translates of this triangle fill out all of $\mathfrak{H}$. The edges are the $\mathrm{SL}_2(\mathbb{Z})$-translates of the geodesic connecting 0 to ∞.

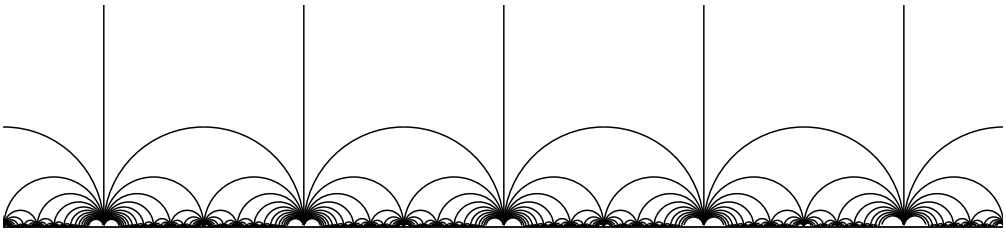

FIGURE 5. Farey tessellation of $\mathfrak{H}$. The edges are the $\mathrm{SL}_2(\mathbb{Z})$-translates of the geodesic from 0 to ∞.

Since Γ has finite-index in $\mathrm{SL}_2(\mathbb{Z})$, there are only finitely many unimodular symbols mod Γ. Thus the unimodular symbols yield a computable version of $\mathcal{S}_2(\Gamma)$, at least potentially: we of course need to know that $\mathcal{S}_2(\Gamma)$ is spanned by them, and that all the relations needed to cut out $\mathcal{S}_2(\Gamma)$ can be written using unimodular symbols (this is actually a separate question). We also have the problem that the Hecke operators can't possibly preserve unimodularity. This is clear from the definition (23); in general a symbol of the form $\{g \cdot 0, g \cdot \infty\}$ won't correspond to an edge of the tessellation.

We solve these difficulties in one stroke.

2.15. Theorem (Manin's trick, a.k.a. the modular symbol algorithm)**.** *For cusps α and β, we have the relation*

$$\{\alpha, \beta\} = \sum \{\alpha_i, \beta_i\},$$

where each term is unimodular.

Proof. Without loss of generality, assume

$$\{\alpha, \beta\} = \left\{0, \frac{p}{q}\right\}.$$

Make simple continued fraction for p/q:

$$\frac{p}{q} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots \cfrac{1}{a_r}}}} = [\![a_1, a_2, \ldots, a_r]\!].$$

We get convergents

$$\frac{p_k}{q_k} \coloneqq [\![a_1, a_2, \ldots, a_k]\!].$$

Then $\{p_k/q_k, p_{k+1}/q_{k+1}\}$ is unimodular, and our desired relation is

$$\left\{0, \frac{p}{q}\right\} = \{0,\infty\} + \left\{\infty, \frac{p_1}{q_1}\right\} + \left\{\frac{p_1}{q_1}, \frac{p_2}{q_2}\right\} + \cdots + \left\{\frac{p_{r-1}}{q_{r-1}}, \frac{p_r}{q_r}\right\}.$$

□

2.16. Example. Let's express $\{0, 71/31\}$ as a sum of unimodular symbols. We have

$$\frac{71}{31} = [\![2,3,2,4]\!].$$

Then the convergents are

$$[\![2]\!] = 2, \quad [\![2,3]\!] = \frac{7}{3}, \quad \text{and} \quad [\![2,3,2]\!] = \frac{16}{7}.$$

Thus

$$\left\{0, \frac{71}{31}\right\} = \{0,\infty\} + \{\infty, 2\} + \left\{2, \frac{7}{3}\right\} + \left\{\frac{7}{3}, \frac{16}{7}\right\} + \left\{\frac{16}{7}, \frac{71}{31}\right\}.$$

2.17. M-symbols. At this point we have

- a finite, computable model of $\mathcal{M}_2(\Gamma)$ and $\mathcal{S}_2(\Gamma)$,
- an algorithm for to compute Hecke operators.

To go further, we specialize to $\Gamma = \Gamma_0(N)$ (actually just $\Gamma_0(p)$ right now). We also introduce another trick that's even faster than the modular symbol algorithm for Hecke operator computations.

2.18. Proposition. *We have a bijection*

$$\Gamma\backslash \mathrm{SL}_2(\mathbb{Z}) \xrightarrow{\sim} \mathbb{P}^1(\mathbb{F}_p)$$

given by the bottom row map

$$\Gamma \begin{bmatrix} a & b \\ c & d \end{bmatrix} \longmapsto (c:d).$$

Proof. The group $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{F}_p)$, and the stabilizer of $(0:1)$ is Γ. □

Thus we can identify cosets in $\Gamma\backslash \mathrm{SL}_2(\mathbb{Z})$ with $\mathbb{P}^1(\mathbb{F}_p)$. This implies unimodular symbols mod Γ are in bijection with $\mathbb{P}^1(\mathbb{F}_p)$.

What about the relations? We need 2-term and 3-term relations. Let $S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, and let $R = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

2.19. Claim. *The relations are those of the form*

$$(c:d) + (c:d)S = 0 \tag{24}$$

$$(c:d) + (c:d)R + (c:d)R^2 = 0. \tag{25}$$

We have (24) from the orientation reversing identity (2-term relation). We have (25) because the boundary of a triangle is zero (3-term relation). Why? Lift $(c:d)$ to a matrix in $\mathrm{SL}_2^{\pm}(\mathbb{Z})$ to get a unimodular symbol. Then (24) visibly flips the orientation, and (25) finds one of the two Farey triangles with this as an edge (we get the other for another choice of lift). Note we use the $\pm$ here because the determinant of a Farey edge is either ± 1.

Computing, we get

$$(c:d) + (-d:c)0 \tag{26}$$

$$(c:d) + (-d:c+d) + (-c-d:c) = 0. \tag{27}$$

2.20. Theorem (M-symbols)**.** *The $\mathbb{Q}$-vector space generated by $\mathbb{P}^1(\mathbb{F}_p)$ modulo* (26) *and* (27) *is isomorphic to $\mathcal{M}_2(\Gamma_0(p))$.*

2.21. An example. It's time to actually compute something. Let's take $p = 11$ and figure out what's happening. The finite projective space $\mathbb{P}^1(\mathbb{F}_{11})$ has $12 = 11 + 1$ points. We take

$$(0:1), (1:0), (1:1), (1:2), \ldots, (1:A)$$

as representatives, and because we're lazy we abbreviate $(c:d)$ to cd. What are the relations? We start with a 12-dimensional $\mathbb{Q}$-vector space. The 2-term relation gives

$$\begin{aligned} 10 &= -01 & 13 &= -17 \\ 11 &= -1A & 14 &= -18 \\ 12 &= -15 & 16 &= -19. \end{aligned}$$

This cuts us down to a 6-dimensional space. The 3-term relation gives

$$\begin{aligned} 10 + 01 + 1A &= 0 & 12 + 14 + 17 &= 0 \\ 11 + 19 + 15 &= 0 & 13 + 16 + 18 &= 0 \end{aligned}$$

which implies

$$\begin{aligned} 10 + 01 - 11 &= 0 & 12 + 14 - 13 &= 0 \\ 11 - 16 - 12 &= 0 & 13 + 16 - 14 &= 0. \end{aligned}$$

Combining these with the 2-term relations cuts the space down to a 3-dimensional space. Namely, we can write everything in terms of 10, 12, and 14. Note that the modular symbol 10 gives a path on $X_0(11)$ connecting the two cusps, and is an example of a relative class that is truly relative (i.e. doesn't give a cycle unless one is allowed to have boundary in the cusps).

Now we want to compute the Hecke operators on the modular symbols and get our hands on the Hecke eigenvalues. There are two ways to proceed. The first one is obvious: one can simply lift M-symbols to modular symbols, and use the modular symbol algorithm. Indeed, an M-symbol $(c:d)$ can be enlarged to a matrix γ in $\mathrm{SL}_2(\mathbb{Z})$ with bottom row c, d, by undoing the bottom-row map from Proposition 2.18. Then the columns of γ can be interpreted as cusps α, β, and thus as a modular symbol $\{\alpha, \beta\}$. We can then use Manin's trick (Theorem 2.15) to work with the Hecke images of these modular symbols, and at the end can convert back to our chosen basis of M-symbols.

The second way is to work directly with M-symbols. In fact, in a sense one can precompute exactly what happens after one follows the technique above. In other words, one can write down the unimodular output of Manin's trick in a simple way, without explicitly computing the continued fractions.

2.22. Definition. Let $\mathcal{Y}_n$ be the set of integral matrices

$$\mathcal{Y}_n = \left\{ g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \det(g) = n,\ a > b \geq 0,\ d > c \geq 0 \right\}.$$

Note that $\#\mathcal{Y}_n$ is finite. These matrices, called by Mazur a set of *Heilbronn matrices*, satisfy the following amazing property, whose proof we omit. For details, we refer to [6, §II.2.4].

2.23. Claim. *If $\ell \neq p$ is prime, then*

$$T_\ell = \sum_{g \in \mathcal{Y}_\ell} (c:d)g.$$

2.24. Example. Let's compute the Hecke operator T_2 on $\mathcal{M}_2(11)$. We have

$$\mathcal{Y}_2 = \left\{ \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \right\}.$$

Thus

$$\begin{aligned}
(1:0)T_2 &= 10 + 10 + 10 + 16 \\
&= 3 \cdot 10 - 12 \\
(1:2)T_2 &= 14 + 15 + 11 + 17 \\
&= 14 - 12 - 12 - 14 \\
&= -2 \cdot 12 \\
(1:4)T_2 &= 18 + 16 + 12 + 18 \\
&= -14 - 12 + 12 - 14 \\
&= -2 \cdot 14.
\end{aligned}$$

Therefore the matrix of T_2 (with respect to the ordering $10, 12, 14$) is

$$T_2 = \begin{bmatrix} 3 & 0 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

The eigenvalues are 3, -2, -2. This is exactly what we expect. Indeed, it is known that up to isomorphism there are three distinct elliptic curves over $\mathbb{Q}$ of conductor 11. Although they are not isomorphic, they are in fact *isogenous*;[7] this (weaker) equivalence relation implies that the L-functions of any of them are equal. From [6, page 110], or from the LMFDB[8] [15] we get an equation for one of these curves:

$$E\colon y^2 + y = x^3 - x^2 - 10x - 20.$$

One can check, either by enumerating points, or by looking at [6, 15] that for E one has $a_2 = -2$. (Recall that $a_p \coloneqq p + 1 - \#E(\mathbb{F}_p)$.) Why does this eigenvalue appear twice? The reason is that the cohomology of the the modular curve $H^1(X_\Gamma, \mathbb{C})$ is not quite the same as $S_2(\Gamma)$, but instead we have

$$H^1(X_\Gamma, \mathbb{C}) \simeq S_2(\Gamma) \oplus \overline{S}_2(\Gamma),$$

where the bar denotes complex conjugation Thus we see the eigenvalue for a rational cuspform twice. What about the other eigenvalue 3? This eigenvalue comes from the Eisenstein series.

[7]Two elliptic curves E_1, E_2 over a field K are isogenous if there exists a nonconstant morphism $f\colon E_1 \to E_2$ defined over K that takes the identity (in the group law) of one into the other.

[8]The *L-functions and modular forms database*. This is an excellent resource for learning about many aspects of automorphic forms and their relationship to arithmetic geometry.

Indeed, we expect the eigenvalue $\ell+1$ for the operator T_ℓ. One can continue and find other Hecke eigenvalues.

2.25. Exercise. (1) Check by hand that the matrix of T_3 is

$$T_3 = \begin{bmatrix} 4 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

and that this agrees with a_3 for our elliptic curve defined above.

(2) Use a computer and your own program to check that the eigenvalues for the Hecke operators T_ℓ, $\ell \leq 97$, $\ell \neq 11$ agree with what you expect for the elliptic curve E.

2.26. Exercise. Consider the fundamental domain for $\Gamma_0(11)$ in Figure 3. Find all the identifications needed to glue it up into the torus (if you didn't do it already as part of Exercise 2.7). What do the images of the modular symbols 12 and 14 look like when you draw the corresponding geodesics on the torus? Do these classes generate the *integral* homology of the torus?

2.27. Exercise. Use a computer to experiment with modular symbols and to do more computations. You can use packages that compute modular symbols in SAGE, MAGMA, and Pari, and/or you can do computer-assisted computations. The former means to use built-in commands that build spaces of modular symbols and compute Hecke operators. The latter means to work as much by hand as possible but to use the computer for especially annoying subcomputations (like linear algebra, including finding characteristic polynomials). I strongly encourage the latter approach to learn what's going on. Verify more examples of Theorem 1.20, e.g. $\Gamma_0(37)$ is interesting. Another interesting example is $\Gamma_0(23)$, which contains the first newform defined over a quadratic extension of q.

What about nonprime levels? To generalize our constructions to $\Gamma_0(N)$ with N not necessarily prime, we just have to understand how to represent the cosets $\Gamma_0(N)\backslash\Gamma$. One is naturally led to the set $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$, the *projective line mod* N. By definition, this is the collection of tuples (a,b) mod N with $\gcd(a,b,N)=1$, modulo the action of the group of units $(\mathbb{Z}/N\mathbb{Z})^\times$. As an example, for $N=4$, one can show that representatives for the points in $\mathbb{P}^1(\mathbb{Z}/4\mathbb{Z})$ we can take

$$\{(1,0),(0,1),(1,1),(1,3),(1,2),(2,1)\}.$$

2.28. Exercise. Verify the analogue of Proposition 2.18 for nonprime N.

Starting from this point, one proceeds as before, with the same relations.

2.29. Exercise. Compute some Hecke operators on the modular symbols for level 14. What are the dimensions of $M_2(14)$ and $S_2(14)$? How many isogeny classes of elliptic curves over $\mathbb{Q}$ of conductor 14 are there? Search over equations and try to find a relevant elliptic curve (hint: try to match as many $a_p(f)$ as you can).

Lecture 3. Higher weight

3.1. We now know how to use modular symbols and M-symbols to compute with weight 2 modular forms of any level. How do we deal with higher weight? It turns out that almost all the same ideas work, with just a little more complexity: we need to replace the trivial

coefficients $\mathbb{C}$ from before with something more complicated, namely a nontrivial system of local coefficients.

Let X, Y be variables, and let $\mathbb{Q}[X,Y]_{k-2}$ be the space of homogeneous polynomials in X and Y of degree $k-2$. Recall that $\mathcal{M}_2$ is our space of modular symbols $\{\alpha,\beta\}$ ($\mathbb{Q}$-vector space on the symbols $\{\alpha,\beta\}$ modulo the 2-term and 3-term relations). We combine $\mathcal{M}_2$ with $\mathbb{Q}[X,Y]_{k-2}$ by defining

$$\mathcal{M}_k := \mathbb{Q}[X,Y]_{k-2} \otimes_{\mathbb{Q}} \mathcal{M}_2.$$

For this to make sense, we have to explain how the group acts on the coefficients. To ease the notation, we omit the tensor product $\otimes$ and just write $P \otimes \{\alpha,\beta\}$ as $P\{\alpha,\beta\}$. Let $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$. Then for $P \in \mathbb{Q}[X,Y]_{k-2}$ and $\{\alpha,\beta\} \in \mathcal{M}_2$, we have

$$\begin{aligned} (gP)(X,Y) &:= P\left(g^{-1}\begin{bmatrix} X \\ Y \end{bmatrix}\right) = P(dX - bY, -cX + aY), \\ g\{\alpha,\beta\} &:= \{g\alpha, g\beta\}. \end{aligned}$$

We combine these to act on symbols:

$$g(P\{\alpha,\beta\}) = gP\{g\alpha, g\beta\}.$$

We can now define higher weight modular symbols and cuspidal modular symbols just as we did for weight 2. We just need to mod out by the relations imposed by the Γ-action:

3.2. Definition.

$$\mathcal{M}_k(\Gamma) := \mathcal{M}_k/(P\{\alpha,\beta\} - g(P\{\alpha,\beta\}))$$

To get higher weight cuspidal symbols $\mathcal{S}_k(\Gamma)$, we take the kernel of a boundary map as we did before. Let

$$\begin{aligned} \mathcal{B}_2 &:= \mathbb{Q}\text{-vector space on symbols } \{\alpha\} \text{ for } \alpha \in \mathbb{P}^1(\mathbb{Q}), \\ \mathcal{B}_k &:= \mathbb{Q}[X,Y]_{k-2} \otimes_{\mathbb{Q}} \mathcal{B}_2, \\ \mathcal{B}_k(\Gamma) &:= \mathcal{B}_k/(P\{\alpha\} - gP\{g\alpha\}). \end{aligned}$$

The boundary map $\partial\colon \mathcal{M}_k(\Gamma) \to \mathcal{B}_2(\Gamma)$ is given by

$$\partial(P\{\alpha,\beta\}) = P\{\beta\} - P\{\alpha\}.$$

We define the weight k cuspidal modular symbols to be the kernel of this boundary map:

3.3. Definition.

$$\mathcal{S}_k(\Gamma) := \ker(\partial).$$

3.4. Pairing. As before, we have a pairing of cuspforms and modular symbols. Let

$$\begin{aligned} S_k(\Gamma) &:= \mathbb{C}\text{-vector space of weight } k \text{ holomorphic cuspforms} \\ \overline{S}_k(\Gamma) &= \mathbb{C}\text{-vector space of weight } k \text{ antiholomorphic cuspforms} \\ &= \{\overline{f}\colon f \in S_k(\Gamma)\}. \end{aligned}$$

The integration pairing is now

$$S_k(\Gamma) \oplus \overline{S}_k(\Gamma) \times \mathcal{M}_k(\Gamma) \to \mathbb{C}$$

$$\langle (f_1, f_2), P\{\alpha,\beta\}\rangle = \int_\alpha^\beta f_1(z)P(z,1)\,dz + \int_\alpha^\beta f_2(z)P(\overline{z},1)\,d\overline{z}.$$

3.5. Theorem (Shokurov)**.** *The pairing*

$$\langle \cdot, \cdot \rangle \colon S_k(\Gamma) \oplus \overline{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \otimes_{\mathbb{Q}} \mathbb{C} \to \mathbb{C}$$

is a nondegenerate pairing of $\mathbb{C}$*-vector spaces.*

One application of the pairing is computing special values of L-functions at critical integers. Let f be a weight k holomorphic cuspform. The L-function $L(f,s)$ has functional equation of shape $s \mapsto k - s$. The integers $j = 1, \ldots, k-1$ are called *critical* [7]. They are analogues of $s = 1$ in the weight 2 case.

We have

$$L(f,j) = \frac{(-2\pi i)^j}{(j-1)!} \langle f, X^{j-1} Y^{k-2-(j-1)} \{0, \infty\} \rangle.$$

3.6. M**-symbols for** $\Gamma_0(N)$**.** Now we specialize to $\Gamma = \Gamma_0(N)$ and work with M-symbols. The construction is very similar to what we did before.

Recall that we have defined the mod N projective line $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ to be the set of pairs $(c : d)$, where c, d are integers mod N satisfying $\gcd(c, d, N) = 1$ and where $(c : d)$ is the equivalence class modulo the action of $(\mathbb{Z}/N\mathbb{Z})^\times$. Instead of the rational numbers, we take our coefficients to be homogeneous polynomials in two variables X, Y of degree $k-2$. Thus our M-symbols are linear combinations of symbols

$$P(c : d), \quad P \in \mathbb{Q}[X,Y]_{k-2}, (c : d) \in \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}).$$

We have a right action of $\mathrm{SL}_2(\mathbb{Z})$ on such symbols by

$$(P(c : d))g = (g^{-1}P)((c : d)g), \quad g \in \mathrm{SL}_2(\mathbb{Z}).$$

To finish the job, we need to impose the 2- and 3-term relations. In fact there is a further relation we need to impose, since now the group $\mathrm{SL}_2(\mathbb{Z})$ acts nontrivially on the coefficients (we didn't see this before for weight 2). Define three matrices by

$$S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}, \quad \text{and} \quad J = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

3.7. Definition. The space $\mathcal{M}_k(N)$ of *weight* k M*-symbols mod* N is the $\mathbb{Q}$-vector space generated by $x = X^i Y^{k-2-i}(c : d) \in \mathbb{P}^1$, modulo the relations

$$\begin{aligned} x + xS &= 0, \\ x + xR + xR^2 &= 0, \\ x - xJ &= 0, \end{aligned}$$

for all x as above.

To compute Hecke operators, we just do what we did before. In particular, we use the $\mathcal{Y}_\ell$ matrices as in Claim 2.23 to compute the action of T_ℓ. However, one must beware that now g also acts on the coefficients (the action was trivial before).

3.8. Exercise. Using modular symbols show that there is a Hecke eigenform of weight 4 and level 5 with q-expansion $q - 4q^2 + 2q^3 + 8q^4 - 5q^5 - 8q^6 + 6q^7 - 23q^9 + O(q^{10})$.

Lecture 4. Further exercises

4.1. Exercise. This exercise covers some of the basics of the geometry of the upper-halfplane, if you haven't seen it before.

(1) Show that $G = \mathrm{SL}_2(\mathbb{R})$ acts transitively on $\mathfrak{H}$ via fractional linear transformations.
(2) Show that the stabilizer of i is isomorphic to $K = \mathrm{SO}(2)$. This means $\mathfrak{H} \simeq G/K$.
(3) Show that the G action on $\mathfrak{H}$ is by isometries with respect to the hyperbolic metric $(dx^2 + dy^2)/y^2$.
(4) Compute the area of the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ with respect to the hyperbolic metric.

4.2. Exercise. Show that the principal congruence subgroup $\Gamma(N) \subset \mathrm{SL}_2(\mathbb{Z})$ is torsion-free if $N \geq 3$. (Hint: the conjugates of the subgroups generated by the matrices S and R from the text contain all the elements of finite order in $\mathrm{SL}_2(\mathbb{Z})$. Look at traces.)

4.3. Exercise. To get a presentation for a group using a fundamental domain, one can use the following theorem:

4.4. Theorem. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ be a discrete group acting properly discontinuously on $\mathfrak{H}$. Let $V \subset \mathfrak{H}$ be an open connected subset such that*

$$\mathfrak{H} = \bigcup_{\gamma \in \Gamma} \gamma V,$$
$$\Sigma = \{\gamma \mid V \cap \gamma V \neq \emptyset\} \text{ is finite.}$$

Then a presentation for Γ can by constructed by taking generators to be symbols $[\gamma]$ for $\gamma \in \Sigma$ subject to the relations $[\gamma][\gamma'] = [\gamma\gamma']$ if $V \cap \gamma V \cap \gamma' V \neq \emptyset$.

Use Theorem 4.4 to get a presentation of $\mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{\pm I\}$. (Hint: take V to be a slight "thickening" of the standard fundamental domain described in Exercise 2.7.)

4.5. Exercise. This problem explores some geometry similar to the modular group acting on the upper-halfplane, and is relevant to studying other automorphic forms. Let $\mathfrak{H}_3$ be hyperbolic three-space. An "upper halfspace" model for $\mathfrak{H}_3$ can by gotten by taking the points $(z, r) \in \mathbb{C} \times \mathbb{R}_{>0}$ and using the metric $ds^2 = (dx^2 + dy^2 + dr^2)/r^2$ (here we are writing $z = x+iy$). We can also think of $\mathfrak{H}^3$ as being the subset of quaternions $\mathbf{H} = \{x+iy+rj+tk \mid x, y, r, t \in \mathbb{R}\}$ with $r > 0$ and $t = 0$. Write $P = P(z, r)$ for the quaternion corresponding to $(z, r) \in \mathfrak{H}_3$.

Let $G = \mathrm{SL}_2(\mathbb{C})$. For $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G$, define a transformation of $\mathfrak{H}_3$ by

$$M \cdot P = (aP + b)(cP + d)^{-1}.$$

In this defintion the operations on the right are to be computed in $\mathbf{H}$.

(1) Show that this is a left action of G on $\mathfrak{H}_3$.
(2) Show that the action is transitive.
(3) Show that the stabilizer of $(0, j)$ is isomorphic to

$$\mathrm{SU}(2) = \{M \in G \mid M\bar{M}^t = I\}.$$

4.6. Exercise. This exercise continues the study of hyperbolic 3-space and discrete group actions.

(1) Let $\Gamma = \mathrm{SL}_2(\mathbb{Z}[i]) \subset \mathrm{SL}_2(\mathbb{C})$. Then Γ acts on $\mathfrak{H}_3$. Show that the set

$$D = \{(x+iy, r) \in \mathfrak{H}_3 \mid 0 \leq |x|, y \leq 1/2, x^2+y^2+r^2 \geq 1\}$$

is a fundamental domain for the action of Γ on $\mathfrak{H}_3$. (Hint: consider using the matrices $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and matrices of the form $\left(\begin{smallmatrix} 1 & a+bi \\ 0 & 1 \end{smallmatrix}\right)$ to move points in $\mathfrak{H}_3$ around. Look at the proof for a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ acting on $\mathfrak{H}$ (in say, Serre, or lots of other places) to get a feel for what to do.)

(2) If you try to construct a fundamental domain of $\Gamma' = SL_2(\mathbb{Z}[\sqrt{-5}])$ using something like the above, it doesn't work. What goes wrong? (This is related to the nontrivial class number of this ring.)

4.7. Exercise. This exercise continues the study of hyperbolic 3-space. Let D be the fundamental domain of $\Gamma = \mathrm{SL}_2(\mathbb{Z}[i])$ from Exercise 4.6.

(1) The domain D has a cellular structure. Find the subgroups of Γ that stabilize the different cells on the boundary.

(2) The groups you just computed are the analogues of the subgroups of $\mathrm{SL}_2(\mathbb{Z})$ that give rise to the 2- and 3-term relations used in defining the space of M-symbols. Define M-symbols for $\mathrm{SL}_2(\mathbb{Z}[i])$ and do some computations with them (including computing Hecke eigenvalues). This was done by Cremona in his thesis; the results were published in [5]. You can look there if you get stuck or want some pointers. Yasaki has implemented these symbols in MAGMA, so you can easily play with them.

4.8. Exercise. Let $\theta(z)$ be the classical theta function

$$\theta(z) = \sum_{n \in \mathbb{Z}} q^{n^2}.$$

(1) Show that $\theta(z)^m = 1 + \sum_{k \geq 1} \rho_m(k) q^k$, where $\rho_m(k)$ is the number of ways of representing k as a sum of m squares.

(2) One can show that $\theta(z)^4$ is a modular form of weight 2 for the group $\Gamma_0(4)$. Furthermore, one knows that the space $M_2(\Gamma_0(4))$ is spanned by the two weight two Eisenstein series $E_2(z) - 2E_2(2z)$ and $E_2(z) - 4E_2(4z)$, where $E_2(z) = 1 - 24\sum_{n \geq 1} \sigma_1(n) q^n$. (In particular these combinations of E_2 are actually modular.) Write $\theta(z)^4$ in terms of these Eisenstein series.

(3) Use part (b) to prove a famous formula of Jacobi:

$$\rho_4(n) = 8 \sum_{\substack{d \mid n \\ d \not\equiv 0 \bmod 4}} d.$$

(4) Deduce Lagrange's theorem: every positive integer can be written as a sum of four squares.

4.9. Exercise. The notation for this problem is taken from Exercise 4.8. This time we consider $\theta(z)^8 \in M_4(\Gamma_0(4))$. This space of modular forms is spanned by $E_4(az)$ for $a = 1, 2, 4$. Prove

$$\rho_8(n) = 16 \sum_{\substack{d \mid n \\ d \not\equiv 2 \bmod 4}} d^3 + 12 \sum_{\substack{d \mid n \\ d \equiv 2 \bmod 4}} d^3.$$

4.10. Exercise. Suppose $Q\colon \mathbb{Z}^k \to \mathbb{Z}$ is a positive-definite quadratic form, where k is even. Then the theta series $\Theta_Q(q) = \sum_{x\in\mathbb{Z}^k} q^{Q(x)}$ is known to be a modular form of weight $k/2$ of some level and some nebentype. (For the case Q is even unimodular, this modular form has full level and trivial nebentype). This exercise explores how to pin down this modular form.

Let us write Q using an integral $k \times k$ matrix $C = (C_{ij})$ so that $Q(x) = x^t C x$ (here we view $x \in \mathbb{Z}^r$ as a column vector). Let A be the matrix $C + C^t$. Thus A is integral, even ($A_{ii} \in 2\mathbb{Z}$) and symmetric ($A_{ij} = A_{ji}$). Furthermore $Q(x) = \frac{1}{2}x^t A x$. The *level* N of Q is the smallest integer N such that $A^* = NA^{-1}$ is again even.

- Check that $N \mid \det A$ and $\det A \mid N^k$.

The *discriminant* D of Q is $(-1)^{k/2}\det A$.

- Check that D is congruent to 0 or 1 mod 4.

Then one can show that Θ_Q is a modular form of weight $k/2$, level N, and nebentype χ_D, where χ_D is the quadratic character $(\frac{D}{\bullet})$

- Compute the level and nebentype for the classical theta function from Exercises 4.8 and 4.9.
- Compute the level and nebentype for the modular forms for some interesting quadratic forms. You can take lattices from the LMFDB [15] look at their associated quadratic forms.
- Use modular form data from the LMFDB to find relations between your theta series and other modular forms. Compute interesting representation numbers for your quadratic forms in terms of divisor sums and other data.

4.11. Exercise. Let $f(z)$ be the eta-product $(\eta(z)\eta(11z))^2$, where η is the Dedekind eta-function (see Example 1.9 for the product formula for η).

(1) Compute the q-expansion of f up to q^{100}.
(2) Verify that the coefficients of this q-expansion agree with the $a(n)(E)$ data produced by the elliptic curve $y^2+y = x^3-x^2-10x-20$ of conductor 11 up to q^{100}. (Thus this eta-product gives a compact way to write down this modular form. Most modular forms, unfortunately, have no such product expansion.)

4.12. Exercise. This exercise explores another geometric aspect of the Hecke operators, namely their connection with (affine) buildings. Fix an integer q. Let $\mathcal{T}$ be the infinite tree of degree $q+1$. Thus $\mathcal{T}$ is a graph with infinitely many vertices and no cycles; each vertex of $\mathcal{T}$ is joined to $q+1$ others. Define the distance $d(v,v')$ between two vertices v, v' to be the length of the shortest path connecting them, where each edge is defined to have length 1. Finally define two sequences of correspondences $\theta_k, T_k, k \geq 0$ on the set of vertices of $\mathcal{T}$ by

$$\theta_k(v) = \sum_{d(v,v')=k} v'$$

and

$$T_k = \theta_k + T_{k-2} \quad (k \geq 2),$$

with the initial conditions $T_0 = \theta_0$, $T_1 = \theta_1$.

(1) Show that the θ_k satisfy

$$\theta_1\theta_1 = \theta_2 + (q+1)\theta_0,$$
$$\theta_1\theta_k = \theta_{k+1} + q\theta_{k-1} \quad (k \geq 2).$$

(2) Show that the T_k satisfy

$$T_k T_1 = T_{k+1} + qT_{k-1} \quad (k \geq 1).$$

(3) Explain what this has to do with Hecke operators. Hint: let $q = p$, a prime. Fix a lattice $L_0 \subset \mathbb{R}^2$ (say $L_0 = \mathbb{Z}^2$). Let $\mathcal{L}$ be the set of all sub and superlattices of L_0 such that $[L : L_0] = p^k$, $k \in \mathbb{Z}$. Define an equivalence relation on $\mathcal{L}$ by $L \sim p^k L$ for any $k \in \mathbb{Z}$. Let $[L]$ denote the class of a lattice. Then there is a bijection between lattice classes and the vertices of $\mathcal{T}$. The edges correspond to chains of the form $L \subsetneq L' \subsetneq pL$. where $[L' : L] = p$.

4.13. Exercise. This exercise and the following ones explore the analogue of modular symbols method for $\mathrm{SL}_3(\mathbb{Z})$. The original results are due to Ash–Grayson–Green [2] and Ash–Rudolph [1]. For more details one should refer to these papers.[9]

Fix a prime p and let $\Gamma_0(p) \subset \mathrm{SL}_3(\mathbb{Z})$ be the subgroup of matrices (a_{ij}) with $a_{i1} = 0 \bmod p$ for $i = 2, 3$. Just as H^1 of the modular curve provides a geometric incarnation of the holomorphic modular forms, we can use cohomology to give a geometric incarnation of certain automorphic forms for SL_3. We replace the upper-halfplane with the symmetric space $X = \mathrm{SO}(3)\backslash \mathrm{SL}_3(\mathbb{R})$ and the modular curve with the quotient $X_0(p) := X/\Gamma_0(p)$. The relevant cohomology group is $H^3(X_0(p); \mathbb{C})$.

Ash–Grayson–Green proved that the subspace $H^3_{\mathrm{cusp}}(X_0(p); \mathbb{C}) \subset H^3(X_0(p); \mathbb{C})$ of cohomology corresponding to cuspidal automorphic forms is isomorphic to the space of functions $W(p) = \{f \colon \mathbb{P}^2(\mathbb{F}_p) \to \mathbb{C}\}$ satisfying the following relations:

(1) $f(x, y, z) = f(z, x, y) = f(-x, y, z) = -f(y, x, z)$
(2) $f(x, y, z) + f(-y, x - y, z) + f(y - x, -x, z) = 0$
(3) $f(x, y, 0) = 0$
(4) $\sum_{z \in \mathbb{Z}/p\mathbb{Z}} f(x, y, z) = 0$.

(Beware: we are using a "first-column" map to identify $\mathrm{SL}_3(\mathbb{Z})/\Gamma_0(p)$ with $\mathbb{P}^2(\mathbb{F}_p)$. The argument of f is to be interpreted as homogenous coordinates.) Implement this space and compute the dimension for some primes p. You will find $p = 53$ especially interesting.

4.14. Exercise. This exercise introduces higher-dimensional modular symbols. The basic reference is [1]. (We have swapped rows and columns in our description to be compatible with [2].)

Let A be a matrix in $M_n(\mathbb{Q})$. Regard A as a collection of rational row vectors $A = (v_1, \ldots, v_n)$. Let M be the $\mathbb{Q}$-vector space generated by symbols $[A]$ modulo the following relations:

(1) $[v_1, \ldots, v_n] = (-1)^{|\sigma|}[v_{\sigma(1)}, \ldots, v_{\sigma(n)}]$ where σ is a permutation and $|\sigma|$ denotes its parity.
(2) $[A] = 0$ unless the v_i span $\mathbb{Q}^n$.
(3) $[qv_1, v_2, \ldots, v_n] = [v_1, \ldots, v_n]$ for $q \in \mathbb{Q}^\times$.
(4) If all $v_1, \ldots, v_{n+1}$ are nonzero, then $\sum_i (-1)^i [v_1, \ldots, \hat{v}_i, \ldots, v_{n+1}] = 0$, where the hat means omit v_i.

The symbols $[A]$ are called *modular symbols*. The relations imply that we can assume A is integral. If A is integral and $\det A = \pm 1$, we say $[A]$ is a *unimodular symbol*.

[9] Ask if you don't have access to them. Beware that in [2], discrete groups act on symmetric spaces from the *right*, not the left. So some objects here look different from what you might expect.

If $\gamma \in \mathrm{SL}_n(\mathbb{Z})$ we put $[A] \cdot \gamma = [A\gamma]$. In [1],[10] it is shown that for any finite-index $\Gamma \subset \mathrm{SL}_n(\mathbb{Z})$, the symbol $[A]$ modulo Γ determines a cohomology class $[A]_\Gamma \in H^N(X_\Gamma; \mathbb{Q})$, where $X_\Gamma = X/\Gamma$ (here $X = \mathrm{SO}(n)\backslash \mathrm{SL}_n(\mathbb{R})$), and $N = n(n-1)/2$.

(1) Explain the relationship between this construction and what we did above in the case $n = 2$. (Hint: it's the same thing!)
(2) Ash–Rudolph proved a version of Manin's trick (Theorem 2.15) for these symbols, and in particular gave an explicit algorithm to write any modular symbol as a sum of unimodular symbols. Implement this algorithm for $n = 3$. (Hint: the main step in the algorithm is the) following: Given a integral symbol $[A] = [v_1, \dots, v_n]$ with $D = |\det A| > 1$, find an integral point w such that the n symbols $[A_i(w)] = [v_1, \dots, v_{i-1}, w, v_{i+1}, \dots, v_n]$ satisfy $0 \le |\det A_i(w)| < D$. There are various ways to do this; for $n \le 3$ one can use LLL-reduction [9].)

4.15. Exercise. This exercise puts together Exercises 4.13 and 4.14. We freely use their notation. The goal is to compute the Hecke action on $H^3_{\mathrm{cusp}}(X_0(p); \mathbb{C})$.

Let $\ell \neq p$ be a prime. There are two types of Hecke operators at ℓ, called $T_{\ell,1}$ and $T_{\ell,2}$. The first has coset representatives

$$\mathcal{X}_{\ell,1} = V \cup V' \cup V''$$

where

$$V = \left\{ \begin{pmatrix} \ell & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}, V' = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ cp & \ell & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}, V'' = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ ap & b & \ell \end{pmatrix} \right\},$$

where $0 \le a, b, c < \ell$. The second has coset representatives

$$\mathcal{X}_{\ell,2} = W \cup W' \cup W''$$

$$W = \left\{ \begin{pmatrix} \ell & 0 & 0 \\ 0 & \ell & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}, W' = \left\{ \begin{pmatrix} \ell & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & \ell \end{pmatrix} \right\}, W'' = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ ap & \ell & 0 \\ bp & 0 & \ell \end{pmatrix} \right\},$$

where $0 \le a, b, c < \ell$. Note that these are not Heilbronn matrices, they are just the representatives for the Hecke operators. The action on modular symbols is

$$T_{\ell,k}[A] = \sum_{B \in \mathcal{X}_{\ell,k}} [AB], \quad k = 1, 2. \tag{28}$$

The symbols on the right of (28) are not unimodular, and the algorithm of Exercise 4.14 should be applied to them to make them unimodular. After this, one obtains a expression of the form

$$T_{\ell,k}[A] = \sum [A_i], \tag{29}$$

where the symbols on the right of (29) are unimodular.

Here is how this can be used to compute the Hecke action on the functions in $W(p)$ from Exercise 4.13. There is a pairing between symbols and functions, namely

$$\langle [A], f \rangle := f(A), \tag{30}$$

where the right means evaluation of f on the point of $\mathbb{P}^2(\mathbb{F}_p)$ corresponding to the first column of A. There are finitely many unimodular symbols modulo $\Gamma_0(p)$, and any $f \in W(p)$

[10]Ask if you don't have it.

is uniquely determined by its values on these unimodular symbols. Thus Hecke operators can act on $W(p)$ by taking the adjoint action with respect to (30). Concretely, if we have the relation (29) where $[A]$ is unimodular, and $T^*_{\ell,k}$ is the adjoint operator on $W(p)$, then

$$\langle [A], T^*_{\ell,k} f \rangle = \sum_i \langle [A_i], f \rangle. \tag{31}$$

Finally we can say what to do: combine these facts with the results of Exercises 4.13 and 4.14 to compute the Hecke action on any nonzero classes you found in 4.13. (Hint: $p = 53$ is the first interesting case, and it is known that the Hecke eigenvalues live in $\mathbb{Q}(\sqrt{-11})$.)

References

[1] A. Ash and L. Rudolph, *The modular symbol and continued fractions in higher dimensions*, Invent. math **55** (1979), 241–250.

[2] A. Ash, D. Grayson, and P. Green, *Computations of cuspidal cohomology of congruence subgroups of* SL(3, **Z**), J. Number Theory **19** (1984), no. 3, 412–436.

[3] D. Bump, J. W. Cogdell, E. de Shalit, D. Gaitsgory, E. Kowalski, and S. S. Kudla, *An introduction to the Langlands program*, Birkhäuser Boston, Inc., Boston, MA, 2003, Lectures presented at the Hebrew University of Jerusalem, Jerusalem, March 12–16, 2001, Edited by Joseph Bernstein and Stephen Gelbart.

[4] G. Cornell, J. H. Silverman, and G. Stevens (eds.), *Modular forms and Fermat's last theorem*, Springer-Verlag, New York, 1997, Papers from the Instructional Conference on Number Theory and Arithmetic Geometry held at Boston University, Boston, MA, August 9–18, 1995.

[5] J. E. Cremona, *Hyperbolic tessellations, modular symbols, and elliptic curves over complex quadratic fields*, Compositio Math. **51** (1984), no. 3, 275–324.

[6] J. E. Cremona, *Algorithms for modular elliptic curves*, second ed., Cambridge University Press, Cambridge, 1997.

[7] P. Deligne, *Valeurs de fonctions L et périodes d'intégrales*, Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, With an appendix by N. Koblitz and A. Ogus, pp. 313–346.

[8] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Mathematics, vol. 228, Springer-Verlag, New York, 2005.

[9] D. Doud and R. Ricks, *LLL reduction and a conjecture of Gunnells*, Proc. Amer. Math. Soc. **138** (2010), no. 2, 409–415.

[10] S. Gelbart, *An elementary introduction to the Langlands program*, Bull. Amer. Math. Soc. (N.S.) **10** (1984), no. 2, 177–219.

[11] S. S. Gelbart, *Class field theory, the Langlands program, and its application to number theory*, Automorphic forms and the Langlands program, Adv. Lect. Math. (ALM), vol. 9, Int. Press, Somerville, MA, 2010, pp. 21–67.

[12] L. J. P. Kilford, *Modular forms*, second ed., Imperial College Press, London, 2015, A classical and computational introduction.

[13] A. W. Knapp, *Elliptic curves*, Mathematical Notes, vol. 40, Princeton University Press, Princeton, NJ, 1992.

[14] S. Lang, *Introduction to modular forms*, Springer-Verlag, Berlin-New York, 1976, Grundlehren der mathematischen Wissenschaften, No. 222.

[15] LMFDB, *The L-functions and Modular Forms Database*, `www.lmfdb.org`, 2013.

[16] J. S. Milne, *Modular Functions and Modular Forms*, available from `www.jmilne.org`, 1997–2017.

[17] J.-P. Serre, *A course in arithmetic*, Springer-Verlag, New York-Heidelberg, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7.

[18] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan, vol. 11, Princeton University Press, Princeton, NJ, 1994, Reprint of the 1971 original, Kanô Memorial Lectures, 1.

[19] N. J. A. Sloane, *Online Encyclopedia of Integer Sequences*, available at `oeis.org`.
[20] D. Zagier, *Elliptic modular forms and their applications*, The 1-2-3 of modular forms, Universitext, Springer, Berlin, 2008, pp. 1–103.

Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003-9305

E-mail address: `gunnells@math.umass.edu`

URL: `http://www.math.umass.edu/~gunnells/`