Chapter 3

Modular Forms of Weight Two

We saw in Chapter 2 (especially Section 2.2) that we can compute each space $M_k(\mathrm{SL}_2(\mathbb{Z}))$ explicitly. This involved computing Eisenstein series E_4 and E_6 to some precision, then forming the basis $\{E_4^a E_6^b : 4a + 6b = k, 0 \leq a, b \in \mathbb{Z}\}$ for $M_k(\mathrm{SL}_2(\mathbb{Z}))$. In this chapter we instead consider the problem of computing $S_2(\Gamma_0(N))$, for any positive integer N. Again we have a decomposition

$$M_2(\Gamma_0(N)) = S_2(\Gamma_0(N)) \oplus \mathrm{Eis}_2(\Gamma_0(N)),$$

where $\mathrm{Eis}_2(\Gamma_0(N))$ is a space spanned by generalized Eisenstein series and $S_2(\Gamma_0(N))$ is the space of cusp forms, i.e., elements of $M_2(\Gamma_0(N))$ that vanish at *all* cusps.

In Chapter 5 we compute the space $\mathrm{Eis}_2(\Gamma_0(N))$ much like how we computed $M_k(\mathrm{SL}_2(\mathbb{Z}))$. On the other hand, elements of $S_2(\Gamma_0(N))$ often cannot be written as sums or products of generalized Eisenstein series. In fact, the structure of $M_2(\Gamma_0(N))$ is, in general, much more complicated than that of $M_k(\mathrm{SL}_2(\mathbb{Z}))$. For example, when p is a prime $\mathrm{Eis}_2(\Gamma_0(p))$ has dimension 1, whereas $S_2(\Gamma_0(p))$ has dimension about $p/12$.

Fortunately an idea of Birch called "modular symbols" provides a powerful method for computing $S_2(\Gamma_0(N))$, and indeed much more that is relevant to understanding special values of L-functions. Modular symbols are also a powerful theoretical tool. In this chapter, we explain how $S_2(\Gamma_0(N))$ is related to modular symbols, and how to use this relationship to explicitly compute a basis for $S_2(\Gamma_0(N))$. In Chapter 8 we will introduce more general modular symbols and explain how to use them to compute $S_k(\Gamma_0(N))$, $S_k(\Gamma_1(N))$ and $S_k(N, \chi)$ for any integers $k \geq 2$ and N and Dirichlet character χ.

Section 3.1 contains a brief summary of basic facts about modular forms of weight 2, modular curves, Hecke operators, and integral homology. Section 3.2 introduces modular symbols, and describes how to compute with them. In Section 3.4 we talk about how to cut out the subspace of modular symbols corresponding to cuspforms using the boundary map. Section 3.5 is about a straightforward method to compute a basis for $S_2(\Gamma_0(N))$ using modular symbols, and Section 3.6 outlines a more sophisticated algorithm for constructing cuspforms that uses Atkin-Lehner theory.

Before reading this chapter you should have read Chapter 1 and at least skimmed Chapter 2, though this chapter is mostly independent of Chapter 2. We also assume familiarity with algebraic curves, Riemann surfaces, and homology groups of compact Riemann surfaces.

3.1 Weight 2 modular forms and Hecke operators

The group $\Gamma_0(N)$ acts on $\mathfrak{h}^*$ by linear fractional transformations, and the quotient $\Gamma_0(N)\backslash\mathfrak{h}^*$ is a Riemann surface, which we denote by $X_0(N)$. Shimura showed in [Shi94, §6.7] that $X_0(N)$ has a canonical structure of algebraic curve over $\mathbb{Q}$.

Recall from Section 1.3 that a cusp form of weight 2 for $\Gamma_0(N)$ is a function f on $\mathfrak{h}$ such that $f(z)dz$ defines a holomorphic differential on $X_0(N)$. Equivalently, a cusp form is a holomorphic function f on $\mathfrak{h}$ such that

(a) the expression $f(z)dz$ is invariant under replacing z by $\gamma(z)$ for each $\gamma \in \Gamma_0(N)$, and

(b) $f(z)$ vanishes at every cusp for $\Gamma_0(N)$.

The space $S_2(\Gamma_0(N))$ of weight 2 cusp forms on $\Gamma_0(N)$ is a finite dimensional complex vector space, of dimension equal to the genus g of $X_0(N)$. Viewed topologically, as a 2-dimensional real manifold, $X_0(N)(\mathbb{C})$ is a g-holed torus (see Figure 3.1.1 on page 42).

Condition (b) in the definition of $f(z)$ means that $f(z)$ has a Fourier expansion about each element of $\mathbb{P}^1(\mathbb{Q})$. Thus, at ∞ we have

$$\begin{aligned} f(z) &= a_1 e^{2\pi i z} + a_2 e^{2\pi i 2z} + a_3 e^{2\pi i 3z} + \cdots \\ &= a_1 q + a_2 q^2 + a_3 q^3 + \cdots, \end{aligned}$$

where, for brevity, we write $q = q(z) = e^{2\pi i z}$.

Example 3.1.1. Let E be the elliptic curve defined by the equation $y^2 + xy = x^3 + x^2 - 4x - 5$. Let $a_p = p + 1 - \#\tilde{E}(\mathbb{F}_p)$, where $\tilde{E}$ is the reduction of E mod p (note that for the bad primes we have $a_3 = -1$, $a_{13} = 1$). For n composite, define a_n using the relations at the end of Section 3.6. Then

$$\begin{aligned} f &= q + a_2 q^2 + a_3 q^3 + a_4 q^4 + a_5 q^5 + \cdots \\ &= q + q^2 - q^3 - q^4 + 2q^5 + \cdots \end{aligned}$$

is the q-expansion of a modular form on $\Gamma_0(39)$. The Shimura-Taniyama conjecture, which is now a theorem (see [BCDT01]) asserts that any q-expansion constructed as above from an elliptic curve over $\mathbb{Q}$ is a modular form.

Just as is the case for level 1 modular forms (see Section 2.4) there are commuting Hecke operators $T_1, T_2, T_3, \ldots$ that act on $S_2(\Gamma_0(N))$. To define them conceptually, we introduce an interpretation of the modular curve $X_0(N)$ as an object whose points *parameterize* elliptic curves with extra structure.

Proposition 3.1.2. *The complex points of the open subcurve* $Y_0(N) = \Gamma_0(N)\backslash\mathfrak{h}$ *are in natural bijection with isomporphism classes of pairs* (E, C)*, where* E *is an elliptic curve over* $\mathbb{C}$ *and* C *is a cyclic subgroup of* $E(\mathbb{C})$ *of order* N*. The class of the point* $\lambda \in \mathfrak{h}$ *corresponds to the pair*

$$\left(\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\lambda), \quad \left(\frac{1}{N}\mathbb{Z}\right)/(\mathbb{Z}+\mathbb{Z}\lambda)\right).$$

Proof. See Exercise 3.1 and [[**TODO: reference to the literature.**]] □

Suppose n and N are coprime positive integers. In view of Proposition 3.1.2, there are two natural maps π_1 and π_2 from $Y_0(n \cdot N)$ to $Y_0(N)$; the first, π_1, sends $(E, C) \in Y_0(N)(\mathbb{C})$ to (E, C'), where C' is the unique cyclic subgroup of C of order N, and the second, π_2, sends (E, C) to $(E/D, C/D)$, where D is the unique cyclic subgroup of C of order n. These maps extend in a unique way to algebraic maps from $X_0(n \cdot N)$ to $X_0(N)$:

$$\begin{tikzcd} & X_0(n\cdot N) \arrow[dl, "\pi_2"'] \arrow[dr, "\pi_1"] & \\ X_0(N) & & X_0(N). \end{tikzcd} \tag{3.1.1}$$

The nth *Hecke operator* T_n is $\pi_{1*} \circ \pi_2^*$, where π_2^* and π_{1*} denote pullback and pushforward of differentials respectively. (There is a similar definition of T_n when $\gcd(n, N) \neq 1$.) Using our interpretation of $S_2(\Gamma_0(N))$ as differentials on $X_0(N)$ this gives an action of Hecke operators on $S_2(\Gamma_0(N))$. One can show that these induce the maps of Proposition 2.4.6 on q-expansions.

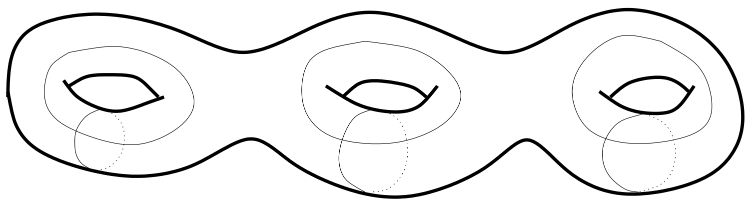
Example 3.1.3. There is a basis of $S_2(39)$ so that

$$T_2 = \begin{pmatrix} 1 & 1 & 0 \\ -2 & -3 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad T_5 = \begin{pmatrix} -4 & -2 & -6 \\ 4 & 4 & 4 \\ 0 & 0 & 2 \end{pmatrix}.$$

Notice that these matrices commute. Also, the characteristic polynomial of T_2 is $(x-1)\cdot(x^2+2x-1)$.

3.1.1 Homology

The first homology group $\mathrm{H}_1(X_0(N), \mathbb{Z})$ is the group of closed 1-cycles modulo boundaries of 2-cycles (formal sums of images of 2-simplexes). Recall that topologically $X_0(N)$ is a g-holed torus, where g is the genus of $X_0(N)$. The group $\mathrm{H}_1(X_0(N), \mathbb{Z})$ is thus a free abelian group of rank $2g$ (see, e.g., [GH81,



$$\mathrm{H}_1(X_0(39), \mathbb{Z}) \cong \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$$

Figure 3.1.1: The homology of $X_0(39)$.

Ex. 19.30]), with two generators corresponding to each hole, as illustrated in the case $N = 39$ in Figure 3.1.1.

Homology is closely connected to modular forms, since the Hecke operators T_n also act on $\mathrm{H}_1(X_0(N), \mathbb{Z})$. The action is by pullback of homology classes by π_2 followed by taking the image under π_1, where π_1 and π_2 are as in (3.1.1).

Integration defines a pairing

$$\langle\ ,\ \rangle : S_2(\Gamma_0(N)) \times H_1(X_0(N), \mathbb{Z}) \to \mathbb{C}. \qquad (3.1.2)$$

Explicitly, for a path x,

$$\langle f, x\rangle = 2\pi i \cdot \int_x f(z)dz.$$

Theorem 3.1.4. *The pairing (3.1.2) is nondegenerate and Hecke equivariant in the sense that for every Hecke operator T_n, we have $\langle fT_n, x\rangle = \langle f, T_n x\rangle$. Moreover, it induces a perfect pairing*

$$\langle\ ,\ \rangle : S_2(\Gamma_0(N)) \times H_1(X_0(N), \mathbb{R}) \to \mathbb{C}. \qquad (3.1.3)$$

Proof. [[**Todo: Add reference to the literature.**]] □

As we will see, modular symbols allow us to make explicit the action of the Hecke operators on $\mathrm{H}_1(X_0(N), \mathbb{Z})$; the above pairing then translates this into a wealth of information about cusp forms.

3.2 Modular symbols

The modular symbols formalism provides a presentation of $\mathrm{H}_1(X_0(N), \mathbb{Z})$ in terms of paths between elements of $\mathbb{P}^1(\mathbb{Q})$. Furthermore, we will show that there is an explicit finite list of generators and relations for $\mathrm{H}_1(X_0(N), \mathbb{Z})$.

The *modular symbol* defined by a pair $\alpha, \beta \in \mathbb{P}^1(\mathbb{Q})$ is denoted $\{\alpha, \beta\}$. As illustrated in Figure 3.2.1, we view this modular symbol as the homology class, relative to the cusps, of a (geodesic) path from α to β in $\mathfrak{h}^*$. The homology group $\mathrm{H}_1(X_0(N), \mathbb{Z}; \text{cusps})$ of $X_0(N)$ *relative to the cusps* is an enlargement of the usual homology group, in that we allow paths with endpoints in the cusps instead of restricting to closed loops.

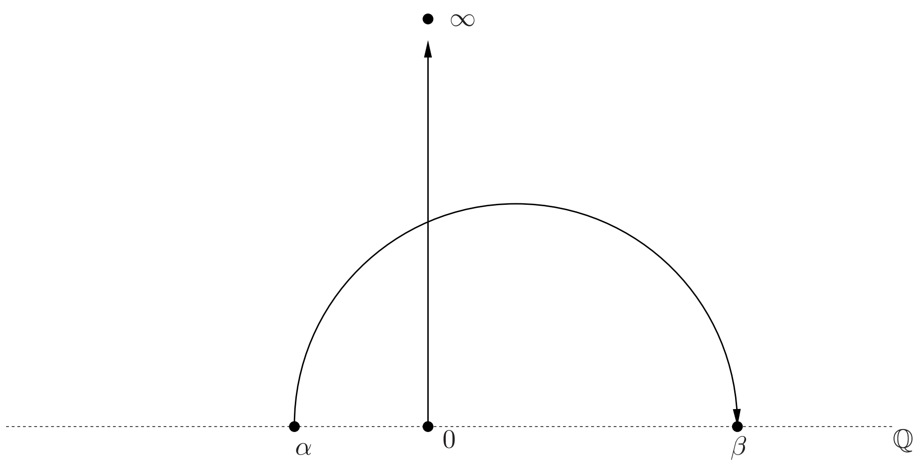


Figure 3.2.1: The modular symbols $\{\alpha, \beta\}$ and $\{0, \infty\}$.

Note that modular symbols satisfy the following homology relations: if $\alpha, \beta, \gamma \in \mathbb{Q} \cup \{\infty\}$, then

$$\{\alpha, \beta\} + \{\beta, \gamma\} + \{\gamma, \alpha\} = 0.$$

Furthermore, the space of modular symbols is torsion free, so, e.g., $\{\alpha, \alpha\} = 0$ and $\{\alpha, \beta\} = -\{\beta, \alpha\}$.

Denote by $\mathbb{M}_2$ the free abelian group with basis the set of symbols $\{\alpha, \beta\}$ modulo the 3-term homology relations above and modulo any torsion. There is a left action of $\mathrm{GL}_2(\mathbb{Q})$ on $\mathbb{M}_2$, whereby a matrix g acts by

$$g\{\alpha, \beta\} = \{g(\alpha), g(\beta)\},$$

and g acts on α and β by a linear fractional transformation. The space $\mathbb{M}_2(\Gamma_0(N))$ of *modular symbols for* $\Gamma_0(N)$ is the quotient of $\mathbb{M}_2$ by the submodule generated by the infinitely many elements of the form $x - g(x)$, for x in $\mathbb{M}_2$ and g in $\Gamma_0(N)$, and modulo any torsion. A *modular symbol for* $\Gamma_0(N)$ is an element of this space. We frequently denote the equivalence class that defines a modular symbol by giving a representative element.

Example 3.2.1. Some modular symbols are 0 no matter what the level N is! For example, since $\gamma = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, we have

$$\{\infty, 0\} = \{\gamma(\infty), \gamma(0)\} = \{\infty, 1\},$$

so

$$0 = \{\infty, 1\} - \{\infty, 0\} = \{\infty, 1\} + \{0, \infty\} = \{0, \infty\} + \{\infty, 1\} = \{0, 1\}.$$

There is a natural homomorphism

$$\varphi : \mathbb{M}_2(\Gamma_0(N)) \to \mathrm{H}_1(X_0(N), \{\text{cusps}\}, \mathbb{Z}) \tag{3.2.1}$$

that sends a formal linear combination of geodesic paths in the upper half plane to their image as paths on $X_0(N)$. In [Man72] Manin proved that (3.2.1) is an isomorphism (this is a fairly involved topological argument).

Manin also identified the subspace of $\mathbb{M}_2(\Gamma_0(N))$ that is sent isomorphically onto $\mathrm{H}_1(X_0(N), \mathbb{Z})$. Let $\mathbb{B}_2(\Gamma_0(N))$ denote the free abelian group whose basis is the finite set $C(\Gamma_0(N)) = \Gamma_0(N)\backslash\mathbb{P}^1(\mathbb{Q})$ of cusps for $\Gamma_0(N)$. The *boundary map*

$$\delta : \mathbb{M}_2(\Gamma_0(N)) \to \mathbb{B}_2(\Gamma_0(N))$$

sends $\{\alpha, \beta\}$ to $\{\beta\} - \{\alpha\}$, where $\{\beta\}$ denotes the basis element of $\mathbb{B}_2(\Gamma_0(N))$ corresponding to $\beta \in \mathbb{P}^1(\mathbb{Q})$. The kernel $\mathbb{S}_2(\Gamma_0(N))$ of δ is the subspace of *cuspidal* modular symbols. An element of $\mathbb{S}_2(\Gamma_0(N))$ can be thought of as a linear combination of paths in $\mathfrak{h}^*$ whose endpoints are cusps, and whose images in $X_0(N)$ are a linear combination of loops.

Theorem 3.2.2 (Manin)**.** *The map φ given above induces a canonical isomorphism*

$$\mathbb{S}_2(\Gamma_0(N)) \cong \mathrm{H}_1(X_0(N), \mathbb{Z}).$$

Example 3.2.3. We illustrate modular symbols in the case when $N = 11$. Using SAGE we find that $\mathbb{M}_2(11)$ has basis $\{\infty, 0\}$, $\{-1/8, 0\}$, $\{-1/9, 0\}$:

```
sage: M = ModularSymbols(11, 2)
sage: print [b.modular_symbol_rep() for b in M.basis()]
[{Infinity,0}, {-1/8,0}, {-1/9,0}]
```

The integral homology $\mathrm{H}_1(X_0(11), \mathbb{Z})$ corresponds to the abelian subgroup generated by $\{-1/7, 0\}$ and $\{-1/5, 0\}$.

3.3 Computing with modular symbols

3.3.1 Manin's trick

In this section, we describe a trick of Manin that shows that the space of modular symbols can be computed.

The group $\Gamma_0(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$ (see Exercise 1.6). Let $r_0, r_1, \ldots, r_m$ be distinct right coset representatives for $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$, so that

$$\mathrm{SL}_2(\mathbb{Z}) = \Gamma_0(N)r_0 \cup \Gamma_0(N)r_1 \cup \cdots \cup \Gamma_0(N)r_m,$$