Explicit computations with modular Galois representations

Bosman, J.G.

Chapter 2

Computations with modular forms

In this chapter we will discuss several aspects of computations with modular forms. Let us warn the reader on beforehand that we will focus on how to compute in practice, not on theoretical aspects of computability. What in theory can be proven to be computable, can often not be computed in practice and what in practice can be computed, can often not be proven to be computable in theory.

2.1 Modular symbols

Modular symbols provide a way of doing symbolic calculations with modular forms, as well as the homology of modular curves. In this section as well, our intention is to give the reader an idea of what is going on rather than a complete and detailed account of the material. For more details and further reading on the subject of modular symbols, the reader could take a look at [51], [72] and [53]. A computational approach to the material can be found in [78] and [79].

2.1.1 Definitions

Let A be the free abelian group on the symbols $\{\alpha,\beta\}$ with $\alpha,\beta \in \mathbb{P}^1(\mathbb{Q})$. Consider the subgroup $I \subset A$ generated by all elements of the forms

$$\{\alpha,\beta\}+\{\beta,\gamma\}+\{\gamma,\alpha\}, \quad \{\alpha,\beta\}+\{\beta,\alpha\}, \quad \text{and} \quad \{\alpha,\alpha\}.$$

We define the group

$$\mathbb{M}_2 := (A/I)/\text{torsion}$$

as the quotient of A/I by its torsion subgroup. By a slight abuse of notation, we will denote the class of $\{\alpha,\beta\}$ in this quotient also by $\{\alpha,\beta\}$. We have an action $\mathrm{GL}_2^+(\mathbb{Q})$ on $\mathbb{M}_2$ by

$$\gamma\{\alpha,\beta\} := \{\gamma\alpha,\gamma\beta\},$$

where γ acts on $\mathbb{P}^1(\mathbb{Q})$ by fractional linear transformations.

For $k \geq 2$, we consider also the abelian group $\mathbb{Z}[x,y]_{k-2} \subset \mathbb{Z}[x,y]$ of homogeneous polynomials of degree $k-2$ and we let matrices in $\mathrm{GL}_2^+(\mathbb{Q})$ with integer coefficients act on it on the left by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} P(x,y) := P(dx-by, -cx+ay).$$

We define

$$\mathbb{M}_k := \mathbb{Z}[x,y]_{k-2} \otimes \mathbb{M}_2,$$

and we equip $\mathbb{M}_k$ with the component-wise action of integral matrices in $\mathrm{GL}_2^+(\mathbb{Q})$ (that is $\gamma(P \otimes \alpha) = \gamma(P) \otimes \gamma(\alpha)$).

Definition 2.1. Let $k \geq 2$ be an integer. Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a subgroup of finite index and let $I \subset \mathbb{M}_k$ be the subgroup generated by all elements of the form $\gamma x - x$ with $\gamma \in \Gamma$ and $x \in \mathbb{M}_k$. Then we define the space of *modular symbols* of weight k for Γ to be the quotient of M_k/I by its torsion subgroup and we denote this space by $\mathbb{M}_k(\Gamma)$:

$$\mathbb{M}_k(\Gamma) := (\mathbb{M}_k/I)/\text{torsion}.$$

In the special case $\Gamma = \Gamma_1(N)$, which we will mostly be interested in, $\mathbb{M}_k(\Gamma)$ is called the space of modular symbols of weight k and level N. The class of $\{\alpha,\beta\}$ in $\mathbb{M}_k(\Gamma)$ will be denoted by $\{\alpha,\beta\}_\Gamma$ or, if no confusion exists, by $\{\alpha,\beta\}$.

The group $\Gamma_0(N)$ acts naturally on $\mathbb{M}_k(\Gamma_1(N))$ and hence induces an action of $(\mathbb{Z}/N\mathbb{Z})^\times$ on $\mathbb{M}_k(\Gamma_1(N))$. We denote this action by the diamond symbol $\langle d \rangle$. The operator $\langle d \rangle$ on $\mathbb{M}_k(\Gamma_1(N))$ is called a *diamond operator*. This leads to the notion of modular symbols with character.

Definition 2.2. Let $\varepsilon : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ be a Dirichlet character. Denote by $\mathbb{Z}[\varepsilon] \subset \mathbb{C}$ the subring generated by all values of ε. Let $I \subset \mathbb{M}_k(\Gamma_1(N)) \otimes \mathbb{Z}[\varepsilon]$ be the $\mathbb{Z}[\varepsilon]$-submodule generated by all elements of the form $\langle d \rangle x - \varepsilon(d)x$ with $d \in (\mathbb{Z}/N\mathbb{Z})^\times$ and $x \in \mathbb{M}_k(\Gamma_1(N))$. Then we define the space $\mathbb{M}_k(N,\varepsilon)$ of modular symbols of weight k, level N and character ε as the $\mathbb{Z}[\varepsilon]$-module

$$\mathbb{M}_k(N,\varepsilon) := \big(\mathbb{M}_k(\Gamma_1(N)) \otimes \mathbb{Z}[\varepsilon]/I\big)/\text{torsion}.$$

We denote the elements of $\mathbb{M}_k(N,\varepsilon)$ by $\{\alpha,\beta\}_{N,\varepsilon}$ or simply by $\{\alpha,\beta\}$. If ε is trivial, then we have $\mathbb{M}_k(N,\varepsilon) \cong \mathbb{M}_k(\Gamma_0(N))$.

Let $\mathbb{B}_2$ be the free abelian group on the symbols $\{\alpha\}$ with $\alpha \in \mathbb{P}^1(\mathbb{Q})$ with action of $\mathrm{SL}_2(\mathbb{Z})$ by $\gamma\{\alpha\} = \{\gamma\alpha\}$ and define $\mathbb{B}_k := \mathbb{Z}[x,y]_{k-2} \otimes \mathbb{B}_2$ with component-wise $\mathrm{SL}_2(\mathbb{Z})$-action. Elements of $\mathbb{B}_k$ are called *boundary modular symbols*. For a subgroup $\Gamma < \mathrm{SL}_2(\mathbb{Z})$ of finite index, we define $\mathbb{B}_k(\Gamma)$ as

$$\mathbb{B}_k(\Gamma) := (\mathbb{B}_k/I)/\text{torsion}$$

where I is the subgroup of $\mathbb{B}_k$ generated by all elements $\gamma x - x$ with $\gamma \in \Gamma$ and $x \in \mathbb{B}_k$. We define $\mathbb{B}_k(N,\varepsilon)$ to be the quotient of $(\mathbb{B}_k(\Gamma_1(N)) \otimes \mathbb{Z}[\varepsilon])/I$ by its torsion submodule, where I is the $\mathbb{Z}[\varepsilon]$-submodule of $\mathbb{B}_k(\Gamma_1(N)) \otimes \mathbb{Z}[\varepsilon]$ generated by the elements $\gamma x - \varepsilon(\gamma)x$ with $\gamma \in \Gamma_0(N)$.

We have *boundary homomorphisms* $\delta : \mathbb{M}_k(\Gamma) \to \mathbb{B}_k(\Gamma)$ and $\delta : \mathbb{M}_k(N,\varepsilon) \to \mathbb{B}_k(N,\varepsilon)$, defined by

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

The spaces of *cuspidal modular symbols*, denoted by $\mathbb{S}_k(\Gamma)$ and $\mathbb{S}_k(N,\varepsilon)$ respectively are defined as the kernel of δ.

2.1.2 Properties

One can interpret the symbol $\{\alpha,\beta\}$ as a smooth path in $\mathfrak{H}^*$ from the cusp α to the cusp β, lying in $\mathfrak{H}$ except for the endpoints α and β. It can be shown that this interpretation induces an isomorphism

$$\mathbb{M}_2(\Gamma) \cong H_1(X_\Gamma, \text{cusps}, \mathbb{Z}).$$

Here the homology is taken of the topological pair $(X_1(N), \text{cusps})$. We also get an isomorphism

$$\mathbb{S}_2(\Gamma) \cong H_1(X_\Gamma, \mathbb{Z}).$$

So we immediately see that there is a perfect pairing

$$(\mathbb{S}_2(\Gamma(N)) \otimes \mathbb{C}) \times \left(S_2(\Gamma(N)) \oplus \overline{S}_2(\Gamma(N))\right) \to \mathbb{C}$$

defined by

$$(\{\alpha,\beta\}, f \oplus g) \mapsto \int_\alpha^\beta \left(f\frac{dq}{q} + g\frac{d\overline{q}}{\overline{q}} \right).$$

More generally, there is a pairing

$$\mathbb{M}_k(\Gamma_1(N)) \times \left(S_k(\Gamma_1(N)) \oplus \overline{S}_k(\Gamma_1(N))\right) \to \mathbb{C} \tag{2.1}$$

defined by

$$(P \otimes \{\alpha,\beta\}, f \oplus g) \mapsto 2\pi i \int_\alpha^\beta (f(z)P(z,1)dz - g(z)P(\overline{z},1)d\overline{z}),$$

which becomes perfect if we restrict and tensor the left factor to $\mathbb{S}_k(\Gamma(N)) \otimes \mathbb{C}$. This pairing induces a pairing

$$(\mathbb{M}_k(N,\varepsilon)) \times \left(S_k(N,\varepsilon) \oplus \overline{S}_k(N,\varepsilon)\right) \to \mathbb{C}$$

which is perfect when the left factor is restricted and tensored to $\mathbb{S}_k(N,\varepsilon) \otimes_{\mathbb{Z}[\varepsilon]} \mathbb{C}$. From now on we will denote all these pairings with the notation

$$(x, f) \mapsto \langle x, f \rangle.$$

The star involution

On the spaces $\mathbb{M}_k(\Gamma_1(N))$ and $\mathbb{M}_k(N,\varepsilon)$ we have an involution ι^* defined by

$$\iota^*(P(x,y)\otimes\{\alpha,\beta\}) := -P(x,-y)\otimes\{-\alpha,-\beta\},$$

which is called the *star involution*. It preserves cuspidal subspaces. We define $\mathbb{S}_k(\Gamma_1(N))^+$ and $\mathbb{S}_k(\Gamma_1(N))^-$ subspaces of $\mathbb{S}_k(\Gamma_1(N))$ where ι^* acts as $+1$ and -1 respectively and we use similar definitions for $\mathbb{S}_k(N,\varepsilon)^\pm$. It can be shown that the pairing (2.1) induces perfect pairings

$$(\mathbb{S}_k(\Gamma_1(N))^+\otimes\mathbb{C})\times S_k(\Gamma_1(N))\to\mathbb{C}$$

and

$$(\mathbb{S}_k(\Gamma_1(N))^-\otimes\mathbb{C})\times \overline{S}_k(\Gamma_1(N))\to\mathbb{C}$$

and similarly for the spaces with character. This allows us to work sometimes in modular symbols spaces of half the dimension of the full cuspidal space.

2.1.3 Hecke operators

Hecke operators on modular symbols are defined in a similar way as on modular forms (see Subsection 1.1.4). Let $k\geq 2$ and $N\geq 1$ be given. Then for $\gamma\in\mathrm{GL}_2^+(\mathbb{Q})\cap\mathrm{M}_2(\mathbb{Z})$ we define an operator T_γ on $\mathbb{M}_k(\Gamma_1(N))$ by letting $\gamma_1,\ldots,\gamma_r$ be double coset representatives for $\Gamma_1(N)\setminus\Gamma_1(N)\gamma\Gamma_1(N)$ and putting

$$T_\gamma(x) := \sum_{i=1}^{r}\gamma_i x \quad \text{for } x\in\mathbb{M}_k(\Gamma_1(N)). \tag{2.2}$$

It follows from [72, Theorem 4.3] that this operator is well-defined. For a prime number p we put $T_p = T_\gamma$ for $\gamma = \left(\begin{smallmatrix}1&0\\0&p\end{smallmatrix}\right)$ and for positive integers n we define T_n by means of the relations (1.13). The operators T_n are called Hecke operators.

The Hecke operators preserve the subspace $\mathbb{S}_k(\Gamma_1(N))$ and induce an action on the spaces $\mathbb{M}_k(N,\varepsilon)$ and $\mathbb{S}_k(N,\varepsilon)$. Furthermore, from [72, Theorem 4.3] one can conclude that the diamond and Hecke operators are self-adjoint with respect to the pairings defined in the previous subsection:

$$\langle Tx,f\rangle = \langle x,Tf\rangle. \tag{2.3}$$

for any modular symbol x, cusp form f and diamond or Hecke operator T for which this relation is well-defined. Furthermore, the Hecke operators commute with the star involution ι^*.

In conclusion, we see how we can write cusp forms spaces as the dual of modular symbols spaces. The computation of Hecke operators on these modular symbols spaces would enable us to compute q-expansions of cusp forms: q-coefficients of newforms can be computed once we can compute the eigenvalues of Hecke operators. But because of (2.3) this reduces to the computation of the eigenvalues of Hecke operators on modular symbols spaces. In computations one often works with the spaces $\mathbb{S}_k(N,\varepsilon)^+\otimes_{\mathbb{Z}[\varepsilon]}\mathbb{Q}(\varepsilon)$ because these have smaller dimension than $\mathbb{S}_k(\Gamma_1(N))\otimes\mathbb{Q}$. Since we also know how all cusp forms arise from newforms of possibly lower level (see Theorem 1.5), this allows us to compute the q-expansions of a basis for the spaces $S_k(\Gamma_1(N))$ and $S_k(N,\varepsilon)$. For precise details on how these computations work, please read [79, Chapter 9].

2.1.4 Manin symbols

If we want to do symbolic calculations with modular symbols, then the above definitions are not quite applicable since the groups of which we take quotients are not finitely generated. The *Manin symbols* enable us to give finite presentations for the spaces of modular symbols.

First we need some definitions and lemmas. For a positive integer N we define a set

$$E_N := \left\{(c,d)\in(\mathbb{Z}/N\mathbb{Z})^2 : \gcd(N,c,d)=1\right\}.$$

Define the following equivalence relation on E_N:

$$(c,d)\sim(c',d') \quad \overset{\text{def}}{\Longleftrightarrow} \quad \text{there is an } a\in(\mathbb{Z}/N\mathbb{Z})^\times \text{ such that } (c,d)=(ac',ad')$$

and the denote the quotient by P_N:

$$P_N := E_N/\sim. \tag{2.4}$$

The following lemma is easily verified:

Lemma 2.1. *Let N be a positive integer. Then the maps*

$$\Gamma_1(N)\setminus\mathrm{SL}_2(\mathbb{Z})\to E_N : \overline{\begin{pmatrix} a & b \\ c & d\end{pmatrix}}\mapsto(\bar{c},\bar{d}) \quad \textit{and}$$
$$\Gamma_0(N)\setminus\mathrm{SL}_2(\mathbb{Z})\to P_N : \overline{\begin{pmatrix} a & b \\ c & d\end{pmatrix}}\mapsto\overline{(c,d)}$$

are well-defined and bijective.

This lemma enables us to write down an explicit set of coset representatives for the orbit spaces $\Gamma_1(N)\setminus\mathrm{SL}_2(\mathbb{Z})$ and $\Gamma_0(N)\setminus\mathrm{SL}_2(\mathbb{Z})$. The following lemma provides us a first step in reducing the set of generators for the spaces of modular symbols:

Lemma 2.2. *Each space $\mathbb{M}_2(\Gamma_1(N))$ or $\mathbb{M}_2(N,\varepsilon)$ is generated by the symbols $\{a/c,b/d\}$ with $a,b,c,d\in\mathbb{Z}$ and $ad-bc=1$, where in this notation a fraction with denominator equal to zero denotes the cusp at infinity.*

Calculating the continued fraction expansion at each cusp in $\mathbb{Q}$ gives us immediately an algorithm to write a given element of $\mathbb{M}_2$ in terms of the generators in the lemma. Furthermore, note that

$$\left\{\frac{a}{c},\frac{b}{d}\right\}=\begin{pmatrix} a & b \\ c & d\end{pmatrix}\{\infty,0\},$$

so that we can write each element of $\mathbb{M}_2$ as a sum of $\gamma\{\infty,0\}$ with $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Let's consider the space $\mathbb{M}_2(\Gamma_1(N))$. As we saw, it is generated by the elements $\gamma\{\infty,0\}$ where γ runs through $\mathrm{SL}_2(\mathbb{Z})$. Now, two matrices γ define the same element this way if they are in the same coset of the quotient $\Gamma_1(N)\setminus\mathrm{SL}_2(\mathbb{Z})$. According to Lemma 2.1 such a coset can be uniquely identified with a pair $(c,d) \in (\mathbb{Z}/N\mathbb{Z})^2$. The corresponding element in $\mathbb{M}_2(\Gamma_1(N))$ is also denoted by (c,d). This element (c,d) is called a *Manin symbol*. Clearly, there are only a finite number of Manin symbols so we now know a finite set of generators for $\mathbb{M}_2(\Gamma_1(N))$.

For arbitrary k we define the Manin symbols in $\mathbb{M}_k(\Gamma_1(N))$ as the symbols of the form $P \otimes (c,d)$ where P is a monomial in $\mathbb{Z}[x,y]_{k-2}$ and (c,d) a Manin symbol in $\mathbb{M}_2(\Gamma_1(N))$. In this case as well there are finitely many Manin symbols and they generate the whole space.

In the modular symbols spaces with character ε, we have $\gamma(\alpha) = \varepsilon(\alpha)$ for $\gamma \in \Gamma_0(N)$. Now for each element of P_N we choose according to Lemma 2.1 a corresponding element $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and hence an element in $\mathbb{M}_2(N,\varepsilon)$, which we call again a Manin symbol. Note that this Manin symbol depends on the choice of γ, but because of the relation $\gamma(x) = \varepsilon(x)$ these chosen Manin symbols always form a finite set of generators for $\mathbb{M}_2(N,\varepsilon)$ as a $\mathbb{Z}[\varepsilon]$-module. Likewise, $\mathbb{M}_k(N,\varepsilon)$ is generated by elements $P \otimes (c,d)$ with P a monomial in $\mathbb{Z}[x,y]_{k-2}$ and (c,d) a Manin symbol in $\mathbb{M}_2(N,\varepsilon)$.

If we want to do symbolic calculations, then besides generators we also need to know the relations between the Manin symbols. For $\mathbb{M}_k(\Gamma_1(N))$ one can do the following.

Proposition 2.1. *Let N be a positive integer and let A be the free abelian group on the Manin symbols of the space $\mathbb{M}_k(\Gamma_1(N))$. Let $I \subset A$ be the subgroup generated by the following elements:*

$$\begin{aligned}
&P(x,y)\otimes(c,d) + P(-y,x)\otimes(-d,-c),\\
&P(x,y)\otimes(c,d) + P(-y,x-y)\otimes(-d,-c-d) + P(-x+y,-x)\otimes(-c-d,-c),\\
&P(x,y)\otimes(c,d) - P(-x,-y)\otimes(c,d),
\end{aligned}$$

where $P(x,y)\otimes(c,d)$ runs through all Manin symbols. Then $\mathbb{M}_k(\Gamma_1(N))$ is naturally isomorphic to the quotient of A/I by its torsion subgroup.

For the modular symbols spaces $\mathbb{M}_k(N,\varepsilon)$ we have a similar proposition.

Proposition 2.2. *Let N and ε be given. Let A be the free $\mathbb{Z}[\varepsilon]$-module on the Manin symbols of $\mathbb{M}_k(N,\varepsilon)$. Let $I \subset A$ be the submodule generated by the elements given in Proposition 2.1 plus for each $n \in (\mathbb{Z}/N\mathbb{Z})^\times$ the elements*

$$P(x,y)\otimes\overline{(nc,nd)} - \varepsilon(n)P(x,y)\otimes\overline{(c,d)}.$$

Then $\mathbb{M}_k(N,\varepsilon)$ is naturally isomorphic to the quotient of A/I by its torsion submodule.

These presentations enable us to perform symbolic calculations very efficiently.

A remark on the computation of Hecke operators is in order here. The formula (2.2) does not express the Hecke action on Manin symbols in terms of Manin symbols. Instead, one uses other formulas to compute Hecke operators. The following theorem, due to Merel, allows us to express Hecke operators more directly in terms of Manin symbols:

Theorem 2.1 (see [53, Theorem 2]). *On the spaces $\mathbb{M}_k(\Gamma_1(N))$ and $\mathbb{M}_k(N,\varepsilon)$ the Hecke operator T_n satisfies the following relation:*

$$T_n(P(x,y)\otimes(u,v)) = \sum_{\substack{a>b\geq 0\\ d>c\geq 0\\ ad-bc=n}}{}' P(ax+by,cx+dy)\otimes(au+cv,bu+dv),$$

where the prime in the sum notation means that terms with $\gcd(N,au+cv,bu+dv)\neq 1$ *have to be omitted.*

One would also like to express $\mathbb{S}_k(\Gamma_1(N))$ and $\mathbb{S}_k(N,\varepsilon)$ in terms of the Manin symbols. The following proposition will help us.

Proposition 2.3 (See [53, Proposition 4]). *Let integers $N\geq 1$ and $k\geq 2$ be given. Define an equivalence relation on the vector space $\mathbb{Q}[\Gamma_1(N)\setminus\mathbb{Q}^2]$ by*

$$[\overline{\lambda x}]\sim \operatorname{sign}(\lambda)^k[\overline{x}] \quad \textit{for } \lambda\in\mathbb{Q}^\times \textit{ and } x\in\mathbb{Q}^2.$$

Then the map

$$\mu : \mathbb{B}_k(\Gamma_1(N))\to\mathbb{Q}[\Gamma_1(N)\setminus\mathbb{Q}^2]/\sim$$

given by

$$\mu : P\otimes\left\{\frac{a}{b}\right\}\mapsto P(a,b)\left[\overline{\begin{pmatrix}a\\ b\end{pmatrix}}\right] \quad \textit{(a,b coprime integers)}$$

is well-defined and injective.

The vector space $\mathbb{Q}[\Gamma_1(N)\setminus\mathbb{Q}^2]/\sim$ is finite dimensional. The above proposition shows that $\mathbb{S}_k(\Gamma_1(N))$ is the kernel of $\mu\delta$, which is a map that can be computed in terms of Manin symbols. The computation of $\mathbb{S}_k(N,\varepsilon)$ can be done in a similar way, see [79, Section 8.4].

2.2 Basic numerical evaluations

In this section we will describe how to perform basic numerical evaluations, such as the evaluation of a cusp form at a point in $\mathfrak{H}$ and the evaluation of an integral of a cusp form between to points in $\mathfrak{H}^*$. Again, the paradigm will be performing actual computations.

2.2.1 Period integrals: the direct method

In this subsection we will stick to the case $k = 2$, referring to [79, Chapter 10] for a more general approach (see also [18, Section 2.10] for a treatment of $\Gamma_0(N)$). So fix a positive integer N and an $f \in S_2(\Gamma_1(N))$. Our goal is to efficiently evaluate $\langle x, f\rangle$ for $x \in \mathbb{S}_2(\Gamma_1(N))$.

Let us indicate why it suffices to look at newforms f. Because of Theorem 1.5, it suffices to look at $f = \alpha_d(f')$ with $f' \in S_k(\Gamma_1(M))$ a newform for some $M \mid N$ and $d \mid N/M$. By [72, Theorem 4.3] we have

$$\langle x, f\rangle = \langle x, \alpha_d(f')\rangle = d^{1-k}\left\langle \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} x, f' \right\rangle$$

so that computing period integrals for f reduces to computing period integrals of the newform f'.

Let us now make the important remark that for each $z \in \mathfrak{H}$ we can numerically compute $\int_\infty^z f dq/q$ by formally integrating the q-expansion of f:

$$\int_\infty^z f\frac{dq}{q} = \sum_{n\geq 1} \frac{a_n(f)}{n} q^n \quad \text{where } q = \exp(2\pi i z). \tag{2.5}$$

The radius of convergence of this series is 1 and the coefficients are small (that is, estimated by $\tilde{O}(n^{(k-3)/2})$). So if $\Im z \gg 0$ then we have $|q| \ll 1$ and the series converges rapidly. To be more concrete, for $\Im z > M$ we have $|q^n| < \exp(-2\pi M n)$ so if we want to compute $\int_\infty^z f dq/q$ to a precision of p decimals, we need to compute about $\frac{p \log 10}{2\pi M} \approx 0.37\frac{p}{M}$ terms of the series.

To compute a period integral we remark that for any $\gamma \in \Gamma_1(N)$ and any $z \in \mathfrak{H}^*$ any continuous, piecewise smooth path δ in $\mathfrak{H}^*$ from z to γz, the homology class of δ pushed forward to $X_1(N)(\mathbb{C})$ depends only on γ [51, Proposition 1.4]. Let us denote this homology class by

$$\{\infty, \gamma\infty\} \in \mathbb{S}_2(\Gamma_1(N)) \cong H_1(X_1(N)(\mathbb{C}), \mathbb{Z})$$

and remark that all elements of $H_1(X_1(N)(\mathbb{C}), \mathbb{Z})$ can be written in this way. As we also have $S_2(\Gamma_1(N)) \cong H^0(X_1(N)_\mathbb{C}, \Omega^1)$, this means we can calculate $\int_{\{\infty,\gamma\infty\}} f\frac{dq}{q}$ by choosing a smart path in $\mathfrak{H}^*$:

$$\int_\infty^{\gamma\infty} f\frac{dq}{q} = \int_z^{\gamma z} f\frac{dq}{q} = \int_\infty^{\gamma z} f\frac{dq}{q} - \int_\infty^{z} f\frac{dq}{q}.$$

If we write $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ then a good choice for z is

$$z = -\frac{d}{c} + \frac{i}{|c|}.$$

In this case we have $\Im z = \Im \gamma z = 1/|c|$ so in view of (2.5), to compute the integral to a precision of p decimals we need about $\frac{pc \log 10}{2\pi} \approx 0.37pc$ terms of the series.

Another thing we can use is the Hecke compatibility from (2.3). Put

$$W_f := \big(\mathbb{S}_2(\Gamma_1(N))/I_f\mathbb{S}_2(\Gamma_1(N))\big)\otimes\mathbb{Q},$$

where I_f is the Hecke ideal belonging to f. The space W_f has the structure of a vector space over $\mathbb{T}/I_f \cong K_f$ of dimension 2. This means that computing any period integral of f, we only need to precompute 2 period integrals. So one tries to find a K_f-basis of W_f consisting of elements $\{\infty,\gamma\infty\}$ where $\gamma\in\Gamma_1(N)$ has a very small c-entry. In practice it turns out that we do not need to search very far.

2.2.2 Period integrals: the twisted method

In this subsection we have the same set-up as in the previous subsection. There is another way of computing period integrals for $f\in S_2(\Gamma_1(N))$ which sometimes beats the method described in the previous subsection. The method described in this subsection is similar to [18, Section 2.11] and makes use of winding elements and twists.

The *winding element* of $\mathbb{M}_2(\Gamma_1(N))$ is simply defined as the element $\{\infty,0\}$ (some authors define it as $\{0,\infty\}$ but this is only a matter of sign convention). Integration over this element is easy to perform because we can break up the path in a very neat way:

$$\begin{aligned}\int_\infty^0 f\frac{dq}{q} &= \int_\infty^{i/\sqrt{N}} f\frac{dq}{q} + \int_{i/\sqrt{N}}^0 f\frac{dq}{q} = \int_\infty^{i/\sqrt{N}} f\frac{dq}{q} + \int_{i/\sqrt{N}}^\infty W_N(f)\frac{dq}{q}\\ &= \int_\infty^{i/\sqrt{N}} (f-W_N(f))\frac{dq}{q}.\end{aligned}$$

Now, choose an odd prime $\ell\nmid N$ and a primitive Dirichlet character $\chi:\mathbb{Z}\to\mathbb{C}$ of conductor ℓ. If $f\in S_k(\Gamma_1(N))$ is a newform then $f\otimes\chi$ is a newform in $S_k(\Gamma_1(N\ell^2))$, where

$$f\otimes\chi = \sum_{n\geq1} a_n(f)\chi(n)q^n.$$

The following formula to express χ as a linear combination of additive characters is well-known:

$$\chi(n) = \frac{g(\chi)}{\ell}\sum_{\nu=1}^{\ell-1}\overline{\chi}(-\nu)\exp\left(\frac{2\pi i\nu n}{\ell}\right),$$

where $g(\chi)$ is the Gauss sum of χ (see (1.7)). It follows now immediately that

$$f\otimes\chi = \frac{g(\chi)}{\ell}\sum_{\nu=1}^{\ell-1}\chi(-\nu)f\left(z+\frac{\nu}{\ell}\right) = \frac{g(\chi)}{\ell}\sum_{\nu=1}^{\ell-1}\chi(-\nu)f\left|\begin{pmatrix}\ell & \nu\\ 0 & \ell\end{pmatrix}\right.. \qquad (2.6)$$

For $f\in S_2(\Gamma_1(N))$ we now get the following useful formula for free:

$$\langle\{\infty,0\},f\otimes\chi\rangle = \frac{g(\chi)}{\ell}\left\langle\sum_{\nu=0}^{l-1}\chi(-\nu)\left\{\infty,\frac{\nu}{\ell}\right\},f\right\rangle. \qquad (2.7)$$

The element $\sum_{\nu=0}^{l-1}\chi(-\nu)\left\{\infty,\frac{\nu}{\ell}\right\}$ of $\mathbb{M}_k(\Gamma_1(N))\otimes\mathbb{Z}[\chi]$ or of some other modular symbols space where it is well-defined is called a *twisted winding element* or, more precisely the *χ-twisted winding element*. Because of formula (2.7), we can calculate the pairings of newforms in $S_2(\Gamma_1(N))$ with twisted winding elements quite efficiently as well.

We can describe the action of the Atkin-Lehner operator $W_{N\ell^2}$ on $f\otimes\chi$:

$$W_{N\ell^2}(f\otimes\chi)=\frac{g(\chi)}{g(\overline{\chi})}\varepsilon(\ell)\chi(-N)\lambda_N(f)\tilde{f}\otimes\overline{\chi},$$

where $\tilde{f}=\sum_{n\geq 1}\overline{a_n(f)}q^n$ (see for example [3, Section 3]). So in particular we have the following integral formula for a newform $f\in S_2(N,\varepsilon)$:

$$\begin{aligned}\int_{\infty}^{0} f\otimes\chi\frac{dq}{q} &= \int_{\infty}^{i/(\ell\sqrt{N})}(f\otimes\chi-W_{N\ell^2}(f\otimes\chi))\frac{dq}{q}\\ &= \int_{\infty}^{i/(\ell\sqrt{N})}\left(f\otimes\chi-\frac{g(\chi)}{g(\overline{\chi})}\chi(-N)\varepsilon(\ell)\lambda_N(f)\tilde{f}\otimes\overline{\chi}\right)\frac{dq}{q}.\end{aligned} \tag{2.8}$$

So to calculate

$$\left\langle\sum_{\nu=0}^{l-1}\chi(-\nu)\left\{\infty,\frac{\nu}{\ell}\right\},f\right\rangle$$

we need to evaluate the series (2.5) at z with $\Im z=1/(\ell\sqrt{N})$ which means that for a precision of p decimals we need about $\frac{p\ell\sqrt{N}\log 10}{2\pi}\approx 0.37p\ell\sqrt{N}$ terms of the series. In the spirit of the previous subsection, we try several ℓ and χ, as well as the untwisted winding element $\{\infty,0\}$, until we can make a K_f-basis for W_f. It follows from [71, Theorems 1 and 3] that we can always find such a basis. Also here, it turns out that in practice we do not need to search very far.

2.2.3 Computation of q-expansions at various cusps

The upper half plane $\mathfrak{H}$ is covered by neighbourhoods of the cusps. If we want to evaluate a cusp form $f\in S_k(\Gamma_1(N))$ or an integral of a cusp form at a point in such a neighbourhood then it is useful to be able to calculate the q-expansion of f at the corresponding cusp. We shall mean by this the following: A cusp a/c can be written as $\gamma\infty$ with $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbb{Z})$. Then a q-expansion of f at a/c is simply the q-expansion of $f|_k\gamma$. This notation is abusive, since it depends on the choice of γ. The q-expansion will be an element of the power series ring $\mathbb{C}[[q^{1/w}]]$ where w is the width of the cusp a/c and $q^{1/w}=\exp(2\pi iz/w)$.

If the level N is square-free this can be done symbolically. However for general N it is not known how to do this but we shall give some attempts that do at least give numerical computations of q-expansions. We use that we can compute the q-expansions of newforms in $S_k(\Gamma_1(N))$ at ∞ using modular symbols methods.

The case of square-free N

The method we present here is due to Asai [2]. Let N be square-free and let $f \in S_k(\Gamma_1(N))$ be a newform of character ε. The main reason for being able to compute q-expansions at all cusps in this case is because the group generated by $\Gamma_0(N)$ and all w_Q (see (1.18)) acts transitively on the cusps.

So let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ be given. Put

$$c' = \frac{c}{\gcd(N,c)}, \quad \text{and} \quad Q = \frac{N}{\gcd(N,c)}.$$

Let $r \in \mathbb{Z}$ be such that $d \equiv cr \bmod Q$ and define $b', d' \in \mathbb{Z}$ by

$$Qd' = d - cr \quad \text{and} \quad b' = b - ar.$$

Then we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} Qa & b' \\ Nc' & Qd' \end{pmatrix} \begin{pmatrix} Q^{-1} & rQ^{-1} \\ 0 & 1 \end{pmatrix}.$$

We know how $\left(\begin{smallmatrix} Qa & b' \\ Nc' & Qd' \end{smallmatrix}\right)$ acts on q-expansions by Theorems 1.6 and 1.8. The action of $\left(\begin{smallmatrix} Q^{-1} & rQ^{-1} \\ 0 & 1 \end{smallmatrix}\right)$ on q-expansions is simply

$$\sum_{n \geq 1} a_n q^n \mapsto Q^{1-k} \sum_{n \geq 1} a_n \zeta_Q^{rn} q^{n/Q} \quad \text{with } \zeta_Q = \exp(\tfrac{2\pi i}{Q}).$$

This shows how the q-expansion of $f|_k\gamma$ can be derived from the q-expansion of f.

Let us now explain how to do it for oldforms as well. By induction and Theorem 1.5 we may suppose $f = \alpha_p(f')$ with $p \mid N$ prime, $f' \in S_k(\Gamma_1(N/p))$ and that we know how to compute the q-expansions of f' at all the cusps. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be given. Then we have

$$f|_k\gamma = p^{1-k} f'\big|_k \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \gamma = p^{1-k} f'\big|_k \begin{pmatrix} pa & pb \\ c & d \end{pmatrix}.$$

We will now distinguish on two cases: $p \mid c$ and $p \nmid c$. If $p \mid c$ then we have a decomposition

$$\begin{pmatrix} pa & pb \\ c & d \end{pmatrix} = \begin{pmatrix} a & pb \\ c/p & d \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$$

and we know how both matrices on the right hand side act on q-expansions. If $p \nmid c$, choose b', d' with $pad' - b'c = 1$. Then we have

$$\begin{pmatrix} pa & pb \\ c & d \end{pmatrix} = \begin{pmatrix} pa & b' \\ c & d' \end{pmatrix} \beta$$

with $\beta \in \mathrm{GL}_2^+(\mathbb{Q})$ upper triangular, so also in this case we know how both matrices on the right hand side act on q-expansions.

The general case

In a discussion with Peter Bruin, the author figured out an attempt to drop the assumption that N be square-free and compute q-expansions of cusp forms numerically in this case. The idea is to generalise the W_Q operators from Subsection 1.1.7.

So let N be given. Let Q be a divisor of N and put $R = \gcd(Q, N/Q)$. Let w_Q be any matrix of the form

$$w_Q = \begin{pmatrix} RQa & b \\ RNc & Qd \end{pmatrix} \quad \text{with } a,b,c,d \in \mathbb{Z}$$

such that $\det w_Q = QR^2$ (the conditions guarantee us that such matrices do exist). One can then verify

$$\Gamma_1(NR^2) < w_Q^{-1}\Gamma_1(N)w_Q,$$

so that slashing with w_Q defines a linear map

$$S_k(\Gamma_1(N)) \oplus \overline{S}_k(\Gamma_1(N)) \xrightarrow{|w_Q} S_k(\Gamma_1(NR^2)) \oplus \overline{S}_k(\Gamma_1(NR^2))$$

which is injective since the slash operator defines a group action on the space of all functions $\mathfrak{H} \to \mathbb{C}$.

On the other hand, w_Q defines an operation on $\mathbb{M}_k$ which can be shown to induce a linear map

$$w_Q : \mathbb{S}_k(\Gamma_1(NR^2)) \otimes \mathbb{Q} \to \mathbb{S}_k(\Gamma_1(N)) \otimes \mathbb{Q}$$

that satisfies the following compatibility with respect to the integration pairing between modular symbols and cusp forms (see [72, Theorem 4.3]):

$$\langle w_Q x, f \rangle = \langle x, f|_k w_Q \rangle. \tag{2.9}$$

Let $(x_1, \ldots, x_r)$ and $(y_1, \ldots, y_s)$ be bases of $\mathbb{S}_k(\Gamma_1(N)) \otimes \mathbb{Q}$ and $\mathbb{S}_k(\Gamma_1(NR^2)) \otimes \mathbb{Q}$ respectively. Then one can write down a matrix A in terms of these basis that describes the map w_Q since we can express any symbol $P \otimes \{\alpha, \beta\}$ in terms of Manin symbols. The matrix A^t then defines the action of w_Q in terms of the bases of the cusp forms spaces that are dual to $(x_1, \ldots x_r)$ and $(y_1, \ldots, y_s)$.

Now, let $(f_1, \ldots, f_r)$ be a basis of $S_k(\Gamma_1(N)) \oplus \overline{S}_k(\Gamma_1(N))$ and let $(g_1, \ldots, g_s)$ be a basis of $S_k(\Gamma_1(NR^2)) \oplus \overline{S}_k(\Gamma_1(NR^2))$ (for instance we could take bases consisting of eigenforms for the Hecke operators away from N). Define matrices

$$B := \big(\langle x_i, f_j \rangle\big)_{i,j} \quad \text{and} \quad C := \big(\langle y_i, g_j \rangle\big)_{i,j}.$$

These can be computed numerically as the entries are period integrals. Then the matrix $C^{-1}A^tB$ describes the map $\cdot|_k w_Q$ in terms of the bases $(f_1, \ldots, f_r)$ and $(g_1, \ldots, g_s)$. Hence if we can invert C efficiently, then we can numerically compute the q-expansion of $f|_k w_Q$ with $f \in S_k(\Gamma_1(N))$.

Let now a matrix $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ be given. Put

$$c' := \gcd(N,c) \quad \text{and} \quad Q := N/c'.$$

Because of $\gcd(c/c',Q) = 1$ we can find $\alpha \in (\mathbb{Z}/Q\mathbb{Z})^\times$ with $\alpha c/c' \equiv 1 \bmod Q$. If we lift α to $(\mathbb{Z}/N\mathbb{Z})^\times$ then we have $\alpha c \equiv c' \bmod N$. Let now $d' \in \mathbb{Z}$ be a lift of αd. We have $\gcd(c',d') = \gcd(c',d',N) = 1$ so we can find $a',b' \in \mathbb{Z}$ that satisfy $a'd' - b'c' = 1$. According to Lemma 2.1, we have

$$\gamma = \gamma_0 \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \quad \text{with } \gamma_0 \in \Gamma_0(N).$$

Put $R = \gcd(c',Q)$. Then we have $\gcd(NR, Q^2Ra') = QR\gcd(c',Qa') = QR^2$ so there exist $b'',d'' \in \mathbb{Z}$ with

$$w_Q := \begin{pmatrix} QRa' & b'' \\ NR & Qd'' \end{pmatrix}$$

having determinant QR^2. One can now verify that we have $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) = w_Q\beta$ with $\beta \in \mathrm{GL}_2^+(\mathbb{Q})$ upper triangular. So in the decomposition

$$\gamma = \gamma_0 w_Q \beta$$

we can compute the slash action of all three matrices on the right hand side in terms of q-expansions, hence also of γ.

In conclusion we see that in this method we have to increase the level and go to $S_k(\Gamma_1(NR^2))$ for the square divisors R^2 of N to compute q-expansions of cusp forms in $S_k(\Gamma_1(N))$ at arbitrary cusps.

2.2.4 Numerical evaluation of cusp forms

For $f \in S_k(\Gamma_1(N))$ and a point $P \in \mathfrak{H}$ we wish to compute $f(P)$ to a high numerical precision. Before we do this let us say some words on how P should be represented. Looking at Figure 1.1 on page 2 we convince ourselves that representing P as $x+iy$ with $x,y \in \mathbb{R}$ is not a good idea, as this would be numerically very unstable when P is close to the real line. Instead, we represent P as

$$P = \gamma z \quad \text{with } \gamma \in \mathrm{SL}_2(\mathbb{Z}),\ z = x+iy,\ x \ll \infty \text{ and } y \gg 0. \tag{2.10}$$

For instance, one could demand $z \in \mathscr{F}$, although this is not strictly necessary.

So let $P = \gamma z$ be given, with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $\Im z > M$, say. Let $w = w(\gamma)$ be the width of the cusp $\gamma\infty$ with respect to $\Gamma_1(N)$. To compute $f(P)$ we make use of a q-expansion of f at $\gamma\infty$:

$$f(P) = (cz+d)^k (f|_k\gamma)(z) = (cz+d)^k \sum_{n\geq 1} a_n q^{n/w} \quad \text{where } q^{1/w} = \exp(2\pi i z/w).$$

The radius of convergence is 1 and the coefficients are small (estimated by $\tilde{O}(n^{(k-1)/2})$). So to compute $f(P)$ to a precision of p decimals we need about $\frac{pw\log 10}{2\pi M} \approx 0.37\frac{pw}{M}$ terms of the q-expansion of $f|_k\gamma$.

Of course, we have some freedom in choosing γ and z to write down P. We want to find γ such that $P = \gamma z$ with $\Im z/w(\gamma)$ as large as possible. In general, one can always write $P = \gamma z$ with $z \in \mathscr{F}$ so one obtains

$$\max_{\gamma\in \mathrm{SL}_2(\mathbb{Z})} \frac{\Im \gamma^{-1}P}{w(\gamma)} \geq \frac{\sqrt{3}}{2N}. \tag{2.11}$$

We see that in order to calculate $f(P)$ to a precision of p decimals it suffices to use about $\frac{pN\log 10}{\sqrt{3}\pi} \approx 0.42pN$ terms of the q-expansions at each cusp. Although for most points P there is a better way of writing it as γz in this respect than taking $z \in \mathscr{F}$, it seems hard to improve the bound $\frac{\sqrt{3}}{2N}$ in general.

We wish to adjust the representation sometimes from $P = \gamma z$ to $P = \gamma' z'$ where $\gamma' \in \mathrm{SL}_2(\mathbb{Z})$ is another matrix, for instance because during our calculations $\Re z$ has become too large or $\Im z$ has become too small (but still within reasonable bounds). We can make $\Re z$ smaller by putting $z' := z - n$ for appropriate $n \in \mathbb{Z}$ and putting $\gamma' := \gamma\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$. Making $\Im z$ larger is very easy as well. We want to find $\gamma'' = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\Im \gamma'' z = \frac{\Im z}{|cz+d|^2}$$

is large. But this simply means that we have to find a small vector $cz+d$ in the lattice $\mathbb{Z}z+\mathbb{Z}$, something which can be done easily if $\Re z \ll \infty$ and $\Im z \gg 0$. If c and d are not coprime we can divide both by their greatest common divisor to obtain a smaller vector. The matrix γ'' can now be completed and we put $z' := \gamma'' z$ and $\gamma' := (\gamma'')^{-1}$.

2.2.5 Numerical evaluation of integrals of cusp forms

In this subsection we will describe for $f \in S_2(\Gamma_1(N))$ and $P \in \mathfrak{H}$ how to evaluate the integral $\int_\infty^P f\,dq/q$. As in the previous subsection, we assume P to be given by means of (2.10). The path of integration will be broken into two parts: first we go from ∞ to a cusp α near P and then we go from α to P.

Integrals over paths between cusps

The pairing (2.1) gives a map

$$\Theta : \mathbb{M}_2(\Gamma_1(N)) \to \mathrm{Hom}_{\mathbb{C}}\left(S_2(\Gamma_1(N)), \mathbb{C}\right),$$

which is injective when restricted to $\mathbb{S}_2(\Gamma_1(N))$. The image of Θ is a lattice of full rank, hence the induced map

$$\mathbb{S}_2(\Gamma_1(N)) \otimes \mathbb{R} \to \mathrm{Hom}_{\mathbb{C}}\left(S_2(\Gamma_1(N)), \mathbb{C}\right)$$

is an isomorphism. In particular we obtain a map

$$\Phi : \mathbb{M}_2(\Gamma_1(N)) \to \mathbb{S}_2(\Gamma_1(N)) \otimes \mathbb{R},$$

which is an interesting map to compute if we want to calculate integrals of cusp forms along paths between cusps. The map Φ is called a *period mapping*.

The Manin-Drinfel'd theorem (see [51, Corollary 3.6] and [26, Theorem 1]) tells us that $\operatorname{im}(\Phi) \subset \mathbb{S}_2(\Gamma_1(N)) \otimes \mathbb{Q}$. This is equivalent to saying that each degree 0 divisor of $X_1(N)$ which is supported on cusps is a torsion point of $J_1(N)$. The proof given in [26] already indicates how to compute Φ with symbolic methods: let p be a prime that is 1 mod N. Then the operator $p+1-T_p$ on $\mathbb{M}_2(\Gamma_1(N))$ has its image in $\mathbb{S}_2(\Gamma_1(N))$. The same operator is invertible on $\mathbb{S}_2(\Gamma_1(N)) \otimes \mathbb{Q}$. So we simply have

$$\Phi = (p+1-T_p)^{-1}(p+1-T_p),$$

where the rightmost $p+1-T_p$ denotes the map $\mathbb{M}_2(\Gamma_1(N)) \to \mathbb{S}_2(\Gamma_1(N))$ and the leftmost $p+1-T_p$ denotes the invertible operator on $\mathbb{S}_2(\Gamma_1(N)) \otimes \mathbb{Q}$. For other methods to compute Φ, see [79, Section 10.6]. So we can express the integral of $f\,dq/q$ between any two cusps α and β in terms of period integrals, which we have already seen how to compute:

$$\int_\alpha^\beta f \frac{dq}{q} = \langle \Phi(\{\alpha, \beta\}), f \rangle.$$

Integrals over general paths

We can imitate the previous subsection pretty much. Write $P \in \mathfrak{H}$ as $P = \gamma z$ with $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\Im z / w(\gamma\infty)$ is as large as possible. Then we have

$$\int_\infty^P f \frac{dq}{q} = \int_\infty^{\gamma\infty} f \frac{dq}{q} + \int_{\gamma\infty}^{\gamma z} f \frac{dq}{q} = \int_\infty^{\gamma\infty} f \frac{dq}{q} + \int_\infty^z (f|_2\gamma) \frac{dq}{q}. \tag{2.12}$$

The integral $\int_\infty^{\gamma\infty} f \frac{dq}{q}$ is over a path between two cusps so we can compute it by the above discussion and the integral $\int_\infty^z (f|_2\gamma) \frac{dq}{q}$ can be computed using the q-expansion of $f|_2\gamma$:

$$\int_\infty^z (f|_2\gamma) \frac{dq}{q} = w \sum_{n \geq 1} \frac{a_n}{n} q^{n/w},$$

where $w = w(\gamma)$, $q^{1/w} = \exp(2\pi i z/w)$ and $f|_2\gamma = \sum a_n q^{n/w}$. Because of (2.11), computing about $\frac{pN \log 10}{\sqrt{3}\pi} \approx 0.42pN$ terms of the series should suffice to compute $\int_\infty^P f \frac{dq}{q}$ for any $P \in \mathfrak{H}$.

Note also that we can use formula (2.12) to compute the pseudo-eigenvalue $\lambda_Q(f)$ by plugging in $\gamma = w_Q$ and a z for which both $\operatorname{im} z$ and $\operatorname{im} w_Q z$ are high and for which $\int_\infty^z W_q(f) dq/q$ is not too close to zero.

2.3 Computation of modular Galois representations

In this section, we will give a short overview of the project [28] to which the research of this thesis belongs. Here we omit many details which can be found in [28]. However, we will not give precise references to sections or theorems, since at the time of writing the present section, the paper [28] is undergoing a huge revision. In the first few subsections we will explain the theoretical ideas and in Subsection 2.3.3 we will discuss how to perform actual computations.

A motivational question is: how fast can the q-coefficients of a modular form be computed? Our main example here will be the Ramanujan tau function, but we remark that most techniques that we discuss here can be generalised.

From the recurrence properties on page 6 it follows that we can compute $\tau(n)$ if we can factor n and compute $\tau(p)$ for all prime factors $p \mid n$. Also, in [4] it was shown that we can factor numbers $n = pq$ where p and q are distinct unknown primes if we can compute $\tau(n)$ and $\tau(n^2)$, provided at least one of these numbers is non-zero. The idea is as follows: put $\alpha = \tau(p)/p^{11}$ and $\beta = \tau(q)/q^{11}$. We can compute α and β because their product is $\tau(n)/n^{11}$ and their sum is $(\tau(n)^2 - \tau(n^2) - n^{11})/n^{11}$. The primes p and q can now be obtained by looking at the denominators of α and β.

Because of the above discussion, it seems reasonable to focus on computing $\tau(p)$ for p prime. A strategy for this is computing $\tau(p) \bmod \ell$ for many small primes ℓ. If the product of all these primes ℓ exceeds $4p^{11/2}$ then by the bound $|\tau(p)| \leq 2p^{11/2}$ we know exactly what $\tau(p)$ is. The main theorem of [28] is the following:

Theorem 2.2. *There exists a probabilistic algorithm that on input two prime numbers p and ℓ with $p \neq \ell$ can compute $\tau(p) \bmod \ell$ in expected time polynomial in $\log p$ and ℓ.*

Corollary 2.1. *There exists a probabilistic algorithm that on input a prime number p can compute $\tau(p)$ in expected time polynomial in $\log p$.*

2.3.1 Computing representations for $\tau(p) \bmod \ell$

We saw in Subsection 1.1.2 that for some values of ℓ, called exceptional primes, there exist simple formulas for $\tau(p) \bmod \ell$. So assume from now that ℓ is non-exceptional. We can work with the residual representations $\overline{\rho}_\ell := \overline{\rho}_{\Delta,\ell}$, see Subsections 1.3.4 and 1.3.5. For $p \neq \ell$ we have

$$\tau(p) \equiv \operatorname{tr}(\overline{\rho}(\mathrm{Frob}_p)) \bmod \ell.$$

If we put $K_\ell := \overline{\mathbb{Q}}^{\ker(\overline{\rho}_\ell)}$ then $\overline{\rho}_\ell$ factors through $\mathrm{Gal}(K_\ell/\mathbb{Q})$. Our main task is to give a polynomial P_ℓ whose splitting field is K_ℓ. Since $\operatorname{im} \rho_\ell$ acts faithfully and transitively on $\mathbb{F}_\ell^2 - \{0\}$ (remember that ℓ is non-exceptional), we will demand that P_ℓ has degree $\ell^2 - 1$ and that the number field K'_ℓ defined by P_ℓ is the subfield of K_ℓ that is fixed by the stabiliser of a point in

$\mathbb{F}_\ell^2 - \{0\}$.

We can find ρ_ℓ inside the Jacobian of $X_1(\ell)$. If $\mathbb{T} \subset \mathrm{End}(J_1(\ell))$ is the algebra generated by the diamond and Hecke operators acting on $J_1(\ell)$ then we have a homomorphism

$$\theta = \theta_{\Delta,\ell} : \mathbb{T} \to \mathbb{F}_\ell, \quad \theta : \langle d \rangle \mapsto d^{10} \bmod \ell, \quad \theta : T_n \mapsto \tau(n) \bmod \ell.$$

If $I \subset \mathbb{T}$ denotes the kernel of θ, then ρ_ℓ can be defined as $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acting on $V_\ell := J_1(\ell)(\overline{\mathbb{Q}})[I]$, which is a 2-dimensional $\mathbb{F}_\ell$-linear subspace of $J_1(\ell)(\overline{\mathbb{Q}})[\ell]$. One can express this space in terms of modular symbols since we have isomorphisms

$$J_1(\ell)(\mathbb{C})[\ell] \cong H_1(X_1(\ell)(\mathbb{C}), \mathbb{F}_\ell) \cong \mathbb{S}_2(\Gamma_1(\ell)) \otimes \mathbb{F}_\ell$$

and the action of $\mathbb{T}$ on $\mathbb{S}_2(\Gamma_1(\ell))$ can be computed.

Let g be the genus of $X_1(\ell)$. If we choose an effective divisor D of degree g on $X_1(\ell)$ then we have a morphism

$$\phi : X_1(\ell)^g \to J_1(\ell), \quad (Q_1, \ldots, Q_g) \mapsto \sum_{i=1}^{g} Q_i - D$$

which induces a birational morphism

$$\phi' : \mathrm{Sym}^g X_1(\ell) \to J_1(\ell). \tag{2.13}$$

Suppose that D is such that ϕ is étale over V_ℓ. Take a function $f \in \mathbb{Q}(X_1(\ell))$ such that for any $(Q_1, \ldots, Q_g) \in \phi^{-1}(V_\ell - \{0\})$ it has no poles at the Q_i and such that the induced map $f_* : \mathrm{Sym}^g(X_1(\ell)) \to \mathrm{Sym}^g(\mathbb{P}^1_{\mathbb{Q}})$ is injective on $\phi'^{-1}(V_\ell - \{0\})$.

The field K'_ℓ is the field of definition of a point $P \in V_\ell - \{0\}$. Put $\phi'^{-1}(P) = (Q_1, \ldots, Q_g)$. Then certainly K'_ℓ contains $e_i(f(Q_1), \ldots, f(Q_g))$ for all i, where e_i is the i-th elementary symmetric polynomial in g variables. But in fact we have an equality

$$K'_\ell = \mathbb{Q}\big(e_1(f(Q_1), \ldots, f(Q_g)), \ldots, e_g(f(Q_1), \ldots, f(Q_g))\big).$$

This can be seen as follows: the field on the right hand side, say L, is the field of definition of $f_*(\phi'^{-1}(P))$. The group $\mathrm{Gal}(\overline{\mathbb{Q}}/L)$ acts on $\mathrm{Sym}^g \mathbb{P}^1(\overline{\mathbb{Q}})$ and fixes $f_*(\phi'^{-1}(P))$. But f_* is injective on $\phi'^{-1}(V_\ell - \{0\})$ so $\mathrm{Gal}(\overline{\mathbb{Q}}/L)$ fixes P as well. So L contains, hence is equal to, K'_ℓ.

In practice, it often suffices to take $D = g \cdot [0]$ (remember from Subsection 1.2.3 that the cusp 0 is defined over $\mathbb{Q}$) and any non-constant f. The field K'_ℓ will almost always be equal to $\mathbb{Q}(f(Q_1) + \cdots + f(Q_g))$. If we assume that all of this is correct, then P_ℓ will be equal to

$$P_\ell = \prod_{P \in V_\ell - \{0\}} \big(x - \sum_i f(Q_i)\big) \quad \text{where } (Q_1, \ldots, Q_g) = \phi'^{-1}(P). \tag{2.14}$$

In theory however, to show that a good divisor D and a good function f can be found, one has to work with $X_1(5\ell)_{\mathbb{Q}(\zeta_\ell)}$ instead of $X_1(\ell)$. In this thesis, we will ignore these theoretical complications. The main reasons for this are that we want to compute actual polynomials and we want to explain ideas rather than technical details.

To compute the polynomial P_ℓ we will use numerical methods. The idea is to approximate the coefficients of P_ℓ. This could be done in several ways, for instance approximating them p-adically for one or more primes p or approximating them in $\mathbb{R}$. In [17] and [40] one can find methods to compute with modular curves over $\mathbb{F}_p$ which can be used to compute P_ℓ mod p for primes p. Note that this is a special case of p-adically approximating P_ℓ. In Subsection 2.3.3 we will describe how to approximate P_λ over the reals, in a way that is practically convenient.

Heights

If the used precision for the approximation of P_ℓ is high enough, we can compute the exact coefficients in $\mathbb{Q}$. To know how high this precision should actually be, we use *height bounds*.

Definition 2.3. Let K be a number field and take $\alpha \in K$. Then the (logarithmic) field height of α is defined as

$$\mathrm{ht}_K(\alpha) := \sum_v [K_v : \mathbb{Q}_v] \log \max(1, |\alpha|_v).$$

Here, the sum is taken over all places of K and the absolute value is normalised by demanding $|p|_v = 1/p$ for v finite lying above p and $|x|_v = |\sigma(x)|$ for v infinite belonging to the embedding $\sigma : K \hookrightarrow \mathbb{C}$. The absolute (logarithmic) height of α is defined as

$$\mathrm{ht}(\alpha) := \frac{\mathrm{ht}_K(\alpha)}{[K : \mathbb{Q}]}.$$

The absolute height of an algebraic number is independent of the number field we put around it. Also note that for a rational number p/q written in lowest terms we have $\mathrm{ht}(p/q) = \log \max(|p|, |q|)$.

Definition 2.4. Let K be a number field and consider a point $P = (\alpha_0 : \ldots : \alpha_n) \in \mathbb{P}^n(K)$. Then the (logarithmic) field height of P is defined as

$$\mathrm{ht}_K(P) := \sum_v [K_v : \mathbb{Q}_v] \log \max_i |\alpha_i|_v,$$

using the same conventions for valuations as in Definition 2.3. The absolute (logarithmic) height of P is defined as

$$\mathrm{ht}(P) := \frac{\mathrm{ht}_K(P)}{[K : \mathbb{Q}]}.$$

It is a fact that this definition is consistent in the sense that the height does not depend on the scaling of projective coordinates. Again, the absolute height of $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$ does not depend on the chosen number field. If we write $P \in \mathbb{P}^n(\mathbb{Q})$ as $(p_0 : \ldots : p_n)$ with p_i coprime integers, then $\mathrm{ht(P)} = \log \max_i |p_i|$.

For $P = a_n x^n + \cdots + a_0 \in K[x]$ with K a number field we define the height of P as the height of $(a_0 : \ldots : a_n) \in \mathbb{P}^n(K)$. If $P \in \mathbb{Q}[x]$ is an irreducible polynomial of degree d and $\alpha \in \overline{\mathbb{Q}}$ is a root of P then we have the following estimations between the height of P and the field height of α in $\mathbb{Q}(\alpha)$:

$$\mathrm{ht}(P) - d\log 2 \leq d\,\mathrm{ht}(\alpha) \leq \mathrm{ht}P + \log(d+1)/2.$$

This means that bounding the height of P_λ is equivalent with bounding the height of its roots. One can embed $X_1(\ell)$ into projective space. Bounding the roots of P_λ boils then down to bounding the Q_i occurring in formula (2.14), or rather the version of this formula that can be proven to be correct. Using a vast amount of highly non-trivial Arakelov geometry, Bas Edixhoven and Robin de Jong succeeded in bounding the Q_i and using this to show that $\mathrm{ht}(P_\ell)$ is bounded polynomially in ℓ.

Their method relies on the fact that Δ is a modular form of level one. In fact, this method works for any newform of level one. At the time of writing this section, it is not known how to produce bounds for more general levels but some progress on this is expected to be made soon.

Suppose now that a height bound for a rational number $x = p/q$ (written in lowest terms with $q > 0$) is known, say $\mathrm{ht}(x) < C$. Using non-archimedean local approximations of x one can find a large integer $M > 0$ with $\gcd(q, M) = 1$ and with $x \bmod M$ congruent to a given number a. Using real approximations, one can find a small $\varepsilon > 0$ and a ξ such that $|x - \xi| \leq \varepsilon |x| < \varepsilon \frac{\exp C}{q}$. If one doesn't use non-archimedian approximations, one can take $M = 1$ and if one doesn't use real approximations one can put $\xi = 0$ and $\varepsilon = 1$. If the approximations are close enough to satisfy $\log \frac{M}{2\varepsilon} > 2C$ then they determine the number x: suppose that $x' = p'/q'$ is another rational number satisfying the same approximation conditions as x. Then we have

$$2\varepsilon \frac{\exp(2C)}{qq'} > \varepsilon \exp(C)\left(\frac{1}{q} + \frac{1}{q'}\right) > |x - \xi| + |x' - \xi| \geq |x - x'| \geq \frac{M}{|qq'|}, \tag{2.15}$$

leading to a contradiction with $\log \frac{M}{2\varepsilon} > 2C$.

We want to actually compute x from its approximations and height bound. Note that the above reasoning is still valid if we weaken the condition $p/q \equiv a \bmod M$ to $p \equiv qa \bmod M$, dropping the assumption $\gcd(q, M) = 1$. We will change our notation a bit and assume that the approximation ξ is given in terms of a rational number $\xi = m/n$ with $n > 0$ (so typically n will be a power of 2 or 10). We thus assume

$$\left|\frac{p}{q} - \frac{m}{n}\right| < \frac{1}{2n} \tag{2.16}$$

and the condition that we need to determine p/q uniquely is

$$\log \frac{Mn}{q} > C.$$

We can use the extended Euclidean algorithm [73, Section 4.2] with $(na-m, nM)$ as input to generate a sequence of triples (q_i, r_i, s_i) satisfying $(na-m)q_i + nMr_i = s_i$ with $|s_i|$ decreasing and $|q_i r_{i+1} - q_{i+1} r_i| = 1$ for all i. Put $r = (p-qa)/M$ and $s = pn - qm$. From (2.16) it follows that the triple (q,r,s) satisfies $(na-m)q + nMr = s$ with $|s| < \frac{q}{2} < \frac{Mn}{2\exp(C)}$. By [73, Theorem 4.9] the first index i for which the bound $|s_i| \leq \lceil \frac{Mn}{2\exp(C)} \rceil - 1$ holds satisfies $|q_i| \leq \lceil \exp(C) \rceil - 1$ and $r_i/q_i = r/q$, thus also $p/q = (q_i a + Mr_i)/q_i$.

2.3.2 Computing $\tau(p)$ mod ℓ from P_ℓ

The image of $\overline{\rho}_\ell$ is a group G between $\mathrm{SL}_2(\mathbb{F}_\ell)$ and $\mathrm{GL}_2(\mathbb{F}_\ell)$. The stabiliser subgroup of a basis of $\mathbb{F}_\ell^2$ in G is trivial, so K_ℓ can be obtained by adjoining two roots of P_ℓ; make sure that the second root is not in the field generated by the first root. There are methods to compute this [45, Corollary 6]. Also, we have obtained P_ℓ from approximations in $J_1(\ell)$. From this we can deduce a bijection between the roots of P_ℓ and $V_\ell - \{0\}$ that induces an isomorphism $\mathrm{Gal}(K_\ell/\mathbb{Q}) \cong G$ which defines $\overline{\rho}_\ell$.

Let p be a prime different from ℓ. We want to compute the conjugacy class $[\mathrm{Frob}_p]$ inside $\mathrm{Gal}(K_\ell/\mathbb{Q})$. This would give us $\overline{\rho}_\ell(\mathrm{Frob}_p)$ and thus $\tau(p)$ mod ℓ. To do this, one first computes the maximal order $\mathscr{O}_{K_\ell}$ of K_ℓ [11, Theorem 1.4]. For a prime $\mathfrak{p}$ of K_ℓ above p we have that $\mathrm{Frob}_{\mathfrak{p}/p}$ is equal to the unique $\sigma \in G$ that satisfies $\sigma(\mathfrak{p}) = \mathfrak{p}$ and $\sigma(x) \equiv x^p \bmod \mathfrak{p}$ for all $x \in \mathscr{O}_{K_\ell}$. We have a decomposition

$$\mathscr{O}_{K_\ell}/(p) \cong \prod_{\mathfrak{p}|p} \mathscr{O}_{K_\ell}/\mathfrak{p}.$$

So $\mathrm{Frob}_{\mathfrak{p}/p}$ is the element that fixes $\mathscr{O}_{K_\ell}/\mathfrak{p}$ in this decomposition and acts there as $x \mapsto x^p$. To check whether $\sigma \in G$ is equal to $\mathrm{Frob}_{\mathfrak{p}/p}$ for at least one $\mathfrak{p} \mid p$ we do the following. Both σ and $x \to x^p$ are $\mathbb{F}_p$-linear maps from $\mathscr{O}_{K_\ell}/(p)$ to itself; compute them. We have $\sigma = \mathrm{Frob}_{\mathfrak{p}/p}$ if and only if the image of the map $\sigma - (x \to x^p)$ is contained in $\mathfrak{p}$. From this it follows that σ is equal to at least one of the $\mathrm{Frob}_{\mathfrak{p}/p}$ if and only if the ideal in $\mathscr{O}_{K_\ell}/(p)$ generated by the image of $\sigma - (x \to x^p)$ is not the unit ideal. So given p, we can obtain $[\mathrm{Frob}_p]$ by checking the above for all $\sigma \in G$.

2.3.3 Explicit numerical computations

Let now an arbitrary positive integer N be given and let $f \in S_2(\Gamma_1(N))$ be a newform with character ε (remember from Subsection 1.3.4 that we can reduce to the weight 2 case). Also, let ℓ be a prime number and let λ be a prime of K_f lying above ℓ. Assume that the representation $\overline{\rho}_{f,\lambda}$ is absolutely irreducible and let $\mathbb{T}$ be the Hecke algebra acting on $J_1(N)$. In Subsection 1.3.4 we saw that there is a subspace V_λ of $J_1(N)(\overline{\mathbb{Q}})[\ell]$ on which both $\mathbb{T}$ and $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ act, such that the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ defines $\overline{\rho}_{f,\lambda}$.

Approximation of torsion points

The Jacobian $J_1(N)_{\mathbb{C}}$ can be described as follows. Pick a basis $f_1,\dots,f_g$ of $S_2(\Gamma_1(N))$. Put

$$\Lambda := \left\{ \int_\gamma (f_1,\dots,f_g)\frac{dq}{q} : [\gamma] \in H_1(X_1(N)(\mathbb{C}),\mathbb{Z}) \right\} \subset \mathbb{C}^g.$$

This is a lattice in $\mathbb{C}^g$ of full rank. By the Abel-Jacobi theorem we have an isomorphism

$$J_1(N)(\mathbb{C}) \xrightarrow{\sim} \mathbb{C}^g/\Lambda, \quad \big[\sum_i([Q_i]-[R_i])\big] \mapsto \sum_i \int_{R_i}^{Q_i}(f_1,\dots,f_g)\frac{dq}{q}.$$

Let again a divisor $D=\sum_{i=1}^g [R_i]$ on $X_1(N)$ be given. Identifying $J_1(N)(\mathbb{C})$ with $\mathbb{C}^g/\Lambda$ in this way, the map (2.13) becomes a birational morphism

$$\phi' : \mathrm{Sym}^g X_1(N)(\mathbb{C}) \to \mathbb{C}^g/\Lambda, \quad (Q_1,\dots,Q_g) \mapsto \sum_{i=1}^g \int_{R_i}^{Q_i}(f_1,\dots,f_g)\frac{dq}{q}.$$

The homology group $H_1(X_1(N)(\mathbb{C}),\mathbb{Z})$ is canonically isomorphic to the modular symbols space $\mathbb{S}_2(\Gamma_1(N))$. The period lattice Λ can thus be computed numerically using the methods from Subsections 2.2.1 and 2.2.2. Since we can compute the action of $\mathbb{T}$ on $\mathbb{S}_2(\Gamma_1(N)) \cong \Lambda$, we can write down the points in $\frac{1}{\ell}\Lambda/\Lambda \subset \mathbb{C}^g/\Lambda$ that correspond to the points of V_λ. The aim is now to compute the divisors on $X_1(N)_{\mathbb{C}}$ that map to these points along ϕ'. In our computations, we assume without proof that ϕ is étale above V_λ.

We start calculating with a small precision. Let $P \in V_\lambda(\mathbb{C}) \subset \mathbb{C}^g/\Lambda$ be given. First we try out a lot of random points $Q=(Q_1,\dots,Q_g) \in X_1(N)(\mathbb{C})$. Here, each Q_i will be written as $Q_i=\gamma_i w_i$, with γ_i in a set of representatives for $\Gamma_1(N)\setminus\mathrm{SL}_2(\mathbb{Z})$ and $w_i \in \mathscr{F}$. We can compute $\phi'(Q)$ using methods from Subsection 2.2.5. We work with the point Q for which $\phi'(Q)$ is closest to P. If we in fact already know some points Q with $\phi'(Q)$ approximately equal to a point in $V_\lambda(\mathbb{C})$, then we could also take one of those points as a starting point Q to work with.

The next thing to do is adjust Q so that $\phi'(Q)$ comes closer to P. We'll make use of the Newton-Raphson approximation method. Let $\phi'' : \mathfrak{H}^g \to \mathbb{C}^g/\Lambda$ be the function defined by

$$\phi''(z_1,\dots z_g) = \phi'(\gamma_1 z_1,\dots \gamma_g z_g).$$

We observe that for a small vector $h=(h_1,\dots h_g) \in \mathbb{C}^g$ we have

$$\phi''(w_1+h_1,\dots,w_g+h_g) = \phi'(Q)+hD+O(\|h\|^2)$$

with

$$D = \left.\begin{pmatrix} \frac{\partial\phi''_1}{\partial z_1} & \cdots & \frac{\partial\phi''_g}{\partial z_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial\phi''_1}{\partial z_g} & \cdots & \frac{\partial\phi''_g}{\partial z_g} \end{pmatrix}\right|_{(w_1,\dots,w_g)}.$$

From the definition of ϕ' we can immediately deduce

$$\frac{\partial \phi_i''}{\partial z_j}(w_1, \ldots, w_g) = 2\pi i(f_i|_2\gamma_j)(w_j),$$

where we apologise for the ambiguous i. We can thus compute the matrix D using the methods of Subsection 2.2.4. Now choose a small vector $v = (v_1, \ldots, v_g) \in \mathbb{C}^g$ such that $\phi'(Q) + v$ is closer to P than $\phi'(Q)$ is. For example, v can be chosen among all vectors of a bounded length so that $\phi'(Q) + v$ is closest to P. If we write

$$h = vD^{-1},$$

then we expect $\phi''(w_1 + h_1, \ldots, w_g + h_g)$ to be approximately equal to $\phi'(Q) + v$. If this is not the case, then we try the same thing with a smaller v. It could be that this still fails, for instance because we are too close to the non-étale locus of the map ϕ. In that case, we start with a new random point Q.

We repeat the above adjustments until we are (almost) as close as we can get, considering our calculation precision. It might happen that the w_i become too wild, i.e. $|\Re w_i|$ becomes too large or $\Im w_i$ becomes too small. If this is the case we adjust the way we write Q_i as $\gamma_i w_i$ using the method described in Subsection 2.2.4. We can always replace the γ_i then by a small matrix in the same coset of $\Gamma_1(N) \setminus \mathrm{SL}_2(\mathbb{Z})$.

Once we have for each $P \in V_\lambda$ a point Q such that $\phi'(Q)$ is approximately equal to P, we can start increasing the precision. We double our calculation precision and repeat the above adjustments ($\phi'(Q) + v$ will in this case be equal to P). We repeat this a few times until we have very good approximations.

Computation of polynomials

Now, we have to choose a function in $h \in \mathbb{Q}(X_1(N))$. Since h multiplies heights of points roughly by $\deg(h)$, we want to find a function of small degree. Take any k and a basis $h_1, \ldots, h_n$ of $S_k(\Gamma_1(N))$ such that the q-expansions of the h_i lie in $\mathbb{Z}[[q]]$ and such that the exponents of the first non-zero terms of these q-expansions form a strictly increasing sequence. We propose to use $h = W_N(h_{n-1})/W_N(h_n)$ as a function to use (assuming $n \geq 2$). Remember from Subsection 1.2.4 that $S_k(\Gamma_1(N))$ is the space of global sections of the line bundle $\mathscr{L} = \omega^{\otimes k}(-\text{cusps})$ on $X_1(N)$, base changed to $\mathbb{C}$. Remember also that the cusp ∞ is not defined over $\mathbb{Q}$, but the cusp 0 is. Since we demand the q-expansions to have rational coefficients, the sections $W_N(h_1), \ldots, W_N(h_n)$ are defined over $\mathbb{Q}$ and they have increasing order at 0. One can now verify that for $h = W_N(h_{n-1})/W_N(h_n)$ we have

$$\deg(h) \leq \deg(\mathscr{L}) - \mathrm{ord}_\infty(h_{n-1}) \leq \deg(\mathscr{L}) - \dim H^0(\mathscr{L}) + 2 \leq g + 1.$$

For $k = 2$ and $g \geq 2$ we have $\mathscr{L} \cong \Omega^1(X_1(N))$ and we get g as an upper bound for $\deg(h)$. Using methods from Subsection 2.2.4, we can evaluate h numerically. The author is not aware of a sophisticated method for finding a function $h \in \mathbb{Q}(X_1(N))$ of minimal degree in general; this minimal degree is called the *gonality* of the curve $X_1(N)$. Published results on these matters seem to either be limited to $X_0(N)$ or to concern only *lower* bounds for the gonality of modular curves, see for example [1], [5, Chapter 3] or [60].

Now put

$$\alpha_P = \sum_{i=1}^{g} h(Q_i), \quad \text{for } P \in V_\lambda(\mathbb{C}) - \{0\} \text{ and where } \phi'(Q_1, \ldots, Q_g) = P.$$

We work out the product in

$$P_\lambda(x) := \prod_{P \in V_\lambda(\mathbb{C}) - \{0\}} (x - \alpha_i) = \sum_{k=0}^{n} a_k x^k, \quad \text{where } n = \deg P_\lambda.$$

The coefficients a_k are rational numbers that we have computed numerically. Since the height of P_λ is expected to be not too large, the denominators of the a_k should have a relative small common denominator. The LLL algorithm can be used to compute integers $p_0, \ldots, p_{n-1}, q$ such that $|p_k - a_k q|$ is small for all k, see [49, Proposition 1.39]. If the sequence (a_k) is arbitrary, then we'll be able to find p_k and q such that $|p_k - a_k q|$ is roughly of order $q^{-1/n}$ for each k, but not much better than that. So if it happens that we find p_k and q with $|p_k - a_k q|$ much smaller than $q^{-1/n}$ for all k, then we guess that a_k is equal to p_k/q. If we cannot find such p_k and q then we will double the precision and repeat all the calculations described above.

Heuristically, the calculation precision that is needed to find the true value of a_k is about $(1 + 1/n)\,\mathrm{ht}(P_\lambda)/\log(10)$ decimals. Another way of finding rational approximations of the a_k is by approximating them using continued fractions. For this method, the precision needed to find the true value of a_k would be about $2\,\mathrm{ht}(P_\lambda)/\log(10)$ decimals.

Since the degree of P_λ will be quite large, we won't be able to do many further calculations with it. In particular it may be hard to verify whether all the guesses we made were indeed correct. Instead, we will look at the following variant. If $\mathfrak{m}$ is the Hecke ideal of f mod λ, then V_λ is a vector space over $\mathbb{T}/\mathfrak{m}$. The representation $\overline{\rho}_{f,\lambda}$ induces an action $\tilde{\rho}_\lambda$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the set $\mathbb{P}(V_\lambda)$ of lines in V_λ. We can attach a polynomial $\tilde{P}_\lambda$ to this projectivised representation $\tilde{\rho}_\lambda$, analogously to the way this was done for $\overline{\rho}$. This polynomial will have smaller a degree than P_λ. We put

$$\tilde{P}_\lambda(x) = \prod_{L \in \mathbb{P}(V_\ell)} \big(x - \sum_{P \in L - \{0\}} \alpha_P\big) = \sum_{k=0}^{m} b_k x^k, \quad \text{where } m = \deg \tilde{P}_\lambda.$$

As above, if the calculation precision is sufficient we can use lattice reduction algorithms to compute the exact values of the b_k.

Reduction of polynomials

Although the polynomial $\tilde{P}$ will not have a very huge height, its height is still too large to do any useful computations with it. The first step in making a polynomial of smaller height defining the same number field is computing the maximal order of that number field. Let q be the common denominator of the coefficients and put $p_k = b_k q$. Consider the polynomial

$$Q(x) = q \cdot \tilde{P}_\lambda(x) = qx^m + p_{m-1}x^{m-1} + \cdots + p_0.$$

We make ourselves confident that we correctly computed $Q(x)$ (although we won't prove anything at this point yet). For instance, we verify that $Q(x)$ is irreducible and that its discriminant has the prime factors of $N\ell$ in it. We can also compute for several primes p not dividing $\mathrm{Disc}(Q(x))$ the decomposition type of $Q(x) \bmod p$ and verify that it could be equal to the cycle type of $\tilde{\rho}(\mathrm{Frob}_p)$. If not, we again double the precision and repeat the above calculations.

Let now α be a root of $\tilde{P}_\lambda(x)$ and write down the order

$$\mathscr{O} := \mathbb{Z} + \sum_{k=1}^{m-1} \left(\mathbb{Z} \cdot \sum_{j=0}^{k-1} a_{m-j}\alpha^{k-j} \right),$$

which is an order that is closer to the maximal order than $\mathbb{Z}[q\alpha]$ (see [48, Subsection 2.10]). Being confident in the correctness of $Q(x)$, we know where the number field K defined by it ramifies and thus we can compute its maximal order (see [11, Section 6 and Theorems 1.1 and 1.4]). Having done this, we embed $\mathscr{O}_K$ as a lattice into $\mathbb{C}^m$ in the usual way and we use the LLL algorithm to compute a basis of small vectors in $\mathscr{O}_K$. We can then search for an element of small length in $\mathscr{O}_K$ that generates K over $\mathbb{Q}$. Its defining polynomial $\tilde{P}'_\lambda$ will have small coefficients. See also [16].

In the computation of the polynomials P_λ and $\tilde{P}'_\lambda$ we made several guesses and assumptions that we cannot prove to be correct. In Chapters 3 and 4, we work out in special cases how we can use established parts of Serre's conjecture to prove afterwards for polynomials of the style $\tilde{P}'_\lambda$ that they indeed belong to the modular Galois representations that we claim they belong to. In the unlikely case that such tests may fail we can of course make adjustments like choosing another function h or another divisor D.

Further refinements

The Jacobian $J_1(N)$ has large dimension (for N prime it is $(N-5)(N-7)/24$). It could be that our newform f is an element of $S_2(\Gamma)$ with $\Gamma_1(N) \lneq \Gamma < \Gamma_0(N)$. In that case we work with the curve X_Γ, which is given its $\mathbb{Q}$-structure by defining it as a quotient of $X_1(N)$. The Jacobian J_Γ of X_Γ is isogenous to an abelian subvariety of $J_1(N)$ that contains V_λ, so this works perfectly well.

In the case $\Gamma = \Gamma_0(N)$ we can sometimes go a step further. The operator W_N on $X_0(N)$ is defined over $\mathbb{Q}$. If f is invariant under W_N, one can work with the curve $X_0^+(N) := X_0(N)/\langle W_N \rangle$. Its Jacobian $J_0^+(N)$ is isogenous to an abelian subvariety of $J_1(N)$ that contains V_λ, so also here it works. Some words on the computation of the homology of $X_0^+(N)$ are in order. The action of W_N on $X_0(N)$ induces an action on $H_1(X_0(N)(\mathbb{C}), \mathbb{Z})$ and on

$H_1(X_0(N)(\mathbb{C}), \text{cusps}, \mathbb{Z})$. Since paths between cusps on $X_0^+(N)(\mathbb{C})$ lift to paths between cusps on $X_0(N)(\mathbb{C})$ we have a surjection

$$H_1(X_0(N), \text{cusps}, \mathbb{Z}) \twoheadrightarrow H_1(X_0^+(N)(\mathbb{C}), \text{cusps}, \mathbb{Z}).$$

The kernel of this surjection consists of the elements $[\gamma] \in H_1(X_0(N), \text{cusps}, \mathbb{Z})$ satisfying $W_N([\gamma]) = -[\gamma]$. So modular symbols methods allow us to compute $H_1(X_0^+(N)(\mathbb{C}), \text{cusps}, \mathbb{Z})$ as a quotient of $\mathbb{M}_2(\Gamma_0(N))$. Let $\mathbb{B}_2^+(\Gamma_0(N))$ be the free abelian group on the cusps of $X_0^+(N)(\mathbb{C})$ and define

$$\delta : H_1(X_0^+(N)(\mathbb{C}), \text{cusps}, \mathbb{Z}) \to \mathbb{B}_2^+(\Gamma_0(N)), \quad \{\alpha, \beta\} \mapsto \{\beta\} - \{\alpha\}.$$

Then $H_1(X_0^+(N)(\mathbb{C})) = \ker(\delta)$.