Modular Forms of Weight One Over Finite Fields

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D. D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op vrijdag 2 september 2005
klokke 14.15 uur
door

Gabor Jürgen Wiese

geboren te Warendorf, Duitsland,
in 1976.

Samenstelling van de promotiecommissie:

promotor:	prof. dr. S. J. Edixhoven
referent:	prof. dr. K. Buzzard (Imperial College London)
overige leden:	prof. dr. H. W. Lenstra prof. dr. L. Merel (Université Paris 7) dr. B. de Smit prof. dr. J. Top (Rijksuniversiteit Groningen) prof. dr. S. M. Verduyn Lunel

Modular Forms of Weight One Over Finite Fields

(Modulaire vormen van gewicht één over eindige lichamen)

Gabor Wiese
Mathematisch Instituut
Universiteit Leiden
Postbus 9512
2300 RA Leiden
The Netherlands
gabor@math.leidenuniv.nl
http://www.math.leidenuniv.nl/∼gabor/

THOMAS STIELTJES INSTITUTE
FOR MATHEMATICS

Contents

Introduction

The absolute Galois group $G_{\mathbb{Q}}$ of the field of rational numbers is arguably the central object of algebraic number theory, as it governs all number fields and their arithmetic. However, its structure remains very mysterious. A natural approach is to study its linear representations, i.e. continuous homomorphisms $G_{\mathbb{Q}} \to \mathrm{GL}_n(K)$ for some integer $n \geq 1$, where K is a topological field. Among other things, the Langlands program describes the case of complex representations, i.e. those with $K = \mathbb{C}$, via automorphic representations. Only for $n = 1$ all complex Galois representations are known explicitly, as they are described by the Kronecker-Weber theorem resp. by class field theory, when $\mathbb{Q}$ is replaced by an arbitrary number field.

One of the aims of this thesis is to study and develop methods for computing explicitly with odd semi-simple continuous representations of dimension $n = 2$ over $\overline{\mathbb{F}_p}$ for a prime p, i.e.

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$$

where $\overline{\mathbb{F}_p}$ is equipped with the discrete topology. Odd means that the image of any complex conjugation has determinant -1. For both complex representations and representations over $\overline{\mathbb{F}_p}$ continuity implies that the image is a finite group. However, in $\mathrm{GL}_2(\mathbb{C})$ there are relatively few finite subgroups up to conjugation, whereas the theory is much richer over $\overline{\mathbb{F}_p}$.

Odd semi-simple 2-dimensional continuous Galois representations over $\overline{\mathbb{F}_p}$ arise from certain modular forms by a theorem of Deligne, Deligne-Serre and Shimura. The arithmetic of such a modular representation is closely connected with the coefficients of the modular form it comes from. A conjecture by Serre (henceforth simply the *Serre conjecture*) claims a converse, namely, that the irreducible among those representations can be obtained from precisely described modular forms. Thus, *the irreducible odd 2-dimensional Galois representations with coefficients in $\overline{\mathbb{F}_p}$ are believed to be completely governed by modular forms.* As modular forms are very accessible for explicit computations, the Serre conjecture provides us also with a tentative computational approach to all such 2-dimensional representations of $G_{\mathbb{Q}}$ over $\overline{\mathbb{F}_p}$.

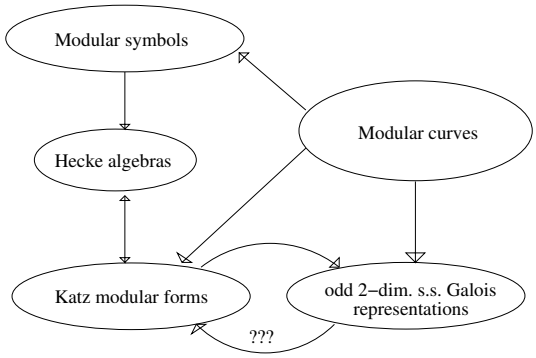
The modular forms used in the original version of the Serre conjecture were classical Hecke eigenforms, that is, they are holomorphic functions from the upper half plane to the complex numbers, satisfying certain transformation and growth properties and they are eigen-

forms for the so-called Hecke operators. These conditions imply that after a suitable normalisation these forms have a Fourier series at $i\infty$ of the form $q + \sum_{n\geq 2} a_n q^n$ with $q = e^{2\pi i\tau}$, where the a_n are algebraic integers. The associated Galois representation over $\overline{\mathbb{F}_p}$ only depends on the reduction of the a_n modulo a chosen prime above p. So, it is natural in the context of the Serre conjecture to try to define modular forms directly over finite fields.

A good theory of modular forms over any ring in which the level is invertible was set up by Katz in terms of the algebraic geometry of modular curves. It is this theory that we will be using in this thesis. For weight at least 2 when working with the group $\Gamma_1(N)$ for $N \geq 5$ the Katz forms over $\overline{\mathbb{F}_p}$ coincide with the reductions of the forms described in the previous paragraph. The case of forms of weight one, however, plays a special rôle, as then the Katz theory is much richer than the classical one. One can extend the Serre conjecture to include weight one Katz forms over $\overline{\mathbb{F}_p}$, which ought to correspond to Galois representations unramified at p. This aspect was discussed by Edixhoven in [EdixWeight].

In view of their number theoretic significance it is essential to be able to compute (Katz) modular forms over finite fields explicitly. One aim of this thesis is to establish methods for computing the associated Hecke algebra with fast methods, preferably in terms of linear algebra over finite fields. Using work by Eichler and Shimura, one can compute classical modular forms of weight at least 2 with linear algebra methods over the integers by using integral modular symbols or integral group cohomology. Hence, reduction modulo a prime above p yields a method for computing (Katz) modular forms over $\overline{\mathbb{F}_p}$. However, the theory of modular symbols and group cohomology also makes sense over $\mathbb{F}_p$. So, a natural question to ask is whether one can compute modular forms over $\overline{\mathbb{F}_p}$ directly with linear algebra methods over $\mathbb{F}_p$. More precisely, the question arises in which cases the Hecke algebra over $\mathbb{F}_p$ of (Katz) modular forms over $\overline{\mathbb{F}_p}$ coincides with the one of modular symbols over $\mathbb{F}_p$.

The relationships between the objects described is illustrated in the figure. The modular curves can be seen as the unifying element of the objects concerned. Considering the modular curves as Riemann surfaces, analytic cohomology for a certain sheaf gives rise to the modular symbols. The étale cohomology of the modular curve over $\overline{\mathbb{Q}}$ for a similarly defined étale sheaf leads to the Galois representation. Finally, global sections of a certain invertible sheaf on the modular curve base changed to $\mathbb{F}_p$ yield the Katz modular forms over $\overline{\mathbb{F}_p}$.



We now give an overview of the thesis and mention important results.

Chapter I is a reprinting of the article "Dihedral Galois Representations and Katz Modular Forms" ([W-Dih]). In that article we prove the extended form of the Serre conjecture for

dihedral Galois representations. More precisely, the principal result is the following theorem (cf. Theorem (1.1.1)).

Theorem. *Let p be a prime and $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ an irreducible odd Galois representation such that the image of $G_{\mathbb{Q}} \xrightarrow{\rho} \mathrm{GL}_2(\overline{\mathbb{F}_p}) \overset{\mathrm{proj}}{\twoheadrightarrow} \mathrm{PGL}_2(\overline{\mathbb{F}_p})$ is a dihedral group D_n for some n. As in [Serre1] define N_ρ to be the conductor of ρ and ϵ_ρ to be the prime-to-p part of $\det \circ \rho$ (that is the restriction to $(\mathbb{Z}/N_\rho\mathbb{Z})^*$ when $\det \circ \rho$ is considered as a character of $(\mathbb{Z}/(N_\rho p)\mathbb{Z})^*$). Define the minimal weight $k(\rho)$ as in [EdixWeight].*

Then there exists a normalised Katz eigenform $f \in S_{k(\rho)}(\Gamma_1(N_\rho), \epsilon, \overline{\mathbb{F}_p})$ (i.e. it has level N_ρ, weight $k(\rho)$ and character ϵ_ρ) such that its associated Galois representation $\rho_f : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ is isomorphic to ρ.

The modularity of dihedral representations was apparently already known to Hecke, at least for $p > 2$. So the question is whether the weight and the level of the modular form can be chosen as predicted. For modular, irreducible, but not necessarily dihedral representations this is known if $p \geq 3$ by the work of many mathematicians, but for $p = 2$ there are open *exceptional* cases. Our result hence shows that this is also true for $p = 2$, at least when the representation is dihedral. The proof relies on the use of Katz modular forms and does not work when one only uses reductions of holomorphic modular forms.

Chapters II, III and IV concern the computation of the Hecke algebra of Katz modular forms over finite fields. In other words, we need a faithful module for that Hecke algebra which can be easily described and calculated. The one used in the Magma implementations is the module of modular symbols, but also a certain group cohomology group can be employed.

In Chapter II we study this group cohomology group and modular symbols (for their definition see (2.5.1)) by relating them to certain cohomology groups of modular curves. From a geometric point of view the cohomology groups of modular curves are the natural object to consider. However, they are a priori not very accessible. But for modular curves that are obtained as quotients of the upper half plane by groups like $\Gamma = \Gamma_1(N)$ with $N \geq 5$, they agree with certain group cohomology groups for Γ, which have an elementary description. For more general groups $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ there are differences. The usefulness of modular symbols (or rather the modular symbols formalism) stems from the fact that a good implementation (in Magma by William Stein) exists. Besides the modular curves, which we consider as Riemann surfaces, we also use - in slight generalisation - analytic stacks, called modular stacks. The latter notion only differs from the former, when the modular curve in question is obtained as the quotient of the upper half plane by a non-freely acting subgroup $\Gamma \leq \mathrm{PSL}_2(\mathbb{Z})$ of finite index. In the stack setting the cohomology groups under consideration also naturally arise in the theory of group cohomology, whereas for modular curves geometric methods such as Poincaré duality are available. Using both points of view allows us to establish an explicit description of the first cohomology group of any modular curve and the push-forward

of any locally constant sheaf of R-modules on the modular stack for an arbitrary ring R (cf. Theorem (2.4.6)).

Theorem. *For any ring R, any congruence subgroup $\Gamma \leq \mathrm{PSL}_2(\mathbb{Z})$ and any $R[\Gamma]$-module V with associated locally constant sheaf $\mathbb{V}$ on the analytic stack $[\Gamma\backslash\mathbb{H}]$, we have*

$$H^1(\Gamma\backslash\mathbb{H}, \pi_*\mathbb{V}) \cong M/\big(M^{\langle\sigma\rangle} + M^{\langle\tau\rangle}\big)$$

with $M = \mathrm{Coind}_{\Gamma}^{\mathrm{PSL}_2(\mathbb{Z})}(V)$, $\sigma = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $\tau = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and π the natural projection map from the stack $[\Gamma\backslash\mathbb{H}]$ to the modular curve $\Gamma\backslash\mathbb{H}$, seen as a Riemann surface.

We can precisely describe the difference between the objects in question, which yields the following criterion for them to be equal (cf. Theorem (2.6.1)). For the precise definitions see Chapter II.

Theorem. *Let R be a ring, $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup and $k \geq 2$ an integer. Suppose that the orders of all stabilisors for the action of $\Gamma/\Gamma \cap \langle -\mathbf{1}\rangle$ on the upper half plane $\mathbb{H}$ are invertible in R.*

Then the module of modular symbols over R for Γ of weight k is isomorphic with the group cohomology group over R for Γ of weight k and the cohomology group over R of weight k of the modular curve $\Gamma\backslash\mathbb{H}$. Similar results also hold for the respective parabolic and the boundary subspaces.

We are also able to describe the torsion of the modules in question over the integers (see Proposition (2.4.8)). Finally, a study of these objects for $\Gamma_1(N)$ as a $(\mathbb{Z}/N\mathbb{Z})^*$-module is carried out, which will be necessary in order to pass to characters in Chapter III.

The principal aim of Chapter III is to compare the Hecke algebra of modular forms over $\mathbb{F}_p$ for $\Gamma_1(N)$ with $p \nmid N$ to the Hecke algebra defined on the parabolic group cohomology group $H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))$, where $V_{k-2}(\mathbb{F}_p)$ is the $\mathbb{F}_p[\Gamma_1(N)]$-module of homogeneous polynomials of degree $k-2$ in two variables. The main idea is to work in weight 2 with level Np which forces us to restrict to weights $2 \leq k \leq p+1$.

We introduce the following notation. Let M be any $\mathbb{F}_p$-vector space on which the Hecke operators T_l and the p-part of the diamond operators $\langle\cdot\rangle_p$ act. By $M[k-2]$ we mean M with the action of the Hecke operator T_l "twisted" to be $l^{k-2}T_l$ (in particular T_p acts as zero). Furthermore, by $M(k-2)$ be denote the subspace on which $\langle l\rangle_p$ acts as l^{k-2}.

For $3 \leq k \leq p$ there is the following proposition by Serre (cf. Proposition (3.3.8)).

Proposition. (Serre) *Let p be a prime, $N \geq 5$ and $3 \leq k \leq p$ integers such that $p \nmid N$. Moreover, let L denote the $\mathbb{Z}_p[\zeta_p]$-module consisting of the modular forms in $S_2(\Gamma_1(Np), \mathbb{Q}_p(\zeta_p))$ all of whose q-expansions are integral. Let $\overline{L} = L \otimes \mathbb{F}_p$.*

Then there is an isomorphism

$$\overline{L}(k-2) \cong S_k(\Gamma_1(N), \mathbb{F}_p) \oplus S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)[k-2],$$

which respects the Hecke action.

We establish a parallel result on group cohomology (cf. Proposition (3.2.5)), which for the non-parabolic spaces is already present in [Ash-Stevens].

Proposition. *Let p be a prime, $N \geq 5$ and $3 \leq k \leq p$ integers such that $p \nmid N$.*
We have the exact sequence

$$0 \to H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p)) \to H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p)(k-2) \\ \to H^1_{\mathrm{par}}(\Gamma_1(N), V_{p+3-k-2}(\mathbb{F}_p))[k-2] \to 0,$$

in which the Hecke action is respected.

Via the Jacobian one can obtain a connection between Katz modular forms over $\mathbb{F}_p$ and the corresponding group cohomology group, following the strategy of the proof of [EdixJussieu], Theorem 5.2. In that way we are able to prove the following result (cf. Corollary (3.3.14)).

Theorem. *Let p be a prime, $N \geq 5$ and $k \in \{2, \ldots, p+1\}$ integers such that $p \nmid N$. Let $\mathfrak{P}$ be a maximal ideal of the $\mathbb{F}_p$-Hecke algebra $\mathbb{T}$ of $S_k(\Gamma_1(N), \mathbb{F}_p)$ corresponding to a normalised cuspidal eigenform f which is* ordinary*, i.e. the p-th coefficient $a_p(f)$ of the standard q-expansion of f is non-zero.*
Then $H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))_{\mathfrak{P}}$ is a faithful module for $\mathbb{T}_{\mathfrak{P}}$.

Studying $S_k(\Gamma_1(N), \overline{\mathbb{F}_p})$ as a $(\mathbb{Z}/N\mathbb{Z})^*$-module this result can be extended to characters (cf. Proposition (3.3.20)). It should be mentioned that methods from p-adic Hodge theory (cf. Corollary (3.3.7) and [EdixJussieu], Theorem 5.2) show that the ordinariness assumption is not necessary when $k < p$.

In Chapter IV we explain how the methods from Chapters II and III can be used algorithmically. Using a method from [EdixJussieu] we obtain the following corollary of the case $k = p$ of the preceding theorem (cf. Corollary (4.5.5)).

Corollary. *The Hecke algebra of weight one Katz modular forms for $\Gamma_1(N)$ over $\overline{\mathbb{F}_p}$ with $p \nmid N$ can be computed using cuspidal modular symbols over $\mathbb{F}_p$.*

Chapter V reports on computer calculations performed with the algorithms from Chapter IV. One result is the following (cf. Theorem (5.1.1)).

Theorem. *All groups $\mathrm{SL}_2(\mathbb{F}_{2^r})$ occur as Galois groups over $\mathbb{Q}$ for r from 1 up to 77.*

This extends computations by Mestre, who covered $r \leq 16$.

Chapters I and V are independent of any other chapter. Chapters II, III and IV build on each other.

Notations and Conventions

Let R be a ring which is commutative and has a unit element. All base rings in this thesis are assumed to satisfy these properties.

If M is a left $R[G]$-module for a group G, we denote the *(left) coinvariants* by

$$_GM = M/I_GM,$$

with the *augmentation ideal* I_G defined by the exact sequence

$$0 \to I_G \to R[G] \xrightarrow{g\mapsto 1} R \to 0.$$

The augmentation ideal is the ideal of $R[G]$ generated by all elements of the form $(1-g)$ for $g \in G$. If M is a right $R[G]$-module, we denote the *(right) coinvariants* by

$$M_G = M/MI_G.$$

For the *right* resp. *left invariants* we use the notation M^G resp. ^GM.

If g is an element of finite order n in G, we define the *norm of* g as the element $N_g = 1 + g + \cdots + g^{n-1}$ in $R[G]$. Similarly, if G is a finite group we mean by N_G the formal sum over the group elements of G inside $R[G]$.

If ϕ is an endomorphism of M, respecting the submodule $N \subseteq M$, the notation $\ker_N(\phi)$ means the kernel of ϕ considered as an endomorphism of N.

We let $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$ denote the monoid of 2×2-matrices with entries in $\mathbb{Z}$ and non-zero determinant. We have the following important matrices in $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$:

$$T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right), \quad \sigma = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right), \quad \tau := T\sigma = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right),$$
$$\tau^2 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right), \quad T' = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right), \quad \eta = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right).$$

For a 2×2-matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ over a ring R one defines *Shimura's main involution*

$$M^\iota = \mathrm{Tr}(M) - M = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right).$$

If M has invertible determinant, we have $M^\iota = M^{-1}\det(M)$. The matrix M^ι is also called the *adjoint matrix*. Moreover, we have the identity $M^\iota = (\sigma^{-1}M\sigma)^\top$.

We consider the standard subgroups $\Gamma(N)$, $\Gamma_1(N)$ and $\Gamma_0(N)$ of $\mathrm{SL}_2(\mathbb{Z})$ consisting of those matrices in $\mathrm{SL}_2(\mathbb{Z})$ which reduce to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ resp. to $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ resp. to $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)$ modulo N.

If G is a subgroup of $\mathrm{SL}_2(\mathbb{Z})$, we denote by $\overline{G} = G/(\langle -\mathbf{1}\rangle \cap G)$ the corresponding subgroup of $\mathrm{PSL}_2(\mathbb{Z})$.

If $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup, i.e. contains some $\Gamma(N)$, then throughout this thesis the notation $S_k(\Gamma, R)$ means Katz modular forms of weight k for the group Γ over the $\mathbb{Z}[1/N]$-algebra R (see e.g. [EdixBoston]). A similar notation is used with a character.

Chapter I

Dihedral Galois Representations and Katz Modular Forms

This chapter has appeared as [W-Dih]. All changes to the published version are indicated by footnotes. The notation slightly differs from the one used in the other chapters of this thesis.

We show that any two-dimensional odd dihedral representation ρ over a finite field of characteristic $p > 0$ of the absolute Galois group of the rational numbers can be obtained from a Katz modular form of level N, character ϵ and weight k, where N is the conductor, ϵ is the prime-to-p part of the determinant and k is the so-called minimal weight of ρ. In particular, $k = 1$ if and only if ρ is unramified at p. Direct arguments are used in the exceptional cases, where general results on weight and level lowering are not available.

1.1. Introduction

In [Serre1] Serre conjectured that any odd irreducible continuous Galois representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ for a prime p comes from a modular form in characteristic p of a certain level N_ρ, weight $k_\rho \geq 2$ and character ϵ_ρ. Later Edixhoven discussed in [EdixWeight] a slightly modified definition of weight, the so-called *minimal weight*, denoted $k(\rho)$, by invoking Katz' theory of modular forms. In particular, one has that $k(\rho) = 1$ if and only if ρ is unramified at p.

The present note contains a proof of this conjecture for *dihedral representations*. We define those to be the continuous irreducible Galois representations that are induced from a character of the absolute Galois group of a quadratic number field. Let us mention that this is equivalent to imposing that the representation is irreducible and its projective image is

isomorphic to a dihedral group D_n for some n.[1]

(1.1.1) Theorem. *Let p be a prime and $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ an odd dihedral representation. As in [Serre1] define N_ρ to be the conductor of ρ and ϵ_ρ to be the prime-to-p part of $\det \circ \rho$ (considered as a character of $(\mathbb{Z}/(N_\rho p)\mathbb{Z})^*$)[2]. Define $k(\rho)$ as in [EdixWeight].*

Then there exists a normalised Katz eigenform $f \in \mathcal{S}_{k(\rho)}(\Gamma_1(N_\rho), \epsilon_\rho, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$, whose associated Galois representation ρ_f is isomorphic to ρ.

We will on the one hand show directly that ρ comes from a Katz modular form of level N_ρ, character ϵ_ρ and minimal weight $k(\rho) = 1$, if ρ is unramified at p. If on the other hand ρ is ramified at p, we will finish the proof by applying the fundamental work by Ribet, Edixhoven, Diamond, Buzzard and others on "weight and level lowering" (see Theorem (1.4.2)).

Let us recall that in weight at least 2 every Katz modular form on ${\Gamma_1}^3$ is classical, i.e. a reduction from a characteristic zero form of the same level and weight. Hence multiplying by the Hasse invariant, if necessary, it follows from Theorem (1.1.1) that every odd dihedral representation as above also comes from a classical modular form of level N_ρ and Serre's weight k_ρ. However, if one also wants the character to be ϵ_ρ, one has to exclude in case $p = 2$ that ρ is induced from $\mathbb{Q}(i)$ and in case $p = 3$ that ρ is induced from $\mathbb{Q}(\sqrt{-3})$ (see [Buzzard], Corollary 2.7, and [Diamond], Corollary 1.2).

Edixhoven's theorem on weight lowering ([EdixWeight], Theorem 4.5) states that modularity in level N_ρ and the modified weight $k(\rho)$ follows from modularity in level N_ρ and Serre's weight k_ρ, unless one is in a so-called *exceptional case*. A representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ is called *exceptional* if the semi-simplification of its restriction to a decomposition group at p is the sum of two copies of an unramified character. Because of work by Coleman and Voloch the only open case left is that of characteristic 2 (see the introduction of [EdixWeight]).

Exceptionality at 2 is a common phenomenon for mod 2 dihedral representations. One way to construct examples is to consider the Hilbert class field H of a quadratic field K that is unramified at 2 and has a non-trivial class group. One lets ρ_K be the dihedral representation obtained by induction to $G_{\mathbb{Q}}$ of a mod 2 character of the Galois group of $H|K$. If the prime 2 stays inert in $\mathcal{O}_K$, then $2\mathcal{O}_K$ splits completely in H and the order of $\rho_K(\mathrm{Frob}_2)$ is 2, where Frob_2 is a Frobenius element at 2. Consequently, ρ_K is exceptional. An example for this behaviour is provided by $K = \mathbb{Q}(\sqrt{229})$. If the prime 2 splits in $\mathcal{O}_K$ and the primes of $\mathcal{O}_K$ lying above 2 are principal, then $\rho_K(\mathrm{Frob}_2)$ is the identity and hence ρ_K is exceptional. This happens for example for $K = \mathbb{Q}(\sqrt{2089})$.

Let us point out that some of the weight one forms that we obtain cannot be lifted to characteristic zero forms of weight one and the same level, so that the theory of modular forms by Katz becomes necessary. Namely, if $p = 2$ and the dihedral representation in

[1] A small mistake concerning $n = 2$ has been corrected (pointed out by K. Buzzard).

[2] By the prime-to-p part we mean the restriction to $(\mathbb{Z}/N_\rho\mathbb{Z})^*$.

[3] More precisely: $\Gamma_1(N)$ with $N \geq 5$.

question has odd conductor N and is induced from a real quadratic field K of discriminant N, whose fundamental units have norm -1, then there does not exist an odd characteristic zero representation with conductor dividing N that reduces to ρ. The representation coming from the quadratic field $\mathbb{Q}(\sqrt{229})$ used above, can also here serve as an example.[4]

The fact that dihedral representations come from *some* modular form is well-known (apparently already due to Hecke[5]). So the subtle issue is to adjust the level, character and weight. It should be noted that Rohrlich and Tunnell solved many cases for $p = 2$ with Serre's weight k_ρ by rather elementary means in [R-T], however, with the more restrictive definition of a dihedral representation to be such that its image in $\mathrm{GL}_2(\overline{\mathbb{F}_2})$, and not in $\mathrm{PGL}_2(\overline{\mathbb{F}_2})$, is isomorphic to a dihedral group.

Let us also mention that it is possible to do computations of weight one forms in positive characteristic on a computer (see [W-App]) and thus to collect evidence for Serre's conjecture in some cases.

This note is organised as follows. The number theoretic ingredients on dihedral representations are provided in Section 2. In Section 3 some results on oldforms, also in positive characteristic, are collected. Section 4 is devoted to the proof of Theorem (1.1.1). Finally, in Section 5 we include a result on the irreducibility of certain mod p representations.

I wish to thank Peter Stevenhagen for helpful discussions and comments and especially Bas Edixhoven for invaluable explanations and his constant support.

1.2. Dihedral representations

We shall first recall some facts on Galois representations. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}(V)$ be a continuous representation with V a 2-dimensional vector space over an algebraically closed discrete field k.

Let L be the number field such that $\mathrm{Ker}(\rho) = G_L$ (by the notation G_L we always mean the absolute Galois group of L). Given a prime Λ of L dividing the rational prime l, we denote by $G_{\Lambda,i}$ the i-th ramification group in lower numbering of the local extension $L_\Lambda|\mathbb{Q}_l$. Furthermore, one sets

$$n_l(\rho) = \sum_{i \geq 0} \frac{\dim(V/V^{G_{\Lambda,i}})}{(G_{\Lambda,0} : G_{\Lambda,i})}.$$

This number is an integer, which is independent of the choice of the prime Λ above l. With this one defines the *conductor* of ρ to be $\mathfrak{f}(\rho) = \prod_l l^{n_l(\rho)}$, where the product runs over all primes l different from the characteristic of k. If k is the field of complex numbers, $\mathfrak{f}(\rho)$ coincides with the *Artin conductor*.

[4] It was pointed out by Frank Calegari that the form in question does come from a holomorphic eigenform of weight one and level 229. The projective image of its complex representation is S_4 and thus not dihedral. This phenomenon cannot happen when the class number of the real quadratic field is at least 5.

[5] Hecke probably knew this for odd p. The case $p = 2$ can be dealt with by Serre's trick (see Lemma (1.2.1))

Let ρ be a dihedral representation. Then ρ is induced from a character $\chi : G_K \to k^*$ for a quadratic number field K such that $\chi \neq \chi^\sigma$, with $\chi^\sigma(g) = \chi(\sigma^{-1}g\sigma)$ for all $g \in G_K$, where σ is a lift to $G_{\mathbb{Q}}$ of the non-trivial element of $G_{K|\mathbb{Q}}$. For a suitable choice of basis we then have the following explicit description of ρ: If an unramified prime l splits in K as $\Lambda\sigma(\Lambda)$, then $\rho(\mathrm{Frob}_l) = \begin{pmatrix} \chi(\mathrm{Frob}_\Lambda) & 0 \\ 0 & \chi^\sigma(\mathrm{Frob}_\Lambda) \end{pmatrix}$. Moreover, $\rho(\sigma)$ is represented by the matrix $\begin{pmatrix} 0 & 1 \\ \chi(\sigma^2) & 0 \end{pmatrix}$. As ρ is continuous, its image is a finite group, say, of order m.

(1.2.1) Lemma. *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ be an odd dihedral representation that is unramified at p. Define K, χ, σ and m as above. Let N be the conductor of ρ. Let ζ_m a primitive m-th root of unity and $\mathfrak{P}$ a prime of $\mathbb{Q}(\zeta_m)$ above p.*

Then one of the following two statements holds.

(a) There exists an odd dihedral representation $\widehat{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}[\zeta_m])$, which has Artin conductor N and reduces to ρ modulo $\mathfrak{P}$.

(b) One has that $p = 2$ and K is real quadratic. Moreover, there is an infinite set S of primes such that for each $l \in S$ the trace of $\rho(\mathrm{Frob}_l)$ is zero, and there exists an odd dihedral representation $\widehat{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}[\zeta_m])$, which has Artin conductor Nl and reduces to ρ modulo $\mathfrak{P}$.

Proof. Suppose that the quadratic field K equals $\mathbb{Q}(\sqrt{D})$ with D square-free. The character $\chi : G_K \to k^*$ can be uniquely lifted to a character $\widetilde{\chi} : G_K \to \mathbb{Z}[\zeta_m]^*$ of the same order, which reduces to χ modulo $\mathfrak{P}$. Denote by $\widetilde{\rho}$ the continuous representation $\mathrm{Ind}_{G_K}^{G_{\mathbb{Q}}} \widetilde{\chi}$. For the choice of basis discussed above the matrices representing ρ can be lifted to matrices representing $\widetilde{\rho}$, whose non-zero entries are in the m-th roots of unity. Then for a subgroup H of the image $\rho(G_{\mathbb{Q}})$, one has that $(\overline{\mathbb{F}_p}^2)^H$ is isomorphic to $(\mathbb{Z}[\zeta_m]^2)^H \otimes \overline{\mathbb{F}_p}$. Hence the conductor of ρ equals the Artin conductor of $\widetilde{\rho}$, as $\widetilde{\rho}$ is unramified at p. Alternatively, one can first remark that the conductor of χ equals the conductor of $\widetilde{\chi}$ and then use the formulae $\mathfrak{f}(\rho) = \mathrm{Norm}_{K|\mathbb{Q}}(\mathfrak{f}(\chi))D$ and $\mathfrak{f}(\widetilde{\rho}) = \mathrm{Norm}_{K|\mathbb{Q}}(\mathfrak{f}(\widetilde{\chi}))D$.

Thus condition (a) is satisfied if $\widetilde{\rho}$ is odd. Let us now consider the case when $\widetilde{\rho}$ is even. This immediately implies $p = 2$ and that the quadratic field K is real, as is the number field L whose absolute Galois group G_L equals the kernel of ρ, and hence also the kernel of $\widetilde{\chi}$. We shall now adapt "Serre's trick" from [R-T], p. 307, to our situation.

Let $\mathfrak{f}$ be the conductor of $\widetilde{\chi}$. As L is totally real, $\mathfrak{f}$ is a finite ideal of $\mathcal{O}_K$. Via class field theory, $\widetilde{\chi}$ can be identified with a complex character of $\mathrm{Cl}_K^{\mathfrak{f}}$, the ray class group modulo $\mathfrak{f}$. Let ∞_1, ∞_2 be the infinite places of K. Consider the class

$$c = [\{(\lambda) \in \mathrm{Cl}_K^{4D\mathfrak{f}\infty_1\infty_2} \mid \mathrm{Norm}(\lambda) < 0, \lambda \equiv 1 \bmod 4D\mathfrak{f}\}]$$

in the ray class group of K modulo $4D\mathfrak{f}\infty_1\infty_2$. By Cebotarev's density theorem the primes of $\mathcal{O}_K$ are uniformly distributed over the conjugacy classes of $\mathrm{Cl}_K^{4D\mathfrak{f}\infty_1\infty_2}$. Hence, there are infinitely many primes Λ of degree 1 in the class c. Take S to be the set of rational primes

lying under them. Let a prime Λ from the class c be given. It is principal, say $\Lambda = (\lambda)$, and coprime to $4D\mathfrak{f}$. By construction we have $c^2 = [\Lambda^2] = 1$. As $\mathrm{Cl}_K^{\mathfrak{f}}$ is a quotient of $\mathrm{Cl}_K^{4D\mathfrak{f}\infty_1\infty_2}$, the class of Λ in $\mathrm{Cl}_K^{\mathfrak{f}}$ has order 1 or 2. Since $p = 2$, the character χ has odd order and we conclude that $\chi(\Lambda) = 1$.

We have $\lambda \equiv 1 \bmod 4D\mathfrak{f}$ and $\mathrm{Norm}(\lambda) = -l$ for some odd prime l. Hence, the extension $K(\sqrt{\lambda})$ has two real and two complex embeddings and is unramified at 2 and at the primes dividing $D\mathfrak{f}$. We represent $K(\sqrt{\lambda})$ by the quadratic character $\xi : G_K \to \{\pm 1\}$. For the complex conjugation, the "infinite Frobenius element", Frob_{∞_1}, we have that $\xi(\mathrm{Frob}_{\infty_1})\xi^\sigma(\mathrm{Frob}_{\infty_1}) = -1$. We now consider the representation $\widehat{\rho}$ obtained by induction from the character $\widehat{\chi} = \widetilde{\chi}\xi$. Using the same basis as in the discussion at the beginning of this section, an element g of G_K is represented by the matrix $\left(\begin{smallmatrix} \widetilde{\chi}(g)\xi(g) & 0 \\ 0 & \widetilde{\chi}^\sigma(g)\xi^\sigma(g) \end{smallmatrix}\right)$. In particular, we obtain that the determinant of Frob_∞ over $\mathbb{Q}$ equals -1, whence $\widehat{\rho}$ is odd. Moreover, as l splits in K, one has that $\rho(\mathrm{Frob}_l)$ is the identity matrix, so that the trace of $\rho(\mathrm{Frob}_l)$ is zero.

The reduction of $\widehat{\rho}$ equals ρ, as ξ is trivial in characteristic 2. Moreover, outside Λ the conductor of $\widehat{\chi}$ equals the conductor of $\widetilde{\chi}$. At the prime Λ the local conductor of $\widehat{\chi}$ is Λ, as the ramification is tame. Consequently, the Artin conductor of $\widehat{\rho}$ equals Nl. □

Also without the condition that it is unramified at p, one can lift a dihedral representation to characteristic zero, however, losing control of the Artin conductor.

(1.2.2) Lemma. *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ be an odd dihedral representation. Define K, χ, m, ζ_m and $\mathfrak{P}$ as in the previous lemma.*

There exists an odd dihedral representation $\widehat{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}[\zeta_m])$, whose reduction modulo $\mathfrak{P}$ is isomorphic to ρ.

Proof. We proceed as in the preceding lemma for the definitions of $\widetilde{\chi}$ and $\widetilde{\rho}$. If $\widetilde{\rho}$ is even, then $p = 2$ and K is real. In that case we choose some $\lambda \in \mathcal{O}_K - \mathbb{Z}$, which satisfies $\mathrm{Norm}(\lambda) < 0$. The field $K(\sqrt{\lambda})$ then has signature $(2, 1)$ and gives a character $\xi : G_K \to \mathbb{Z}[\zeta_m]^*$. As in the proof of the preceding lemma one obtains that the representation $\widehat{\rho} = \mathrm{Ind}_{G_K}^{G_{\mathbb{Q}}} \widetilde{\chi}\xi$ is odd and reduces to ρ modulo $\mathfrak{P}$. □

1.3. On oldforms

In this section we collect some results on oldforms. We try to stay as much as possible in the characteristic zero setting. However, we also need a result on Katz modular forms.

(1.3.1) Proposition. *Let N, k, r be positive integers, p a prime and ϵ a Dirichlet character of modulus N. The homomorphism*

$$\phi_{p^r}^N : \big(\mathcal{S}_k(\Gamma_1(N), \epsilon, \mathbb{C})\big)^{r+1} \hookrightarrow \mathcal{S}_k(\Gamma_1(Np^r), \epsilon, \mathbb{C}), \quad (f_0, f_1, \ldots, f_r) \mapsto \sum_{i=0}^{r} f_i(q^{p^i})$$

is compatible with all Hecke operators T_n with $(n,p)=1$.

Let $f \in \mathcal{S}_k(\Gamma_1(N),\epsilon,\mathbb{C})$ be a normalised eigenform for all Hecke operators. Then the forms $f(q), f(q^{p^2}), \ldots, f(q^{p^r})$ in the image of $\phi^N_{p^r}$ are linearly independent, and on their span the action of the operator T_p in level Np^r is given by the matrix

$$\begin{pmatrix} a_p(f) & 1 & 0 & 0 & \ldots & 0 \\ -\delta p^{k-1}\epsilon(p) & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ & & \vdots & & & \\ 0 & \ldots & 0 & 0 & 0 & 1 \\ 0 & \ldots & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $\delta = 1$ if $p \nmid N$ and $\delta = 0$ otherwise.

Proof. The embedding map and its compatibility with the Hecke action away from p is explained in [DiamondIm], Section 6.1. The linear independence can be checked on q-expansions. Finally, the matrix can be elementarily computed. □

(1.3.2) Corollary. *Let p be a prime, $r \geq 0$ some integer and $f \in \mathcal{S}_k(\Gamma_1(Np^r),\epsilon,\mathbb{C})$ an eigenform for all Hecke operators. Then there exists an eigenform for all Hecke operators $\tilde{f} \in \mathcal{S}_k(\Gamma_1(Np^{r+2}),\epsilon,\mathbb{C})$, which satisfies $a_l(\tilde{f}) = a_l(f)$ for all primes $l \neq p$ and $a_p(\tilde{f}) = 0$.*

Proof. One computes the characteristic polynomial of the operator T_p of Proposition (1.3.1) and sees that it has 0 as a root if the dimension of the matrix is at least 3. Hence one can choose the desired eigenform $\widetilde{f}$ in the image of $\phi^{Np^r}_{p^2}$. □

As explained in the introduction, Katz' theory of modular forms ought to be used in the study of Serre's conjecture. Following [EdixBoston], we briefly recall this concept, which was introduced by Katz in [Katz]. However, we shall use a "non-compactified" version.

Let $N \geq 1$ be an integer and R a ring, in which N is invertible. One defines the category $[\Gamma_1(N)]_R$, whose objects are pairs $(E/S/R,\alpha)$, where S is an R-scheme, E/S an elliptic curve (i.e. a proper smooth morphism of R-schemes, whose geometric fibres are connected smooth curves of genus one, together with a section, the "zero section", $0 : S \to E$) and $\alpha : (\mathbb{Z}/N\mathbb{Z})_S \to E[N]$, the *level structure*, is an embedding of S-group schemes. The morphisms in the category are cartesian diagrams

$$\begin{array}{ccc} E' & \longrightarrow & E \\ \downarrow & \square & \downarrow \\ S' & \longrightarrow & S, \end{array}$$

which are compatible with the zero sections and the level structures. For every such elliptic curve $E/S/R$ we let $\underline{\omega}_{E/S} = 0^*\underline{\Omega}_{E/S}$. For every morphism $\pi : E'/S'/R \to E/S/R$ the induced map $\underline{\omega}_{E'/S'} \to \pi^*\underline{\omega}_{E/S}$ is an isomorphism.

A *Katz cusp form* $f \in \mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ assigns to every object $(E/S/R, \alpha)$ of $[\Gamma_1(N)]_R$ an element $f(E/S/R, \alpha) \in \underline{\omega}^{\otimes k}_{E/S}(S)$, compatibly for the morphisms in the category, subject to the condition that all q-expansions (which one obtains by adjoining all N-th roots of unity and plugging in a suitable Tate curve) only have positive terms.

For the following definition let us remark that if $m \geq 1$ is coprime to N and is invertible in R, then any morphism of group schemes of the form $\phi_{Nm} : (\mathbb{Z}/Nm\mathbb{Z})_S \to E[Nm]$ can be uniquely written as $\phi_N \times_S \phi_m$ with $\phi_N : (\mathbb{Z}/N\mathbb{Z})_S \to E[N]$ and $\phi_m : (\mathbb{Z}/m\mathbb{Z})_S \to E[m]$.

(1.3.3) Definition. *A Katz modular form $f \in \mathcal{S}_k(\Gamma_1(Nm), R)_{\mathrm{Katz}}$ is called* independent of m *if for all elliptic curves $E/S/R$, all $\phi_N : (\mathbb{Z}/N)_S \hookrightarrow E[N]$ and all $\phi_m, \phi'_m : (\mathbb{Z}/m)_S \hookrightarrow E[m]$ one has the equality*

$$f(E/S/R, \phi_N \times_S \phi_m) = f(E/S/R, \phi_N \times_S \phi'_m) \in \underline{\omega}^{\otimes k}_{E/S}(S).$$

(1.3.4) Proposition. *Let N, m be coprime positive integers and R a ring, which contains the Nm-th roots of unity and $\frac{1}{Nm}$. A Katz modular form $f \in \mathcal{S}_k(\Gamma_1(Nm), R)_{\mathrm{Katz}}$ is independent of m if and only if there exists a Katz modular form $g \in \mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ such that*

$$f(E/S/R, \phi_{Nm}) = g(E/S/R, \phi_{Nm} \circ \psi)$$

for all elliptic curves $E/S/R$ and all $\phi_{Nm} : (\mathbb{Z}/Nm\mathbb{Z})_S \hookrightarrow E[Nm]$. Here ψ denotes the canonical embedding $(\mathbb{Z}/N\mathbb{Z})_S \hookrightarrow (\mathbb{Z}/Nm\mathbb{Z})_S$ of S-group schemes. In that case, f and g have the same q-expansion at ∞.

Proof. If $m = 1$, there is nothing to do. If necessary replacing m by m^2, we can hence assume that m is at least 3.

Let us now consider the category $[\Gamma_1(N; m)]_R$, whose objects are triples $(E/S/R, \phi_N, \psi_m)$, where S is an R scheme, E/S an elliptic curve, $\phi_N : (\mathbb{Z}/N\mathbb{Z})_S \hookrightarrow E[N]$ an embedding of group schemes and $\psi_m : (\mathbb{Z}/m\mathbb{Z})^2_S \cong E[m]$ an isomorphism of group schemes. The morphisms are cartesian diagrams compatible with the zero sections, the ϕ_N and the ψ_m as before.

We can pull back the form $f \in \mathcal{S}_k(\Gamma_1(Nm), R)_{\mathrm{Katz}}$ to a Katz form h on $[\Gamma_1(N; m)]_R$ as follows. First let $\beta : (\mathbb{Z}/m\mathbb{Z})_S \hookrightarrow (\mathbb{Z}/m\mathbb{Z})^2_S$ be the embedding of S-group schemes defined by mapping onto the first factor. Using this, f gives rise to h by setting

$$h((E/S/R, \phi_N, \psi_m)) = f((E/S/R, \phi_N, \psi_m \circ \beta)) \in \underline{\omega}^{\otimes k}_{E/S}(S).$$

As f is independent of m, it is clear that h is independent of ψ_m and thus invariant under the natural $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$-action.

As $m \geq 3$, one knows that the category $[\Gamma_1(N; m)]_R$ has a final object $(E^{\mathrm{univ}}/Y_1(N; m)_R/R, \alpha^{\mathrm{univ}})$. In other words, h is an $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$-invariant global section of $\underline{\omega}^{\otimes k}_{E^{\mathrm{univ}}/Y_1(N;m)_R}$. Since this R-module is equal to $\mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ (see e.g.

Equation 1.2 of [EdixBoston], p. 210), we find some $g \in \mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ such that $f(E/S/R, \phi_{Nm}) = g(E/S/R, \phi_{Nm} \circ \psi)$ for all $(E/S/R, \phi_{Nm})$.

Plugging in the Tate curve, one sees that the standard q-expansions of f and g coincide. □

(1.3.5) Corollary. *Let N, m be coprime positive integers, p a prime not dividing Nm and $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \overline{\mathbb{F}_p}$ a character. Let $f \in \mathcal{S}_k(\Gamma_1(Nm), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ be a Katz cuspidal eigenform for all Hecke operators.*

If f is independent of m, then there exists an eigenform for all Hecke operators $g \in \mathcal{S}_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ such that the associated Galois representations ρ_f and ρ_g are isomorphic.

Proof. From the preceding proposition we get a modular form $g \in \mathcal{S}_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$, noting that the character is automatically good. Because of the compatibility of the embedding map with the operators T_l for primes $l \nmid m$, we find that g is an eigenform for these operators. As the operators T_l for primes $l \nmid m$ commute with the others, we can choose a form of the desired type. □

1.4. Proof of the principal result

We first cover the weight one case.

(1.4.1) Theorem. *Let p be a prime and $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ an odd dihedral representation of conductor N, which is unramified at p. Let ϵ denote the character* $\det \circ \rho$.

Then there exists a Katz eigenform f in $\mathcal{S}_1(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$, whose associated Galois representation is isomorphic to ρ.

Proof. Assume first that part (a) of Lemma (1.2.1) applies to ρ, and let $\widehat{\rho}$ be a lift provided by that lemma. A theorem by Weil-Langlands (Theorem 1 of [Serre2]) implies the existence of a newform g in $\mathcal{S}_1(\Gamma_1(N), \det \circ \widehat{\rho}, \mathbb{C})$, whose associated Galois representation is isomorphic to $\widehat{\rho}$. Now reduction modulo a suitable prime above p yields the desired modular form. In particular, one does not need Katz' theory in this case.

If part (a) of Lemma (1.2.1) does not apply, then part (b) does, and we let S be the infinite set of primes provided. For each $l \in S$ the theorem of Weil-Langlands yields a newform $f^{(l)}$ in $\mathcal{S}_1(\Gamma_1(Nl), \mathbb{C})$, whose associated Galois representation reduces to ρ modulo $\mathfrak{P}$, where $\mathfrak{P}$ is the ideal from the lemma. Moreover, the congruence $a_q(f^{(l)}) \equiv 0 \bmod \mathfrak{P}$ holds for all primes $q \in S$ different from l.

From Corollary (1.3.2) we obtain Hecke eigenforms $\widetilde{f}^{(l)} \in \mathcal{S}_1(\Gamma_1(Nl^3), \mathbb{C})$ such that $a_l(\widetilde{f}^{(l)}) = 0$ and $a_q(\widetilde{f}^{(l)}) = a_q(f^{(l)}) \equiv 0 \bmod \mathfrak{P}$ for all primes $q \in S$, $q \neq l$. Reducing modulo the prime ideal $\mathfrak{P}$, we get eigenforms $g^{(l)} \in \mathcal{S}_1(\Gamma_1(Nl^3), \epsilon, \overline{\mathbb{F}_p})$, whose associated Galois representations are isomorphic to ρ. One also has $a_q(g^{(l)}) = 0$ for all $q \in S$.

The coefficients $a_q(f^{(l)})$ for all primes $q \mid N$ appear in the L-series of the complex representation $\rho_{f^{(l)}}$ associated to $f^{(l)}$. As the image of $\rho_{f^{(l)}}$ is isomorphic to a fixed finite group G, not depending on l, there are only finitely many possibilities for the value of $a_q(f^{(l)})$. Hence the same holds for the $g^{(l)}$. Consequently, there are two forms $g_1 = g^{(l_1)}$ and $g_2 = g^{(l_2)}$ for $l_1 \neq l_2$ that have the same coefficients at all primes $q \mid N$. For primes $q \nmid Nl_1l_2$ one has that the trace of $\rho_{f^{(l_1)}}(\mathrm{Frob}_q)$ is congruent to the trace of $\rho_{f^{(l_2)}}(\mathrm{Frob}_q)$, whence $a_q(g_1) = a_q(g_2)$. Let us point out that this includes the case $q = p = 2$, as the complex representation is unramified at p.

In the next step we embed g_1 and g_2 into $\mathcal{S}_1(\Gamma_1(Nl_1^3l_2^3), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ via the method in the statement of Proposition (1.3.4). As the q-expansions coincide, g_1 and g_2 are mapped to the same form h. But as h comes from g_2, it is independent of l_1 and analogously also of l_2. Since $\rho_h = \rho$, Theorem (1.4.1) follows immediately from Corollary (1.3.5). □

We will deduce the cases of weight at least two from general results. The current state of the art in "level and weight lowering" seems to be the following theorem.

(1.4.2) Theorem. (Ribet, Edixhoven, Diamond, Buzzard,...) *Let p be a prime and $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ a continuous irreducible representation, which is assumed to come from some modular form. Define k_ρ and N_ρ as in [Serre1]. If $p = 2$, additionally assume either (i) that the restriction of ρ to a decomposition group at 2 is not contained within the scalar matrices or (ii) that ρ is ramified at 2.*

Then there exists a normalised eigenform $f \in \mathcal{S}_{k_\rho}(\Gamma_1(N_\rho), \overline{\mathbb{F}_p})$ giving rise to ρ.

Proof. The case $p \neq 2$ is Theorem 1.1 of [Diamond], and the case $p = 2$ with condition (i) follows from Propositions 1.3 and 2.4 and Theorem 3.2 of [Buzzard], multiplying by the Hasse invariant if necessary.

We now show that if $p = 2$ and ρ restricted to a decomposition group $G_{\mathbb{Q}_2}$ at 2 is contained within the scalar matrices, then ρ is unramified at 2. Let $\phi : G_{\mathbb{Q}} \to \overline{\mathbb{F}_2}^*$ be the character such that $\phi^2 = \det \circ \rho$. As ϕ has odd order, it is unramified at 2 because of the Kronecker-Weber theorem. If ρ restricted to $G_{\mathbb{Q}_2}$ is contained within the scalar matrices, then we have that $\rho|_{G_{\mathbb{Q}_2}}$ is $\begin{pmatrix} \phi|_{G_{\mathbb{Q}_2}} & 0 \\ 0 & \phi|_{G_{\mathbb{Q}_2}} \end{pmatrix}$, whence ρ is unramified at 2. □

Proof of theorem (1.1.1). Let ρ be the dihedral representation from the assertion. If ρ is unramified at p, one has $k(\rho) = 1$, and Theorem (1.1.1) follows from Theorem (1.4.1).

If ρ is ramified at p, then let $\widehat{\rho}$ be a characteristic zero representation lifting ρ, as provided by Lemma (1.2.2). The theorem by Weil-Langlands already used above (Theorem 1 of [Serre2]) implies the existence of a newform in weight one and characteristic zero giving rise to $\widehat{\rho}$. So from Theorem (1.4.2) we obtain that ρ comes from a modular form of Serre's weight k_ρ and level N_ρ. Let us note that using Katz modular forms the character is automatically the conjectured one ϵ_ρ.

The weights k_ρ and $k(\rho)$ only differ in two cases (see [EdixWeight], Remark 4.4). The first case is when $k(\rho) = 1$. The other case is when $p = 2$ and ρ is not finite at 2. Then

one has $k(\rho) = 3$ and $k_\rho = 4$. In that case one applies Theorem 3.4 of [EdixWeight] to obtain an eigenform of the same level and character in weight 3, or one applies Theorem 3.2 of [Buzzard] directly. □

1.5. An irreducibility result

We first study the relation between the level of an eigenform in characteristic p and the conductor of the associated Galois representation.

(1.5.1) Lemma. *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ be a continuous representation of conductor N, and let k be a positive integer. If $f \in \mathcal{S}_k(\Gamma_1(M), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ is a Hecke eigenform giving rise to ρ, then N divides M.*

Proof. By multiplying with the Hasse invariant, if necessary, we can assume that the weight is at least 2. Hence the form f can be lifted to characteristic zero (see e.g. [DiamondIm], Theorem 12.3.2) in the same level. Thus there exists a newform g, say of level L, whose Galois representation ρ_g reduces to ρ. Now Proposition 0.1 of [Livné] yields that N divides L. As L divides M, the lemma follows. □

We can derive the following proposition, which is of independent interest.

(1.5.2) Proposition. *Let $f \in \mathcal{S}_k(\Gamma_0(N), \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ be a normalised Hecke eigenform for a square-free level N with $p \nmid N$ in some weight $k \geq 1$.*

(a) If $p = 2$, the associated Galois representation is either irreducible or trivial.

(b) For any prime p the associated Galois representation is either irreducible or corresponds to a direct sum $\alpha \oplus \chi_p^{k-1}\alpha^{-1}$, where χ_p is the mod p cyclotomic character and α is a character factoring through $G(\mathbb{Q}(\zeta_p)|\mathbb{Q})$ for a primitive p-th root of unity ζ_p.

Proof. Let us assume that the representation ρ associated to f is reducible. Since ρ is semi-simple, it is isomorphic to the direct sum of two characters $\alpha \oplus \beta$. As the determinant is the $(k-1)$-th power of the mod p cyclotomic character χ_p, we have that $\beta = \chi_p^{k-1}\alpha^{-1}$. Since the conductor of χ_p^{k-1} is 1, it follows that the conductor of α equals that of β. Consequently, the conductor of ρ is the square of the conductor of α. Lemma (1.5.1) implies that the conductor of ρ divides N. As we have assumed this number to be square-free, we have that ρ can only ramify at p.

The number field L with $G_L = \mathrm{Ker}(\rho)$ is abelian. As only p can be ramified, it follows that L is contained in $\mathbb{Q}(\zeta_{p^n})$ for some p^n-th root of unity. Since the order of α is prime to p, we conclude that α factors through $G(\mathbb{Q}(\zeta_p)|\mathbb{Q})$. In characteristic $p = 2$ this implies that ρ is the trivial representation, as χ_2 is the trivial character. □

Chapter II

Modular Symbols Over Rings

The Eichler-Shimura-Theorem (Theorem (3.3.1)) establishes an isomorphism between the direct sum of two copies of the space of holomorphic cusp forms for a congruence subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ of finite index and the parabolic subspace of the analytic cohomology of the associated modular curve X_Γ for a certain sheaf of $\mathbb{C}$-vector spaces. In this setting the Hecke algebra defined on the cohomology group coincides with the usual one on cusp forms, so that the knowledge of the Hecke operators on the cohomology group determines the cusp forms completely. One of the principal themes of this thesis is to obtain similar results over finite fields in certain cases.

This chapter is concerned with the analytic cohomology groups used in the Eichler-Shimura theorem, but over general rings. Whereas from a geometric point of view the cohomology of modular curves is the most natural object to study, it only becomes explicitly accessible via the natural comparison with group cohomology. Another explicit approach is provided by the modular symbols formalism. It is of practical interest, as it has been implemented by William Stein into Magma. We compute the differences between these objects for general congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ and give a criterion when they agree.

A link with the theory of modular forms will be established in Chapter III.

We start this chapter by introducing modular curves as Riemann surfaces, analytic modular stacks and the sheaves and some of their properties to be used in the sequel. We begin our study with the cohomology of modular stacks and relate it to group cohomology. Next, we derive an explicit description of the cohomology of modular curves for the push-forward of any locally constant sheaf on the modular stack by comparing it via the Leray spectral sequence to stack cohomology and using the Mayer-Vietoris sequence for group cohomology. Moreover, torsion properties are discussed. The following section is devoted to introducing the modular symbols formalism and to prove an explicit description in terms of the so-called Manin symbols. Next, we will be able to give a precise description of when the spaces in question agree, resp. what their differences are. The final section treats modular symbols for

$\Gamma_1(N)$ as a $(\mathbb{Z}/N\mathbb{Z})^*$-module and a slight generalisation to some other subgroups.

(2.0.3) Notation. *Recall that for a subgroup H of $\mathrm{SL}_2(\mathbb{Z})$ we denote $\overline{H} = H/(\langle -\mathbf{1}\rangle \cap H)$, which we consider as a subgroup of $\mathrm{PSL}_2(\mathbb{Z})$.*

Throughout this chapter we let Γ and G be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ such that

$$\Gamma \lhd G \leq \mathrm{SL}_2(\mathbb{Z}).$$

For a ring R and an integer $k \geq 2$ we let

$$V_{k-2}(R) := \mathrm{Sym}^{k-2}(R^2)$$

which carries the natural left $\mathrm{SL}_2(\mathbb{Z})$-action. Moreover, we will use a character of the form

$$\epsilon : G \overset{\mathrm{proj}}{\twoheadrightarrow} \Gamma\backslash G \to R^*$$

and denote by R^ϵ the $R[G]$-module which is defined to be a copy of R with G-action through ϵ^{-1}. Also define

$$V_{k-2}^\epsilon(R) := V_{k-2}(R) \otimes_R R^\epsilon$$

for the diagonal G-action. In case that G contains the matrix $-\mathbf{1}$, we will always assume that $\epsilon(-\mathbf{1}) = (-1)^k$, so that $V_{k-2}^\epsilon(R)$ is an $R[\overline{G}]$-module.

2.1. Modular curves and modular stacks

We assume Notation (2.0.3), as we do in all this chapter. The group Γ acts from the left on the extended upper half plane $\overline{\mathbb{H}} = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$ by fractional linear transformations. We can associate to it the compact Riemann surface $X_\Gamma := \Gamma\backslash\mathbb{H} \cup \Gamma\backslash\mathbb{P}^1(\mathbb{Q})$. It contains the open Riemann surface $Y_\Gamma := \Gamma\backslash\mathbb{H}$. Both X_Γ and Y_Γ are called the *modular curve* of Γ. We denote the inclusion by $j_\Gamma : Y_\Gamma \hookrightarrow X_\Gamma$. We remark that $-\mathbf{1}$ acts trivially, so that we could have used $\overline{\Gamma}$ in the definitions.

Analogously, we also define the analytic Deligne-Mumford stacks $[X_{\overline{\Gamma}}]$ and $[Y_{\overline{\Gamma}}]$ as the stacks obtained by taking the quotient for the $\overline{\Gamma}/\overline{\Gamma(N)}$-action on $X_{\Gamma(N)}$ resp. $Y_{\Gamma(N)}$, when $\Gamma(N) \leq \Gamma$ with $N \geq 3$. These stacks will be referred to as the *modular stacks of* Γ. Again we have the open embedding $j_{[\Gamma]} : [Y_{\overline{\Gamma}}] \hookrightarrow [X_{\overline{\Gamma}}]$.

Moreover, there are natural projections $\pi_\Gamma : [X_{\overline{\Gamma}}] \twoheadrightarrow X_\Gamma$ and $\pi_\Gamma : [Y_{\overline{\Gamma}}] \twoheadrightarrow Y_\Gamma$. These commute with the embeddings j_Γ and $j_{[\Gamma]}$. If the group $\overline{\Gamma}$ acts freely on $\mathbb{H}$ and if the stabiliser subgroup of Γ for any cusp only contains unipotent elements, then both π_Γ are isomorphisms.

(2.1.1) Remark. *Analytic Deligne-Mumford stacks have e.g. been defined in [Toen], Definition 5.2, building on the definition of the analytic site (loc. cit. p. 171). Moreover, it is stated that quotient stacks of analytic spaces by finite groups are analytic Deligne-Mumford stacks, which implies that the $[Y_{\overline{\Gamma}}]$ and $[X_{\overline{\Gamma}}]$ above are.*

In the category of sheaves on the analytic site there are enough injectives (see e.g. [Milne], Proposition III.1.1), so that a derived functor cohomology exists. This cohomology coincides with the derived functor cohomology on analytic spaces, if the analytic stack is an analytic space (for a discussion see [Milne], p. 118). As we will use the Leray spectral sequence, we point out that it is a formal consequence, as the direct image of an injective sheaf is injective and both the direct image functor and the global sections functor are left exact (see e.g. [Milne], Theorem B.1).

There is a category equivalence between the locally constant sheaves of R-modules on $[Y_{\overline{\Gamma}}]$ and $R[\overline{\Gamma}]$-modules, given by the functor

$$\mathcal{F} \mapsto H^0(\mathbb{H}, f^*\mathcal{F}),$$

where $f : \mathbb{H} \overset{\text{proj}}{\twoheadrightarrow} [Y_{\overline{\Gamma}}]$ is the quotient morphism. As $\mathbb{H}$ is simply connected, the sheaf $f^*\mathcal{F}$ is constant and consequently $H^0(\mathbb{H}, f^*\mathcal{F}) = (f^*\mathcal{F})_y = \mathcal{F}_{f(y)}$ for any point $y \in \mathbb{H}$. It follows that

$$\big(H^0(\mathbb{H}, f^*\mathcal{F})\big)^{\overline{\Gamma}} = H^0([Y_{\overline{\Gamma}}], \mathcal{F}).$$

As stack cohomology is the derived functor cohomology of $H^0([Y_{\overline{\Gamma}}], \cdot)$ and group cohomology for $R[\overline{\Gamma}]$-modules is the derived functor cohomology of taking $\overline{\Gamma}$-invariants, we obtain

$$H^i([Y_{\overline{\Gamma}}], \mathcal{F}) \cong H^i(\overline{\Gamma}, H^0(\mathbb{H}, f^*\mathcal{F})) \cong H^i(\overline{\Gamma}, \mathcal{F}_x)$$

for any $i \geq 0$, $\mathcal{F}$ a locally constant sheaf of R-modules on $[Y_{\overline{\Gamma}}]$ and $x \in [Y_{\overline{\Gamma}}]$. We say that $H^0(\mathbb{H}, f^*\mathcal{F}) = \mathcal{F}_x$ is the $R[\overline{\Gamma}]$-module associated to *the locally constant sheaf $\mathcal{F}$ and vice versa.*

2.2. The module $V^{\epsilon}_{k-2}(R)$ and the sheaf $\mathbb{V}^{\epsilon}_{k-2}(R)$

In Notation (2.0.3) we have defined $V_{k-2}(R)$ and $V^{\epsilon}_{k-2}(R)$. Via the correspondence outlined in Remark (2.1.1) the $\overline{\Gamma}$-module $V_{k-2}(R)$ corresponds to a locally constant sheaf on $[Y_{\overline{\Gamma}}]$ which we denote by $\mathbb{V}_{k-2,\overline{\Gamma}}(R)$. Similarly, we write $\mathbb{V}^{\epsilon}_{k-2,\overline{G}}(R)$ for the locally constant sheaf on $[Y_{\overline{G}}]$ corresponding to the $\overline{G}$-module $V^{\epsilon}_{k-2}(R)$. We will usually drop $\overline{\Gamma}$ and $\overline{G}$ from the notation.

(2.2.1) Remark. *Let us assume that $-\mathbf{1} \notin \Gamma$. Then we define the* universal elliptic curve *$\pi^{\text{univ}} : [\mathbb{E}^{\text{univ}}_{\overline{\Gamma}}] \twoheadrightarrow [Y_{\overline{\Gamma}}]$, as the stack obtained by taking the $\overline{\Gamma}$-quotient of $\mathbb{E}$ in the exact sequence*

$$0 \to \mathbb{Z}^2 \times \mathbb{H} \xrightarrow{((n,m),\tau)\mapsto(n\tau+m,\tau)} \mathbb{C} \times \mathbb{H} \longrightarrow \mathbb{E} \to 0,$$

where all spaces are equipped with the natural projection to $\mathbb{H}$ and $\mathbb{C} \times \mathbb{H}$ carries the $\overline{\Gamma}$-action $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).(z,\tau) = (\frac{z}{c\tau+d}, \frac{a\tau+b}{c\tau+d})$. Alternatively, $[\mathbb{E}^{\text{univ}}_{\overline{\Gamma}}]$ can also be obtained as the quotient stack for the group $\overline{\Gamma}/\overline{\Gamma(N)}$ of the universal elliptic curve $\mathbb{E}^{\text{univ}}_{\Gamma(N)}$ over $Y_{\Gamma(N)}$, when $\Gamma(N) \leq \Gamma$ and $N \geq 3$.

When $k \geq 2$ is an integer, the sheaf $\mathbb{V}_{k-2,\overline{\Gamma}}(R)$ on the modular stack $[Y_{\overline{\Gamma}}]$ agrees with $\mathrm{Sym}^{k-2}(R^1\pi^{\mathrm{univ}}_* R_{[\mathbb{E}^{\mathrm{univ}}_{\overline{\Gamma}}]})$, where $R_{[\mathbb{E}^{\mathrm{univ}}_{\overline{\Gamma}}]}$ denotes the constant sheaf R on $[\mathbb{E}^{\mathrm{univ}}_{\overline{\Gamma}}]$.

Replacing $\mathbb{Z}^2$ by $\mathbb{Z}^2 \otimes_{\mathbb{Z}} \mathbb{Z}^\epsilon$ and $\mathbb{C}$ by $\mathbb{C} \otimes_{\mathbb{Z}} \mathbb{Z}^\epsilon$ one can also make a universal elliptic curve over $Y_{\overline{G}}$, when ϵ is a quadratic character of G with kernel Γ.

In the sequel we will often use the following different description of $V_{k-2}(R)$.

(2.2.2) Lemma. *Let $R[X,Y]_n$ denote the R-module of homogeneous polynomials of degree n in the variables X and Y over R. The map*

$$\mathrm{Sym}^n(R^2) \to R[X,Y]_n, \quad \left(\begin{smallmatrix} a_1 \\ b_1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} a_n \\ b_n \end{smallmatrix}\right) \mapsto (a_1X + b_1Y) \ldots (a_nX + b_nY)$$

defines an isomorphism of left $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-modules, when we equip the polynomials with the action $(M.P)(X,Y) = P\big((X,Y)M\big)$.

Proof. The map is well defined and every monomial is obviously hit. As $\mathrm{Sym}^n(R^2)$ is freely generated by the classes of $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \otimes \cdots \otimes \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, the map is in fact an isomorphism. □

(2.2.3) Remark. *The polynomials of degree n are often equipped with a slightly different left $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-action, namely by*

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).P\left(\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)\right) := P\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^\iota \left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)\right) = P\left(\left(\begin{smallmatrix} dX - bY \\ -cX + aY \end{smallmatrix}\right)\right).$$

This action is considered e.g. in [MerelUniversal] and the Magma implementation of modular symbols. These two actions are isomorphic due to the identity $(x,y)(M^\iota)^\top = (x,y)\sigma^{-1}M\sigma$.

(2.2.4) Proposition. *Suppose that $n!$ is invertible in R. Then there is a perfect pairing $V_n(R) \times V_n(R) \to R$ of R-modules, which induces an isomorphism $V_n(R) \to V_n(R)^\vee$ of $R[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-modules, if we equip $V_n(R)^\vee$ with the left action $(M.\phi)(w) = \phi(M^\iota w)$. When M is invertible, we have $(M.\phi)(w) = \det(M)^n \phi(M^{-1}w)$.*

Proof. One defines the perfect pairing on $V_n(R)$ by first constructing a perfect pairing on R^2, which we consider as column vectors. We set

$$R^2 \times R^2 \to R, \quad \langle v, w \rangle := \det(v|w) = v_1w_2 - v_2w_1.$$

If M is a matrix in $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$, one checks easily that $\langle Mv, w \rangle = \langle v, M^\iota w \rangle$. This pairing extends to a pairing on the n-th tensor power of R^2 by letting

$$\langle v_1 \otimes \cdots \otimes v_n, w_1 \otimes \cdots \otimes w_n \rangle = \langle v_1, w_1 \rangle \cdot \ldots \cdot \langle v_n, w_n \rangle.$$

Due to our assumption on the invertibility of $n!$, we may view $\mathrm{Sym}^n(R^2)$ as a submodule in the n-th tensor power, and hence obtain the desired pairing. Consequently, one has the isomorphism of R-modules

$$V_n(R) \to V_n(R)^\vee, \quad v \mapsto (w \mapsto \langle v, w \rangle),$$

which is $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-equivariant for the actions considered. □

(2.2.5) Lemma. *Let $n \geq 1$ be an integer. We suppose that $n!N$ is not a zero divisor in R. The left t-invariants are ${}^{\langle t \rangle}V_n(R) = \langle X^n \rangle$ for $t = \left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right)$ and the left t'-invariants are ${}^{\langle t' \rangle}V_n(R) = \langle Y^n \rangle$ for $t' = \left(\begin{smallmatrix} 1 & 0 \\ N & 1 \end{smallmatrix}\right)$.*

Proof. The action of t gives $t.(X^{n-i}Y^i) = X^{n-i}(NX + Y)^i$ and consequently $(t-1).(X^{n-i}Y^i) = \sum_{j=0}^{i-1} r_{i,j} X^{n-j}Y^j$ with $r_{i,j} = N^{i-j}\binom{i}{j}$, which is not a zero divisor by assumption. For $x = \sum_{i=0}^{n} a_i X^{n-i}Y^i$ we have

$$(t-1).x = \sum_{j=0}^{n-1} X^{n-j}Y^j \Big(\sum_{i=j+1}^{n} a_i r_{i,j} \Big).$$

If $(t-1).x = 0$, we conclude for $j = n-1$ that $a_n = 0$. Next, for $j = n-2$ it follows that $a_{n-1} = 0$, and so on, until $a_1 = 0$. This proves the first part. The second follows from symmetry. □

(2.2.6) Proposition. *Let $n \geq 1$ be an integer.*

(a) If $n!N$ is not a zero divisor in R, then the R-module of left $\Gamma(N)$-invariants ${}^{\Gamma(N)}V_n(R)$ is zero.

(b) If $n!$ is invertible in R and N is not a zero divisor in R, then the R-module of left $\Gamma(N)$-coinvariants ${}_{\Gamma(N)}V_n(R)$ is zero.

(c) Suppose that Γ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ such that reduction modulo p defines a surjection $\Gamma \twoheadrightarrow \mathrm{SL}_2(\mathbb{F}_p)$. Suppose moreover that $1 \leq n \leq p$ if $p > 2$, and $n = 1$ if $p = 2$. Then one has ${}^{\Gamma}V_n(\mathbb{F}_p) = 0 = {}_{\Gamma}V_n(\mathbb{F}_p)$.

Proof. As $\Gamma(N)$ contains the matrices t and t', Lemma (2.2.5) already finishes Part (a). Under the assumptions of Part (b) Proposition (2.2.4) implies a self-duality, so that (b) follows from (a). The only part of (c) that is not yet covered is when the degree is $n = p > 2$. In that case we have an exact sequence of $\Gamma(N)$-modules

$$0 \to V_1(\mathbb{F}_p) \to V_p(\mathbb{F}_p) \to V_{p-2}(\mathbb{F}_p) \to 0.$$

In fact, $V_p(\mathbb{F}_p)$ is naturally isomorphic with the space U_1 considered on p. 46, so one can proceed as there. It suffices to take (co-)invariants to obtain the desired result. □

We also have a character version of this.

(2.2.7) Proposition. *In Notation (2.0.3) we assume that R is an integral domain and we let $N \geq 1$ be an integer which is non-zero in R.*

(a) If $n = 0$ and ϵ is non-trivial, or if $n > 0$ and $n! \neq 0$ in R, then the R-module of left G-invariants ${}^{G}V^{\epsilon}_n(R)$ is zero.

(b) If $n = 0$, ϵ is non-trivial and R is a field, or if $n > 0$ and $n!$ is invertible in R, then the R-module of left G-coinvariants ${}_G V_n^\epsilon(R)$ is zero.

Proof. If $n > 0$, this follows directly from Proposition (2.2.6) by taking Γ-invariants. If $n = 0$, we only have to remark that the G-invariants of R^ϵ are zero, if the character is non-trivial. The same holds for the coinvariants in the case of a field. □

2.3. Cohomology of modular stacks and group cohomology

Parabolic and boundary spaces

Let $\mathcal{F}$ be a sheaf on $[Y_{\overline{\Gamma}}]$. We apply the Leray spectral sequence to $j = j_{[\Gamma]} : [Y_{\overline{\Gamma}}] \hookrightarrow [X_{\overline{\Gamma}}]$. The first four terms of its associated five term exact sequence are

$$0 \to H^1([X_{\overline{\Gamma}}], j_*\mathcal{F}) \to H^1([Y_{\overline{\Gamma}}], \mathcal{F}) \to H^0([X_{\overline{\Gamma}}], R^1 j_*\mathcal{F}) \to H^2([X_{\overline{\Gamma}}], j_*\mathcal{F}).$$

In analogy with the result of Proposition (2.4.1) we call

$$H^1_{\mathrm{par}}([Y_{\overline{\Gamma}}], \mathcal{F}) := H^1([X_{\overline{\Gamma}}], j_*\mathcal{F})$$

the *parabolic stack cohomology group* (for $[Y_{\overline{\Gamma}}]$ and $\mathcal{F}$). Furthermore, $H^0([X_{\overline{\Gamma}}], R^1 j_*\mathcal{F})$ is called the *boundary stack cohomology group*.

If $\mathcal{F} = \mathbb{V}_{k-2,\overline{\Gamma}}(R)$ (resp. $\mathcal{F} = \mathbb{V}^\epsilon_{k-2,\overline{G}}(R)$ on $[Y_{\overline{G}}]$), then we speak of the *(parabolic* resp. *boundary) stack cohomology group of weight k over R for Γ (resp. for G with character ϵ).*

Comparison with group cohomology

Let now $\mathbb{V}$ be a locally constant sheaf of R-modules on $[Y_{\overline{\Gamma}}]$ which corresponds to an $R[\overline{\Gamma}]$-module V. Then we have by Remark (2.1.1)

$$H^i([Y_{\overline{\Gamma}}], \mathbb{V}) \cong H^i(\overline{\Gamma}, V).$$

We define the *parabolic group cohomology group* as the left hand term and the *boundary group cohomology group* as the right hand term in the exact sequence

$$0 \to H^1_{\mathrm{par}}(\overline{\Gamma}, V) \to H^1(\overline{\Gamma}, V) \xrightarrow{\mathrm{res}} \bigoplus_{g \in \overline{\Gamma}\backslash \mathrm{PSL}_2(\mathbb{Z})/U} H^1(\overline{\Gamma} \cap gUg^{-1}, \mathrm{Res}^{\overline{\Gamma}}_{\overline{\Gamma}\cap gUg^{-1}} V),$$

where $U = \langle T \rangle$. We notice that $\overline{\Gamma} \cap gUg^{-1}$ is the stabiliser in $\overline{\Gamma}$ of $g\infty$.

Again, if $V = V_{k-2}(R)$, then we speak about the *(parabolic/boundary) group cohomology group of weight k over R for* Γ and similarly in the case where $\overline{\Gamma}$ is replaced by $\overline{G}$ with a character ϵ.

(2.3.1) Proposition. *For $\mathbb{V}$ a locally constant sheaf of R-modules on $[Y_{\overline{\Gamma}}]$ corresponding to an $R[\overline{\Gamma}]$-module V, the stack cohomology group for $\mathbb{V}$ and $[Y_{\overline{\Gamma}}]$ agrees with the group cohomology group for V and $\overline{\Gamma}$. This result also holds for the parabolic and the boundary spaces.*

Proof. As we have already seen that the "full" spaces agree, it suffices to prove that the boundary spaces coincide, i.e. that

$$H^0([X_{\overline{\Gamma}}], R^1 j_* \mathbb{V}) \cong \bigoplus_{g \in \overline{\Gamma} \backslash \mathrm{PSL}_2(\mathbb{Z})/U} H^1(\overline{\Gamma} \cap gUg^{-1}, V).$$

The sheaf $R^1 j_* \mathbb{V}$ is a skyscraper sheaf, whose support lies on the cusps, whence one has $H^0([X_{\overline{\Gamma}}], R^1 j_* \mathbb{V}) \cong \bigoplus_c (R^1 j_* \mathbb{V})_c$, where the sum runs over the cusps of $[X_{\overline{\Gamma}}]$. However, these cusps are in bijective correspondence with the double cosets $\overline{\Gamma} \backslash \mathrm{PSL}_2(\mathbb{Z})/U$ under the mapping $g \mapsto g\infty$. Moreover, we have that $(R^1 j_* \mathbb{V})_c$ equals $H^1(\overline{\Gamma} \cap gUg^{-1}, V)$, if the cusp c is obtained from g under the mapping just described. □

Computing group cohomology

In order to compute the group cohomology for $\overline{\Gamma}$, it suffices to compute the cohomology of $\mathrm{PSL}_2(\mathbb{Z})$-modules because of Shapiro's Lemma, which for any $R[\overline{\Gamma}]$-module V gives an isomorphism

$$H^1(\mathrm{PSL}_2(\mathbb{Z}), \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V) \cong H^1(\overline{\Gamma}, V).$$

An elementary proof of the fact that Shapiro's Lemma respects the parabolic subspace was communicated to me by Adriaan Herremans. Here, however, I shall use the representation theoretic machinery, more precisely Mackey's formula.

(2.3.2) Proposition. *Let V be a left $R[\overline{\Gamma}]$-module for a subgroup $\overline{\Gamma} \leq \mathrm{PSL}_2(\mathbb{Z})$ of finite index. The group $H^1_{\mathrm{par}}(\overline{\Gamma}, V)$ is isomorphic under the isomorphism of Shapiro's Lemma to $H^1_{\mathrm{par}}(\mathrm{PSL}_2(\mathbb{Z}), \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V)$.*

Proof. It suffices to show that $H^1(U, \mathrm{Res}_U^{\mathrm{PSL}_2(\mathbb{Z})} \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V)$ is equal to the direct sum $\bigoplus_{g \in \overline{\Gamma} \backslash \mathrm{PSL}_2(\mathbb{Z})/U} H^1(\overline{\Gamma} \cap gUg^{-1}, \mathrm{Res}^{\overline{\Gamma}}_{\overline{\Gamma} \cap gUg^{-1}} V)$. Applying Mackey's formula (see e.g. [Brown], Proposition III.5.6(b))

$$\mathrm{Res}_U^{\mathrm{PSL}_2(\mathbb{Z})} \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V = \bigoplus_{g \in U \backslash \mathrm{PSL}_2(\mathbb{Z})/\overline{\Gamma}} \mathrm{Coind}^U_{U \cap g\overline{\Gamma}g^{-1}} {}^g \mathrm{Res}^{\overline{\Gamma}}_{\overline{\Gamma} \cap g^{-1}Ug} V,$$

the isomorphism

$$H^1(U \cap g\overline{\Gamma}g^{-1}, {}^g V) \cong H^1(g^{-1}Ug \cap \overline{\Gamma}, V)$$

and sending g to g^{-1} the proposition follows from Shapiro's Lemma . □

(2.3.3) Corollary. *The boundary space $H^0([X_{\overline{\Gamma}}], R^1 j_* \mathbb{V})$ has the group cohomological description $H^1(\langle T\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}(V))$.* □

We now explicitly compute the first group cohomology of $R[\mathrm{PSL}_2(\mathbb{Z})]$-modules. A first, however, not complete description is provided by the Mayer-Vietoris sequence, using that $\mathrm{PSL}_2(\mathbb{Z})$ is the free product of the cyclic group of order 2 generated by the class of σ and the cyclic group of order 3 generated by the class of τ. The result will be important for the sequel and we record it in the following proposition.

(2.3.4) Proposition. *Let M be a left $R[\mathrm{PSL}_2(\mathbb{Z})]$-module. Then the Mayer-Vietoris sequence gives the exact sequence*

$$\begin{aligned} 0 \to & M^{\mathrm{PSL}_2(\mathbb{Z})} \to M^{\langle\sigma\rangle} \oplus M^{\langle\tau\rangle} \to M \xrightarrow{m\mapsto f_m} \\ & H^1(\mathrm{PSL}_2(\mathbb{Z}), M) \to H^1(\langle\sigma\rangle, M) \oplus H^1(\langle\tau\rangle, M) \to 0, \end{aligned}$$

where the 1-cocycle f_m uniquely given by $f_m(\sigma) = (1-\sigma)m$ and $f_m(\tau) = 0$, and for all $i \geq 2$ isomorphisms

$$H^i(\mathrm{PSL}_2(\mathbb{Z}), M) \cong H^i(\langle\sigma\rangle, M) \oplus H^i(\langle\tau\rangle, M).$$

Proof. Let us write $G := \mathrm{PSL}_2(\mathbb{Z})$, $G_1 := \langle\sigma\rangle$ and $G_2 := \langle\tau\rangle$. By [Brown], II.8.8, we have the split exact sequence of $R[G]$-modules

$$0 \to R[G] \to R[G/G_1] \oplus R[G/G_2] \to R \to 0.$$

Application of the functor $\mathrm{Hom}_R(\cdot, M)$ gives rise to the exact sequence of $R[G]$-modules

$$0 \to M \to \mathrm{Hom}_{R[G_1]}(R[G], M) \oplus \mathrm{Hom}_{R[G_2]}(R[G], M) \to \mathrm{Hom}_R(R[G], M) \to 0.$$

The central terms, as well as the term on the right, can be identified with coinduced modules. Hence, the statements follow by taking the long exact sequence of cohomology and invoking Shapiro's Lemma. □

We now derive an explicit description of the group cohomology of $\mathrm{PSL}_2(\mathbb{Z})$.

(2.3.5) Proposition. *Let M be a left $R[\mathrm{PSL}_2(\mathbb{Z})]$-module. Then we have the exact sequence*

$$0 \to M^{\mathrm{PSL}_2(\mathbb{Z})} \to M \to \ker_M(1+\sigma) \times \ker_M(1+\tau+\tau^2) \to H^1(\mathrm{PSL}_2(\mathbb{Z}), M) \to 0.$$

Proof. We determine the 1-cocycles of M. Apart from $f(1) = 0$, they must satisfy

$$0 = f(\sigma^2) = \sigma f(\sigma) + f(\sigma) = (1+\sigma) f(\sigma) \text{ and}$$

$$0 = f(\tau^3) = \cdots = (1+\tau+\tau^2) f(\tau).$$

Since these are the only relations in $\mathrm{PSL}_2(\mathbb{Z})$, a cocycle is uniquely given by the choices

$$f(\sigma) \in \ker_M(1+\sigma) \text{ and } f(\tau) \in \ker_M(1+\tau+\tau^2).$$

The 1-coboundaries are precisely those cocycles f which satisfy $f(\sigma) = (1-\sigma)m$ and $f(\tau) = (1-\tau)m$ for some $m \in M$, which proves

$$H^1(\mathrm{PSL}_2(\mathbb{Z}), M) \cong \ker_M(1+\sigma) \times \ker_M(1+\tau+\tau^2)/\big(((1-\sigma)m, (1-\tau)m) \mid m \in M\big).$$

Rewriting yields the proposition. □

(2.3.6) Remark. *As $U = \langle T \rangle < \mathrm{PSL}_2(\mathbb{Z})$ is an infinite cyclic group, one has*

$$H^1(U, \mathrm{Res}^U_G M) \cong M/(1-T)M.$$

An explicit presentation of the parabolic group cohomology is the following.

(2.3.7) Proposition. *The parabolic group cohomology group sits in the exact sequence*

$$0 \to M^{\langle T \rangle}/M^{\mathrm{PSL}_2(\mathbb{Z})} \to \ker_M(1+\sigma) \cap \ker_M(1+\tau+\tau^2) \xrightarrow{\phi} H^1_{\mathrm{par}}(\mathrm{PSL}_2(\mathbb{Z}), M) \to 0,$$

where ϕ maps an element m to the 1-cocycle f uniquely determined by $f(\sigma) = f(\tau) = m$.

Proof. Using Proposition (2.3.5), we have the exact commutative diagram

$$\begin{array}{ccccc}
M^{\langle T \rangle}/M^{\mathrm{PSL}_2(\mathbb{Z})} & \xrightarrow{(\sigma^{-1}-1)} & \ker N_\sigma \cap \ker N_\tau & \longrightarrow & H^1_{\mathrm{par}}(\mathrm{PSL}_2(\mathbb{Z}), M) \\
\downarrow{\scriptstyle \sigma^{-1}} & & \downarrow & & \downarrow \\
M/M^{\mathrm{PSL}_2(\mathbb{Z})} & \xrightarrow{(1-\sigma, 1-\tau)} & \ker N_\sigma \times \ker N_\tau & \twoheadrightarrow & H^1(\mathrm{PSL}_2(\mathbb{Z}), M) \\
\downarrow{\scriptstyle (1-T)\sigma} & & \downarrow{\scriptstyle (a,b)\mapsto b-a} & & \downarrow \\
(1-T)M & \longrightarrow & M & \twoheadrightarrow & H^1(U, M).
\end{array}$$

As the bottom left vertical arrow is surjective, the claim follows from the snake lemma. □

2.4. Cohomology of modular curves

Parabolic and boundary spaces

Let $\mathcal{F}$ be a sheaf on Y_Γ. We proceed exactly as for stacks, now with $j = j_\Gamma$ instead of $j_{[\Gamma]}$ and get the exact sequence

$$\begin{aligned} 0 \to H^1(X_\Gamma, j_*\mathcal{F}) \to H^1(Y_\Gamma, \mathcal{F}) \to H^0(X_\Gamma, R^1 j_*\mathcal{F}) \\ \to H^2(X_\Gamma, j_*\mathcal{F}) \to H^2(Y_\Gamma, \mathcal{F}) \to 0, \end{aligned}$$

since $R^2 j_* \mathcal{F} = 0$ and $H^1(X_\Gamma, R^1 j_* \mathcal{F}) = 0$.

We consider the exact sequence of sheaves on X_Γ

$$0 \to j_! \mathcal{F} \to j_* \mathcal{F} \to C \to 0,$$

in which the last term is defined as the cokernel. The *parabolic cohomology group (for Y_Γ and $\mathcal{F}$)* is image of the map $H^i_c(Y_\Gamma, \mathcal{F}) \to H^i(Y_\Gamma, \mathcal{F})$. It is denoted by $H^i_{\mathrm{par}}(Y_\Gamma, \mathcal{F})$. Moreover, we call $H^0(X_\Gamma, R^1 j_* \mathcal{F})$ the *boundary cohomology group (for Y_Γ and $\mathcal{F}$)*.

(2.4.1) Proposition. *We have $H^1_{\mathrm{par}}(Y_\Gamma, \mathcal{F}) \cong H^1(X_\Gamma, j_* \mathcal{F})$.*

Proof. The sheaf C is a skyscraper sheaf, as it is only supported on the cusps. Hence, $H^1(X_\Gamma, C) = 0$ and the long exact sequence associated to the short exact sequence of sheaves above yields that the upper map is surjective in the commutative diagram

$$\begin{array}{ccc} H^1_c(Y_\Gamma, \mathcal{F}) & \twoheadrightarrow & H^1(X_\Gamma, j_* \mathcal{F}) \\ & \searrow & \downarrow \\ & & H^1(Y_\Gamma, \mathcal{F}), \end{array}$$

in which the vertical map comes from the Leray sequence above. As it is injective, the proposition follows. □

Explicit description of the cohomology

Let V be some $R[\overline{\Gamma}]$-module. Via Remark (2.1.1), associated to it we have a locally constant sheaf $\mathbb{V}$ on the stack $[Y_{\overline{\Gamma}}]$, which we can push forward under the projection $\pi = \pi_\Gamma : [Y_{\overline{\Gamma}}] \twoheadrightarrow Y_\Gamma$.

The spaces $H^i(Y_\Gamma, \pi_* \mathbb{V}_{k-2,\overline{\Gamma}}(R))$, $H^i_{\mathrm{par}}(Y_\Gamma, \pi_* \mathbb{V}_{k-2,\overline{\Gamma}})$, $H^0(X_\Gamma, R^1 j_*(\pi_* \mathbb{V}_{k-2,\overline{\Gamma}}))$ are called the *(parabolic/boundary) cohomology group of weight k over R for Y_Γ*. We make a similar definition with the sheaf $\mathbb{V}^\epsilon_{k-2,\overline{G}}(R)$ on $[Y_{\overline{G}}]$.

(2.4.2) Proposition. *The boundary cohomology group for Y_Γ and $\pi_* \mathbb{V}$ equals the boundary stack cohomology group for $[Y_{\overline{\Gamma}}]$ and $\mathbb{V}$.*

Proof. We only need to show that

$$\left(R^1 j_* \mathbb{V}\right)_x \cong \left(R^1 j_* (\pi_* \mathbb{V})\right)_{\pi(x)}$$

for x in $[X_{\overline{\Gamma}}] - [Y_{\overline{\Gamma}}]$. That is clear, since X_Γ and $[X_{\overline{\Gamma}}]$ do not differ in a (suitably small) neighbourhood of the cusp x, when x is taken out. □

Considering the Leray spectral sequence in order to compare the cohomology of modular curves with group cohomology was suggested by Bas Edixhoven. Indeed, it even allows us to give a simple description of the cohomology of modular curves. We shall first prove a result on some second cohomology group.

(2.4.3) Lemma. *Let $\mathbb{V}$ be a locally constant sheaf on $[Y_{\overline{\Gamma}}]$. Denote by Y_Γ^0 the analytic subspace of Y_Γ obtained as the quotient by Γ of the upper half plane minus all non-trivially stabilised points (for $\overline{\Gamma}$). Denote by j^0 the embedding $Y_\Gamma^0 \hookrightarrow Y_\Gamma$.*

Then the sheaf $(j^0)_*(j^0)^*\pi_*\mathbb{V}$ is a locally constant sheaf on Y_Γ.

Proof. Write $j = j^0$ for short. Let $x \in Y_\Gamma$, which we may assume to lie in the complement of Y_Γ^0 and take $y \in [Y_{\overline{\Gamma}}]$ with $\pi(y) = x$. As $\mathbb{V}$ is locally constant, we can choose an open set $V \subset [Y_{\overline{\Gamma}}]$ containing y such that $\mathbb{V}|_V$ is constant. The quotient map π is open (*universally submersive*, see e.g. [Toen], p. 31, for algebraic stacks). So $W = \pi(V)$ is an open neighbourhood in Y_Γ containing x. For $W_1 \subseteq W$ open with $x \in W_1$ and $V_1 = \pi^{-1}(W_1)$, we have $j_*j^*\pi_*\mathbb{V}(W_1) = (\pi_*\mathbb{V})(W_1 - \{x\}) = \mathbb{V}(V_1 - \pi^{-1}(\{x\}))$, since π is a local isomorphism outside the points x resp. $\pi^{-1}(\{x\})$. Our assumption on V hence implies that $j_*j^*\pi_*\mathbb{V}|_W$ is constant. □

(2.4.4) Proposition. *Let $\mathbb{V}$ be a locally constant sheaf on $[Y_{\overline{\Gamma}}]$.*

(a) We have $H^2(Y_\Gamma, \pi_*\mathbb{V}) = 0$.

(b) We have $H^2_c(Y_\Gamma, \pi_*\mathbb{V}) = H^0([Y_{\overline{\Gamma}}], \mathbb{V}^\vee)^\vee$.

(c) For all $i \geq 2$ we have $H^i_c(Y_\Gamma, \pi_*\mathbb{V}) \cong H^i(X_\Gamma, j_*\pi_*\mathbb{V})$, where j denotes the embedding $Y_\Gamma \hookrightarrow X_\Gamma$.

Proof. We use the notations of Lemma (2.4.3). In the exact sequence of sheaves on Y_Γ

$$0 \to K \to \pi_*\mathbb{V} \to (j^0)_*(j^0)^*\pi_*\mathbb{V} \to C \to 0$$

both the kernel and the cokernel are skyscraper sheaves. As their higher cohomology vanishes, we obtain

$$H^i(Y_\Gamma, \pi_*\mathbb{V}) \cong H^i(Y_\Gamma, (j^0)_*(j^0)^*\pi_*\mathbb{V}) \quad \text{for all } i \geq 2$$

and similarly for compactly supported cohomology. We may apply Poincaré duality to $H^2(Y_\Gamma, (j^0)_*(j^0)^*\pi_*\mathbb{V})$ and $H^2_c(Y_\Gamma, (j^0)_*(j^0)^*\pi_*\mathbb{V})$. It yields that the first space is isomorphic to $H^0_c(Y_\Gamma, ((j^0)_*(j^0)^*\pi_*\mathbb{V})^\vee)^\vee$, which is zero, as Y_Γ is non-compact and connected and the sheaf $((j^0)_*(j^0)^*\pi_*\mathbb{V})^\vee$ is locally constant, proving (a). Poincaré duality furthermore gives

$$H^2_c(Y_\Gamma, (j^0)_*(j^0)^*\pi_*\mathbb{V}) \cong H^0(Y_\Gamma, ((j^0)_*(j^0)^*\pi_*\mathbb{V})^\vee)^\vee \cong H^0(Y_\Gamma^0, (\pi_*\mathbb{V})^\vee|_{Y_\Gamma^0})^\vee.$$

The latter space is isomorphic to $H^0([Y_{\overline{\Gamma}}]^0, \mathbb{V}^\vee|_{[Y_{\overline{\Gamma}}]^0})^\vee$, which in turn itself is equal to $H^0([Y_{\overline{\Gamma}}], \mathbb{V}^\vee)^\vee$, proving (b).

Part (c) follows immediately from the exact sequence of sheaves on X_Γ

$$0 \to j_!\pi_*\mathbb{V} \to j_*\pi_*\mathbb{V} \to C' \to 0,$$

as the cokernel is again a skyscraper sheaf. □

We now compare the cohomology groups of the modular stack to that of the modular curve via the Leray spectral sequence. It gives rise to the short exact sequence

$$0 \to H^1(Y_\Gamma, \pi_*\mathbb{V}) \to H^1([Y_{\overline{\Gamma}}], \mathbb{V}) \to H^0(Y_\Gamma, R^1\pi_*\mathbb{V}) \to 0,$$

as $H^2(Y_\Gamma, \pi_*\mathbb{V}) = 0$ by Proposition (2.4.4). The sheaf $R^1\pi_*\mathbb{V}$ is a skyscraper sheaf, supported only on non-trivially stabilised points. More precisely, if Γ_x denotes the stabiliser group of Γ at the point $x \in \mathbb{H}$, then

$$(R^1\pi_*\mathbb{V})_x = H^1(\Gamma_x, V).$$

(2.4.5) Proposition. *We have the exact sequence of R-modules*

$$\begin{aligned} 0 \to H^1(Y_\Gamma, \pi_*\mathbb{V}) &\to H^1([Y_{\overline{\Gamma}}], \pi_*\mathbb{V}) \\ &\to H^1(\langle\sigma\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V) \oplus H^1(\langle\tau\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V) \to 0. \end{aligned}$$

Proof. We first note that any non-trivially stabilised point x of $\mathbb{H}$ is conjugate by some $g \in \mathrm{PSL}_2(\mathbb{Z})$ to either i or ζ_3, whence the stabiliser group then is $g\langle\sigma\rangle g^{-1} \cap \overline{\Gamma}$ or $g\langle\tau\rangle g^{-1} \cap \overline{\Gamma}$. As in the proof of Proposition (2.3.2) we can apply Mackey's formula to obtain

$$H^1(\langle\sigma\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V) \cong \bigoplus_{g \in \overline{\Gamma}\backslash \mathrm{PSL}_2(\mathbb{Z})/\langle\sigma\rangle} H^1(g\langle\sigma\rangle g^{-1} \cap \overline{\Gamma}, V)$$

and a similar result for τ. So we get

$$H^0(Y_\Gamma, R^1\pi_*\mathbb{V}) \cong H^1(\langle\sigma\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V) \oplus H^1(\langle\tau\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V),$$

which finishes the proof. □

We have already earlier encountered the very same obstruction term, namely in the Mayer-Vietoris sequence (see Proposition (2.3.4)). This establishes the following theorem.

(2.4.6) Theorem. *For any ring R, any congruence subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ and any $R[\overline{\Gamma}]$-module V with associated sheaf $\mathbb{V}$ on $[Y_{\overline{\Gamma}}]$, we have*

$$H^1(Y_\Gamma, \pi_*\mathbb{V}) \cong M/\big(M^{\langle\sigma\rangle} + M^{\langle\tau\rangle}\big)$$

with $M = \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}(V)$ and $\pi : [Y_{\overline{\Gamma}}] \twoheadrightarrow Y_\Gamma$ the natural projection.

We let

$$\mathcal{H}_k(\Gamma, R) = M/\big(M^{\langle\sigma\rangle} + M^{\langle\tau\rangle}\big)$$

as in the theorem with $M = \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_{k-2}(R))$ and define $\mathcal{CH}_k(\Gamma, R)$ as the kernel of the boundary map

$$M/\big(M^{\langle\sigma\rangle} + M^{\langle\tau\rangle}\big) \xrightarrow{m \mapsto (1-\sigma)m} M/(1-T)M.$$

We call $\mathcal{CH}_k(\Gamma, R)$ the *parabolic* subspace and the space on the right the *boundary space*.

Moreover, we let $\mathcal{H}_k(G, \epsilon, R) := M/\big(M^{\langle\sigma\rangle} + M^{\langle\tau\rangle}\big)$ for $M = \mathrm{Coind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_{k-2}^{\epsilon}(R))$ and similarly as above we define a parabolic and a boundary space.

Merel's study of homology

A study of the homology of modular curves (as Riemann surfaces) has been carried out by [MerelHecke] also in order to compute modular forms. We shall see that Merel's explicit description is a special case of ours.

The first homology group relative to the cusps features in the long exact sequence

$$0 \to H_1(X_\Gamma, R) \to H_1(X_\Gamma, \mathrm{cusps}, R) \to R[\mathrm{cusps}] \to R \to 0.$$

From this sequence it follows that $H_1(X_\Gamma, \mathrm{cusps}, R)$ is a free R-module (as $H_1(X_\Gamma, R)$ is free, which is well known for compact Riemann surfaces).

(2.4.7) Proposition. *We have isomorphisms*

$$H_1(X_\Gamma, \mathrm{cusps}, R) \cong H^1(Y_\Gamma, R) \cong H_1(Y_\Gamma, R)^\vee.$$

Proof. The first isomorphism is a simple application of the general duality theorem [Dold], Proposition VIII.7.2, noting that in this case Čech cohomology coincides with singular cohomology (see e.g. [Dold], Proposition VIII.6.12). The second isomorphism is a consequence of the universal coefficient theorem. □

In view of Proposition (2.4.7), the description of the relative homology group of [MerelHecke], Proposition 4,

$$H_1(X_\Gamma, \mathrm{cusps}, R) \cong \mathcal{H}_2(\Gamma, R)$$

is now immediate.

Torsion-freeness and base change properties

Merel's original computation of $H_1(X_\Gamma, \mathrm{cusps}, \mathbb{Z})$ as $\mathcal{H}_2(\Gamma, \mathbb{Z})$ was to compute the torsion-freeness of the latter module and to show that its rank is right. More generally, Herremans has computed the torsion in the $\Gamma_1(N)$-Manin symbols over $\mathbb{Z}$ ([Herremans], Proposition 9). We will, however, give a geometric and more general proof of torsion-freeness, which Bas Edixhoven has explained to the author.

(2.4.8) Proposition. *Assume that R is an integral domain of characteristic 0 such that $R/pR \cong \mathbb{F}_p$. Let $N \geq 1$ be an integer such that $p \nmid N$. We assume that $\Gamma_1(N) \leq \Gamma$ and that the stabilisers for the action of $\overline{G}$ on $\mathbb{H}$ have order invertible in R, or that $k = 2$ and ϵ is trivial.*

We denote by $\overline{\epsilon}$ the reduction modulo p of ϵ. Recall that π_G denotes the projection $[Y_{\overline{G}}] \twoheadrightarrow Y_G$. Write $\mathcal{F}(R) = \pi_{G,*}\mathbb{V}^{\epsilon}_{k-2,\overline{G}}(R)$ and similarly for $\mathcal{F}(\mathbb{F}_p)$. Then the following statements hold:

(a) We have $H^1_c(Y_G, \mathcal{F}(R)) \otimes_R \mathbb{F}_p \cong H^1_c(Y_G, \mathcal{F}(\mathbb{F}_p))$.

(b) We have an isomorphism $H^1(Y_G, \mathcal{F}(R)) \otimes_R \mathbb{F}_p \cong H^1(Y_G, \mathcal{F}(\mathbb{F}_p))$. If $k = 2$ and ϵ is trivial, $H^1(Y_G, \mathcal{F}(R))[p] = 0$ holds. Otherwise, the p-torsion $H^1(Y_G, \mathcal{F}(R))[p]$ is isomorphic with ${}^{\mathrm{SL}_2(\mathbb{F}_p)}V^{\overline{\epsilon}}_{k-2}(\mathbb{F}_p)$.

(c) We have $H^1_{\mathrm{par}}(Y_G, \mathcal{F}(R)) \otimes_R \mathbb{F}_p \cong H^1_{\mathrm{par}}(Y_G, \mathcal{F}(\mathbb{F}_p))$.

Proof. Let us first notice that the sequence

$$0 \to \mathbb{V}^{\epsilon}_{k-2,\overline{G}}(R) \xrightarrow{\cdot p} \mathbb{V}^{\epsilon}_{k-2,\overline{G}}(R) \to \mathbb{V}^{\overline{\epsilon}}_{k-2,\overline{G}}(\mathbb{F}_p) \to 0$$

of sheaves on $[Y_{\overline{G}}]$ is exact. Applying the left exact functor $\pi_{G,*}$ we obtain the short exact sequence of sheaves on Y_G

$$0 \to \mathcal{F}(R) \xrightarrow{\cdot p} \mathcal{F}(R) \to \mathcal{F}(\mathbb{F}_p) \to 0,$$

because we have seen before that $R^1\pi_{G,*}\mathbb{V}^{\epsilon}_{k-2,\overline{G}}(R)$ is a skyscraper sheaf supported only on non-trivially stabilised points and there the stalk is $H^1(\overline{G}_x, V^{\epsilon}_{k-2}(R))$, which is 0 by assumption, as either the order of $\overline{G}_x$ is invertible or $V^{\epsilon}_{k-2}(R) = R$. The associated long exact sequence gives rise to the short exact sequence

$$0 \to H^i(Y_G, \mathcal{F}(R)) \otimes \mathbb{F}_p \to H^i(Y_G, \mathcal{F}(\mathbb{F}_p)) \to H^{i+1}(Y_G, \mathcal{F}(R))[p] \to 0$$

for every $i \geq 0$. A similar exact sequence also follows by taking compactly supported cohomology.

We have $H^2(Y_G, \mathcal{F}(R)) = 0$ and $H^2_c(Y_G, \mathcal{F}(R))[p] = 0$. The former was proved in Proposition (2.4.4). The latter can also be deduced from that proposition, as $H^2_c(Y_G, \mathcal{F}(R))$ is a free R-module, since it is isomorphic to $H^0([Y_{\overline{G}}], \mathbb{V}_{k-2,\overline{G}}(R)^\vee)^\vee$. This proves (a) and the base change part of (b).

We finish Part (b) by the isomorphism $H^0(Y_G, \mathcal{F}(\mathbb{F}_p)) \cong {}^{\overline{G}}V^{\overline{\epsilon}}_{k-2}(\mathbb{F}_p)$ and the fact that $H^0(Y_G, \mathcal{F}(R)) \cong {}^{\overline{G}}V^{\epsilon}_{k-2}(R)$ is zero, unless $k = 2$ and ϵ is trivial by Propositions (2.2.6) and (2.2.7).

Part (c) is a direct consequence of (a) and (b), since parabolic cohomology is the image of compactly supported cohomology in the usual one. □

(2.4.9) Remark. *We can use the short exact sequence $0 \to j_!\mathcal{F} \to j_*\mathcal{F} \to C \to 0$ to compare compactly supported cohomology with parabolic cohomology. Namely, the associated long exact sequence gives rise to the exact sequence*

$$0 \to {}^{\overline{\Gamma}}V_{k-2}(R) \to \bigoplus_{\mathrm{cusps}} R \to H^1_c(Y_\Gamma, \mathbb{V}_{k-2,\overline{\Gamma}}(R)) \to H^1_{\mathrm{par}}(Y_\Gamma, \mathbb{V}_{k-2,\overline{\Gamma}}(R)) \to 0.$$

We omit the details, as this will not be used in the sequel.

2.5. Modular symbols

Definition

Modular symbols can be thought of as geodesic paths between two cusps resp. as the associated homology class relative to the cusps. We shall, however, give a combinatorial definition, as is implemented in Magma and like the one in Stein's thesis [SteinThesis], except that we do not factor out torsion, but intend a common treatment for all rings. We keep the Notation (2.0.3).

(2.5.1) Definition. *We define the R-modules*

$$\mathcal{M}_2(R) := R[\{\alpha,\beta\}|\alpha,\beta\in\mathbb{P}^1(\mathbb{Q})]/\langle\{\alpha,\alpha\},\{\alpha,\beta\}+\{\beta,\gamma\}+\{\gamma,\alpha\}|\alpha,\beta,\gamma\in\mathbb{P}^1(\mathbb{Q})\rangle$$

and

$$\mathcal{B}_2(R) := R[\mathbb{P}^1(\mathbb{Q})],$$

which we equip with the natural left $\mathrm{PSL}_2(\mathbb{Z})$*-action. Furthermore, we let*

$$\mathcal{M}_k^\epsilon(R) := \mathcal{M}_2(R)\otimes_R V_{k-2}^\epsilon(R)$$

and

$$\mathcal{B}_k^\epsilon(R) := \mathcal{B}_2(R)\otimes_R V_{k-2}^\epsilon(R)$$

for the left diagonal $\overline{G}$-action. If ϵ is the trivial character, we usually drop it from the notation.

(a) We call the (left-)coinvariants

$$\mathcal{M}_k(G,\epsilon,R) := {}_{\overline{G}}\mathcal{M}_k^\epsilon(R) = \mathcal{M}_k(R)/\langle(x-gx)|g\in\overline{G},x\in\mathcal{M}_k^\epsilon(R)\rangle$$

the space of G-modular symbols of weight k over R (for the character ϵ).

(b) We call the (left-)coinvariants

$$\mathcal{B}_k(G,\epsilon,R) := {}_{\overline{G}}\mathcal{B}_k^\epsilon(R) = \mathcal{B}_k(R)/\langle(x-gx)|g\in\overline{G},x\in\mathcal{B}_k^\epsilon(R)\rangle$$

the space of G-boundary symbols of weight k over R (for the character ϵ).

(c) We define the boundary map *as the map*

$$\mathcal{M}_k(G,\epsilon,R)\to\mathcal{B}_k(G,\epsilon,R)$$

which is induced from the map $\mathcal{M}_2(R)\to\mathcal{B}_2(R)$ sending $\{\alpha,\beta\}$ to $\{\beta\}-\{\alpha\}$.

(d) The kernel of the boundary map is denoted by $\mathcal{CM}_k(G,\epsilon,R)$ and is called the space of cuspidal G-modular symbols of weight k over R (for the character ϵ).

(e) The image of the boundary map inside $\mathcal{B}_k(G,\epsilon,R)$ is denoted by $\mathcal{E}_k(G,\epsilon,R)$ and is called the space of G-Eisenstein symbols of weight k over R (for the character ϵ).

The definitions can be summarised in the exact sequence

$$0 \to \mathcal{CM}_k(G,\epsilon,R) \to \mathcal{M}_k(G,\epsilon,R) \to \mathcal{E}_k(G,\epsilon,R) \to 0.$$

In the standard situation that $\Gamma = \Gamma_1(N)$ and $G = \Gamma_0(N)$, we can make the identification

$$\Gamma_1(N)\backslash\Gamma_0(N) \to (\mathbb{Z}/N\mathbb{Z})^*, \qquad \begin{pmatrix} a & b \\ Nc & d \end{pmatrix} \mapsto a.$$

In the definitions above it seems natural to write e.g. $\mathcal{M}_k(G,\epsilon,R)$ and not $\mathcal{M}_k(\Gamma,\epsilon,R)$, which would be closer to the usual notation for modular forms, namely $S_k(\Gamma,\epsilon,R)$.

(2.5.2) Remark. *The map*

$$\mathcal{M}_2(\mathbb{Z}) \to \mathrm{Div}^0(\mathbb{P}^1(\mathbb{Q})), \quad \{\alpha,\beta\} \mapsto \beta - \alpha$$

is an isomorphism of left $\mathrm{PSL}_2(\mathbb{Z})$*-modules.*

Indeed, surjectivity is clear. The elements $\{\alpha,\alpha\}$ and $\{\alpha,\beta\}+\{\beta,\gamma\}+\{\gamma,\alpha\}$ are in the kernel. These generate all relations of the form $\{\alpha_1,\alpha_2\}+\{\alpha_2,\alpha_3\}+\cdots+\{\alpha_n,\alpha_1\}$ for $n \geq 1$. But the kernel is generated by these.

Ash and Stevens (in [Ash-Stevens]) call $\mathrm{Hom}_\Gamma(\mathrm{Div}^0(\mathbb{P}^1(\mathbb{Q})), R)$ the space of modular symbols. This is thus precisely the R-dual of the module considered here.

We end this section by a remark on changing the coefficient ring.

(2.5.3) Remark. *Let $R \to S$ be a ring homomorphism. As tensoring, as well as taking coinvariants, is right exact, we have*

$$\mathcal{M}_k(G,R) \otimes_R S \cong \mathcal{M}_k(G,S) \quad \textit{and} \quad \mathcal{B}_k(G,R) \otimes_R S \cong \mathcal{B}_k(G,S).$$

If $R \to S$ is flat, also $\mathcal{CM}_k(G,R) \otimes_R S \cong \mathcal{CM}_k(G,S)$ holds. Similar statements are true with a character ϵ. □

Manin symbols

Manin symbols provide an explicit description of modular symbols. We stay in the general setting over a ring R. Most proofs that modular and Manin symbols coincide (e.g. Merel in [MerelUniversal]) use Manin's original homological approach [Manin] or its generalisation by [Šokurov]. In this section we show, using a combinatorial proposition due to Martin, that the identification is purely algebraic.

Martin has the following purely algebraic proposition, the proof of which is combinatorial in nature. It is Proposition 4.3 in his thesis [Martin].

(2.5.4) Proposition. (Martin) *We consider the homomorphism*

$$\psi : \mathbb{Z}[\mathrm{SL}_2(\mathbb{Z})] \to \mathbb{Z}[\mathbb{P}^1(\mathbb{Q})], M \mapsto M.\infty - M.0.$$

Its kernel is given by $\mathbb{Z}[\mathrm{SL}_2(\mathbb{Z})](1+\sigma) + \mathbb{Z}[\mathrm{SL}_2(\mathbb{Z})](1-T-T')$.

Translating this proposition into the theory of Manin symbols, one obtains the following proposition.

(2.5.5) Proposition. *The homomorphism of R-modules*

$$R[\mathrm{PSL}_2(\mathbb{Z})] \xrightarrow{\phi} \mathcal{M}_2(R), g \mapsto \{g.0, g.\infty\}$$

is surjective and its kernel is given by $R[\mathrm{PSL}_2(\mathbb{Z})](1+\sigma) + R[\mathrm{PSL}_2(\mathbb{Z})](1+\tau+\tau^2)$.

Proof. For the surjectivity we follow [Cremona], p. 14. It suffices to prove that the element $\{\infty, \alpha\}$ with α a rational number is hit. Let $\alpha = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots \frac{1}{a_k}}}$ be the continued fractions expansion of α and let $P_l = \left(\begin{smallmatrix} a_1 & 1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} a_2 & 1 \\ 1 & 0 \end{smallmatrix}\right)\cdots\left(\begin{smallmatrix} a_l & 1 \\ 1 & 0 \end{smallmatrix}\right)$ for all $l \in \{1, \ldots, k\}$. We may write $P_l = \left(\begin{smallmatrix} p_l & p_{l-1} \\ q_l & q_{l-1} \end{smallmatrix}\right)$ with $p_0 = 1$ and $q_0 = 0$. By construction we have $\alpha = \frac{p_k}{q_k}$ and $\infty = \frac{p_0}{q_0}$. Consequently, we obtain

$$\left(\begin{smallmatrix} -p_1 & p_0 \\ -q_1 & q_0 \end{smallmatrix}\right)\{0,\infty\} + \left(\begin{smallmatrix} p_2 & p_1 \\ q_2 & q_1 \end{smallmatrix}\right)\{0,\infty\} + \cdots + \left(\begin{smallmatrix} (-1)^k p_k & p_{k-1} \\ (-1)^k q_k & q_{k-1} \end{smallmatrix}\right)\{0,\infty\} = \{\infty, \alpha\}.$$

We notice that due to the extra minus signs the determinants of all matrices equal 1.

Let us now notice that tensoring by R we may work with R-modules instead of $\mathbb{Z}$-modules. Moreover, as $(1-\sigma)(1+\sigma) = 1-(-\mathbf{1})$, we may replace $\mathrm{SL}_2(\mathbb{Z})$ by $\mathrm{PSL}_2(\mathbb{Z})$ in Proposition (2.5.4).

Next we show that $\ker(\phi) = \ker(\psi)$, using the homomorphism

$$\pi : \mathcal{M}_2(R) \to R[\mathbb{P}^1(\mathbb{Q})], \quad \{\alpha, \beta\} \mapsto \beta - \alpha.$$

As $\psi = \pi \circ \phi$, the inclusion $\ker(\phi) \subseteq \ker(\psi)$ follows. For the other one we assume that $\sum_M u_M[M] \in \ker(\psi)$, i.e.

$$0 = \sum_M u_M(M.0 - M.\infty) = (\sum_M u_M M.0) - (\sum_M u_M M.\infty).$$

But then $\sum_M u_M M\{0,\infty\} = (\sum_M u_M\{M.0, \infty\}) - (\sum_M u_M\{M.\infty, \infty\}) = 0$, establishing the converse inclusion.

Now it only remains to establish the claimed form of the kernel. We have the identities $\tau = T\sigma$, $T' = \tau^2\sigma^{-1}$ and consequently $1 - T - T' = ((1+\sigma) - (1+\tau+\tau^2))\sigma^{-1}$. The latter one implies for all $R[\mathrm{PSL}_2(\mathbb{Z})]$-modules M the identity

$$(1-T-T')M + (1+\sigma)M = (1+\tau+\tau^2)M + (1+\sigma)M,$$

which finishes the proof. □

If V is any left $R[\mathrm{PSL}_2(\mathbb{Z})]$-module, the induced module $\mathrm{Ind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}(V)$ is by definition the left $\mathrm{PSL}_2(\mathbb{Z})$-module $R[\mathrm{PSL}_2(\mathbb{Z})]\otimes_{R[\overline{\Gamma}]} V$, where $R[\mathrm{PSL}_2(\mathbb{Z})]$ is equipped with the natural right $R[\overline{\Gamma}]$-action and the left $R[\mathrm{PSL}_2(\mathbb{Z})]$-action. Sending $g\otimes_\Gamma v$ to $g^{-1}\otimes v$ establishes an isomorphism of $\mathrm{Ind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}(V)$ with ${}_{\overline{\Gamma}}(R[\mathrm{PSL}_2(\mathbb{Z})]\otimes_R V)$, where now $\overline{\Gamma}$ acts diagonally from the left. The left $R[\mathrm{PSL}_2(\mathbb{Z})]$-action is the one obtained by inversion from the natural right action. We will in the sequel consider the module ${}_{\overline{\Gamma}}(R[\mathrm{PSL}_2(\mathbb{Z})]\otimes_R V)$ with this right action.

(2.5.6) Theorem. *Let $M := \mathrm{Ind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_{k-2}^{\epsilon}(R))$ be the induced module with the right $R[\mathrm{PSL}_2(\mathbb{Z})]$-action described directly before the theorem. Then the following statements hold:*

(a) The homomorphism ϕ from Proposition (2.5.5) induces the exact sequence of R-modules

$$0\to M(1+\sigma)+M(1+\tau+\tau^2)\to M\to \mathcal{M}_k(G,\epsilon,R)\to 0.$$

(b) The homomorphism $R[\mathrm{PSL}_2(\mathbb{Z})]\to R[\mathbb{P}^1(\mathbb{Q})]$ sending $[g]$ to $g.\infty$ induces the exact sequence of R-modules

$$0\to M(1-T)\to M\to \mathcal{B}_k(G,\epsilon,R)\to 0.$$

(c) Under the identifications of (a) and (b) the boundary map is the map

$$M/\big(M(1+\sigma)+M(1+\tau+\tau^2)\big)\to M/\big(M(1-T)\big)$$

induced from $m\mapsto m(1-\sigma)$ on M.

Proof. (a) We derive this from Proposition (2.5.5), which gives the exact sequence

$$0\to R[\mathrm{PSL}_2(\mathbb{Z})](1+\sigma)+R[\mathrm{PSL}_2(\mathbb{Z})](1+\tau+\tau^2)\to R[\mathrm{PSL}_2(\mathbb{Z})]\to \mathcal{M}_2(R)\to 0.$$

Let $N := R[\mathrm{PSL}_2(\mathbb{Z})]\otimes_R V_{k-2}^{\epsilon}(R)$, which we equip with the right $\mathrm{PSL}_2(\mathbb{Z})$-action $([g]\otimes v).[\sigma]=[g\sigma]\otimes v$. As $V_{k-2}^{\epsilon}(R)$ is a free R-module we obtain the exact sequence of left $R[\overline{G}]$-modules

$$0\to N(1+\sigma)+N(1+\tau+\tau^2)\to N\to \mathcal{M}_k^{\epsilon}(R)\to 0.$$

Passing to left $\overline{G}$-coinvariants yields (a).

(b) It is easy to compute that the described map fits into the exact sequence

$$0\to R[\mathrm{PSL}_2(\mathbb{Z})](1-T)\to R[\mathrm{PSL}_2(\mathbb{Z})]\to R[\mathbb{P}^1(\mathbb{Q})]\to 0.$$

Now we can proceed precisely as in (a) and obtain (b).

(c) It is clear that this map corresponds to the boundary map. It is well defined because of $(1+\tau+\tau^2)(1-\sigma)=(1+\tau+\tau^2)(1-T)$. □

In the literature on Manin symbols one usually finds a more explicit version of the module M. This is the contents of the following proposition.

(2.5.7) Proposition. *(a) Consider the R-module $X := R[\Gamma\backslash \mathrm{SL}_2(\mathbb{Z})] \otimes_R V_{k-2}(R) \otimes_R R^\epsilon$ equipped with the right $\mathrm{SL}_2(\mathbb{Z})$-action $(\Gamma h \otimes V \otimes r)g = (\Gamma hg \otimes g^{-1}v \otimes r)$ and with the left $\Gamma\backslash G$-action $g(\Gamma h \otimes v \otimes r) = (\Gamma gh \otimes v \otimes \epsilon(g)r)$.*

Then X is isomorphic as a right $R[\mathrm{SL}_2(\mathbb{Z})]$-module and a left $R[\Gamma\backslash G]$-module to $\mathrm{Ind}_\Gamma^{\mathrm{SL}_2(\mathbb{Z})}(V_k^\epsilon(R))$, and, moreover, ${}_{\Gamma\backslash G}X$ is isomorphic to $\mathrm{Ind}_G^{\mathrm{SL}_2(\mathbb{Z})}(V_k^\epsilon(R))$.

If $-\mathbf{1} \in G$ and $-\mathbf{1} \notin \Gamma$, then the latter module is isomorphic to $\mathrm{Ind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_k^\epsilon(R))$.

(b) Consider the module $X := R[\overline{G}\backslash \mathrm{PSL}_2(\mathbb{Z})] \otimes_R V_{k-2}(R) \otimes_R R^\epsilon$ equipped with the right $\mathrm{PSL}_2(\mathbb{Z})$-action $(\Gamma h \otimes V \otimes r)g = (\Gamma hg \otimes g^{-1}v \otimes r)$.

If $(-1)^k = 1$, then X is isomorphic to $\mathrm{Ind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_k^\epsilon(R))$ as a right $R[\mathrm{PSL}_2(\mathbb{Z})]$-module.

Proof. (a) Mapping $g \otimes v \otimes r$ to $g \otimes g^{-1}v \otimes r$ defines an isomorphism of right $R[\mathrm{SL}_2(\mathbb{Z})]$-modules and of left $R[\Gamma\backslash G]$-modules

$$ {}_\Gamma(R[\mathrm{SL}_2(\mathbb{Z})] \otimes_R V_{k-2}(R) \otimes_R R^\epsilon) \to X. $$

As we have seen above, the left hand side module is naturally isomorphic to the induced module $\mathrm{Ind}_\Gamma^{\mathrm{SL}_2(\mathbb{Z})}(V_k^\epsilon(R))$ (equipped with its right $R[\mathrm{SL}_2(\mathbb{Z})]$-action described before). This establishes the first statement. The second one follows from ${}_{\Gamma\backslash G}\big({}_\Gamma M\big) = {}_G M$ for any G-module M. The third statement is due to the fact that ${}_{\langle -\mathbf{1}\rangle}(R[\mathrm{SL}_2(\mathbb{Z})] \otimes_R V_{k-2}^\epsilon(R))$ is naturally isomorphic to $R[\mathrm{PSL}_2(\mathbb{Z})] \otimes_R V_{k-2}^\epsilon(R)$, since $-\mathbf{1}$ acts trivially on the second factor.

(b) This works analogously to the discussion in (a) with $\mathrm{SL}_2(\mathbb{Z})$ replaced by $\mathrm{PSL}_2(\mathbb{Z})$ because we can now view $V_{k-2}(R)$ as a $\mathrm{PSL}_2(\mathbb{Z})$-module. □

Transportable Modular Symbols

In this section I present Stein's and Verrill's definition of transportable modular symbols, and reprove their principal theorem (see [SteinVerrill], Theorem 2.4). The difference is that I prove the result over any ring R, whereas the original proof was for modular symbols over $\mathbb{Z}$ *modulo torsion*. This section is not used in the sequel, but can serve as an illustration that working with the torsion can make things much easier.

Transportable modular symbols are used to compute periods of modular symbols resp. modular forms. The aim is to *transport* a path from the cusp $\{\alpha\}$ to $\{\infty\}$ to a path from z to γz for a well chosen z in the upper half plane (for some $\gamma \in \Gamma$) representing the same homology class.

We shall not restate the original definition of transportable modular symbols, but the equivalent variant of [SteinVerrill], Lemma 2.3 (The equivalence works over any ring, not only $\mathbb{Q}$).

(2.5.8) Definition. *A modular symbol $x \in \mathcal{M}_k(\Gamma, R)$ is called* transportable *if it can be written in the form $\sum_{i=1}^m \{\infty, \gamma_i\infty\} \otimes P_i$ with $\gamma_i \in \Gamma$ and $P_i \in V_{k-2}(R)$ such that $\sum_{i=1}^m P_i = \sum_{i=1}^m \gamma_i^{-1} P_i$.*

To make the last formula a little more understandable (and set a decisive step towards proving the principal theorem in this context), let us note that by a straight forward calculation a symbol of the form $\sum_{i=1}^m \{\infty, \gamma_i\infty\} \otimes P_i$ is cuspidal (i.e. in the kernel of the boundary map) if and only if $\sum_{i=1}^m P_i = \sum_{i=1}^m \gamma_i^{-1} P_i$ holds.

(2.5.9) Theorem. (Stein, Verrill) *A modular symbol is transportable if and only if it is cuspidal.*

Proof. Choose a system of representatives $\mathcal{R}$ of $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$, representing $\Gamma\infty$ by ∞. Let us suppose that $x \in \mathcal{CM}_k(\Gamma, R)$. Writing $\{\alpha, \beta\} \otimes P = \{\infty, \beta\} \otimes P - \{\infty, \alpha\} \otimes P$ and using the Γ-invariance, we write

$$x = \sum_{\beta \in \mathcal{R}} \sum_{\gamma \in \Gamma} \{\infty, \gamma\beta\} \otimes P_{\gamma,\beta}.$$

By assumption x is in the kernel of the boundary map, i.e.

$$\sum_{\beta \in \mathcal{R}} \sum_{\gamma \in \Gamma} \gamma\{\beta\} \otimes P_{\gamma,\beta} = \sum_{\beta \in \mathcal{R}} \sum_{\gamma \in \Gamma} \{\infty\} \otimes P_{\gamma,\beta} \in {}_\Gamma\mathcal{B}_k(R).$$

For $\infty \neq \beta \in \mathcal{R}$ it follows

$$\sum_{\gamma \in \Gamma} \gamma^{-1} P_{\gamma,\beta} = 0,$$

which in turn yields

$$\begin{aligned}
\sum_{\gamma \in \Gamma} \{\infty, \gamma\beta\} \otimes P_{\gamma,\beta} &= \sum_{\gamma \in \Gamma} (\{\infty, \gamma\infty\} + \{\gamma\infty, \gamma\beta\}) \otimes P_{\gamma,\beta} \\
&= \sum_{\gamma \in \Gamma} \{\infty, \gamma\infty\} \otimes P_{\gamma,\beta} + \sum_{\gamma \in \Gamma} \{\infty, \beta\} \otimes \gamma^{-1} P_{\gamma,\beta} \\
&= \sum_{\gamma \in \Gamma} \{\infty, \gamma\infty\} \otimes P_{\gamma,\beta},
\end{aligned}$$

finishing the proof. □

2.6. Comparison between the spaces

In group cohomology one conceptually has to work with coinduced modules. However, if the index is finite, which is the case in all our considerations, one can identify induced and coinduced modules. In the section about Manin symbols we have considered the induced modules as right modules by inverting the natural left action. This was done in order to stay

close to other treatments, e.g. [SteinThesis]. Here, however, we will go back to the natural left action. An analog of Theorem (2.5.6) for left actions is obtained by formally rewriting all right actions into left ones.

We still assume Notation (2.0.3).

(2.6.1) Theorem. *The boundary spaces of modular symbols, group cohomology and of the cohomology of modular curves agree, i.e.*

$$\mathcal{B}_k(G,\epsilon,R) \cong H^1(\langle T\rangle, \mathrm{Coind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_k^\epsilon(R))) \cong H^0(X_G, R^1 j_{G,*}(\pi_{G,*}\mathbb{V}_k^\epsilon(R))).$$

Assuming further that the orders of all stabilisers of $\overline{G}$ acting on $\mathbb{H}$ are invertible in R, then also the full spaces of modular symbols, group cohomology and of the cohomology of modular curves are isomorphic, i.e.

$$\mathcal{M}_k(G,\epsilon,R) \cong H^1(\overline{G}, V_k^\epsilon(R)) \cong H^1(Y_G, \pi_{G,*}\mathbb{V}_k^\epsilon(R)),$$

as are their parabolic resp. cuspidal subspaces

$$\mathcal{CM}_k(G,\epsilon,R) \cong H^1_{\mathrm{par}}(\overline{G}, V_k^\epsilon(R)) \cong H^1_{\mathrm{par}}(Y_G, \pi_{G,*}\mathbb{V}_k^\epsilon(R)).$$

Proof. Because of Proposition (2.3.1), Theorem (2.5.6) (b), Corollary (2.3.3), Proposition (2.4.2) and Remark (2.3.6), the boundary spaces agree.

Using Mackey's formula as in Proposition (2.4.5) we get

$$H^1(\langle\sigma\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V_{k-2}^\epsilon(R)) \cong \bigoplus_{g\in\overline{\Gamma}\backslash\mathrm{PSL}_2(\mathbb{Z})/\langle\sigma\rangle} H^1(g\langle\sigma\rangle g^{-1}\cap\overline{\Gamma}, V_{k-2}^\epsilon(R))$$

and a similar result for τ. The right hand side is zero due to the assumption on the stabiliser order. A similar result holds for the corresponding first homology group, which using cyclicity gives $\hat{H}^0(\langle\sigma\rangle, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V_{k-2}^\epsilon(R)) = 0$.

The vanishing of the first cohomology group implies via Theorem (2.4.6) and Proposition (2.3.4) that the full spaces of group cohomology and the cohomology of modular curves agree. The former always coincides with the cohomology of the modular stack by Proposition (2.3.1). The vanishing of the $\hat{H}^0$-term means written out that

$$(1+\sigma)\mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V_{k-2}^\epsilon(R) = \left(\mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V_{k-2}^\epsilon(R)\right)^{\langle\sigma\rangle}$$

and similarly for τ, which via Theorems (2.5.6) and (2.4.6) establishes the comparison between modular symbols and the cohomology of modular curves.

As we have seen that the boundary spaces and the full spaces coincide, the same follows for the parabolic resp. cuspidal spaces, as the boundary maps are compatible. □

If $k = 2$ and ϵ is trivial, it actually suffices for the comparison between group cohomology and the cohomology of modular curves to assume that the stabiliser orders are no zero

divisors, as then $V_{k-2}^{\epsilon}(R) = R$ and the H^1-terms in the proof above vanish (but not the $\hat{H}^0$-terms in general).

The stabilisers of the action of $\mathrm{PSL}_2(\mathbb{Z})$ on $\mathbb{H}$ all have order dividing 6. The following proposition investigates, when precisely stabilisers of order 2 or 3 occur.

(2.6.2) Proposition. *(a) The following statements are equivalent:*

(i) $\Gamma_0(N)$ *contains no conjugate of* σ.

(ii) The action of $\overline{\Gamma_0(N)}$ *on* $\mathbb{H}$ *does not have any stabiliser of even order.*

(iii) N *is divisible by a prime* q *which is* 3 *modulo* 4 *or by* 4.

(b) The following statements are equivalent:

(i) $\Gamma_0(N)$ *contains no conjugate of* τ.

(ii) The action of $\overline{\Gamma_0(N)}$ *on* $\mathbb{H}$ *does not have any stabiliser of order divisible by* 3.

(iii) N *is divisible by a prime* q *which is* 2 *modulo* 3 *or by* 9.

(c) If $N > 3$, *then* $\overline{\Gamma_1(N)}$ *acts freely on* $\mathbb{H}$.

Proof. Writing out (i) in the two cases as $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \sigma \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$ (resp. with τ) gives the equations $c^2 + d^2 = (c + id)(c - id) \equiv 0 \mod N$ resp. $c^2 + d^2 + cd = (c - \zeta_3 d)(c - \overline{\zeta_3} d) \equiv 0 \mod N$, with $(c, d) = 1$. Let l be a prime dividing N. It is clear that l cannot be inert in the extension $\mathbb{Q}(i)$ resp. $\mathbb{Q}(\zeta_3)$. If 4 divides N, then it follows that 2 divides $c + id$, which contradicts the fact that $(c, d) = 1$. Concluding similarly for 9 in case (b) establishes the implication (iii) $\Rightarrow$ (i) for (a) and (b). Conversely, we suppose that N is divisible only by split primes, i.e. $l_j = (c_j + id_j)(c_j - id_j)$ resp. $l_j = (c_j - \zeta_3 d_j)(c_j - \zeta_3 d_j)$, and possibly by $2 = (1 + i)(1 - i)$ resp. $3 = (1 - \zeta_3)(1 - \overline{\zeta_3})$. Multiplying out, it follows that N takes the form $c^2 + d^2$ resp. $c^2 + d^2 + cd$ with $(c, d) = 1$. Choosing $a, b \in \mathbb{Z}$ s.t. $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is in $\mathrm{SL}_2(\mathbb{Z})$ it follows that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \sigma \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$ (resp. with τ) is an element of $\overline{\Gamma_0(N)}$, establishing (i) $\Rightarrow$ (iii).

The equivalence of (i) and (ii) follows from the well known fact that the only non-trivial stabiliser groups of points in the usual fundamental domain are the groups generated by σ resp. τ.

(c) If a conjugate of σ (resp. τ) is in $\Gamma_1(N)$, one has the equations $ac + bd \equiv \epsilon 1 \mod N$ and $-(ac + bd) \equiv \epsilon 1 \mod N$ (resp. $ad + ac + bd \equiv \epsilon 1 \mod N$ and $-(bc + ac + bd) \equiv \epsilon 1 \mod N$) with $\epsilon = \pm 1$ (since we can replace σ (resp. τ) by $-\sigma$ (resp. $-\tau$)). This yields $\pm 2 \equiv 0 \mod N$ (resp. $2 \equiv \pm 1 \mod N$). □

2.7. Characters and the Δ-action

In this section we study the action of the group $\Delta := \Gamma \backslash G$ on various spaces in the Notation (2.0.3). That action is given by the diamond operators. We will be especially interested in how far this action is semi-simple, i.e., if modular symbols decompose into eigenspaces for characters.

Some computations in group (co-)homology

We first provide some results of group (co-)homology that will be used later.

(2.7.1) Proposition. *Let R be a ring, Δ a finite group, S a finite left Δ-set and V a left $R[\Delta]$-module. Suppose that for all $s \in S$ the stabiliser group Δ_s has order invertible in R. Then we have for all $i \geq 1$*

$$H_i(\Delta, R[S] \otimes_R V) = 0 = H^i(\Delta, R[S] \otimes_R V)$$

for the diagonal left Δ-action on $R[S] \otimes_R V$.

Proof. We prove this for homology. The proof for cohomology is obtained by dualising the arguments.

Choosing a system of representatives $s_1, \ldots, s_n$ of the Δ-orbits of S, we obtain a direct sum decomposition respecting the Δ-action

$$R[S] \otimes_R V \cong \bigoplus_{j=1}^{n} R[\Delta/\Delta_{s_j}] \otimes_R V.$$

From the projection formula we get

$$R[\Delta/\Delta_{s_j}] \otimes_R V \cong \mathrm{Ind}_{\Delta_{s_j}}^{\Delta} \mathrm{Res}_{\Delta_{s_j}}^{\Delta} V.$$

Shapiro's lemma now gives for all $i \geq 0$

$$H_i(\Delta, R[S] \otimes_R V) \cong \bigoplus_{j=1}^{n} H_i(\Delta_{s_j}, \mathrm{Res}_{\Delta_{s_j}}^{\Delta} V).$$

For $i \geq 1$, the right hand side, however, is zero, as multiplication by the group order of Δ_{s_j} is invertible in R. □

(2.7.2) Proposition. *Let A be a finite abelian group and K a field with trivial A-action.*

(a) If the characteristic of K is zero, then $H_i(A, K) = 0$ for all $i \geq 1$.

(b) If the characteristic of K is a prime p, then we have $\dim_K H_1(A, K) = n$ and $\dim_K H_2(A, K) = \frac{n(n+1)}{2}$, where n is the number of cyclic factors of the p-Sylow subgroup of A.

Proof. (a) is clear, as the group order is invertible in K. For (b) one can e.g. use that the dimensions of $H_1(A_p, K)$ resp. $H_2(A_p, K)$ are the minimal number of generators resp. relations of A_p. □

For a character we have the following more general statement on the first homology groups.

(2.7.3) Proposition. *Let K be a field, Δ a finite abelian group and $\epsilon : \Delta \to K^*$ a character. If K has characteristic $p > 0$, we also assume that ϵ is not the trivial character.*

Then we have for all $q \geq 0$ that $\dim_K H_q(\Delta, K^\epsilon) = \dim_K H^q(\Delta, K^\epsilon) = 0$.

Proof. We only prove this for cohomology. The statement on homology can be obtained by dualising the argument. The statement in characteristic 0 is clear, as the order of Δ is finite, so we assume that K has characteristic p.

For $\delta \in \Delta$ the endomorphism on $H^q(\Delta, K^\epsilon)$ which is induced from the action of δ on K^ϵ is the identity (it is well-defined, since Δ is abelian). Hence, for $\delta \in \Delta$ such that $\epsilon(\delta) \neq 1$ the non-zero element $\epsilon(\delta) - 1$ kills $H^q(\Delta, K^\epsilon)$, from which the claim follows. □

The Δ-action on the boundary space

(2.7.4) Proposition. *Let $N \geq 1$ be an integer which is invertible in R and assume Γ contains $\Gamma_1(N)$. Let $\overline{\Delta} := \overline{\Gamma}\backslash\overline{G}$. Then we have*

$$H_1\big(\overline{\Delta}, \mathcal{B}_k(\Gamma, R) \otimes_R R^\epsilon\big) = 0.$$

Proof. Let us write for short $M := \mathrm{Ind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}(V_{k-2}^\epsilon)$ and set $U := \langle T, -\mathbf{1}\rangle \leq \mathrm{SL}_2(\mathbb{Z})$. By Theorem (2.5.6) (b) we have

$$M_{\overline{U}} \cong M/M(1-T) \cong \mathcal{B}_k(\Gamma, R).$$

We first assume $(-1)^k = 1$, which is the case if $-\mathbf{1} \in \Gamma$, as then $1 = \epsilon(-\mathbf{1}) = (-1)^k$. Then by Proposition (2.5.7) (b) we have

$$M \cong R[\overline{\Gamma}\backslash\mathrm{PSL}_2(\mathbb{Z})] \otimes V_{k-2} \otimes R^\epsilon$$

with the actions described in that proposition. In particular, R^ϵ is a trivial right $\overline{U}$-module (by the restriction of the right $\mathrm{PSL}_2(\mathbb{Z})$-action). As T^N acts trivially on $R[\overline{\Gamma}\backslash\mathrm{PSL}_2(\mathbb{Z})]$, we obtain

$$M_{\langle T^N\rangle} \cong R[\overline{\Gamma}\backslash\mathrm{PSL}_2(\mathbb{Z})] \otimes (V_{k-2})_{\langle T^N\rangle} \otimes R^\epsilon.$$

The stabilisers of the $\overline{\Delta}$-action on the set $\overline{\Gamma}\backslash\mathrm{PSL}_2(\mathbb{Z})$ are trivial, whence by Proposition (2.7.1), we have $H_1(\overline{\Delta}, M_{\langle T^N\rangle}) = 0$. As the group $\overline{U}/\langle T^N\rangle$ has order N, which is invertible by assumption, $M_{\overline{U}}$ is a direct summand of $M_{\langle T^N\rangle}$, yielding the claim of the proposition in the case under consideration.

We assume now that $-\mathbf{1} \notin \Gamma$. Then $M = {}_{\langle -\mathbf{1}\rangle}N$ with $N := \mathrm{Ind}_{\Gamma}^{\mathrm{SL}_2(\mathbb{Z})}(V_{k-2}^\epsilon)$. We proceed as above. By Proposition (2.5.7) (a) we have

$$N \cong R[\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})] \otimes V_{k-2} \otimes R^\epsilon$$

with the actions described in that proposition. In particular, R^ϵ is a trivial right U-module (by the restriction of the right $\mathrm{SL}_2(\mathbb{Z})$-action). As T^N acts trivially on $R[\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})]$, we obtain

$$N_{\langle T^N\rangle} \cong R[\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})] \otimes (V_{k-2})_{\langle T^N\rangle} \otimes R^\epsilon.$$

The stabilisers of the Δ-action on the set $\Gamma\backslash \mathrm{SL}_2(\mathbb{Z})$ are trivial, whence by Proposition (2.7.1), we have $H_1(\Delta, N_{\langle T^N\rangle}) = 0$. By the Hochschild-Serre spectral sequence we get a surjection $H_1(\Delta, N_{\langle T^N\rangle}) \twoheadrightarrow H_1(\overline{\Delta}, {}_{\langle -\mathbf{1}\rangle}N_{\langle T^N\rangle})$, whence the right space is also zero. Hence, we have $H_1(\overline{\Delta}, M_{\langle T^N\rangle}) = 0$, and we can finish as above. □

Modular symbols with and without character

In this section we will specialise to the case of fields instead of general rings. However, an extension of the results to rings under natural restrictions is easily possible.

We start by comparing boundary and Eisenstein modular symbols.

(2.7.5) Proposition. *(a) We have the exact sequence*

$$0 \to \mathcal{E}_2(\Gamma, R) \to \mathcal{B}_2(\Gamma, R) \to R \to 0.$$

(b) Let $N \geq 1$ be an integer such that $\Gamma_1(N) \leq \Gamma$ and let K be a field. If the characteristic of K is $p > 0$, then we assume $p \nmid N$ and $k \leq p + 2$. If $k = 2$, then we suppose that ϵ is not the trivial character.

Then we have

$$\mathcal{E}_k(G, \epsilon, K) \cong \mathcal{B}_k(G, \epsilon, K).$$

Proof. For any ring R we have the exact sequence

$$\mathcal{M}_2(R) \xrightarrow{\{\alpha,\beta\}\mapsto\{\beta\}-\{\alpha\}} \mathcal{B}_2(R) \xrightarrow{\{\alpha\}\mapsto 1} R \to 0.$$

We only need to show that $\sum_{\{\alpha\}} r_\alpha\{\alpha\}$ is in the image of the boundary map, if $\sum_{\{\alpha\}} r_\alpha = 0$. But then $\sum_{\{\alpha\}} r_\alpha\{\alpha\} = \sum_{\{\alpha\}} r_\alpha(\{\alpha\} - \{\infty\})$, which clearly lies in the image. Taking $\overline{\Gamma}$-coinvariants, we obtain part (a), as R is a trivial $\overline{\Gamma}$-module.

Let us now assume the situation described in (b). From the exact sequence above for $R = K$, we immediately obtain the following exact sequence by tensoring with $V^\epsilon_{k-2}(K)$

$$\mathcal{M}_2(K) \otimes_K V^\epsilon_{k-2}(K) \xrightarrow{\{\alpha,\beta\}\mapsto\{\beta\}-\{\alpha\}} \mathcal{B}_2(K) \otimes_K V^\epsilon_{k-2}(K) \xrightarrow{\{\alpha\}\mapsto 1} V^\epsilon_{k-2}(K) \to 0.$$

Propositions (2.2.6) and (2.2.7) now finish the proof. □

We now compute the difference of the Eisenstein spaces.

(2.7.6) Lemma. *Under the assumptions of Proposition (2.7.5)(b) we have the exact sequence*

$$0 \to H_1(\overline{\Delta}, {}_{\overline{\Gamma}}V^\epsilon_{k-2}(K)) \to {}_{\overline{\Delta}}(\mathcal{E}_k(\Gamma, K) \otimes_K K^\epsilon) \to \mathcal{E}_k(G, \epsilon, K) \to 0$$

with $\overline{\Delta} := \overline{\Gamma}\backslash\overline{G}$.

Proof. We start with the exact sequence

$$0 \to \mathcal{E}_k(\Gamma, K) \otimes_K K^\epsilon \to (\mathcal{B}_k(\Gamma, K)) \otimes_K K^\epsilon \to ({}_{\overline{\Gamma}}V_{k-2}(K)) \otimes_K K^\epsilon \to 0,$$

which gives rise to the long exact sequence

$$H_1(\overline{\Delta}, \mathcal{B}_k(\Gamma, K) \otimes_K K^\epsilon) \to H_1(\overline{\Delta}, {}_{\overline{\Gamma}}V^\epsilon_{k-2}(K)) \to$$
$${}_{\overline{\Delta}}(\mathcal{E}_k(\Gamma, K) \otimes_K K^\epsilon) \to \mathcal{B}_k(G, \epsilon, K) \xrightarrow{\phi} {}_{\overline{\Delta}}({}_{\overline{\Gamma}}V^\epsilon_{k-2}(K)) \to 0.$$

In Proposition (2.7.4) we proved that the first term is zero. The kernel of ϕ equals $\mathcal{E}_k(G, \epsilon, K)$ by exactness, which proves the lemma. □

(2.7.7) Proposition. *(a) We have the exact sequence*

$$0 \to H_1(\overline{\Delta}, K) \to {}_{\overline{\Delta}}\mathcal{E}_2(\Gamma, K) \to \mathcal{E}_2(G, K) \to 0.$$

(b) Under the assumptions of Proposition (2.7.5)(b) we have

$${}_{\overline{\Delta}}\big(\mathcal{E}_2(\Gamma, K) \otimes_K K^\epsilon\big) \cong \mathcal{E}_k(G, \epsilon, K).$$

Proof. The Proposition follows directly from Lemma (2.7.6) and Propositions (2.2.6), (2.2.7) and (2.7.3). □

Next we compare the spaces of cuspidal modular symbols.

(2.7.8) Theorem. *(a) We have the exact sequence*

$$H_1(\overline{\Delta}, \mathcal{E}_2(\Gamma, K)) \to {}_{\overline{\Delta}}\mathcal{CM}_2(\Gamma, K) \to \mathcal{CM}_2(G, K) \to H_1(\overline{\Delta}, K) \to 0$$

and $H_1(\overline{\Delta}, K) \hookrightarrow H_0(\overline{\Delta}, \mathcal{E}_2(\Gamma_1, K))$ *and* $H_2(\overline{\Delta}, K) \twoheadrightarrow H_1(\overline{\Delta}, \mathcal{E}_2(\Gamma_1, K))$.

(b) Let $N \geq 1$ be an integer such that $\Gamma_1(N) \leq \Gamma$ and let K be a field. If the characteristic of K is $p > 0$, then we assume $p \nmid N$ and $k \leq p + 2$. If $k = 2$, then we suppose that ϵ is not the trivial character. Then we have

$${}_{\overline{\Delta}}(\mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon) \cong \mathcal{CM}_k(G, \epsilon, K).$$

Proof. We compare the long exact sequence associated to the short exact sequence of $\overline{\Delta}$-modules

$$0 \to \mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon \to \mathcal{M}_k(\Gamma, K) \otimes_K K^\epsilon \to \mathcal{E}_k(\Gamma, K) \otimes_K K^\epsilon \to 0$$

with the short exact sequence

$$0 \to \mathcal{CM}_k(G, \epsilon, K) \to \mathcal{M}_k(G, \epsilon, K) \to \mathcal{E}_k(G, \epsilon, K) \to 0.$$

Using the snake lemma and Lemma (2.7.6) we obtain the exact sequence

$$H_1(\overline{\Delta}, \mathcal{E}_k(\Gamma, K) \otimes_K K^\epsilon) \to {}_{\overline{\Delta}}(\mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon) \to \\ \mathcal{CM}_k(G, \epsilon, K) \to H_1(\overline{\Delta}, {}_{\overline{\Gamma}}V^\epsilon_{k-2}(K)) \to 0,$$

from which all statements follow, except the second one of (a), via Propositions (2.7.4) and (2.7.5). In order to finish part (a), we show

$$H_2(\Delta, K) \twoheadrightarrow H_1(\overline{\Delta}, \mathcal{E}_k(\Gamma, K)) \quad \text{and} \quad H_1(\overline{\Delta}, K) \hookrightarrow H_0(\overline{\Delta}, \mathcal{E}_2(\Gamma_1, K)).$$

Both follow from the long exact sequence associated to the short exact sequence from Proposition (2.7.5) (a) and the fact that $H_1(\overline{\Delta}, \mathcal{B}_2(\Gamma, K)) = 0$, as provided by Proposition (2.7.4). □

The obstruction terms occurring in Theorem (2.7.8) have been calculated in Propositions (2.7.2).

The Δ-action on modular symbols

We first need a technical computation on induced and coinduced modules.

(2.7.9) Lemma. *Let R be a ring, $\Gamma \lhd G$ be subgroups of finite index in a group $\mathcal{S}$, let $\Delta := G/\Gamma$ and let V be a right $R[G]$-module. Then the diagram of right $R[\mathcal{S}]$-modules*

$$\begin{array}{ccccc} \mathrm{Ind}_G^{\mathcal{S}} V & \xrightarrow{\sim} & \mathrm{Coind}_G^{\mathcal{S}} V & \longrightarrow & \mathrm{Coind}_\Gamma^{\mathcal{S}} V \\ \uparrow & & & & \downarrow \int \\ \mathrm{Ind}_\Gamma^{\mathcal{S}} V & & \xrightarrow{N_\Delta \cdot} & & \mathrm{Ind}_\Gamma^{\mathcal{S}} V \end{array}$$

commutes, where the norm is taken for the natural left Δ-action.

Proof. For convenience we have exchanged right and left actions in the proof, which can easily be undone by inverting. We consider $\mathrm{Ind}_\Gamma^{\mathcal{S}} V = R[\mathcal{S}] \otimes_{R[\Gamma]} V$ with the left $R[\mathcal{S}]$-action on the left factor and the right $R[\Delta]$-action $(\sigma \otimes_\Gamma v)\delta = \sigma\delta \otimes_\Gamma \delta^{-1} v$. This action is compatible with the right $R[\Gamma]$-action on $R[\mathcal{S}]$ and the given left $R[\Gamma]$-action on V for which the tensor product has been taken. We regard $\mathrm{Coind}_\Gamma^{\mathcal{S}} V = \mathrm{Hom}_\Gamma(R[\mathcal{S}], V)$ with the left $R[\mathcal{S}]$-action $(g.f)(\sigma) = f(g^{-1}\sigma)$ and the right $R[\Delta]$-action $(f.\delta)(\sigma) = \delta^{-1} f(g\delta^{-1})$. This last action is the restriction of the G-action defining $\mathrm{Hom}_G(R[\mathcal{S}], V) = \big(\mathrm{Hom}_R(R[\mathcal{S}], V)\big)^G$.

Now we can check commutativity. We first go up, then right and then down, and verify in the end that we obtain N_Δ in this way. We choose a system of representatives $g_1, \ldots, g_n$ for the residue classes $\mathcal{S}/G$. Then the $g_i\delta$ are a system of representatives for the residue classes of $\mathcal{S}/\Gamma$ when $i = 1, \ldots, n$ and $\delta \in \Delta$. So, let $x = \sum_\delta \sum_i g_i\delta \otimes_\Gamma v_{\delta,i}$ be an element of $\mathrm{Ind}_G^{\mathcal{S}} V$. It is first mapped to $\sum_\delta \sum_i g_i\delta \otimes_G v_{\delta,i} = \sum_i g_i \otimes_G (\sum_\delta \delta v_{\delta,i})$. Its image in the centre of the top row is the map f which is uniquely defined by sending g_i to $\sum_\delta \delta v_{\delta,i}$. We

have that $f(g_i\tilde{\delta}) = \tilde{\delta}^{-1}f(g_i)$. Hence, the image of x in the right upper corner is the map which is uniquely given by sending $g_i\tilde{\delta}$ to $\tilde{\delta}^{-1}\sum_\delta \delta v_{\delta,i}$. Mapping this element down to the right bottom corner gives

$$\sum_{\tilde{\delta}}\sum_{\delta}\sum_{i} g_i\tilde{\delta}\otimes_\Gamma \tilde{\delta}^{-1}\delta v_{\delta,i}.$$

This element, however, agrees with

$$x.N_\Delta = \sum_{\hat{\delta}}\sum_{\delta}\sum_{i} g_i\delta\hat{\delta}\otimes_\Gamma \hat{\delta}^{-1}v_{\delta,i},$$

for $\tilde{\delta} = \delta\hat{\delta}$, as claimed. □

(2.7.10) Proposition. *Let $N \geq 1$ be an integer such that $\Gamma_1(N) \leq \Gamma$ and let K be a field. If $k = 2$, then we suppose that ϵ is not the trivial character. If the characteristic of K is $p > 0$, then we assume also $p \nmid N$, $k \leq p+2$ and that all stabiliser subgroups of $\overline{G}$ for its action on $\mathbb{H}$ have order invertible in K (cf. Theorem (2.6.2)).*

Then the norm map $N_{\overline{\Delta}}$ induces isomorphisms

$$\mathcal{M}_k(G,\epsilon,K) = {}_{\overline{\Delta}}(\mathcal{M}_k(\Gamma,K)\otimes_K K^\epsilon) \cong {}^{\overline{\Delta}}(\mathcal{M}_k(\Gamma,K)\otimes_K K^\epsilon),$$

$$\mathcal{B}_k(G,\epsilon,K) = {}_{\overline{\Delta}}(\mathcal{B}_k(\Gamma,K)\otimes_K K^\epsilon) \cong {}^{\overline{\Delta}}(\mathcal{B}_k(\Gamma,K)\otimes_K K^\epsilon)$$

and

$$\mathcal{CM}_k(G,\epsilon,K) = {}_{\overline{\Delta}}(\mathcal{CM}_k(\Gamma,K)\otimes_K K^\epsilon) \cong {}^{\overline{\Delta}}(\mathcal{CM}_k(\Gamma,K)\otimes_K K^\epsilon).$$

Proof. With $V = V_{k-2}^\epsilon(K)$ and $\mathcal{S} = \mathrm{PSL}_2(\mathbb{Z})$ Lemma (2.7.9) gives the commutative diagram of $K[\mathrm{PSL}_2(\mathbb{Z})]$-modules

$$\begin{array}{ccccc}
(\mathrm{Ind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}V)/\mathfrak{a} & \longrightarrow & (\mathrm{Coind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})}V)/\mathfrak{a} & \longrightarrow & {}^{\overline{\Delta}}\big((\mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}V)/\mathfrak{a}\big) \\
\uparrow & & & & \downarrow \\
{}_{\overline{\Delta}}\big((\mathrm{Ind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}V)/\mathfrak{a}\big) & & \xrightarrow{\cdot N_{\overline{\Delta}}} & & {}^{\overline{\Delta}}\big((\mathrm{Ind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})}V)/\mathfrak{a}\big)
\end{array}$$

with $\mathfrak{a} = (1+\sigma, 1+\tau+\tau^2) \lhd K[\mathrm{PSL}_2(\mathbb{Z})]$. Due to the assumptions we may combine the comparison result Theorem (2.6.1) with the description in terms of Manin symbols (Theorem (2.5.6)). This allows us to reinterpret the diagram as

$$\begin{array}{ccccc}
\mathcal{M}_k(G,\epsilon,K) & \xrightarrow{\sim} & H^1(\overline{G}, V_{k-2}^\epsilon(K)) & \xrightarrow{\mathrm{res}} & {}^{\overline{\Delta}}\big(H^1(\overline{\Gamma}, V_{k-2}^\epsilon(K))\big) \\
\wr\uparrow & & & & \downarrow\wr \\
{}_{\overline{\Delta}}\big(\mathcal{M}_k(\Gamma,K)\otimes_K K^\epsilon\big) & & \xrightarrow{\cdot N_{\overline{\Delta}}} & & {}^{\overline{\Delta}}\big(\mathcal{M}_k(\Gamma,K)\otimes_K K^\epsilon\big).
\end{array}$$

In the diagram the left vertical arrow is the definition, the upper left horizontal and the right vertical arrow come from the comparison and res is the restriction from group cohomology, which features in the exact sequence

$$0 \to H^1(\overline{\Delta}, V^{\overline{\Gamma}}) \to H^1(\overline{G}, V) \xrightarrow{\mathrm{res}} {}^{\overline{\Delta}}H^1(\overline{\Gamma}, V) \to H^2(\overline{\Delta}, V^{\overline{\Gamma}})$$

coming from the Hochschild-Serre spectral sequence. The first part of the proposition follows from Propositions (2.2.6) and (2.2.7) for $k > 2$ and from the assumptions for $k = 2$, which imply that the map res above is an isomorphism.

The result on the cuspidal subspace will follow from the result on the boundary space. For that we proceed as above with $\mathfrak{a} = (1-T) \lhd K[\mathrm{PSL}_{\mathbb{Z}}]$. This reduces us to show that the map

$$H^1(U, \mathrm{Coind}_{\overline{G}}^{\mathrm{PSL}_2(\mathbb{Z})} V) \to {}^{\overline{\Delta}}H^1(U, \mathrm{Coind}_{\overline{\Gamma}}^{\mathrm{PSL}_2(\mathbb{Z})} V)$$

coming from the restriction via Shapiro's lemma is an isomorphism. We claim that the restriction map

$$H^1(\overline{G} \cap g\overline{U}g^{-1}, V) \xrightarrow{\mathrm{res}} H^1(\overline{\Gamma} \cap g\overline{U}g^{-1}, V)^{\overline{G} \cap g\overline{U}g^{-1}}$$

is an isomorphism for all $g \in \mathrm{PSL}_2(\mathbb{Z})$. An easy calculation shows that $g\langle T^N \rangle g^{-1}$ is an element of $\Gamma_1(N)$ and is hence in $\overline{\Gamma}$. Consequently we have the inclusions

$$g\langle T^N \rangle g^{-1} \subseteq \overline{\Gamma} \cap g\overline{U}g^{-1} \subseteq \overline{G} \cap g\overline{U}g^{-1} \subseteq g\overline{U}g^{-1}.$$

As the total index is N, the index of $\overline{\Gamma} \cap g\overline{U}g^{-1}$ in $\overline{G} \cap g\overline{U}g^{-1}$ divides N and is thus coprime with p. Using again the five term sequence associated to the Hochschild-Serre spectral sequence immediately implies that the restriction map above is an isomorphism as claimed.

Given a fixed $g \in \overline{G}\backslash \mathrm{PSL}_2(\mathbb{Z})/\overline{U}$ we are reduced to consider the diagonal restriction

$$H^1(\overline{G} \cap g\overline{U}g^{-1}, V) \xrightarrow{\mathrm{res}} \bigoplus_h H^1(\overline{\Gamma} \cap h\overline{U}h^{-1}, V)^{\overline{G} \cap g\overline{U}g^{-1}},$$

where h runs through a system of representatives of $\overline{\Gamma}h\overline{U}$ such that $\overline{G}h\overline{U} = \overline{G}g\overline{U}$. The group $\overline{\Delta}$ permutes this set and only the diagonal is invariant. □

Let us point out that the $\overline{\Delta}$-action on the set $\overline{\Gamma}\backslash \mathrm{PSL}_2(\mathbb{Z})/\overline{U}$ is not free if N is not square-free.

(2.7.11) Corollary. *Under the assumptions of Proposition (2.7.10) all of the following Tate cohomology groups are zero* $\widehat{H}^0(\overline{\Delta}, \mathcal{M}_k(\Gamma, K) \otimes_K K^\epsilon)$, $\widehat{H}_0(\overline{\Delta}, \mathcal{M}_k(\Gamma, K) \otimes_K K^\epsilon)$, $\widehat{H}^0(\overline{\Delta}, \mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon)$, $\widehat{H}_0(\overline{\Delta}, \mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon)$, $\widehat{H}^0(\overline{\Delta}, \mathcal{B}_k(\Gamma, K) \otimes_K K^\epsilon)$ *and* $\widehat{H}_0(\overline{\Delta}, \mathcal{B}_k(\Gamma, K) \otimes_K K^\epsilon)$.

Proof. This is immediate from the definition of the Tate cohomology groups, which can be summarised in the exact sequence $0 \to \widehat{H}_0 \to H_0 \xrightarrow{\mathrm{Norm}} H^0 \to \widehat{H}^0 \to 0$, and Proposition (2.7.10). □

Separating the p-Sylow action

We let $\overline{\Delta} := \overline{\Gamma}\backslash\overline{G}$ and assume that it is an abelian group. Here we are interested in modular symbols as a $\overline{\Delta}$-module. We will treat the case of p-primary and p-group action separately,

when p is the characteristic of the coefficient field $R = K$. This is what we need the freedom in choosing the groups Γ and G different from only $\Gamma_1(N)$ and $\Gamma_0(N)$ for.

By Sylow theory there is a group Γ_p such that

$$\Gamma \lhd \Gamma_p \lhd G$$

with $\Gamma\backslash\Gamma_p$ a p-group and $\Gamma_p\backslash G$ of order prime to p. The restriction of ϵ to $\Gamma\backslash\Gamma_p$ is necessarily trivial. We define the character

$$\tilde{\epsilon} : G \to \Gamma_p\backslash G \subseteq \Gamma\backslash G \to K^*$$

using $\Gamma\backslash G \cong \Gamma_p\backslash G \times \Gamma\backslash\Gamma_p$. We clearly have $\tilde{\epsilon}(-\mathbf{1}) = \epsilon(-\mathbf{1}) = (-1)^k$.

(2.7.12) Remark. *Let us point out that the condition on the characters only stems from the fact that we want to work with* $\mathrm{PSL}_2(\mathbb{Z})$ *instead of* $\mathrm{SL}_2(\mathbb{Z})$. *This choice unfortunately prevents us from repeating the above arguments with a subgroup* $\tilde{\Gamma}$ *such that* $\Gamma \lhd \tilde{\Gamma} \lhd G$ *with* $\Gamma\backslash\tilde{\Gamma}$ *of order prime to* p *and* $\tilde{\Gamma}\backslash G$ *a* p*-group. In that case the necessarily trivial character* $G \overset{\mathrm{proj}}{\twoheadrightarrow} \tilde{\Gamma}\backslash G \to K^*$ *would not be allowed if* $(-1)^k \neq 1$.

The action of the group $\Gamma_p\backslash G$ is semi-simple, and any module splits into a direct sum of character eigenspaces, if the ground field contains the character values. The behaviour is thus as in characteristic zero.

There is quite a strong criterion to show that a module for a p-group is coinduced.

(2.7.13) Proposition. *Let* K *be a finite field of characteristic* p, *let* Δ_p *be a finite* p*-group and let* A *be a* $K[\Delta_p]$*-module. If* $\hat{H}^n(\Delta_p, A) = 0$ *for one* n, *then* A *is a coinduced* $K[\Delta_p]$*-module.*

Proof. This is [NSW], Proposition 1.7.3 (ii). □

(2.7.14) Corollary. *Let* $k \geq 3$, $N \geq 1$ *be integers and* K *a finite field of characteristic* p. *We assume* $p \nmid N$ *and* $k \leq p + 2$. *Furthermore, let* $\Gamma_1(N) \leq \Gamma \lhd \Gamma_p \leq \mathrm{SL}_2(\mathbb{Z})$ *be subgroups such that* $\Gamma\backslash\Gamma_p$ *is a* p*-group. We furthermore suppose that* $\overline{G}$ *has no stabilisers of order* p *for its action on* $\mathbb{H}$. *Let* $\Delta_p := \Gamma\backslash\Gamma_p$. *Then* $\mathcal{M}_k(\Gamma, K)$, $\mathcal{CM}_k(\Gamma, K)$ *and* $\mathcal{B}_k(\Gamma, K)$ *are coinduced* $K[\Delta_p]$*-modules.*

Proof. This follows directly from Corollary (2.7.11) and Proposition (2.7.13). □

Chapter III

Hecke Algebras of mod p Modular Forms and Modular Symbols

In this chapter we prove that under certain conditions the Hecke algebra of cuspidal modular forms over $\overline{\mathbb{F}_p}$ can be obtained by considering only group cohomology, generalising results from [EdixJussieu]. When these conditions apply, one obtains much more information than e.g. [Ash-Stevens], who have studied group cohomology in order to prove that all systems of eigenvalues of modular forms mod p in level N for $p \nmid N$ and weight $k \geq 3$ occur in the group cohomology of level Np and weight 2.

We start this chapter by introducing Hecke operators on the group cohomology groups considered in Chapter II. Moreover, the compatibility of the Hecke operators with Shapiro's lemma is studied.

The principal idea in this chapter is to relate modular forms and modular symbols of level N with $p \nmid N$ and weight $2 \leq k \leq p+1$ to level Np and weight 2. In the second section we will develop this level raising for the cohomology groups.

The third section is concerned with Hecke algebras of modular symbols and a comparison to Hecke algebras of modular forms. The Eichler-Shimura-Theorem for holomorphic modular forms will be recalled first. Next results of p-adic Hodge theory will be used to exhibit a faithful module for the Hecke algebra of cusp forms over $\overline{\mathbb{F}_p}$, when the weight is between 2 and $p-1$. As modular forms of weight 1 can be embedded into weight p, it is desirable to extend the weight range. This, however, does not seem to be possible with p-adic Hodge theory. In order to cover weights up to $p+1$, we relate them to weight 2 and higher level, so that the Jacobian of the modular curve can be used. This method allows us to prove that locally at ordinary primes of the Hecke algebra a faithful module is provided by group cohomology with coefficients in $\mathbb{F}_p$ (see Corollary (3.3.14)). We end the chapter by a discussion of the action of $\Gamma_0(N)/\Gamma_1(N)$ on cusp forms, which allows us to extend our results to modular forms with characters.

In Chapter II we have discussed modular symbols and related spaces over quite general rings. In this chapter we will mostly take a finite field of characteristic p as base field and principally work with the group cohomological description.

3.1. Hecke action

The definitions in this section are based on [DiamondIm].

We directly define Hecke operators on group cohomology. Although we do not expose the theory here, we should not fail to mention that Hecke operators conceptually come from correspondences on the underlying modular curves that also have a very explicit description in the moduli interpretation.

By the comparison result Theorem (2.6.1) the definition can be transferred to modular symbols in the case of the group $\Gamma_1(N)$ for $N \geq 5$. Taking coinvariants one can extend the definition of Hecke operators also to spaces for $\Gamma_0(N)$ with a character.

Hecke operators on group cohomology

Let $\alpha \in \mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$ and Γ a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We use the notations $\Gamma_\alpha := \Gamma \cap \alpha^{-1}\Gamma\alpha$ and $\Gamma^\alpha := \Gamma \cap \alpha\Gamma\alpha^{-1}$, where we consider α^{-1} as an element of $\mathrm{GL}_2(\mathbb{Q})$. Both groups are commensurable with Γ.

Suppose that V is an R-module with a $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-(semi-group)-action. We define the *Hecke operator* T_α acting on group cohomology as the composite

$$H^1(\Gamma, V) \xrightarrow{\mathrm{res}} H^1(\Gamma^\alpha, V) \xrightarrow{\mathrm{conj}_\alpha} H^1(\Gamma_\alpha, V) \xrightarrow{\mathrm{cores}} H^1(\Gamma, V).$$

The first map is the usual *restriction*, and the third is the so-called *corestriction*, which one also finds in the literature under the name *transfer* (cf. [Weibel], [Brown]). We explicitly describe the second map on cochains (cf. [DiamondIm], p. 116):

$$\mathrm{conj}_\alpha : H^1(\Gamma^\alpha, V) \to H^1(\Gamma_\alpha, V), \quad c \mapsto \big(g_\alpha \mapsto \alpha^\iota . c(\alpha g_\alpha \alpha^{-1})\big).$$

The following formula can also be found in [DiamondIm], p. 116, and [Shimura], Section 8.3.

(3.1.1) Proposition. *Suppose that* $\Gamma\alpha\Gamma = \bigcup_{i=1}^n \Gamma\delta_i$ *is a disjoint union. Then the Hecke operator* T_α *acts on* $H^1(\Gamma, V)$ *by sending the non-homogeneous cocyle* c *to* $T_\alpha c$ *defined by*

$$(T_\alpha c)(g) = \sum_{i=1}^n \delta_i^\iota c(\delta_i g \delta_{j(i)}^{-1})$$

for $g \in \Gamma$. *Here* $j(i)$ *is the index such that* $\delta_i g \delta_{j(i)}^{-1} \in \Gamma$.

Proof. We only have to describe the corestriction explicitly. For that we notice that one has $\Gamma = \bigcup_{i=1}^{n} \Gamma_\alpha g_i$ with $\alpha g_i = \delta_i$. Furthermore the corestriction of a non-homogeneous cocycle $u \in H^1(\Gamma_\alpha, V)$ is the cocycle $\mathrm{cores}(u)$ uniquely given by

$$\mathrm{cores}(u)(g) = \sum_{i=1}^{n} g_i^{-1} u(g_i g g_{j(i)}^{-1})$$

for $g \in \Gamma$. Combining with the explicit description of the map conj_α yields the result. □

For a positive integer n, one defines the *Hecke operator* T_n to be T_α for $\alpha = \left(\begin{smallmatrix} 1 & 0 \\ 0 & n \end{smallmatrix}\right)$.

If $\Gamma_1(N) \subseteq \Gamma$ and the integer d is coprime to N, one defines the *diamond operator* $\langle d \rangle$ to be T_α for any matrix $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, whose reduction modulo N is $\left(\begin{smallmatrix} d^{-1} & 0 \\ 0 & d \end{smallmatrix}\right)$. The diamond operator gives a group action by $(\mathbb{Z}/N\mathbb{Z})^*$. If the level is NM with $(N, M) = 1$, then we can separate the diamond operator into two parts $\langle d \rangle = \langle d \rangle_M \times \langle d \rangle_N$, corresponding to $\mathbb{Z}/NM\mathbb{Z} \cong \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$.

Hecke operators and Shapiro's lemma

(3.1.2) Lemma. *Let N, M be coprime positive integers, and let V be an $R[\Gamma_1(N)]$-module. Define the R-module*

$$\mathcal{W}(M, V) := \{f \in \mathrm{Hom}_R(R[(\mathbb{Z}/M\mathbb{Z})^2], V) \mid f((u, v)) = 0 \ \forall (u, v) \text{ s.t. } \langle u, v \rangle \neq \mathbb{Z}/M\mathbb{Z}\}.$$

We equip it with the left $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-(semi-group)-action $(g.f)((u, v)) = gf((u, v)g)$.

Then the homomorphism

$$\mathcal{W}(M, V) \to \mathrm{Hom}_{R[\Gamma_1(NM)]}(R[\Gamma_1(N)], V), \quad f \mapsto \big(g \mapsto (g.f)((0, 1))\big)$$

is an isomorphism of left $\Gamma_1(N)$-modules (by restricting the action on $\mathcal{W}(M, V)$). In particular, $\mathcal{W}(M, V)$ is isomorphic to $\mathrm{Coind}_{\Gamma_1(NM)}^{\Gamma_1(N)}(V)$ as a left $\Gamma_1(N)$-module.

Proof. As N and M are coprime, reduction modulo M defines a surjection from $\Gamma_1(N)$ onto $\mathrm{SL}_2(\mathbb{Z}/M\mathbb{Z})$. This implies that the map

$$\Gamma_1(NM)\backslash\Gamma_1(N) \xrightarrow{A \mapsto (0,1)A \bmod M} (\mathbb{Z}/M\mathbb{Z})^2$$

is injective, and its image is the set of the (u, v) with $\mathbb{Z}/M\mathbb{Z} = \langle u, v \rangle$. From this the claimed isomorphism follows directly. □

(3.1.3) Lemma. *Let N be a positive integer and l a prime. We have the coset decomposition*

$$\Gamma_1(N) \left(\begin{smallmatrix} 1 & 0 \\ 0 & l \end{smallmatrix}\right) \Gamma_1(N) = \bigcup_a \bigcup_b \Gamma_1(N) \sigma_a \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$$

when a runs through the integers such that $a > 0$, $(a, N) = 1$, $ad = l$ and b through a system of representatives of $\mathbb{Z}/d\mathbb{Z}$. Here $\sigma_a \in \mathrm{SL}_2(\mathbb{Z})$ is a matrix reducing to $\left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & a \end{smallmatrix}\right)$ modulo N.

Proof. This is [Shimura], Proposition 3.36. □

We can now prove the compatibility of the Hecke operators with the isomorphism from Shapiro's lemma when we take the $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-action on the coinduced module from Lemma (3.1.2). A proof of this fact in the more general, but rather heavy language of weakly compatible Hecke pairs can be found in [Ash-Stevens] (Lemma 2.2(b)).

The Shapiro map is the isomorphism on cohomology groups

$$\mathrm{Sh} : H^1(\Gamma_1(N), \mathcal{W}(M,V)) \to H^1(\Gamma_1(NM), V)$$

induced by the homomorphism

$$\mathcal{W}(M,V) \to V, \quad f \mapsto f((0,1)).$$

(3.1.4) Proposition. *Let N, M be coprime positive integers, and let V be an $R[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-module. For all primes l and all integers $d \geq 1$ with $(d,N) = 1$ we have*

$$T_l \circ \mathrm{Sh} = \mathrm{Sh} \circ T_l \quad \textit{and} \quad \langle d \rangle_N \circ \mathrm{Sh} = \mathrm{Sh} \circ \langle d \rangle_N.$$

Proof. First we prove the statement for T_l. We choose a matrix σ_a for $(a,N) = 1$ such that it reduces to $\left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & a \end{smallmatrix}\right)$ modulo N. If $(a,M) = 1$, then we also impose that σ_a reduces to $\left(\begin{smallmatrix} a^{-1} & 0 \\ 0 & a \end{smallmatrix}\right)$ modulo M. If not, then we want $\sigma_a \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ modulo M. Lemma (3.1.3) implies that coset representatives of $\Gamma_1(NM)\backslash\Gamma_1(NM)\left(\begin{smallmatrix} 1 & 0 \\ 0 & l \end{smallmatrix}\right)\Gamma_1(NM)$ can be chosen as a subset of representatives of $\Gamma_1(N)\backslash\Gamma_1(N)\left(\begin{smallmatrix} 1 & 0 \\ 0 & l \end{smallmatrix}\right)\Gamma_1(N)$. With the above choice of σ_a that is the subset such that $\langle u, v\rangle = \mathbb{Z}/M\mathbb{Z}$ with $\left(\begin{smallmatrix} u & * \\ v & * \end{smallmatrix}\right) = \sigma_a \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$. For those we have by definition for $f \in \mathcal{W}(M,V)$ that $(\left(\begin{smallmatrix} u & * \\ v & * \end{smallmatrix}\right)^\iota f)((0,1)) = 0$.

Let now $c \in H^1(\Gamma_1(N), \mathcal{W}(M,V))$ be a cocycle. Then by Proposition (3.1.1) and the definition of the $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-action on $\mathcal{W}(M,V)$ we have for $g \in \Gamma_1(NM)$

$$(\mathrm{Sh}(T_n c))(g) = \sum_\delta \delta^\iota (c(\delta g \widetilde{\delta}^{-1})((0,1)\delta^\iota)),$$

where the sum runs over the above coset representatives for $\Gamma_1(N)$ and $\widetilde{\delta}$ is chosen among these representatives such that $\delta g \widetilde{\delta}^{-1} \in \Gamma_1(NM)$. Moreover, we have

$$(T_n(\mathrm{Sh}(c)))(g) = \sum_\delta \delta^\iota (c(\delta g \widetilde{\delta}^{-1})((0,1))),$$

where now the sum only runs through the subset described above. By what we have remarked right above $(0,1)\delta^\iota$ is $(0,1)$ if and only if $(a,M) = 1$. In all other cases $(0,1)\delta^\iota = (u,v)$ with $\langle u, v\rangle \neq \mathbb{Z}/M\mathbb{Z}$. This proves the compatibility for T_l.

The same arguments as above also show the compatibility of the diamond operator, except that we only have one coset representative. □

(3.1.5) Proposition. *Let N, M be coprime positive integers, and let V be an $R[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-module. For $(n, M) = 1$ we define the $R[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-isomorphism*

$$\mathrm{mult}_n : \mathcal{W}(M, V) \to \mathcal{W}(M, V), \quad f \mapsto ((u, v) \mapsto f\big((nu, nv)\big)).$$

Then we have

$$\langle n \rangle_M \circ \mathrm{Sh} = \mathrm{Sh} \circ \mathrm{mult}_n.$$

Proof. Let $\sigma \in \mathrm{SL}_2(\mathbb{Z})$ be a matrix reducing to $\left(\begin{smallmatrix} n^{-1} & 0 \\ 0 & n \end{smallmatrix}\right)$ modulo M and to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ modulo N. This means in particular that $\sigma \in \Gamma_1(N)$. Hence, for a cocycle $c \in H^1(\Gamma_1(N), V)$ we have

$$\sigma^{-1} c(\sigma g \sigma^{-1}) = c(g) + (g - 1)c(\sigma^{-1}),$$

so that the equality $c(\sigma g \sigma^{-1}) = \sigma c(g)$ holds in $H^1(\Gamma_1(N), V)$.

We can now check the claim. First we have

$$(\langle n \rangle_p \circ \mathrm{Sh})(c)(g) = \sigma^\iota(\sigma.c(g)((0, 1))) = c(g)((0, 1)\sigma).$$

This agrees with $(\mathrm{Sh} \circ \mathrm{mult}_n)(c)(g) = c(g)((0, n))$. □

3.2. Level raising for parabolic group cohomology

The contents of this section is already partly present in [Ash-Stevens]. However, in that paper the parabolic subspace is not treated.

Decomposition of $\mathcal{W}(p, \mathbb{F}_p)$ as $\mathbb{F}_p[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-module

We will now relate the $\mathbb{F}_p[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-modules $\mathcal{W}(p, \mathbb{F}_p)$ and $V_d(\mathbb{F}_p)$ for $0 \leq d \leq p - 1$, which are in fact precisely the simple $\mathbb{F}_p[\mathrm{SL}_2(\mathbb{F}_p)]$-modules (see e.g. [Alperin], p. 15).

(3.2.1) Lemma. *Evaluation of polynomials on $\mathbb{F}_p^2$ induces the natural isomorphism of left $\mathbb{F}_p[\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}]$-modules*

$$\mathbb{F}_p[X, Y]/(X^p - X, Y^p - Y) \cong \mathbb{F}_p^{\mathbb{F}_p^2}.$$

Proof. The map is well-defined because of Fermat's little theorem and the compatibility for the natural action is clear. As the dimensions on both sides agree, it suffices to prove injectivity. Let $f \in \mathbb{F}_p[X, Y]$ be a polynomial having degree $\leq p - 1$ in both variables such that $f(a, b) = 0$ for all $a, b \in \mathbb{F}_p$. Then for fixed a the polynomial $f(a, Y)$ is identically zero, as it is zero for all the p specialisations of Y. Hence, considering f as a polynomial in Y with coefficients in $\mathbb{F}_p[X]$, it follows that all those coefficients are identically zero for the same argument. Consequently, the polynomial f is zero as an element of $\mathbb{F}_p[X, Y]$ proving the claim. □

We can thus identify $\mathcal{W}(p,\mathbb{F}_p)$ with $\{f \in \mathbb{F}_p[X,Y]/(X^p-X, Y^p-Y) \mid f((0,0))=0\}$. Let $U_d(\mathbb{F}_p)$ be the subspace consisting of polynomial classes of degree $d \in \{0,\dots,p-2\}$, i.e. those that satisfy $f(lx,ly) = l^d f(x,y)$ for all $l \in \mathbb{F}_p$. Note that the degree is naturally defined modulo $p-1$. It is clear that the natural $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-action respects the degree. By collecting the monomials we obtain

$$\mathcal{W}(p,\mathbb{F}_p) = \bigoplus_{d=0}^{p-2} U_d(\mathbb{F}_p).$$

Furthermore, we dispose of the perfect bilinear pairing

$$\mathcal{W}(p,\mathbb{F}_p) \times \mathcal{W}(p,\mathbb{F}_p) \to \mathbb{F}_p, \quad \langle f,g\rangle = \sum_{(a,b)\in\mathbb{F}_p^2} f(a,b)g(a,b).$$

(3.2.2) Lemma. *Let $d,e \geq 0$ be integers. With $(p-1) \nmid d$ or $(p-1) \nmid e$ we have*

$$\sum_{(a,b)\in\mathbb{F}_p^2} a^d b^e = 0.$$

Proof. As the statement is symmetric in d and e, we may suppose that $(p-1) \nmid e$ and in particular $e \neq 0$. Then $\sum_{(a,b)\in\mathbb{F}_p^2} a^d b^e = \sum_{a=0}^{p} a^d (\sum_{b=1}^{p-1} b^e)$. The latter sum, however, is zero, as one can for instance see by choosing a generator σ of $\mathbb{F}_p^*$ and rewriting $\sum_{b=1}^{p-1} b^e = \sum_{i=1}^{p-1} (\sigma^e)^i$. As σ^e clearly is a zero of the polynomial $X^{p-1}-1$, it is a zero of the polynomial $\sum_{i=1}^{p-1} X^i$, since $\sigma^e \neq 1$ using $(p-1) \nmid e$. □

If $(p-1) \nmid (d+e)$, Lemma (3.2.2) implies that $U_d(\mathbb{F}_p)$ pairs to zero with $U_e(\mathbb{F}_p)$. Hence, the restricted pairing $U_d(\mathbb{F}_p) \times U_{p-1-d}(\mathbb{F}_p) \to \mathbb{F}_p$ is perfect for $0 \leq d \leq p-1$, as the dimensions of $U_{p-1-d}(\mathbb{F}_p)$ and $U_d(\mathbb{F}_p)$ are equal. Furthermore, $\mathbb{F}_p[X,Y]_d$ pairs to zero with $\mathbb{F}_p[X,Y]_{p-1-d}$. This follows from Lemma (3.2.2) and an easy calculation. Consequently the induced pairing $U_d(\mathbb{F}_p)/V_d(\mathbb{F}_p) \times V_{p-1-d}(\mathbb{F}_p) \to \mathbb{F}_p$ is perfect.

Weight $k \in \{2,\dots,p+1\}$ in weight 2

Let $M \in \mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$ such that its reduction modulo p is invertible. Then it is clear that the above pairing respects the action of M, i.e. $\langle Mf, Mg\rangle = \langle f,g\rangle$. Consequently, we receive an isomorphism of $\mathbb{F}_p$-vector spaces

$$U_d(\mathbb{F}_p)/V_d(\mathbb{F}_p) \to V_{p-1-d}(\mathbb{F}_p)^\vee$$

respecting the left action defined before. Composing with the map from Proposition (2.2.4), we obtain an isomorphism

$$U_d(\mathbb{F}_p)/V_d(\mathbb{F}_p) \to V_{p-1-d}(\mathbb{F}_p).$$

We now study how the $\mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$-action behaves with respect to this isomorphism.

(3.2.3) Lemma. *Let $0 < d \leq p-1$ and let $M \in \mathrm{Mat}_2(\mathbb{Z})_{\neq 0}$ such that its reduction modulo p is in $GL_2(\mathbb{F}_p)$. Then the following diagram commutes:*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & V_d(\mathbb{F}_p) & \longrightarrow & U_d(\mathbb{F}_p) & \longrightarrow & V_{p-1-d}(\mathbb{F}_p) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle M.} & & \downarrow{\scriptstyle M.} & & \downarrow{\scriptstyle \det(M)^d M.} & & \\ 0 & \longrightarrow & V_d(\mathbb{F}_p) & \longrightarrow & U_d(\mathbb{F}_p) & \longrightarrow & V_{p-1-d}(\mathbb{F}_p) & \longrightarrow & 0. \end{array}$$

Proof. This follows from the compatibilities of the two pairings with the group actions described above. □

(3.2.4) Lemma. *Let $M = \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)$ and $0 < d \leq p-1$. Then the following diagram commutes:*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & V_d(\mathbb{F}_p) & \longrightarrow & U_d(\mathbb{F}_p) & \longrightarrow & V_{p-1-d}(\mathbb{F}_p) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle M^\iota.} & & \downarrow{\scriptstyle M^\iota.} & & \downarrow{\scriptstyle 0} & & \\ 0 & \longrightarrow & V_d(\mathbb{F}_p) & \longrightarrow & U_d(\mathbb{F}_p) & \longrightarrow & V_{p-1-d}(\mathbb{F}_p) & \longrightarrow & 0. \end{array}$$

Proof. We have $M^\iota = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$. A basis of $U_d(\mathbb{F}_p)$ is given by the monomials of degree d, which correspond to the embedding of $V_d(\mathbb{F}_p)$, together with the monomials $X^i Y^{p-1+d-i}$ for $d \leq i \leq p-1$. As the latter monomials all contain at least one factor of X, they are killed by applying the matrix. □

We hence find formulae similar to those that hold in a comparable situation for the action on modular forms of level Np (see Proposition (3.3.8), resp. [Gross], p. 475). The following Proposition, except for the parabolic part, is also [Ash-Stevens], Theorem 3.4.

We introduce the following notation. Let M be any $\mathbb{F}_p$-vector space on which the Hecke operators T_l and the p-part of the diamond operators $\langle\cdot\rangle_p$ act. By $M[d]$ we mean M with the action of the Hecke operator T_l "twisted" to be $l^d T_l$ (in particular T_p acts as zero). Furthermore, by $M(d)$ be denote the subspace on which $\langle l\rangle_p$ acts as $l^d = \chi_p(l)^d$ with χ_p the mod p cyclotomic character.

(3.2.5) Proposition. *Let p be a prime, $N \geq 5$ and $0 < d \leq p-1$ integers such that $p \nmid N$. We have isomorphisms respecting the Hecke operators*

$$H^1(\Gamma_1(Np), \mathbb{F}_p)(d) \cong H^1(\Gamma_1(N), U_d(\mathbb{F}_p)) \quad \textit{and}$$

$$H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p)(d) \cong H^1_{\mathrm{par}}(\Gamma_1(N), U_d(\mathbb{F}_p)).$$

Moreover, there are the exact sequences

$$H^1(\Gamma_1(N), V_d(\mathbb{F}_p)) \hookrightarrow H^1(\Gamma_1(N), U_d(\mathbb{F}_p)) \twoheadrightarrow H^1(\Gamma_1(N), V_{p-1-d}(\mathbb{F}_p))[d]$$

and

$$H^1_{\mathrm{par}}(\Gamma_1(N), V_d(\mathbb{F}_p)) \hookrightarrow H^1_{\mathrm{par}}(\Gamma_1(N), U_d(\mathbb{F}_p)) \twoheadrightarrow H^1_{\mathrm{par}}(\Gamma_1(N), V_{p-1-d}(\mathbb{F}_p))[d],$$

which respect the Hecke operators.

Proof. The first statement follows from Propositions (3.1.4) and (3.1.5) together with the definition of $U_d(\mathbb{F}_p)$. The twisting of the Hecke action in the exact sequences is clear from the definition of the Hecke operators on group cohomology using Lemmas (3.2.3) and (3.2.4).

For $d = p-1$ we have $U_0(\mathbb{F}_p) = V_0(\mathbb{F}_p) \oplus V_{p-1}(\mathbb{F}_p)$, from which the statements follow. So we now assume $d < p-1$, in particular $p \neq 2$. For the top sequence we only need to check that it is exact on the left and on the right. By Proposition (2.2.6) we have $H^0(\Gamma_1(N), V_{p-1-d}(\mathbb{F}_p)) = 0$. The H^2-terms are trivial as the cohomological dimension of $\Gamma_1(N)$ is one, since the group acts freely on the upper half plane and is hence a free group.

The exactness of the second sequence follows from the snake lemma, once we have established the exactness of

$$0 \to \bigoplus_{c \text{ cusps}} H^1(D_c, V_d(\mathbb{F}_p)) \to \bigoplus_{c \text{ cusps}} H^1(D_c, U_d(\mathbb{F}_p)) \to \bigoplus_{c \text{ cusps}} H^1(D_c, V_{p-1-d}(\mathbb{F}_p)) \to 0,$$

where D_c is the stabiliser group of the cusp $c = g\infty$ with $g \in \mathrm{SL}_2(\mathbb{Z})$. Hence, $D_c = g\langle \pm T \rangle g^{-1} \cap \Gamma_1(N)$. This group is infinite cyclic generated by $g \left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right) g^{-1}$ for some $r \in \mathbb{Z}$. Hence, we have $H^2(D_c, V_d(\mathbb{F}_p)) = 0$. If r is 0 modulo p, the sequence

$$0 \to \bigoplus_{c \text{ cusps}} H^0(D_c, V_d(\mathbb{F}_p)) \to \bigoplus_{c \text{ cusps}} H^0(D_c, U_d(\mathbb{F}_p)) \to \bigoplus_{c \text{ cusps}} H^0(D_c, V_{p-1-d}(\mathbb{F}_p)) \to 0$$

is clearly right exact, as the action of D_c on the modules is trivial. If r is invertible in $\mathbb{F}_p$, it follows as in Lemma (2.2.5) that both $H^0(D_c, V_d(\mathbb{F}_p))$ and $H^0(D_c, V_{p-1-d}(\mathbb{F}_p))$ are 1-dimensional. To finish the proof, it thus suffices to prove that $H^0(D_c, U_d(\mathbb{F}_p))$ is (at least) 2-dimensional. The elements $X^d \in U_d(\mathbb{F}_p)$ and $Y^d(1-X^{p-1}) \in U_d(\mathbb{F}_p)$ are invariant under T. Indeed,

$$\begin{aligned} T.Y^d(1-X^{p-1}) &= (X+Y)^d(1-X^{p-1}) \\ &= Y^d(1-X^{p-1}) + \sum_{i=1}^{d} \left(\tfrac{d}{i}\right) Y^{d-i}X^i(1-X^{p-1}) = Y^d(1-X^{p-1}), \end{aligned}$$

as in $U_d(\mathbb{F}_p)$ we have $X^i(1-X^{p-1}) = X^{i-1}(X-X^p) = 0$ for $i > 0$. □

3.3. Hecke algebras

In this section we will compare the Hecke algebra of modular forms with that of modular symbols and establish isomorphisms in certain cases. Whenever we have an R-module M, on which Hecke operators T_n act for all n, we let

$$\mathbb{T}_R(M) := R[T_n \mid n \in \mathbb{N}] \subseteq \mathrm{End}_R(M),$$

i.e. the R-subalgebra of the endomorphism algebra generated by the Hecke operators.

The Hecke algebra of modular forms and Eichler-Shimura

We recall a theorem by Eichler and Shimura.

(3.3.1) Theorem. (Eichler-Shimura) *For $k \geq 2$ and $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ a congruence subgroup, there is an isomorphism of $\mathbb{T}_{\mathbb{Z}}(S_k(\Gamma, \mathbb{C}))$-modules, the* Eichler-Shimura isomorphism*,*

$$H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{C})) \cong S_k(\Gamma, \mathbb{C}) \oplus \overline{S_k(\Gamma, \mathbb{C})}.$$

Proof. [DiamondIm], Theorem 12.2.2. □

(3.3.2) Corollary. *In the situation of Theorem (3.3.1) we have natural ring isomorphisms*

$$\mathbb{T}_{\mathbb{Z}}\big(S_k(\Gamma, \mathbb{C})\big) \cong \mathbb{T}_{\mathbb{Z}}\big(H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{Z}))/\mathrm{torsion}\big).$$

Proof. It is clear that the $\mathbb{C}$-vector space $H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{C}))$ contains the natural $\mathbb{Z}$-structure $H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{Z}))/\mathrm{torsion}$. This follows for instance from Remark (2.5.3) together with the comparison result Theorem (2.6.1). Any $\mathbb{Z}$-structure, however, gives an isomorphic Hecke algebra. Finally, Theorem (3.3.1) implies that the Hecke algebra of $H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{C}))$ is isomorphic to the Hecke algebra of $S_k(\Gamma, \mathbb{C})$. □

The formula in this corollary is the reason why many people prefer to factor out the torsion of modular symbols.

(3.3.3) Proposition. *Let $N \geq 5$, $k \geq 2$ integers and $p \nmid N$ a prime. Then we have*

$$\mathbb{T}_{\mathbb{Z}}\big(S_k(\Gamma_1(N), \mathbb{C})\big) \otimes_{\mathbb{Z}} \mathbb{F}_p \cong \mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big).$$

Proof. By [DiamondIm], Theorem 12.3.2, we have

$$S_k(\Gamma_1(N), \mathbb{Z}[1/N]) \otimes_{\mathbb{Z}[1/N]} \mathbb{F}_p \cong S_k(\Gamma_1(N), \mathbb{F}_p).$$

We note that in this case there is no difference between Katz modular forms and those that are reductions of classical modular forms whose q-expansion is in $\mathbb{Z}[1/N]$. By the q-expansion principle we hence have the two perfect pairings

$$\mathbb{T}_{\mathbb{Z}}\big(S_k(\Gamma_1(N), \mathbb{C})\big) \otimes_{\mathbb{Z}} \mathbb{Z}[1/N] \times S_k(\Gamma_1(N), \mathbb{Z}[1/N]) \to \mathbb{Z}[1/N], \quad (T, f) \mapsto a_1(Tf)$$

and

$$\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big) \times S_k(\Gamma_1(N), \mathbb{F}_p) \to \mathbb{F}_p, \quad (T, f) \mapsto a_1(Tf).$$

Tensoring the first one with $\mathbb{F}_p$ allows us to compare it to the second one, from which the proposition follows. □

(3.3.4) Corollary. *Let p be a prime and $N \geq 5$, $2 \leq k \leq p+2$ integers s.t. $p \nmid N$. Then the $\mathbb{F}_p$-algebra homomorphism*

$$\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big) \twoheadrightarrow \mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))\big),$$

sending the operator T_l to T_l for all primes l is a surjection.

Proof. From Corollary (3.3.2) we obtain because of p-torsion-freeness (Proposition (2.4.8) together with the comparison result Theorem (2.6.1)) an isomorphism of $\mathbb{F}_p$-algebras

$$\mathbb{T}_{\mathbb{Z}}\big(S_k(\Gamma_1(N), \mathbb{C})\big) \otimes \mathbb{F}_p \cong \mathbb{T}_{\mathbb{Z}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{Z}_p))\big) \otimes_{\mathbb{Z}_p} \mathbb{F}_p.$$

By Proposition (3.3.3) the term on the left hand side is equal to $\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big)$ so that it suffices to have a surjection

$$\mathbb{T}_{\mathbb{Z}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{Z}_p))\big) \otimes \mathbb{F}_p \twoheadrightarrow \mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))\big),$$

which follows from Proposition (2.4.8). Indeed, the isomorphism

$$H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{Z}_p)) \otimes \mathbb{F}_p \cong H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))$$

is compatible with Hecke operators, and allows to define a homomorphism from the Hecke algebra on the left hand term to the one on the right hand term, which is automatically surjective by the definition of the Hecke algebra. □

(3.3.5) Proposition. *Let $N \geq 1$, $k \geq 2$ be integers and K a field. If the characteristic of K is $p > 0$, then we assume $p \nmid N$. Furthermore, let $\Gamma_1(N) \leq \Gamma \lhd G \leq \mathrm{SL}_2(\mathbb{Z})$ be subgroups and $\epsilon : G \overset{\mathrm{proj}}{\twoheadrightarrow} \Gamma\backslash G \to R^*$ a character such that $\epsilon(-\mathbf{1}) = (-1)^k$ if $-\mathbf{1} \in G$. Denote by $\mathbb{T}$ the K-Hecke algebra of $S_k(\Gamma, K)$ and by $\mathbb{T}_\epsilon$ the K-Hecke algebra of $S_k(G, \epsilon, K)$. Furthermore, let*

$$I = (\langle\delta\rangle - \epsilon(\delta) \mid \delta \in \Gamma\backslash G) \lhd \mathbb{T}.$$

Then $\mathbb{T}/I$ and $\mathbb{T}_\epsilon$ are isomorphic K-algebras.

Proof. As we work with Katz modular forms (for that we need the condition $p \nmid N$), we dispose of the q-expansion principle. Hence we have isomorphisms respecting the Hecke action $(\mathbb{T}_\epsilon)^\vee \cong S_k(G, \epsilon, K) = \mathbb{T}^\vee[I] \cong (\mathbb{T}/I)^\vee$, whence the proposition follows. □

Applying p-adic Hodge Theory

In this section we present an analog of the Eichler-Shimura isomorphism, formulated in terms of p-adic Hodge theory. This was already used in [EdixJussieu], Theorem 5.2, to derive an algorithm for computing modular forms. However, p-adic Hodge theory always has the restriction that the weight be smaller than p.

(3.3.6) Theorem. (Fontaine, Messing, Faltings) *Let p be a prime and $N \geq 5$, $2 \leq k < p$ be integers s.t. $p \nmid N$. Then the Galois representation $H^1_{\text{ét, par}}(Y_1(N)_{\overline{\mathbb{Q}}_p}, \mathrm{Sym}^{k-2}(\mathbb{V}))^\vee$ is crystalline, where $\mathbb{V} = R^1\pi_*\mathbb{F}_p$ with $\pi : \mathbb{E} \to Y_1(N)$ the universal elliptic curve. The corresponding ϕ-module D sits in the exact sequence*

$$0 \to S_k(\Gamma_1(N), \mathbb{F}_p) \to D \to S_k(\Gamma_1(N), \mathbb{F}_p)^\vee \to 0,$$

which is equivariant for the action of the Hecke operators.

This can be compared to Theorem 1.1 and Theorem 1.2 of [FaltingsJordan]. Part (a) of the following corollary is part of [EdixJussieu], Theorem 5.2.

(3.3.7) Corollary. *Let $N \geq 5$, $p \nmid N$ and $2 \leq k < p$.*

(a) The parabolic group cohomology group $H^1_{\mathrm{par}}(\Gamma_1(N), \mathbb{V}_{k-2}(\mathbb{F}_p))$ is a faithful module for $\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big)$.

(b) Let $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \overline{\mathbb{F}_p}^*$ be a character. Define the ideal

$$I = (\langle l\rangle - \epsilon(l) \mid (l, N) = 1) \lhd \mathbb{T}_{\overline{\mathbb{F}_p}}\big(S_k(\Gamma_1(N), \overline{\mathbb{F}_p})\big).$$

Then $\big(H^1_{\mathrm{par}}(\Gamma_1(N), \mathbb{V}_{k-2}(\mathbb{F}_p)) \otimes_{\mathbb{F}_p} \overline{\mathbb{F}_p}\big)/I$ is a faithful module for the Hecke algebra $\mathbb{T}_{\overline{\mathbb{F}_p}}\big(S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})\big)$.

Proof. (a) From Theorem (3.3.6) we know that D is a faithful Hecke module. Hence, so is $H^1_{\text{ét, par}}(Y_{\Gamma_1(N)}, \mathrm{Sym}^{k-2}(\mathbb{V}))$. This module can be identified with its analog in analytic cohomology which is isomorphic to $H^1_{\mathrm{par}}(\Gamma_1(N), \mathbb{V}_{k-2}(\mathbb{F}_p))$ (see Chapter II).

(b) If the Hecke operator T acts as zero on $\big(H^1_{\text{ét, par}}(Y_{\Gamma_1(N)}, \mathrm{Sym}^{k-2}(\mathbb{V})) \otimes_{\mathbb{F}_p} \overline{\mathbb{F}_p}\big)/I$, then it acts as zero on $(D \otimes \overline{\mathbb{F}_p})/I$, hence also on $S_k(\Gamma_1(N), \overline{\mathbb{F}_p})^\vee/I = \mathbb{T}_{\overline{\mathbb{F}_p}}\big(S_k(\Gamma_1(N), \overline{\mathbb{F}_p})\big)/I$, from which $T \in I$ follows. The statement now follows from Proposition (3.3.5). □

Modular forms of weight 2 and level Np

We recall some work of Serre as explained in [Gross], cf. also [EdixWeight], Section 6.

Let us now introduce notation that is used throughout the sequel of this chapter. We consider the modular curve $X_1(Np)$ over $\mathbb{Q}_p(\zeta_p)$ for a prime $p > 2$ not dividing $N \geq 5$. It has a regular stable model X over the ring $\mathbb{Z}_p[\zeta_p]$, see e.g. [Katz-Mazur]. Let J denote the Néron model over $\mathbb{Z}_p[\zeta_p]$ of $J_1(Np)$, the Jacobian of $X_1(Np)$ over $\mathbb{Q}_p(\zeta_p)$. We let, following [Gross], Section 8,

$$L = H^0(X, \Omega_{X/\mathbb{Z}_p[\zeta_p]}),$$

where $\Omega_{X/\mathbb{Z}_p[\zeta_p]}$ is the *dualising sheaf of* X of [Deligne-Rapoport], Section I.2. By [Gross], Equation 8.2, we have for the special fibre $X_{\mathbb{F}_p}$ that

$$\overline{L} := H^0(X_{\mathbb{F}_p}, \Omega_{X_{\mathbb{F}_p}/\mathbb{F}_p}) = L \otimes_{\mathbb{Z}_p[\zeta_p]} \mathbb{F}_p.$$

On L and $\overline{L}$ we have the action of the p-part $\langle\cdot\rangle_p$ of the diamond action. The principal result on $\overline{L}$ that we will need is the following, which is Proposition 8.13 and Proposition 8.18 in [Gross].

(3.3.8) Proposition. (Serre) *Assume $3 \leq k \leq p$, $N \geq 5$ and $p \nmid N$. Then there is an isomorphism of $\mathbb{T}_{\mathbb{F}_p}(\overline{L})$-modules*

$$\overline{L}(k-2) \cong S_k(\Gamma_1(N), \mathbb{F}_p) \oplus S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)[k-2].$$

Moreover, the sequence of Hecke modules

$$0 \to S_2(\Gamma_1(N), \mathbb{F}_p)[p-1] \to \overline{L}(p-1) \to S_{p+1}(\Gamma_1(N), \mathbb{F}_p) \to 0$$

is exact.

In our attempt to compare Hecke algebras of modular forms with those of modular symbols in characteristic p, we generalise the strategy of the second part of the proof of [EdixJussieu], Theorem 5.2. Hence, we wish to bring the Jacobian into the play, since it will enable us to pass from characteristic zero geometry to characteristic p.

(3.3.9) Lemma. *Under the assumptions and notations above we have isomorphisms*

$$\overline{L} \cong \mathrm{Cot}_0(J^0_{\mathbb{F}_p}) \cong \mathrm{Cot}_0(J^0_{\mathbb{F}_p}[p]).$$

Proof. The first isomorphism is e.g. [EdixWeight], Equation 6.7.2. The second one follows from the fact that multiplication by p on $J^0_{\mathbb{F}_p}$ induces multiplication by p on the tangent space at 0, which is the zero map. Hence, the tangent space at 0 of $J^0_{\mathbb{F}_p}[p]$ is equal to the one of $J^0_{\mathbb{F}_p}$. □

Parabolic cohomology and the p-torsion of the Jacobian

To establish an explicit link between parabolic cohomology and modular forms, we identify the parabolic cohomology group for $\Gamma_1(N)$ with $\mathbb{F}_p$-coefficients as the p-torsion of the Jacobian of the corresponding modular curve. Here we may view the Jacobian as a complex abelian variety.

(3.3.10) Proposition. *Let $N \geq 3$ be an integer, and p a prime. Then we have an isomorphism of $\mathbb{T}_{\mathbb{Z}}(S_2(\Gamma_1(Np), \mathbb{C})) \otimes \mathbb{F}_p$-modules*

$$H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p) \cong J(\mathbb{C})[p] = J(\overline{\mathbb{Q}}_p)[p].$$

Proof. The second equality follows from the fact that torsion points are algebraic. We start with the exact *Kummer sequence* of analytic sheaves over $X_1(Np)$

$$0 \to \mu_p \to \mathbb{G}_m \xrightarrow{p} \mathbb{G}_m \to 0.$$

Its long exact sequence in analytic cohomology yields

$$0 \to H^1(X_1(Np), \mu_p) \to H^1(X_1(Np), \mathbb{G}_m) \xrightarrow{p} H^1(X_1(Np), \mathbb{G}_m).$$

Using that $H^1(X_1(Np), \mathbb{G}_m) = J(\mathbb{C})$, we already obtain that $H^1(X_1(Np), \mu_p) \cong J(\mathbb{C})[p]$. As $\mathbb{C}$ contains the p-th roots of unity, we may replace the sheaf μ_p by the constant sheaf $\mathbb{F}_p$. Moreover, the group $H^1(X_1(Np), \mathbb{F}_p)$ coincides with $H^1_{\mathrm{par}}(Y_1(Np), \mathbb{F}_p)$ (see Proposition (2.4.1)), which in turn is equal to $H^1(\Gamma_1(Np), \mathbb{F}_p)$, using that $\mathbb{H} \twoheadrightarrow Y_1(Np)$ is a Galois covering under the assumption $N \geq 3$. □

Comparing Hecke algebras over $\mathbb{F}_p$

(3.3.11) Proposition. *Let $N \geq 5$ be an integer, $p \nmid N$ a prime and $0 \leq d \leq p-1$ an integer. There exists a surjection $\mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p)(d)\big) \twoheadrightarrow \mathbb{T}_{\mathbb{F}_p}(\overline{L}(d))$ such that the diagram of $\mathbb{F}_p$-algebras*

$$\begin{array}{ccc} & & \mathbb{T}_{\mathbb{F}_p}(\overline{L}(d)) \\ & \nearrow & \uparrow \\ \mathbb{T}_{\mathbb{Z}}\big(S_2(\Gamma_1(Np), \mathbb{C})(d)\big) \otimes \mathbb{F}_p & & \\ & \searrow & \\ & & \mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p)(d)\big) \end{array}$$

commutes. All maps are uniquely determined by sending the Hecke operator T_l to T_l.

Proof. Let us first remark how the diagonal arrows are made. The lower one comes from the isomorphism (see Proposition (2.4.8) and Theorem (2.6.1))

$$H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{Z}) \otimes \mathbb{F}_p \cong H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p).$$

The upper one is due to the fact that L is a lattice in $S_2(\Gamma_1(Np, \mathbb{C}))$, using arguments as in Corollary (3.3.2). We use that the order of $\mathbb{F}_p^*$ is invertible in $\mathbb{F}_p$, so that we can everywhere use the eigencomponents of the action of the p-part of the diamond operator $\langle\cdot\rangle_p$.

We obtain the vertical arrow by showing that the kernel of the lower diagonal map is contained in the kernel of the upper diagonal map. In other words, we will show that if $T \in \mathbb{T}_{\mathbb{Z}}\big(S_2(\Gamma_1(Np), \mathbb{C})(d)\big) \otimes \mathbb{F}_p$ acts as zero on $H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p)(d)$, then it acts as zero on $\overline{L}(d)$.

So assume that T acts as zero on $H^1_{\mathrm{par}}(\Gamma_1(Np), \mathbb{F}_p)(d)$. By Proposition (3.3.10), it acts as zero on $J_{\overline{\mathbb{Q}}_p}(\overline{\mathbb{Q}}_p)[p](d)$, hence on $J_{\overline{\mathbb{Q}}_p}[p](d)$. But then it also acts as zero on $J_{\mathbb{Z}_p[\zeta_p]}[p](d)$, as it acts as zero on the generic fibre using that $J[p]$ is flat over $\mathbb{Z}_p[\zeta_p]$ ([BLR], Lemma 7.3.2, as J is semi-abelian) and that $J_{\mathbb{Q}_p}[p]$ is reduced. But consequently, it also acts as zero on the special fibre $J[p](d)$, whence also on the cotangent space $\mathrm{Cot}_0(J^0[p])(d)$. Now Lemma (3.3.9) finishes the proof. □

(3.3.12) Theorem. *Let $2 < k \leq p+1$, $N \geq 5$ such that $p \nmid N$. We write for short $\mathbb{T}_{\mathrm{par},N,k} := \mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))\big)$, $\mathbb{T}_{\mathrm{mod},N,k} := \mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big)$ and similarly for the twisted ones. Then there is the commutative diagram of $\mathbb{F}_p$-algebras*

$$\begin{array}{ccccc}
\mathbb{T}_{\mathbb{F}_p}\big(\overline{L}(k-2)\big) & \longrightarrow & \mathbb{T}_{\mathrm{mod},N,k} & \times & \mathbb{T}_{\mathrm{mod},N,p+3-k,[k-2]} \\
\uparrow & & \downarrow & & \downarrow \\
\mathbb{T}_{\mathrm{par},Np,2} & \longrightarrow & \mathbb{T}_{\mathrm{par},N,k} & \times & \mathbb{T}_{\mathrm{par},N,p+3-k,[k-2]}.
\end{array}$$

The vertical arrows are obtained from Proposition (3.3.11) resp. Corollary (3.3.4), and the horizontal ones from Proposition (3.3.8) and Proposition (3.2.5). The vertical arrows are surjective. If $2 < k \leq p$, then the upper horizontal arrow is injective.

Proof. The commutativity is clear, as T_l is sent to $T_l \times T_l$ along the horizontal arrows, and T_l is sent to T_l along the vertical arrows. The surjectivity of the vertical arrows has been proved at the places cited above.

The injectivity of the upper homomorphism is the fact that $\overline{L}(k-2)$ is the direct sum of $S_k(\Gamma_1(N), \mathbb{F}_p)$ and $S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)[k-2]$, if $2 < k \leq p$. □

(3.3.13) Corollary. *Let $2 < k \leq p+1$, $N \geq 5$ such that $p \nmid N$. Let $\mathfrak{P}$ be a maximal ideal of $\mathbb{T}_{\mathbb{F}_p}\big(\overline{L}(k-2)\big)$ which is not in the support of $S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)$. Then we have an isomorphism*

$$\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)_{\mathfrak{P}}\big) \cong \mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))_{\mathfrak{P}}\big).$$

Proof. The assumption means that $(S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)[k-2])_{\mathfrak{P}} = 0$. Because of Corollary (3.3.4) we know that $\mathfrak{P}$ is not in the support of $H^1_{\mathrm{par}}(\Gamma_1(N), V_{p+1-k}(\mathbb{F}_p))[k-2]$ either, whence $(H^1_{\mathrm{par}}(\Gamma_1(N), V_{p+1-k}(\mathbb{F}_p))[k-2])_{\mathfrak{P}} = 0$. Hence, the sequence of Proposition (3.2.5) localised at $\mathfrak{P}$ is split, and all maps in the localisation of the diagram of Theorem (3.3.12) are isomorphisms. □

(3.3.14) Corollary. *Let $2 < k \leq p+1$, $N \geq 5$ such that $p \nmid N$. Let $\mathfrak{P}$ be a maximal ideal of $\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)\big)$ corresponding to a normalised eigenform $f \in S_k(\Gamma_1(N), \mathbb{F}_p)$ which is* ordinary, *i.e. $a_p(f) \neq 0$. Then we have an isomorphism*

$$\mathbb{T}_{\mathbb{F}_p}\big(S_k(\Gamma_1(N), \mathbb{F}_p)_{\mathfrak{P}}\big) \cong \mathbb{T}_{\mathbb{F}_p}\big(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))_{\mathfrak{P}}\big).$$

Proof. As the operator T_p always acts as zero on $S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)[k-2]$ the maximal ideal $\mathfrak{P}$ cannot be in the support of $S_{p+3-k}(\Gamma_1(N), \mathbb{F}_p)[k-2]$, whence we are in the situation of Corollary (3.3.13). □

(3.3.15) Remark. *In contrast to Proposition (3.3.8) the exact sequence of Proposition (3.2.5) is in general non-split for $d = k-2$ with $2 < k \leq p$. However, it is split for $k = 2$.*

Action through characters

(3.3.16) Lemma. *In the situation of Proposition (3.3.18) we have*

$$H^1(\Delta, S_k(\Gamma, K)) = 0.$$

This also holds for $k = 2$ away from Eisenstein ideals.

Proof. Without loss of generality we may assume that Δ is a p-group and hence that $\overline{G}$ acts freely on $\mathbb{H}$ and that $\pi : X_\Gamma \twoheadrightarrow X_G$ is a Galois cover with group Δ of proper K-schemes. The group action of Δ on cohomology is through the Diamond operators. The Hochschild-Serre spectral sequence gives an injection

$$0 \to H^1(\Delta, H^0(X_\Gamma, \pi^*\omega^{\otimes k}(-\mathrm{cusps}))) \to H^1(X_G, \omega^{\otimes k}(-\mathrm{cusps})).$$

Using Serre duality and the Kodaira-Spencer isomorphism we obtain

$$H^1(X_G, \omega^{\otimes k}(-\mathrm{cusps})) \overset{\text{S-D}}{\cong} H^0(X_G, \Omega^1 \otimes (\omega^{\otimes k}(-\mathrm{cusps}))^\vee)^\vee \overset{\text{K-S}}{\cong} H^0(X_G, \omega^{\otimes 2-k})^\vee$$

which is zero, since the degree of $\omega^{\otimes 2-k}$ is negative (as $k \geq 3$). The map π is étale and we have $H^0(X_\Gamma, \pi^*\omega^{\otimes k}(-\mathrm{cusps})) \cong S_k(\Gamma, K)$, from which the claim follows. For $k = 2$ we have $H^1(X_G, \omega^{\otimes 2}(-\mathrm{cusps})) \cong H^0(X_G, \mathcal{O})^\vee$, which is 1-dimensional. As a Hecke module it cannot be in the support of a non-Eisenstein prime. □

(3.3.17) Corollary. *In the situation of Proposition (3.3.18) suppose that K has characteristic p and that Δ is a p-group. Then $S_k(\Gamma, K)$ is an induced $K[\Delta]$-module.*

Proof. This follows from Lemma (3.3.16) and Proposition (2.7.13). □

(3.3.18) Proposition. *Let $k \geq 3$, $N \geq 1$ be integers and K a field. Furthermore, let $\Gamma_1(N) \leq \Gamma \lhd G \leq \mathrm{SL}_2(\mathbb{Z})$ be subgroups and $\epsilon : G \overset{\text{proj}}{\twoheadrightarrow} \Gamma\backslash G \to K^*$ a character such that $\epsilon(-\mathbf{1}) = (-1)^k$ if $-\mathbf{1} \in G$. Let $\Delta := \Gamma\backslash G$. We assume that Δ is abelian and that $\overline{\Gamma}$ acts without stabilisers on $\mathbb{H}$. If the characteristic of K is $p > 0$, then we assume $p \nmid N$ and that $\overline{G}$ has no stabilisers of order p for its action on $\mathbb{H}$.*

Then the norm N_Δ induces an isomorphism

$$\big(S_k(\Gamma, K) \otimes K^\epsilon\big)_\Delta \xrightarrow{N_\Delta} \big(S_k(\Gamma, K) \otimes K^\epsilon\big)^\Delta = S_k(G, \epsilon, K).$$

When $k = 2$, then the statements also hold if one localises away from Eisenstein maximal ideals (i.e. those not corresponding to irreducible Galois representations).

Proof. If the characteristic of K is zero, the finite abelian group Δ acts semi-simply, and hence the claim follows. If the characteristic is p, it suffices to prove the statement for the p-Sylow subgroup Δ_p of Δ, as again Δ/Δ_p acts semi-simply. Corollary (3.3.17) implies that $S_k(\Gamma, K)$ is a cohomologically trivial (for Tate cohomology) $K[\Delta_p]$-module. Consequently, the norm induces an isomorphism. □

(3.3.19) Remark. *If the characteristic of K is p, then the result of Proposition (3.3.18) also holds for $k = 1$. For the Δ-action commutes with the derivation Θ used in Proposition (4.5.2). As the Δ-invariants agree with the Δ-coinvariants in weights p and $p+2$, it follows that the same holds in weight one by the exact sequence in Part (a) of that proposition.*

(3.3.20) Proposition. *We keep the assumptions of Proposition (3.3.18). If the characteristic of K is $p > 0$, we also assume $k \leq p+2$.*

If $\mathcal{CM}_k(\Gamma, K)$ is a faithful $\mathbb{T}_K(S_k(\Gamma, K))$-module, then $\mathcal{CM}_k(G, \epsilon, K)$ is a faithful $\mathbb{T}_K(S_k(\Gamma, \epsilon, K))$-module. For $k = 2$ similar statements hold away from Eisenstein primes.

Proof. Dualising the result of Proposition (3.3.18) gives an isomorphism

$$\big(\mathbb{T}(S_k(\Gamma, K)) \otimes K^\epsilon\big)_\Delta \xrightarrow{N_\Delta} \big(\mathbb{T}(S_k(\Gamma, K)) \otimes K^\epsilon\big)^\Delta,$$

which in particular yields that the implication

$$T(\sum_{\delta \in \Delta} \epsilon(\delta)^{-1}\langle\delta\rangle) = 0 \quad \Rightarrow \quad T \in I,$$

where I is the ideal defined in Proposition (3.3.5). In view of that proposition, we only need to show that if T acts as zero on $\mathcal{CM}_k(G, \epsilon, K)$, then T is in I.

That can be seen as follows. We now assume that $\overline{\Delta} = \Delta$, i.e. that $G = \overline{G}$. For that we may have to replace G by a subgroup of index 2. This may be done since neither the space of modular symbols nor the space of modular forms changes.

From Proposition (2.7.10) we know

$$\mathcal{CM}_k(G, \epsilon, K) \cong {}_\Delta(\mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon) \overset{N_\Delta}{\cong} {}^\Delta(\mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon).$$

If T acts as zero on

$$\mathcal{CM}_k(G, \epsilon, K) = N_\Delta(\mathcal{CM}_k(\Gamma, K) \otimes_K K^\epsilon) = (\sum_{\delta \in \Delta} \epsilon(\delta)^{-1}\langle\delta\rangle)\mathcal{CM}_k(\Gamma, K),$$

then

$$(T \sum_{\delta \in \Delta} \epsilon(\delta)^{-1}\langle\delta\rangle)\mathcal{CM}_k(\Gamma, K) = 0$$

and by the assumed faithfulness of $\mathcal{CM}_k(\Gamma, K)$, it follows that $T(\sum_{\delta \in \Delta} \epsilon(\delta)^{-1}\langle\delta\rangle) = 0$, whence $T \in I$, as required. □

Chapter IV

Computations of mod p Modular Forms

In this chapter we explain how the results of Chapters II and III can be used algorithmically to compute modular forms over finite fields with methods from linear algebra, most notably modular symbols, under certain restrictions.

As modular forms in the situation when we consider them are uniquely determined by their q-expansions, we only need to compute the corresponding Hecke algebra, since the space of modular forms is its dual. If one is only interested in eigenforms, not the whole Hecke algebra structure is needed, and we can do with fewer conditions. However, the knowledge of the Hecke algebra structure is necessary for the computation of weight one forms, and it is interesting to study e.g. the Gorenstein property in view of a possible identification between the $\mathbb{Z}_p$-Hecke algebra with a deformation ring. We also explain how weight one Hecke algebras can be computed using weight p, following [EdixJussieu]. Moreover, the principal algorithms of my Magma package `Weight1.mg`, which builds on William Stein's package `ModularSymbols`, are presented. Fortunately, Stein's package has already provided modular symbols over finite fields for a long time, and one could say that this chapter is about their interpretation.

We start this chapter by recalling the relation between modular forms and Hecke algebras. Next we present an algorithm which splits a module over a commutative algebra over a finite field into local pieces up to Galois conjugacy. The third section compares systems of eigenvalues of modular forms with those of modular symbols. In the fourth section an algorithm for the computation of Hecke algebras of weight $k \geq 2$ over finite fields using modular symbols is treated. Then we explain how weight one and weight p are related for finite fields of characteristic p, from which we derive an algorithm for the computation of weight one forms, following [EdixJussieu]. The final section sketches a certain generalisation of Merel's universal Fourier expansions.

4.1. Modular forms and Hecke algebras

Let K be a perfect field and $\overline{K}$ an algebraic closure. Let furthermore $S(\overline{K})$ be some space of modular forms defined over K and $\mathbb{T}_K$ the associated Hecke algebra over K, such that the pairing

$$\mathbb{T}_K \times S(\overline{K}) \to \overline{K}, \quad (T, f) \mapsto a_1(Tf)$$

is non-degenerate, where $a_n(f)$ denotes the n-th coefficient of the standard q-expansion of f. This is the case for instance for holomorphic modular forms ($K = \mathbb{C}$) for $\Gamma_1(N)$, $\Gamma_0(N)$ and all $N \geq 1$, or for Katz modular forms over $K = \mathbb{F}_{p^r}$ for $\Gamma_1(N)$, $\Gamma_0(N)$ and all $N \geq 1$ such that $p \nmid N$. The pairing gives rise to the following Hecke equivariant isomorphisms

$$S(\overline{K}) \cong \mathrm{Hom}_K(\mathbb{T}_K, \overline{K}) \cong \mathrm{Hom}_{\overline{K}}(\mathbb{T}_K \otimes_K \overline{K}, \overline{K}),$$

where the first arrow is given by $f \mapsto (T \mapsto a_1(Tf))$. Let us recall the important formula $a_1(T_n f) = a_n(f)$, which follows from the action of the Hecke operators on q-expansions. Normalised Hecke eigenforms in $S(\overline{K})$ correspond under the first isomorphism to K-algebra homomorphisms $\mathbb{T}_K \to \overline{K}$. Eigenforms that are Galois conjugate (i.e. the coefficients of the standard q-expansion are conjugate by $G(\overline{K}|K)$) correspond to Galois conjugate K-algebra homomorphisms $\mathbb{T}_K \to \overline{K}$. Two K-algebra homomorphisms $\mathbb{T}_K \to \overline{K}$ are Galois conjugate if and only if they have the same kernel. It is common to refer to a K-algebra homomorphism $f : \mathbb{T}_K \to \overline{K}$ as the *system of eigenvalues* $(\lambda_n)_n$ *of* $\mathbb{T}_K$ with $\lambda_n = f(T_n)$.

We have established bijections

$$\mathrm{Spec}(\mathbb{T}_K) \overset{1-1}{\longleftrightarrow} \mathrm{Hom}_{\mathrm{K-alg}}(\mathbb{T}_K, \overline{K}) \overset{1-1}{\longleftrightarrow} \{ \text{ normalised eigenforms in } S(\overline{K}) \}/G(\overline{K}|K)$$

and

$$\begin{array}{ccccc} \mathrm{Spec}(\mathbb{T}_K \otimes \overline{K}) & \overset{1-1}{\longleftrightarrow} & \mathrm{Hom}_{\overline{\mathrm{K}}-\mathrm{alg}}(\mathbb{T}_K \otimes \overline{K}, \overline{K}) & \overset{1-1}{\longleftrightarrow} & \\ \{ \text{ normalised eigenforms in } S(\overline{K}) \} & \overset{1-1}{\longleftrightarrow} & \{ \text{ systems of eigenvalues of } \mathbb{T}_K \}. & & \end{array}$$

The Hecke algebra $\mathbb{T}_K$ is finite dimensional (an *Artin algebra*) and commutative. So all its prime ideals are maximal, and using the Chinese Remainder Theorem the algebra decomposes as a product of its localisations:

$$\mathbb{T}_K \cong \prod_{\mathfrak{m} \in \mathrm{Spec}(\mathbb{T}_K)} (\mathbb{T}_K)_{\mathfrak{m}} \cong \prod_{\mathfrak{m} \in \mathrm{Spec}(\mathbb{T}_K)} \mathbb{T}_K/\mathfrak{m}^{\infty} \cong \prod_{\mathfrak{m} \in \mathrm{Spec}(\mathbb{T}_K)} \mathbb{T}_K/(1 - e_{\mathfrak{m}})\mathbb{T}_K.$$

If r is an integer r such that $\mathfrak{m}^r = \mathfrak{m}^{r+1}$, then we write $\mathfrak{m}^{\infty}$ for $\mathfrak{m}^r$. The $e_{\mathfrak{m}}$ in the formula are idempotents corresponding to the decomposition.

4.2. Computing local factors of Hecke algebras

Let K be a perfect field, $\overline{K}$ an algebraic closure and A a finite dimensional commutative K-algebra. We will write A_L for $A \otimes_K L$, where $L|K$ is an extension inside $\overline{K}$. The image of $a \in A$ in $A_{\overline{K}}$ is denoted as $\overline{a}$.

In the context of Hecke algebras we would like to

(1) compute a local decomposition of A, resp.

(2) compute a local decomposition of $A_{\overline{K}}$ keeping track of the $G(\overline{K}|K)$-conjugacy.

In this section we present an algorithm solving both points. This algorithm is implemented in my Magma package `Weight1.mg`. It is based on the following lemma.

(4.2.1) Lemma. *(a) A is local if and only if the minimal polynomial of a (in $K[X]$) is a prime power for all $a \in A$.*

(b) Let V be an A-module such that for all $a \in A$ the minimal polynomial of a on V is a prime power in $K[X]$, i.e. V is a primary space for all $a \in A$. Then the image of A in $\mathrm{End}(V)$ is a local algebra.

(c) Let V be an $A_{\overline{K}}$-module and let $a_1, \ldots, a_n$ be generators of the algebra A. Suppose that for $i \in \{1, \ldots, n\}$ the minimal polynomial of $\overline{a_i}$ on V is a power of $(X - \lambda_i)$ in $\overline{K}[X]$ for some $\lambda_i \in \overline{K}$. Then the image of $A_{\overline{K}}$ in $\mathrm{End}(V)$ is a local algebra.

Proof. (a) Suppose first that A is local and take $a \in A$. Let $\phi_a : K[X] \to A$ be the homomorphism of K-algebras defined by sending X to a. Let (f) be the kernel with f monic, so that by definition f is the minimal polynomial of a. Hence, $K[X]/(f) \hookrightarrow A$, whence $K[X]/(f)$ is local, implying that f cannot have two different prime factors.

Conversely, if A were not local, we would have an idempotent $e \notin \{0, 1\}$. The minimal polynomial of e is $X(X-1)$, which is not a prime power.

(b) follows directly. For (c) one can use the following. Suppose that $(a - \lambda)^r V = 0$ and $(b - \mu)^s V = 0$. Then $((a + b) - (\lambda + \mu))^{r+s} V = 0$, as one sees by rewriting $((a + b) - (\lambda + \mu)) = (a - \lambda) + (b - \mu)$ and expanding out. From this it also follows that $(ab - \lambda\mu)^{2(r+s)} V = 0$ by rewriting $ab - \lambda\mu = (a - \lambda)(b - \mu) + \lambda(b - \mu) + \mu(a - \lambda)$. □

Let us call a pair (V, L) consisting of a finite extension $L|K$ with $L \subset \overline{K}$ and an A_L-module V an *a-pair* for $a \in A$ if the coefficients of the minimal polynomial of $\overline{a}$ acting on $V \otimes_L \overline{K}$ generate L over K.

Let us furthermore call a set $\{(V_1, L_1), \ldots (V_n, L_n)\}$ consisting of a-pairs an *a-decomposition* of an a-pair (V, L) if

(i) $V \otimes_L \overline{K} \cong \bigoplus_{i=1}^n \widetilde{V_i}$ with $\widetilde{V_i} \cong \bigoplus_{\sigma \in G_L/G_{L_i}} \sigma(V_i \otimes_{L_i} \overline{K})$ and

(ii) the minimal polynomial of $\overline{a}$ restricted to V_i is a power of $(X - \lambda_i)$ for some $\lambda_i \in L_i$ for all i and

(iii) the minimal polynomial of $\overline{a}$ restricted to $\widetilde{V_i}$ is coprime to the minimal polynomial of $\overline{a}$ restricted to $\widetilde{V_j}$ whenever $i \neq j$.

The $\widetilde{V_i}$ correspond to the local factors of the L-algebra $\langle a\rangle$ and the $\sigma(V_i \otimes_{L_i} \overline{K})$ to the local factors of the $\overline{K}$-algebra $\langle\overline{a}\rangle$. So the (V_i, L_i) are a choice out of a $G(L_i|L)$-conjugacy class. The third condition above assures that for $i \neq j$ no $(\sigma V_i, \sigma L_i)$ for $\sigma \in G(L_i|L)$ is conjugate to a $(\tau V_j, \tau L_j)$ for any $\tau \in G(L_j|L)$.

An a-decomposition of an a-pair can be computed by the following algorithm.

(4.2.2) Algorithm. *We define the function* `DecomposePair` *as follows.*
input: $(V, L), a$, where (V, L) is an a-pair.
output: A list `output` $[(V_1, L_1), \ldots, (V_n, L_n)]$ *containing an a-decomposition of (V, L).*

1. *Create an empty list* `output`*, which after the running will contain an a-decomposition.*
2. *Compute $f \in L[X]$, the minimal polynomial of $\overline{a}$ restricted to V.*
3. *Factor $f = \prod_{i=1}^{n} p_i^{e_i}$ with $p_i \in L[X]$ pairwise coprime.*
4. *For all i in $\{1, \ldots, n\}$ do*
   1. *Compute $\widetilde{V_i}$ as the kernel of $p_i(\overline{a}|_V)^{e_i}$.*
   2. *Compute L_i, the splitting field over L of p_i.*
   3. *Factor $p_i(X) = \prod_{\sigma \in G_L/G_{L_i}} (X - \sigma\lambda_i)$, for some $\lambda_i \in L_i$.*
   4. *Compute V_i as the kernel of $(\overline{a}|_{\widetilde{V_i}} - \lambda_i)^{e_i}$.*
   5. *Join (V_i, L_i) to the list* `output`*.*
5. *Return* `output` *and stop.*

The decomposition of an A_K-module V corresponding to the local factors of $A_{\overline{K}}$ and keeping track of conjugacy can be computed by the following algorithm, when the $a_1, \ldots, a_n$ in the input generate A.

(4.2.3) Algorithm. *We define the function* `Decompose` *as follows.*
input: $(V, K), [a_1, \ldots, a_n]$ with $[a_1, \ldots, a_n]$ a list of elements of A and (V, K) an a_i-pair for all $i = 1, \ldots, n$.
output: A list `output` $= [(V_1, K_1), \ldots, (V_n, K_n)]$ *consisting of pairs with K_i a finite extension of K and V_i an A_{K_i}-module. See Proposition (4.2.4) for an interpretation.*

1. *Compute* `dec` *as* `DecomposePair(`(V, K)`,` a_1`)`*.*
2. *If $n = 1$, then return* `dec`*.*
3. *Create the empty list* `output`*.*
4. *For all d in* `dec` *do*
   1. *Compute* `dec1` *as* `Decompose(`d`,` $[a_2, \ldots, a_n]$`)`*.*
   2. *Join* `dec1` *to the list* `output`*.*
5. *Return the list* `output` *and stop.*

From Lemma (4.2.1) the following is clear.

(4.2.4) Proposition. *Let A be a commutative finite dimensional K-algebra with generators $a_1, \ldots, a_n$. Let V be an A-module. Suppose that* `Decompose((V,K), [a_1,...,a_n])` *gives the output $\{(V_1, K_1), \ldots, (V_m, K_m)\}$.*

Then $V \otimes_K \overline{K} = \bigoplus_{i=1}^m \widetilde{V_i}$ with $\widetilde{V_i} = \bigoplus_{\sigma \in G_k/G_{K_i}} \sigma V_i$. The $\widetilde{V_i}$ correspond to the local factors of A and the σV_i correspond to the local factors of $A_{\overline{K}}$. □

(4.2.5) Corollary. *We keep the notation from Proposition (4.2.4). If V is a faithful A-module, then the local factors of A are isomorphic to the images of A in $\mathrm{End}(\widetilde{V_i})$. Moreover the local factors of $A_{\overline{K}}$ correspond to the images of $A_{\overline{K}}$ in $\mathrm{End}(\sigma V_i)$.*

4.3. Computing eigenforms of weight $k \geq 2$ over finite fields

[Ash-Stevens] have already noticed that all systems of eigenvalues of modular forms modulo a suitable prime ideal above p also occur in group cohomology. We shall reprove that result in a slightly more precise form using the properties of group cohomology and modular symbols established before. We will also explain how this gives rise to an algorithm for computing eigenforms over finite fields with methods from linear algebra over finite fields.

(4.3.1) Proposition. *(a) Let p be a prime, $k \geq 2$, $N \geq 5$ with $p \nmid N$ integers and $f \in S_k(\Gamma_1(N), \overline{\mathbb{F}_p})$ a normalised eigenform for $2 \leq k$. Then its system of eigenvalues occurs in any of the spaces $H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))$, $\mathcal{CM}_k(\Gamma_1(N), \mathbb{F}_p)$ and $\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}_p)$.*

(b) Let $k \geq 2$ be an integer and $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{F}^*$ be a character with $\epsilon(-1) = (-1)^k$ for $\mathbb{F}|\mathbb{F}_p$ a finite extension. Then the system of eigenvalues of any normalised eigenform $f \in S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})$ occurs in any of the spaces

$$H^1_{\mathrm{par}}(\Gamma_0(N), V^\epsilon_{k-2}(\mathbb{F})), \qquad {}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F})) \otimes_{\mathbb{F}} \mathbb{F}^\epsilon\right),$$
$$\mathcal{CM}_k(\Gamma_0(N), \epsilon, \mathbb{F}), \qquad {}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(\mathcal{CM}_k(\Gamma_1(N), \mathbb{F}) \otimes_{\mathbb{F}} \mathbb{F}^\epsilon\right),$$
$${}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}) \otimes_{\mathbb{F}} \mathbb{F}^\epsilon\right), \qquad {}_{(\mathbb{Z}/N\mathbb{Z})^*}\left(\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}) \otimes_{\mathbb{F}} \mathbb{F}^\epsilon\right).$$

(c) Assume the situation of (a) and that $2 \leq k \leq p+2$. Then all systems of eigenvalues occurring on any of the spaces cited in (a) come from a normalised eigenform $f \in S_k(\Gamma_1(N), \overline{\mathbb{F}_p})$.

(d) Assume the situation of (b), $2 \leq k \leq p+2$ and that $\overline{\Gamma_0(N)}$ does not have any stabiliser of order p for its action on $\mathbb{H}$ (compare Proposition (2.6.2)). If $k = 2$, then we also assume ϵ to be non-trivial. Then all systems of eigenvalues occurring on any of the spaces cited in (b) come from a normalised eigenform $f \in S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})$.

(e) Assume the situation of (a). Suppose there is a system of eigenvalues coming from any of the spaces cited in (a) but not from a modular form. Then for any prime $l \nmid Np$ the Hecke operator $(T_l)^2$ acts on it with eigenvalue $(l+1)^2 l^{k-2}$.

Proof. Part (a) follows from Corollary (3.3.2) and the comparison result Theorem (2.6.1). It is clear that a system of eigenvalues of a modular form with character ϵ also occurs in any of the spaces given in (b).

(c) The statement is an immediate consequence of Corollary (3.3.4) and Theorem (2.6.1).

(d) follows from (c) using Proposition (2.7.10).

(e) We can assume $k > p+2$ by (c). By Corollary (3.3.2) a system of eigenvalues living on the torsion free quotient comes from a modular form. So, a system of eigenvalues as in the assumption must live on

$$H^1(Y_{\overline{\Gamma_1(N)}}, V_{k-2}(\mathbb{Z}_p))[p] \cong H^0(Y_{\overline{\Gamma_1(N)}}, V_{k-2}(\mathbb{F}_p)) \cong V_{k-2}(\mathbb{F}_p)^{\mathrm{SL}_2(\mathbb{F}_p)},$$

where the first isomorphism comes from Proposition (2.4.8). Applying the definition of the Hecke operator T_l for a prime $l \nmid Np$, we see that it acts on the right hand side by sending a polynomial f of degree $k-2$ to $(l+1)\big(\left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right).f\big)(x,y)$. So $(T_l)^2$ acts as $(l+1)^2\big(\left(\begin{smallmatrix} l^2 & 0 \\ 0 & 1 \end{smallmatrix}\right).f\big)(x,y)$. But $\left(\begin{smallmatrix} l^2 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} l & 0 \\ 0 & l \end{smallmatrix}\right)\left(\begin{smallmatrix} l & 0 \\ 0 & l^{-1} \end{smallmatrix}\right)$ and the latter matrix acts trivially, which implies the statement. □

(4.3.2) Remark. *Proposition (4.3.1)(a) and (c) also hold more generally for $k \geq 3$ and $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ a subgroup of finite index whose stabilisers have order invertible modulo p. With these assumptions Proposition (3.3.3) and Corollary (3.3.4) are also true. However, for $k = 2$ there could be lifting problems to characteristic 0 (i.e. Carayol's Lemma does not hold).*

For computational purposes it is essential to have a finite set of generators for the Hecke algebra. This is provided by the following proposition.

(4.3.3) Proposition. *Let $N \geq 1$ and $k \geq 2$ be integers such that $p \nmid N$, $\mathbb{F}|\mathbb{F}_p$ a finite extension and let $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{F}^*$ be a character with $\epsilon(-1) = (-1)^k$. Set*

$$B = \frac{N}{12} \prod_{l|N, l \text{ prime}} (1 + \frac{1}{l}).$$

Let $\mathbb{T}^{(k)}$ be the Hecke algebra for $S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})$. Then the Hecke operators

$$T_1^{(k)}, T_2^{(k)}, \ldots, T_{kB}^{(k)}$$

generate $\mathbb{T}^{(k)}$ as an $\mathbb{F}$-vector space.

The number kB is called the Hecke bound *of $S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})$.*

Proof. This follows from the proof of [EdixJussieu], Proposition 4.2. □

As we have seen that systems of eigenvalues of modular symbols are closely related to those of modular forms, we quickly sketch how to compute them up to Galois conjugacy.

(4.3.4) Algorithm. *We define the function* `Eigenforms` *as follows.*

input: N, k, p, ϵ, where $N \geq 1$, $k \geq 2$ are integers, p is a prime and ϵ is a Dirichlet character of modulus N with values in some finite extension $\mathbb{F}$ of $\mathbb{F}_p$.

output: A list `output` $[(V_1, L_1), \ldots, (V_n, L_n)]$.

1. *Generate the space* `M` *of cuspidal modular symbols for $\Gamma_1(N)$, weight k and character ϵ over $\mathbb{F}$.*
2. *Compute the Hecke bound kB as in Proposition (4.3.3).*
3. *Compute the list* `L = [`T_1`, ..., `T_{kB}`]` *consisting of the listed Hecke operators on* `M`.
4. `output := Decompose((M,`$\mathbb{F}$`),L)`
   The function `Decompose` *was defined in Algorithm (4.2.3). We may replace all e_i in Step 4 of Algorithm (4.2.2) by 1, as we are only interested in systems of eigenvalues.*
5. *Return* `output` *and stop.*

The (V_i, L_i) in the list `output` correspond precisely to the different Galois conjugacy classes of systems of eigenvalues $(\lambda_n)_n$ on the cuspidal modular symbols. That means that the restriction of the Hecke operator T_n to V_i is a scalar matrix with eigenvalue λ_n.

4.4. Computing Hecke algebras of weight $k \geq 2$ over finite fields

We now address the question of computing the Hecke algebra of modular forms over finite fields.

The following theorem is a very satisfactory result, if the weight is smaller than p or equal to $p+1$. In the former case it is mainly due to Edixhoven ([EdixJussieu], Theorem 5.2).

(4.4.1) Theorem. *Let p be a prime and $N \geq 5$, $k \geq 2$ integers such that $p \nmid N$. Suppose $k < p$ or $k = p+1$.*

(a) The Hecke algebra over $\mathbb{F}_p$ of $S_k(\Gamma_1(N), \overline{\mathbb{F}_p})$ can be computed by the Hecke action on any one of $H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F}_p))$, $\mathcal{CM}_k(\Gamma_1(N), \mathbb{F}_p)$ or $\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}_p)$.

(b) Let $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{F}^*$ a character with $\epsilon(-1) = (-1)^k$ for a finite extension $\mathbb{F}$ of $\mathbb{F}_p$. If $k = 2$, then we assume that ϵ is non-trivial. We assume further that $\overline{\Gamma_0(N)}$ does not have any stabiliser of order p for its action on $\mathbb{H}$ (compare Proposition (2.6.2)).

Then the $\mathbb{F}$-Hecke algebra of $S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})$ can be computed by the Hecke action on any one of the spaces

$$H^1_{\mathrm{par}}(\Gamma_0(N), V^{\epsilon}_{k-2}(\mathbb{F})), \qquad {}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(H^1_{\mathrm{par}}(\Gamma_1(N), V_{k-2}(\mathbb{F})) \otimes_{\mathbb{F}} \mathbb{F}^{\epsilon}\right),$$
$$\mathcal{CM}_k(\Gamma_0(N), \epsilon, \mathbb{F}), \qquad {}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(\mathcal{CM}_k(\Gamma_1(N), \mathbb{F}) \otimes_{\mathbb{F}} \mathbb{F}^{\epsilon}\right),$$
$${}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}) \otimes_{\mathbb{F}} \mathbb{F}^{\epsilon}\right), \qquad {}^{(\mathbb{Z}/N\mathbb{Z})^*}\left(\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}) \otimes_{\mathbb{F}} \mathbb{F}^{\epsilon}\right).$$

Proof. In the case where $k < p$ both parts are immediate from Corollary (3.3.7), the comparison result Theorem (2.6.1) and Proposition (2.7.10).

If $k = p + 1$, the result follows from Corollary (3.3.14), Proposition (3.3.20) because all modular forms in weight $p + 1$ are ordinary (see e.g. [EdixWeight], Proposition 3.3). □

Next we cover the weight p case.

(4.4.2) Theorem. *The statements of Theorem (4.4.1) also hold for weight $k = p > 2$ localised at ordinary ($a_p(f) \neq 0$) systems of eigenvalues.*

Proof. This follows as above from Corollary (3.3.14), Proposition (3.3.20), the comparison result Theorem (2.6.1) and Proposition (2.7.10). □

(4.4.3) Lemma. *Let $\overline{\Delta} := (\mathbb{Z}/N\mathbb{Z})^*/\langle -1\rangle$. The Hecke operator T_l for any prime $l \nmid Np$ acts on $\mathcal{E}_2(\Gamma_1(N), \mathbb{F}_p)$, $H_0(\overline{\Delta}, \mathcal{E}_2(\Gamma_1(N), \mathbb{F}_p))$ and $H_1(\overline{\Delta}, \mathcal{E}_2(\Gamma_1(N), \mathbb{F}_p))$ by multiplication by $(l + 1)$.*

Proof. For the boundary space $\bigoplus_{g\in\overline{\Gamma_1(N)}\backslash \mathrm{PSL}_2(\mathbb{Z})/\overline{U}} H^1(\overline{\Gamma_1(N)} \cap g\overline{U}g^{-1}, \mathbb{F}_p)$ the corresponding statement is immediately verified from the definition of the Hecke operator T_l and the fact that the index of $\overline{\Gamma_1(N)} \cap \overline{\Gamma^0(l)} \cap g\overline{U}g^{-1}$ in $\overline{\Gamma_1(N)} \cap g\overline{U}g^{-1}$ is $l + 1$. As it holds on the boundary space, it holds on the Eisenstein subspace, which is Hecke stable. As the $\overline{\Delta}$-action is through the diamond operators, which commute with T_l, the result also follows for the two homology groups listed. □

(4.4.4) Proposition. *Let $N \geq 5$ an integer and $p \nmid N$ a prime. We assume further that $\overline{\Gamma_0(N)}$ does not have any stabiliser of order p for its action on $\mathbb{H}$ (compare Proposition (2.6.2)). Let f be a normalised eigenform in $S_2(\Gamma_0(N), \overline{\mathbb{F}_p})$ corresponding to a maximal ideal $\mathfrak{P}$ of the Hecke algebra $\mathbb{T}$. If the associated Galois representation of f is not Eisenstein (i.e. is not reducible), then the localisation at $\mathfrak{P}$ of $\mathbb{T}$ can be computed by the Hecke action on $\mathcal{CM}_2(\Gamma_0(N), \mathbb{F}_p)_{\mathfrak{P}}$.*

Proof. If $p \neq 2$ we invoke Theorem (4.4.1) and if $p = 2$ [EdixJussieu], Theorem 5.2, in order to obtain the result for $\Gamma_1(N)$ without a character. By Theorem (2.7.8) it suffices to prove that the Hecke action on $H_0(\overline{\Delta}, \mathcal{E}_2(\Gamma_1(N), \mathbb{F}_p))$ and $H_1(\overline{\Delta}, \mathcal{E}_2(\Gamma_1(N), \mathbb{F}_p))$ cannot give rise to an irreducible representation. That, however, is clear by Lemma (4.4.3). □

(4.4.5) Proposition. *Let p be a prime, $N \geq 5$, $k \geq 2$ integers such that $p \nmid N$, and let $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{F}^*$ be a character with $\epsilon(-1) = (-1)^k$ for a finite extension $\mathbb{F}$ of $\mathbb{F}_p$. We assume further that $\overline{\Gamma_0(N)}$ does not have any stabiliser of order p for its action on $\mathbb{H}$ (compare Proposition (2.6.2)). Let $f \in S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})$ a normalised eigenform for $k \geq 2$ with an irreducible Galois representation. Let M be any of the spaces $H^1_{\mathrm{par}}(\Gamma_0(N), V^{\epsilon}_{k-2}(\mathbb{F}))$, $\mathcal{CM}_k(\Gamma_0(N), \epsilon, \mathbb{F})$ or ${}^{(\mathbb{Z}/N\mathbb{Z})^*}\big(\mathcal{CH}_k(\Gamma_1(N), \mathbb{F}) \otimes \mathbb{F}^{\epsilon}\big)$.*

If the local factor at f of $\mathbb{T}(S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p}))$ has the same dimension as the corresponding local factor of $\mathbb{T}(M)$, then these two local algebras are isomorphic.

Proof. As we know that differences occurring in the passage to characters only correspond to reducible Galois representations, it suffices to prove similar statements for $\Gamma := \Gamma_1(N)$ without a character.

We show that there is a surjection of algebras from the local factor of $\mathbb{T}(S_k(\Gamma, \overline{\mathbb{F}_p}))$ to that of $\mathbb{T}(M)$. We have

$$\mathbb{T}_{\mathbb{Z}}(S_k(\Gamma, \mathbb{C})) \otimes_{\mathbb{Z}} \mathbb{F}_p \cong \mathbb{T}_{\mathbb{Z}_p}(H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{Z}_p))/p\text{−torsion}) \otimes_{\mathbb{Z}_p} \mathbb{F}_p.$$

Locally at a prime ideal $\mathfrak{P}$ corresponding to an irreducible representation we have

$$H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{Z}_p))/p\text{−torsion})_{\mathfrak{P}} \cong H^1_{\mathrm{par}}(\Gamma, V_{k-2}(\mathbb{Z}_p)))_{\mathfrak{P}},$$

since the p-torsion part cannot correspond to an irreducible representation due to the calculation in the proof of part (e) of Proposition (4.3.1). Moreover, it is easy to check that irreducible representations cannot live on the possible differences $H^1(\langle\sigma\rangle, \mathrm{Coind}_{\Gamma}^{\mathrm{SL}_2(\mathbb{Z})} V_{k-2}(\mathbb{F}_p))$ and similarly for τ. □

We next sketch an algorithm for computing the local factors of the Hecke algebra of modular symbols up to Galois conjugacy.

(4.4.6) Algorithm. *We define the function* `HeckeAlgebras` *as follows.*

input: N, k, p, ϵ, where $N \geq 1$, $k \geq 2$ are integers, p is a prime and ϵ is a Dirichlet character of modulus N with values in some finite extension $\mathbb{F}$ of $\mathbb{F}_p$.

output: A list `output` *$[(V_1, L_1), \ldots, (V_n, L_n)]$.*

1. *Generate the space* `M` *of cuspidal modular symbols for $\Gamma_1(N)$, weight k and character ϵ over $\mathbb{F}$.*
2. *Compute the Hecke bound kB as in Proposition (4.3.3).*
3. *Compute the list* `L = [`T_1`, ..., `T_{kB}`]` *consisting of the listed Hecke operators on* `M`.
4. `output := Decompose((M,`$\mathbb{F}$`),L)`
   The function `Decompose` *has been defined in Algorithm (4.2.3).*
5. *Return* `output` *and stop.*

The (V_i, L_i) in the list `output` correspond precisely to the different Galois conjugacy classes. That means that the corresponding local Hecke algebra is generated by the restrictions of the Hecke operators to the V_i.

In order to obtain proved results if the conditions of Theorems (4.4.1) and (4.4.2) do not apply, we must compute with cuspidal modular symbols over $\mathbb{Q}$ and choose a lattice. We may then work with the reduction modulo p of the Hecke operators written with respect to a lattice basis. Note that this method in the primitive form given only applies to situations when the character takes values in $\{\pm 1\}$. This approach works as any lattice gives rise to an isomorphic Hecke algebra. However, working over $\mathbb{Q}$, choosing a lattice and computing

Hecke operators with integral coefficients is very slow. Moreover, if we work with non-free groups such as $\Gamma_0(N)$ then we only get those forms that are reductions mod p of holomorphic modular forms and possibly not all Katz forms. One of the advantages of working with all the torsion is that we get Katz modular forms, whenever Theorems (4.4.1) or (4.4.2) apply. E.g. with $k = 2$ and $p = 3$, mod 3 modular symbols also compute those mod 3 eigenforms that cannot be lifted to characteristic 0 with a character of the same order.

4.5. Embedding weight one into weight p

In this section we describe how weight one and weight p modular forms are related. Recall that we are working throughout with Katz modular forms, which becomes really essential in this section. All ideas are taken from [EdixJussieu] (Section 4) and have also been described in [W-App]. For more details the reader is referred there.

Let $\mathbb{F}$ be a finite field of prime characteristic p of $\overline{\mathbb{F}_p}$ and fix a level $N \geq 1$ with $p \nmid N$ and a character $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{F}^*$ with $\epsilon(-1) = (-1)^k$. We have two injections of $\mathbb{F}$-vector spaces

$$F, A : S_1(\Gamma_1(N), \epsilon, \mathbb{F}) \to S_p(\Gamma_1(N), \epsilon, \mathbb{F}),$$

given on q-expansions by $a_n(Ag) = a_n(g)$ and $a_n(Fg) = a_{n/p}(g)$ (with $a_n(Fg) = 0$ if $p \nmid n$), that are compatible with all Hecke operators T_l for primes $l \neq p$. The former comes from the *Frobenius* and the latter is multiplication by the *Hasse invariant*. One has $T_p^{(p)} F = A$ and $AT_p^{(1)} = T_p^{(p)} A + \epsilon(p)F$, where we have indicated the weight as a superscript (see e.g. [EdixJussieu], Equation 4.1.2).

The key to an effective description of the image of F is the following theorem by Katz, which is the main theorem of [KatzDerivation].

(4.5.1) Theorem. (Katz) *Let k be an integer.*

(1) There exists a homomorphism

$$A\Theta : S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}}) \to S_{k+p+1}(\Gamma_1(N), \epsilon, \overline{\mathbb{F}}),$$

whose effect on q-expansions is $q\frac{d}{dq}$ (i.e. $a_n(A\Theta f) = na_n(f)$), whence it is called a derivation.

(2) Suppose $p \nmid k$. If $f \in S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}})$ does not come from a lower weight, then $A\Theta f$ has weight $k + p + 1$, and does not come from a lower weight. In particular, $A\Theta f \neq 0$.

(3) If $f \in S_{pk}(\Gamma_1(N), \epsilon, \overline{\mathbb{F}})$ and $A\Theta f = 0$, then $f = h^p$ for a unique $h \in S_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}})$.

Let $\mathbb{T}^{(k)}$ be the Hecke algebra over $\mathbb{F}$ of weight k for a fixed level N and a fixed character ϵ. We will also indicate the weight of Hecke operators by superscripts. We denote by $A^{(p)}$ the $\mathbb{F}_p$-subalgebra of $\mathbb{T}^{(p)}$ generated by all Hecke operators $T_n^{(p)}$ for $p \nmid n$.

(4.5.2) Proposition. *(a) There is a homomorphism Θ, also called a* derivation*, given on q-expansions by $a_n(\Theta f) = n a_n(f)$ such that the sequence*

$$0 \to S_1(\Gamma_1(N), \epsilon, \mathbb{F}) \xrightarrow{F} S_p(\Gamma_1(N), \epsilon, \mathbb{F}) \xrightarrow{\Theta} S_{p+2}(\Gamma_1(N), \epsilon, \mathbb{F})$$

is exact.

(b) Suppose $f \in S_1(\Gamma_1(N), \epsilon, \mathbb{F})$ such that $a_n(f) = 0$ for all n with $p \nmid n$. Then $f = 0$. In particular $AS_1(\Gamma_1(N), \epsilon, \mathbb{F}) \cap FS_1(\Gamma_1(N), \epsilon, \mathbb{F}) = 0$.

(c) The Hecke algebra $\mathbb{T}^{(1)}$ in weight one can be generated by all $T_l^{(1)}$, where l runs through the primes different from p.

(d) The weight one Hecke algebra $\mathbb{T}^{(1)}$ is the algebra generated by the $A^{(p)}$-action on the module $\mathbb{T}^{(p)}/A^{(p)}$.

Proof. (a) Theorem (4.5.1) (3) gives the exact sequence

$$0 \to S_1(\Gamma_1(N), \epsilon, \mathbb{F}) \xrightarrow{F} S_p(\Gamma_1(N), \epsilon, \mathbb{F}) \xrightarrow{A\Theta} S_{2p+1}(\Gamma_1(N), \epsilon, \mathbb{F})$$

by taking Galois invariants. However, as explained in [EdixJussieu], Section 4, the image $A\Theta S_p(\Gamma_1(N), \epsilon, \mathbb{F})$ in weight $2p+1$ can be divided by the Hasse invariant, whence the weight is as claimed.

(b) The condition implies by looking at q-expansions that $A\Theta f = 0$, whence by Theorem (4.5.1) (3) f comes from a lower weight than 1, but below there is just the 0-form (see also [EdixJussieu], Proposition 4.4).

(c) It is enough to show that $T_p^{(1)}$ is linearly dependent on the span of all $T_n^{(1)}$ for $p \nmid n$. If it were not, then there would be a modular form of weight 1 satisfying $a_n(f) = 0$ for $p \nmid n$, but $a_p(f) \neq 0$, contradicting (b).

(d) Dualising the exact sequence in (a) yields that $\mathbb{T}^{(p)}/A^{(p)}$ and $\mathbb{T}^{(1)}$ are isomorphic as $A^{(p)}$-modules, which implies the claim. □

(4.5.3) Proposition. *The $\overline{\mathbb{F}}$-algebra $A^{(p)}$ defined above can already be generated as an $\overline{\mathbb{F}}$-vector space by the set*

$$\{\ T_n^{(p)}\ \mid\ p \nmid n,\ n \leq (p+2)B\ \},$$

where B is the number from Proposition (4.3.3).

Proof. Assume that some $T_m^{(p)}$ for $m > (p+2)B$ and $p \nmid m$ is linearly independent of the operators in the set of the assertion. This means that there is a modular form $f \in S_p(\Gamma_1(N), \epsilon, \overline{\mathbb{F}})$ satisfying $a_n(f) = 0$ for all $n \leq (p+2)B$, but $a_m(f) \neq 0$. One gets $a_n(\Theta f) = 0$ for all $n \leq (p+2)B$, but $a_m(\Theta f) \neq 0$. This contradicts Proposition (4.3.3). □

These two propositions provide us with an effective method for computing the Hecke algebra in weight one, once we dispose of a faithful module for the Hecke algebra in weight p.

If we are only interested in forms of weight one, we would like to be able to throw away parts that cannot come from weight one. The following considerations will also enable us in certain cases to compute weight one eigenforms without computing all the Hecke algebra.

(4.5.4) Proposition. *Let $V \subset S_p(\Gamma_1(N), \epsilon, \overline{\mathbb{F}})$ be the eigenspace of a system of eigenvalues for the operators $T_l^{(p)}$ for all primes $l \neq p$*

If the system of eigenvalues does not come from a weight one form, then V is at most 1-dimensional. Conversely, if there is a normalised weight one eigenform g with that system of eigenvalues for $T_l^{(1)}$ for all primes $l \neq p$, then $V = \langle Ag, Fg \rangle$ and that space is 2-dimensional. On it $T_p^{(p)}$ acts with eigenvalues u and $\epsilon(p)u^{-1}$ satisfying $u+\epsilon(p)u^{-1} = a_p(g)$. In particular, the eigenforms in weight p which from weight one are ordinary.

Proof. If V is at least 2-dimensional, then we can choose a normalised eigenform f for all operators and we then have $V = \mathbb{F}f \oplus \{h \mid a_n(h) = 0 \ \forall p \nmid n\}$. As a form h in the right summand is annihilated by Θ, it is equal to Fg for some form g of weight one by Proposition (4.5.2) (a). By Part (b) of that proposition we know that $\langle Ag, Fg \rangle$ is 2-dimensional. If V were more than 2-dimensional, then there would be two different modular forms in weight 1, which are eigenforms for all $T_l^{(1)}$ with $l \neq p$. This, however, contradicts Part (c).

Any normalised eigenform $f \in V$ for all Hecke operators in weight p has to be of the form $Ag + \mu Fg$ for some $\mu \in \overline{\mathbb{F}}$. The eigenvalue of $T_p^{(p)}$ on f is the p-th coefficient, hence $u = a_p(g) + \mu$, as $a_p(Fg) = a_1(g) = 1$. Now we have

$$\begin{aligned}(a_p(g)+\mu)(Ag+\mu Fg) = \quad & T_p^{(p)}(Ag+\mu Fg) = & T_p^{(p)}Ag + \mu Ag \\ = \quad & AT_p^{(1)}g - \epsilon(p)Fg + \mu Ag = & (a_p(g)+\mu)Ag - \epsilon(p)Fg,\end{aligned}$$

which implies $-\epsilon(p) = (a_p(g)+\mu)\mu = u^2 - ua_p(g)$ by looking at the p-th coefficient. From this one obtains the claim on u. □

(4.5.5) Corollary. *Let $N \geq 5$ an integer not disible by the prime p. The Hecke algebra of $S_1(\Gamma_1(N), \mathbb{F}_p)$ can be computed using modular symbols over $\mathbb{F}_p$.*

Proof. Due to the ordinarity (Proposition (4.5.4)), this follows from Theorem (4.4.2) and Proposition (4.5.2)(d). □

4.6. Computing Hecke algebras of weight one over finite fields

If one is only interested in eigenforms, one can use Algorithm (4.3.4) for weight p with mod p modular symbols and look for pairs of eigenforms differing only at p. These correspond to weight one eigenforms by Proposition (4.5.4). However, a weight p eigenform f with

$a_p(f)^2 = \epsilon(p)$ might or might not correspond to weight one. Here an extra check will be needed.

We next describe the function `SystemsOfEigenvalues` implemented in the Magma package `Weight1.mg`.

(4.6.1) Algorithm. *We define the function* `SystemsOfEigenvalues` *as follows.*

input: N, p, ϵ *with* N *a positive integer,* p *a prime and* ϵ *a Dirichlet character over a field extension* $\mathbb{F}$ *of* $\mathbb{F}_p$.

option: `Wt1APriori` $\in \{$ `true, false` $\}$.

output: A decomposition of the cuspidal modular symbols for $\Gamma_1(N)$ *of weight* p *and character* ϵ *corresponding to the conjugacy classes of local factors of the algebra* $A^{(p)}$. *If the option* `Wt1APriori` *is set, local factors that cannot correspond to weight one forms by Proposition (4.5.4) are discarded.*

1. *Generate the space* `M` *of cuspidal modular symbols for* $\Gamma_1(N)$, *weight* p *and character* ϵ *over* $\mathbb{F}$.
2. *Compute* $b := (p+2)B$ *as in Proposition (4.5.3).*
3. *Compute the list* `L` *consisting of the Hecke operators* T_n *acting on* `M` *for* $p \nmid n$ *and* $1 \leq n \leq b$.
4. *If* `not Weight1APriori`, *call* `output := Decompose` *with the list* L *and the pair* $(M, \mathbb{F})$. *Return* `output` *and stop. The function* `Decompose` *was defined in Algorithm (4.2.3).*
5. *If* `Weight1APriori`, *then proceed as follows.*
   1. *Compute the operator* T_p *acting on* `M`.
   2. *Compute the minimal polynomial* $F_p \in \mathbb{F}[X]$ *of* T_p *and factor it* $\prod_{i=1}^{n} p_i(X)^{e_i}$ *over* $\mathbb{F}[X]$ *into coprime prime powers.*
   3. *Create a set* S *of prime factors as follows.*
      If $p_i(X) = X + a$ *with* $a^2 = \overline{\epsilon(p)}$, *then join* $p_i(X)^{e_i}$ *to the set* S.
      If $p_i(X)$ *satisfies* $p_i(a) = 0 \Rightarrow p_i(\overline{\epsilon(p)}a^{-1}) = 0$, *then join* $p_i(X)^{e_i}$ *to the set* S.
      If for a pair $i \neq j$ *one has* $p_i(a) = 0 \Rightarrow p_j(\overline{\epsilon(p)}a^{-1}) = 0$, *then join* $p_i(X)^{\mathrm{lcm}(e_i,e_j)}$ *to the set* S *and discard* $p_j(X)$.
      Discard all other $p_i(X)$.
   4. *Create an empty list* `output`.
   5. *For each* $p(X)^e$ *in* S *do:*
      1. *Compute the kernel* W *of* $p(T_p)^e$ *acting on* `M`.
      2. *Create a list* L' *by restricting the entries of* L *to* W.
      3. *Join* `Decomposition` $((W, \mathbb{F}), L')$ *to the list* `output`.
   6. *Return* `output` *and stop.*

One now obtains the local weight one Hecke algebras for the given level and character as follows. Let V_i be one of the spaces in the output of `SystemsOfEigenvalues`. Then the quotient

$$\langle T_n|_{V_i} \mid 1 \leq n \leq pB\rangle / \langle T_n|_{V_i} \mid 1 \leq n \leq (p+2)B, p \nmid n\rangle$$

is an $A^{(p)}$-module. Whenever the space of cuspidal modular symbols used in the function `SystemsOfEigenvalues` is a faithful $\mathbb{T}^{(p)}$-module, the algebra generated by $T_l^{(p)}$ on that quotient for $l \neq p$ and $1 \leq l \leq B$ is the local factor of the Hecke algebra of weight one that we were looking for.

4.7. Universal q-expansions

In [MerelUniversal] Merel has among other things established a "universal Fourier expansion" of holomorphic modular forms in terms of Hecke operators acting on modular symbols. We sketch how a generalisation of Merel's result can be deduced from the theory developed here.

The *plus-space*, denoted by the superscript $^+$, is the subspace fixed by the action of the matrix $\eta := \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (supposing that there is such an action).

(4.7.1) Proposition. *Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a subgroup of finite index. Suppose $\eta\Gamma\eta = \Gamma$ and that for some $x \in \mathcal{CM}_k(\Gamma, R)^+$ the homomorphism*

$$\mathbb{T} \to \mathcal{CM}_k(\Gamma, R)^+, \quad T \mapsto Tx$$

is an isomorphism, i.e. that $\mathcal{CM}_k(\Gamma, R)^+$ is a free $\mathbb{T}$-module of rank 1. Then we have a universal q-expansion*, i.e. an isomorphism*

$$\mathrm{Hom}(\mathcal{CM}_k(\Gamma, R)^+, R) \to S_k(\Gamma, R), \quad \phi \mapsto \sum_{n \geq 1} \phi(T_n x) q^n.$$

Proof. We have the isomorphism

$$\mathrm{Hom}(\mathbb{T}, R) \to S_k(\Gamma, R), \quad \psi \mapsto \sum_{n \geq 1} \psi(T_n) q^n,$$

which we only need to combine with the isomorphism between $\mathbb{T}$ and $\mathcal{CM}_k(\Gamma, R)^+$ to obtain the proposition. □

From a version of the Eichler-Shimura Theorem involving the plus-space it follows that the condition in Proposition (4.7.1) is satisfied, when $R = \mathbb{C}$. In [EPW], p. 30, an isomorphism between $\mathbb{T}_{\mathfrak{P}}$ and a space isomorphic to $\mathcal{CM}_k(\Gamma_1(N), \mathbb{Z}_p)^+_{\mathfrak{P}}$ locally at p-ordinary and p-distinguished primes $\mathfrak{P}$ of the Hecke algebra is derived from a fundamental theorem by Wiles ([Wiles], Theorem 2.1). Hence, also in that situation there is a universal Fourier expansion.

Chapter V

Some Computational Results

This chapter gives an overview over some computations that were carried out using the algorithms presented in Chapter IV. Moreover, we give some motivation why these computations are useful.

5.1. Weight one modular forms over $\overline{\mathbb{F}_2}$ for $\Gamma_0(N)$

Mestre's computations

The first computer calculations of weight one modular forms known to the author were carried out by Jean-François Mestre and written down in a letter to Serre from October 1987. The letter has appeared as Appendix A of [EdixJussieu]. We have verified the computations and reported on them in Appendix B [W-App] of loc. cit.

Further computations

All modular forms of weight 1 and 2 for $\Gamma_0(N)$ for odd N in the range from 11 to 3445 have been computed with the Magma package `Weight1.mg`. Up to Galois conjugacy we found 2998 cuspidal Katz eigenforms of weight 1 and 14009 systems of eigenvalues of weight 2. Among the latter there might be some that do not correspond to cusp forms (see Chapter IV). For each of them we computed the field $\mathbb{F}$ generated by the coefficients, the local factor of the corresponding Hecke algebra and a lower bound for the image of the associated Galois representation (an upper bound is provided by $\mathrm{SL}_2(\mathbb{F})$).

We found the following distribution of the degree of the coefficient field. In the tables we count all eigenforms (resp. systems of eigenvalues), not only up to Galois conjugacy.

Weight 1:

Degree	1	2	3	4	5	8	10
# forms	1417	1102	1200	972	535	96	320
Percentage	21.1	16.4	17.9	14.5	7.9	1.4	4.7

The biggest degree occurring is 29.

Weight 2:

Degree	1	2	3	4	5	8	10
# forms	3765	5036	5115	6036	3160	2976	3430
Percentage	9.8	13.1	13.3	15.7	8.2	7.7	8.9

The biggest degree occurring is 127.

Non-liftable Katz forms of weight one

Eigenforms of weight one whose associated Galois representation has image equal to some $\mathrm{SL}_2(\mathbb{F}_{2^r})$ with $r \geq 3$ cannot be lifted to holomorphic weight one forms, because $\mathrm{PGL}_2(\mathbb{C})$ does not have a finite subgroup having $\mathrm{SL}_2(\mathbb{F}_{2^r})$ with $r \geq 3$ as a quotient. The first such form was found by Mestre. More calculations have been carried out by Lloyd Kilford, Edray Goins and the author for forms over $\overline{\mathbb{F}_2}$.

Level	Group	History	Level	Group	History
1429	$\mathrm{SL}_2(\mathbb{F}_8)$	Mestre	2879	$\mathrm{SL}_2(\mathbb{F}_8)$	W.
1567	$\mathrm{SL}_2(\mathbb{F}_8)$	Mestre	3271	$\mathrm{SL}_2(\mathbb{F}_{512})$	W.
1613	$\mathrm{SL}_2(\mathbb{F}_8)$	Mestre	3517	$\mathrm{SL}_2(\mathbb{F}_8)$	Kilford
1693	$\mathrm{SL}_2(\mathbb{F}_8)$	Mestre	3709	$\mathrm{SL}_2(\mathbb{F}_8)$	Kilford
1997	$\mathrm{SL}_2(\mathbb{F}_8)$	W.	4817	$\mathrm{SL}_2(\mathbb{F}_8)$	Kilford
2017	$\mathrm{SL}_2(\mathbb{F}_8)$	W.	4889	$\mathrm{SL}_2(\mathbb{F}_8)$	Kilford
2089	$\mathrm{SL}_2(\mathbb{F}_8)$	W.	6133	$\mathrm{SL}_2(\mathbb{F}_{1024})$	Kilford
2633	$\mathrm{SL}_2(\mathbb{F}_{32})$	Kilford	6709	$\mathrm{SL}_2(\mathbb{F}_{16})$	Kilford
2647	$\mathrm{SL}_2(\mathbb{F}_{16})$	W.	7237	$\mathrm{SL}_2(\mathbb{F}_8)$	Kilford
2767	$\mathrm{SL}_2(\mathbb{F}_{64})$	W.			

Non-Gorenstein Hecke algebras

The first non-Gorenstein Hecke algebras were found by Lloyd Kilford ([Kilford]) for $\Gamma_0(p)$ in weight 2 over $\overline{\mathbb{F}_2}$ with $p \in \{431, 503, 2089\}$. All corresponding modular forms are dihedral, the associated Galois representation is only ramified in p and also occurs for a weight one form. However, the local Hecke algebra in weight one is Gorenstein in all the cases. These non-Gorenstein forms of weight 2 should be considered as old forms due to the presence of the weight one forms, which can be embedded into weight 2 in two different ways as explained in Chapter IV. We did not find any weight one Hecke algebra which is non-Gorenstein and whose system of eigenvalues does not already live in a strictly lower level.

Realisation of $\mathrm{SL}_2(\mathbb{F}_{2^r})$ as Galois groups over $\mathbb{Q}$

Weight 2 eigenforms for $\Gamma_0(N)$ over $\overline{\mathbb{F}_2}$ give rise to a Galois representation, whose image is a finite subgroup in $\mathrm{SL}_2(\overline{\mathbb{F}_2})$. This leads to a realisation of the occurring images as Galois groups over $\mathbb{Q}$. It seems only to have been known previously that $\mathrm{SL}_2(\mathbb{F}_{2^r})$ occurs for $r \leq 16$, again due to computations by Mestre (see [SerreGalois], p. 53). The abstract realisations that we obtain in that way only ramify in N and usually also 2.

Up to conjugation all subgroups of $\mathrm{SL}_2(\mathbb{F}_{2^r})$ are

- $\mathrm{SL}_2(\mathbb{F}_{2^s})$ with $s \mid r$,
- dihedral D_n with $n \mid 2^r - 1$ or $n \mid 2^r + 1$,
- cyclic C_n with $n \mid 2^r - 1$ or $n \mid 2^r + 1$,
- subgroups of the upper triangular matrices.

In square-free levels only D_n or $\mathrm{SL}_2(\mathbb{F}_{2^s})$ are possible by Proposition (1.5.2), unless the representation is trivial. It is easy to determine a lower bound for the image, as conjugacy classes can be distinguished by their traces. The traces of all Frobenius elements for primes $p \nmid 2N$ and p less than the Hecke bound are used in the program. Although it is unlikely, it is not excluded that not enough conjugacy classes are hit to obtain the whole image. Hence, we can only talk about lower bounds.

We have 2506 groups of the type $\mathrm{SL}_2(\mathbb{F}_{2^s})$ with $s > 1$ in a table. For instance, there is a Galois extension $K|\mathbb{Q}$ with group $\mathrm{SL}_2(\mathbb{F}_{2^{127}})$, which ramifies only in 3313 and (probably) 2.

There are relatively few forms with a full $\mathrm{SL}_2(\mathbb{F}_{2^r})$ as image in weight 1:

Degree	all	2	3	4	5	6
# forms	1581	551	400	243	107	99
Percentage SL_2	2.8	5.9	2.0	0.4	1.0	1.0

The central row indicates the number of eigenforms for that degree and the lower row contains the percentage of SL_2-forms. We point out that we count $D_3 = \mathrm{SL}_2(\mathbb{F}_2)$ as dihedral and not among the SL_2.

In weight 2 there are many more forms with a full $\mathrm{SL}_2(\mathbb{F}_{2^r})$ as image:

Degree	all	2-4	5-7	8-10	11-13	14-16
# forms	10244	5732	1746	907	511	262
Percentage SL_2	47	33	60	57	61	74

If we restrict only to prime levels, in weight 2 the percentage is roughly 60%.

The computations have yielded the following result.

(5.1.1) Theorem. *All groups $\mathrm{SL}_2(\mathbb{F}_{2^r})$ occur as Galois groups over $\mathbb{Q}$ for r from 1 up to 77.*

5.2. Icosahedral Galois representations and Serre's conjecture

Shepherd-Barron and Taylor ([ShBT]) have proved that any irreducible Galois representation

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_4)$$

which is unramified at 3 and 5 is modular. In view of Serre's conjecture only the level and the weight question for such representations are hence not fully answered, when the representation is exceptional (see Section 1.1).

From the point of view of Maaß forms these representations are interesting, when the number field K cut out by the projectivisation $\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_4) \twoheadrightarrow \mathrm{PGL}_2(\mathbb{F}_4)$ is totally real. For then there exists an even representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{C})$ whose projectivisation cuts out the same field K. Such representations are generally conjectured to come from certain Maaß forms. If indeed this is true, then the corresponding Maaß form has coefficients in the algebraic integers and should reduce modulo a prime above 2 to a Katz modular form over $\overline{\mathbb{F}_2}$ of weight one. It would be interesting to carry out some computations of Maaß forms in that direction.

In [Doud-Moore] Doud and Moore use a targeted Hunter search to obtain a complete list of all even icosahedral complex Galois representations of prime conductor $p < 10000$. Moreover, they supply polynomials generating the corresponding A_5-extensions of $\mathbb{Q}$. None of the representations in that range is exceptional, whence the cited result by Shepherd-Barron and Taylor together with level and weight lowering gives the existence of the corresponding modular forms for $\Gamma_0(p)$ and weight 1 over $\overline{\mathbb{F}_2}$. We have verified some of the cases also computationally.

Another list of polynomials generating totally real A_5-extensions of $\mathbb{Q}$ was supplied by Jürgen Klüners. It contains a totally real A_5-extension ramifying only in the prime $p = 8311$ such that the associated representation $G_{\mathbb{Q}} \to \mathrm{SL}_2(\mathbb{F}_4)$ has conductor p, but there does not exist a complex icosahedral representation of prime conductor. That representation is not exceptional and the associated weight one form over $\overline{\mathbb{F}_2}$ was found. The first exceptional case of a totally real A_5-representation $G_{\mathbb{Q}} \to \mathrm{PGL}_2(\mathbb{C})$ coming from Klüners' table has prime conductor $p = 10267$. As predicted by Serre's conjecture there is a weight one form in level 10267 whose coefficients match with the traces of the Frobenius elements Frob_p for primes $3 \leq p \leq 3413$. This is the first test known to the author of Serre's conjecture for a totally real A_5-extension in a case not covered by level lowering and weight lowering.

Bibliography

[Alperin] J. L. Alperin. *Local representation theory*, Cambridge Univ. Press, Cambridge, 1986.

[Ash-Stevens] A. Ash and G. Stevens. *Modular forms in characteristic l and special values of their L-functions*, Duke Math. J. **53** (1986), no. 3, 849–868.

[BLR] S. Bosch, W. Lütkebohmert and M. Raynaud. *Néron Models*, Springer Verlag, Berlin, 1990.

[Brown] K. S. Brown. *Cohomology of groups*, Springer, New York, 1982.

[Buzzard] K. Buzzard. *On level lowering for mod 2 representations.* Mathematical Research Letters **7** (2000), 95–110.

[Cremona] J. E. Cremona. *Algorithms for modular elliptic curves.* Second edition. Cambridge University Press, Cambridge, 1997.

[Deligne-Rapoport] P. Deligne and M. Rapoport. *Les schémas de modules de courbes elliptiques*, in *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, 143–316. Lecture Notes in Math. **349**, Springer, Berlin, 1973.

[Diamond] F. Diamond. *The refined conjecture of Serre.* Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993), Internat. Press, Cambridge, MA, 1995, 22–37.

[DiamondIm] F. Diamond and J. Im. *Modular forms and modular curves*, in *Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994)*, 39–133, Amer. Math. Soc., Providence, RI, 1995.

[Dold] A. Dold. *Lectures on algebraic topology*, Springer, New York, 1972.

[Doud-Moore] D. Doud and M. W. Moore. *Even Icosahedral Galois Representations of Prime Conductor*. Preprint, 2004.

[EdixBoston] S. J. Edixhoven. *Serre's Conjecture*, in *Modular forms and Fermat's last theorem (Boston, MA, 1995)*, 209–242, Springer, New York, 1997.

[EdixWeight] S. J. Edixhoven. *The weight in Serre's conjectures on modular forms*, Invent. Math. **109** (1992), no. 3, 563–594.

[EdixJussieu] S. J. Edixhoven. *Comparison of integral structures on spaces of modular forms of weight two, and computation of spaces of forms mod* 2 *of weight one*, Accepted for publication in the Journal de l'Institut de Mathématiques de Jussieu. `http://arxiv.org/abs/math.NT/0312019`

[EPW] M. Emerton, R. Pollack and T. Weston. *Variation of Iwasawa Invariants in Hida Families*, Preprint, 2004.

[FaltingsJordan] G. Faltings, B. W. Jordan. *Crystalline Cohomology and* $\mathrm{GL}_2(\mathbb{Q})$, Israel J. Math. **90** (1995), no. 1–3, 1–66.

[Gross] B. H. Gross. *A tameness criterion for Galois representations associated to modular forms (mod* p*)*, Duke Math. J. **61** (1990), no. 2, 445–517.

[Hartshorne] R. Hartshorne. *Algebraic geometry*, Springer, New York, 1977.

[Herremans] A. Herremans. *A combinatorial interpretation of Serre's conjecture on modular Galois representations*, Ann. Inst. Fourier (Grenoble) **53** (2003), no. 5, 1287–1321.

[Katz] N. M. Katz. p*-adic properties of modular schemes and modular forms.* Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972). Lecture Notes in Mathematics, Vol. 350, Springer, Berlin, 1973, 69–190.

[KatzDerivation] N. M. Katz. *A result on modular forms in characteristic* p. Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976), 53–61. Lecture Notes in Math., Vol. **601**, Springer, Berlin, 1977.

[Katz-Mazur] N. M. Katz and B. Mazur. *Arithmetic moduli of elliptic curves*, Ann. of Math. Stud. **108**, Princeton Univ. Press, Princeton, NJ, 1985.

[Kilford] L. J. P. Kilford. *Some non-Gorenstein Hecke algebras attached to spaces of modular forms.* J. Number Theory **97** (2002), no. 1, 157–164.

[Livné] R. Livné. *On the conductors of mod l Galois representations coming from modular forms.* J. of Number Theory **31** (1989), 133–141.

[Manin] Manin, Ju. I. *Parabolic points and zeta functions of modular curves.* Izv. Akad. Nauk SSSR Ser. Mat. **36** (1972), 19–66.

[Martin] F. Martin. *Périodes de formes modulaires de poids* 1, Thèse de doctorat, Université Paris 7, 2001.

[MerelHecke] L. Merel. *Opérateurs de Hecke pour* $\Gamma_0(N)$ *et fractions continues*, Ann. Inst. Fourier (Grenoble) **41** (1991), no. 3, 519–537.

[MerelUniversal] L. Merel. *Universal Fourier expansions of modular forms*, in *On Artin's conjecture for odd* 2*-dimensional representations*, 59–94, Lecture Notes in Math., 1585, Springer, Berlin, 1994.

[Milne] J.S. Milne. *Étale cohomology*. Princeton Mathematical Series, **33**. Princeton University Press, Princeton, N.J., 1980.

[NSW] J. Neukirch, A. Schmidt and K. Wingberg. *Cohomology of Number Fields*, Grundlehren der mathematischen Wissenschaften **323**, Springer, Berlin, 2000.

[R-T] D. E. Rohrlich and J. B. Tunnell. *An elementary case of Serre's conjecture.* Pacific Journal of Mathematics **181**, No. 3 (1997), 299–309.

[Serre1] J.-P. Serre. *Sur les représentations de degré* 2 *de* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Duke Mathematical Journal **54**, No. 1 (1987), 179–230.

[Serre2] J.-P. Serre. *Modular forms of weight one and Galois representations*. Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), Academic Press, London, 1977, 193–268.

[SerreGalois] J.-P. Serre. *Topics in Galois theory*. Research Notes in Mathematics, 1. Jones and Bartlett Publishers, Boston, MA, 1992.

[ShBT] N. I. Shepherd-Barron and R. Taylor. $\mathrm{mod}2$ *and* $\mathrm{mod}5$ *icosahedral representations*. J. Amer. Math. Soc. **10** (1997), no. 2, 283–298.

[Shimura] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Forms*. Princeton University Press, 1994.

[Šokurov] V. V. Šokurov. *Shimura integrals of cusp forms.* Izv. Akad. Nauk SSSR Ser. Mat. **44** (1980), no. 3, 670–718, 720.

[SteinThesis] W. A. Stein. *Explicit approaches to modular abelian varieties*, UC Berkeley, Ph.D. thesis.

[SteinVerrill] W. A. Stein and H. A. Verrill. *Transportable Modular Symbols*, LMS J. Comput. Math. **4** (2001), 170–181.

[Toen] B. Toen. *K-Théorie et cohomologie des champs algébriques: théorèmes de Riemann-Roch, $\mathcal{D}$-modules et théorèmes "GAGA"*. Thèse de doctorat, Université Paul Sabatier, Toulouse, 1999.

[Weibel] C. A. Weibel. *An introduction to homological algebra*, Cambridge Univ. Press, Cambridge, 1994.

[W-App] G. Wiese. *Computing Hecke algebras of weight* 1 *in Magma.* Appendix B of [EdixJussieu].

[W-Dih] G. Wiese. *Dihedral Galois Representations and Katz Modular Forms*. Documenta Mathematica, Vol. 9 (2004), 123–133.

[Wiles] A. J. Wiles. *Modular Elliptic Curves and Fermat's Last Theorem*, Ann. of Math. (2) 141 (1995), no. 3, 443–551.

Index

Acknowledgements

This thesis would not have been written without the help and support of many people and institutions. I gratefully acknowledge financial support by my parents, l'association Égide, the European Research Training Networks "Arithmetic Algebraic Geometry" (AAG) and "European Algebraic Geometry Education and Research" (EAGER), het Leidsch Universiteits Fonds, het Stieltjes Instituut and het Mathematisch Instituut van de Universiteit Leiden.

Je tiens à remercier l'équipe de géométrie algébrique de l'Université de Rennes 1 où j'ai commencé ma thèse, en particulier Frédéric, Daniel et Philippe, ces deux derniers m'ayant appris deux façons totalement différentes de parler français.

Ich möchte mich auch bei Christopher Deninger und dem SFB 478 für das Zurverfügungstellen eines ständigen Arbeitsplatzes in Münster bedanken.

Het was een groot plezier aan het Mathematisch Instituut van de Universiteit Leiden werkzaam te zijn. De vriendelijke en vriendschappelijke atmosfeer tussen professoren, medewerkers, studenten en "ondersteunend personeel" is een bijzondere vermelding waard. Dank u wel allemaal!

I would also like to thank my friends and colleagues for discussions, seminars and also just sharing lunch together. Among many, I would like to mention Bas, Bill, Christiaan, Christophe, Clemens, Denis, Jeanine, Joost, Julia, Jürgen, Lara, Lenny, Markus, Michael, Reinier, Remke, Robert, Roman, Tobi, Tobias and Willem Jan.

A special thanks is due to Luca for sharing the house, the problems and much more.

Bijzondere dank verdient ook mijn kamergenoot Theo voor een leuke tijd samen, al zijn mooie uitleg over algebraïsche meetkunde en Latex en het zetten van ontelbaar veel kopjes koffie.

Heel erg bedanken wil ik ook mijn paranimfen Jan en Ron, en ook alle anderen van DOCOS tafeltennis voor leuke en gezellige tafeltennisavonden.

Schließlich möchte ich mich ganz herzlich bei meinen Eltern bedanken für ihre Unterstützung in allen möglichen Hinsichten während meines ganzen Lebens.

Le plus grand merci est réservé à ma femme Cécile pour son amour ferme et son soutien pendant tout ce long temps de séparation et de voyages hebdomadaires.

Samenvatting

In deze samenvatting zal ik eerst een zo begrijpelijk mogelijke, elementaire inleiding geven tot het gebied van de wiskunde waarover mijn proefschrift gaat. Daarna volgt een overzicht van de inhoud van deze dissertatie.

Modulaire vormen spelen al sinds hun introductie in de 19de eeuw een belangrijke rol in de getaltheorie. In het begin werden zij met behulp van de complexe analyse bestudeerd, omdat de bijbehorende Fouriercoëfficiënten vaak getaltheoretische interpretaties bezitten. Bijvoorbeeld bestaat er een modulaire vorm waarvan de n-de Fouriercoëfficiënt gelijk is aan het aantal mogelijkheden het getal n als som van acht kwadraten te schrijven. Sinds de jaren zestig is de taal van de algebraïsche meetkunde, in het bijzonder die van de *aritmetische algebraïsche meetkunde*, in veel gebieden van de getaltheorie heel nuttig gebleken. Op grond van inzichten van Shimura, Weil, Serre en Deligne werd deze nieuwe taal met veel succes ook op de theorie van modulaire vormen toegepast en er werden enige diepe samenhangen ontdekt. Als hoogtepunt tot nu toe is het bewijs van het vermoeden van Fermat te noemen, dat in 1994 door Andrew Wiles gevonden werd. Dit vermoeden zegt dat de vergelijking

$$a^n + b^n = c^n$$

met gehele machten $n \geq 3$ geen oplossing heeft voor positieve natuurlijke getallen a, b, c.

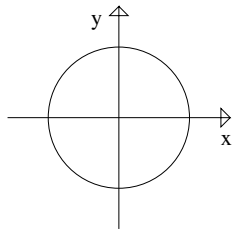
De samenhang tussen getaltheorie en meetkunde wil ik met behulp van een eenvoudig voorbeeld aanduiden. Laten we de vergelijking

$$a^2 + b^2 = c^2$$

beschouwen. Anders dan in het vermoeden van Fermat heeft deze vergelijking wel oplossingen, namelijk de bekende Pythagoreïsche drietallen, zoals $3^2 + 4^2 = 5^2$ of $5^2 + 12^2 = 13^2$. Als we $x = \frac{a}{c}$ en $y = \frac{b}{c}$ schrijven, dan verkrijgen we middels een eenvoudige manipulatie de vergelijking

$$x^2 + y^2 - 1 = 0.$$

Beschouwen we nu eerst alle reële oplossingen (d.w.z. we staan getallen toe die oneindig veel cijfers achter de komma mogen hebben en niet periodiek hoeven te zijn).



Met behulp van de parametrisatie $x = \cos(\varphi)$ en $y = \sin(\varphi)$ zien we dat de reële oplossingen precies de eenheidscirkel vormen (d.w.z. de cirkel om de oorsprong van het coördinatenvlak met straal 1). Nu zijn we heel duidelijk in de wereld van de meetkunde! Onze vraag naar de Pythagoreïsche drietallen kan nu worden vertaald in de vraag naar punten op het eenheidscirkel waarvan de coördinaten breuken (d.w.z. *rationale getallen*) zijn.

We zullen zien dat de vergelijking $a^2 + b^2 = c^2$ makkelijker te bestuderen is, als men niet alleen met breuken werkt maar ook met het getal i dat als een wortel van -1 gedefiniëerd is, d.w.z. als een oplossing van de vergelijking

$$X^2 + 1 = 0.$$

Volgens de hoofdstelling van de algebra heeft namelijk iedere zulke vergelijking over de complexe getallen even veel oplossingen (met multipliciteiten) als haar graad aangeeft; in dit geval dus twee, namelijk i en $-i$.

We zullen in de getallen van Gauß rekenen, dit zijn alle getallen die men door optellen en vermenigvuldigen van gehele getallen en het getal i verkrijgt. Het is makkelijk in te zien dat men iedere getal van Gauß als $a + ib$ met gehele getallen a en b kan schrijven. Laten we ons herinneren dat een positief natuurlijk getal ongelijk 1 een *priemgetal* heet, als zijn enige positieve delers 1 en het getal zelf zijn. Ieder geheel getal ongelijk 0 kan op de volgorde na op eenduidige manier geschreven worden als plus of min een product van priemgetallen, bijv. is $12 = 2 \cdot 2 \cdot 3$. Zoiets is ook voor de getallen van Gauß geldig. De enige getallen van Gauß die ieder willekeurig getal van Gauß delen zijn $1, -1, i, -i$; deze worden de Gaußeenheden genoemd. Een Gaußpriemgetal is een getal van Gauß ongelijk 1 die in de kwadrant rechts boven met de positieve x-as en zonder de y-as ligt en alleen door de Gaußeenheden en door zichzelf keer een Gaußeenheid gedeeld word. Ieder getal van Gauß kan op de volgorde na op eenduidige manier geschreven worden als product van Gaußpriemgetallen en een Gaußeenheid.

Bovendien is het volgende geldig: Als een priemgetal p bij deling door 4 rest 3 heeft (bijv. $p = 3$, $p = 7$ of $p = 11$), dan is $p = p + i \cdot 0$ ook een Gaußpriemgetal. We zeggen in dat geval dat p *inert* is. Als p gedeeld door 4 rest 1 heeft, dan kan men gehele getallen u, v vinden, zodat $p = u^2 + v^2$ geldig is, en dus kan men p in de getallen van Gauß factoriseren:

$$p = (u + iv)(u - iv) = u^2 - (iv)^2 = u^2 - (i)^2v^2 = u^2 - (-1)v^2 = u^2 + v^2.$$

Daarom is in dat geval p geen Gaußpriemgetal, maar $u + iv$ en $u - iv$ zijn dat wel (op een eenheid na). We zeggen dat p in de getallen van Gauß *gespleten* is. Een bijzondere rol speelt het priemgetal 2. Het is

$$2 = -i(1 + i)^2,$$

dus een Gaußeenheid keer het kwadraat van een Gaußpriemgetal. Het gehele priemgetal 2 heet daarom in de getallen van Gauß *vertakt*.

Laten we teruggaan naar de vergelijking $a^2 + b^2 = c^2$. In de getallen van Gauß kunnen we nu schrijven:

$$a^2 + b^2 = (a + ib)(a - ib) = c^2.$$

Als we aannemen dat a, b, c geen gemeenschappelijke delers hebben (we zoeken dus alleen primitieve Pythagoreïsche drietallen; door delen door de gemeenschappelijke factor kan men ieder Pythagoreïsch drietal in een primitieve veranderen), dan zijn $a + ib$ en $a - ib$ getallen van Gauß waarvan de gemeenschappelijke delers alleen de Gaußeenheden zijn. Wegens de unieke priemfactorisatie in de getallen van Gauß moet dan $a + ib$ zelf een kwadraat zijn. Dus er moet gelden

$$\epsilon(a + ib) = (u + iv)^2 = u^2 - v^2 + i2uv,$$

met een Gaußeenheid ϵ en gehele getallen u, v. Is $\epsilon = \pm 1$, dan verkrijgen we dus $a = \pm(u^2 - v^2)$ en $b = \pm 2uv$. Is $\epsilon = \pm i$, dan is het precies andersom, namelijk $a = \pm 2uv$ en $b = \mp(u^2 - v^2)$. Het is ook makkelijk te verifiëren dat door

$$a = u^2 - v^2, \quad b = 2uv, \quad c = u^2 + v^2$$

Pythagoreïsche drietallen gegenereerd worden, namelijk:

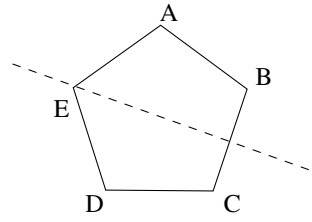
$$a^2 + b^2 = (u^2 - v^2)^2 + (2uv)^2 = u^4 + 2u^2v^2 + v^4 = (u^2 + v^2)^2 = c^2.$$

Dus hebben we alle primitieve Pythagoreïsche drietallen bepaald door het getalbereik waarin we rekenen slim uit te breiden. Dit is een van de belangrijkste methoden van de algebraïsche getaltheorie. In het algemeen bestudeert men onder vermenigvuldiging en optellen afgesloten uitbreidingen van de gehele getallen resp. de breuken, die door bijvoegen van oplossingen van vergelijkingen van de vorm

$$X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 = 0$$

met gehele getallen a_i ontstaan. Zulke oplossingen noemt men *gehele algebraïsche getallen*. In plaats van unieke priemfactorisatie heeft men echter in het algemeen alleen nog unieke priemideaalfactorisatie. Begrippen als inertie, splijten en vertakking bestaan ook in deze algemene context. Dit wordt *aritmetiek van getallenlichamen* genoemd.

Voordat we het over symmetriegroepen van getallenlichamen (de *Galoisgroepen*) hebben, behandelen we een voorbeeld van symmetriegroepen uit de platte euclidische meetkunde. We beschouwen de regelmatige vijfhoek (pentagoon). Welke afstandsbehoudende omkeerbare transformaties bestaan er, die de pentagoon op zichzelf afbeelden? Het zijn de rotaties over $n \cdot 72$ graden met $n \in \{0, 1, \ldots, 4\}$ en de spiegelingen door de assen die door een hoekpunt lopen en loodrecht op de tegenoverliggende zijde staan. Bij elkaar bestaan er dus 10 zulke transformaties. Het samenstellen van twee zulke levert altijd een derde



op. Bovendien kan men de transformaties weer omkeren (de rotatie over $n \cdot 72$ graden door de rotatie over $(5-n) \cdot 72$ graden, en de spiegeling door hem nog een keer te doen). Zoiets noemt men een groep. We hebben dus net de symmetriegroep van de regelmatige vijfhoek beschreven. In het algemeen noemt men de symmetriegroep van de regelmatige n-hoek de n-de *diëdergroep*. Zij heeft $2n$ elementen.

In de getaltheorie bekijkt men de symmetriegroepen van getallenlichamen en noemt deze *Galoisgroepen*. Laten we met het voorbeeld van boven doorgaan. De rationale getallen van Gauß zijn alle getallen $a+ib$ waarbij nu a en b breuken zijn. Een symmetrie van de rationale getallen van Gauß is een omkeerbare afbeelding van de rationale getallen van Gauß naar zichzelf die vermenigvuldiging en optellen behoudt. Zij is dan automatisch de identiteit op de breuken. Naast de identieke symmetrie bestaat er één andere. Deze wordt gegeven door het getal $a+ib$ op het getal $a-ib$ af te beelden, dus door complexe conjugatie. Past men deze afbeelding twee keer toe dan verkrijgt men weer de identiteit. De Galoisgroep van de rationale getallen van Gauß bevat precies deze twee elementen.

Maar er zijn ook getallenlichamen waarvan de symmetriegroep dezelfde vermenigvuldiging heeft als de symmetriegroep van de vijfhoek (in het algemeen geldt dit voor iedere regelmatige n-hoek). Bijvoorbeeld is dit het geval voor het getallenlichaam dat men verkrijgt door aan de breuken nog alle oplossingen van de vergelijking

$$X^5 - 2X^4 + 2X^3 - X^2 + 1$$

toe te voegen en ook nog alle getallen die door vermenigvuldiging en optellen hieruit ontstaan.

De symmetriegroep van de verzameling van alle algebraïsche getallen samen heet de *absolute Galoisgroep van de rationale getallen* en wordt door het symbool $G_{\mathbb{Q}}$ aangeduid. Uit deze groep kan men in principe alle informatie over alle getallenlichamen en hun aritmetiek aflezen! Dus is $G_{\mathbb{Q}}$ het centrale object van de algebraïsche getaltheorie. Helaas is de structuur van $G_{\mathbb{Q}}$ heel mysterieus (zij heeft bijv. overaftelbaar veel elementen, d.w.z. veel meer dan er gehele getallen bestaan) en zij is zeer slecht begrepen.

Op deze plaats speelt de diepe samenwerking van algebraïsche meetkunde en algebraische getaltheorie in de theorie van de modulaire vormen een heel belangrijke rol. Er is namelijk een theorema van Shimura, Deligne en Serre dat bij een modulaire vorm (die een eigenvorm is, dat betekent bijv. als de vorm als oneindige reeks $e^{2\pi i\tau} + \sum_{n=2}^{\infty} a_n e^{2\pi i n\tau}$ geschreven is dat dan $a_n \cdot a_m = a_{nm}$ geldt voor n en m zonder gemeenschappelijke factor) voor een gegeven priemgetal p een *Galoisrepresentatie* (dat is een continuë afbeelding, d.w.z. zij respecteert de meetkunde en het samenstellen)

$$G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$$

maakt (deze is oneven en semi-simpel). De rechterkant van de formule moet nog worden uitgelegd. Hier is $\overline{\mathbb{F}_p}$ de verzameling van alle oplossingen van vergelijkingen

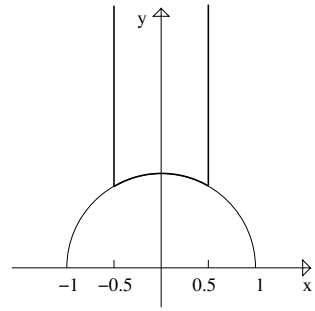
$$X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 = 0,$$

waar we nu van de coëfficiënten alleen de rest bekijken die zij bij het delen door p geven. Daarenboven is $\mathrm{GL}_2(\overline{\mathbb{F}_p})$ de groep van omkeerbare lineaire afbeeldingen van het vlak met coördinaten in $\overline{\mathbb{F}_p}$. Eenvoudig gezegd betekent dit dat we platte stukken van $G_{\mathbb{Q}}$ in karakteristiek p beschouwen. Hiervoor bestaat geen goede, intuïtieve aanschouwing, en de taal wordt alleen in analogie met de gewone, reële meetkunde gebruikt. De topologie, d.w.z. de manier waarop we de meetkunde op $\overline{\mathbb{F}_p}$ definiëren, namelijk diskreet, heeft als gevolg dat de "platte stukken" van $G_{\mathbb{Q}}$ eindig zijn, dus alleen maar uit eindig veel elementen bestaan. Dit betekent dan dat de modulaire vorm, waarmee we begonnen waren, een getallenlichaam oplevert. Het belangrijke is nu dat de aritmetiek van het getallenlichaam (ten minste gedeeltelijk) aan de coëfficiënten van de modulaire vorm kan worden afgelezen (die kunnen we berekenen; we kunnen ze zelf direct in $\overline{\mathbb{F}_p}$ nemen)! Op deze manier verlenen ons de modulaire vormen een klein inzicht in de mysterieuze absolute Galoisgroep $G_{\mathbb{Q}}$!

Laten we een voorbeeld bekijken. Er is een modulaire vorm van niveau 229 en gewicht 1 waarvan de coëfficiënten in de verzameling $\{0, 1\}$ liggen (met de optelling en vermenigvuldiging $1 + 0 = 1, 1 + 1 = 0, 1 \cdot 1 = 1, 1 \cdot 0 = 0$ dus in het eindige lichaam $\mathbb{F}_2$). Zij K het getallenlichaam dat uit de breuken door bijvoegen van een wortel van het priemgetal 229 gemaakt wordt. Zij l een priemgetal ongelijk aan 2 en 229. Dan is de l-de coëfficiënt van de modulaire vorm gelijk aan 0 dan en slechts dan als l in K inert is (d.w.z. dat geen kwadraat bij delen door l dezelfde rest heeft als 229) of dat l in twee hoofdidealen splijt. Anders is de coëfficiënt gelijk aan 1.

We hebben dus gezien, dat een modulaire vorm "platte stukken" van $G_{\mathbb{Q}}$ in karakteristiek p oplevert. De beroemde wiskundige Jean-Pierre Serre (in 2003 de eerste winnaar van de nieuwe Abelprijs die de Nobelprijs voor de wiskunde zal worden) heeft het vermoeden uitgesproken dat andersom alle "platte stukken" van $G_{\mathbb{Q}}$ in karakteristiek p door modulaire vormen kunnen worden beschreven. Hij heeft zelfs een formule aangegeven waarmee men naar de modulaire vormen moet zoeken (d.w.z. het niveau, het karakter en het gewicht). Als dit vermoeden waar is, dan kunnen we alle zulke platte stukken van $G_{\mathbb{Q}}$ met de computer berekenen, omdat we modulaire vormen kunnen berekenen! Serres vermoeden is dus zowel van groot structureel als van computationeel belang. Echter is het niet bekend of Serres vermoeden waar is. Maar enkele maanden geleden werd een belangrijk geval opgelost zodat het onderzoek tegenwoordig sterk in beweging is.

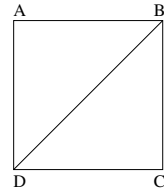
Nu zullen we kort modulaire krommen bespreken. Deze kunnen als het meetkundige aspect van modulaire vormen worden beschouwd. Bovendien geven zij de verbinding tussen modulaire vormen, modulaire symbolen (zie beneden) en Galoisrepresentaties. Modulaire krommen zijn voorlopig complexe krommen, dus vlakken in de aanschouwing. De eenvoudigste modulaire kromme is gegeven als de punten in het coördinatenvlak, waarvan de x-coördinaat tussen $-\frac{1}{2}$ en $\frac{1}{2}$ ligt en die op of boven de eenheidscirkel liggen. Nu moet men de linkerrand op de rechterrand plakken (letterlijk: we knippen dit gebied met een schaar uit; dan plakken we de twee lange lijnen aan elkaar; tenslotte plakken we nog de linkerhelft van de boog aan de rechterhelft; zo verkrijgt men een cilinder met een ietwat vreemde bodem).



Het op deze manier verkregen vlak is boven open. Men kan hem door bijvoegen van een punt, van een *spits*, compactificeren (ook dit is aanschouwelijk te maken: we duwen de cilinder aan de bovenkant tot een punt samen; dan zien we de spits heel duidelijk). Het zo ontstane vlak is een compact Riemannoppervlak, d.w.z. dat kleine stukken meetkundig er hetzelfde uitzien als het complexe getallenvlak. Modulaire vormen vindt men op de modulaire krommen terug als differentiaalvormen (deze heeft men bijv. nodig om op Riemannoppervlakken te integreren). Het belangrijke voor de getaltheorie is dat de modulaire krommen ook een vrij diepe algebraïsche structuur hebben, d.w.z. dat hun punten oplossingen van vergelijkingen met gehele coëfficiënten zijn, maar dan in meerdere variabelen. Ook de differentiaalvormen hebben een algebraïsche analogon, die de *Katz modulaire vormen* oplevert, die in dit proefschrift gebruikt worden. Ook de Galoisrepresentaties worden met behulp van de algebraïsche beschrijving van de modulaire kromme gemaakt.

Voor de studie van oppervlakken (en ook hogerdimensionale variëteiten) gebruikt men de (co-)homologietheorie. We zullen kort de homologietheorie van Riemannoppervlakken met triviale coëfficiënten beschrijven. Maar in het proefschrift worden ook (co-)homologietheorieën van schema's (dat zijn algebraïsche generalisaties van Riemannoppervlakken), stacks (dat zijn nog verdere generalisaties) en van groepen gebruikt en dan in het algemeen met niet-triviale coëfficiënten.

Men kan ieder Riemannoppervlak trianguleren, d.w.z. hem in eindig veel driehoeken opdelen (de zijden mogen krom zijn maar geen knikken bevatten). Voor het opdelen in driehoeken worden zijden getekend. Iedere driehoek heeft drie zijden en twee elkaar aanrakende driehoeken hebben ten minste een gemeenschappelijke zijde. Bovendien bekijken we de verzameling van snijpunten van zijden.

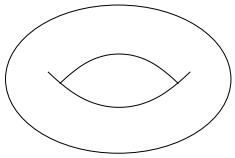


We beschrijven nu een triangulatie van de fietsband (de *torus*). Dit doen we constructief. We beginnen met de rechthoek uit het plaatje die we in twee driehoeken opgedeeld hebben. Door plakken zal het aantal zijden dalen. Eerst plakken we de zijde AD aan de zijde BC. Op deze manier verkrijgen we een cilinder. Nu plakken we het deksel op de bodem (we stellen ons de cilinder uit rubber voor).

De Eulerkarakteristiek van een oppervlak is $\chi = d - z + p$, waar d het aantal driehoeken, z het aantal zijden en p het aantal hoekpunten aanduidt. De Eulerkarakteristiek is onafhankelijk van de triangulatie. Bovendien geldt de beroemde formule

$$\chi = 2 - 2g,$$

waar g het geslacht van het oppervlak is, d.w.z. het aantal gaten.



In het voorbeeld van de fietsband vinden we inderdaad $g = 1$. We hebben namelijk nog altijd de twee driehoeken, waarmee we begonnen zijn. Omdat we de zijde AD met BC en ook AB met DC geïdentificeerd hebben, is het aantal zijden van onze triangulatie van de torus 3. Bovendien vallen alle vier hoekpunten onder het plakken samen tot één punt. Dus verkrijgen we inderdaad $\chi = 2-3+1 = 0$.

De modulaire kromme die we boven beschreven hebben heeft geen gat. Dus geldt voor haar $g = 0$. We kunnen ook de modulaire kromme makkelijk trianguleren. We vouwen haar weer uiteen en gebruiken maar één driehoek. Dit bestaat uit de twee hoekpunten beneden links en beneden rechts samen met een denkbeeldig punt helemaal boven (dit is het punt dat door het samenduwen van de cilinder ontstaan is). Dan hebben we na het plakken nog drie hoekpunten, twee zijden (de verticale en het stuk van de boog) en de driehoek. Dus verkrijgen we $\chi = 1 - 2 + 3 = 2$. Wat algemenere modulaire krommen, bijv. de in het proefschrift gebruikte modulaire kromme $X_1(N)$, hebben meestal veel gaten.

De homologiegroepen staan in nauwe relatie tot de Eulerkarakteristiek (de Eulerkarakteristiek wordt met behulp van de homologiegroepen afgeleid). De nulde en de tweede homologiegroep zijn vrije groepen van rang gelijk aan het aantal samenhangscomponenten. In ons geval is de rang van allebei dus 1. De eerste homologiegroep is ook een vrije groep. Haar rang is $2g$ met g het geslacht.

Nadat we nu geprobeerd hebben een eerste, heel erg vereenvoudigd idee te geven van de objecten die in het proefschrift behandeld worden, zullen we nu de inhoud ervan beschrijven.

Het eerste hoofdstuk is inmiddels als artikel verschenen. Er wordt een aangepaste versie van Serres vermoeden behandeld. Diepe resultaten van verschillende wiskundigen zeggen dat voor oneven karakteristiek p Serres formules voor het niveau, het karakter en het gewicht van de gepostuleerde modulaire vorm inderdaad juist zijn. Dit wil zeggen dat als er een modulaire vorm bestaat die een gegeven "plat stuk" van $G_{\mathbb{Q}}$ geeft, dan bestaat er ook een modulaire vorm die aan Serres formule voldoet. Het geval $p = 2$ is echter nog gedeeltelijk open.

In het artikel beperk ik me tot "platte stukken" in karakteristiek p van $G_{\mathbb{Q}}$ (dus tweedimensionale Galoisrepresentaties) waarvan de symmetriegroep een Diëdergroep, dus een symmetriegroep van een regelmatige n-hoek is. Voor deze toon ik het aangepaste Serrevermoeden aan zonder uitzondering, dus inclusief het geval $p = 2$. Dat zulk een Galoisrepresentatie van een modulaire vorm komt was in principe al Erich Hecke bekend, ten minste als $p \neq 2$ is. In het bewijs maak ik oneindig veel zulke modulaire vormen, zodat ik dan met behulp van het ladenprincipe (verdeel 10 letters over 5 laden, dan is er een lade waarin er ten minste twee liggen) er twee kan kiezen, die zich met methoden van de algebraïsche meetkunde tot de gewenste modulaire vorm laten combineren.

In het Hoofdstuk II bereken en vergelijk ik verschillende soorten cohomologiegroepen die alle met de modulaire kromme $X_1(N)$ (dit is een iets algemener Riemannoppervlak dan de hiervoor beschreven modulaire kromme) samenhangen, met het formalisme van de modulaire

symbolen dat van de homologie afgeleid is. In deze berekeningen is de coëfficiëntenring willekeurig. Er worden expliciete beschrijvingen in termen van lineaire algebra gegeven.

We bekijken modulaire symbolen om praktische redenen: zij zijn in het ver verspreide computeralgebrasysteem Magma geïmplementeerd. Ik heb computerprogramma's geschreven die hierop werken.

In het derde hoofdstuk worden nieuwe gevallen bewezen, wanneer de Katz modulaire vormen over $\overline{\mathbb{F}_p}$ met behulp van de expliciete beschrijvingen van de cohomologiegroepen uit Hoofdstuk II direct over het eindige lichaam $\mathbb{F}_p$ kunnen worden berekend. Dit betekent een snelheidswinst in vergelijking tot methoden die gehele getallen gebruiken. Met behulp van een idee van Edixhoven verkrijgen we zo ook een algoritme voor de berekening van Katz modulaire vormen van gewicht één (deze zijn niet direct berekenbaar) met behulp van modulaire symbolen over $\mathbb{F}_p$.

Het bewijs gebruikt het opmerkelijke parallel gedrag tussen de modulaire vormen van gewicht 2 en niveau Np over $\mathbb{F}_p$ en de eerste cohomologiegroepen van de Riemannoppervlakken $X_1(Np)$ met $\mathbb{F}_p$-coëfficiënten. In allebei vindt men namelijk de modulaire vormen resp. de eerste cohomologiegroepen terug, die bij het niveau N en het gewicht $k \in \{2, \ldots, p+1\}$ horen.

De overgang van de complexe meetkunde naar de algebraïsche over $\mathbb{F}_p$ vindt met behulp van de Jakobiaan van de modulaire kromme plaats. De eerste kohomologiegroep kan namelijk met de p-torsie van de complexe Jakobiaan geïdentificeerd worden. Gaat men dan naar het Néronmodel van de Jakobiaan, dan kan men eigenschappen van de generieke vezel (zelfs van het Riemannoppervlak) naar de speciale vezel (dus naar $\mathbb{F}_p$) overdragen.

Het vierde hoofdstuk bevat een beschrijving van de algoritmen die voortkomen uit de theorie van de twee voorgaande hoofdstukken. Tenslotte wordt in het vijfde hoofdstuk over computerberekeningen gerapporteerd die met behulp van de voorgestelde algoritmen zijn uitgevoerd. Er wordt bijvoorbeeld geconstateerd dat de platte stukken van $G_{\mathbb{Q}}$ in karakteristiek 2 opmerkelijk sterk groeien. Bovendien worden ook observaties gemaakt, die enkele interessante theoretische samenhangen suggereren. Hun studie zou het onderwerp van toekomstige projecten kunnen zijn.

Zusammenfassung

In dieser kurzen Zusammenfassung möchte ich eine möglichst allgemein verständliche Einführung in das Gebiet der vorliegenden Arbeit und einen Überblick über diese geben.

Modulformen spielen seit ihrer Einführung im 19. Jahrhundert eine zentrale Rolle in der Zahlentheorie. Zu Anfang wurden sie mit Hilfe der Funktionentheorie untersucht, da die zugehörigen Fourierkoeffizienten häufig interessante zahlentheoretische Bedeutungen haben. So gibt es z. B. eine Modulform, deren n-ter Fourierkoeffizient angibt, wie oft die natürliche Zahl n als Summe von 8 Quadraten dargestellt werden kann. Seit den 60er Jahren hat sich die Sprache der algebraischen Geometrie, besonders der *arithmetischen algebraischen Geometrie*, in vielen Bereichen der Zahlentheorie als sehr nützlich erwiesen. Auf Grund von Einsichten von Shimura, Weil, Serre und Deligne wurde diese neue Sprache mit viel Erfolg auch auf die Theorie der Modulformen angewandt und hat einige tief liegende Zusammenhänge zu Tage gebracht. Als spektakulärer bisheriger Höhepunkt ist der Beweis der Fermatschen Vermutung zu nennen, den Andrew Wiles 1994 gefunden hat. Diese Vermutung besagt, dass die Gleichung

$$a^n + b^n = c^n$$

für ganze Exponenten $n \geq 3$ keine Lösung in positiven natürlichen Zahlen a, b, c hat.

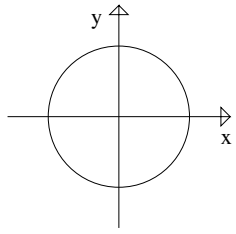
Den Zusammenhang zwischen Zahlentheorie und Geometrie möchte ich an einem einfachen Beispiel andeuten. Nehmen wir die Gleichung

$$a^2 + b^2 = c^2.$$

Im Gegensatz zum Fermatproblem hat diese Gleichung Lösungen, nämlich die wohlbekannten Pythagoräischen Tripel, z. B. $3^2 + 4^2 = 5^2$ oder $5^2 + 12^2 = 13^2$. Schreiben wir $x = \frac{a}{c}$ und $y = \frac{b}{c}$, dann erhalten wir durch einfache Umformungen die Gleichung

$$x^2 + y^2 - 1 = 0.$$

Wir können zunächst alle reellen Lösungen betrachten (d. h. wir erlauben Zahlen mit beliebiger, also auch unendlicher und nicht periodischer Dezimalschreibweise).



Mittels der Parametrisierung $x = \cos(\varphi)$ und $y = \sin(\varphi)$ sehen wir, dass die reellen Lösungen gerade den Einheitskreis bilden (d. h. den Kreis von Radius 1 um den Ursprung der Koordinatenebene). Damit sind wir nun ganz offensichtlich in der Welt der Geometrie! Unsere Frage nach den Pythagoräischen Tripeln übersetzt sich dann in die Frage nach Punkten auf dem Einheitskreis, deren Koordinaten Bruchzahlen (diese nennen wir *rationale Zahlen*) sind.

Es stellt sich heraus, dass sich die Gleichung $a^2+b^2 = c^2$ leichter untersuchen lässt, wenn man nicht nur Bruchzahlen zulässt, sondern auch die Zahl i, welche als eine Wurzel von -1 definiert ist, also als eine Lösung der Gleichung

$$X^2 + 1 = 0.$$

Nach dem Hauptsatz der Algebra hat nämlich jede solche Gleichung über den komplexen Zahlen so viele Lösungen (mit Multiplizitäten gezählt) wie der Grad ist; hier also zwei, und die Lösungen sind i und $-i$.

Wir werden in den Gaußschen Zahlen rechnen. Das sind alle Zahlen, die man durch Addition und Multiplikation von ganzen Zahlen und der Zahl i erhält. Es ist einfach einzusehen, dass sich jede Gaußche Zahl schreiben lässt als $a + ib$ mit ganzen Zahlen a und b. Erinnern wir uns, dass eine positive natürliche Zahl ungleich 1 *Primzahl* heißt, wenn die einzigen positiven Teiler 1 und die Zahl selbst sind. Jede ganze Zahl ungleich 0 lässt sich auf bis auf die Reihenfolge eindeutige Weise als plus oder minus einem Produkt von Primzahlen schreiben, z. B. ist $12 = 2 \cdot 2 \cdot 3$. Etwas ganz Ähnliches gilt in den Gaußschen Zahlen. Die einzigen Gaußschen Zahlen, die jede beliebige Gaußsche Zahl teilen, sind $1, -1, i, -i$; diese heißen Gaußsche Einheiten. Eine Gaußsche Primzahl ist eine Gaußsche Zahl ungleich 1, die im Quadraten rechts oben einschließlich dem positiven Teil der x-Achse ohne die y-Achse liegt und nur von den Gaußschen Einheiten und von sich selbst mal einer Gaußschen Einheit geteilt wird. Jede Gaußsche Zahl lässt sich auf bis auf die Reihenfolge eindeutige Weise als Produkt einer Gaußschen Einheit mit einem Produkt von Gaußschen Primzahlen schreiben. Es gilt ferner Folgendes: Wenn die ganze Primzahl p beim Teilen durch 4 den Rest 3 ergibt (z. B. $p = 3$, $p = 7$ oder $p = 11$), dann ist $p = p + i \cdot 0$ auch eine Gaußsche Primzahl. Wir sagen dann, dass p *träge* ist. Lässt p aber beim Teilen durch 4 den Rest 1, dann kann man ganze Zahlen u, v finden, derart dass $p = u^2 + v^2$ gilt, und daher kann man p in den Gaußschen Zahlen faktorisieren:

$$p = (u + iv)(u - iv) = u^2 - (iv)^2 = u^2 - (i)^2v^2 = u^2 - (-1)v^2 = u^2 + v^2.$$

Daher ist in diesem Fall p keine Gaußsche Primzahl, stattdessen aber $u + iv$ und $u - iv$ (evtl. bis auf eine Einheit). Wir sagen, dass p in den Gaußschen Zahlen *zerlegt* ist. Eine besondere Rolle spielt die Primzahl 2, sie ist

$$2 = -i(1 + i)^2,$$

d. h. eine Gaußsche Einheit mal einem Quadrat einer Gaußschen Primzahl. Die ganze Primzahl 2 heißt deswegen in den Gaußschen Zahlen *verzweigt*.

Kommen wir zurück zur Gleichung $a^2 + b^2 = c^2$. In den Gaußschen Zahlen können wir diese nun so schreiben:

$$a^2 + b^2 = (a + ib)(a - ib) = c^2.$$

Wenn wir annehmen, dass a, b, c keine gemeinsamen Teiler haben (wir suchen dann nur primitive Pythagoräische Tripel; durch das Herausteilen des gemeinsamen Faktors kann jedes Pythagoräische Tripel auf ein primitives zurückgeführt werden), dann sind $a + ib$ und $a - ib$ teilerfremde Gaußsche Zahlen, d. h. dass ihre gemeinsamen Teiler nur die Gaußschen Einheiten sind. Wegen der eindeutigen Primfaktorzerlegung in den Gaußschen Zahlen muss dann aber $a + ib$ schon selbst ein Quadrat sein, also muss gelten

$$\epsilon(a + ib) = (u + iv)^2 = u^2 - v^2 + i2uv,$$

mit einer Gaußschen Einheit ϵ und ganzen Zahlen u, v. Ist $\epsilon = \pm 1$, dann erhalten wir also $a = \pm(u^2 - v^2)$ und $b = \pm 2uv$. Ist $\epsilon = \pm i$, dann ist es gerade umgekehrt $a = \pm 2uv$ und $b = \mp(u^2 - v^2)$. Andersherum ist es ganz einfach nachzuprüfen, dass die Zuordnung

$$a = u^2 - v^2, \quad b = 2uv, \quad c = u^2 + v^2$$

Pythagoräische Tripel erzeugt, nämlich:

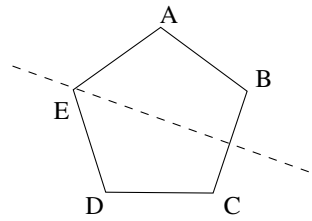
$$a^2 + b^2 = (u^2 - v^2)^2 + (2uv)^2 = u^4 + 2u^2v^2 + v^4 = (u^2 + v^2)^2 = c^2.$$

Damit haben wir alle primitiven Pythagoräischen Tripel bestimmt, indem wir den Zahlenbereich, in dem wir rechnen, geschickt erweitert haben. Dies ist eine Hauptmethode der algebraischen Zahlentheorie. Allgemeiner studiert man unter Multiplikation und Addition abgeschlossene Erweiterungen der ganzen Zahlen bzw. der Bruchzahlen, die durch Hinzufügen von Lösungen von Gleichungen der Form

$$X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 = 0$$

mit ganzen Zahlen a_i entstehen. Solche Lösungen nennt man *ganze algebraische Zahlen*. An die Stelle der eindeutigen Primfaktorzerlegung tritt dann jedoch im Allgemeinen nur noch die eindeutige Primidealzerlegung. Begriffe wie Trägheit, Zerlegung und Verzweigung hat man jedoch auch im erweiterten Sinn. Sie werden zusammengefasst im Begriff *Arithmetik der Zahlkörper*.

Bevor wir zu Symmetriegruppen von Zahlkörpern (den sogenannten *Galoisgruppen*) kommen, behandeln wir ein Beispiel von Symmetriegruppen aus der ebenen euklidischen Geometrie. Wir betrachten das regelmäßige Fünfeck (Pentagon).



Welche abstandserhaltenden umkehrbaren Transformationen gibt es, die das Pentagon in sich selbst überführen? Es sind dies die Drehungen um $n \cdot 72$ Grad mit $n \in \{0, 1, \ldots, 4\}$ und die Spiegelungen an den Achsen, die durch einen Eckpunkt gehen und senkrecht auf der gegenüber dem Eckpunkt liegenden Seite stehen. Insgesamt gibt es also 10 solche Transformationen. Das Hintereinanderausführen von zwei solchen liefert eine dritte. Außerdem kann man die Transformationen wieder rückgängig machen (die Rotation um $n \cdot 72$ Grad durch Rotation um $(5-n) \cdot 72$ Grad, und die Spiegelung durch nochmaliges Ausführen). So etwas nennt man eine Gruppe. Wir haben also gerade die Symmetriegruppe des regelmäßigen Fünfecks beschrieben. Im Allgemeinen nennt man die Symmetriegruppe des regelmäßigen n-Ecks die n-te *Diedergruppe*. Diese hat $2n$ Elemente.

In der Zahlentheorie betrachtet man Symmetriegruppen von Zahlkörpern und nennt diese Galoisgruppen. Schließen wir an das Beispiel von oben an. Die gebrochenen Gaußschen Zahlen sind alle Zahlen $a + ib$, wobei a, b nun Bruchzahlen sind. Eine Symmetrie der gebrochenen Gaußschen Zahlen ist eine umkehrbare Selbstabbildung, die die Multiplikation und die Addition respektiert. Sie ist dann automatisch die Identität auf den Bruchzahlen. Neben der identischen Symmetrie gibt es eine weitere. Diese ist dadurch gegeben, dass die Zahl $a + ib$ auf die Zahl $a - ib$ abgebildet wird. Führt man diese Abbildung zweimal nacheinander aus, so erhält man wieder die Identität. Die Galoisgruppe der gebrochenen Gaußschen Zahlen enthält genau diese zwei Elemente.

Es gibt aber auch Zahlkörper, deren Symmetriegruppe denselben Gesetzen folgt wie die Symmetriegruppe des Fünfecks (allgemeiner gilt dies für jedes regelmässige n-Eck). Z. B. ist dies der Fall beim Zahlkörper, den man erhält, indem man zu den Bruchzahlen noch alle Lösungen der Gleichung

$$X^5 - 2X^4 + 2X^3 - X^2 + 1$$

hinzufügt und alle Zahlen, die man aus diesen durch Multiplikation und Addition erhält.

Die Symmetriegruppe der Menge aller algebraischen Zahlen überhaupt nennt man die *absolute Galoisgruppe der rationalen Zahlen* und bezeichnet sie mit dem Symbol $G_{\mathbb{Q}}$. Aus ihr kann man im Prinzip alle Informationen zu allen Zahlkörpern und deren Arithmetik ablesen! Daher ist $G_{\mathbb{Q}}$ das zentrale Objekt der algebraischen Zahlentheorie. Allerdings ist die Struktur von $G_{\mathbb{Q}}$ sehr mysteriös (sie hat z. B. überabzählbar viele Elemente, d. h. viel mehr als es ganze Zahlen gibt) und sie ist nur sehr schlecht verstanden.

An dieser Stelle kommt nun das tiefliegende Zusammenspiel von algebraischer Geometrie und algebraischer Zahlentheorie in der Theorie der Modulformen voll zum Tragen. Es gibt nämlich einen Satz von Shimura, Deligne und Serre, der einer Modulform (die eine Eigenform ist, d. h. u. a. wenn sie geschrieben wird als unendliche Reihe $e^{2\pi i\tau} + \sum_{n=2}^{\infty} a_n e^{2\pi i n\tau}$, dass dann $a_n \cdot a_m = a_{nm}$ gilt für n und m ohne gemeinsamen Faktor) für eine vorgegebene Primzahl p eine *Galoisdarstellung* (das ist eine stetige Abbildung von Gruppen, d. h. sie

respektiert die Geometrie und das Hintereinanderausführen)

$$G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$$

zuordnet (genauer: diese ist ungerade und halbeinfach). Die rechte Seite der Formel ist noch zu erklären. Hier ist $\overline{\mathbb{F}_p}$ die Menge aller Lösungen von Gleichungen

$$X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 = 0,$$

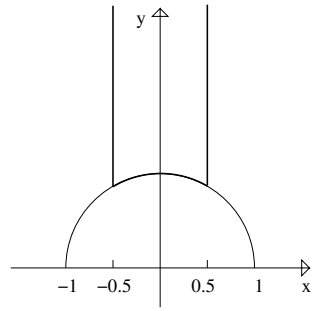
wobei wir nun von den Koeffizienten nur den Rest betrachten, den sie beim Teilen durch p lassen. Ferner ist $\mathrm{GL}_2(\overline{\mathbb{F}_p})$ die Gruppe der umkehrbaren linearen Abbildungen der Ebene mit Koordinaten in $\overline{\mathbb{F}_p}$. Vereinfacht ausgedrückt bedeutet dies, dass uns eine solche Galoisdarstellung "ebene Stücke" von $G_{\mathbb{Q}}$ in Charakteristik p liefert. Davon existiert keine gute, anschauliche Vorstellung, und die Sprache wird nur in Analogie zur gewöhnlichen reellen Geometrie gebraucht. Die Topologie, d. h. die Art, wie wir uns die Geometrie auf $\overline{\mathbb{F}_p}$ definieren, nämlich diskret, hat zur Folge, dass die "ebenen Stücke" von $G_{\mathbb{Q}}$ endlich sind, d. h. nur aus einer endlichen Anzahl Elementen bestehen. Das wiederum bedeutet, dass uns die Modulform, von der wir ausgegangen sind, einen Zahlkörper liefert. Das Wichtige dabei ist, dass die Arithmetik dieses Zahlkörpers (zumindest zum Teil) an den Koeffizienten der Modulform abzulesen ist (diese können wir ausrechnen; wir können sie sogar gleich in $\overline{\mathbb{F}_p}$ nehmen)! Damit gewähren uns Modulformen einen kleinen Einblick in die mysteriöse absolute Galoisgruppe $G_{\mathbb{Q}}$!

Dies veranschaulichen wir uns an einem Beispiel. Es gibt eine Modulform von Stufe 229 und Gewicht 1, deren Koeffizienten a_n in der Menge $\{0, 1\}$ liegen (mit der Addition und Multiplikation $1 + 0 = 1, 1 + 1 = 0, 1 \cdot 1 = 1, 1 \cdot 0 = 0$, mit anderen Worten dem endlichen Körper $\mathbb{F}_2$). Sei K der Zahlkörper, der aus den Bruchzahlen durch Hinzunehmen einer Quadratwurzel der Primzahl 229 gebildet wird. Sei l eine Primzahl, die nicht 2 und nicht 229 ist. Dann ist der l-te Koeffizient unserer Modulform gleich 0 genau dann, wenn l träge in K ist (mit anderen Worten, wenn keine Quadratzahl beim Teilen durch l denselben Rest lässt wie 229) oder l in zwei Hauptideale zerfällt. Sonst ist der Koeffizient gleich 1.

Wir haben also gesehen, dass eine Modulform uns ebene Stücke von $G_{\mathbb{Q}}$ in Charakteristik p gibt. Der berühmte Mathematiker Jean-Pierre Serre (2003 der erste Gewinner des neuen Abel-Preises, der der "Nobelpreis" für Mathematik werden soll) hat die Vermutung ausgesprochen, dass umgekehrt alle ebenen Stücke von $G_{\mathbb{Q}}$ in Charakteristik p durch Modulformen beschrieben werden können. Er hat sogar noch eine Formel angegeben, wo die Modulformen zu suchen sind (die Stufe, den Charakter und das Gewicht). Ist diese Vermutung war, dann können wir alle solche ebenen Stücke von $G_{\mathbb{Q}}$ mit dem Computer berechnen, denn wir können Modulformen berechnen! Serres Vermutung ist daher sowohl von ungeheuerer struktureller als auch von rechnerischer Bedeutung. Allerdings ist nicht bekannt, ob Serres Vermutung wahr ist. Aber vor wenigen Monaten wurde ein wichtiger Fall gelöst, so dass die Forschung gerade stark in Bewegung ist.

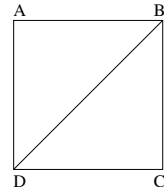
Als nächstes wollen wir kurz Modulkurven betrachten. Diese können als der geometrische Aspekt von Modulformen angesehen werden. Außerdem bilden sie das Verbin-

dungsglied zwischen Modulformen, Modulsymbolen (siehe unten) und Galoisdarstellungen. Modulkurven sind zunächst komplexe Kurven, d. h. Flächen in der Anschauung. Die allereinfachste ist gegeben als die Punkte in der Koordinatenebene, deren x-Koordinate zwischen $-\frac{1}{2}$ und $\frac{1}{2}$ liegt und die auf bzw. über dem Einheitskreis liegen. Dabei muss man nun den linken Rand mit dem rechten Rand verkleben (das ist ganz wörtlich vorstellbar: wir schneiden diesen Bereich mit der Schere aus; dann kleben wir die beiden langen Geraden zusammen; schließlich kleben wir noch die linke Hälfte des Bogenstücks mit der rechten zusammen; dann hat man einen Zylinder mit einem etwas komischen Boden). Die so erhaltene (Ober-)Fläche ist oben offen. Man kann diese durch Hinzufügen eines Punktes, einer sogenannten *Spitze*, kompaktifizieren (auch das ist ganz bildlich: wir drücken den Zylinder oben zu einem Punkt zusammen; dann sehen wir die Spitze ganz deutlich). Die so entstandene Fläche ist eine kompakte Riemannsche Fläche, d. h. kleine Stücke haben dieselbe Geometrie wie die komplexe Zahlenebene. Modulformen findet man auf den Modulkurven wieder als sog. Differentialformen (diese gebraucht man z. B. zum Integrieren auf der Riemannschen Fläche). Der entscheidende Punkt für die Zahlentheorie ist, dass die Modulkurven eine recht tiefliegende algebraische Struktur haben, d. h. dass ihre Punkte auch Lösungen von Gleichungen mit ganzen Koeffizienten sind, allerdings in vielen Variablen. Auch die Differentialformen haben ein algebraisches Analogon, das uns die sogenannten *Katz-Modulformen* liefert, die in der vorliegenden Arbeit benutzt werden. Auch die Galoisdarstellungen werden mit Hilfe der algebraischen Beschreibung der Modulkurven konstruiert.



Dem Studium von Flächen (und höherdimensionalen Varietäten) dient die (Ko-)Homologietheorie. Die Homologietheorie von Riemannschen Flächen mit trivialen Koeffizienten wollen wir kurz vorstellen. In der vorliegenden Arbeit werden aber auch Kohomologietheorien von Schemas (das sind weitreichende algebraische Verallgemeinerungen von Riemannschen Flächen), Stacks (das sind noch andere Verallgemeinerungen) und von Gruppen und dann im allgemeinen mit nicht-trivialen Koeffizienten benutzt.

Eine Riemannsche Fläche kann man triangulieren, d. h. sie in endlich viele Dreiecke aufteilen (dabei dürfen die Seiten "krumm" sein, aber keine Knicke enthalten). Zur Aufteilung in Dreiecke werden Seiten gezogen, d. h. jedes Dreieck hat drei Seiten und zwei aneinander grenzende Dreiecke haben (mindestens) eine gemeinsame Seite. Außerdem betrachten wir die Menge der Schnittpunkte der Seiten.



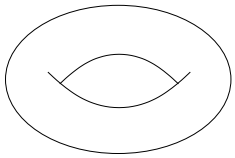
Wir beschreiben nun eine Triangulation des Fahrradreifens (des sog. *Torus*). Dabei gehen wir konstruktiv vor. Wir beginnen mit dem nebenstehenden Rechteck, das wir in zwei Dreiecke aufgeteilt haben. Durch Zusammenkleben wird sich die Anzahl der Seiten verkleinern. Zunächst kleben wir die Seite AD an die Seite BC. Auf diese Weise erhalten wir einen Zylinder.

Nun kleben wir den Deckel an den Boden (dazu stellen wir uns Gummi als Baumaterial vor).

Die Eulercharakteristik einer Fläche ist definiert als $\chi = d - s + p$, wobei d die Anzahl der Dreiecke, s die Anzahl der Seiten und p die Anzahl der Punkte bezeichnen. Die Eulercharakteristik ist unabhängig von der Triangulation. Außerdem gilt die berühmte Beziehung

$$\chi = 2 - 2g,$$

wobei g das Geschlecht der Fläche ist, d. h. die Anzahl der Löcher.



Im Beispiel des Fahrradreifens finden wir in der Tat $g = 1$. Wir haben nämlich noch immer die beiden Dreiecke, von denen wir ausgegangen waren. Da wir die Seite AD mit BC und ferner auch AB mit DC identifiziert haben, ist die Anzahl der Seiten unserer Triangulation des Torus 3. Außerdem fallen durch obige Verklebungen alle vier Punkte zusammen zu einem. Somit ergibt sich tatsächlich $\chi = 2 - 3 + 1 = 0$.

Die oben beschriebene Modulkurve hat kein Loch. Daher gilt für sie $g = 0$. Wir können auch die Modulkurve einfach triangulieren. Dazu falten wir sie wieder auseinander. Wir benutzen nur ein Dreieck. Dieses besteht aus den beiden Punkten unten links und unten rechts und einem gedachten Punkt ganz oben (der Punkt der durch das Zusammendrücken des Zylinders entstanden ist). Dann haben wir nach dem Zusammenkleben noch drei Punkte, zwei Seiten (die senkrechte und das Bogenstück) und das Dreieck. Damit ergibt sich $\chi = 1 - 2 + 3 = 2$. Allgemeinere Modulkurven, wie z. B. die in der Arbeit behandelte Modulkurve $X_1(N)$, haben in der Regel viele Löcher.

Die Homologiegruppen stehen zur Eulercharakteristik in enger Beziehung (die Eulercharakteristik wird aus diesen abgeleitet). Die nullte und die zweite Homologiegruppe sind freie Gruppen vom Rang gleich der Anzahl der Zusammenhangskomponenten. In unserem Fall ist der Rang beider also 1. Die erste Homologiegruppe ist wiederum eine freie Gruppe. Ihr Rang ist $2g$, wobei g wie oben die Anzahl der Löcher ist.

Nachdem wir versucht haben, eine erste, sehr stark vereinfachte Idee von den in der vorliegenden Arbeit untersuchten Objekten zu geben, wenden wir uns nun dem Inhalt zu.

Das erste Kapitel ist bereits als eigenständiger Artikel erschienen. Es geht in ihm um eine leicht modifizierte Version von Serres Vermutung. In tief liegenden Arbeiten einer Vielzahl Mathematiker wurde gezeigt, dass für ungerade Charakteristik p Serres Formeln für die Stufe, den Charakter und das Gewicht der vorhergesagten Modulform richtig sind. Genauer, wenn irgendeine Modulform existiert, die ein vorgegebenes ebenes Stück von $G_{\mathbb{Q}}$ gibt, dann gibt es auch eine da, wo Serres Formeln diese voraussagen. Der Fall $p = 2$ ist jedoch zum Teil noch offen.

In dem Artikel beschränke ich mich auf "ebene Stücke" in Charakteristik p von $G_{\mathbb{Q}}$ (also zweidimensionale Galoisdarstellungen), deren Symmetriegruppe eine Diedergruppe, also eine Symmetriegruppe eines regelmässigen n-Ecks ist. Von diesen zeige ich die modifizierte

Serre-Vermutung ohne Ausnahme, d. h. einschließlich $p = 2$. Dass solche Galoisdarstellungen von irgendeiner Modulform kommen, war im Prinzip schon Erich Hecke bekannt, zumindest wenn $p \neq 2$ ist. Im Beweis mache ich unendlich viele solche Modulformen, so dass ich dann mittels des Schubfachprinzips (verteile 10 Briefe auf 5 Schubfächer, dann liegen in einem mindestens zwei Breife) zwei wählen kann, die sich mit Hilfe algebraisch geometrischer Methoden zu der gewünschten Form kombinieren lassen.

Im Kapitel II berechne und vergleiche ich verschiedene Arten von (Ko-)Homologiegruppen, die alle mit der Modulkurve $X_1(N)$ (einer etwas allgemeineren als der oben vorgestellten Riemannschen Fläche) zusammenhängen, mit dem Modulsymbolformalismus, der an die Homologie angelehnt ist. Dabei ist der Koeffizientenring beliebig. Es werden jeweils explizite Beschreibungen in Termen von linearer Algebra abgeleitet.

Modulsymbole betrachten wir aus praktischen Gesichtspunkten: sie sind im weit verbreiteten Computeralgebrasystem Magma implementiert. Ich habe Computerprogramme erstellt, die hierauf beruhen.

Im dritten Kapitel werden neue Fälle bewiesen, in denen die Katz-Modulformen über $\mathbb{F}_p$ mit Hilfe der expliziten Beschreibungen der Kohomologiegruppen aus Kapitel II direkt über dem endlichen Körper $\mathbb{F}_p$ berechnet werden können. Dieses bringt einen Geschwindigkeitszuwachs im Vergleich zu Methoden, die mit ganzen Zahlen rechnen. Unter Benutzung einer Idee von Edixhoven erhalten wir auch einen Algorithmus zur Berechnung von Katz-Modulformen von Gewicht eins (diese sind nicht direkt zugänglich!) mittels Modulsymbolen über $\mathbb{F}_p$.

Ausgenutzt wird im Beweis eine erstaunliche Parallelität im Verhalten der Modulformen von Gewicht 2 und Stufe Np über $\mathbb{F}_p$ und der ersten Kohomologiegruppen der Riemannschen Fläche $X_1(Np)$ mit $\mathbb{F}_p$-Koeffizienten. In beiden spiegeln sich nämlich die Modulformen bzw. die ersten Kohomologiegruppen wider, die zu Stufe N und Gewicht $k \in \{2, \ldots, p+1\}$ gehören.

Der Übergang von komplexer Geometrie zu algebraischer Geometrie über $\mathbb{F}_p$ wird dabei mit Hilfe der Jakobischen der Modulkurve bewerkstelligt. Die erste Kohomologiegruppe kann nämlich mit der p-Torsion der komplexen Jakobischen identifiziert werden. Geht man dann zum Néronmodell der Jakobischen über, so gelingt es, Eigenschaften von der generischen Faser (sogar der Riemannschen Fläche) zur speziellen Faser (also nach $\mathbb{F}_p$) zu übertragen.

Das vierte Kapitel enthält eine Beschreibung der Algorithmen, die sich aus der Theorie der beiden vorangehenden Kaptiel ergeben. Schließlich wird im fünften Kapitel von Computerberechnungen berichtet, die mit Hilfe der vorgestellten Algorithmen ausgeführt wurden. Dabei wurde zum Beispiel festgestellt, dass die ebenen Stücke von $G_{\mathbb{Q}}$ in Charakteristik 2 erstaunlich schnell sehr groß werden. Desweiteren wurden noch andere Beobachtungen gemacht, die einige interessante theoretische Zusammenhänge suggerieren. Das Studium dieser kann Gegenstand zukünftiger Arbeiten sein.

Curriculum Vitae

Gabor Wiese werd op 6 februari 1976 te Warendorf in Duitsland geboren. Na zijn eindexamen (Abitur) aan het Gymnasium Laurentianum te Warendorf werkte hij als Zivildienstleistender tijdens 13 maanden in begeleiding van geestelijk gehandicapte kinderen. In october 1996 begon hij zijn studie aan de Universität Heidelberg, waar hij in 1998 het Vordiplom in de vakken Wiskunde en Natuurkunde haalde. Het studiejaar 1999/2000 studeerde hij aan de universiteit van Cambridge in Engeland en haalde het *Certificate of Advanced Study in Mathematics*. Terug in Heidelberg studeerde hij in november 2001 af bij Prof. K. Wingberg met een scriptie getiteld "Zur Fontaine-Mazur-Vermutung für Erweiterungen mit beschränkter Verzweigung von CM-Körpern".

In december 2001 ging hij naar Rennes in Frankrijk om zijn promotieonderzoek met Bas Edixhoven als begeleider te beginnen. In augustus 2002 volgde hij Edixhoven naar Leiden, waar hij na een periode als predoc van het Europese Netwerk "Arithmetic Algebraic Geometry" een aanstelling als Assistent in Opleiding (AiO) kreeg. Tijdens de eerste 6 maanden werd hij betaald door het Europese Netwerk "European Algebraic Geometry Education and Research (EAGER)". In de eerste acht maanden in Leiden bereikte hij het opmerkelijke cijfer van vijf keer verhuizen.

Zijn promotieonderzoek leidde tot dit proefschrift.

Sinds juli 2005 is hij getrouwd met Cécile Wiese née Vantard.

Na zijn promotie zal hij als postdoc aan de Universität Regensburg gaan werken.