Computing modular forms

John Voight
jvoight@gmail.com

v.0.1
January 15, 2026

Preface

These notes are in a preliminary version, intended as a preview for the 2026 Arizona Winter School for students in projects, including a more complete description of possible projects. Hopefully it will also be useful to others who want to get a first impression of the lecture contents and prepare a bit in terms of background.

In the meantime, I am really sorry for the many mistakes in these notes! If you see something that looks wrong, it probably is wrong or a typo or something else not helpful, and I'm sorry about that. Please write to me at jvoight@gmail.com and I will fix it in the next version, hopefully before the lectures start!

Goal

The goal is to give a broad survey of what it means to compute modular forms, with the intended audience graduate students, so including enough background material that it can be a launching point for further work, either theoretical or algorithmic.

Acknowledgements

Thanks to Eran Assaf and Edgar Costa for helpful discussions.

Contents

Contents ii

1 Introduction 1

I Classical modular forms 3

2 Classical modular curves 5
2.1 Beginnings . . . 5
2.2 Congruence subgroups . . . 7
2.3 Modular forms . . . 8
2.4 Modular forms as sections of a line bundle . . . 10
2.5 Eisenstein series . . . 11
2.6 Hecke operators . . . 12
2.7 Newforms and oldforms . . . 15
2.8 Graded ring and equations . . . 15
2.9 The computational problem . . . 15

3 Modular symbols 17
3.1 Farey symbols . . . 17
3.2 Coset representatives . . . 20
3.3 Modular symbols in homology . . . 21
3.4 Integration pairing . . . 23
3.5 Manin symbols and reduction . . . 25
3.6 Hecke operators . . . 26
3.7 Cohomology . . . 27
Exercises . . . 27

4 The trace formula 29
4.1 Eichler–Selberg trace formula . . . 29
Exercises . . . 29

5 Additional methods 31
5.1 Multiplication and graded ring . . . 31
5.2 Eisenstein series . . . 31

Exercises . . . 31

II Modular forms on reductive groups **33**

6 Bianchi modular forms **35**
6.1 Compactification . . . 35
6.2 Modular symbols . . . 35
6.3 Eisenstein series . . . 35
Exercises . . . 35

7 Reductive algebraic groups **37**
7.1 Beyond GL_2 . . . 37
7.2 Algebraic groups . . . 38
7.3 Subgroups, identity component . . . 41
7.4 Restriction of scalars . . . 44
7.5 Unipotent and reductive . . . 44
7.6 Tori, characters, and rank . . . 45
7.7 Parabolics and Levi decomposition . . . 46
7.8 Classification over $\mathbb{C}$ . . . 46
7.9 Classification over $\mathbb{R}$ . . . 46
7.10 Integral models . . . 47
Exercises . . . 47

8 Symmetric spaces **49**
8.1 Overview . . . 49
8.2 Symmetric domain . . . 50
8.3 Spaces of lattices and quadratic forms . . . 52
8.4 Hyperbolic spaces . . . 54
8.5 Representations . . . 55
8.6 Arithmetic groups . . . 55
8.7 Hermitian symmetric domains . . . 57
8.8 Borel–Serre compactification . . . 57
8.9 Reductive Borel–Serre compactification . . . 59
8.10 Compactifications in the Hermitian case . . . 59
Exercises . . . 60

9 Adelic quotients **61**
9.1 Overview . . . 61
9.2 Level structure . . . 61
9.3 Double cosetification . . . 62
9.4 Hecke operators . . . 63

10 Modular forms on reductive groups **65**
10.1 Automorphic forms . . . 65
10.2 Quick review on cohomology . . . 66

10.3 Cohomology . . . 66
10.4 $(\mathfrak{g}, K)$-cohomology . . . 68
10.5 Cuspidal and Eisenstein cohomology . . . 68
10.6 Hecke operators . . . 68
10.7 Degeneracy maps . . . 69

11 Computing modular forms **71**
11.1 Computational problem . . . 71
11.2 Algorithm overview . . . 72
11.3 Discussion . . . 72

III Algebraic modular forms **73**

12 Theta series and Brandt matrices **75**
12.1 Theta series and lattice methods . . . 75
12.2 Brandt matrices . . . 75

13 Definite setting **77**
13.1 The totally definite setting . . . 77
13.2 Algebraic modular forms . . . 78
13.3 Algorithmic details . . . 79
13.4 Change of level . . . 80
13.5 Orthogonal groups and lattices . . . 81
13.6 Unitary groups . . . 81
13.7 Exceptional isomorphisms . . . 81

IV Geometry and cohomology **83**

14 Fundamental domains **85**
14.1 Computing fundamental domains . . . 85
14.2 Computing in cohomology . . . 86
14.3 Shimura curves . . . 86
Exercises . . . 86

15 Generalized modular symbols **87**
15.1 Cycles from parabolics . . . 87
15.2 Hecke correspondences . . . 90
15.3 Shapiro's lemma . . . 90
15.4 Computations in practice . . . 90

16 Complexes **91**

17 Extensions **93**

V Projects (with Eran Assaf) 95

18 Projects 97
18.1 Project 1: Picard from Hecke 97
18.2 Project 2: Siegel stability 99
18.3 Project 3: Binary Hermitian 102
18.4 Project 4: Eisenstein in polynomial time 104
18.5 Project 5: Symplectic lattices 106
18.6 Project 6: Bad, bad Hecke 108

Bibliography 111

Chapter 1

Introduction

We begin with a brief history of computing modular forms.

Part I

Classical modular forms

Chapter 2

Classical modular curves

In this chapter, we quickly review classical modular curves with a focus on computational and algorithmic aspects.

2.1 Beginnings

There are very many references for classical modular groups, including Apostol [Apo90, Chapter 2], Diamond–Shurman [DS05, Chapter 2], and Serre [Ser73, Chapter VII]. For continuity, we largely follow Voight [Voi21, Chapters 33–35].

2.1.1. The **upper half-plane** is the set

$$\mathbf{H}^2 := \{z = x + iy \in \mathbb{C} : \mathrm{Im}(z) = y > 0\} \tag{2.1.2}$$

equipped with the hyperbolic metric $\mathrm{d}s := |\mathrm{d}z|/\operatorname{Im} z$. The group of orientation-preserving isometries of $\mathbf{H}^2$ is $\mathrm{PSL}_2(\mathbb{R}) := \mathrm{SL}_2(\mathbb{R})/\{\pm 1\}$, acting via linear fractional transformation.

The subgroup $\mathrm{SL}_2(\mathbb{Z}) \leq \mathrm{SL}_2(\mathbb{R})$ is discrete and therefore the quotient $\mathrm{PSL}_2(\mathbb{Z}) := \mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$ is a Fuchsian group [Voi21, §§34.2, 34.7], acting properly on $\mathbf{H}^2$. The matrices

$$S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \tag{2.1.3}$$

generate $\mathrm{SL}_2(\mathbb{Z})$ with $S^2 = (ST)^3 = -1$ a minimal set of relations.

A fundamental set for the action of $\mathrm{PSL}_2(\mathbb{Z})$ on $\mathbf{H}^2$ is

$$\text{¤} := \{z \in \mathbf{H}^2 : |\operatorname{Re} z| \leq 1/2 \text{ and } |z| \geq 1\}. \tag{2.1.4}$$

2.1.5. Indeed, there is an explicit *reduction algorithm*: given as input $z \in \mathbf{H}^2$ (provided by a complex approximation or using ball arithmetic), it produces as output a word $\gamma \in \langle S, T\rangle$ such that $\gamma z \in$ ¤, as follows.

1. We translate z applying a power of T so that $|\operatorname{Re} z| \leq 1/2$.

2. If $|z| \geq 1$, we are done; otherwise, if $|z| < 1$, then we apply S and

$$\operatorname{Im}\left(\frac{-1}{z}\right) = \frac{\operatorname{Im} z}{|z|^2} > \operatorname{Im} z. \tag{2.1.6}$$

We then repeat this process, returning to step 1.

In this way, we obtain (from the second step) a sequence of elements $z = z_1, z_2, \ldots$ with $\operatorname{Im} z_1 < \operatorname{Im} z_2 < \ldots$. This process terminates after finitely many steps since the orbit of z under $\mathrm{SL}_2(\mathbb{Z})$ is discrete and therefore its intersection with the compact set

$$\{z' \in \mathbf{H}^2 : |\mathrm{Re}(z')| \leq 1/2 \text{ and } \operatorname{Im} z \leq \operatorname{Im} z' \leq 1\}$$

is finite.

The group $\{\pm 1\} \leq \mathrm{SL}_2(\mathbb{R})$ of scalar matrices acts trivially on the upper half-plane, so we also work with $\mathrm{PSL}_2(\mathbb{R}) := \mathrm{SL}_2(\mathbb{R})/\{\pm 1\}$ and similarly $\mathrm{PSL}_2(\mathbb{Z})$. Let $Y = \mathrm{PSL}_2(\mathbb{Z})\backslash\mathbf{H}^2$. Gluing together the fundamental set, we obtain a homeomorphism

$$Y = \mathrm{PSL}_2(\mathbb{Z})\backslash\mathbf{H}^2 \simeq \mathbb{P}^1(\mathbb{C}) \smallsetminus \{\infty\} \simeq \mathbb{C}. \tag{2.1.7}$$

2.1.8. We compactify Y, closing up the **cusp** ∞, as follows. The orbit of the limit point ∞ under $\mathrm{SL}_2(\mathbb{Z})$ is $\mathbb{P}^1(\mathbb{Q}) \subseteq \operatorname{bd}\mathbf{H}^2$. Let $\mathbf{H}^{2*} := \mathbf{H}^2 \cup \mathbb{P}^1(\mathbb{Q})$; then the action by linear fractional transformations extends to $\mathbb{P}^1(\mathbb{Q})$. We define a topology on $\mathbf{H}^{2*}$, called the **Satake topology**: a basis of open sets of ∞ consists of the sets $\{z \in \mathbf{H}^2 : \mathrm{Im}(z) > \sigma\}$ for $\sigma > 0$ and a basis of open sets at a cusp $c \in \mathbb{P}^1(\mathbb{Q})$ consists of open disks tangent to the real axis at c. This induces a topology on the quotient, giving a homeomorphism

$$X := \mathrm{PSL}_2(\mathbb{Z})\backslash\mathbf{H}^{2*} \simeq \mathbb{P}^1(\mathbb{C}). \tag{2.1.9}$$

Away from the orbits of i and ω with nontrivial stabilizer, the complex structure on $\mathbf{H}^2$ descends giving the structure of a Riemann surface. We prefer to keep track of stabilizers, in which case X has the structure of a good, compact complex 1-orbifold, in this case a stacky projective line.

2.1.10. The quotient Y has an interpretation as a moduli space of (oriented) complex lattices: the map

$$\begin{aligned} Y = \mathrm{PSL}_2(\mathbb{Z})\backslash\mathbf{H}^2 &\to \{\Lambda \subset \mathbb{C} \text{ lattice}\}/\sim \\ \mathrm{PSL}_2(\mathbb{Z})\tau &\mapsto [\mathbb{Z} + \mathbb{Z}\tau]; \end{aligned} \tag{2.1.11}$$

is a bijection, which is to say Y parametrizes oriented complex lattices up to homothety. To a lattice Λ, we associate the complex elliptic curve $\mathbb{C}/\Lambda$, thus the space Y also parametrizes isomorphism classes of complex elliptic curves. The compact X can be seen as parametrizing *generalized* elliptic curves: the elliptic curve $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$ via the map $q\colon z \mapsto \exp(2\pi i z)$ has image $\mathbb{C}^\times/\langle q(\tau)\rangle$ (Tate curve) so as $\tau \to \infty$ we get $\mathbb{C}^\times$.

2.2 Congruence subgroups

The setup in the previous section extends to finite-index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ play a central role, and of particular importance are those subgroups defined by congruence conditions on the entries.

Definition 2.2.1. Let $N \in \mathbb{Z}_{\geq 1}$. The **principal congruence subgroup** $\Gamma(N) \trianglelefteq \mathrm{SL}_2(\mathbb{Z})$ of **level** N is

$$\begin{aligned}\Gamma(N) &:= \ker(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})) \\ &= \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N}\right\}.\end{aligned}$$

2.2.2. By strong approximation for $\mathrm{SL}_2(\mathbb{Z})$ [Voi21, Theorem 28.2.6] (boiling down to the fact that elementary matrices generate, just like in linear algebra over a field), the reduction map $\pi_N \colon \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is surjective for all $N \geq 1$, so there is an exact sequence

$$1 \to \Gamma(N) \to \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) \to 1. \tag{2.2.3}$$

Definition 2.2.4. A subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ is a **congruence subgroup** if $\Gamma \geq \Gamma(N)$ for some $N \geq 1$; if so, the minimal such N is called the **level** of Γ.

2.2.5. We describe congruence subgroups algorithmically as follows. Let $K_N^1 \leq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ be a subgroup. Then the inverse image $\pi_N^{-1}(K_N^1)$ is a congruence subgroup of level dividing N. And conversely, if Γ has level N then $\Gamma = \pi_N^{-1}(K_N^1)$ where $K_N^1 = \pi_N(\Gamma)$).

By lookup/hash in $O(\log[\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) : K_N^1])$ time we can determine if $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ belongs to $\pi_N^{-1}(K_N^1)$. We could also construct coset table, and there are better methods for specific groups.

Remark 2.2.6. Sorry for the notation, but later when we work adelically, it will be important to take $K_N \leq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, and then $K_N^1 = H \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Other authors use either G or H, but we need these for a reductive group and homology and cohomology, and K will be our level structure!)

2.2.7. Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, and let $\mathrm{P}\Gamma := \Gamma/(\Gamma \cap \{\pm 1\})$. We define

$$Y = Y(\Gamma) := \mathrm{P}\Gamma \backslash \mathbf{H}^2. \tag{2.2.8}$$

The set of orbits of $\mathrm{P}\Gamma$ on $\mathbb{P}^1(\mathbb{Q}) \subseteq \mathrm{bd}\,\mathbf{H}^2$ is again finite, called the **cusps** of Γ. We therefore again compactify

$$X = X(\Gamma) := \mathrm{P}\Gamma \backslash \mathbf{H}^{2*}. \tag{2.2.9}$$

Again X has the structure of a good, compact complex 1-orbifold, with finitely many points with nontrivial stabilizer. (If we did not mod out by $\{\pm 1\}$, potentially every point would have additional stabilizer, and in certain contexts we may want to consider that as well! But we will need to get even stackier, and so the notation will get scriptier.)

2.2.10. In addition to the congruence groups $\Gamma(N)$ themselves, we will make use of two other important congruence subgroups for $N \geq 1$, which we call the **standard** congruence subgroups:

$$\begin{aligned}\Gamma_0(N) &:= \left\{\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \pmod{N}\right\} \\ &= \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{N}\right\} \\ \Gamma_1(N) &:= \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{N}\right\}\end{aligned} \tag{2.2.11}$$

Visibly, $\Gamma(N) \leq \Gamma_1(N) \leq \Gamma_0(N)$. We accordingly write

$$\begin{aligned}Y_0(N) &:= \Gamma_0(N)\backslash \mathbf{H}^2 \\ X_0(N) &:= \Gamma_0(N)\backslash \mathbf{H}^{2*}\end{aligned} \tag{2.2.12}$$

where $\mathbf{H}^{2*} := \mathbf{H}^2 \cup \mathbb{P}^1(\mathbb{Q})$, and similarly $Y_1(N) \subseteq X_1(N)$ and $Y(N) \subseteq X(N)$.

We will find it useful both to work with a general congruence subgroup since the theory and algorithms extend this far and because they are still interesting for applications, and to restrict to the standard congruence subgroups for concreteness especially in examples.

2.3 Modular forms

There are also a wealth of references for the classical modular forms, including Apostol [Apo90, Chapters 1–2], Diamond–Shurman [DS05], Miyake [Miy06, Chapter 4], Lang [Lang95, §1], and Serre [Ser73, Chapter VII]. For laziness we largely follow Voight [Voi21, Chapter 40].

Definition 2.3.1. Let $k \in \mathbb{Z}$ and let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a Fuchsian group. A map $f\colon \mathbf{H}^2 \to \mathbb{C} \cup \mathbb{P}^1(\mathbb{C})$ is **weight** k **invariant** for Γ if

$$f(\gamma z) = (cz + d)^k f(z) \quad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \text{ and all } z \in \mathbf{H}^2. \tag{2.3.2}$$

2.3.3. If f is weight k invariant and f' is weight k' invariant, then ff' is weight $k + k'$ invariant, and if $k' = k$ then $f + f'$ is weight k invariant. Therefore, the set of weight k-invariant functions has the structure of a $\mathbb{C}$-vector space and the set of all functions of some weight k forms a ring.

2.3.4. For $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ with $\det g > 0$ and $z \in \mathbf{H}^2$ we define the **automorphy factor**

$$\jmath(g, z) = cz + d;$$

and for $k \in \mathbb{Z}$ and a map $f\colon \mathbf{H}^2 \to \mathbb{C}$

$$f[g]_k(z) := (\det g)^{k-1} \jmath(g, z)^{-k} f(gz). \tag{2.3.5}$$

Sorry for all of that, at least when $g \in \mathrm{SL}_2(\mathbb{Z})$ the determinant disappears, and f is weight k-invariant for Γ if and only if $f[\gamma]_k(z) = f(z)$ for all $\gamma \in \Gamma$.

2.3.6. The automorphic factor $\jmath$ satisfies the **cocycle relation**

$$\jmath(gg'; z) = \jmath(g; g'z)\jmath(g'; z) \tag{2.3.7}$$

for all $g, g' \in \mathrm{GL}_2(\mathbb{R})_{>0}$. Hence $f(\gamma(\gamma' z)) = \jmath(\gamma\gamma'; z)^k f(z)$ for $\gamma, \gamma' \in \Gamma$, and so weight k invariance under Γ can be checked on a set of generators for Γ.

Since $\mathrm{PSL}_2(\mathbb{Z})$ is generated by S, T, it follows that a map f is weight k invariant for $\mathrm{PSL}_2(\mathbb{Z})$ if and only if

$$\begin{aligned} f(z+1) &= f(z) \\ f(-1/z) &= z^k f(z) \end{aligned} \tag{2.3.8}$$

hold for all $z \in \mathbf{H}^2$.

2.3.9. Let $f\colon \mathbf{H}^2 \to \mathbb{C}$ be a meromorphic map that is weight k invariant under a Fuchsian group $\Gamma \ni T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then $f(z+1) = f(z)$, so f admits a Fourier expansion in $q = \exp(2\pi i z)$

$$f(z) = \sum_{n=-\infty}^{\infty} a_n q^n. \tag{2.3.10}$$

If $a_n \in \mathbb{C}$ and $a_n = 0$ for all but finitely many $n < 0$, then we say that f is **meromorphic** at ∞; if further $a_n = 0$ for $n < 0$, we say f is **holomorphic at** ∞.

2.3.11. For what comes later, we record the following equivalent condition: a weight k invariant meromorphic function $f\colon \mathbf{H}^2 \to \mathbb{C}$ is holomorphic at ∞ if the $\mathrm{SL}_2(\mathbb{Z})$-invariant function $z \mapsto (\operatorname{Im} z)^{k/2}|f(z)|$ has at most polynomial growth in y as $y \to \infty$; we then say that f is of **moderate growth**.

2.3.12. More generally, let $\Gamma \leq \mathrm{PSL}_2(\mathbb{Z})$ be a subgroup of finite index. For $\gamma \in \mathrm{PSL}_2(\mathbb{Z})$, we define

$$f[\gamma]_k(z) := \jmath(\gamma; z)^{-k} f(\gamma z). \tag{2.3.13}$$

Then $f[\gamma]_k(z)$ is weight k invariant under the group $\gamma^{-1}\Gamma\gamma$. We say that f is **meromorphic at the cusps** if for every $\gamma \in \mathrm{PSL}_2(\mathbb{Z})$, the function $f[\gamma]_k$ is meromorphic at ∞, in the above sense. Since the set of cusps is finite, in fact it suffices to consider finitely many such representatives.

Finally, we say that f is **holomorphic at the cusps** if $f[\gamma]_k(z)$ is holomorphic at ∞ for all $\gamma \in \mathrm{PSL}_2(\mathbb{Z})$, and **vanishes at the cusps** if $f[\gamma]_k(\infty) = 0$ for all γ.

Definition 2.3.14. A **(holomorphic) modular form of weight** k is a holomorphic map $f\colon \mathbf{H}^2 \to \mathbb{C}$ that is weight k invariant under Γ and holomorphic at the cusps. A **cusp form of weight** k is a holomorphic modular form of weight k that vanihses at the cusps.

If $-1 \in \Gamma$, then the only holomorphic modular form of odd weight k is 0.

Let $M_k(\Gamma)$ be the $\mathbb{C}$-vector space of modular forms of weight k for Γ, and let $S_k(\Gamma) \subseteq M_k(\Gamma)$ be the subspace of cusp forms.

2.4 Modular forms as sections of a line bundle

In this section, we reinterpret $M_k(\Gamma)$ and $S_k(\Gamma)$ as the $\mathbb{C}$-vector space of sections of a line bundle on $X(\Gamma)$, and conclude that these spaces are finite-dimensional. This is an important guide to the generalization of modular forms we pursue in these notes.

2.4.1. Since

$$\frac{\mathrm{d}(\gamma z)}{\mathrm{d}z} = \frac{1}{(cz+d)^2} \tag{2.4.2}$$

the weight k invariance (2.3.2) of a map f for $k \in 2\mathbb{Z}_{\geq 0}$ can be rewritten

$$f(\gamma z)\,\mathrm{d}(\gamma z)^{\otimes k/2} = f(z)\,\mathrm{d}z^{\otimes k/2} \tag{2.4.3}$$

so equivalently, the differential $f(z)\,\mathrm{d}z^{\otimes k/2}$ is (straight up) *invariant* under Γ, and that kind of justifies the name.

Although such forms are globally defined, we can make the same definition locally, packing them together into the line bundle $\Omega_Y^{k/2}$ of holomorphic $k/2$-forms over $Y = Y(\Gamma)$ when k is even. In other words, the weight k invariance for modular forms pops out from the condition of being a differential $k/2$-form. When we compactify to $X = X(\Gamma)$, the condition of being holomorphic at ∞ concerns the q-expansion: since $\mathrm{d}q = 2\pi i q\,\mathrm{d}z$, we check that modular forms correspond to $k/2$-differential forms with a pole of order at most 1 at the cusps, and cusp forms are exactly those which are holomorphic at ∞, giving global sections of $\Omega_X^{k/2}$. This perspective also explains why we naturally "start" at weight $k \geq 2$ and continue most easily with even weights. To summarize, for $k \geq 2$ even, we have

$$\begin{aligned} M_k(\Gamma) &\cong H^0(X(\Gamma), \Omega^{k/2}(\Delta)) \\ S_k(\Gamma) &\cong H^0(X(\Gamma), \Omega^{k/2}) \end{aligned} \tag{2.4.4}$$

where Δ is the divisor given by the sum of the cusps.

We can go a bit further to handle odd weights (when $-1 \notin \Gamma$ otherwise there are none): we need somehow to extract a "square root" of the bundle $\Omega^1_{X(\Gamma)}$ of holomorphic 1-forms. There is such a bundle ω_X over X, called the Hodge bundle, defined as follows. We let $\pi\colon \mathcal{E} \to X$ be the universal elliptic curve over

X: this exists as a complex orbifold and as a complex manifold if Γ is torsion-free. Then the **Hodge bundle** is the pushforward $\omega := \pi_* \Omega^1_{\mathcal{E}/X}$: the fiber at a point $[E]$ is $H^0(X, \Omega^1_E)$ the one-dimensional space of invariant differentials on the modular curve. The Kodaira–Spencer map gives an isomorphism

$$\Omega^1_{X(\Gamma)}(\Delta) \simeq \omega^{\otimes 2}. \tag{2.4.5}$$

This reinterpretation has several immediate consequences. First, it implies that the spaces $M_k(\Gamma)$ and $S_k(\Gamma)$ are finite-dimensional $\mathbb{C}$-vector spaces: indeed, the Riemann–Roch theorem then provides their dimensions. Again, this makes sense on $X(\Gamma)$ as an orbifold (see Voight–Zureick-Brown [VZB15, Chapters 5–6]) or is literally true if Γ is torsion free. In this case, we can eek out something further for odd weights, but in general this will guide our conceptual template for generalizations of modular forms: we will have sections of natural vector bundles on arithmetic quotients, with analytic conditions reinterpreted as extension or vanishing properties along boundary components.

2.5 Eisenstein series

Key examples of modular forms coming from averaging at the cusp ∞, as follows. We first consider the case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$.

Let $\Gamma_\infty \leq \Gamma = \mathrm{SL}_2(\mathbb{Z})$ be the stabilizer of ∞; then Γ_∞ is the extension of the infinite cyclic group generated by $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ by $\{\pm 1\}$. We consider the cosets $\Gamma_\infty \backslash \Gamma$: for $t \in \mathbb{Z}$ and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we have $T^t \gamma = \begin{pmatrix} a + tc & b + td \\ c & d \end{pmatrix}$ with the same bottom row. Thus the function $(cz + d)^2$ is well-defined on the coset $\Gamma_\infty \gamma$. Thus we can form the sum

$$E_k(z) = \sum_{\Gamma_\infty \gamma \in \Gamma_\infty \backslash \Gamma} (cz + d)^{-k} = \frac{1}{2} \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1}} \frac{1}{(cz+d)^k} \tag{2.5.1}$$

the factor 2 coming from the choice of sign. Since every nonzero $(m, n) \in \mathbb{Z}^2$ can be written $(m, n) = r(c, d)$ with $r = \gcd(m, n) > 0$ and $\gcd(c, d) = 1$, we find that

$$G_k(z) = 2\zeta(k) E_k(z)$$

where $\zeta(k) = \sum_{n=1}^{\infty} n^{-k}$ and

$$G_k(z) := \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \frac{1}{(m + nz)^k}. \tag{2.5.2}$$

The series $E_k(z)$ and $G_k(z)$ converges absolutely, and we call them **Eisenstein series** of weight k.

Lemma 2.5.3. *The Eisenstein series $E_k(z)$ is a holomorphic modular form of weight $k \in 2\mathbb{Z}_{\geq 2}$ for $\mathrm{SL}_2(\mathbb{Z})$ with Fourier expansion*

$$E_k(z) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n \tag{2.5.4}$$

where $B_k \in \mathbb{Q}$ are the Taylor coefficients of

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!} = 1 - \frac{1}{2}x + \frac{1}{6}\frac{x^2}{2!} - \frac{1}{30}\frac{x^4}{4!} + \frac{1}{42}\frac{x^6}{6!} + \dots$$

and

$$\sigma_{k-1}(n) = \sum_{\substack{d|n \\ d>0}} d^{k-1}.$$

The coefficients B_k are called **Bernoulli numbers**.

It is often preferable to normalize Eisenstein series to be primitive with integer coefficients (when working with series rings over $\mathbb{Z}$), or to take the coefficient of q^1 to be equal to 1 (when looking at congruences). Sorry about that!

2.6 Hecke operators

If modular forms are functions on lattices, then Hecke operators are obtained by summing functions over "neighboring" lattices, namely those of a fixed index. This provides a good intuition for Hecke operators; we then show how to formulate it in terms of double cosets, then providing a general definition.

Let $n \geq 1$. Let $\mathrm{M}_2(\mathbb{Z})_n$ be the set of integer matrices with determinant n. Then by the theory of elementary divisors,

$$\mathrm{GL}_2(\mathbb{Z})\,\mathrm{M}_2(\mathbb{Z})_n\,\mathrm{GL}_2(\mathbb{Z}) = \bigsqcup_{\substack{a|d \\ ad=n}} \mathrm{GL}_2(\mathbb{Z}) \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \mathrm{GL}_2(\mathbb{Z}) \tag{2.6.1}$$

This arises as part of the fundamental theorem of finitely generated modules over a PID, here just the case over $\mathbb{Z}$. We interpret the actions of $\mathrm{GL}_2(\mathbb{Z})$ on left and right as acting by row and column operations. Indeed, given a matrix in $\mathrm{M}_2(\mathbb{Z})_n$ for a computable PID (e.g., a computable Euclidean domain) there is an explicit algorithm to reduce to the above representative, called **Smith normal form**.

In fact, this algorithm only uses elementary matrices (which have determinant 1) to get to a diagonal matrix; so just also allowing the scalar matrix -1, in the same breath we get

$$\mathrm{SL}_2(\mathbb{Z})\,\mathrm{M}_2(\mathbb{Z})_n\,\mathrm{SL}_2(\mathbb{Z}) = \bigsqcup_{\substack{a|d \\ ad=n}} \mathrm{SL}_2(\mathbb{Z}) \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \mathrm{SL}_2(\mathbb{Z}). \tag{2.6.2}$$

A matrix $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z})$ is **primitive** if $\gcd(a,b,c,d) = 1$. Let $\mathrm{M}_2(\mathbb{Z})_{n,\mathrm{prim}} \subseteq \mathrm{M}_2(\mathbb{Z})_n$ be the subset of primitive matrices, so that

$$\mathrm{M}_2(\mathbb{Z})_n = \bigsqcup_{d|n} d\,\mathrm{M}_2(\mathbb{Z})_{n/d^2,\mathrm{prim}}.$$

Next, if instead we only act on the *left* by $\mathrm{SL}_2(\mathbb{Z})$, then we only get *row* operations, and we can only reduce a matrix to **Hermite normal form**, i.e.,

$$\mathrm{SL}_2(\mathbb{Z})\,\mathrm{M}_2(\mathbb{Z})_{n,\mathrm{prim}}\,\mathrm{SL}_2(\mathbb{Z}) = \bigsqcup_{\alpha \in \Theta_n} \alpha\,\mathrm{SL}_2(\mathbb{Z}) \tag{2.6.3}$$

where

$$\Theta_n := \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z}) : ad = n, \gcd(a,d) = 1, 0 \le b < d \right\}. \tag{2.6.4}$$

Then Θ_n is in bijection with the set $\{\Lambda \subseteq \mathbb{Z}^2 : \mathbb{Z}^2/\Lambda \simeq \mathbb{Z}/N\mathbb{Z}\}$. This matches cyclic N-isogenies for elliptic curves.

Example 2.6.5. For $n = p$ prime, Θ_p consists of the $p+1$ matrices

$$\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & p \end{pmatrix}, \dots, \begin{pmatrix} 1 & p-1 \\ 0 & p \end{pmatrix}.$$

For $\theta \in \Theta_n$, the row span of θ defines a lattice of index n in $\mathbb{Z}^2$ and conversely every such lattice occurs uniquely this way.

With this motivation in mind, we now make general definitions: for more, see Diamond–Shumran [DS05, Chapter 5]. Let $\Gamma \le \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, and write $\mathrm{GL}_2(\mathbb{Q})_{>0}$ for the subgroup of matrices of positive determinant.

Lemma 2.6.6. *For each $\beta \in \mathrm{GL}_2(\mathbb{Q})_{>0}$, the double coset $\Gamma\beta\Gamma$ is a disjoint union of cosets*

$$\Gamma\beta\Gamma = \bigsqcup_{\alpha \in \Theta(\beta)} \Gamma\alpha \tag{2.6.7}$$

with $\#\Theta(\beta) < \infty$.

Proof. This is most conveniently seen by thinking adelically, but can also be proved in a direct and somewhat less conceptual manner! □

Example 2.6.8. When Γ is a standard congruence subgroup, for $\beta = \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix}$ we have again that $\Theta(\beta) = \Theta_n$.

Definition 2.6.9. For $\alpha \in \mathrm{GL}_2(\mathbb{Q})_{>0}$, the **Hecke operator** T_β acting on the space of weight k modular forms $M_k(\Gamma)$ is the operator defined by

$$\begin{aligned} M_k(\Gamma) &\to M_k(\Gamma) \\ T_\beta(f) &= \sum_{\alpha \in \Theta(\beta)} f[\alpha]_k \end{aligned} \tag{2.6.10}$$

Example 2.6.11. For Γ a standard congruence subgroup and $f \in \mathrm{M}_k(\Gamma)$, we have

$$(T_n f)(z) = \sum_{\alpha = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in \Theta_n} d^{-2} f\left(\frac{az+b}{d}\right).$$

The Hecke operators are well-defined independent of the choice of representatives in $\Theta(\beta)$ precisely because they are weight k invariant. They pairwise commute and satisfy a natural recursion. They preserve the space of cusp forms and commute with the Petersson inner product (so preserve the Eisenstein subspace); they generate a commutative, semisimple subalgebra of $\mathrm{End}_{\mathbb{C}} M_k(\Gamma)$ and so by the spectral theorem this space decomposes into eigenspaces; an element in an eigenspace is called a **eigenform**. We say an eigenform f is **normalized** if $a_1(f) = 1$.

Lemma 2.6.12. *If Γ is a standard congruence subgroup and $f(q) = \sum_{n=0}^{\infty} a_n q^n$ is a normalized eigenform, then $f|T_n = a_n f$.*

Proof. We have

$$(T_m f)(z) = \sum_{n=1}^{\infty} \sum_{d \mid \gcd(m,n)} d a_{mn/d^2} q^n$$

so for example

$$(T_p f)(z) = \sum_{n=1}^{\infty} (a_{np} - p a_{n/p}) q^n$$

for p prime where $a_{n/p} = 0$ if $p \nmid n$; so if $(T_m f) = c_m f$ then from $a_1 = 1$ we conclude $c_m = a_m$. This is still OK for normalized Eisenstein series! □

Remark 2.6.13. We lose the property that the q-expansion can be read off immediately from its eigenvalues for a general congruence subgroup, but it is still possible to recover its coefficients by considering twists.

Remark 2.6.14. We revisit the definition of Hecke operators for a general congruence subgroup by thinking adelically in section 9.4.

2.7 Newforms and oldforms

2.8 Graded ring and equations

A complete description of the ring of (holomorphic) modular forms for $\mathrm{SL}_2(\mathbb{Z})$. By 2.3.3, the $\mathbb{C}$-vector space

$$M_*(\Gamma) = \bigoplus_{k\in\mathbb{Z}} M_k(\Gamma)$$

under multiplication has the structure of a graded $\mathbb{C}$-algebra; we call $M(\Gamma)$ the **graded ring of modular forms** for Γ.

2.9 The computational problem

Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup and let $k \in \mathbb{Z}_{\geq 0}$. By *computing modular forms* for Γ, we mean the following suite of tasks:

1. compute the dimensions $\dim M_k(\Gamma)$, $\dim E_k(\Gamma)$, and $\dim S_k(\Gamma)$, the decomposition into Hecke irreducible subspaces;
2**. compute the Hecke action on $M_k(\Gamma)$;
3. extract Hecke eigensystems using linear algebra;
4. perform a decomposition into oldforms and newforms; and
5. write out q-expansions.

In what follows, we will focus on the case $k \geq 2$; but we can ask and pursue the algorithmic question in general.

Step 2 has a double star because it is the big one, and to a large extent the rest of the text zooms into making this step precise (itself requiring many further steps). In this step, sometimes we are not interested in Eisenstein series, in which case we restrict the second step to the cuspidal subspace. In any event, the *Hecke action* means for each p to give a matrix $[T_p]$ representing the action of T_p (in a fixed basis of a vector space). We may content ourselves with only partial data, for example just the traces $\mathrm{Tr}[T_p]$.

Step 3 is one that we will see in generalizations is almost universal: given commuting normal operators on a finite-dimensional complex vector space, compute a common basis of eigenvectors. In truth, the linear algebra aspects are often the bottleneck in our computations, so a lot of engineering goes into refining the precise linear algebra questions and then adapting and optimizing the algorithms to work efficiently in practice.

Before we write out q-expansions in step 5, we are only dealing with Hecke data (sequences of numbers)—an idea we will carry as far as we can in our generalizations. For standard congruence subgroups, step 5 is immediate by Lemma 2.6.12, but for a general congruence subgroup more is required.

The computations do not stop here! Given a modular form, we may also want to compute other interesting data about it, including the associated Galois representations and L-functions. From the perspective taken in these notes,

we consider this a call to action and save for another time their algorithmic elaboration.

Chapter 3

Modular symbols

Having laid the basic notation and setup for modular curves in the previous section, here we show how to compute classical modular forms using modular symbols.

Modular forms are dual to homology by the integration pairing, and it is indeed easier to work with homology: we can construct (relative) one-dimensional cycles as paths on the modular curve as Riemann surface whose endpoints lie on the set of cusps. From this one can systematically work with the relative homology group and then the action of Hecke operators, from which we can hope to recover modular forms.

The rough sketch is as follows:

1. we find a nice basis in relative homology,
2. we take the kernel with respect to a boundary operator,
3. then we apply Hecke operators, and the crux is reducing the result in homology to the given basis.

3.1 Farey symbols

Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup—in fact, the algorithm works for an arbitrary finite index subgroup. For concreteness and because it is some nice number theory, we begin by discussing how to compute a fundamental domain for the action of $\Gamma \circlearrowright \mathbf{H}^{2*}$ with a reduction theory, adapted from 2.1.5.

We will make fundamental domains drawing edges between cusps and elliptic points, like $\mathrm{SL}_2(\mathbb{Z})$.

The original **Farey fractions** of order $n \in \mathbb{Z}_{\geq 1}$ is set of rational numbers

$$\{0 \leq p/q \leq 1 : \gcd(p,q) = 1, q < n\} \tag{3.1.1}$$

and the **Farey sequence** just writes the Farey fractions in increasing order. For example: $0/1, 1/1$; $0/1, 1/2, 1/1$, $0/1, 1/3, 1/2, 2/3, 1/1$, etc. We notice that the sequence of order $n+1$ is obtained from order n by inserting a unique frac-

tion in between two **Farey neighbors**, namely the "frosh's sum of fractions"

$$\frac{a}{b} F \frac{c}{d} = \frac{a+c}{b+d}$$

and more properly called the **mediant**. This is a fun way to algorithmically generate all rational numbers by increasing height, never computing a gcd!

Lemma 3.1.2. *The following statements hold.*

(a) *If $a/b < c/d$ are Farey fractions of order n, then they are Farey neighbors if and only if $bc - ad = 1$.*
(b) *If $a/b < c/d$ are Farey neighbors and $h/k = (a/b) F (c/d)$, then $a/b < h/k < c/d$ and $bh - ak = ck - dh = 1$.*

Proof. As a hint for (b), we have

$$\frac{a+c}{b+d} - \frac{a}{b} = \frac{bc-ad}{b(b+d)}$$

and

$$\frac{c}{d} - \frac{a+c}{b+d} = \frac{bc-ad}{d(b+d)} > 0;$$

and

$$k = (bc-ad)k = b(ck-dh) + d(bh-ak)$$

from which the result can be derived. □

This is also clearly related to continued fractions in a lovely way, either say more here or maybe peek at Lemma 3.5.1! And if we draw geodesics in the upper half-plane between Farey neighbors, we get *Ford circles*.

Now our fundamental domain algorithm is due to Kulkarni [Kul91], following Kurth–Long [KL08].

Definition 3.1.3. A **generalized Farey sequence** is a sequence of rationals

$$x_0 = \frac{a_0}{b_0} < \cdots < \frac{a_n}{b_n} = x_n \tag{3.1.4}$$

with consecutive terms $a_i/b_i < a_{i+1}/b_{i+1}$ satisfying

$$b_i a_{i+1} - a_i b_{i+1} = 1. \tag{3.1.5}$$

We include formally $-\infty = \frac{-1}{0}$ and $\infty = \frac{1}{0}$.

Definition 3.1.6. A **Farey symbol** is a generalized Farey sequence starting at $-\infty$, ending at ∞, together with neighbors labelled with one of the following:

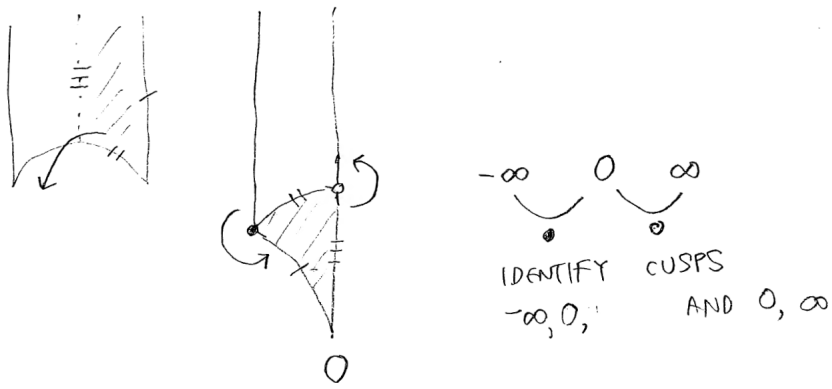
- free, labelled with a positive integer;
- even, labelled with •;
- odd, labelled with ∘.

and such that each free integer appears twice.

There is an algorithm that takes as input $\mathrm{P}\Gamma$ (in this presentation, coming from $K_N^1 \leq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$) and returns as output a Farey symbol such that:

(i) The x_i represent the cusps of $\mathrm{P}\Gamma$, and
(ii) The interior of the Ford circles drawn by neighbors (with some fiddling at odd vertices) is a star-like fundamental domain with Ford circles as boundary, and
(iii) Gluing together the fundamental domain by identifying oriented free edges yields the quotient $X(\Gamma)$.

Example 3.1.7. For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ so $\mathrm{P}\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, we have the Farey sequence $-\infty <_\bullet 0 <_\circ \infty$:



The algorithm proceeds roughly as follows. Working out the case of index 2 (Exercise 3.1), we may suppose that $[\mathrm{PSL}_2(\mathbb{Z}) : \mathrm{P}\Gamma] \geq 3$.

1. Initialize: if $\pm ST = \pm \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{P}\Gamma$, start with with $-\infty < 0 < 1 < \infty$ (unlabelled); else start with $-\infty < -1 < 0 < \infty$.
2. Choose an unlabelled edge $\frac{a_i}{b_i}, \frac{a_{i+1}}{b_{i+1}}$ be unlabeled. See if it can be non-trivially paired with itself (even or odd according to the order of the element) or another unlabelled $\frac{a_j}{b_j}, \frac{a_{j+1}}{b_{j+1}}$: check if the side pairing matrix

$$\gamma = \begin{pmatrix} a_{j+1}b_{i+1} + a_j b_i & -a_j a_i - a_{j+1}a_{i+1} \\ b_j b_i + b_{j+1}b_{i+1} & -a_{i+1}b_{j+1} - a_i b_j \end{pmatrix}$$

has $\gamma \in \Gamma$. If so, label accordingly. If not, split

$$\frac{a_i}{b_i}, \frac{a_i + a_{i+1}}{b_i + b_{i+1}}, \frac{a_{i+1}}{b_{i+1}}$$

(still unlabelled) and repeat.

3. Return when there are no unlabelled sides.

Example 3.1.8. We execute this algorithm for $\Gamma = \Gamma(2)$ the principal congruence subgroup. $-\infty, 0, \infty$. Check $\frac{0}{1} \sim \frac{1}{1}$: $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \notin \Gamma(2)$. Check $\frac{1}{1} \sim \frac{1}{0}$: $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \notin \Gamma(2)$. So we split! $-\infty, 0, 1, \infty$.

Continuing in this way, we get

$$-\infty <_1 -1 <_2 0 <_2 1 <_1 \infty$$

with sidepairing elements labelled 1: $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and 2: $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$. The cusps are then represented by $0, 1, \infty$.

This algorithm also gives a wealth of other information about $X(\Gamma)$: the number of cusps, generators for Γ (as side-pairing elements), the number of elliptic points (the number of $\circ$ and $\bullet$ labels, respectively), the width of the cusps, etc. There is also an explicit reduction theory similar to what we had over $\mathrm{SL}_2(\mathbb{Z})$.

3.2 Coset representatives

It will always take time at least linear in the $\mathrm{SL}_2(\mathbb{Z})$-index of Γ to write out the cosets. Looking ahead to Hecke operators, we also want an algorithm that that takes $\delta \in \mathrm{SL}_2(\mathbb{Z})$ and computes (very efficiently!) an element $\gamma \in \Gamma$ and the coset represenative δ_j such that $\delta = \gamma\delta_j$.

The Farey symbol for Γ provides geometrically meaningful coset representatives for $\Gamma\backslash \mathrm{SL}_2(\mathbb{Z})$. Let Δ be the hyperbolic triangle with vertices i, ω, ∞. (Together with its complex conjugate, the dark side, they form the usual fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$.) By construction, Δ is contained in the fundamental polygon $\square$ constructed from the Farey symbol, and the set of $\gamma \in \Gamma$ such that $\gamma\Delta \in \square$ are a set of coset representatives for Γ. These can be read off from the Farey symbol, a first approximation is that they are γ_i^{-1} where $\gamma_i = \begin{pmatrix} a_i & a_{i+1} \\ b_i & b_{i+1} \end{pmatrix}$ for $x_i = a_i/b_i < x_{i+1} = a_{i+1}/b_{i+1}$ in the sequence; but we also have to add the stabilizers at each cusp and fiddle with odd vertices.

For Γ coming from $K_N^1 \leq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, we can apply the the Todd–Coxeter algorithm.

3.2.1. For $\Gamma = \Gamma_0(N)$, the map

$$\begin{aligned} \Gamma\backslash \mathrm{SL}_2(\mathbb{Z}) &\to \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}) \\ \Gamma \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto (c : d) \end{aligned} \tag{3.2.2}$$

is a bijection. Algorithmically, this gives a clean way to enumerate coset representatives and to store them in a hash table for fast reduction.

3.3 Modular symbols in homology

Modular symbols were introduced by Manin [Man72] and developed by Merel [Mer94], and they were turned into a practical computational tool in work of Cremona [Cre97] and Stein [Ste07], as well as others. We give a homological definition that is well-suited for algorithms.

Just as in the previous section, where Farey symbols as paths in the upper half-plane described the boundary of a fundamental domain, here we will consider the resulting homology: we will find a presentation for $H_1(X(\Gamma), \mathbb{Z})$ in terms of paths between cusps, and then we will use an action of Hecke operators to see systems of Hecke eigenvalues. This will only give us access to modular forms of weight $k \geq 2$.

We work abstractly, axiomatizing how these paths should behave, for algorithmic purposes; we then prove that this correctly computes homology.

3.3.1. Let $\mathcal{D}$ be the free abelian group on ordered pairs (x, y) with $x, y \in \mathbb{P}^1(\mathbb{Q})$. The group $\mathrm{GL}_2(\mathbb{Q})$ acts on $\mathcal{D}$ by

$$\gamma \cdot (x, y) = (\gamma x, \gamma y)$$

where γ acts on $\mathbb{P}^1(\mathbb{Q}) = \mathrm{bd}\, \mathbf{H}^{2*}$ by linear fractional transformations. Thinking of paths from x to y, let $\mathcal{R} \leq \mathcal{D}$ be the subgroup generated by the relations

$$\begin{aligned} (x, y) + (y, z) + (z, x) &= 0 \\ (x, x) &= 0 \end{aligned} \tag{3.3.2}$$

for all $x, y, z \in \mathbb{P}^1(\mathbb{Q})$. The resulting quotient $\mathcal{M}_2 := \mathcal{D}/\mathcal{R}$ is the **universal module** of modular symbols; the equivalence class of (x, y) in the quotient is denoted $\{x, y\}$. In particular, we have $\{y, x\} = -\{x, y\}$ (paths are oriented).

Remark 3.3.3. Most of the time we further take the quotient by the torsion subgroup of $\mathcal{M}_2$, presumably caused by the elliptic points? For the purposes of modular forms in characteristic zero these are in the way, for example in the integration pairing they will vanish, but it may be useful to keep them around for a bit. Otherwise, it seems like we should be honest about it and replace coefficients by $\mathbb{Q}$ or $\mathbb{Z}[1/6]$?

Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. Then for $\gamma \in \Gamma$, the path from x to y and its image under γ are identified!

Definition 3.3.4. The **modular symbols module** for Γ is the Γ-coinvariants

$$\mathcal{M}_2(\Gamma) := (\mathcal{D}/\mathcal{R})_\Gamma \tag{3.3.5}$$

i.e., the quotient of $\mathcal{M}_2(\Gamma)$ by the submodule generated by

$$\{m - gm : m \in \mathcal{M}_2, g \in \Gamma\}.$$

We write $\{x, y\}_\Gamma$ for the image of $\{x, y\}$ in $\mathcal{M}_2(\Gamma)$ when we need to distinguish the two, but we will usually drop the subscript.

3.3.6. Homology is about closed loops, not paths! So time to close the loop—with the boundary map.

Let $\mathcal{B}_2(\Gamma)$ be the free abelian group on the set of cusps of Γ, the Γ-orbits in $\mathbb{P}^1(\mathbb{Q})$. Define

$$\partial_2 \colon \mathcal{M}_2(\Gamma) \to \mathcal{B}_2(\Gamma), \qquad \partial_2\{x, y\} = (y) - (x). \tag{3.3.7}$$

The given relations ensure that ∂_2 is well-defined (Exercise 3.3).

Remark 3.3.8. We could also define the boundary map $\partial_2 \colon \mathcal{M}_2 \to \mathcal{B}_2(\Gamma)$, and it would be the group of degree 0 divisors on $\mathbb{P}^1(\mathbb{Q})$.

Definition 3.3.9. The kernel

$$\mathcal{S}_2(\Gamma) := \ker \partial_2$$

is the space of **cuspidal modular symbols**.

The group of cuspidal modular symbols consists of linear combinations of paths in $\mathbf{H}^{2*}$ whose endpoints are in $\operatorname{bd} \mathbf{H}^2 = \mathbb{P}^1(\mathbb{Q})$ and whose images in $X(\Gamma)$ are linear combinations of loops.

Proposition 3.3.10 (Manin)**.** *The natural map ("take the path in homology") provides canonical group isomorphisms*

$$\mathcal{M}_2(\Gamma) \cong H_1(X(\Gamma), \operatorname{bd} X(\Gamma); \mathbb{Z})$$
$$\mathcal{S}_2(\Gamma) \cong H_1\big(X(\Gamma), \mathbb{Z}\big).$$

Proof. See Manin [Man72, Theorem 1.9]. In a nutshell, we get a relative homology group via a relative chain complex. □

Since the Farey neighbors describe the boundary of $X(\Gamma)$, a generating set for $\mathcal{M}_2(\Gamma)$ is given by the symbols $\{x_i, x_{i+1}\}$ for $x_i < x_{i+1}$ Farey neighbors! And a basis is obtained by taking the quotient by the relations specified by the labels in the Farey symbol for Γ. Then a basis for $\mathcal{S}_2(\Gamma)$ can be computed using straightforward linear algebra.

We extend these definitions to more general coefficients and weights as follows.

Let R be a (computable) commutative ring, let $k \geq 2$, and let $W_k(R) := \operatorname{Sym}^{k-2}(R) = R[X, Y]_{k-2}$ be the binary forms of degree $k - 2 \geq 0$ with coefficients in R. Then $W_k(R)$ has a natural left action by $\mathrm{SL}_2(\mathbb{Z})$, namely if $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $v \in W_k(R)$ then

$$(gv)(x, y) = v(g^{-1}(x, y)) = v(dX - bY, -cX + aY).$$

(Sorry about the left versus right, this is an inevitable consequence of the choice $(fg)(x)$ is first g then f!) Let

$$\mathcal{M}_k(R) := W_k(R) \otimes \mathcal{M}_2. \tag{3.3.11}$$

We abbreviate $\mathcal{M}_k := \mathcal{M}_k(\mathbb{Z})$. We extend the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{M}_k(R)$ by its natural action on each component

$$\gamma(v \otimes \{x, y\}) := \gamma v \otimes \{\gamma x, \gamma y\}$$

for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $v \otimes \{x, y\} \in \mathcal{M}_k(R)$, and extending linearly.

We similarly define

$$\mathcal{B}_k(R) := W_k(R) \otimes \mathcal{B}_2 \tag{3.3.12}$$

with the natural action $\gamma(v \otimes (x)) = \gamma(v) \otimes (\gamma x)$; and the boundary map

$$\begin{aligned} \partial_k \colon \mathcal{M}_k(R) &\to \mathcal{B}_k(R) \\ v \otimes \{x, y\} &\mapsto v \otimes y - v \otimes x. \end{aligned} \tag{3.3.13}$$

Definition 3.3.14. The space of **modular symbols of weight** k **over** R is the Γ-coinvariants

$$\mathcal{M}_k(\Gamma, R) := \mathcal{M}_k(R)_\Gamma.$$

The space of **cuspidal modular symbols of weight** k **over** R is the kernel

By naturality, we have $\mathcal{M}_k(\Gamma, Z) \otimes_{\mathbb{Z}} R \cong \mathcal{M}_k(\Gamma, R)$; and from Manin's theorem (Proposition 3.3.10) and universal coefficients or something, we have

$$\mathcal{M}_k(\Gamma, R) \cong H_1(X(\Gamma), \operatorname{bd} X(\Gamma); \operatorname{Sym}^{k-2} R^2)$$

and similarly $\mathcal{S}_k(\Gamma, R) \cong H_1(X(\Gamma); \operatorname{Sym}^{k-2} R^2)$.

3.4 Integration pairing

In weight 2, cusp forms are exactly holomorphic differentials on $X(\Gamma)$. As we saw in section 2.4, if $f \in S_2(\Gamma)$ then the differential $f(z)\,\mathrm{d}z$ is Γ-invariant (2.4.3) and extends holomorphically across the cusps.

3.4.1. One of the good thing about a differential is that you can integrate against them! In general, if you have a compact Riemann surface X of genus g and $H^0(X, \Omega^1)$ is the g-dimensional $\mathbb{C}$-vector space of holomorphic differential 1-forms on X, then there is a well-defined integration pairing

$$\begin{aligned} H^0(X, \Omega^1) \times H_1(X, \mathbb{Z}) &\to \mathbb{C} \\ \langle \omega, v \rangle &= \int_v \omega. \end{aligned} \tag{3.4.2}$$

that is right nondegenerate: if $\langle \omega, v \rangle = 0$ for all $v \in H_1(X, \mathbb{Z})$ then $\omega = 0$. But $H_1(X, \mathbb{Z})$ is a free $\mathbb{Z}$-module of rank $2g$, not g, so extending scalars to $\mathbb{C}$ could not give a perfect pairing. We could start by extending scalars to $\mathbb{R}$, but we are missing the *anti*-holomorphic differentials; once we add those, the pairing becomes perfect.

Remark 3.4.3. We can interpret this in terms of de Rham cohomology $H^1_{\mathrm{dR}}(X, \mathbb{C})$ and the Hodge filtration $H^{0,1} \oplus H^{1,0}$.

3.4.4. Returning to the context of modular forms and modular symbols, let

$$\overline{S_2(\Gamma)} = \{\overline{f} : f \in S_2(\Gamma)\}$$

be the complex conjugate of $S_2(\Gamma)$, with $\overline{f}(z) = \overline{f(z)}$; intrinsically, we can think of these as weight k-invariant anti-holomorphic functions which vanish at the cusps.

Then we again have an integration pairing

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon \left(S_2(\Gamma) \oplus \overline{S_2(\Gamma)}\right) \times \mathcal{S}_2(\Gamma) &\to \mathbb{C} \\ \langle (f, \overline{g}), \{x, y\} \rangle &= \int_x^y f(z)\,\mathrm{d}z + \int_x^y \overline{g}(z)\,\mathrm{d}\overline{z} \end{aligned} \tag{3.4.5}$$

where the integral is taken along any path (for example, the hyperbolic geodesic in $\mathbf{H}^2$) from x to y: Manin's theorem (Proposition 3.3.10, well-defined homology class) and then Cauchy's theorem (independent of choice of path, extended to anti-holomorphic functions) ensures that this is pairing is well-defined.

The integration pairing (3.4.5) is left and right nondegenerate. In particular, it gives a natural isomorphism

$$\mathrm{Hom}_{\mathbb{R}}(\mathcal{S}_2(\Gamma), \mathbb{R}) \xrightarrow{\sim} S_2(\Gamma) \oplus \overline{S_2(\Gamma)}.$$

The natural modification of this pairing also works in weight k.

Proposition 3.4.6. *The integration pairing*

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon \left(S_k(\Gamma) \oplus \overline{S_k(\Gamma)}\right) \times \mathcal{M}_k(\Gamma) &\to \mathbb{C} \\ \langle (f, \overline{g}), v \otimes \{x, y\} \rangle &= \int_x^y v(z,1) f(z)\,\mathrm{d}z + \int_x^y v(\overline{z},1)\overline{g}(z)\,\mathrm{d}\overline{z}. \end{aligned} \tag{3.4.7}$$

is well defined and right nondegenerate; its left kernel is $\mathcal{B}_k(\Gamma)$, giving a natural isomorphism

$$\mathrm{Hom}_{\mathbb{R}}(\mathcal{M}_k(\Gamma), \mathbb{R}) \xrightarrow{\sim} M_k(\Gamma) \oplus \overline{S_k(\Gamma)}. \tag{3.4.8}$$

Proof. See Merel [Mer94, §1.5], at least for $\mathcal{S}_k(\Gamma) \xrightarrow{\sim} S_k(\Gamma) \oplus \overline{S_k(\Gamma)}$. □

Remark 3.4.9. It should be better to normalize the pairing (3.4.5) by multiplying by $2\pi i$, so we are really integrating with respect to $\mathrm{d}q/q$; likely something similar for (3.4.7).

The conclusion: modular symbols are dual to modular forms!

Example 3.4.10. $\Gamma_0(11)$ has Farey symbol $-\infty <_1 0 <_2 \frac{1}{2} <_3 \frac{1}{3} <_3 1 <_1 \infty$, with side-pairing elements

$$\gamma_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \gamma_2 = \begin{pmatrix} 7 & -2 \\ 11 & -3 \end{pmatrix}, \qquad \gamma_3 = \begin{pmatrix} 8 & -3 \\ 11 & -4 \end{pmatrix}.$$

The cusps are $0 \overset{1}{\leftrightarrow} 1 \overset{3}{\leftrightarrow} 1/3 \overset{2}{\leftrightarrow} 1/2 \overset{3}{\leftrightarrow} 2/3$ and $\infty = -\infty$.

The space $\mathcal{M}_2(\Gamma)$ is spanned by $\{\infty, 0\}, \{0, 1/3\}, \ldots, \{1, \infty\}$ and has basis $\{\infty, 0\}, \{0, 1/3\}, \{1/3, 1/2\}$. The cuspidal space has basis $\{0, 1/3\}$, $\{1/3, 1/2\}$. (This might look weird, because we were supposed to have loops, right? But we just computed that the cusps $0, 1/3$ and $1/3, 1/2$ are identified!) We have $\operatorname{rk}_{\mathbb{Z}} \mathcal{S}_2(\Gamma) = 2 = 2g$, so $g(X_0(11)) = 1$.

3.5 Manin symbols and reduction

To compute with modular symbols, we need a basis and a reduction algorithm to that basis. The Farey sequence and its reduction algorithm provides such a method, but there is an *even more* efficient way of doing the reduction due to Manin which realizes the reduction theory as continued fractions.

Lemma 3.5.1 (Manin's trick)**.** *The $\mathbb{Q}$-vector space $\mathcal{M}_2(\mathbb{Q})$ is generated by the set of symbols $\gamma\{\infty, 0\}$ with $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.*

We sometimes call such symbols **unimodular symbols**. The proof will in fact provide an algorithm for writing a modular symbol as a linear combination of unimodular symbols.

Proof. Using $\{x, y\} = \{x, 0\} - \{y, 0\}$, it suffices to treat $\{x, 0\}$ with $x \in \mathbb{Q}$. Compute the negative-regular or Hirzebruch–Jung continued fraction of x [Voi21, Exercise 35.4] and let a_i/b_i be its convergents, obtained from the Euclidean algorithm. (Or you can also use the usual continued fraction and adjust some signs, no worries.) Consecutive convergents satisfy $a_{i+1}b_i - a_i b_{i+1} = 1$, so the relation

$$\left\{\frac{a_{i+1}}{b_{i+1}}, 0\right\} = \left\{\frac{a_{i+1}}{b_{i+1}}, \frac{a_i}{b_i}\right\} + \left\{\frac{a_i}{b_i}, 0\right\} = \gamma_i\{\infty, 0\} + \left\{\frac{a_i}{b_i}, 0\right\} \tag{3.5.2}$$

with $\gamma_i = \begin{pmatrix} a_{i+1} & a_i \\ b_{i+1} & b_i \end{pmatrix}$ iterates to express $\{x, 0\}$ as a sum of symbols given by translates of $\{\infty, 0\}$ by an element of $\mathrm{SL}_2(\mathbb{Z})$. (This should be the right way around, since we usually swap the order and then the determinant is -1? Yikes.) □

3.5.3. Compute a set of representatives $\{\delta_i\}_i$ for $\Gamma \backslash \mathrm{SL}_2(\mathbb{Z})$ as in section 3.2. It follows from Manin's trick (Lemma 3.5.1) that the symbols $\delta_i\{\infty, 0\}$ generate $\mathcal{M}_2(\Gamma)$. The relations among them come from the standard generators

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad ST = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

of $\mathrm{SL}_2(\mathbb{Z})$ satisfying $S^2 = (ST)^3 = -1$. Writing the right action on cosets as $[i]\gamma = [j]$ when $\Gamma\gamma_i\gamma = \Gamma\gamma_j$, the resulting **Manin relations** take the form

$$\begin{aligned} [i] + [i]S &= 0 \\ [i] + [i]R + [i]R^2 &= 0. \end{aligned} \tag{3.5.4}$$

More generally, with coefficients $V_k(R)$ the same presentation is used with the action extended diagonally to the coefficients.

3.5.5. Reduction of Manin symbols then proceeds as follows. Given a symbol $\{x, y\}$, we use continued fractions as in (3.5.2) to express it as a linear combination of translates of $\{\infty, 0\}$, then reduces each translate to a coset representative δ_i as in section 3.2, then applies the quotient to the coefficients.

3.6 Hecke operators

We recall the definition of Hecke operators from section 2.6 defined using double cosets $\beta \in \mathrm{GL}_2(\mathbb{Q})_{>0}$ and with the explicit expression by Hermite normal forms for the standard congruence groups.

3.6.1. Working now (dually) in homology, we express the Hecke operators as follows. For $v \otimes \{x, y\} \in \mathcal{M}_k(\Gamma, R)$ and $\beta \in \mathrm{GL}_2(\mathbb{Q})_{>0}$, we define

$$T_\beta(v \otimes \{x, y\}) := \sum_{\alpha \in \Theta(\beta)} \alpha(v \otimes \{x, y\}) = \sum_{\alpha \in \Theta(\beta)} \alpha v \otimes \{\alpha x, \alpha y\}.$$

Compatibility of the Hecke operators with the integration pairing

$$\langle T_n f,\ \xi \rangle = \langle f,\ T_n \xi \rangle \tag{3.6.2}$$

comes from a straightforward change-of-variables argument.

Example 3.6.3. We return to the Example 3.4.10. The cosets for $\Gamma = \Gamma_0(11)$ are specified by $\mathbb{P}^1(\mathbb{Z}/11\mathbb{Z})$ which we write as $[0], [1], \ldots, [10], [\infty]$. This gives a generating set.

The relations are as follows. For $[i]S = [-1/i]$ we get

$$[0] + [\infty] = [1] + [10] = [2] + [5] = [3] + [7] = [4] + [8] = [6] + [9] = 0.$$

For $[i]R = [-1/(i+1)]$, we obtain

$$[0] + [10] + [\infty] = [2] + [7] + [4] = [1] + [5] + [9] = [3] + [8] + [6] = 0.$$

Thus we get a basis $[0], [2], [3]$ matching the Farey symbols, and the following identifications at other cosets: $[1] = 0 = [10]$, $[4] = [3] - [2]$, $[5] = -[2] = [6]$, $[7] = -[3]$, $[8] = [2] - [3]$, $[9] = [2]$, $[\infty] = -[0]$. This also gives the boundary map, and the basis for the kernel is $[2], [3]$.

Great, now back to Hecke operators. We start with

$$\begin{aligned} T_2[2] = T_2\{1/2, 0\} &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \{1/2, 0\} + \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \{1/2, 0\} + \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \{1/2, 0\} \\ &= \{1, 0\} + \{1/4, 0\} + \{3/4, 1/2\} = [1] + [4] + \{3/4, 0\} - \{1/2, 0\} \\ &= [1] + [4] + [7] + [10] - [2] = 0 + [3] - [2] - [3] + 0 - [2] = -2[2]. \end{aligned} \tag{3.6.4}$$

(Computed with the standard continued fraction, $3/4 = 0 + 1/(1 + 1/3)$ so convergents $0 = 0/1$, $0 + 1/1 = 1/1$ and $3/4$, giving

$$\{3/4, 0\} = \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix} \{\infty, 0\} + \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \{\infty, 0\} + \{0, 0\} \tag{3.6.5}$$

$$= [(4 : -1)] + [(1 : -1)] = [7] + [10].) \tag{3.6.6}$$

Should compute with the negative continued fraction to see better the connection with $\mathrm{SL}_2(\mathbb{Z})$!

Similarly we compute that $T_2[3] = -2[3]$, so $[T_2] = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$. From the eigenform

$$f(q) = q \prod (1 - q^n)^2 (1 - q^{11n})^2 = q - 2q^2 - q^3 + \dots$$

which we look up in the LMFDB, we find a match with the coefficient of q^2!

Remark 3.6.7. A small implementation point: reduction is the bottleneck. If you repeatedly expand T_p and reduce every term from scratch, performance collapses. The standard speedup is to precompute reduction data once (continued fraction steps, coset reductions, and the matrix giving coordinates of each raw Manin symbol in the final basis) and then reuse it for every prime p you need.

3.7 Cohomology

We have emphasized homology because it is concrete and efficient for computation, but the same structure appears (often more conceptually cleanly) in *co*homology; this ends up being just another wrapper for the same conclusion, as we might expect by duality.

The Eichler–Shimura isomorphism provides an identification

$$H^1(X(\Gamma), \mathbb{C}) \xrightarrow{\sim} S_2(\Gamma) \oplus \overline{S_2(\Gamma)},$$

and more generally, for the weight representation W_k and the associated local system $\mathscr{W}_k$,

$$H^1\big(X(\Gamma), \mathscr{W}_k\big) \xrightarrow{\sim} S_k(\Gamma) \oplus \overline{S_k(\Gamma)}.$$

Working over $Y(\Gamma)$ rather than $X(\Gamma)$ introduces an Eisenstein contribution, reflecting the difference between absolute and relative cohomology at the cusps.

Exercises

3.1. Show that there is a unique subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ of index 2, identify it as a congruence subgroup, and show that a Farey symbol is $-\infty <_\bullet 0 <_\bullet \infty$.

3.2. Analyze the running time of the algorithm for computing Farey symbols by counting bit operations and the number of calls to a function which identifies membership in Γ.

3.3. Prove the following.

(a) In 3.3.6, show that $\mathcal{B}_2(\Gamma)$ is isomorphic the Γ-coinvariants of the free abelian group on $\mathbb{P}^1(\mathbb{Q})$.
(b) Show that the boundary map (3.3.7) is well-defined.
(c) Repeat (a) and (b) for $\mathcal{B}_k(\Gamma, R)$.

3.4. Carry out the same steps in Example 3.4.10 and Example 3.6.3 for $N = 4$ and $N = 13$. Most importantly, have fun doing so!

3.5. What is the deal with the Hecke operators and the torsion submodule $\mathcal{M}_2(\Gamma)_{\text{tors}}$? It's Eisenstein torsion?

Chapter 4

The trace formula

4.1 Eichler–Selberg trace formula

Exercises

4.1. Verify the computational complexities and comparisons between modular symbols and the trace formula computed in [BB+20, Table 5.2.3].

Chapter 5

Additional methods

We mention several other methods for computing modular forms that are quite useful.

5.1 Multiplication and graded ring

One great way to get modular forms of higher weight is to multiply forms of lower weight. A main result of [VZB15, Main Theorem 1.4.1, Corollary 1.5.1] is that the graded ring of even weight modular forms is generated in weights $2, 4, 6$; the results also provide an initial ideal for the relations among them. This gives a way to express forms of larger weight as monomials in smaller weight.

This point of view for example is useful when working with p-adic modular forms, where weights can quickly get large.

Cusp forms of weight 2 define the canonical map from a modular form to projective space, and more generally modular forms of weight 2 define a *log* canonical map; this gives a way to compute equations for modular curves. More generally, the proj $\operatorname{Proj} M_*(\Gamma)$ of the graded ring of even weight modular forms gives equations for the stacky modular curve in weighted projective space.

5.2 Eisenstein series

There are also explicit formulas for Eisenstein series, obtained from averaging over G Eisenstein series for the principal congruence subgroup in a manner analogous to the above. See sections 7.3–7.4 of Cohen–Stromberg, Brunault–Neururer, and Zywina, as well as earlier work of Borisov–Gunnells.

Then Khuri-Makdisi shows that products of Eisenstein series generate.

Exercises

5.1.

Part II

Modular forms on reductive groups

Chapter 6

Bianchi modular forms

In this chapter, we make our first step towards generalize notions from the previous part. The main goal is to begin preparations for the general case by considering a key example.

6.1 Compactification

Take F imaginary quadratic.

There is a bijection between the cusps, the orbits of $\mathrm{GL}_2(\mathbb{Z}_F)$ (or $\mathrm{SL}_2(\mathbb{Z}_F)$ acting on $\mathbb{P}^1$ and the ideal class group. This is in fact something that is true for any number field, so we could do it in that level of generality.

6.2 Modular symbols

Then $D = \mathbf{H}^3$ and one can similarly define modular symbols, now with elements in $\mathbb{P}^1(K)$. Things look just like in Chapter 3: a finitely generated module and matrices for Hecke operators.

6.3 Eisenstein series

The notion of Eisenstein series extends in a natural way to the Bianchi groups $\mathrm{PSL}_2(\mathbb{Z}_F)$ where F is an imaginary quadratic field: see Elstrodt–Grunewald–Mennicke [EGM98, Chapter 3].

Exercises

6.1.

Chapter 7

Reductive algebraic groups

In this chapter, we give the basics of linear algebraic groups. Our goal is modest: we introduce and outline the basic infrastructure, giving references for further reading. We hope to provide intuition and enough detail so that we can define modular forms in a general context and then describe algorithms and then effective computation: (not just in principle algorithmic, but also something concrete that can actually be carried out that finishes on realistic input).

7.1 Beyond GL_2

In Part I, we worked with modular forms as differential forms on modular curves, naturally presented as a quotient of the hyperbolic upper half-plane and moduli spaces for elliptic curves with level structure. In the previous section, we considered the natural generalization for hyperbolic three-space. These arose naturally from the groups GL_2 over $F = \mathbb{Q}$ and F an imaginary quadratic field.

We generalize by allowing subgroups of invertible matrices over a number field defined by polynomial equations. We want to apply techniques from algebraic geometry, so we want to think not only of the group of points over a field but also its scheme structure, which amounts to thinking about the group of points over extensions.

Let $\mathbf{G} \leq \mathrm{GL}_n$ be a **linear algebraic group** defined over a number field F, a closed subgroup of the algebraic group GL_n defined by the vanishing of polynomials with coefficients in $\mathbb{F}$ in the entries and the inverse of the determinant of a matrix of indeterminates. Equivalently, $\mathbf{G}$ is an affine algebraic variety over F equipped with multiplication and inversion maps that are morphisms of algebraic varieties. The most basic examples of linear algebraic groups are the multiplicative group $\mathbb{G}_{\mathrm{m}}$ and the additive group $\mathbb{G}_{\mathrm{a}}$ and the general linear group GL_n and special linear group SL_n.

Linear algebraic groups which are built from $\mathbb{G}_{\mathrm{m}}$ and $\mathbb{G}_{\mathrm{a}}$, for example subgroups of upper-triangular matrices, do not provide the generalization of mod-

ular forms we seek here: they are *solvable*. Indeed, subgroups of unipotent matrices (upper-triangular with 1s on the diagonal), with composition factors $\mathbb{G}_{\mathrm{a}}$, have a representation theory that is decidedly less fun as we know from linear algebra.

Accordingly, we say that the group $\mathbf{G}$ is **reductive** if its maximal connected *unipotent* normal subgroup is trivial and is **semisimple** if its maximal connected *solvable* normal subgroup is trivial. The group GL_n is reductive but not semisimple, as it has center given by scalar matrices; the group SL_n is semisimple. Reductive linear algebraic groups $\mathbf{G}$ over an algebraically closed field are classified by their *root datum*, with the following key examples (with $n \geq 1$):

A_n	$\mathrm{SL}_{n+1}, \mathrm{SU}_n$
B_n	SO_{2n+1}
C_n	Sp_{2n}
D_n	SO_{2n}
$\mathsf{E}_6, \mathsf{E}_7, \mathsf{E}_8$	exceptional
$\mathsf{F}_4, \mathsf{G}_2$	exceptional

The fact that there are redundancies in this list is a feature, not a bug! We may also consider forms of any of these groups, such as the (special) orthogonal group of a quadratic form or the units (of reduced norm 1) of a central simple algebra. We once again feel the lesson that the classification of semisimple Lie algebras has an unfair impact on understanding the mathematical big picture!

7.2 Algebraic groups

This section collects the minimum structure theory of algebraic groups that we will use later when we pass from modular forms to Hecke modules on locally symmetric spaces. References for the general theory are [Bo91, Spr98, Hum75, Mil17]. Add Getz–Hahn.

Let k be a field. Via its coordinate entries, the matrix ring

$$\mathrm{M}_{n,k} := \operatorname{Spec} k[x_{ij}]_{1 \leq i,j \leq n} \simeq \mathbb{A}_k^{n^2} \tag{7.2.1}$$

is naturally an affine space over k. The Zariski open subset of matrices with nonzero determinant defines an affine variety $\mathrm{GL}_{n,k} = \operatorname{Spec} k[\mathrm{GL}_{n,k}]$ with coordinate ring $\mathscr{O}_{\mathrm{GL}_{n,k}} = k[\mathrm{GL}_{n,k}] = k[x_{ij}]_{i,j}[D^{-1}]$ and $D := \det(x_{ij})_{i,j} \in k[x_{i,j}]_{i,j}$ is the determinant (a polynomial of degree n with $n!$ terms). In this way, $\mathrm{GL}_{n,k}$ is a smooth affine k-scheme but its point sets define a group. We will sometimes drop the subscript k, writing just GL_n.

Historically, when k is algebraically closed, linear algebraic groups were defined as closed subgroups of $\mathrm{GL}_{n,k}$. As we will work over non-algebraically closed fields, but will still want to work with point sets as groups, we adopt a categorical perspective and then reconcile this with the concreteness of matrix groups.

Definition 7.2.2. An **affine group scheme** over k is a group object in the category of affine k-schemes of finite type.

Equivalently, an affine group scheme is an affine k-scheme $\mathrm{G} = \operatorname{Spec} k[\mathrm{G}]$ whose coordinate ring $k[\mathrm{G}]$ is a finitely generated commutative Hopf algebra over k: giving a Hopf algebra structure on $k[\mathrm{G}]$ (coidentity, comultiplication, coinverse) is equivalent to giving identity $e\colon \operatorname{Spec} k \to \mathrm{G}$, multiplication $m\colon \mathrm{G}\times \mathrm{G} \to \mathrm{G}$, and inverse morphisms $i\colon \mathrm{G} \to \mathrm{G}$ [Mil17, §1–2] that satisfy the axioms of a group. Thus for an affine group scheme G over k and every k-algebra R, the set

$$\mathrm{G}(R) := \operatorname{Hom}_k(\operatorname{Spec} R, \mathrm{G}) \tag{7.2.3}$$

is a group, functorially in R. From this point of view, we can also think of an affine group scheme over k as a functor from the category of k-algebras to the category of groups that is representable by a finitely generated commutative k-algebra. The Hopf algebra structure is recovered automatically from the group-valued functor by the Yoneda lemma.

Definition 7.2.4. A **homomorphism** $\varphi\colon \mathrm{G} \to \mathrm{G}'$ of affine group schemes is a morphism of schemes that is a homomorphism of group objects (i.e., φ commutes with the identity, multiplication, and inverse maps in the three commutative diagrams).

Equivalently, the map of sets $\mathrm{G}(R) \to \mathrm{G}'(R)$ is a group homomorphism for every k-algebra R.

Example 7.2.5. The matrix multiplication map $m\colon \mathrm{GL}_n \times \mathrm{GL}_n \to \mathrm{GL}_n$, on points defined by

$$\big((x_{ij})_{i,j}, (y_{ij})_{i,j}\big) \mapsto (z_{ij})_{i,j}$$

where $(x_{ij}) \cdot (y_{ij}) = (z_{ij})$ is defined by $z_{ij} = \sum_{r=1}^{n} x_{ir} y_{rj}$, is given by the polynomial map

$$\begin{aligned} m^*\colon k[\mathrm{GL}_n] &\to k[\mathrm{GL}_n] \otimes_k k[\mathrm{GL}_n] \\ m^*(x_{ij}) &= \sum_{r=1}^{n} x_{ir} \otimes x_{rj} \end{aligned} \tag{7.2.6}$$

The identity map is $e\colon \operatorname{Spec} k \to \mathrm{GL}_n$ by $k[\mathrm{GL}_n] \to k$ by $e^*(x_{ij})_{i,j} \mapsto I_n$. The inverse map $i\colon \mathrm{GL}_n \to \mathrm{GL}_n$ is similarly defined by cofactor expansion (Cramer's rule).

Definition 7.2.7. An **affine algebraic group** over k is a reduced affine group scheme of finite type over k.

Theorem 7.2.8 (Cartier)**.** *If* $\operatorname{char} k = 0$*, then every affine group scheme of finite type over* k *is smooth (in particular, reduced).*

Proof. See Milne. Briefly, left translation by $g \in \mathrm{G}$ identifies $\Omega^1_{\mathrm{G}/k}$ with its pullback along translation by g, so $\Omega^1_{\mathrm{G}/k}$ is a locally free $\mathscr{O}_{\mathrm{G}}$-module of (constant) rank equal to $\dim \operatorname{Lie}(\mathrm{G})$. In characteristic zero, the Jacobian criterion

shows that locally free implies smoothness. See more generally Stacks Project, Tag 0BF6. □

Thus in characteristic zero we avoid some pathologies coming from nonreduced affine group schemes like α_p over $\mathbb{F}_p$.

With this setup, we now make the key definition.

Definition 7.2.9. A **(linear) representation** of an affine algebraic group is a morphism $\rho\colon G \to \mathrm{GL}_n$ for some n; we say ρ is **faithful** if ρ is a closed immersion.

A **linear algebraic group** over k is an affine algebraic group $\mathbf{G}$ which admits a faithful linear representation.

Theorem 7.2.10. *Every affine algebraic group over k is a linear algebraic group over k.*

Proof. See Milne, Theorem 9.1: the regular representation has faithful finite-dimensional subrepresentations. See also Waterhouse's Introduction to Affine Group Schemes. □

Example 7.2.11. Examples of linear algebraic groups abound. For example, we have the subgroup of diagonal matrices

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix} \leq \mathrm{GL}_{n,k}$$

defined by $x_{ij} = 0$ for all $i \neq j$. We note how the description of the points and the scheme go together. These are already defined over the prime field e.g. $k = \mathbb{F}_p$ or $k = \mathbb{Q}$.

We similarly have tori $\mathbb{G}_{\mathrm{m}}^r$ and additive groups $\mathbb{G}_{\mathrm{a}}^r$.

Example 7.2.12. Similarly, we have upper triangular matrices, unipotent matrices (upper triangular with 1s on the diagonal), as well as $\mathrm{SL}_{n,k}$ defined by the closed condition $D = 1$.

Example 7.2.13. Important for these notes are those subgroups that preserve a bilinear form up to rescaling, for example the usual orthogonal group

$$\mathrm{O}_n(k) := \{A \in \mathrm{GL}_n(k) : AA^{\mathsf{t}} = A^{\mathsf{t}}A = 1\} \tag{7.2.14}$$

(preserving the dot product). Of course (7.2.14) only defines the set of points; but we could equally well replace k by R any k-algebra with the same condition to get $\mathrm{O}_n(R)$, and by the Yoneda lemma this is enough to recover it uniquely. Here, we can take R the coordinate ring of $\mathrm{GL}_{n,k}$ representing the universal

matrix $A = (x_{ij})_{i,j}$ to see explicitly the defining equations. For example, for $n = 2$,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{7.2.15}$$

gives the equations $a^2 + b^2 = c^2 + d^2 = 1$ and $ac + bd = 0$ defining $\mathrm{O}_{n,k}$ inside $\operatorname{Spec} k[a, b, c, d, D^{-1}]$.

In general, if $T \in \mathrm{M}_n(k)$ then we can similarly define

$$\mathbf{G}(R) := \{(A, u) \in \mathrm{GL}_n(R) \times \mathrm{GL}_1(R) : ATA^{\mathsf{t}} = uT\} \tag{7.2.16}$$

and $\mathbf{G}$ is a linear algebraic group.

Example 7.2.17. Every finite group is a linear algebraic group, via the regular representation. As a finite set of points over k, such groups are disconnected whenever the group is nontrivial. But they are allowed!

Example 7.2.18. Let B be a finite-dimensional algebra over k. Then via the regular representation, $B^\times$ is a linear algebraic group. If B is a central simple k-algebra, then there is further a reduced norm map $\operatorname{nrd}\colon B^\times \to \mathbb{G}_{\mathrm{m}}$ (a power of which is the algebra norm) giving a group $B^1 = \ker \operatorname{nrd}$. The special case where B is a quaternion algebra over k is of particular interest in what follows.

In fact, more generally, $\mathrm{Aut}_k(B)$ is also a linear algebraic group.

Remark 7.2.19. An irreducible projective group variety over k in fact has a commutative multiplication law: it is an abelian variety.

In principle, to specify a reductive group $\mathbf{G}$ in bits, one could give a set of defining equations for $\mathbf{G}$ inside GL_n or to give the classical groups (or their forms) by name. Using the classification, will see below that we can cover the spread reasonably well by considering labels.

7.3 Subgroups, identity component

Let $\mathbf{G}$ be a linear algebraic group over k.

Definition 7.3.1. A **closed (linear algebraic) subgroup** $\mathbf{H} \leq \mathbf{G}$ is a closed, reduced subscheme which inherits a group scheme structure from $\mathbf{G}$.

In other words, $k[\mathbf{H}] = k[\mathbf{G}]/I$ where I is a radical ideal preserved by the coidentity, comultiplication, and coinverse maps. These are all properties we can check over the algebraic closure, so it is enough to show that ($\mathbf{H}$ is reduced and) $\mathbf{H}(k^{\mathrm{al}})$ is Zariski closed in $\mathbf{G}(k^{\mathrm{al}})$.

A linear algebraic group over k is then (essentially by definition) a closed subgroup of $\mathrm{GL}_{n,k}$ for some k.

Example 7.3.2. We have the upper-triangular **Borel**

$$\mathrm{B}_n := \left\{ \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix} \right\} \leq \mathrm{GL}_n \,.$$

Inside B_n, the **diagonal torus**

$$\mathrm{T}_n := \left\{ \begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix} \right\}$$

and the upper-triangular unipotent subgroup

$$\mathrm{U}_n := \left\{ \begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \right\}.$$

We will soon see that B_n is built from T_n and U_n. For now, we just note the dimensions

$$\begin{aligned} \dim \mathrm{B}_n &= 1 + 2 + \cdots + n = \frac{n(n+1)}{2}, \\ \dim \mathrm{T}_n &= n, \\ \dim \mathrm{U}_n &= 1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}. \end{aligned}$$

Example 7.3.3. Let $\mathrm{H} \leq \mathrm{G}$ be a closed subgroup. Then the **normalizer** and **centralizer**

$$\begin{aligned} \mathrm{N}_{\mathrm{G}}(\mathrm{H}) &:= \{g \in G : g^{-1}Hg = H\} \leq G, \\ \mathrm{C}_{\mathrm{G}}(\mathrm{H}) &:= \{g \in G : g^{-1}hg = h \text{ for all } h \in H\} \leq G, \end{aligned} \tag{7.3.4}$$

are closed subgroups.

Indeed, for $h \in H$, the conjugation map

$$\begin{aligned} c_h \colon G &\to G \\ g &\mapsto g^{-1}hg \end{aligned} \tag{7.3.5}$$

is a morphism of affine algebraic groups, so $c_h^{-1}(H) = \{g \in G : g^{-1}hg \in H\}$ is closed, thus $\bigcap_{h \in H} \varphi_h^{-1}(H) = N_G(H)$ is closed. And similarly, $C_G(H)$ is closed.

In particular, the center $\mathrm{Z}(\mathrm{G}) = \mathrm{C}_{\mathrm{G}}(\mathrm{G})$ is a closed subgroup.

Let $e = e(\operatorname{Spec} k) \in \mathrm{G}(k)$ be the identity element. Let

$$\mathrm{G} = \mathrm{G}_1 \cup \mathrm{G}_2 \cup \cdots \cup \mathrm{G}_r \tag{7.3.6}$$

be the irredundant decomposition into irreducible components.

Lemma 7.3.7. *$e \in \mathbf{G}_i$ for a unique i.*

Proof. In a nutshell, the product of the components $\mathbf{G}_i$ containing e is irreducible so contained in some $\mathbf{G}_j$, but $\mathbf{G}_i \subseteq \mathbf{G}_j$ for all i so since irredundant must have only one such component. □

Definition 7.3.8. The unique irreducible component of $\mathbf{G}$ containing e is called the **identity component** of $\mathbf{G}$, denoted $\mathbf{G}^0 \leq \mathbf{G}$.

Since $\mathbf{G}^0$ is connected, we also say $\mathbf{G}^0$ is the **connected component of the identity**.

Proposition 7.3.9. *The following statements hold.*

(a) $\mathbf{G}^0 \triangleleft \mathbf{G}$ *is a closed, normal subgroup of finite index.*
(b) *There is a decomposition*

$$\mathbf{G} = \bigsqcup_{g\mathbf{G}^0 \in \mathbf{G}/\mathbf{G}^0} g\mathbf{G}^0$$

as a finite, disjoint union into irreducible/connected components.
(c) *Every closed subgroup of* $\mathbf{G}$ *of finite index contains* $\mathbf{G}^0$.

Proof. For (a), we show closed under inverse and multiplication, for all $g \in \mathbf{G}^0$ we have $e \in g^{-1}\mathbf{G}^0 \simeq \mathbf{G}^0$ irreducible so $g^{-1}\mathbf{G}^0 \subseteq \mathbf{G}^0$ and hence equality holds, and hence $g^{-1}h \in \mathbf{G}^0$ for all $g, h \in \mathbf{G}^0$. To check normal, same with $e \in g\mathbf{G}^0 g^{-1} = \mathbf{G}^0$ for all $g \in \mathbf{G}$.

For (b), we have $\mathbf{G} = \bigcup_{g\in\mathbf{G}} g\mathbf{G}^0$ a union of irreducible components, so by Noetherianity it is a finite union, and cosets are disjoint.

For (c), given $\mathbf{H} \leq \mathbf{G}$ closed of finite index, we have $\mathbf{G} = \bigsqcup_{g\mathbf{H}\in\mathbf{G}/\mathbf{H}} gH = H \sqcup G'$ where G' is the rest; intersecting with G^0 gives $e \in G^0 = (H \cap G^0) \sqcup (G' \cap G_0)$, but G^0 is connected and $e \in H \cap G^0$, so $G^0 = H \cap G^0$ thus $G^0 \subseteq H$. □

Proposition 7.3.10. *Let $\varphi\colon G \to G'$ be a homomorphism of linear algebraic groups. Then the following statements hold.*

(a) $\ker\varphi \leq G$ *is a closed subgroup.*
(b) $\operatorname{img}\varphi = \phi(G) \leq G'$ *is a closed subgroup.*
(c) $\varphi(G^0) = \varphi(G)^0$.

Proof. For (a), we have $\ker\varphi = \phi^{-1}(\{e'\})$. For (b), we have $\phi(G)$ constructible so closed (add reference). For (c), we have $\varphi(G^0)$ closed by (b) and irreducible, so $\varphi(G^0) \subseteq \varphi(G)^0$; but $\varphi(G^0) \leq \varphi(G)$ is finite index since $G^0 \leq G$ is finite index, so in fact $\varphi(G)^0 \subseteq \varphi(G^0)$ by Proposition 7.3.9(c), so equality holds. □

7.4 Restriction of scalars

One may work with a group $\mathbf{G}$ defined over a number field F by taking the restriction of scalars. If F is a number field and $\mathbf{G}$ is an algebraic group over F, then $\mathrm{Res}_{F|\mathbb{Q}}\mathbf{G}$ is a $\mathbb{Q}$-group characterized by

$$(\mathrm{Res}_{F|\mathbb{Q}}\mathbf{G})(R) = \mathbf{G}(R \otimes_{\mathbb{Q}} F)$$

for all $\mathbb{Q}$-algebras R. In particular $(\mathrm{Res}_{F|\mathbb{Q}}\mathbf{G})(\mathbb{Q}) = \mathbf{G}(F)$. This allows us to work with groups over $\mathbb{Q}$ without any real loss of generality, for example with Hilbert modular groups [Mil17, §2], and it is a standard tool in arithmetic subgroup theory [PR94].

Remark 7.4.1. For the locally symmetric spaces we define, we are really working with the complex points of something which we descend to a number field using the theory of canonical models due to Shimura and Deligne. This construction does not see a difference between the group $\mathbf{G}$ and its restriction of scalars, and especially in the formalism of Shimura varieties it is standard just to work over $\mathbb{Q}$.

7.5 Unipotent and reductive

Definition 7.5.1. An element $g \in \mathrm{GL}_n(k^{\mathrm{al}})$ is **unipotent** if $(g-1)^m = 0$ for some $m \geq 1$, equivalently all eigenvalues of g are 1.

Definition 7.5.2. A linear algebraic group $\mathbf{U}$ over k is **unipotent** if $\mathbf{U}(k^{\mathrm{al}})$ consists of unipotent elements in some (equivalently any) faithful representation.

We say that a linear algebraic group is **solvable** if its group of k^{al}-points is solvable. We then have the following key structural theorem of Lie–Kolchin.

Theorem 7.5.3 (Lie–Kolchin)**.** *Suppose k is algebraically closed of characteristic 0 and let $\mathbf{H} \leq \mathrm{GL}_n$ be a solvable, connected algebraic subgroup. Then $\mathbf{H}$ stabilizes a complete flag in k^n, or equivalently $\mathbf{H}$ is conjugate to a subgroup of upper triangular matrices.*

Proof. We use the Borel fixed-point theorem: a connected solvable linear algebraic group acting by morphisms on a complete variety has a fixed point [Bo91, §11] (see also Humphries [Hum75, §15]). The induced action of $\mathbf{H}$ on the projective space $\mathbb{P}^{n-1}$ therefore fixes a point of $\mathbb{P}(V)$, i.e., stabilizes a line $L \subseteq k^n$. The induced action on the quotient is again by a connected, solvable group, so we finish by induction on n. □

Definition 7.5.4. The **solvable radical** $\mathrm{rad}(\mathbf{G})$ is the maximal connected solvable normal subgroup, and the **unipotent radical** $\mathrm{urad}(\mathbf{G})$ is the maximal connected unipotent normal subgroup.

These exist and are unique [Bo91, §13].

Definition 7.5.5. The group $\mathbf{G}$ is **reductive** if $\mathrm{urad}(\mathbf{G}) = \{1\}$ and **semisimple** if $\mathrm{rad}(\mathbf{G}) = \{1\}$.

Example 7.5.6. GL_n is reductive: it has no nontrivial connected unipotent normal subgroup. It is not semisimple because its center contains the scalar matrices, a copy of $\mathbb{G}_{\mathrm{m}}$. In contrast, SL_n is semisimple for $n \geq 2$. See [Hum75, §14–16] or [Bo91, §11–14].

One pretty good reason why reductive groups are the natural class for automorphic forms is that they have a workable representation theory and a clean structure theory via maximal tori and roots [Bo91, Spr98].

7.6 Tori, characters, and rank

Definition 7.6.1. A **torus** over k is an algebraic group $\mathbf{T}$ such that $\mathbf{T}_{k^{\mathrm{al}}} \simeq \mathbb{G}_{\mathrm{m}}^r$ for some r. A torus over k is **(k-)split** if $\mathbf{T} \simeq \mathbb{G}_{\mathrm{m}}^r$ over k.

The basic invariant of a torus is the character lattice

$$X^*(\mathbf{T}) = \mathrm{Hom}(\mathbf{T}_{k^{\mathrm{al}}}, \mathbb{G}_{\mathrm{m}}),$$

a free abelian group of rank r equipped with a natural $\mathrm{Gal}_k := \mathrm{Gal}(k^{\mathrm{al}} \,|\, k)$-action [Mil17, §8].

The **rank** of a connected reductive group over k is the dimension of a maximal split torus over k. If $\mathbf{G}$ has no nontrivial split torus over k, we say that $\mathbf{G}$ is **anisotropic** over k.

This simple invariant controls many important properties of arithmetic quotients and the resulting modular forms.

Proposition 7.6.2. *Let G be a simple Lie group with real rank 1 and finite center. Then the Lie algebra of G is isomorphic to the Lie algebra of one of the following Lie groups with $n \geq 2$:*

- $\mathrm{SO}(n, 1)$,
- $\mathrm{SU}(n, 1)$,
- $\mathrm{Sp}(n, 1)$ *(a quaternionic unitary group), or*
- *a certain form of the exceptional Lie group F_4.*

The first three of these fit together nicely: they are the group of real, complex, or quaternionic matrices preserving a quadratic, Hermitian, or quaternionic Hermitian form, respectively, with signature $n, 1$.

Proof. Cartan's classification [Tit67]. □

7.7 Parabolics and Levi decomposition

To begin, we suppose that k is algebraically closed.

Definition 7.7.1. A **Borel subgroup** $\mathrm{B} \subset \mathrm{G}$ is a maximal connected solvable subgroup.

Definition 7.7.2. A **parabolic subgroup** $\mathrm{P} \subset \mathrm{G}$ is a closed subgroup containing a Borel subgroup.

Every parabolic subgroup has a large unipotent radical and admits a Levi decomposition.

Theorem 7.7.3 (Levi decomposition)**.** *Let* G *be a connected linear algebraic group over a field* k *of characteristic* 0*, and let* $\mathrm{P} \subset \mathrm{G}$ *be a parabolic subgroup. Write* $\mathrm{U} = \mathrm{urad}(\mathrm{P})$*. Then there exists a reductive subgroup* $\mathrm{L} \subset \mathrm{P}$ *such that* $\mathrm{P} = \mathrm{L} \ltimes \mathrm{U}$*.*

Proof sketch in the matrix case. For $\mathrm{G} = \mathrm{GL}_n$ and P the stabilizer of a flag, P is the group of block upper triangular matrices, U is the subgroup of block upper unitriangular matrices, and L is the block diagonal subgroup. This gives $\mathrm{P} = \mathrm{L} \ltimes \mathrm{U}$ directly. The general case reduces to this by choosing a faithful representation in which P stabilizes a flag (see Borel [Bo91, §14–15] or Springer[Spr98, §8–9]). □

Equivalently, P is parabolic if and only if the homogeneous variety G/P is projective; see [Bo91, §11].

Now allowing a general field k, we make the following definition.

Definition 7.7.4. A closed subgroup $\mathrm{P} \leq \mathrm{G}$ defined over k is **parabolic** if G/P is a projective variety over k. Its **unipotent radical** is the maximal connected normal unipotent k-subgroup $\mathrm{U}_\mathrm{P} \trianglelefteq \mathrm{P}$ and the **Levi quotient** is the reductive k-group $\mathrm{L}_\mathrm{P} := \mathrm{P}/\mathrm{U}_\mathrm{P}$.

Parabolics defined over the ground field are the ones that parametrize boundary strata of arithmetic quotients, and they are central in reduction theory.

7.8 Classification over $\mathbb{C}$

See e.g. Helgason [Hel01, §X.6].

7.9 Classification over $\mathbb{R}$

Classification of real Lie groups, see Onischchik–Vinberg or Knapp.

7.10 Integral models

For integral models and group schemes (when we want a clean notion of "integral points" over $\mathbb{Z}$ or $\mathbb{Z}_F$) see [Con14, Mil17].

We can start by choosing an embedding $\mathbf{G} \hookrightarrow \mathrm{GL}_n$ and then considering $\mathbf{G}(F) \cap \mathrm{GL}_n(\mathbb{Z}_F)$. This group depends on the choice of embedding.

Lemma 7.10.1. *The commensurability class of* $\mathbf{G}(F) \cap \mathrm{GL}_n(R)$ *in* $\mathbf{G}(F)$ *is well-defined.*

Proof. More precisely, if $\mathbf{G}$ is defined over F, then after clearing denominators in the defining equations one gets some group scheme $\mathcal{G} \subset \mathrm{GL}_n$ over $R = \mathbb{Z}_F$ with generic fibre $\mathcal{G}_F \simeq \mathbf{G}$. Changing the embedding gives another model over R, which changes only finitely many fibers, and this does not affect its commensurability class [BHC62, Bor69, PR94]. □

Exercises

Unless otherwise specified, in the exercises we work over a field k.

7.1. For the orthogonal group show that $D^2 = 1$ where D is the determinant, so $\mathrm{O}_n \subseteq \mathrm{M}_n$ is closed but not connected. Draw a picture or cartoon for $n = 2$.

7.2. A matrix $A \in \mathrm{GL}_n(k)$ is **monomial** if each row, column has a unique nonzero entry.

(a) Show that the set of monomial matrices N is a closed subgroup of GL_n.
(b) Show that $N^0 = T_n$, the diagonal torus.
(c) Show that $N/N^0 \simeq S_n$, with the cosets represented by a permutation matrix.

Chapter 8

Symmetric spaces

8.1 Overview

We saw that the modular curve $Y(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H}^2$ is naturally the space of lattices in dimension 2. Similarly, associated to GL_n and SL_n there is a space of lattices in dimension n and functions on this space—or more precisely, functions on a space of framed lattices which have an invariance with respect to $\mathrm{SL}_n(\mathbb{Z})$—are modular forms.

In this chapter, we similarly describe the replacement for the upper half-plane for a general reductive group $\mathbf{G}$. In this overview, we consider the case $F = \mathbb{Q}$, shunting to the chapter for the details over number fields. We first describe the general setup, then descriptions in terms of lattices.

Let $G_\infty := \mathbf{G}(\mathbb{R})$ denote the real points of $\mathbf{G}$, a real Lie group with finitely many connected components. Let $K_\infty \subset G_\infty$ be a maximal compact subgroup. Let D be the global Riemannian space attached to G_∞; we have $D = G_\infty/K_\infty$ when $\mathbf{G}$ is semisimple.

Then G_∞ is a real Lie group with finitely many connected components. Let $K_\infty \subset G_\infty$ be a maximal compact subgroup and let $A_\infty \subset G_\infty$ be the connected component of the group of real points of a maximal $\mathbb{Q}$-split torus in the center of G; if $\mathbf{G}$ is semisimple, then $A_\infty = \{1\}$. Then the quotient $D = G_\infty/A_\infty K_\infty$ is a global Riemannian symmetric space. For example, if $\mathbf{G} = \mathrm{SL}_2$ then $K_\infty = \mathrm{SO}_2(\mathbb{R}) \subset \mathrm{SL}_2(\mathbb{R}) = G_\infty$ and the quotient D can be identified with the upper half-plane $\mathcal{H}$. If $\mathbf{G} = \mathrm{Sp}_{2n}$ then $K_\infty = \mathrm{U}_n(\mathbb{R}) \subset \mathrm{Sp}_{2n}(\mathbb{R}) = G_\infty$ and the corresponding symmetric space D is the Siegel upper half-space of real dimension $n(n+1)$, complex dimension $n(n+1)/2$.

For the classical groups, these symmetric spaces can be related to *framed* lattices (lattices with basis) in many cases.

G/K is contractible and Γ acts with finite stabilizers, so $Y(\Gamma) := \Gamma\backslash G/K$ is an orbifold $K(\Gamma, 1)$-space (almost an Eilenberg-MacLane space)—at least that seems to be true. If Γ is torsion free, then $Y(\Gamma)$ is genuinely such, which is why we expect to understand the quotient Y in the same way as understanding Γ.

8.2 Symmetric domain

We mostly follow [BHC62, BS73]. Our goal is to define a natural, nice space on which G acts. We start with the real points.

Let F be a number field and let

$$F_\infty := F \otimes_{\mathbb{Q}} \mathbb{R}. \tag{8.2.1}$$

Then $F_\infty \simeq \prod_{v|\infty} F_v$ where $F_v \simeq \mathbb{R}, \mathbb{C}$ according as if v is a real or complex place.

Let $\mathbf{G}$ be a reductive algebraic group over F.

Remark 8.2.2. In many treatments, it is assumed that $\mathbf{G}$ is connected. In general, by Proposition 7.3.9 $\mathbf{G}$ is a finite disjoint union of cosets of the connected component of the identity $\mathbf{G}^0$, and accordingly we will just end up with a disjoint union of symmetric spaces.

For each $v \mid \infty$, define

$$G_v := \mathbf{G}(F_v); \tag{8.2.3}$$

we put them together into

$$G_\infty := \mathbf{G}(F_\infty) \simeq \prod_{v|\infty} \mathbf{G}(F_v). \tag{8.2.4}$$

We have $G_v \simeq \mathbf{G}(\mathbb{R})$ or $\mathbf{G}(\mathbb{C})$, thinking of F as a subfield of $\mathbb{R}$ or $\mathbb{C}$ under the embedding given by the place v. In particular, G_v is a real Lie group or complex Lie group (so also a real Lie group, but with a preferred complex structure).

Example 8.2.5. Of course if $\mathbf{G} = \mathrm{GL}_n$ then $G_v \simeq \mathrm{GL}_n(\mathbb{R}), \mathrm{GL}_n(\mathbb{C})$ and $G_\infty \simeq \mathrm{GL}_n(\mathbb{R})^r \times \mathrm{GL}_n(\mathbb{C})^c$.

Example 8.2.6. If Q is a nondegenerate quadratic form over F of rank n and $\mathbf{G} = \mathrm{O}(Q)$, then for each real v we have a signature $G_v \simeq \mathrm{O}(p_v, q_v)$ with $p_v + q_v = n$.

Example 8.2.7. If $\mathbf{G}$ is the group associated to $B^\times$ where B is a quaternion algebra over F, then at each ramified real place we have $G_v \simeq \mathbb{H}^\times$; at each split real place we have $G_v \simeq \mathrm{GL}_2(\mathbb{R})$; and at each complex place we have $G_v \simeq \mathrm{GL}_2(\mathbb{C})$.

In other words, if $\mathbf{G}$ is a form then we can get genuinely different Lie groups depending on the embedding, and this is a feature!

Real and complex reductive Lie groups can be understood through their Lie algebras, which are classified.

Recalling our goal, we might as well start contracting what we can from this space. Indeed, a key part of the classification is as follows: the Lie group G_v has a maximal compact subgroup (in the real or complex topology), unique up to conjugation in G_v. Such a maximal connected compact arises as the

fixed subgroup of a Cartan involution θ. When G_v is connected, this is the end of the story; but when G_v is not connected, we allow K_v to be any subgroup containing the maximal *connected* compact subgroup, necessarily finite index in the maximal compact subgroup. It is helpful to have such a choice, indeed we can think of it as being the archimedean component of the level structure which we choose over the finite adeles.

When G_v is semisimple, this is already enough. But if G_v is reductive, then there is an additional factor: roughly speaking, G_v/K_v decomposes as a product of a Euclidean space by a nonpositively curved part which is of genuine interest (coming from the derived group). The Euclidean factor gives directions in which the *center* of G acts noncompactly; the action of the central torus commutes with everything and just translates along these flat directions, and here we throw away this global scale parameter (it will not feature in the modular forms we want to compute).

We let

$$K_\infty := \prod_{v|\infty} K_v \leq G_\infty. \tag{8.2.8}$$

Let $\mathbf{S}_\mathbf{G}$ be the maximal F-split torus in the center $\mathbf{Z}(\mathbf{G})$, and let

$$A_{\mathbf{G},\infty} = A_\infty := \mathbf{S}_\mathbf{G}(F_\infty)^0 \tag{8.2.9}$$

be the connected component of the group of F_∞-points of this split central torus. We have $A_\infty \leq Z(G_\infty)$ inside the center of G_∞. (Might need to reconsider notation here, since we might as well define $\mathbf{A}_\mathbf{G}$ and write $A^0_{\mathbf{G},\infty}$?)

Example 8.2.10. Suppose F has r real embeddings and c complex embeddings. For $\mathbf{G} = \mathrm{GL}_{2,F}$,

$$D \xrightarrow{\sim} (\mathcal{H}_2)^r \times (\mathcal{H}_3)^c$$

where $\mathcal{H}_2, \mathcal{H}_3$ are hyperbolic spaces of dimension $2, 3$ (real dimension $2r + 3c$).

Remark 8.2.11. To work over $\mathbb{Q}$ instead (as in section 7.4), we could replace $\mathbf{G}$ over F by $\mathrm{Res}_{F|\mathbb{Q}}(\mathbf{G})$. But then the maximal F-split torus in $\mathbf{Z}(\mathbf{G})$ can differ from the maximal $\mathbb{Q}$-split torus in $\mathbf{Z}(\mathrm{Res}_{F|Q}(\mathbf{G}))$. For example, for $\mathbf{G} = \mathrm{GL}_{1,F} = \mathbb{G}_{\mathrm{m},F}$ itself an F-split torus we have $\mathbf{S}_\mathbf{G} = \mathbf{G}$ and $A_{\mathbf{G},\infty} = (F_\infty^\times)^0 \simeq (\mathbb{R}^\times_{>0})^r \times (\mathbb{C}^\times)^c$. However, for $\mathbf{G} = \mathrm{Res}_{F|\mathbb{Q}}(\mathrm{GL}_{1,F})$ we have $\mathbf{G}(\mathbb{Q}) = F^\times$ and so the maximal $\mathbb{Q}$-split torus is just $\mathbf{S}_\mathbf{G} = \mathbb{G}_{\mathrm{m},\mathbb{Q}}$ and accordingly $\mathbf{S}_\mathbf{G}(\mathbb{R})^0 = \mathbb{R}_{>0}$.

For $\mathbf{G} = \mathrm{GL}_{2,F}$ with F totally real, we get a Hilbert modular variety. Taking the restriction of scalars $\mathbf{G} = \mathrm{Res}_{F|\mathbb{Q}}(\mathrm{GL}_{2,F})$ we instead get a torus bundle over a Hilbert modular variety (so in general not an algebraic variety). For $\mathbf{G} = \mathrm{SL}_{2,F}$ and its restriction of scalars, we again get the Hilbert modular variety.

Because A_∞ is central, right multiplication by A_∞ on G_∞ commutes with right multiplication by K_∞, and it descends to a well-defined right action on G_∞/K_∞

$$(gK_\infty) \cdot a = gaK_\infty$$

for $a \in A_\infty$.

We then define

$$\begin{aligned} D_v &:= G_v/K_v \\ D_\infty &:= \prod_{v|\infty} D_v \end{aligned} \tag{8.2.12}$$

and call D_∞ a **global symmetric domain** attached to $\mathbf{G}$ (for the subgroup K_∞).

Proposition 8.2.13. *Every connected component of D_∞ is diffeomorphic to Euclidean space.*

Proof. Without loss of generality we may suppose that K_∞ is a maximal compact subgroup. Choose a Cartan involution θ of G_∞ with fixed subgroup K_∞, and let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the corresponding decomposition of Lie algebras. The global Cartan decomposition gives a diffeomorphism

$$\begin{aligned} \mathfrak{p} &\xrightarrow{\sim} G_\infty/K_\infty \\ X &\mapsto \exp(X)K_\infty. \end{aligned} \tag{8.2.14}$$

Thus $G_\infty/K_\infty \simeq \mathbb{R}^p$ where $p = \dim \mathfrak{p}$.

The central split torus A_∞ is stable under θ, acting freely and properly on G_∞/K_∞ by right translation; its orbits correspond to translations along an affine subspace of $\mathfrak{p}$. Hence the quotient $(G_\infty/K_\infty)/A_G \simeq G_\infty/(A_G K_\infty)$ is again diffeomorphic to a Euclidean space, now of dimension $p - \dim A_\infty$. □

In particular, each component of D_∞ is connected and contractible. Now G_∞ is locally compact (as a real or complex Lie group), so it has a Haar measure unique up to scaling. This descends to D_∞, giving each component the structure of a uniquely geodesic space.

8.3 Spaces of lattices and quadratic forms

In this section, we consider the case $G_\infty = \mathrm{GL}_n(\mathbb{R})$, for example coming from $\mathbf{G} = \mathrm{GL}_n$ over $F = \mathbb{Q}$ and we characterize the global symmetric space using lattices.

8.3.1. For $g \in G_\infty$, the rows of g span a lattice $\Lambda \subseteq \mathbb{R}^n$ (of rank n); so we can think of G_∞ as being the space of *framed* lattices (i.e., *lattice with basis*). The group $K_\infty = \mathrm{O}(n)$ acts by reflection and rotation of the lattice from the ambient $\mathbb{R}^n$ and the group $A_\infty \simeq \mathbb{R}_{>0}$ corresponds to rescaling the lattice, corresponding to forgetting about the particular choice coordinates on $\mathbb{R}^n$. So we can think of D_∞ as being the space of *shapes* of framed lattices (with the shape being determined by the relative legnths and angles of the vectors in the frame).

If instead we take $\mathbf{G} = \mathrm{SL}_n$, then we get $G_\infty = \mathrm{SL}_n(\mathbb{R})$, $K_\infty = \mathrm{SO}_n(\mathbb{R})$ and $A_\infty = \{1\}$; we can always rescale a lattice to have absolute determinant 1, so now D_∞ gives the space of shapes of *oriented* framed lattices.

To get explicit coordinates on this space, we use the QR decomposition (or more precisely, the LQ decomposition): every matrix $A \in \mathrm{M}_n(\mathbb{R})$ can be written uniquely in the form $A = LQ$ where Q is orthogonal and L is lower triangular (the transpose of the QR decomposition). This is algorithmic and efficient, indeed it is the output of the Gram–Schmidt orthogonalization process.

Thus the (real) dimension is

$$\dim D_{\mathrm{GL}_{n,\mathbb{Q},\infty}} = \dim D_{\mathrm{SL}_{n,\mathbb{Q},\infty}} = \frac{n(n+1)}{2} - 1 \tag{8.3.2}$$

(the minus 1 coming from the determinant or covolume), which gives dimension $2, 5, 9, 14, 20, \ldots$ for $n = 2, 3, 4, 5, \ldots$.

In the next section (**??**), we remove the choice of frame by taking a further quotient by change of basis of the lattice!

Example 8.3.3. In the case $n = 2$, given a framed lattice of rank 2 then we can rotate until the first vector lies along the positive real line, rescale so that it is $(1, 0)$, and then reflect if necessary to get the second basis vector in the upper half-plane, which is D_∞! For SL_2, we are assured without reflection that the second basis vector is automatically in the upper half-plane, by the orientation.

Example 8.3.4. Now for $n = 3$ and $\mathbf{G} = \mathrm{SL}_3$, we have six real parameters: by RQ decomposition, a point of D_∞ can be represented by

$$\begin{pmatrix} t_1 & 0 & 0 \\ x_{12} & t_2 & 0 \\ x_{13} & x_{23} & t_3 \end{pmatrix}, \qquad t_1, t_2, t_3 > 0, \quad x_{12}, x_{13}, x_{23} \in \mathbb{R}.$$

We think of t_1, t_2, t_3 as being *scale* parameters, encoding lengths of the basis vectors; and x_{12}, x_{13}, x_{23} as being *shear* parameters, encoding the angles between the basis vectors.

Geometrically, after rotating $\mathbb{R}^3$ so that the first basis vector lies on the x-axis and the second lies in the xy-plane, the columns define an upper-triangular matrix. For general n, this is what is provided by Gram-Schmidt orthogonalization.

8.3.5. Continuing with the previous paragraph of $\mathbf{G} = \mathrm{GL}_n$ over $F = \mathbb{Q}$, we reinterpret the *shape* by encoding it as a quadratic form. Given $g \in G_\infty = \mathrm{GL}_n(\mathbb{R})$, define

$$T_g := gg^{\mathsf{t}} \tag{8.3.6}$$

Then T_g is symmetric and positive definite with $\det T_g = (\det g)^2$, so T_g defines a quadratic form by $Q_g(x) = x^{\mathsf{t}} T_g x$ for $x \in \mathbb{R}^n$ (column vectors). It is sometimes useful to divide Q by 2, but nevermind!

The reason this encodes shape: multiplying on the right by $k \in K_\infty = \mathrm{O}_n(\mathbb{R})$ yields

$$T_{gk} = (gk)(gk)^{\mathsf{t}} = gkk^{\mathsf{t}}g^{\mathsf{t}} = gg^{\mathsf{t}} = T_g \tag{8.3.7}$$

since $kk^{\mathbf{t}} = 1$ (the condition characterizing belonging to $\mathrm{O}_n(\mathbb{R})$!).

Conversely, given a symmetric positive-definite matrix $T \in \mathrm{M}_n(\mathbb{R})$, there exists $g \in \mathrm{GL}_n(\mathbb{R})$ such that $T = T_g$ by the *Cholesky decomposition*, and g is well-defined up to right multiplication by $\mathrm{O}_n(\mathbb{R})$.

Thus we have an identification

$$D_\infty \xleftrightarrow{\sim} \mathrm{M}_n(\mathbb{R})_{\mathrm{sym},>0}/\mathbb{R}_{>0} \tag{8.3.8}$$

of D_∞ with positive-definite, symmetric real matrices up to scaling (homothety). Strictly speaking, the scalar matrix t acts by rescaling $T_g \mapsto t^2 T_g$, which is good because it should preserve the positive-definiteness.

For $\mathrm{SL}_n(\mathbb{R})$, we have the same, but with unimodular lattices.

In particular, this gives a fast way to remember $\dim D_\infty = n(n+1)/2 - 1$.

Example 8.3.9. When $n = 2$, given $g = \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ we get

$$T_g = \begin{pmatrix} a^2+1 & ab \\ ab & b^2 \end{pmatrix}$$

and the quadratic form

$$(a^2+1)x^2 + 2abxy + b^2y^2.$$

8.4 Hyperbolic spaces

There is a nice sequence of models, generalizing the familiar upper half-plane, for the local symmetric space when the group G_v has real rank one, classified in Proposition 7.6.2.

Example 8.4.1. Real hyperbolic space has three possible models. First, there is the **upper half-space** model

$$\mathbf{H}^n := \mathbb{R}^{n-1} \times \mathbb{R}_{>0} \tag{8.4.2}$$

for which we have seen $\mathbf{H}^2$ and $\mathbf{H}^3$. The action of $\mathrm{SO}(n,1)^0$ is by Möbius transformations of $\mathbb{R}^{n-1} \cup \{\infty\}$ extended to the interior: see work of Dupuy–Hilado–Ingalls–Logan [DHIL24] for the action of the spin group by linear fractional transformations coming from the Clifford algebra.

Second, there is the **hyperboloid model**

$$\mathbf{L}^n := \{(t, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} : Q(x) = 1, t > 0\} \tag{8.4.3}$$

where

$$Q(t, x_1, \ldots, x_n) = t^2 - x_1^2 - x_2^2 - \cdots - x_n^2. \tag{8.4.4}$$

The isometry group $\mathrm{SO}(n,1)^0$ acts linearly on $\mathbb{R}^{n+1}$.

Finally, there is the **Poincaré ball** model

$$\mathbf{D}^n := \{u \in \mathbb{R}^n : \|u\| < 1\}, \tag{8.4.5}$$

and the Klein/projective model; you get it by projecting the hyperboloid through the origin to the affine slice $t = 1$. Accordingly the group acts by projective transformations.

For comparisons of these in the case $n = 2$ see [Voi21, §33.7].

Example 8.4.6. Next is complex hyperbolic space. We replace the quadratic form (8.4.4) with the Hermitian form

$$Q(z_1, \dots, z_n) = |z_1|^2 - |z_2|^2 - \cdots - |z_n|^2 \tag{8.4.7}$$

with $z_i \in \mathbb{C}$, and define

$$\mathbf{H}^n_{\mathbb{C}} := \{(z_1 : \cdots : z_n) \in \mathbb{P}^n(\mathbb{C}) : Q(z) < 0\}.$$

Then $(n, 1)$ acts by projective linear transformations preserving the form. There is a model like the upper half-space model.

There is also quaternionic hyperbolic space, replacing $\mathbb{C}$ by $\mathbb{H}$ to get $\mathrm{Sp}(n, 1)$, and finally a Cayley hyperbolic plane to get an exceptional rank-one space.

8.5 Representations

Algebraic representations, highest weight vectors, and things that are more general. Ultimately these will correspond to the *weight* of the modular form.

8.6 Arithmetic groups

For arithmetic subgroups and reduction theory over number fields see [PR94, Bor69].

The group G_∞ acts by left-multiplication on the global symmetric domain D_∞ (8.2.12) by isometries, analogous to the action of $\mathrm{GL}_2(\mathbb{R})$ on the upper half-plane [Voi21, 33.3.10]. (To get the union of upper and lower half-planes, we can take $K_\infty = \mathrm{SO}(2)$ instead of $\mathrm{O}(2)$, no longer a maximal compact. Are there other examples where it is good to work with subcompacts?)

Remark 8.6.1. Let $\Gamma \leq G_\infty$ be a discrete group. Then Γ acts properly on D_∞?

Let

$$G := \mathbf{G}(F) \tag{8.6.2}$$

be the F-rational points. Under the diagonal embedding we have naturally $G \leq G_\infty$.

Definition 8.6.3. A subgroup $\Gamma \leq G = \mathbf{G}(F)$ is **arithmetic** if Γ is commensurable with $\mathbf{G}(\mathbb{Z}_F) \leq \mathrm{GL}_n(\mathbb{Z}_F)$ for some F-embedding $\mathbf{G} \hookrightarrow \mathrm{GL}_n$.

Recall that two groups are **commensurable** if the intersection is finite index in both; it is an exercise to show that if Γ is commensurable with $\mathrm{GL}_n(\mathbb{Z}_F)$ under some F-embedding then it is so under all.

Let $\Gamma \leq G$ be an arithmetic subgroup. We then form the **locally symmetric domain**

$$Y = Y(\Gamma) := \Gamma \backslash D_\infty. \tag{8.6.4}$$

The domain Y is a good orbifold, analogous to the open modular curve. If further Γ is torsion-free then Y is a manifold.

Later, when we work with compactifications, we will need a stronger condition which precludes boundary stabilizers as well.

Definition 8.6.5. We say that Γ is **neat** if for every $\gamma \in \Gamma$ the subgroup of $(\mathbb{Q}^{\mathrm{al}})^\times$ generated by the eigenvalues of γ (in a faithful representation of $\mathbf{G}$) contains no nontrivial roots of unity.

Immediately neat implies torsion-free, but it is strictly stronger; and the condition of being neat is independent of the representation.

Proposition 8.6.6. *Every arithmetic subgroup Γ contains a neat subgroup of finite index.*

Proof. □

Theorem 8.6.7. *The quotient Y has finite volume.*

Proof. See Borel–Harish-Chandra [BHC62]. □

When $\mathbf{G}$ is semisimple, we sometimes then say that $\Gamma \leq G_\infty$ is a **lattice** (as another way of saying that the quotient has finite volume). So not just discrete (like $\mathbb{Z} \subseteq \mathbb{R}^2$), but big enough (like $\mathbb{Z}^2 \subseteq \mathbb{R}^2$)!

Remark 8.6.8. In general, we will want to reduce to the case where Γ is torsion-free, which feels like some loss of generality. But in truth we work like we would with a stack: we pass to $\Gamma' \triangleq \Gamma$ of finite index and then get what we need geometrically or topologically, then take the quotient Γ/Γ'. In this way, we can pass to Γ' neat without loss of generality.

As with modular curves, we are primarily interested in those finite index subgroups

Definition 8.6.9. Let $\mathfrak{N} \subseteq R$ be a nonzero ideal. The **principal congruence subgroup** of level $\mathfrak{N}$ is the kernel of the map $\mathcal{G}(R) \to \mathcal{G}(R/\mathfrak{N})$.

Definition 8.6.10. An arithmetic group Γ is **congruence** if it contains a principal congruence subgroup for some $\mathfrak{N}$.

Remark 8.6.11. Not all arithmetic groups are congruence: already there are a wealth of noncongruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$. But for $\mathrm{SL}_n(\mathbb{Z})$ with $n \geq 3$ or $\mathrm{Sp}_{2n}(\mathbb{Z})$ for $n \geq 2$, every arithmetic group is congruence; and in many other cases, the failure of arithmetic groups to be congruence is measured (globally) by a finite group. So although congruence subgroups are special among arithmetic groups, for our purposes we have not lost anything.

8.7 Hermitian symmetric domains

8.7.1. The symmetric space D_v for a maximal compact K_v is **of Hermitian type** when it carries a G-invariant Hermitian (complex) structure. This happens exactly when the center of the Lie algebra of K_v is nontrivial, or equivalently when $Z(K_v)$ has a circle factor. The simple real groups of Hermitian type are those whose Lie algebras match the Lie algebras of:

- $\mathrm{SO}(2,n)^0$ with $n \geq 1$ (not simple for $n = 2$), with
$$D = \mathrm{SO}(2,n)^0/(\mathrm{SO}(2) \times \mathrm{SO}(n))$$
a tube domain;
- $\mathrm{SU}(p,q)$ with $p, q \geq 1$, with $D = \mathrm{SU}(p,q)/\mathrm{S}(\mathrm{U}(p) \times \mathrm{U}(q))$, a complex Grassmannian domain;
- $\mathrm{Sp}(2n,\mathbb{R})$ with $n \geq 1$, with $D = \mathrm{Sp}(2n,\mathbb{R})/U(n)$ the Siegel upper half-space;
- the quaternionic orthogonal group $\mathrm{SO}^*(2n)$; or
- two real forms of the exceptional groups E_6 and E_7.

In these cases, $Y(\Gamma)$ is a quasi-projective variety, a **locally symmetric variety** or a *Shimura variety*.

Sorry about the notation $\mathrm{Sp}(2n,\mathbb{R})$ for the subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$ which preserves the standard nondegenerate alternating form: we might write $\mathrm{Sp}(2n)$ but apparently this can be confused for the compact form (quaternionic unitary group).

Example 8.7.2. For $\mathrm{Sp}(2n,\mathbb{R})$ the Siegel upper half-space is now a target for moduli spaces of abelian varieties.

Example 8.7.3. For $\mathrm{SO}(2,n)$, a tube domain, following Bruinier, with the special cases $n = 2, 3$ related to previous.

Hermitian symmetric domains are also the natural target for period maps, as in the case of elliptic curves!

8.8 Borel–Serre compactification

A reference for this section is the original paper by Borel–Serre [BS73]; we also follow Goresky [Gor05, §4].

Example 8.8.1. For $\mathbf{G} = \mathrm{SL}_2$ over $\mathbb{Q}$, we have $D_\infty \simeq \mathbf{H}^2$. For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is one Γ-conjugacy class of proper parabolics, represented by the upper triangular Borel subgroup. The unipotent radical is $\mathrm{U}(\mathbb{R}) \simeq \mathbb{R}$ and $\Gamma \cap \mathrm{U}(\mathbb{Q}) \simeq \mathbb{Z}$. Thus the Borel–Serre boundary stratum is a circle:
$$(\Gamma \cap \mathrm{U}(\mathbb{Q}))\backslash \mathrm{U}(\mathbb{R}) \simeq \mathbb{Z}\backslash\mathbb{R} \simeq S^1. \tag{8.8.2}$$

Geometrically, a cusp neighborhood in $Y = \Gamma\backslash\mathfrak{H}$ is $(S^1) \times (T, \infty)$ and the Borel–Serre compactification replaces it by $(S^1) \times [T, \infty]$, adding a boundary circle.

If you were hoping for the one-point compactification, maybe that is addressed by the reductive Borel–Serre compactification, or something else?

Do we get the same thing for GL_2? The additional element in $\mathrm{GL}_2(\mathbb{Z}) \smallsetminus \mathrm{SL}_2(\mathbb{Z})$ does not seem to matter.

Example 8.8.3. Next $\mathrm{GL}_{2,F}$ where F is an imaginary quadratic field. We saw in section 6.1 that the cusps are in natural bijection with the ideal class group of $\mathbb{Z}_F$. Each cusp is stabilized by a parabolic subgroup, and so the orbits are specified by conjugacy classes of such. For example, the trivial ideal class is stabilized by the upper-triangular subgroup, which is an extension of $\mathbb{Z}_F^\times$ by $\mathbb{Z}_F$.

The Borel–Serre compactification adds the quotient of $\mathbb{Z}_F \otimes \mathbb{R} \simeq \mathbb{C}$ by the translation group of the parabolic, which is a real 2-torus in fact $\mathbb{C}/\mathfrak{a}$ which has the struture of a complex elliptic curve with CM by $\mathbb{Z}_F$. Thus X^{BS} is a compact 3-manifold with boundary a disjoint union indexed by cusps.

Example 8.8.4. Next we consider $\mathbf{G} = \mathrm{SL}_n$ over $\mathbb{Q}$. Then parabolics correspond to flags in $\mathbb{Q}^n$, or equivalently to block upper triangular subgroups. If $\mathbf{P}$ corresponds to a partition $n = n_1 + \cdots + n_r$, then $\mathbf{U_P}$ is the block-upper unipotent group and $\mathbf{L_P}$ is $\mathrm{SL}_n(\mathbb{Z}) \cap \prod_{i=1}^r \mathrm{GL}_{n_i}$. The boundary stratum is a bundle over a product of lower-rank locally symmetric spaces, with compact fiber $\Gamma_U\backslash\mathbf{U_P}(\mathbb{R})$. Corners correspond to refining flags.

Example 8.8.5. For Sp_{2g} over $\mathbb{Q}$, D is Siegel upper half-space $\mathfrak{H}_g$. Parabolics correspond to isotropic flags, U_P is typically a Heisenberg-type unipotent group, and boundary strata again form nilmanifold bundles over Siegel varieties of smaller genus.

In this case, there are other compactifications, including reductive Borel–Serre (see our end remark), Baily–Borel, and toroidal compactifications, and they can be compared explicitly.

Theorem 8.8.6 (Borel–Serre)**.** *The space D_∞^* is a smooth manifold with corners containing D_∞ as a dense open submanifold. The action of Γ on D_∞ extends to a proper action on D_∞^*, and the quotient*

$$X^{\mathrm{BS}}(\Gamma) = X(\Gamma) := \Gamma\backslash D_\infty^*$$

is compact. If Γ is neat, then $X(\Gamma)$ is a compact manifold with corners whose interior is $Y(\Gamma) = \Gamma\backslash D_\infty$. Moreover, the inclusion $D_\infty \hookrightarrow D_\infty^*$ is a Γ-equivariant homotopy equivalence, hence $Y(\Gamma) \hookrightarrow X(\Gamma)$ is a homotopy equivalence.

Proof sketch. The manifold-with-corners structure is local in the cusp directions and comes from the local models $U_P(F_\infty) \times (\mathbb{R}_{\geq 0})^{r_P} \times D_P$ in horospherical

coordinates. Contractibility follows from a Γ-equivariant deformation retraction that pushes points off the boundary in the $(\mathbb{R}_{\geq 0})^{r_P}$-coordinates and uses compatibility under gluing. Compactness of the quotient is where reduction theory enters: one covers D by finitely many Γ-translates of Siegel sets, and the only noncompactness is divergence in A_P/A_G-directions, which become compact after replacing $(\mathbb{R}_{>0})^{r_P}$ by $(\mathbb{R}_{\geq 0})^{r_P}$ and gluing faces across parabolics. □

We conclude with a description of the boundary strata and their structure.

Finally, there is some combinatorics in the boundary, which can be described via buildings (see [BT65]). Let $\mathcal{B}(\mathbf{G})$ be the spherical Tits building of $\mathbf{G}$ over F, whose simplices correspond to proper F-parabolics.

Theorem 8.8.7. *The boundary* $\mathrm{bd}(D^*_\infty) := D^*_\infty \setminus D$ *has the homotopy type of* $\mathcal{B}(\mathbf{G})$. *In particular, it is homotopy equivalent to a wedge of spheres of dimension* $\mathrm{rk}_F(\mathbf{G}) - 1$.

Proof sketch. One covers $\mathrm{bd}\, D^*_\infty$ by closed subsets indexed by maximal proper parabolics, arranged so that all finite intersections are either empty or contractible and correspond exactly to common refinements. The nerve of this cover is the building, hence the boundary has the same homotopy type. Finally, Solomon–Tits computes the homotopy type of the building. □

We will attach natural families of relative cycles to parabolics, generalizing modular symbols. We will return to this perspective in Chapter 3 once we define Hecke operators as double cosets and compare them with correspondences on $X(\Gamma)$.

8.9 Reductive Borel–Serre compactification

There are other compactifications which are also useful; we record a bit here, following Goresky [Gor05] and Borel–Ji [BJ02].

The Borel–Serre boundary strata are quotients of nonreductive groups, bundles with unipotent fibers. The idea of Zucker is to collapse these unipotent fibers, leaving strata that are themselves locally symmetric for Levi quotients.

Example 8.9.1. In the modular curve example, Borel–Serre adds boundary circles at cusps, while reductive Borel–Serre collapses each circle to a point.

8.10 Compactifications in the Hermitian case

When D_∞ is Hermitian (see 8.7.1), then there are two further compactifications: the Baily–Borel (Satake) compactification which is a complex projective algebraic variety, usually singular, and toroidal compactifications (providing various resolutions of singularities of the Baily–Borel).

Exercises

8.1. In Example 8.3.9, show directly that every positive-definite binary quadratic form is of the given form, and interpret this in terms of the Cholesky decomposition.

Chapter 9

Adelic quotients

9.1 Overview

Computational application and conceptual mathematical understanding reinforce and extend one another. For this reason, we cannot neglect the role of the adeles: they not only explain issues with class groups and field of definition, but they help in the abstraction of Hecke operators. All of these can be translated into explicit algorithms.

9.2 Level structure

There are many good references for adeles and ideles; for continuity we follow Voight [Voi21, Chapter 27]. Our main motivation is to separate local computations (at a single prime) from global bookkeeping.

Let

$$\widehat{F} := \prod_{\mathfrak{p}}{}' F_{\mathfrak{p}} \tag{9.2.1}$$

be the finite adeles of F, the restricted direct product under the compact product $\prod_{\mathfrak{p}} R_{\mathfrak{p}}$. The finite adeles pair with the infinite adeles to give the full adele ring

$$\underline{F} := \widehat{F} \times F_\infty = \prod_{v}{}' F_v. \tag{9.2.2}$$

Let $\mathbf{G} \leq \mathrm{GL}_{n,F}$ be an embedded reductive group over F. As above, this gives an integral model $\mathcal{G}$ of $\mathbf{G}$ with for example $\mathcal{G}(R) = \mathbf{G}(F) \cap \mathrm{GL}_n(R)$.

For each prime $\mathfrak{p}$, the group $\mathbf{G}(F_{\mathfrak{p}})$ is locally compact and the subgroup $\mathcal{G}(R_{\mathfrak{p}}) \leq \mathbf{G}(F_{\mathfrak{p}})$ is a compact open subgroup.

We then similarly form

$$\widehat{G} := \prod_{\mathfrak{p}}{}' \mathbf{G}(F_{\mathfrak{p}}) \tag{9.2.3}$$

with respect to the compact product $\prod_{\mathfrak{p}} \mathcal{G}(R_{\mathfrak{p}})$. This restricted direct product is independent of the choice of representation (any two will yield a different

integral structure at only finitely many primes). We have already made friends with G_∞.

A **level** is a compact open subgroup $\widehat{K} \leq \widehat{G}$.

The finite and infinite parts go together to form

$$\underline{G} := \mathrm{G}(\underline{F}) = \widehat{G} \times G_\infty. \tag{9.2.4}$$

It is convenient to have a description which treats archimedean and nonarchimedean places in a parallel fashion; but in truth, they behave complelely differently in the description of automorphic forms, so we also have good notation for the two pieces.

9.3 Double cosetification

Here come the double cosets! The OG double coset is the class group of a number ring R, as follows. A fractional R-ideal $\mathfrak{a}$ is locally principal, so if we choose a generator $\mathfrak{a}_\mathfrak{p} = a_\mathfrak{p} R_\mathfrak{p}$ the element $a_\mathfrak{p} \in F_\mathfrak{p}^\times$ is well-defined up to multiplication by $R_\mathfrak{p}^\times$, and we have $a_\mathfrak{p} = 1$ (or $a_\mathfrak{p} \in R_\mathfrak{p}^\times$) for all but finitely many $\mathfrak{p}$. Thus we get a map from the group of fractional R-ideals to $\widehat{F}^\times / \widehat{R}^\times$, and indeed this is an isomorphism. If we want ideal classes, then we impose the equivalence relation obtained by multiplication by $F^\times$. We obtain an isomorphism $\mathrm{Cl}\, R \xrightarrow{\sim} F^\times \backslash \widehat{F}^\times / \widehat{R}^\times$.

This way of organizing things has a vast generalization. Let $\widehat{K} \leq \widehat{G}$ be a level. We then form

$$Y = Y(\widehat{K}) := G \backslash (\widehat{G}/\widehat{K} \times D_\infty) \tag{9.3.1}$$

where G acts diagonally. Recall that D_∞ is almost G_∞ / K_∞, so the two factors are quite parallel!

The natural projection

$$G \backslash (\widehat{G}/\widehat{K} \times D) \to G \backslash \widehat{G} / \widehat{K} \tag{9.3.2}$$

is continuous. We can think of $G \backslash \widehat{G} / \widehat{K}$ as a *class set* attached to G and $\widehat{K}$: it measures a failure of a local-global principle.

Theorem 9.3.3. *The set $G \backslash \widehat{G} / \widehat{K}$ is finite.*

Example 9.3.4. For example, if $\mathrm{G} = \mathrm{GL}_{n,F}$ and $\widehat{K} = \mathrm{GL}_n(\widehat{R})$, then by strong approximation the determinant map

$$\det \colon \mathrm{GL}_n(F) \backslash \mathrm{GL}_n(\widehat{F}) / \mathrm{GL}_n(\widehat{R}) \to F^\times \backslash \widehat{F}^\times / \widehat{R}^\times = \mathrm{Cl}\, R \tag{9.3.5}$$

is an isomorphism onto the class group of R.

Proof sketch, Theorem 9.3.3. Since $\widehat{K} \leq \widehat{G}$ is open, the quotient $\widehat{G}/\widehat{K}$ is quite intense: it has the discrete topology (every set is open). Indeed, every coset of

$\widehat{K}$ is open (by translation) which (by the quotient topology) implies that every point in $\widehat{G}/\widehat{K}$ is open!

So if we can show that the quotient $G\backslash\widehat{G}/\widehat{K}$ is compact, then it is finite. To show cocompactness, we show that every double coset has a representative in a fixed compact subset of $\widehat{G}$: there exists a compact subset C such that $GC\widehat{K} = G\backslash\widehat{G}/\widehat{K}$, i.e., every class in $G\backslash\widehat{G}/\widehat{K}$ has a representative in C, reminiscent of how we prove the finiteness of the class group.

Once we know Example 9.3.4, then we can conclude using a faithful representation $\rho\colon \mathrm{G} \to \mathrm{GL}_{n,F}$. □

In light of Theorem 9.3.3, we may choose a finite set of representatives

$$G\backslash\widehat{G}/\widehat{K} = \bigsqcup_{i}^{<\infty} G\widehat{\beta}_i\widehat{K} \tag{9.3.6}$$

with $\widehat{\beta}_i \in \widehat{G}$. We then define

$$\Gamma_i := \widehat{\beta}_i\widehat{K}\widehat{\beta}_i^{-1} \cap G. \tag{9.3.7}$$

Then Γ_i is a congruence arithmetic group.

Moreover,

$$Y(\widehat{K}) = \bigsqcup_i Y(\Gamma_i) \tag{9.3.8}$$

where $Y(\Gamma_i) = \Gamma_i\backslash D_\infty$ is a good orbifold as previously considered. (Usually, K_∞ is a maximal compact so D_∞ is connected, in which case these are the connected components of $Y(\widehat{K})$.)

In other words, the introduction of the adeles so far has helped us to avoid an ugly class number 1 hypothesis by indicating how the naturally defined symmetric space, incorporating all places of F, is naturally a disjoint union of locally symmetric spaces indexed by a class set.

9.4 Hecke operators

Not only do adeles help organize according to class groups, it also provides a cleaner definition of Hecke operators.

Hecke operators can be thought of in many equivalent ways: as correspondences, as averaging over neighors. Here we think of them group-theoretically.

Let $\mathcal{H}(\widehat{K})$ be the space of locally constant, compactly supported, $\widehat{K}$-bi-invariant functions on $\widehat{G}$. Then $\mathcal{H}(\widehat{K})$ is a ring under convolution, called the (spherical?) **Hecke algebra**, and is generated by the characteristic functions $T(\widehat{g})$ of double cosets $\widehat{K}\widehat{g}\widehat{K}$ for $\widehat{g} \in \widehat{G}$.

For $\widehat{g} \in \widehat{G}$, the double coset $\widehat{K}\widehat{g}\widehat{K}$ defines a correspondence

$$X(\widehat{K}) \leftarrow X(\widehat{K} \cap \widehat{g}^{-1}\widehat{K}\widehat{g}) \to X(\widehat{K}), \tag{9.4.1}$$

obtained from pull back under the first projection and then pushforward under the second projection. Ultimately, this will induce operators on functions and (co)homology, at the chain level it becomes a finite sum over coset representatives analogous to the usual definition for classical modular forms.

Chapter 10

Modular forms on reductive groups

10.1 Automorphic forms

To a first approximation, automorphic forms are analytic functions on the adelic points of a reductive group.

Let F be a number field and let $\mathbf{G}$ be a reductive group over F with $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$ and finite adele ring $\widehat{F}$. Recalling our notation, we let $G_\infty = \mathbf{G}(F_\infty)$, $\widehat{G} = \mathbf{G}(\widehat{F})$, $\underline{G} = \mathbf{G}(\underline{F}) = G_\infty \times \widehat{G}$, and finally $G = \mathbf{G}(F)$. Then G acts diagonally on the left on $G_\infty \times \widehat{G}$ by left multiplication.

Let $K_\infty \leq G_\infty$ contain the maximal connected compact subgroup and let $\widehat{K} \leq \widehat{G}$ be a compact open subgroup. The globally symmetric space is $D_\infty = G_\infty/(A_\infty K_\infty)$ and the locally symmetric space is $Y(\widehat{K}) = G\backslash(D_\infty \times \widehat{G}/\widehat{K})$.

A K_∞**-type** is a finite-dimensional complex representation $\tau\colon K_\infty \to \mathrm{GL}(V)$. For brevity, we will often abbreviate this by just V but we have not forgotten the representation it affords. A K_∞-type defines a homogeneous vector bundle $\mathscr{V}$ on $Y(\widehat{K})$

$$G\backslash(\widehat{G}/\widehat{K} \times G_\infty/A_\infty \times V)/K_\infty \to Y(\widehat{K}) = G\backslash(\widehat{G}/\widehat{K} \times G_\infty/A_\infty)/K_\infty$$

where G acts diagonally on $\widehat{G}/\widehat{K} \times G_\infty/A_\infty$ and trivially on V, and K_∞ acts on the right trivially on $\widehat{G}/\widehat{K}$ and by

$$(g_\infty, v)k_\infty = (g_\infty k_\infty, \tau(k_\infty)^{-1}v).$$

A section of $\mathscr{V}$ is therefore a V-valued invariant function on $\widehat{G}/\widehat{K} \times G_\infty/A_\infty$, i.e.,

$$f(g(\widehat{x}, g_\infty)k_\infty) = \tau(k_\infty)^{-1} f(\widehat{x}, g_\infty) \tag{10.1.1}$$

for all $g \in G$, $\widehat{x} \in \widehat{G}$, $g_\infty \in G_\infty$, and $k_\infty \in K_\infty$.

We view (10.1.1) as the generalization of weight k invariance for classical modular forms.

Definition 10.1.2. An **automorphic form** of level $\widehat{K}$ and K_∞-type V is a function

$$f\colon G(\underline{F}) = G_\infty/A_\infty \times \widehat{G} \to V$$

which is smooth on G_∞, locally constant on $\widehat{G}$, and satisfies the following properties:

(i) $f(g\underline{x}) = f(\underline{x})$ for all $g \in G$ and $\underline{x} \in \underline{G}$ (left G-invariant);
(ii) $f(\underline{x}\widehat{k}) = f(\underline{x})$ for all $\widehat{k} \in \widehat{K}$ (right $\widehat{K}$-invariant);
(iii) $f(\underline{x}k_\infty) = \tau(k_\infty)^{-1} f(\underline{x})$ for all $k_\infty \in K_\infty$ (right K_∞-equivariant);
(iv) f is $Z(U(\mathfrak{g}_\mathbb{C}))$-*finite*, where $\mathfrak{g}_\mathbb{C} := \mathrm{Lie}(G_\infty) \otimes_\mathbb{R} \mathbb{C}$ (f satisfies a "nice" differential equation); and
(v) f has moderate growth.

Condition (iv) says that f lies in a finite-dimensional generalized eigenspace for the action of $Z(U(\mathfrak{g}_\mathbb{C}))$ and implies that in fact f is real analytic. Both conditions (iv) and (v) are important, but certainly not the focus of these notes.

Definition 10.1.2 generalizes the very useful "sections of a line bundle" point of view; this can be reinterpreted back in terms of $\mathbb{C}$-valued functions on G_∞ by asking that that they are K_∞-finite, then decomposing according to representations of K_∞, and then we are back to the above definition.

Similarly, if we do not want to specify $\widehat{K}$ in advance, we can also consider the larger space of functions without condition (ii); but such functions f are locally constant and so constant on the cosets of a compact open subgroup $\widehat{K} \leq \widehat{G}$, so arise with some level. Hence this larger space is the direct limit (union) of spaces with level.

This definition is really great in its generality, but it does not look altogether amenable to computation: after all, in these notes we are looking for some kind of finite-dimensional vector space with a Hecke action and algorithms!

To zoom in on something that can be made computational: we recall our success in the first part computing modular forms by working with modular symbols with coefficients in a representation. Via the Eichler–Shimura isomorphism, we also reinterpreted this as arising in cohomology. On the one hand it is a bit sad that we are losing our direct connection to sections over $Y(\widehat{K})$; but on the other hand, we should not feel like we need to apologize for cohomology!

10.2 Quick review on cohomology

10.3 Cohomology

The Hecke algebra for $\widehat{K}$ (and $\mathbf{G}$) acts on spaces of automorphic forms and on cohomology; this permits an identification which connects these two worlds through linear algebra.

A **weight** is a finite-dimensional complex representation $\mathbf{G} \to \mathrm{GL}(W)$, usually taken to be irreducible. An irreducible weight is specified by highest

weights at each of the the archimedean places. Again associated to W is a bundle $\mathscr{W}$.

Definition 10.3.1. Let f be an automorphic form of level $\widehat{K}$ that is a Hecke eigenform. We say that f is **cohomological** of weight W and degree $r \in \mathbb{Z}_{\geq 0}$ if the Hecke eigensystem associated to f occurs in $H^r(Y(\widehat{K}), \mathscr{W})$.

In other words, there exists a class $[c] \in H^r(Y(\widehat{K}), \mathscr{W})$ such that $T_{\hat{\alpha}} f = T_{\hat{\alpha}}[c]$ for all Hecke operators $T_{\hat{\alpha}}$. It might be better to define Hecke irreducible subspaces, and implicitly this is respect to some choice of embedding $E^{\mathrm{al}} \hookrightarrow \mathbb{C}$ for this.

For purely psychological reasons, we will also use the abbreviation

$$M_{W,r}(\widehat{K}) := H^r(Y(\widehat{K}), \mathscr{W}) \tag{10.3.2}$$

for the space of cohomological automorphic forms of level $\widehat{K}$, weight W, and degree r (obtained from the span of eigenforms).

Importantly, not all automorphic forms are cohomological—among classical modular forms, they are exactly those of weight $k \geq 2$. We also have not even started talking about Maass forms (or any other more general class!). Nevertheless, to avoid tripping our tongues, in these notes we will call a cohomological automorphic form just a **modular form**. Sorry to the forms of weight 1 and all the others, we mean no disrespect! This abbreviation is only as serious as restricting attention to commutative rings or finite-dimensional algebras but not wanting to repeat the adjective each time!

So in this approach, we algorithmically first try to construct Hecke modules; then, if desired and where possible, we try to reconstruct the automorphic form using some kind of series expansion. This is like going from the eigenvalues a_n to a classical modular form, which even in that case is not always a trivial step (unless we have a standard congruence subgroup)!

We do not attach a single pair of weight and degree to an automorphic form, as the same system of Hecke eigenvalues may occur in more than one such. There is a well-defined set of degrees computed by Vogan–Zuckerman.

In the same way, we should not confuse the K_∞-type with the weight: the weight and degree determine the K_∞-type, and in general the matching between the two is again given in representation-theoretic terms.

From a double coset decomposition (9.3.8), we obtain a decomposition

$$H^r(Y(\widehat{K}), \mathscr{V}) \xrightarrow{\sim} \bigoplus_i H^r(\Gamma_i \backslash D_\infty, \mathscr{V}). \tag{10.3.3}$$

Since D_∞ is contractible, we have an isomorphism $H^r(\Gamma_i \backslash D_\infty, \mathscr{V}) \simeq H^r(\Gamma_i, \mathscr{V})$. This is true even if Γ_i has torsion, since have taken our weight representation to have complex coefficients.

In this way, we have in principle turned our problem of computing modular forms into something which we could imagine inputting into a computer: the groups Γ_i are finitely presented, the weight is an algebraic representation, and we are asking for group cohomology—yas!

Remark 10.3.4. The discrepancy between these two over say finite fields encodes interesting torsion and congruences phenomena which is unfortunately neglected here.

Only finitely many values of q will produce nonzero spaces of modular forms, by a theorem of Borel– Serre: we have $H^r(Y, \mathscr{V}) = 0$ whenever q is at least the **virtual cohomological dimension** of Γ, equal to the dimension of D minus the rank of a maximal F-split torus in $\mathbf{G}$.

10.4 $(\mathfrak{g}, K)$-cohomology

In this section, we sketch the connection between automorphic forms and cohomology using tools from representation theory. The algorithmically inclined reader is encouraged to skip this section!

One way to think about $(\mathfrak{g}, K_\infty)$-cohomology is that it is the representation-theoretic model for the de Rham cohomology of the locally symmetric space (with coefficients), giving rise to harmonic differential forms. So from our point of view, it provides the right generalization of classical modular forms of weight $k \geq 2$.

10.5 Cuspidal and Eisenstein cohomology

We have a decomposition of $H^*(Y, V)$ into a direct sum of cuspidal cohomology and Eisenstein cohomology. Cuspidal cohomology corresponds to cuspidal automorphic forms; the remaining Eisenstein cohomology is built up from automorphic forms on subquotients of $\mathbf{G}$.

Remark 10.5.1. We could remember this cohomological degree by adding separate data: instead of defining only functions (0-forms), we replace the target W by $\mathrm{Hom}(\wedge^q \mathfrak{p}, W)$ where $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and work with the natural K_∞-equivariance. Then "cohomological in degree q" becomes the condition that the resulting W-valued q-form is closed (and harmonic).

10.6 Hecke operators

The cohomology groups $H^r(Y, V) = \bigoplus_i H^r(\Gamma_i, V)$ are equipped with an action of Hecke operators for each double coset $\widehat{K}\widehat{\pi}\widehat{K}$ with $\widehat{\pi} \in \widehat{G}$ by correspondences.

Explicitly, given a characteristic function $T(\widehat{\pi})$ as in section 9.4, we decompose the double coset $\widehat{K}\widehat{\pi}\widehat{K}$ into a disjoint union of right cosets

$$\widehat{K}\widehat{\pi}\widehat{K} = \bigsqcup_j \widehat{\pi}_j \widehat{K}. \tag{10.6.1}$$

We can then define the action of $T(\widehat{\pi})$ on $f \in M_W(\widehat{K})$ by

$$(T(\widehat{\pi})f)(\widehat{g}) = \sum_j f(\widehat{g}\widehat{\pi}_j). \tag{10.6.2}$$

This action is well-defined (independent of the choice of representative $\widehat{\pi}$ and representatives $\widehat{\pi}_j$) by the right $\widehat{K}$-invariance of f. The Hecke operators generate an associative $\mathbb{Z}$-algebra in $\mathrm{End}(H^r(Y, \mathscr{W}))$.

10.7 Degeneracy maps

Chapter 11

Computing modular forms

11.1 Computational problem

The *input* is:

- A number field F,
- A reductive group $\mathbf{G}$ over F,
- A level $\widehat{K} \leq \widehat{G}$,
- A weight W (with $G \circlearrowright W$),
- A degree $r \geq 0$.

As *output*, we mean to compute with the space $M_{W,r}(\widehat{K}) = H^r(Y_{\mathbf{G}}(\widehat{K}), \mathscr{W})$ of modular (cohomological automorphic) forms of weight W, and degree r, and level $\widehat{K}$. As for classical modular forms (section 2.9), our suite of tasks:

1. compute dimension data, organized by spaces of lifts, and the decomposition into Hecke irreducible subspaces;
2**. compute the Hecke action on $M_{W,r}(\widehat{K})$;
3. extract Hecke eigensystems using linear algebra;
4. perform a decomposition into oldforms and newforms; and
5. write out series expansions.

Again the computations do not stop here! But we try to stay within scope.

For the purposes of these notes, the output of the algorithm is a *Hecke module*, meaning a finite-dimensional vector space V over a finitely generated field of characteristic 0 equipped with a procedure that, given a double coset $\widehat{K}\widehat{\pi}\widehat{K}$ with $\widehat{\pi} \in \widehat{G}$, computes the associated Hecke operator as an endomorphism of V, i.e., a matrix with entries in E (the base field of W) in the specified basis for V.

Described in this way, this sure does seem like a difficult problem in general! Nevertheless, we propose the following conjecture.

Conjecture 11.1.1. *There exists an algorithm that, given a reductive group* $\mathbf{G}$ *over a number field* F*, a level* $\widehat{K}$*, a weight* W*, and* $r \geq 0$*, computes the space* $M_{W,r}(\widehat{K})$ *as a Hecke module.*

By inductive calls, we may isolate the cuspidal cohomology from the Eisenstein cohomology (arising from groups of lower rank).

11.2 Algorithm overview

There are roughly four classes of algorithms for computing modular forms.

1. Definite methods (algebraic modular forms)
2. Modular symbols and the Voronoï complex
3. Trace formulas
4. Explicit lifts and multiplication

Over the next two lectures, we will focus on the first two.

11.3 Discussion

Part III

Algebraic modular forms

Chapter 12

Theta series and Brandt matrices

12.1 Theta series and lattice methods

Theta series give modular forms attached to lattices (and more generally to automorphic representations). In computational practice they serve two roles: they provide explicit modular forms to test algorithms against, and they often span interesting subspaces (Eisenstein and sometimes cusp forms). In higher rank, theta lifting is also a way to produce forms with known Hecke eigenvalues.

12.2 Brandt matrices

See [Voi21, Chapter 41].

For $F = \mathbb{Q}$ and $\operatorname{disc} B = p$, the Brandt matrix is the adjacency matrix of the directed ℓ-isogeny graph among isomorphism classes of supersingular elliptic curves over $\mathbb{F}_p^{\mathrm{al}}$ (specified by j-invariant). This connects with the other lectures!

Chapter 13

Definite setting

There is a special and very interesting class of reductive groups $\mathbf{G}$ for which we can compute modular forms, in the sense of these notes: where the real points $G_\infty = \mathbf{G}(\mathbb{R})$ modulo its center is compact. In this case, the associated global symmetric space $D_{\mathbf{G},\infty}$ has dimension zero—it is a finite set of points! (How great is that?!)

Gross [Gro99] investigated the associated modular forms, which are just invariant functions on a finite set—with the geometry essentially gone, he dubbed them *algebraic modular forms*. In this case, the Hecke operators can be computed using purely algebraic means, making it particularly well-suited for computations.

Is this chapter, we set up the basic infrastructure for these forms, followed by several chapters explaining special interesting cases.

13.1 The totally definite setting

References for this section are the fundamental article of Gross [Gro99] and an expository account by Loeffler [Loe08].

Let $\mathbf{G}$ be a reductive group over a number field F. Recall that we associate a symmetric domain $D_\infty = G_\infty/(A_\infty K_\infty)$ obtained from the quotient of the real points $G_\infty = \mathbf{G}(F_\infty)$ by the product of K_∞ (a subgroup containing the maximal connected compact subgroup of G_∞) together with $A_\infty = \mathbf{S}_{\mathbf{G}}(\mathbb{R})^0$ (the connected component of the real points of a maximal F-split torus in $\mathbf{G}$). So to make D_∞ of small dimension, we want the compact K_∞ and the real torus A_∞ to be of large dimension.

The smallest it could be is dimension 0, of course, and that situation is described by the following equivalence. As before, let $G = \mathbf{G}(F)$ and $\widehat{G} = \mathbf{G}(\widehat{F})$.

Proposition 13.1.1. *The following statements are equivalent.*

(i) $D_{\mathbf{G},\infty}$ *has (real) dimension* 0 *(for all* K_∞*).*
(ii) $\mathbf{S}_{\mathbf{G}}(\mathbb{R})$ *is a maximal split torus in* G_∞*.*
(iii) *The derived group* $\mathbf{G}^{\mathrm{der}}$ *has* $\mathbf{G}^{\mathrm{der}}(F_\infty)$ *compact.*

(iv) *Every arithmetic subgroup $\Gamma \leq G$ is finite.*
(v) *The trivial group $\Gamma = \{1\}$ is an arithmetic subgroup of G.*
(vi) *G is a discrete subgroup of $\widehat{G}$.*
(vii) *G is a discrete subgroup of $\widehat{G}$ and the quotient $G\backslash\widehat{G}$ is compact.*

Moreover, if $\mathbf{G}$ *is almost simple, then these are further equivalent to*

(viii) *G_∞ is compact.*

If a reductive group $\mathbf{G}$ enjoys the equivalent properties in Proposition 13.1.1, we say that $\mathbf{G}$ is **totally definite**; others might say *compact at ∞ mod center*.

In light of the classification of reductive groups $\mathbf{G}$, we can triumphantly identify those groups which enjoy the properties in Proposition 13.1.1.

Corollary 13.1.2. *If* $\mathbf{G}$ *is almost simple and totally definite, then F is totally real and for all real places v, the Lie algebra of $G(F_v)$ agrees with the (simple) Lie algebra of one of the following groups ($n \geq 1$): $\mathrm{SU}(n+1)$, $\mathrm{SO}(2n+1)$, $\mathrm{Sp}(n)$, $\mathrm{SO}(2n)$, E_6, E_7, E_8, F_4, G_2.*

Proof. See e.g. Helgason [Hel01, §X.6]. □

13.2 Algebraic modular forms

As in the previous section, let $\mathbf{G}$ be a totally definite reductive group over F, totally real.

When we specialize the definition of automorphic form (Definition 10.1.2) to this case, many nice things happen. First, there is no analytic condition on the finite set D_∞, and when G_∞ is connected and K_∞ is maximal, $D_\infty = \{1\}$ so there really is no symmetric domain to think about, just class sets. In particular, there are no technical conditions on automorphic forms. Moreover, all forms are cohomological (arising in H^0). Let $\widehat{K} \leq \widehat{G}$ be a level and $\rho\colon G \to W$ be a weight.

Definition 13.2.1. An **algebraic modular form** of weight W and level $\underline{K} = \widehat{K} \times K_\infty$ is a modular form of degree 0, i.e., a function

$$f\colon D_\infty \times \widehat{G}/\widehat{K} \to W$$

such that f is locally constant and G-invariant, i.e.,

$$f(g\widehat{x}\widehat{K}) = gf(\widehat{x}\widehat{K}) \tag{13.2.2}$$

for all $g \in G$ and $\widehat{x}\widehat{K} \in \widehat{G}/\widehat{K}$.

As before, we write $M_{\mathbf{G},W}(\underline{K})$ for the space of modular forms of weight $\underline{K}$ and level $\underline{K}$, writing just $M_W(\widehat{K})$ when $\mathbf{G}$ and K_∞ are clear from context.

So an algebraic modular form is just a modular form on a totally definite group. In particular, an algebraic modular form is determined by its values the

set $Y = G\backslash(D_\infty \times \widehat{G}/\widehat{K})$, which is finite since D_∞ is finite (Theorem 9.3.3), so $M_{\mathrm{G},W}(\underline{K})$ is finite-dimensional. When $W = E$ is the trivial representation over a field E, then $M_{\mathrm{G},W}(\underline{K})$ is simply the space of E-valued functions on Y.

More precisely, let $h = \#Y$. Writing

$$D_\infty \times \widehat{G} = \bigsqcup_{i=1}^{h} G\widehat{\beta}_i\widehat{K}, \tag{13.2.3}$$

with $\widehat{\beta}_i \in D_\infty \times \widehat{G}$, it follows from the definition that any $f \in M_W(\widehat{K})$ is determined by the elements $f(\widehat{\beta}_i)$ with $i = 1, \ldots, h$. Let

$$\Gamma_i = G \cap \widehat{\beta}_i\widehat{K}\widehat{\beta}_i^{-1}.$$

Each arithmetic group Γ_i is finite (Proposition 13.1.1). As a result, we can see these in cohomology as follows.

Lemma 13.2.4. *The map*

$$\begin{aligned} M_W(\widehat{K}) &\xrightarrow{\sim} \bigoplus_{i=1}^{h} H^0(\Gamma_i, W) \\ f &\mapsto (f(\widehat{\beta}_1), \ldots, f(\widehat{\beta}_h)) \end{aligned}$$

is an isomorphism of E-vector spaces, where

$$H^0(\Gamma_i, W) = \{v \in W : \gamma v = v \text{ for all } \gamma \in \Gamma_i\}.$$

Proof. Immediate from the definition of group cohomology. □

The space $M(W, \widehat{K})$ comes equipped with the action of Hecke operators, defined in section 10.6. A straightforward calculation shows that the map in Lemma 13.2.4 is equivariant for the Hecke operators.

13.3 Algorithmic details

Having discussed the theory in the previous sections, we now present a general formulation for algebraic groups.

Recall we are computing the space $M_W(\widehat{K})$ of algebraic modular forms of weight W and level $\widehat{K}$ on a group $\mathbf{G}$. To begin with, we must decide upon a way to represent in bits the group $\mathbf{G}$, the open compact subgroup $\widehat{K}$, and the G-representation W so we can work explicitly with these objects.

Then, to compute the space $M_W(\widehat{K})$ as a module for the Hecke operators, we carry out the following tasks:

1. Compute representatives $\widehat{x}_i\widehat{K}$ $(i = 1, \ldots, h)$ for $G\backslash\widehat{G}/\widehat{K}$, as in (13.2.3), compute $\Gamma_i = G \cap \widehat{x}_i\widehat{K}\widehat{x}_i^{-1}$, and initialize

$$H = \bigoplus_{i=1}^{h} H^0(\Gamma_i, W).$$

Choose a basis of (characteristic) functions f of H.

2. Determine a set of Hecke operators $T(\widehat{\pi})$ that generate $\mathcal{H}(\widehat{K})$, as in section 10.6. For each such $T(\widehat{\pi})$:
   a. Decompose the double coset $\widehat{K}\widehat{\pi}\widehat{K}$ into a union of right cosets $\widehat{\pi}_j\widehat{K}$, as in (10.6.1);
   b. For each $\widehat{x}_i$ and $\widehat{\pi}_j$, find $\gamma_{ij} \in G$ and j^* so that
   $$\widehat{x}_i\widehat{\pi}_j\widehat{K} = \gamma_{ij}\widehat{x}_{j^*}\widehat{K}.$$
   c. Return the matrix of $T(\widehat{\pi})$ acting on H via the formula
   $$(T(\widehat{\pi})f)(\widehat{x}_i) = \sum_j \gamma_{ij} f(\widehat{x}_{j^*})$$
   for each f in the basis of H.

13.4 Change of level

Following Greenberg–Voight [GV11] (see also Dembélé–Voight [DV13, §8]), there is a natural map which relates modular forms of higher level to those of lower level by modifying the coefficient module, as follows.

Suppose that $\widehat{K}' \leq \widehat{K}$ is a finite index subgroup. Decomposing as in (13.2.3), we obtain a bijection

$$\begin{aligned} G\backslash\widehat{G}/\widehat{K}' = \bigsqcup_{i=1}^{h} G\backslash\big(G\widehat{x}_i\widehat{K}\big)/\widehat{K}' &\xrightarrow{\sim} \bigsqcup_{i=1}^{h} \Gamma_i\backslash\widehat{K}_i/\widehat{K}_i' \\ G(\gamma\widehat{x}_i\widehat{u})\widehat{K}' &\mapsto \Gamma_i(\widehat{x}_i\widehat{u}\widehat{x}_i^{-1})\widehat{K}_i' \end{aligned}$$

for $\gamma \in G$ and $\widehat{u} \in \widehat{K}$. This yields

$$M(W, \widehat{K}') \xrightarrow{\sim} H^0(\Gamma_i, \mathrm{Hom}(\widehat{K}_i/\widehat{K}_i', W)) \cong \bigoplus_{i=1}^{h} H^0(\Gamma_i, \mathrm{Coind}_{\widehat{K}_i'}^{\widehat{K}_i} W).$$

Via the obvious bijection

$$\widehat{K}_i/\widehat{K}_i' \cong \widehat{K}/\widehat{K}', \tag{13.4.1}$$

letting $W = \mathrm{Coind}_{\widehat{K}'}^{\widehat{K}} W$ we can also write

$$M(W, \widehat{K}') \cong \bigoplus_{i=1}^{h} H^0(\Gamma_i, W_i) \tag{13.4.2}$$

where W_i is the representation W with action twisted by the identification (13.4.1). Moreover, writing $\widehat{K} = (K_{\mathfrak{p}})_{\mathfrak{p}}$ in terms of its local components, for any Hecke operator $T(\widehat{\pi})$ such that

$$\text{if } \widehat{\pi} \notin K'_{\mathfrak{p}} \text{ then } K_{\mathfrak{p}} = K'_{\mathfrak{p}}$$

(noting that $\widehat{\pi} \in K'_{\mathfrak{p}}$ for all but finitely many primes $\mathfrak{p}$), the same definition (10.6.2) applies and by our hypothesis we have a simultaneous double coset decomposition

$$\widehat{K}'\widehat{\pi}\widehat{K}' = \bigsqcup_j \widehat{\pi}_j \widehat{K}' \quad \text{and} \quad \widehat{K}\widehat{\pi}\widehat{K} = \bigsqcup_j \widehat{\pi}_j \widehat{K}.$$

Now, comparing (13.4.2) to the result of Lemma 13.2.4, we see in both cases that modular forms admit a uniform description as h-tuples of Γ_i-invariant maps. For this reason, a special role in our treatment will be played by *maximal* open compact subgroups.

13.5 Orthogonal groups and lattices

Having set up the general theory in the previous section, we now specialize to the case of the orthogonal group; next up after that will be the unitary group, and a few comments on the symplectic group.

The orthogonal groups form come in at least four flavors:

$$\mathrm{O}_n, \mathrm{SO}_n, \mathrm{GO}_n, \mathrm{GSO}_n\,.$$

13.6 Unitary groups

13.7 Exceptional isomorphisms

The Clifford functor provides exceptional isomorphisms between orthogonal modular forms and other (more familiar) spaces of modular forms.

For $n = 2$, we have $\mathfrak{so}_2 \simeq \mathbb{R}$ and get Hecke Grossencharacters [VW25].

For $n = 3$, we have $\mathfrak{so}_3 \simeq \mathfrak{sl}_2$ ($B_1 = A_1$), so we obtain in this way *classical* modular forms (Birch [Bir91], Hein [Hei16], Hein–Tornaría–V [HTV25]).

For $n = 4$, we have $\mathfrak{so}_4 \simeq \mathfrak{sl}_2 \times \mathfrak{sl}_2$ ($D_2 = A_1 \times A_1$), and we obtain Hilbert modular forms.

For $n = 5$, we have $\mathfrak{so}_5 \simeq \mathfrak{sp}_4$ ($B_2 = C_2$), so we obtain Siegel (para)modular forms. In this case, we can get dimensions and Hecke eigenvalues, but not (yet) Fourier expansions.

Part IV

Geometry and cohomology

Chapter 14

Fundamental domains

When G_∞ is not totally definite, our proposed method to compute automorphic forms on arithmetic groups relies upon an algorithm for computing fundamental domains. We now introduce how this method works for Fuchsian groups and propose some generalizations.

14.1 Computing fundamental domains

Let $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ be a Fuchsian group such that $X(\Gamma) = \Gamma\backslash\mathfrak{H}$ has finite hyperbolic area; we say Γ has *cofinite area*. Suppose that $p \in \mathfrak{H}$ has trivial stabilizer $\Gamma_p = \{1\}$. Then the set

$$D(p) = \{z \in \mathfrak{H} : d(z,p) \leq d(gz,p) \text{ for all } g \in \Gamma\},$$

known as a *Dirichlet domain*, is a hyperbolically convex fundamental domain for Γ. Our main theorem is as follows [Voi09].

Theorem 14.1.1. *There exists an algorithm that, given a Fuchsian group Γ of cofinite area and a point $p \in \mathfrak{H}$ with $\Gamma_p = \{1\}$, returns the Dirichlet domain $D(p)$ and a finite presentation of Γ together with a solution to the word problem for Γ.*

The solution to the word problem (for the computed presentation of Γ) means the following: there is an algorithm that, given an element $\gamma \in \mathrm{PSL}_2(\mathbb{R})$, determines if $\gamma \in \Gamma$ and, if so, writes γ as a word in the given generators. This algorithm, as mentioned above, has been implemented in Magma and has been fine-tuned for certain parameters to make it practical.

We now consider the setting of an arbitrary arithmetic group. The space $D = G_\infty/A_G K_\infty$ has a metric d and the notion of Dirichlet domain $D(p)$ extends immediately: it is again a closed and convex domain that admits a finite description [Voi21, §37].

Conjecture 14.1.2. *There exists an algorithm that, given a reductive group* $\mathbf{G}$*, an arithmetic subgroup* $\Gamma \subset G$*, and a point* $z_0 \in D_\infty$ *returns the Dirichlet domain* $D(z_0)$ *and a finite presentation for* Γ *with a solution to the word problem for* Γ*.*

The algorithm simply enumerates elements of Γ until the associated intersection of hyperplane bisectors is described by a face pairing with correct volume. To make this practical, we consider enumerate with respect to suitable majorants (positive definite quadratic forms) that measure the contribution to the fundamental domain. We see the beginning of a kind of "noncommutative Arakelov theory".

14.2 Computing in cohomology

We can directly then compute in $H^1(\Gamma, V) = Z^1(\Gamma, V)/B^1(\Gamma, V)$ as the quotient of 1-cochains by 1-coboundaries. A 1-cochain is represented by its values on the generators of Γ and similarly a spanning set for $B^1(\Gamma, V)$ is constructed from generators of Γ. The Hecke operators act on $H^1(\Gamma, V)$ via double cosets and the reduction theory allows us to write an arbitrary 1-cochain as a linear combination of the given basis of characteristic q-cochains.

14.3 Shimura curves

Exercises

14.1.

Chapter 15

Generalized modular symbols

In this chapter, we make our first step towards a general definition of modular symbol.

Recall how modular symbols were built: they arose from homology classes supported on the boundary of the modular curve. Mazur in a unpublished typewritten letter [Maz] put forward a generalization of modular symbols, explaining how the Borel–Serre compactification provides the same infrastructure in relative (co)homology, with cycles coming from parabolics: Hecke correspondences acting on those cycles generalizes modular symbols. This was followed by work of Ash–Borel [AB90] and Ash–Rudolph [AR79] pursuing this with higher-dimensional modular symbols and continued fractions, leading to computations of Ash–Grayson–Green for $\mathrm{SL}_3(\mathbb{Z})$ [AGG84].

In a nutshell, we compute a basis, apply Hecke operators, and reduce to a preferred set of "small" (unimodular) generators.

What can we hope to get out of this direct approach? We recall that the Borel–Serre compactification is contractible but has a boundary encoding all of the cusp directions; combinatorailly, this boundary is controlled by rational parabolics, and homotopically it is the Tits building. For $\mathsf{G} = \mathrm{SL}_n$, the Solomon–Tits theorem says that the building has the homotopy type of a wedge of $(n-2)$-spheres, so its reduced homology lives in a single top degree and its homology is the Steinberg module. Duality theorems then imply that modular symbols (coinvariants of the Steinberg module) compute in *top* degree cohomology. For SL_n with $n \geq 4$, the computationally interesting cuspidal cohomology occurs in lower degrees. To go deeper in cohomology, we build combinatorial models: see the next chapter.

15.1 Cycles from parabolics

We follow Mazur [Maz], Ash–Borel [AB90], and Ash–Rudolf [AR79].

We recall our notation from the previous chapters. Let G be a reductive algebraic group over a number field F.

$$D_\infty = G_\infty / A_\infty K_\infty \tag{15.1.1}$$

be a global symmetric domain. Let $\widehat{K} \leq \widehat{G} = \mathbf{G}(\widehat{F})$ be compact open subgroup, let $G = \mathbf{G}(F)$, and

$$Y(\widehat{K}) = G\backslash(\widehat{G} \times D_\infty)/\widehat{K} = \bigsqcup_i \Gamma_i \backslash D_\infty \tag{15.1.2}$$

where the groups Γ_i arise from a choice of representatives of $G\backslash\widehat{G}/\widehat{K}$, so well-defined up to conjugacy in G.

Let D_∞^* be the Borel–Serre compactification of D_∞ and define similarly

$$X(\widehat{K}) = G\backslash(\widehat{G}/\widehat{K} \times D_\infty^*) \tag{15.1.3}$$

The boundary $\operatorname{bd} X(\widehat{K})$ is a union of strata indexed by F-rational parabolic subgroups of $\mathbf{G}$. For technical reasons we suppose that $X(\widehat{K})$ is oriented, for example each Γ_i is neat (it is enough that each Γ_i contains no orientation-reversing elements).

In the classical modular curve case, the cusps (i.e., the boundary above) are stable under Hecke correspondences, but boundary (co)homology is governed by Eisenstein (lower-rank) phenomena. To isolate cuspidal information, we must pass to *relative* (co)homology, i.e., classes that restrict trivially to the boundary. This will give us a higher-rank analogue of taking modular symbols modulo boundary, which will amount to working with compactly supported (also sometimes called *parabolic*) cohomology.

To that end, we define nontrivial homology classes. In fact, we can define these more generally for subgroups. Let $\varphi\colon \mathbf{G}' \hookrightarrow \mathbf{G}$ be an injective morphism of reductive groups over F. We choose data for $\mathbf{G}'$ compatible with $\mathbf{G}$ as follows: we define

$$\begin{aligned} K'_\infty &= \phi^{-1}(K_\infty) \leq G'_\infty \\ \widehat{K}' &= \phi^{-1}(\widehat{K}) \leq \widehat{G}'. \end{aligned} \tag{15.1.4}$$

We suppose that K_∞ contains the maximal connected compact subgroup of G'_∞ and that $\widehat{K}'$ is compact and open in $\widehat{G}'$.

Repeating the construction,

$$D'_\infty := G'_\infty / A'_\infty K'_\infty \tag{15.1.5}$$

admits a natural totally geodesic inclusion $D'_\infty \hookrightarrow D_\infty$ and this extends to the Borel–Serre compactifications

$$(D'_\infty)^* \to D_\infty^*. \tag{15.1.6}$$

Consequently for

$$X(\widehat{K}') := G'\backslash(\widehat{G}'/\widehat{K}' \times (D'_\infty)^*) \tag{15.1.7}$$

we obtain an induced map of compactifications

$$X(\widehat{K}') \to X(\widehat{K}) \tag{15.1.8}$$

Ash–Borel give a nice geometric construction of this in important cases via fixed-point loci of involutions.

Let $m' := \dim Y(\widehat{K}')$.

Definition 15.1.9. The **(generalized) modular symbol** attached to φ is the relative homology class

$$[\varphi] := (\phi_*)([X(\widehat{K}')]) \in H_{m'}(X(\widehat{K}), \operatorname{bd} X(\widehat{K}); \mathbb{Q})$$

where the inside $[X(\widehat{K}')]$ denotes the fundamental class.

Keeping track of stabilizers, we could take the class in integral homology.

Let $m := \dim X(\widehat{K})$. By Poincaré–Lefschetz duality, the class $[\varphi]$ corresponds to a compactly-supported cohomology class

$$[\varphi]^\vee \in H^{m-r}H_r(X(\widehat{K}), \operatorname{bd} X(\widehat{K}); \mathbb{Q}) =: H_c^{m-m'}(Y(\widehat{K}), \mathbb{Q}).$$

Finishing our parallel with modular symbols, we will primarily be interested in the **parabolic cohomology**

$$H^*_{\mathrm{par}}(Y(\widehat{K}), \mathbb{Q}) = \ker\big(H^*(Y(\widehat{K}), \mathbb{Q}) \to H^*(\operatorname{bd} X(\widehat{K}), \mathbb{Q})\big)$$

(equivalently, the image of the compactly supported cohomology).

Example 15.1.10. Take $F = \mathbb{Q}$ and $\mathbf{G} = \mathrm{GL}_2$ with $K_\infty = \mathrm{O}_2(\mathbb{R})$ (or $\mathbf{G} = \mathrm{SL}_2$). Then $D_{\mathrm{G},\infty} \simeq \mathbf{H}^2$ is the upper half-plane. Let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, so $Y(\Gamma) = \Gamma\backslash\mathbf{H}^2$.

The Borel–Serre compactification $X(\Gamma)$ is a compact surface with boundary; its boundary $\operatorname{bd} X(\Gamma)$ is a disjoint union of circles, one for each cusp of Γ. (Sorry about that!)

Let $\mathbf{G}' \subset \mathbf{G}$ be the standard split torus (diagonal subgroup). Then $D'_\infty \simeq \mathbb{R}$, and the inclusion $D'_\infty \hookrightarrow D_\infty$ is a geodesic. More precisely, under the identification of the upper half-plane via moving i, this is the oriented vertical line through i; its closure in $D^*_{\mathrm{G},\infty}$ is a compact arc meeting the boundary in two points.

By $\mathrm{GL}_2(\mathbb{Q})$-conjugacy, one then obtains geodesics with endpoints at arbitrary cusps in $\mathbb{P}^1(\mathbb{Q})$.

Indeed, for $\alpha, \beta \in \mathbb{P}^1(\mathbb{Q})$ let $\{\alpha, \beta\} \in H_1(X(\Gamma), \operatorname{bd} X(\Gamma); \mathbb{Z})$ denote the relative homology class of the (oriented) geodesic from α to β. The boundary map in the long exact sequence of the pair yields

$$\partial\{\alpha, \beta\} = [\beta] - [\alpha] \ \in \ H_0\big(\operatorname{bd} X(\Gamma); \mathbb{Z}\big),$$

so the classes with trivial boundary (i.e. in $\ker(\partial)$) are precisely the cuspidal modular symbols. Via Poincaré–Lefschetz duality, these are dual to $H^1_c(Y(\Gamma))$, and after passing to the image in $H^1(Y(\Gamma))$ one recovers parabolic (cuspidal) cohomology.

Example 15.1.11. For $\mathbf{G} = \mathrm{GL}_n$ (or $\mathbf{G} = \mathrm{SL}_n$) over $F = \mathbb{Q}$, we obtain the *top-degree* modular symbols following Ash–Rudolph.

Example 15.1.12. If $\mathbf{G}'$ is (the derived group of) a Levi factor of a parabolic subgroup of $\mathbf{G}$, the resulting classes recover the Ash–Borel generalized modular symbols.

In a limited set of circumstances, modular symbols generate homology, and the image of cuspidal modular symbols by duality generate parabolic cohomology. Still, we wonder if these classes give enough of a handle so that they could be multiplied and their Hecke translates might span spaces in higher weight.

Remark 15.1.13. If we only took action on the boundary of the modular curve (as opposed to paths in the modular curve), that would only give Eisenstein series.

15.2 Hecke correspondences

Hecke operators, defined via correspondences, act on modular symbols: we translate the geometric cycle, decompose into finitely many components modulo Γ, and rewrite in the chosen basis of symbols. In principle, this implies that cohomology reduces to Hecke on a finite symbol module plus reduction.

15.3 Shapiro's lemma

Although it looks like we have a lot to do for each level structure $\widehat{K}$, we recall from modular symbols that in fact we can just change the coefficient module and always work with any supergroup.

More precisely, for $\Gamma' \leq \Gamma$ of finite index and a Γ-module M, Shapiro's lemma gives a canonical isomorphism

$$H^*(\Gamma', M) \cong H^*(\Gamma, \mathrm{Coind}_{\Gamma'}^{\Gamma} M) \tag{15.3.1}$$

at least away from H^0, but there is another way to interpret that as well. In particular, the fact that it is canonical means that the Hecke module structure transports.

This has the interesting consequence that the reduction in Hecke operators only needs to happen once and then for every weight and level, we just act appropriately on the representation.

Explicitly connect coset representatives and how the boundary maps are assembled.

15.4 Computations in practice

We now discuss feasibility and known explicit computations in small rank (and level). Check implementation, torsion, coefficient systems, higher rank, etc.

Chapter 16

Complexes

Chapter 17

Extensions

Part V

Projects (with Eran Assaf)

Chapter 18

Projects

In this chapter joint with Eran Assaf, we list some potential projects. We plan on selecting a subset of these based on the students who are selected—matching interest, preparation, etc.

These notes are of course an important reference for the projects, but we also list other relevant references for further background, context, and details for the specific project directions.

18.1 AWS proposed project 1: Picard from Hecke

Tagline

Compute q-expansions for Picard modular forms (associated to $\mathrm{U}(2,1)$) using Hecke data from $\mathrm{U}(3)$.

Estimated difficulty

Medium? Mostly learning about q-expansions for unitary groups and some hopefully straightforward Magma coding.

Suggested background

- For general background on Picard modular surfaces and Ball quotients, we recommend [Hol98] and [Hol95].
- For specific examples, in particular the Gaussian integers $\mathbb{Z}[i]$, and the Eisenstein integers $\mathbb{Z}[\zeta_3]$, see [Run96], [CvdG13], and [CvdG23].
- For a more general approach, those with more background in algebraic geometry should see [dSG19]. For a more direct approach, see [Shi78].
- For algorithmic and implementation aspects, as well as examples of computations of algebraic modular forms for $\mathrm{U}(3)$, see [Loe08] and [GV14].
- For background on the representation theory of $\mathrm{U}(3)$, see [FH91, §13].

Project overview

In this project, you will compute q-expansions of Picard modular forms, and obtain equations for Picard modular surfaces, which are moduli spaces parameterizing abelian 3-folds with CM.

Let $F = \mathbb{Q}(\sqrt{d}) \supseteq \mathbb{Q}$ be an imaginary quadratic field of discriminant $d < 0$, let $R = \mathbb{Z}_F$. Let V be a 3-dimensional F-vector space equipped with a Hermitian form $\Phi\colon V \times V \to F$ of signature $(2,1)$. Let $\Lambda \subseteq V$ be an R-lattice (a finitely generated, projective R-module such that $F\Lambda = V$), such that $\Phi(\Lambda, \Lambda) \subseteq R$.

For example, we can take $\Phi(x,y) = x_1\overline{y_1} + x_2\overline{y_2} - x_3\overline{y_3}$ and $\Lambda = \mathbb{Z}_F^3$ the standard lattice.

For simplicity, impose the additional restriction that R has class number 1. (This skips adelic stuff for now to get to the point, but we should put that back in at some point.)

Let $\mathrm{U}_{F|\mathbb{Q},3}(V) = \mathrm{U}(V)$ be the unitary algebraic group associated to V, so in particular

$$\mathrm{U}(V)(F) = \{g \in \mathrm{GL}(V)(F) : \Phi(gx, gy) = \Phi(x,y) \text{ for all } x, y \in V\}.$$

Let $\Gamma_\Lambda := \{g \in \mathrm{U}(V)(F) : g(\Lambda) = \Lambda\}$ be the stabilizer of Λ.

Let $k \in \mathbb{Z}^3$ be such that $k_1 \geq k_2 \geq k_3$ be a weight, and let $\rho_k\colon U(V) \to \mathrm{GL}(W_k)$ be the irreducible representation of highest weight k of $\mathrm{U}(V)$. (As a special case, we can take $k \in \mathbb{Z}$ as a short-hand for (k,k,k), in which case $\rho_k = \det^k$.

In this project, you will compute a basis of q-expansions for the space $M_k(\Lambda)$ of Picard modular forms of weight k and level Γ_Λ.

Steps

1. Define the space $M_k(\Lambda)$ of modular forms of weight ρ_k and level Γ_Λ, as well as its cuspidal subspace $S_k(\Lambda)$. Compare with [CvdG13] for the case $\rho_{(j+k,k,k)}$, and with [dSG19] for a more general setup.
2. Let $f \in M_k(\Lambda)$. Show that f has a Fourier–Jacobi expansion $f(u,v) = \sum_{n=0}^{\infty} f_n(u) e^{2\pi i n \delta^{-1} v}$, where $\delta = \sqrt{d}$ and $f_n(u)$ satisfies a quasi-periodicity condition. Compare with [Fin98] and [CvdG23] when $R = \mathbb{Z}[\zeta_3]$.
3. For a prime $p = \mathfrak{p}\overline{\mathfrak{p}}$ which is split in K, let $T_{\mathfrak{p},1}$ be the Hecke operator associated to $\mathrm{diag}(1,1,\pi/\overline{\pi})$ where $\mathfrak{p} = \pi R$ and similarly $T_{\mathfrak{p},2}$ be associated to $\mathrm{diag}(1,\pi/\overline{\pi},\pi/\overline{\pi})$. For an inert prime p, write $T_{p,1}$ for the Hecke operator for $\mathrm{diag}(1,1,p)$. Given a Fourier–Jacobi expansion $f(u,v) = \sum_{n=0}^{\infty} f_n(u) e^{2\pi i n \delta^{-1} v}$, find formulas for the Fourier–Jacobi expansion of $T_{\mathfrak{p},i} f$ for $i = 1, 2$.
4. Assume that $f \in M_k(\Lambda)$ is a Hecke eigenform with eigenvalues $T_{\mathfrak{p},i} f = \lambda_{\mathfrak{p},i} f$ for all primes $\mathfrak{p}$ and $i = 1, 2$. Find an expansion for $f_n(u)$ in terms of the $\lambda_{\mathfrak{p},i}$. See [Fin98] for the scalar valued case and [CvdG23] for $R = \mathbb{Z}[\zeta_3]$.

5. Using Magma, compute Hecke eigenvalues for algebraic modular forms for a compact form of G for various weights k and lattices Λ. Several examples are given in [Loe08] and [GV14]. Use your results to write Fourier–Jacobi expansions for these forms. Implement a function that given k, Λ, and bounds on precision, returns the Fourier–Jacobi expansion up to that bound of every eigenform in $S_k(\Lambda)$.

The real payoff

Graded rings give equations for Picard modular surfaces, which would be interesting to study geometrically and have applications as moduli spaces of abelian threefolds.

Additional possible directions

1. Use the construction of Eisenstein series in [Shi78] to construct a basis for the Eisenstein subspace $E_k(\Lambda) \subseteq M_k(\Lambda)$. Implement a function that given k, Λ and bounds on precision, returns Fourier–Jacobi expansions of a basis for $M_k(\Lambda)$ up to the given bounds.
2. Implement basic arithmetic of Fourier–Jacobi expansions, so that given the Fourier–Jacobi expansions of $f \in M_k(\Lambda)$ and $g \in M_l(\Lambda)$, one can compute $fg \in M_{k+l}(\Lambda)$, and if $k = l$, one can compute $af + bg \in M_k(\Lambda)$ for any $a, b \in \mathbb{Q}$.
3. The graded ring
$$M_*(\Lambda) = \bigoplus_{k \in \mathbb{Z}} M_k(\Lambda),$$
gives rise to the Bailey–Borel compactification of the moduli space, the Picard modular surface $X(\Lambda) = \operatorname{Proj} M_*(\Lambda)$. Implement a function that, given Λ, computes the graded ring $M_*(\Lambda)$, hence a presentation of the Picard modular surface. Compare with the results in [Run96]. Connect explicitly with the moduli theory and the universal abelian variety in cases where it is possible.
4. Compute minimal models of the associated Picard modular surfaces and compute certain geometric properties of these surfaces that cannot be deduced simply from their numerical invariants.
5. Repeat for $\mathrm{U}(3,1)$?

18.2 AWS proposed project 2: Siegel stability under specialization

Tagline

Compute Siegel paramodular forms of weight 2 (associated to abelian surfaces) using Hecke stability under specialization to a modular curve.

Estimated difficulty

Medium-hard? Some coding with some formulas that are a bit complicated.

Suggested background

- Brumer–Pacetti–Poor–Tornaría–Voight–Yuen [BPPTVY19] for background, context and motivation, and for the restriction map.
- Poor–Yuen [PY15] and Poor–Shurman–Yuen [PSY20, PSY24] for computing paramodular cusp forms in weight 2, and their data

  http://siegelmodularforms.org/pages/degree2/paramodular-wt2-all/index.html

- Schaeffer [Sch15] for Hecke stability for classical modular forms.

Project overview

Using algebraic modular forms, either using a definite form of GSp_4 or SO_5 (and using the exceptional isomorphism $\mathsf{B}_2 = \mathsf{C}_2$ giving $\mathrm{PGSp}_4(\mathbb{R}) \simeq \mathrm{SO}_5(\mathbb{R})$, split form), we can compute Hecke eigensystems for (cohomological) Siegel paramodular forms. But only those of weight (k, j) with $k \geq 3$.

At the same time, the paramodular conjecture predicts that typical abelian surfaces A over $\mathbb{Q}$ are modular, matching up with Siegel paramodular forms of weight $(2, 0)$, just below what we can achieve cohomologically. Bummer! This is just like the case of classical modular forms of weight 1, they are apparently too cool to be computed like the others.

Schaeffer [Sch15] gives a method for computing weight 1 classical modular forms using *Hecke stability*. In a nutshell, given a nonzero form $f \in M_1(\Gamma)$, by division we have $M_1(\Gamma) \subseteq f^{-1}M_2(\Gamma)$. But we also know that $M_1(\Gamma)$ is stable under the Hecke operators, and Schaeffer shows that the maximal Hecke stable subspace of $f^{-1}M_2(\Gamma)$ coincides with $M_1(\Gamma)$. Algorithmically, this means we need to compute $M_2(\Gamma)$, find a nonzero f and divide q-expansions, then hit the result with Hecke operators and intersect until it remains stable.

In this project, we try to make this work for Siegel modular forms, computing in weight $2 = 4 - 2$. Poor–Yuen do this [PY15] by first spanning weight 4 forms (in a nutshell, by taking the Hecke span of products of lifts), then dividing by forms in weight 2 (writing $f = h_1/g_1 = h_2/g_2$). We will explore how computing the Hecke stable subspace compares to their approach.

Already it will be interesting to implement multiplication of Siegel modular forms in Magma with an eye to efficiency—this is some fixed (universal) bilinear form. There may be some lessons to learn from multiplying q-expansions of Hilbert modular forms [A+26].

However, series expansions of Siegel modular forms are quite costly to work with, so we will then try to do this by *specializing to a modular curve*. Indeed, there is a homomorphism of graded rings [BPPTVY19, Lemma 6.3.3] and a relation involving the Hecke operators [BPPTVY19, Proposition 6.3.8] (see also

[PSY20, §5]). By specialization, we can compute the Hecke eigenvalue $a_{p,1}$ in time $O(p^2)$ (instead of $O(p^3)$, using the definition).

Can all this be made practical? For example, we should do these computations first all modulo a random prime to avoid coefficient blow up, then after repeat in characteristic 0 once we know the result we are after.

Steps

1. Learn some Magma and develop a basic type for working with Fourier expansions of modular forms on $\mathrm{G} = \mathrm{GSp}_4$ over $F = \mathbb{Q}$. Here is a stub (sorry it isn't great, but it's a place to start!): https://github.com/jvoight/jacobi_siegel.
2. Add basic Siegel modular forms, for example the usual Eisenstein series for $\mathrm{Sp}_4(\mathbb{Z})$. Connect with Igusa invariants [Voi21, §43.5] and verify other computations in level 1 (https://smf.compositio.nl/Fourier/Level1).
3. Implement multiplication and division (when holomorphic at ∞) of Siegel modular forms.
4. Specialize some basic Siegel modular forms to a modular curve, and check that multiplication commutes with specialization in examples.
5. Implement the dimension formula for Siegel paramodular cuspforms of weight $k \geq 4$ (https://math.ou.edu/~rschmidt/dimension_formulas/).
6. Find the first level N not covered by Poor–Yuen with a nonlift cuspform in weight $k = 2$ (looks like they focus on nonsquarefree level, for example, so it might be $N = 388$?), associated to an abelian surface. Find that form!
7. Estimate the running time of some of the above algorithms.

The real payoff

It is like we are at the "Antwerp tables" stage with classical modular forms: we have done a lot already in low level, but there are loads of questions about abelian surfaces that would gain insight from taking larger level.

In that direction: if the algorithms scale, we could predict the isogeny classes of all abelian surfaces with good reduction away from 2 by computing in paramodular levels N a bounded power of 2, working towards an answer to a question of Poonen.

We might also be able to approach the question: what is the (conjecturally) smallest conductor of an abelian fourfold of GSp_4-type? (The answer for classical modular forms is level $N = 13$, for the form 13.2.e.a.)

Additional possible directions

1. Can we get in a *direct* manner the specialization of the Gritsenko lift? Then maybe we don't need Siegel modular forms at all (until the end when we write down the answer)?

2. Implement Fourier expansions of Siegel paramodular Eisenstein series of weight $k \geq 4$ with character [PS25] and maybe look into Eisenstein series of lower weight [Kal77]?
3. Does specialization to CM points provide a further reduction in complexity? See Colman–Ghitza–Ryan [CGR19] for an analytic version, but perhaps the theory of CM abelian surfaces provides further relationships amongst the specialized coeffiicents (analogous but extending what was observed for specialization to a modular curve) so it is only $O(p)$ to compute $a_{p,1}$?!
4. Compute graded rings of Siegel paramodular forms in small level. Start by reverifying ones that are known (levels $N = 2, 3, 5$?). Relate this via isogeny to the moduli space of genus 2 Jacobians with a cyclic N-subgroup (defined over the ground field), can you compute the map algebraically? Even for $N = 2$?
5. Because dimensions are smaller, the linear algebra would be easier if we could work in weight $2 = 3 - 1$. What is the deal with Siegel modular forms of weight $(1, 0)$ (on a potentially smaller arithmetic group?)?

18.3 AWS proposed project 3: Binary Hermitian

Tagline

Identify forms on U(2) amongst Bianchi modular forms and quaternary orthogonal modular forms.

Estimated difficulty

Medium-easy?

Suggested background

- See Loeffler [Loe08] and Greenberg–Voight [GV14] for algebraic modular forms on unitary groups.

Project overview

Let $Q\colon V \to F$ be a (finite-dimensional) definite Hermitian form relative to an imaginary quadratic field $F \supseteq \mathbb{Q}$. Then the algebraic group $\mathbf{G} = \mathrm{U}_{K|\mathbb{Q}}(Q)$ over $\mathbb{Q}$ is totally definite.

If $\dim_K V = 1$, then $G_\infty = \mathbf{G}(\mathbb{R}) \simeq \mathrm{U}(1) \simeq \mathrm{SO}(2)$, and so the associated modular forms arise from class field theory as Hecke quasicharacters (also known as characters and Grossencharacters) [VW25, §7.2].

So we move up to $\dim_K V = 2$. Loeffler [Loe08, §10] suggested by example an interpretation of forms on U(2): they are exactly the Bianchi modular forms whose 2-dimensional Galois representations over K are *conjugate self-dual*: swapping the Frobenius elements at split primes corresponds to a Tate

twist and dual of the original Galois representation. In his example, the Hecke eigenvalues are equal at (split) primes, $a_{\mathfrak{p}}(f) = a_{\overline{\mathfrak{p}}}(f)$, and the associated form looks like a base change. We seek to generalize this observation and see how close we can get to a proof (assuming conjectures if necessary).

Second, a binary Hermitian form gives by extension of scalars a quaternary quadratic form, giving a map from the class set of a binary Hermitian lattice to a quaternary orthogonal lattice. By pullback, this gives a map from a space of quaternary orthogonal modular forms to a space of binary Hermitian modular forms. We would similarly like to identify the image.

Steps

1. Replicate Loeffler's example [Loe08, §9] in Magma.
2. Write out explicitly the condition on Hecke eigenvalues that is implied by the polarization (conjugate self-dual up to twist) and verify that it holds in further examples.
3. Compute more examples (including higher weight) to find a conjugate self-dual form which is not a base change. Can you find an example that is not a twisted base change (i.e., a form obtained from base change after twist by a Hecke character)?
4. Through examples, investigate how the (weight and) level changes in the base change map: in Loeffler's example, the modular form of level 11 becomes level 1 over $\mathbb{Q}(\sqrt{-11})$!
5. Make a conjecture providing an explicit relationship between binary Hermitian forms and Bianchi modular forms, addressing level and weight.
6. Turn now to the map from unitary to orthogonal. Implement in Magma an algorithm that takes as input a binary Hermitian lattice relative to the ring of integers of an imaginary quadratic field and returns an orthogonal lattice over $\mathbb{Z}$.
7. Provide a local characterization of the lattices one can obtain in the previous step. Which ones can you get and which ones cannot be obtained? Give a name to these which is better than "obtainable".
8. Apply this to Loeffler's example and compute the associated space of orthogonal modular forms in Magma. How does it compare to [GV14, Example 7.2]?
9. Prove that the map from the class set of a binary Hermitian lattice to the class set of an orthogonal lattice is injective. Conclude that we get an injective map on the associated spaces of modular forms. Through examples, investigate how the (weight and) level changes in this change of groups.
10. Make a conjecture providing an explicit relationship between obtainable spaces orthogonal modular forms and binary Hermitian forms, addressing level and weight. Prove your conjecture by showing Hecke equivariance?

The real payoff

Current implementations of Bianchi modular forms are limited in terms of level, weight, and character, presenting an obstruction to explicit methods. Extending methods to work more generally may be a significant project, so we want to see what we can already get from totally definite groups. This becomes even more true when we consider working over a general CM extension $K \supseteq F$ and $F \neq \mathbb{Q}$; then associated methods on symmetric spaces will be even more challenging, but nothing changes for the algebraic modular forms!

Whatever we get, it is important for functoriality reasons to know how spaces of modular forms associated to groups of low rank explicltly relate to one another!

Additional possible directions

1. How does the situation change for GU and SU? We have $\mathrm{SU}(2) \simeq \mathbb{H}^1 \to \mathrm{SO}(3)$ which points us back to classical modular forms.
2. Adapt the above to work for a more general CM extension $K \supseteq F$.
3. Generalize the above to rank $n \geq 3$ (both related to GL_n and O_{2n}.
4. More generally, study explicit functoriality of lifts on (algebraic) modular forms. First, given a Lie algebra $\mathfrak{g}$ of a reductive group $\mathbf{G}$ over a number field F, develop an algorithm to compute all sub Lie algebras $\mathfrak{g}' \subseteq \mathfrak{g}$ corresponding to reductive groups $\mathbf{G}'$ over number fields F'. Write out the predictions of Langlands functoriality by increasing complexity (applying the above algorithm to candidate L-homomorphisms, maps between dual groups that are trivial on the Weil group) and find out when these are known.

18.4 AWS proposed project 4: Eisenstein in average polynomial time

Tagline

Compute Hecke eigenvalues for classical modular forms in average polynomial time.

Estimated difficulty

Medium-easy: analyzing algorithmic complexity and providing/adapting an implementation.

Suggested background

- A toy version of the high-precision q-expansion pipeline (relations on X_Γ plus Newton iteration) is described in [MSV21, §2], specifically the subsection *High-accuracy q-expansions*.

- [RZB15] discusses computing modular forms for $X_0(N)$ to determine 2-adic images. This work implicitly demonstrates the necessity of efficient basis generation when verifying congruences between Galois representations and modular forms at high levels.
- [Zyw24, §4] explains how to obtain modular forms for a congruence group from Eisenstein series on $\Gamma(N)$ and in particular indicates the need for efficient lifting.

Project overview

For a GL_2 modular form over a number field F, many algorithms compute the Hecke operator $T_{\mathfrak{p}}$ in time roughly $(\mathrm{Nm}\,\mathfrak{p})^{1+o(1)}$ (as this is the number of cosets in the double coset decomposition).

For many applications we want not just one eigenvalue but to compute the entire sequence of eigenvalues. In the classical case, the direct calculation gives a method to compute a_p for all primes $p < B$ in approximately $O(B^2)$ time. This quadratic scaling becomes prohibitive for applications requiring large datasets, such as L-function or period computations.

In this project, we compute the full set of eigenvalues a_p with $p < B$ in total time $B(\log B)^{O(1)}$. This complexity corresponds to an *average* time per prime that is polynomial in $\log B$ (quasi-linear total time).

The proposed approach expresses the target modular form in terms of "accessible" forms—specifically Eisenstein series—whose Fourier coefficients can be computed efficiently (e.g., via sieve methods). Previous methods proceeded in two phases:

1. Relation finding: Eisenstein series are used to produce low-degree algebraic relations on the modular curve X_Γ.
2. Newton lifting: Once a relation and a short initial truncation of the form are known (computed via standard modular symbols), Newton iteration is used to lift the q-expansion to precision B in quasi-linear time.

Here we will adapt the above by writing the candidate modular form as a *polynomial* in Eisenstein series (of weight 1) for $\Gamma(N)$, skipping the Newton lift.

The initial phase involves estimating the running time for computing a basis of modular forms via Eisenstein series, depending on the genus and congruence subgroup invariants (cf. [RZB15, Zyw24]). Subsequently, we aim to demonstrate that the q-expansion of any modular form can be computed up to $O(q^B)$ in time quasi-linear in B (cf. [MSV21]), thereby yielding the desired average polynomial time algorithm for Hecke eigenvalues.

Steps

1. Extract the implementation due to Zywina of Eisenstein series (`https://github.com/davidzywina/Modular/blob/main/main/Modular.m`) into a standalone function and experiment with it.

2. Given as input a congruence subgroup of level N (specified by $K_N \leq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, analyze the complexity of computing of q-expansions at ∞ for $M_k(\Gamma)$ with $k \geq 2$ using Eisenstein series on $\Gamma(N)$ up to precision $O(q^B)$ using sieve or divisor-sum methods.
3. Now given Hecke eigenvalues of a purported modular form $f \in M_k(\Gamma)$, estimate the running time to certify $f \in M_k(\Gamma)$ written as an explicit polynomial in Eisenstein series. How does this depend on the parameters (genus, number of cusps, etc.)? Then compare this running time to the method of modular symbols and the method of the trace formula [BB+20, Table 5.2.3].
4. Complete the analysis of the algorithm that, on input Hecke eigenvalues for a form $f \in M_k(\Gamma)$, computes the q-expansion for f up to precision $O(q^B)$. Compare this approach to the one explained in [MSV21] using rational functions.

The real payoff

The primary benefit of this approach is the ability to generate massive datasets of Hecke eigenvalues for high-weight or high-level forms where standard modular symbol algorithms are too slow.

Additional possible directions

1. Are there any ways to improve the algorithm or steps that simplify when Γ is a standard congruence subgroup?
2. Can this method be extended to Hilbert or Siegel modular forms, where the geometric complexity is significantly higher?

18.5 AWS proposed project 5: Symplectic lattices

Tagline

Compute algebraic modular forms for (the compact form of) symplectic groups.

Estimated difficulty

Theoretically easy; implementation details: medium.

Suggested background

- [Voi21, Chapters 9, 10, 20] and [Shi63] for the underlying theory of lattices over quaternion orders. See also [Coh93] as a reference for analogous algorithms in the commutative case.
- [GV14] for an overview of lattice methods for algebraic modular forms, with details on the orthogonal and unitary case.

- [Chi14] for specializing the lattice method to work for the symplectic group by considering quaternionic lattices.
- [LP02] for a group-theoretic method to produce algebraic modular forms based on the root data.
- [Sch18] for an improvement over [LP02], and some results for symplectic groups.
- [CD09] works over $F = \mathbb{Q}(\sqrt{5})$ in rank 2, working under the assumption that the associated class set is trivial, without lattices.

Project overview

In this project, you will compute spaces of modular forms for the compact form of the symplectic group $\mathrm{Sp}(n)$ using lattice methods: so instead of usual lattices (for the orthogonal group) or Hermitian lattices (for the unitary group), we work with quaternion lattices! This generalizes Brandt matrices, which work with ideals (lattices of rank 1) to the general case.

Let B be a definite quaternion algebra, let $O \subseteq B$ be an order, and let Λ be an O-lattice, i.e., a finitely generated projective left O-module, of rank n, equipped with a symplectic form on $V = \Lambda \otimes \mathbb{Q}$ taking values in $\mathbb{Z}$ on Λ.

The symplectic group $\mathrm{Sp}_{r,B|\mathbb{Q}}(V) = \mathrm{Sp}(V)$ is a totally definite (compact at ∞) reductive algebraic group with a distinguished arithmetic subgroup $\mathrm{Sp}(\Lambda) \subseteq \mathrm{Sp}(V)$. Thus, it gives rise to spaces of algebraic modular forms $M_W(\Lambda) = M_W(\mathrm{Sp}(\Lambda))$, for any finite dimensional representation $\rho\colon \mathrm{Sp}(V) \to \mathrm{GL}(W)$.

In this project, you will develop and implement an algorithm to compute bases for the finite dimensional spaces $M_W(\Lambda)$, as well as systems of Hecke eigenvalues in these spaces.

Steps

1. Implement a user-defined type in Magma of a lattice over a quaternion order. You may want to restrict to a nice enough class of quaternion orders (e.g. maximal orders or Bass orders). Following [Shi63], implement an algorithm that returns a pseudobasis for such a lattice. The description of an analogous algorithm in the commutative case in [Coh93] may be helpful.
2. Develop and implement an algorithm to compute a Hermite normal form of a quaternion lattice. The theoretical foundation can be found in [Shi63], and again both [Coh93] can prove helpful.
3. Following [Chi14], develop and implement an algorithm to enumerate the p-neighbors of a given quaternionic lattice, and generalize it to compute Hecke operators acting on the space of algebraic modular forms $M_W(\Lambda)$. The analogous case of orthogonal and unitary groups has been done in [GV14] and may prove to be a useful resource.
4. Compare your algorithm and results to previous computations both in [LP02] and in [Sch18].

The real payoff

We would like to computing Siegel modular forms over more totally real fields and in larger rank in a more robust and scalable fashion! It is also important to see in the implementation how lattice methods extend to the quaternion case. Finally, algorithms for working with higher rank lattices over quaternion orders have already found and will find applications outside the modular forms context.

Additional possible directions

1. Implement the algorithm described in [LP02] for computing algebraic modular forms using group-theoretic methods and compare its performance to lattice methods.
2. Extend lattice methods to work with the (compact form of the) algebraic group G_2. What are p-neighbors? How would you enumerate them? The embedding $G_2 \hookrightarrow \mathrm{SO}(7)$ can help, but notice that you may want also a certain cubic form to be preserved. You may want to implement a variation of the algorithm presented in [PS97].

18.6 AWS proposed project 6: Bad, bad Hecke

Tagline

Compute Hecke operators T_p on spaces $M_k(\Gamma)$ with Γ a congruence subgroup for bad primes p (when p divides the level N).

Estimated difficulty

Medium?

Suggested background

- [DS05, §6] for the relation between modular forms and the Jacobians of modular curves, and [DS05, §8] for the Eichler–Shimura relation at the good primes. For those with background in representation theory and the Langlands program, see [Car86] for an indirect proof that this holds also at bad primes for modular forms associated to elliptic curves.
- [RSZB22] for a definition of the modular curve associated to general congruence subgroups.
- [Shi71] for definition and first properties of Hecke operators on spaces of modular forms.
- [KM85] for the description of modular curves as schemes over $\mathbb{Z}$ with a moduli interpretation. See [Kat72] for some basic properties and the action of Frobenius.
- [Gor02, §4] for Hecke operators in characteristic p on spaces of modular forms for $\Gamma_0(N)$.

- [Ass21] for the identification of the Hecke operators T_p for arbitrary congruence groups at good primes $p \mid N$ and their realization on spaces of modular symbols.
- [ES93] and [PS25] for an example application.

Project overview

As described in the notes, let $\widehat{K} \leq \mathrm{GL}_2(\widehat{\mathbb{Z}})$ be an open subgroup of level N with projection $K_N \leq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ and let $X(\widehat{K})$ be the corresponding modular curve. Let $J(\widehat{K}) = J(X(\widehat{K}))$ be its Jacobian. The LMFDB provides many examples of such curves (where $\widehat{K}$ is called H). One defines Hecke operators $T_{\hat{\alpha}} \colon M_k(\widehat{K}) \to M_k(\widehat{K})$ for any $\hat{\alpha} \in \mathrm{GL}_2(\widehat{\mathbb{Q}})$, represented by an element $\alpha \in \mathrm{GL}_2(\mathbb{Q})_{>0}$ which we may assume has $\alpha \in \mathrm{M}_2(\mathbb{Z})$.

In the "usual" case where $\det(\alpha) = p \nmid N$, there is a unique such Hecke operator $T_p = T_{\hat{\alpha}}$ (as explained in the notes): these operates together are all semisimple, normal, and commute with one another.

But in the special case when $\det(\alpha) = p \mid N$, there are more than one operator in general, they are not semisimple in general, and they need not do not commute. At least in the special case where $\widehat{K}$ is the standard congruence group giving $\Gamma_0(N)$, both Hecke operators T_p for $\mathrm{diag}(1,p)$ and U_p for $\mathrm{diag}(p,1)$ do in fact commute with the other Hecke operators.

More generally, as $J(\widehat{K})$ is a scheme defined over $\mathbb{Z}$, it has Frobenius endomorphisms at every prime p, and the Hecke operators T_p correspond to these endomorphisms via the Eichler–Shimura relation [DS05, Section 8.7].

In this project, you will compute the Hecke operator at p corresponding to the Frobenius automorphism by identifying a matrix α when p divides the level of the (general congruence) subgroup $\widehat{K}$. In a nutshell: the moduli functor identifies points as corresponding to elliptic curves with level structure, we can tell what Frobenius does to that level structure, so we just need to reinterpret that action by a matrix.

Steps

1. Let $p \mid N$. Recall that the Eichler–Shimura relation [DS05, §8.7] states that the Hecke operator $T_p \colon J_0(N) \to J_0(N)$ satisfies $T_p \bmod p = \mathrm{Frob}_p$. Using Magma, verify that this holds for various levels N and primes p, by using the canonical models of the modular curves, and directly constructing the Frobenius on the reduction mod p.
2. Let p be a prime dividing N. Let $T_p \colon J_0(N) \to J_0(N)$ be the Hecke operator corresponding to $\mathrm{diag}(1,p)$. Is it still true that $T_p \bmod p = \mathrm{Frob}_p$, where Frob_p is the Frobenius endomorphism of the fiber $J_0(N) \times \mathbb{F}_p$? Test your hypothesis by explicit computation. See [KM85] for the description of $X_0(N)$ as a scheme over $\mathbb{Z}$, [Gor02, §4.3] for a nice overview and summary of operators in characteristic p, and the appendices in [Kat72] for some useful related background.

3. Let $X_{\mathrm{ns}}^{+}(N)$ be the extended non-split Cartan modular curve of level N, as in [RSZB22]. Consider an element $\alpha \in M_2(\mathbb{Z})$ of determinant p such that the image of α modulo N lies in K_N. Is is true that T_α commutes with T_q for $q \nmid N$? Is it true that $T_\alpha \bmod p = \mathrm{Frob}_p$?
4. Using data from the LMFDB to get equations for modular curves and characteristic polynomials of Frobenius, and using Magma (this package might be useful) to compute Hecke operators T_α for various elements $\alpha \in M_2(\mathbb{Z})$ of determinant p on spaces of modular symbols, find which T_α correspond to the actions of Fr_p. Formulate a conjecture.
5. By tracing through the definition of X_H as a scheme over $\mathbb{Z}$ in [KM85], figure out which T_α should correspond to the Frobenius action at $p \mid N$ on this scheme. Use this description to prove your conjecture.

The real payoff

Efficient algorithms to decompose the Jacobians of modular curves, as in [ES93] and [PS25].

Additional possible directions

1. Consider applications to computing the equations of the modular curves X_H. These should be compared to existing algorithms, see [Box21] and [Zyw24].

Bibliography

[Apo90] Tom M. Apostol, *Modular functions and Dirichlet series in number theory*, Springer-Verlag, New York, 1990. 2.1, 2.3

[AB90] A. Ash and A. Borel, *Generalized modular symbols*, in *Cohomology of Arithmetic Groups and Automorphic Forms* (Luminy–Marseille, 1989), Lecture Notes in Mathematics **1447**, Springer, 1990, 57–75. DOI: 10.1007/BFb0085726. 15, 15.1

[AGG84] Avner Ash, Daniel Grayson, and Philip Green, *Computations of cuspidal cohomology of congruence subgroups of* $\mathrm{SL}(3,\mathbb{Z})$, J. Number Theory **19** (1984), no. 3, 412–436. DOI: 10.1016/0022-314X(84)90081-7.

[AGG84] A. Ash, D. Grayson, and P. Green, *Computations of cuspidal cohomology of congruence subgroups of* $\mathrm{SL}(3,\mathbb{Z})$, *Journal of Number Theory* **19** (1984), no. 3, 412–436. 15

[AGM02] Avner Ash, Paul E. Gunnells, and Mark McConnell, *Cohomology of congruence subgroups of* $\mathrm{SL}_4(\mathbb{Z})$, J. Number Theory **94** (2002), no. 1, 181–212. DOI: 10.1006/jnth.2001.2730. Also available as https://arxiv.org/abs/math/0003219.

[AGM02] A. Ash, P. E. Gunnells, and M. McConnell, *Cohomology of congruence subgroups of* $\mathrm{SL}_4(\mathbb{Z})$, *Journal of Number Theory* **94** (2002), no. 1, 181–212.

[AGM08] Avner Ash, Paul E. Gunnells, and Mark McConnell, *Cohomology of congruence subgroups of* $\mathrm{SL}(4,\mathbb{Z})$ *II*, J. Number Theory **128** (2008), no. 8, 2263–2274. DOI: 10.1016/j.jnt.2007.09.002.

[AGM10] Avner Ash, Paul E. Gunnells, and Mark McConnell, *Cohomology of congruence subgroups of* $\mathrm{SL}_4(\mathbb{Z})$ *III*, Math. Comp. **79** (2010), 1811–1831. DOI: 10.1090/S0025-5718-10-02331-8.

[AGM11] Avner Ash, Paul E. Gunnells, and Mark McConnell, *Torsion in the cohomology of congruence subgroups of* $\mathrm{SL}(4,\mathbb{Z})$ *and Galois representations*, J. Algebra **325** (2011), no. 1, 404–415. DOI: 10.1016/j.jalgebra.2010.07.024.

[AR79] Avner Ash and Louis Rudolph, *The modular symbol and continued fractions in higher dimensions*, Invent. Math. **55** (1979), 241–250. DOI: 10.1007/BF01406900. 15

[AR79] A. Ash and L. Rudolph, *The modular symbol and continued fractions in higher dimensions*, *Inventiones mathematicae* **55** (1979), 241–250. DOI: 10.1007/BF01406842. 15.1

[Ash80] Avner Ash, *Cohomology of congruence subgroups of* $\mathrm{SL}(n, \mathbb{Z})$, Math. Ann. **249** (1980), 55–73. DOI: 10.1007/BF01387080.

[Ass21] Eran Assaf, *Computing classical modular forms for arbitrary congruence subgroups*, Arithmetic geometry, number theory, and computation, Simons Symp., Springer, Cham, 2021, 43–104. 18.6

[A+26] Eran Assaf, Angelica Babei, Ben Breen, Sara Chari, Edgar Costa, Juanita Duque-Rosero, Aleksander Horawa, Jean Kieffer, Avinash Kulkarni, Grant Molnar, Abhijit S. Mudigonda, Michael Musty, Sam Schiavone, Shikhin Sethi, Samuel Tripp, John Voight, *Computing with Fourier expansions of Hilbert modular forms*, preprint, 2026. 18.2

[AL70] A. O. L. Atkin and J. Lehner, *Hecke operators on* $\Gamma_0(m)$, *Mathematische Annalen* **185** (1970), 134–160.

[BB+20] A. J. Best, J. Bober, A. R. Booker, E. Costa, J. E. Cremona, M. Derickx, M. Lee, D. Lowry-Duda, D. Roe, A. V. Sutherland, and J. Voight, *Computing classical modular forms*, `arXiv:2002.04717`. Published version: in *Arithmetic Geometry, Number Theory, and Computation* (Simons Symposia), Springer, 2022, pp. 131–213, DOI: 10.1007/978-3-030-80914-0_4. 4.1, 3

[BC18] Karim Belabas and Henri Cohen, *Modular forms in Pari/GP*, Res. Math. Sci. **5** (2018), no. 3, Paper No. 37.

[BHC62] Armand Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), 485–535. 7.10, 8.2, 8.6

[Bir71] B. J. Birch, Elliptic curves over Q: A progress report, 1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969), Amer. Math. Soc., Providence, R.I., 1971, pp. 396-400.

[Bir91] B. J. Birch, *Hecke actions on classes of ternary quadratic forms*, Computational number theory (Debrecen, 1989), 1991, 191–212. 13.7

[BJ02] A. Borel and L. Ji, *Compactifications of symmetric and locally symmetric spaces*, Math. Res. Lett. **9** (2002), 725–739. 8.9

[BJ06] A. Borel and L. Ji, *Compactifications of Symmetric and Locally Symmetric Spaces*, Birkhäuser, 2006.

[BK75] B. J. Birch and W. Kuyk (eds.), *Modular Functions of One Variable IV*, Lecture Notes in Mathematics, Vol. 476, Springer, 1975. Proceedings of the International Summer School, University of Antwerp (RUCA), 1972.

[BM25] B. C. Berndt and P. Moree, *Sums of two squares and the tau-function: Ramanujan's trail*, *Expositiones Mathematicae* **43** (2025), article 125721.

[Bo91] A. Borel, *Linear Algebraic Groups*, 2nd ed., Graduate Texts in Mathematics, vol. 126, Springer-Verlag, New York, 1991. DOI: 10.1007/978-1-4612-0941-6. 7.2, 7.5, 7.5, 7.5.6, 7.5, 7.7

[Bor62] Armand Borel, *Arithmetic properties of linear algebraic groups*, Proc. Internat. Congr. Mathematicians (Stockholm, 1962), Inst. Mittag-Leffler, Djursholm, 1963, 10–22.

[Bor69] Armand Borel, *Introduction aux groupes arithmétiques*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, vol. XV, Actualités Scientifiques et Industrielles, No. 1341, Hermann, Paris, 1969. 7.10, 8.6

[BS73] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Comment. Math. Helv. **48** (1973), 436–491. DOI: 10.1007/BF02566134. 8.2, 8.8

[BT65] Armand Borel and Jacques Tits, *Groupes réductifs*, Institut des Hautes Études Scientiques, Publications Mathématiques 27 (1965), 55–150. 8.8

[Box21] J. Box, *Computing models for quotients of modular curves*, *Res. Number Theory*, **7** (2021), no. 3. DOI: 10.1007/s40993-021-00276-8. 1

[BPPTVY19] Armand Brumer, Ariel Pacetti, Cris Poor, Gonzalo Tornaría, John Voight, and David S. Yuen, *On the paramodularity of typical abelian surfaces*, Algebra Number Theory **13** (2019), no. 5, 1145–1195. 18.2, 18.2

[Car86] H. Carayol, *Sur les représentations l-adiques associées aux formes modulaires de Hilbert*, *Ann. Sci. École Norm. Sup. (4)*, **19** (1986), no. 3, 409–468. 18.6

[Chi14] S. Chisholm, *Algorithmic enumeration of quaternionic lattices*, Ph. D. Thesis, University of Clagary, 2014. `http://hdl.handle.net/11023/1920`. 18.5, 3

[CJ24] D. Clausen and M. Ø. Jansen, *The reductive Borel–Serre compactification as a model for unstable algebraic K-theory*, Selecta Math. (N.S.) **30** (2024), no. 1, Article 10, 1–93.

[CvdG13] F. Cléry and G. van der Geer, *Generators for modules of vector-valued Picard modular forms*, Nagoya Math. J., **212** (2013), 19–57. DOI: 10.1215/00277630-2324006. 18.1, 1

[CvdG23] F. Cléry and G. van der Geer, *Generating Picard modular forms by means of invariant theory*, Pure Appl. Math. Q., **19** (2023), 95–147. DOI: 10.4310/pamq.2023.v19.n1.a6. 18.1, 2, 4

[Coh93] Henri Cohen, *A course in computational algebraic number theory*, Grad. Texts in Math., vol. 138, Springer-Verlag, Berlin, 1993. 18.5, 1, 2

[Coh00] Henri Cohen, *Advanced topics in computational number theory*, Grad. Texts in Math., vol. 193, Springer-Verlag, New York, 2000.

[CGR19] Owen Colman, Alexandru Ghitza, and Nathan C. Ryan, *Analytic evaluation of Hecke eigenvalues for Siegel modular forms of degree two*, Proceedings of the Thirteenth Algorithmic Number Theory Symposium, Open Book Ser., vol. 2, Math. Sci. Publ., Berkeley, CA, 2019, 207–220. 3

[Con14] B. Conrad, *Reductive group schemes*, Notes from the SGA3 summer school (Luminy, 2011), Soc. Math. France, 2014. Available at https://math.stanford.edu/~conrad/papers/luminysga3.pdf. 7.10

[Cre84] John Cremona, *Hyperbolic tessellations, modular symbols, and elliptic curves over complex quadratic fields*, Compositio Math. **51** (1984), no. 3, 275–324.

[Cre92] J. E. Cremona, *Algorithms for Modular Elliptic Curves*, Cambridge University Press, 1992.

[Cre97] John Cremona, *Algorithms for modular elliptic curves*, 2nd ed., Cambridge University Press, Cambridge, 1997. 3.3

[Cre06] Cremona, J. (2006). The Elliptic Curve Database for Conductors to 130000. In: Hess, F., Pauli, S., Pohst, M. (eds) Algorithmic Number Theory. ANTS 2006. Lecture Notes in Computer Science, vol 4076. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11792086_2

[Cre92] J. E. Cremona, *Modular symbols for $\Gamma_1(N)$ and elliptic curves with everywhere good reduction*, *Mathematical Proceedings of the Cambridge Philosophical Society* **111** (1992), no. 2, 199–218.

[Cre15] J. E. Cremona, *The L-functions and Modular Forms Database Project*, arXiv:1511.04289.

[CD09] Clifton Cunningham and Lassina Dembélé, *Computing genus 2 Hilbert-Siegel modular forms over $\mathbb{Q}(\sqrt{5})$ via the Jacquet-Langlands correspondence*, Experimental Math. **18** (2009), no. 3, 337–345. 18.5

[deG17] Willem Adriaan de Graaf, *Computation with linear algebraic groups*, Monogr. Res. Notes Math., CRC Press, Boca Raton, FL, 2017.

[DD08] L. Dembélé and S. Donnelly, *Computing Hilbert modular forms over fields with nontrivial class group*, in *Algorithmic Number Theory* (ANTS VIII), Lecture Notes in Computer Science 5011, Springer, 2008, 371–386.

[Dem07] L. Dembélé, *Quaternionic Manin symbols, Brandt matrices and Hilbert modular forms*, *Mathematics of Computation* **76** (2007), no. 258, 1039–1057.

[DV13] L. Dembélé and J. Voight, *Explicit methods for Hilbert modular forms*, in *Elliptic curves, Hilbert modular forms and Galois deformations*, Birkhäuser, 2013, 135–198. 13.4

[dSG19] E. de Shalit and E.Z. Goren, *Theta operators on unitary Shimura varieties*, *Algebra Number Theory* **13** (2019), no. 8, 1829–1877. DOI: 10.2140/ant.2019.13.1829. 18.1, 1

[DS05] Fred Diamond and Jerry Shurman, *A first course in modular forms*, Grad. Texts in Math., vol. 228, Springer-Verlag, New York, 2005. 2.1, 2.3, 2.6, 18.6, 18.6, 1

[DHIL24] Taylor Dupuy, Anton Hilado, Colin Ingalls, and Adam Logan, *The basic theory of Clifford–Bianchi groups for hyperbolic n-space*, preprint, 2024, `arXiv:2407.19122`. 8.4.1

[EGM98] Jürgen Elstrodt, Fritz Grunewald, and Jens Mennicke, *Groups acting on hyperbolic space: harmonic analysis and number theory*, Springer-Verlag, Berlin, 1998. 6.3

[ES93] T. Ekedahl and J. P. Serre, *Exemples de courbes algébriques à jacobienne complètement décomposable*, *C. R. Acad. Sci. Paris Sér. I Math.*, **317** (1993), no. 5, 509–513. 18.6, 18.6

[Fin98] T. Finis, *Some computational results on Hecke eigenvalues of modular forms on a unitary group*, *Manuscripta Math.*, **96** (1998), no. 2, 149–180. DOI: 10.1007/s002290050059. 2, 4

[FH91] William Fulton and Joe Harris, *Representation theory: a first course*, Graduate Texts in Math., vol. 129, Springer-Verlag, New York, 1991. 18.1

[GRS13] A. Ghitza, N. Ryan, and J. Sulon, *Computations of vector-valued Siegel modular forms*, *Journal of Pure and Applied Algebra* **217** (2013), 2221–2237.

[Gor02] E. Z. Goren, *Lectures on Hilbert modular varieties and modular forms*, in *CRM Monograph Series*, **14** (2002), With the assistance of M. H. Nicole, American Mathematical Society, Providence, RI. DOI: 10.1090/crmm/014. 18.6, 2

[Gor05] M. Goresky, *Compactifications and cohomology of modular varieties*, Clay Math. Proc. (2005), based on CMI/Fields 2003 summer school notes. 8.8, 8.9

[GT99] M. Goresky and Y.-S. Tai, *Toroidal and reductive Borel–Serre compactifications of locally symmetric spaces*, Amer. J. Math. **121** (1999), 1095–1151.

[GV11] Matthew Greenberg and John Voight, *Computing systems of Hecke eigenvalues associated to Hilbert modular forms*, Math. Comp. **80** (2011), 1071–1092. 13.4

[GV14] M. Greenberg and J. Voight, *Lattice methods for algebraic modular forms on classical groups* in *Computations with modular forms*, Contrib. Math. Comput. Sci., **6** (2014), Springer, Cham, 147–179. DOI: 10.1007/978-3-319-03847-6_6. 18.1, 5, 18.3, 8, 18.5, 3

[Gro99] Benedict Gross, *Algebraic modular forms*, Israel J. Math. **113** (1999), 61–93. 13, 13.1

[GS80] Fritz J. Grunewald and Daniel Segal, *Some general algorithms. I: Arithmetic groups*, Ann. of Math. (2) **112** (1980), no. 3, 531–583. DOI: 10.2307/1971091.

[GS80] F. Grunewald and D. Segal, *Strong general algorithms. I: Arithmetic groups*, *Annals of Mathematics* **112** (1980), no. 3, 531–583. DOI: 10.2307/1971091.

[Gun07] Paul E. Gunnells, *Computing in higher rank*, Appendix A in William Stein, *Modular forms, a computational approach*, Grad. Studies in Math., vol. 79, Amer. Math. Soc., Providence, RI, 2007.

[Gun99] Paul E. Gunnells, *Modular symbols for $\mathbb{Q}$-rank one groups and Voronoĭ reduction*, J. Number Theory **75** (1999), no. 2, 198–219. DOI: 10.1006/jnth.1998.2347. Also available as https://arxiv.org/abs/math/9809058.

[GM03] Paul E. Gunnells and Mark McConnell, *Hecke operators and $\mathbb{Q}$-groups associated to self-adjoint homogeneous cones*, J. Number Theory **100** (2003), no. 1, 46–71. DOI: 10.1016/S0022-314X(02)00041-0. Also available as https://arxiv.org/abs/math/9811133.

[GY08] Paul E. Gunnells and Dan Yasaki, *Hecke operators and Hilbert modular forms*, Algorithmic number theory (ANT -VIII, Banff, 2008), eds. Alf van der Poorten and Andreas Stein, Lecture Notes in Comput. Sci., vol. 5011, Springer, Berlin, 2008, 387–401.

[Hel01] Sigurdur Helgason, *Differential geometry, Lie groups, and symmetric spaces*, corrected reprint of the 1978 original, Grad. Studies in Math., vol. 34, American Mathematical Society, Providence, RI, 2001. 7.8, 13.1

[Hei16] Jeffery Hein, *Orthogonal modular forms: an application to a conjecture of Birch, algorithms and computations*, Ph.D. thesis, Dartmouth College, 2016. 13.7

[HTV25] J. Hein, Gonzalo Tornaría, and John Voight, *Computing Hilbert modular forms as orthogonal modular forms*, arXiv:2506.21981 (2025). 13.7

[Hol95] R. P. Holzapfel, *The ball and some Hilbert problems*, with an appendix by J. E. Sarlabous, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1995, DOI: 10.1007/978-3-0348-9051-9. 18.1

[Hol98] R. P. Holzapfel, *Ball and surface arithmetics*, Aspects of Mathematics, **E29**, Friedr. Vieweg & Sohn, Braunschweig, 1998, DOI: 10.1007/978-3-322-90169-9. 18.1

[Hum75] J. E. Humphreys, *Linear Algebraic Groups*, Graduate Texts in Mathematics, vol. 21, Springer-Verlag, New York, 1975. DOI: 10.1007/978-1-4684-9443-3. 7.2, 7.5, 7.5.6

[Kal77] V.L. Kalinin, *Eisenstein series on the symplectic group*, Math. USSR-Sbornik **32** (1977), no. 4, 449. 2

[Kat72] N. M. Katz, *p-adic properties of modular schemes and modular forms*, in *Modular functions of one variable, III*, Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972. *Lecture Notes in Math.*, **350** (1973), 69–190, Springer, Berlin-New York. 18.6, 2

[KM85] N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, in *Annals of Mathematics Studies*, **108** (1985), Princeton University Press, Princeton, NJ. DOI: 10.1515/9781400881710. 18.6, 2, 5

[Kil15] L.J.P. Kilford, *Modular forms. A classical and computational introduction*, 2nd. ed., Imperial College Press, London, 2015.

[Kul91] R. S. Kulkarni, *An arithmetic geometric method in the study of the subgroups of the modular group*, Amer. J. Math. **13** (1991), no. 6, 1053–1133. 3.1

[KL08] Chris A. Kurth and Ling Long, *Computations with finite index subgroups of* $\mathrm{PSL}_2(\mathbb{Z})$ *using Farey symbols*, Advances in algebra and combinatorics, World Sci. Publ., Hackensack, NJ, 2008, 225–242. 3.1

[LP02] Joshua Lansky and David Pollack, *Hecke algebras and automorphic forms*, Compositio Math. **130** (2002), 21–48. 18.5, 4, 1

[Lang95] Serge Lang, *Introduction to modular forms*, appendixes by D. Zagier and Walter Feit, Grundlehren der Mathematischen Wissenschaften, vol. 222, Springer-Verlag, Berlin, 1995. 2.3

[Leh63] D. H. Lehmer, *Table of Ramanujan's function* $\tau(n)$, 1963 manuscript (164 pages of computer printout), UMT file; see the AMS entry for bibliographic record.

[Leh70] D. H. Lehmer, *Note on the distribution of Ramanujan's tau function*, *Mathematics of Computation* **24** (1970), 343–350.

[LL2012] Wen-Ching Winnie Li and Ling Long, *Atkin and Swinnerton-Dyer congruences and noncongruence modular forms*, Algebraic number theory and related topics 2012, RIMS Kôkyûroku Bessatsu, B51, Res. Inst. Math. Sci. (RIMS), Kyoto, 2014, 269–299, `arXiv:1303.6228`.

[LS76] Ronnie Lee and R. H. Szczarba, *On the homology and cohomology of congruence subgroups*, Invent. Math. **33** (1976), no. 1, 15–53. DOI: 10.1007/BF01425503.

[LS76] R. Lee and R. H. Szczarba, *On the homology and cohomology of congruence subgroups*, *Inventiones mathematicae* **33** (1976), no. 1, 15–53.

[Loe08] David Loeffler, *Explicit calculations of automorphic forms for definite unitary groups*, LMS J. Comput. Math **11** (2008), 326–342. 13.1, 18.1, 5, 18.3, 18.3, 1

[Man72] Yuri Manin, *Parabolic points and zeta functions of modular curves*, *Math. USSR-Izv.* **6** (1972), 19–64. 3.3, 3.3

[Magma] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, *Journal of Symbolic Computation* **24** (1997), no. 3–4, 235–265.

[Mas18] N. Mascot, *Certification of modular Galois representations*, *Mathematics of Computation* **87** (2018), 381–423.

[MSV21] N. Mascot, J. Sijsling and J. M. Voight, *A Prym variety with everywhere good reduction over* $\mathbb{Q}(\sqrt{61})$, in Arithmetic geometry, number theory, and computation, 559–581, Simons Symp., Springer DOI: 10.1007/978-3-030-80914-0_20 18.4, 18.4, 4

[Maz] B. Mazur, *Arithmetic in the geometry of symmetric spaces*, unpublished letter/typescript (mid-1970s; to John Millson), available at https://bpb-us-e1.wpmucdn.com/sites.harvard.edu/dist/a/189/files/2023/01/Arithmetic-in-the-geometry-of-symmetric-spaces.pdf. 15, 15.1

[McMac93] McConnell and MacPherson, *Explicit reduction theory for Siegel modular threefolds*, Inv. Math **111** (1993), 575–625.

[Mer94] Loïc Merel, *Universal Fourier expansions of modular forms*, On Artin's conjecture for odd 2-dimensional representations, Lecture Notes in Math., vol. 1585, Springer, Berlin, 1994, 59–94. 3.3, 3.4

[Mil17] J. S. Milne, *Algebraic Groups: The Theory of Group Schemes of Finite Type over a Field*, Cambridge Studies in Advanced Mathematics, vol. 170, Cambridge University Press, 2017. DOI: 10.1017/9781316711736. 7.2, 7.2, 7.4, 7.6, 7.10

[Miy06] Toshitsune Miyake, *Modular forms*, Translated from the 1976 Japanese original by Yoshitaka Maeda, Springer Monographs in Math., Springer-Verlag, Berlin, 2006. 2.3

[PS25] J. Paulhus and A. V. Sutherland, *Completely decomposable modular Jacobians*, arXiv preprint arXiv:2502.16007, 2025. 18.6, 18.6

[PS25] Erin Pierce and Ralf Schmidt, *Siegel Eisenstein series with paramodular level*, preprint, 2025, `arXiv:2509.04395`. 2

[Piz80] A. Pizer, *An algorithm for computing modular forms on* $\Gamma_0(N)$, *Journal of Algebra* **64** (1980), no. 2, 340–390.

[PR94] V. Platonov and A. Rapinchuk, *Algebraic Groups and Number Theory*, Pure and Applied Mathematics, vol. 139, Academic Press, Inc., Boston, MA, 1994. 7.4, 7.10, 8.6

[PS97] W. Plesken and B. Souvignier, *Computing isometries of lattices*, *Computational algebra and number theory (London, 1993)*, J. Symbolic Comput., **24** (1997), no. 3-4, 327–334. DOI: 10.1006/jsco.1996.0130. 2

[PSY20] Cris Poor, Jerry Shurman, and David S. Yuen, *Nonlift weight two paramodular eigenform constructions*, J. Korean Math. Soc. **57** (2020), no. 2, 507–522. 18.2, 18.2

[PSY24] Cris Poor, Jerry Shurman, and David S. Yuen, *Eigenvalues and congruences for the weight 3 paramodular nonlifts of levels 61, 73, and 79*, J. Korean Math. Soc. **61** (2024), no. 5, 997–1033. 18.2

[PY15] Cris Poor and David S. Yuen, *Paramodular cusp forms*, Math. Comp. **84** (2015), no. 293, 1401–1438. 18.2, 18.2

[RS13] A. D. Rahm and M. H. Şengün, *On level one cuspidal Bianchi modular forms*, *LMS Journal of Computation and Mathematics* **16** (2013), 187–199.

[RSZB22] J. Rouse, A. V. Sutherland, and D. Zureick-Brown, *ℓ-adic images of Galois for elliptic curves over $\mathbb{Q}$ (and an appendix with John Voight)*, *Forum Math. Sigma*, **10** (2022), DOI: 10.1017/fms.2022.38. 18.6, 3

[RZB15] J. A. Rouse and D. Zureick-Brown, *Elliptic curves over $\mathbb{Q}$ and 2-adic images of Galois*, *Res. Number Theory* **1** (2015), Paper No. 12, 34 pp DOI: 10.1007/s40993-015-0013-7 18.4, 18.4

[Run96] B. Runge, *Picard modular surfaces*, in Automorphic forms on algebraic groups (Japanese) (Kyoto, 1995), *Sūrikaisekikenkyūsho Kōkyūroku* **965** (1996), 23–40. 18.1, 3

[Sage] SageMath, `http://www.sagemath.org/`.

[Sch15] George J. Schaeffer, *Hecke stability and weight 1 modular forms*, Math. Z. **281** (2015), no. 1-2, 159–191. 18.2, 18.2

[Sch18] S. Schönnenbeck, *Simultaneous computation of Hecke operators*, *J. Algebra*, **501** (2018), 571–597. DOI: 10.1016/j.jalgebra.2018.01.002. 18.5, 4

[Ser73] Jean-Pierre Serre, *A course in arithmetic*, Grad. Texts in Math., vol. 7, Springer-Verlag, New York, 1973. 2.1, 2.3

[Shi63] G. Shimura, *Arithmetic of alternating forms and quaternion hermitian forms*, *J. Math. Soc. Japan*, **15** (1963), 33–65. DOI: 10.2969/jmsj/01510033. 18.5, 1, 2

[Shi71] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, in *Publications of the Mathematical Society of Japan*, **11** (1994), Reprint of the 1971 original; Kanô Memorial Lectures, no. 1, Princeton University Press, Princeton, NJ. 18.6

[Shi78] G. Shimura, *The arithmetic of automorphic forms with respect to a unitary group*, in *Ann. of Math. (2)*, **107** (1978), no. 3, 569–605. DOI: 10.2307/1971129. 18.1, 1

[Sil2009] Joseph H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Grad. Texts in Math., vol. 106, Springer-Verlag, New York, 2009.

[Spr98] T. A. Springer, *Linear Algebraic Groups*, 2nd ed., Progress in Mathematics, vol. 9, Birkhäuser, Boston, 1998. DOI: 10.1007/978-0-8176-4840-4. 7.2, 7.5, 7.7

[Ste07] William Stein, *Modular forms, a computational approach*, with an appendix by Paul E. Gunnells, Grad. Stud. in Math., vol. 79, Amer. Math. Soc., Providence, RI, 2007. 3.3

[Ste08] William A. Stein, *An introduction to computing modular forms using modular symbols*, Algorithmic number theory: lattices, number fields, curves and cryptography, Math. Sci. Res. Inst. Publ., vol. 44, Cambridge Univ. Press, Cambridge, 2008, 641–652.

[Tit67] Jacques Tits, *Tabellen zu den einfachen Lie Gruppen und ihren Darstellungen*, Lect. Notes Math. vol. 40, Springer, 1967. 7.6

[TY25] K. Thalagoda and D. Yasaki, *Bianchi modular forms over class number 4 fields*, arXiv:2502.00141 (2025).

[vzGG13] Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, 3rd ed., Cambridge Univ. Press University Press, Cambridge, 2013.

[Voi21] John Voight, *Quaternion algebras*, Grad. Texts in Math., vol. 288, Springer, Cham, 2021. 2.1, 2.1.1, 2.2.2, 2.3, 3.5, 8.4.1, 8.6, 9.2, 12.2, 14.1, 2, 18.5

[Voi09] John Voight, *Computing fundamental domains for cofinite Fuchsian groups*, J. Théorie Nombres Bordeaux (2009), no. 2, 467–489. 14.1

[VZB15] John Voight and David Zureick–Brown, *The canonical ring of a stacky curve*, Mem. Amer. Math. Soc. **277** (2022), no. 1362. 2.4, 5.1

[VW25] John Voight and Haochen Wu, *An orthogonal perspective on Gauss composition*, preprint, 2025, `arXiv:2511.03987`. 13.7, 18.3

[Wad71] H. Wada, *Tables of Hecke operators. I*, in *Seminar on Modern Methods in Number Theory* (Inst. Statist. Math., Tokyo), Paper No. 39, 1971, pp. 1–10.

[Wad73] H. Wada, *A table of Hecke operators. II*, *Proceedings of the Japan Academy* **49** (1973), no. 6, 380–384.

[Zyw24] D. Zywina Explicit open images for elliptic curves over https://arxiv.org/abs/2206.14959 18.4, 18.4, 1