

Modular forms and modular symbols for noncongruence groups

by

Christopher Kurth

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Mathematics

Program of Study Committee:

Ling Long, Major Professor

Siu-Hung Ng

Jonathan Smith

Gary Lieberman

Wensheng Zhang

Iowa State University

Ames, Iowa

2009

Copyright © Christopher Kurth, 2009. All rights reserved.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	v
CHAPTER 1. Introduction	1
1.1 Preliminaries	1
1.2 Overview of This Thesis	3
CHAPTER 2. Farey Symbols	5
2.1 Computational Methods	5
2.2 Hyperbolic Geometry	6
2.3 Fundamental Domains	6
2.4 Farey Symbols	8
2.4.1 Special Polygons	8
2.4.2 Farey Symbols	12
2.4.3 Generators	14
2.4.4 Group Invariants	14
2.5 Coset Permutation Representation of a Group	15
2.6 Algorithms	17
2.6.1 Calculating a Farey Symbol	17
2.6.2 Group Membership	21
2.6.3 Coset Representatives	25
2.6.4 Congruence Testing	25
2.6.5 Modular Symbols	26

CHAPTER 3. Some modular forms for noncongruence groups	36
3.1 Introduction	36
3.1.1 Modular Forms for Congruence Groups	36
3.1.2 Modular Forms for Noncongruence Groups	38
3.2 Character Groups	39
3.3 The Main Theorem	43
3.4 Equation of the Modular Curve	44
3.5 Constructing the Function	46
3.6 Proof of the Main Theorem	48
3.7 Appendix: Proofs of Some Results	51
3.7.1 Proof of Theorem 35	51
3.7.2 Proof of Lemma 38	55
BIBLIOGRAPHY	59

ACKNOWLEDGEMENTS

I am grateful to many people who helped me during my time at Iowa State while I wrote my thesis. First and foremost, I want to thank my advisor Ling Long for her expert advice, her enthusiasm, her patience with me, and her many efforts to help along my studies and future career. I would like to thank my other professors, especially Richard Ng and Jonathan Smith for all the help they have given me over the years, and my committee for their criticisms and suggestions on my thesis. I also want to acknowledge my appreciation for the hospitality and financial support of the Math Department and Iowa State University.

I want to thank my family for their support, especially my mom for handling the fine details of planning my wedding while I prepared for my defense. Finally, I would like to express my sincere appreciation to my wife, Olga, whose patience, love, and support have kept me going these last two years.

ABSTRACT

Modular forms for congruence groups are a major area of research in number theory and have been studied extensively. Modular forms for noncongruence groups are less understood. In this thesis, we look at noncongruence groups from two points of view. The first is computational: We look at a method of computation with finite index subgroups of the modular group called Farey symbols. This method allows us to work with noncongruence groups as easily as with congruence groups. We present an algorithm that uses Farey symbols to calculate modular symbols, from which we can calculate cusp forms. The second point of view involves the modular forms of noncongruence groups and the unbounded denominator property. We show that large families of noncongruence groups can be constructed which satisfy the unbounded denominator property.

CHAPTER 1. Introduction

1.1 Preliminaries

Modular forms are functions that are “very symmetric” with respect to certain groups of transformations. Let $\mathrm{SL}_2(\mathbb{Z})$ be the set of 2×2 integer-valued matrices with determinant 1. Let $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ be the upper half plane. Then there is an action of $\mathrm{SL}_2(\mathbb{Z})$ on \mathbb{H} given by

$$\gamma z = \frac{az + b}{cz + d} \quad \text{where } \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

If Γ is a finite-index subgroup of $\mathrm{SL}_2(\mathbb{Z})$, the weight k modular forms for Γ are the holomorphic functions $f(z)$ on the upper half plane that satisfy $f(\gamma z) = (cz+d)^k f(z)$ for every $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$ (and also a growth condition on their Fourier coefficients). The vector space of modular forms of weight k will be denoted by $M_k(\Gamma)$.

The element $-I = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ acts trivially on \mathbb{H} , so sometimes it will be convenient to instead consider subgroups of $\mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{I, -I\}$. The action of $\mathrm{PSL}_2(\mathbb{Z})$ on the upper half plane is a faithful action.

In addition to acting on \mathbb{H} , $\mathrm{SL}_2(\mathbb{Z})$ acts on the set $\mathbb{Q} \cup \{\infty\}$, which we will denote $\mathbb{P}^1(\mathbb{Q})$. If Γ is a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ then the action of Γ partitions $\mathbb{P}^1(\mathbb{Q})$ into a finite number of equivalence classes, where $q_1 \sim q_2$ if $q_1 = \gamma q_2$ for some $\gamma \in \Gamma$. These equivalence classes $\{q\}$ are called the **cusps** of Γ and the **width** of the cusp $\{q\}$ is $[\mathrm{Stab}_{\mathrm{PSL}_2(\mathbb{Z})}(q) : \mathrm{Stab}_{\Gamma}(q)]$. We say that the **level** of Γ is the least common multiple of the cusp widths of Γ .

If a point z of \mathbb{H} is fixed by any $\gamma \in \Gamma$ other than I or $-I$ we call z an elliptic point. For

example:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} i = \frac{-1}{i} = i \quad \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \mu_6 = \frac{\mu_6 - 1}{\mu_6} = \mu_6$$

where $\mu_6 = e^{2\pi i/6}$. Again there is an equivalence relation, $z_1 \sim z_2$ if $z_1 = \gamma z_2$ for some $\gamma \in \Gamma$. This partitions the elliptic points into finitely many (possibly zero) equivalence classes called **inequivalent elliptic points** of Γ .

The modular forms for a finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ can be thought of as forms on a certain Riemann surface, called the **modular curve** of Γ . Consider the upper half plane \mathbb{H} modulo the action by elements of Γ on \mathbb{H} . There is a natural Riemann surface structure, and we can get a compact Riemann surface by adding an extra point for each cusp of Γ . This Riemann surface is the modular curve for Γ , and we will denote it by X_Γ . The genus of Γ is defined to be the genus of X_Γ as a compact orientable surface.

The weight k modular forms which vanish at the cusps (as functions locally on the modular curve) are of particular interest. These are called **cusp forms**, and they form a vector space denoted by $S_k(\Gamma)$.

Let N be a positive integer. The following subgroups are standard examples of finite index subgroups of $\mathrm{SL}_2(\mathbb{Z})$:

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

We will use the same notation for these groups considered as subgroups of $\mathrm{PSL}_2(\mathbb{Z})$. It should be clear from context which meaning is intended.

$\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Any subgroup of $\mathrm{SL}_2(\mathbb{Z})$ containing $\Gamma(N)$ is said to be a **congruence group of level N** if N is the smallest N such that $\Gamma(N)$ is contained

in the subgroup (this notion of level agrees with the general one defined above). The congruence groups of level dividing N are easy to describe, as they correspond to the subgroups of $\mathrm{SL}_2(\mathbb{Z})/\Gamma(N) \cong \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. If π is the projection $\pi : \mathrm{SL}_2(\mathbb{Z}) \rightarrow \mathrm{SL}_2(\mathbb{Z})/\Gamma(N)$, and $\bar{\Gamma}$ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})/\Gamma(N)$, then there is a corresponding subgroup:

$$\Gamma = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \pi(\gamma) \in \bar{\Gamma}\}$$

So congruence groups can be described completely by congruence relations, hence the name *congruence group*.

Congruence groups and their modular forms are well-studied and well-understood ([23], [18]). **Noncongruence groups**, finite-index subgroups of $\mathrm{SL}_2(\mathbb{Z})$ which are not congruence groups, are much less understood. One reason for this is that without the congruence structure they can be difficult to even describe. Often they are given only by a set of generators. Additionally, noncongruence groups lack efficient Hecke operators ([2]). The Hecke operators are a commuting family of linear operators on modular forms that play a central role in the study of the Fourier coefficients of congruence modular forms. Unfortunately, while Hecke operators can be defined for noncongruence modular forms, they act trivially on them and do not add anything useful to their study.

1.2 Overview of This Thesis

The goal of this thesis is the study of noncongruence groups from two different points of view. The first is computational: In Chapter 2 we discuss a number of computational methods for working with finite-index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$, primarily Farey symbols, which are well-suited for working with noncongruence groups. We discuss some algorithms involving Farey symbols, including how to use Farey symbols with Margaret Millington's permutation representation to test whether a group is noncongruence. We also present a new algorithm for computing modular symbols from Farey symbols, which allows us to compute Fourier coefficients for cusp forms of some congruence groups.

The second part the paper, Chapter 3, discusses the conjectured unbounded denominator property of noncongruence groups. Modular forms for congruence groups satisfy a bounded

denominator property: If a modular form for a congruence group has algebraic Fourier coefficients, then it can be multiplied by a nonzero scalar to get a modular form with algebraic integer coefficients. This property fails for noncongruence groups, and it is conjectured that the opposite is true: If a modular form for a noncongruence group (and no larger congruence group) has algebraic integer coefficients, then no nonzero scalar multiple of it can have all of its Fourier coefficients be algebraic integers. We consider a special family of noncongruence groups called character groups, and show that for most congruence groups Γ^0 and arbitrarily large N , for all but finitely many primes p there are at least N different index- p character groups of Γ^0 satisfying the unbounded denominator property.

CHAPTER 2. Farey Symbols

2.1 Computational Methods

As modular forms are functions whose symmetries are finite-index subgroups of $SL_2(\mathbb{Z})$, it is important to have methods of calculating with these groups. The congruence groups, as stated above, can be completely described by congruence relations, so they are relatively easy to work with. In fact, most computational methods for working with modular forms apply only to congruence subgroups (cf. [24]). Noncongruence subgroups are harder to work with. Even though they make up the majority of finite-index subgroups of $PSL_2(\mathbb{Z})$ there are few tools for working with them. Two tools that work equally well with both congruence and noncongruence groups are the method of coset permutation representations of Margaret Millington [21], and the method of Farey symbols introduced by Ravi Kulkarni [12]. In this chapter we will primarily discuss Farey symbols and various algorithms for computing with them. We will discuss Millington's coset permutation representation and how to compute it from a Farey symbol in order to practically use Tim Hsu's method [8] for determining if a group is noncongruence. We will also discuss another computation tool, modular symbols, useful for calculating Hecke operators. We will present a new algorithm for computing modular symbols from Farey symbols which works with arbitrary congruence groups (not only $\Gamma_0(N)$), as well as noncongruence groups.

The methods presented here have been or are being implemented by the author in a software package called "KFarey" for the computer algebra system SAGE. The author is interested in implementing other existing algorithms such as the congruence closure algorithm in [9] and the normalizer algorithm in [15], and extending the Farey symbol algorithms to general Hecke groups.

2.2 Hyperbolic Geometry

The Farey symbol method uses the hyperbolic geometry of the upper half plane \mathbb{H} , so we recall the basic properties. There are two kinds of hyperbolic lines: The first is a half circle with center on the real line (so the circle intersects the real line at a 90° angle. The second is a vertical line from a real number to $i\infty$. For any two distinct points on $\mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$, there is a unique hyperbolic line passing through x and y . The hyperbolic arc length of a curve C is defined to be:

$$\int_a^b \frac{\sqrt{(x')^2 + (y')^2}}{y} dt \quad \text{where } C \text{ is parametrized by: } (x(t), y(t)), a \leq t \leq b$$

and the shortest path between distinct points x and y is the hyperbolic line segment joining them. We will write $[x, y]$ for this line segment.

We say a hyperbolic polygon is a polygon in the usual sense with hyperbolic line segments for sides. The hyperbolic area of a region $S \subset \mathbb{H}$ is defined to be:

$$\text{Area} = \int_S \frac{1}{y^2} dx dy$$

The Gauss-Bonnet theorem tells us that the area of a hyperbolic triangle is completely determined by its angles, θ_1 , θ_2 , and θ_3 :

$$\text{Area} = \pi - \theta_1 - \theta_2 - \theta_3$$

The elements of $\text{SL}_2(\mathbb{Z})$ (and more generally of $\text{SL}_2(\mathbb{R})$) act on the upper half plane as orientation preserving translations. This means γ on $\text{SL}_2(\mathbb{R})$ sends hyperbolic lines to hyperbolic lines, and it preserves arc length, area, and angles.

2.3 Fundamental Domains

Farey symbols describe a group using the geometry of a “fundamental domain” of the group, that is, a particular set of orbit representatives of the action of the group on the upper half plane. Formally:

Definition 1. Let Γ be a finite index subgroup of $\text{PSL}_2(\mathbb{Z})$. For our purposes, a **fundamental domain** of Γ is a hyperbolic polygon P (including the boundary) on $\mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$ such that:

1. If z is in the interior of P and $\gamma \in \Gamma$, then $\gamma z \in P$ implies $\gamma z = z$.
2. For every $z \in \mathbb{H}$ there is $\gamma \in \Gamma$ such that $\gamma z \in P$.

Lemma 2. *Let Γ be a finite index subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ and P a hyperbolic polygon. Suppose P is such that:*

1. *If z is in the interior of P and $\gamma \in \Gamma$, then $\gamma z \in P$ implies $\gamma z = z$.*
2. *For each side e of P , there is $\eta \in \Gamma$ such that η maps e to another side of P in an orientation reversing manner.*
3. *The sum of the angles at each set of corners matched by the pairing is either 2π (if the corners are on \mathbb{H}) or 0 (if the corners are on $\mathbb{P}^1(\mathbb{Q})$).*

Then P is a fundamental domain of Γ .

Proof. Let $Q = \bigcup_{\gamma \in \Gamma} \gamma P$. It is sufficient to prove that $Q = \mathbb{H}$. Let Q^c be the connected component of Q containing P . First we claim that every point of Q^c is an interior point, for suppose there is x in Q^c which is not. Then x must equal $\gamma x'$ for some $\gamma \in \Gamma$ and x' which is not an interior point of P . Since x' is not an interior point of P , it is either an edge point or a vertex. If it is an edge point such that $\eta x' \in P$ but $\eta x' \neq x'$ then x' is an interior point of $P \cup \eta^{-1}P$, and x is an interior point of $\gamma(P \cup \eta^{-1}P) \subseteq Q^c$, contradicting that x is not an interior point of Q^c . We proceed in the same way if x' is a vertex, using the fact that vertices of P fit together at an angle of 2π . Thus Q^c is open.

Now we show that Q^c cannot have any limit points not in Q^c , for suppose y is such a point. Then there is a sequence x_i in P and γ_i in Γ such that $\gamma_i x_i$ converges to y . Let

$$P_t = \{z \in P : 1/t \leq \mathrm{Im}(z) \leq t\}$$

for any $t > 1$. P_t is compact, so if the sequence x_i is bounded inside P_t for any t , then x_i converges, say to x . If U is an open disk around y , \overline{U} intersects γP_t for only finitely many $\gamma \in \Gamma$ (Proposition 1.6 of [23]). At least one of them, say γ' , occurs infinitely often in the list, so y is the limit of a subsequence $\gamma' x_{i_k}$. But this limit is $\gamma' x$. This means y is in Q^c .

So we can assume no P_t contains all the terms x_i , hence x_i has a subsequence that converges to a vertex of P at ∞ or to a rational number, say $x_{i_k} \rightarrow x \in \mathbb{Q} \cup \{\infty\}$. We may assume it converges to ∞ , for otherwise we can multiply P by an element of $\mathrm{SL}_2(\mathbb{Z})$ to move x to ∞ . Choose ε such that $0 < \varepsilon < \mathrm{Im}(y)$ and let N be such that for every $k \geq N$, $\mathrm{Im}(x_{i_k}) > \max\{y + \varepsilon, \frac{1}{y - \varepsilon}\}$. Note that for $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have $\mathrm{Im}(\gamma z) = |cz + d|^{-2} \mathrm{Im}(z)$, so if $\mathrm{Im}(x_{i_k}) > N$ then if $|c| > 0$:

$$\begin{aligned} \mathrm{Im}(\gamma x_{i_k}) &= ((c \mathrm{Re}(x_{i_k}) + d)^2 + c^2 \mathrm{Im}(x_{i_k})^2)^{-1} \mathrm{Im}(x_{i_k}) \\ &< (c^2 (\mathrm{Im}(x_{i_k})^2)^{-1} \mathrm{Im}(x_{i_k})) \\ &\leq \mathrm{Im}(x_{i_k})^{-1} \\ &< y - \varepsilon \end{aligned}$$

and if $c = 0$, then $\mathrm{Im}(\gamma(x_{i_k})) > y + \varepsilon$. So the sequence $\gamma_{i_k} x_{i_k}$ is bounded away from y . This is a contradiction, so y cannot exist, and Q^c cannot have limit points not in Q^c . Thus $Q = \mathbb{H}$.

□

2.4 Farey Symbols

2.4.1 Special Polygons

Farey Symbols were introduced by Ravi Kulkarni in 1991 [12] as a compact and efficient way to compute with finite index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$. The idea is to describe the group by a fundamental domain with vertices at certain rational numbers and certain hyperbolic arcs joining these rational numbers. Much of the theory here is summarized from [12].

Throughout this thesis, when a vertex of a hyperbolic arc is in \mathbb{Q} it will always be assumed to be in the form $\frac{a}{b}$ with $a, b \in \mathbb{Z}$, $(a, b) = 1$ and $b > 0$. If the vertex is ∞ , we will write it either $\frac{-1}{0}$ or $\frac{1}{0}$ (depending on if it is the leftmost or rightmost element of a Farey sequence).

Let $\rho = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ and let T be the hyperbolic triangle with vertices ρ , ρ^2 and ∞ . Then T is a fundamental domain for $\mathrm{PSL}_2(\mathbb{Z})$ ([11] Prop. III.1). Let E_e be the edge joining i to ∞ , E_o be the edge joining ρ to ∞ , and E_f be the edge joining i to ρ . Then we call an arc A in the

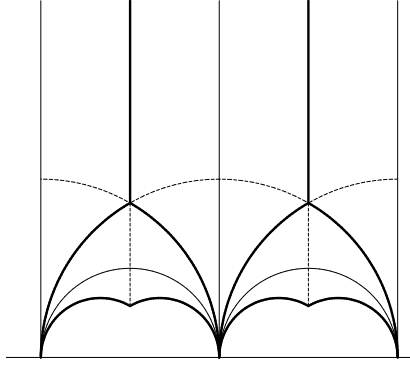


Figure 2.1 Even edges are thin, odd edges are thick, and f-edges are dashed

upper half plane an **even edge** (resp. **odd edge**, resp. **f-edge**) if $A = \gamma E_e$ (resp. $A = \gamma E_o$, resp. $A = \gamma E_f$) for some $\gamma \in \text{PSL}_2(\mathbb{Z})$ (See Figure 2.1). E_e and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} E_e$ together form a hyperbolic arc from 0 to ∞ , and in general even edges come in pairs joining rational numbers $\frac{a}{b}$, and $\frac{a'}{b'}$ with $|a'b - ab'| = 1$ because of the following lemma:

Lemma 3. *If $\gamma \in \text{PSL}_2(\mathbb{Z})$ and a_1/b_1 , a_2/b_2 , a'_1/b'_1 , and a'_2/b'_2 are rational numbers in simplest form such that*

$$\gamma(a_1/b_1) = a'_1/b'_1, \text{ and } \gamma(a_2/b_2) = a'_2/b'_2$$

then

$$a_2b_1 - a_1b_2 = a'_2b'_1 - a'_1b'_2.$$

Proof. If $\gamma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ then

$$\gamma\left(\frac{a_1}{b_1}\right) = \frac{Aa_1 + Bb_1}{Ca_1 + Db_1} = \frac{a'_1}{b'_1} \text{ and } \gamma\left(\frac{a_2}{b_2}\right) = \frac{Aa_2 + Bb_2}{Ca_2 + Db_2} = \frac{a'_2}{b'_2}.$$

So:

$$\begin{aligned} a'_1b'_2 - a'_2b'_1 &= (Aa_1 + Bb_1)(Ca_2 + Db_2) - (Aa_2 + Bb_2)(Ca_1 + Db_1) \\ &= ADa_1b_2 + BCa_2b_1 - ADa_2b_1 - BCa_1b_2 \\ &= (AD - BC)(a_1b_2 - a_2b_1) \\ &= (a_1b_2 - a_2b_1) \end{aligned}$$

□

So the quantity $a_2b_1 - a_1b_2$ is invariant under transformations in $\mathrm{PSL}_2(\mathbb{Z})$. We introduce the notation:

$$\delta(a_1/b_1, a_2/b_2) = \begin{cases} a_2b_1 - a_1b_2 & \text{if } b_1 \neq 0 \\ b_2 & \text{if } b_1 = 0 \end{cases}$$

and we sometimes write $\delta(x, y)$ where it is assumed x and y are fractions written in simplest form with b_1 and b_2 nonnegative. The lemma then says that $\delta(\gamma x, \gamma y) = \delta(x, y)$.

Note that even edges, odd edges and free edges only map to even edges, odd edges and free edges respectively under transformations $\gamma \in \mathrm{PSL}_2(\mathbb{Z})$.

Definition 4. A **special polygon** P is a convex (in the hyperbolic sense) hyperbolic polygon together with a side pairing defined in the following way: The polygon is such that:

1. The boundary of P consists of even and odd edges.
2. The even edges of P come in pairs, each pair forming a hyperbolic arc between elements of $\mathbb{P}^1(\mathbb{Q})$.
3. The odd edges of P come in pairs, each pair meeting a vertex with inner angle $\frac{2\pi}{3}$.

The sides of the polygon are denoted as follows:

1. Each odd edge is called an **odd side**.
2. As even edges come in pairs, either each edge of the pair is called an **even side**, or the union of the two edges (a semicircle) is called a **free side**.

The side pairing on the edges is defined as follows:

1. Each odd side is paired with the odd side it meets at an angle of $\frac{2\pi}{3}$. This is called an **odd pairing**.
2. Each even side is paired with the even side with which it forms a semicircular arc. This is called an **even pairing**.
3. There are an even number of free sides and they are paired off into twos, each called a **free pairing**.

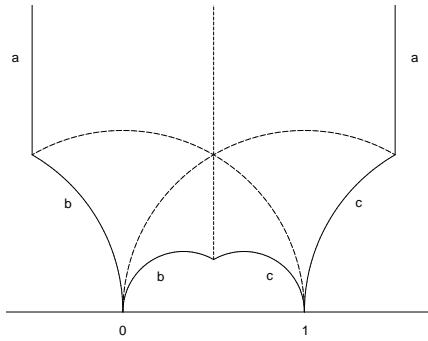


Figure 2.2 A fundamental domain for $\Gamma(2)$

We will always assume that 0 and ∞ are vertices of P (If P does not contain these vertices, an equivalent P can be calculated with the LLT algorithm to be described in Section 2.6.2).

The sides of a special polygon P have a natural orientation obtained by tracing the perimeter of the polygon in a certain direction. If s and s' are paired sides then there is a unique $\gamma \in \text{PSL}_2(\mathbb{Z})$ such that γ maps s to s' in an orientation-reversing manner. We call this the **side pairing transformation** associated with the side pairing, and we let Γ_P be the group generated by all the side pairing transformations of P . Note that it doesn't matter which side we pick for s and which for s' because the two possible γ 's are inverses of each other. Also note that if s is an even side (resp. odd side) then γ is order 2 (resp. order 3).

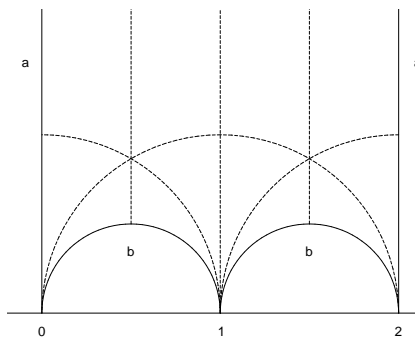
Two theorems of Kulkarni are fundamental here:

Theorem 5 ([12] Theorem 3.2). *If P is a special polygon then P is a fundamental domain for Γ_P . Moreover, the side pairing transformations $\{\gamma_i\}$ are an independent set of generators of Γ_P (i.e. the only relations on the γ_i 's are $\gamma_i^2 = 1$ or $\gamma_i^3 = 1$ for any side pairings coming from even and odd pairings respectively).*

Theorem 6 ([12] Theorem 3.3). *For every $\Gamma \subset \text{PSL}_2(\mathbb{Z})$ of finite index, there is a special polygon P such that $\Gamma = \Gamma_P$.*

Proof. [12] but this also follows from the proof of the algorithm in Section 2.6.1. □

Note that although it is true that any subgroup of $\text{PSL}_2(\mathbb{Z})$ with fundamental domain F is generated by the transformations that map its edges together, the fact that the set of

Figure 2.3 A special polygon for $\Gamma(2)$

generators of a special polygon is an independent set of generators is something special to the special polygon. For example $\Gamma(2)$ has a fundamental domain shown in Figure 2.2. There are six sides, and the three side pairing transformations are

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}$$

But this is not an independent list of generators because

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

A special polygon for $\Gamma(2)$ is shown in Figure 2.3. The pairing transformations from the special polygon are

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

These are independent generators of $\Gamma(2)$, which is isomorphic to the free product $\mathbb{Z} * \mathbb{Z}$.

2.4.2 Farey Symbols

Recall that the classical Farey sequences of order n are constructed by taking all the rational numbers $0 \leq a/b \leq 1$ with denominator at most n and $(a, b) = 1$ and writing them as a finite sequence in ascending order $\{a_0/b_0, \dots, a_n/b_n\}$. Then for each i we have $\delta(a_i/b_i, a_{i+1}/b_{i+1}) = a_{i+1}b_i - a_ib_{i+1} = 1$. We are interested in sequences that satisfy this condition.

Definition 7. A **generalized Farey sequence** is a finite sequence:

$$\left\{ \frac{-1}{0}, x_0, \dots, x_n, \frac{1}{0} \right\}$$

such that:

1. Each $x_i = a_i/b_i$ is a rational number in reduced form with $b_i > 0$. Additionally, we consider $x_{-1} = \frac{-1}{0}$ and $x_{n+1} = \frac{1}{0}$.
2. $\delta(x_i, x_{i+1}) = 1$ for $-1 \leq i \leq n$.

Note that this definition forces x_0 and x_n to be integers. We will always assume $x_i = 0$ for some i .

Definition 8. A **Farey symbol** is a generalized Farey sequence with some additional pairing information. Namely, between each adjacent entries x_{i-1} and x_i we assign a **pairing** p_i which is either a positive integer called a **free pairing** or the symbol “ \circ ” called an **even pairing** or “ \bullet ” called an **odd pairing**. Each integer that appears as a free pairing appears exactly twice in the pairing information.

So if P is a special polygon, let x_0, \dots, x_n be the vertices of P lying in \mathbb{Q} listed in ascending order. Recall these vertices satisfy $\delta(x_i, x_{i+1}) = 1$, so $\{\frac{-1}{0}, x_0, \dots, x_n, \frac{1}{0}\}$ is a generalized Farey sequence. We make a Farey symbol out of the generalized Farey sequence by adding the pairing information in the obvious way.

On the other hand, if F is a Farey symbol we can construct a special polygon for F . For adjacent entries of the Farey sequence, x_{i-1} and x_i , if p_i is a free pairing or an even pairing we let P have as a side $[x_{i-1}, x_i]$, the hyperbolic arc joining x_{i-1} and x_i . Otherwise if it is odd we let γ be the unique element of $\text{PSL}_2(\mathbb{Z})$ such that $\gamma(0) = x_{i-1}$ and $\gamma(\infty) = x_i$ and join x_{i-1} and x_i by the arcs $\gamma([0, \rho])$ and $\gamma([\rho, \infty])$. Thus we get a hyperbolic polygon which is made into a special polygon by adding pairing information in the obvious way.

Example 9. $\Gamma(2)$ has a Farey symbol $-\infty \underset{1}{\frown} \frac{0}{1} \underset{2}{\frown} \frac{1}{1} \underset{2}{\frown} \frac{2}{1} \underset{1}{\frown} \infty$. (cf. Figure 2.3)

2.4.3 Generators

Recall that if P is a special polygon for a group Γ then Γ is generated independently (as a free product) by the transformations mapping each side to its paired side. If F is a Farey symbol:

$$-\infty \underset{p_0}{\frown} a_0/b_0 \underset{p_1}{\frown} a_1/b_1 \underset{p_2}{\frown} \cdots \underset{p_{n-1}}{\frown} a_{n-1}/b_{n-1} \underset{p_n}{\frown} a_n/b_n \underset{p_{n+1}}{\frown} \infty$$

then we can explicitly give formulas for the γ corresponding to a given side pairing.

Theorem 10. *Suppose $(a_i/b_i, a_{i+1}/b_{i+1})$ are two adjacent vertices of F . Then if the pairing between them p_{i+1} is an even pairing, let:*

$$G_{i+1} = \begin{bmatrix} a_{i+1}b_{i+1} + a_i b_i & -a_i^2 - a_{i+1}^2 \\ b_i^2 + b_{i+1}^2 & -a_{i+1}b_{i+1} - a_i b_i \end{bmatrix}.$$

If p_{i+1} is an odd pairing, let:

$$G_{i+1} = \begin{bmatrix} a_{i+1}b_{i+1} + a_i b_{i+1} + a_i b_i & -a_i^2 - a_i a_{i+1} - a_{i+1}^2 \\ b_i^2 + b_i b_{i+1} + b_{i+1}^2 & -a_{i+1}b_{i+1} - a_{i+1} b_i - a_i b_i \end{bmatrix}.$$

And if p_{i+1} is a free pairing that is paired with the side between a_k/b_k and a_{k+1}/b_{k+1} , let:

$$G_{i+1} = \begin{bmatrix} a_{k+1}b_{i+1} + a_k b_i & -a_k a_i - a_{k+1} a_{i+1} \\ b_k b_i + b_{k+1} b_{i+1} & -a_{i+1} b_{k+1} - a_i b_k \end{bmatrix}.$$

Then G_{i+1} is the side transformation corresponding to the pairing p_{i+1} . I.e.

$$\text{If } p_{i+1} \text{ is even: } \quad G_{i+1}(a_i/b_i) = a_{i+1}/b_{i+1} \text{ and } G_{i+1}(a_{i+1}/b_{i+1}) = a_i/b_i$$

$$\text{If } p_{i+1} \text{ is odd: } \quad G_{i+1}^2(a_i/b_i) = a_{i+1}/b_{i+1} \text{ and } G_{i+1}(a_{i+1}/b_{i+1}) = a_i/b_i$$

$$\text{If } p_{i+1} \text{ is free: } \quad G_{i+1}(a_i/b_i) = a_{k+1}/b_{k+1} \text{ and } G_{i+1}(a_{i+1}/b_{i+1}) = a_k/b_k$$

Proof. [12] Theorem 6.1 □

2.4.4 Group Invariants

Several invariants of the group Γ can be read off from the Farey symbol F . Firstly, the number of inequivalent order 2 (resp. order 3) elliptic points, e_2 (resp. e_3), is the number of even (resp. odd) pairings in F . Also, the number of free pairings in F (half the number of

free edges) is equal to r , the rank of $\pi_1(\Gamma \backslash \mathbb{H})$ (the fundamental group of the uncompactified modular curve).

To discuss the cusps of Γ , note that if (x_i, x_{i+1}) is an edge with an even or odd pairing, then x_i and x_{i+1} are equivalent cusps (since $G_{i+1} \in \Gamma$ maps x_i to x_{i+1}). Likewise, if (x_i, x_{i+1}) and (x_j, x_{j+1}) are paired edges then x_i and x_{j+1} are equivalent cusps and x_j and x_{i+1} are equivalent cusps. This defines an equivalence relation on the vertices of P . The equivalence classes are easy to compute, because the defining equivalences occur in a cyclic pattern (When the edges of the fundamental domain are pasted together the corners at equivalent cusps fit together like at the center of a pizza). So the number of cusps t can be counted as the number of equivalence classes.

For an edge $(\frac{a_i}{b_i}, \frac{a_{i+1}}{b_{i+1}})$ let $\gamma = \begin{bmatrix} a_i & a_{i+1} \\ b_i & b_{i+1} \end{bmatrix}$. So $\gamma^{-1}(x_i) = \infty$ and $\gamma^{-1}(x_{i+1}) = 0$. Then define the **width** of a vertex x_i to be the “width” of γP at ∞ . That is:

$$\text{width}(x_i) = \begin{cases} |a_{i-1}b_{i+1} - a_{i+1}b_{i-1}| & \text{if } x_i \text{ is adjacent to no odd edge} \\ |a_{i-1}b_{i+1} - a_{i+1}b_{i-1}| + 1/2 & \text{if } x_i \text{ is adjacent to 1 odd edge} \\ |a_{i-1}b_{i+1} - a_{i+1}b_{i-1}| + 1 & \text{if } x_i \text{ is adjacent to 2 odd edges} \end{cases}$$

The cusp width of a cusp x of Γ is then the sum of the widths of the vertices of P which are Γ -equivalent to x .

$\Gamma \backslash \mathbb{H}$ is a genus g orientable surface with t points missing, one for each cusp. The rank of its fundamental group is $r = 2g + t - 1$, so we can calculate the genus $g = \frac{r-t+1}{2}$. Moreover, using the Hurwitz formula ([23] Prop. 1.40) we get the index of Γ in $\text{PSL}_2(\mathbb{Z})$, $\mu = 3e_2 + 4e_3 + 12g + 6t - 12$. An even simpler formula for the index comes from noting that $n+2 = 2r + e_2 + e_3$ where $n+1$ is as in Definition 7. This, combined with the previous formula, implies $\mu = 3n + e_3$.

2.5 Coset Permutation Representation of a Group

Another method of representing groups that will be useful to us in determining if a group is congruence is the coset permutation representation of Millington [20] [21]. Let Γ be a subgroup of $\text{PSL}_2(\mathbb{Z})$ with $[\text{PSL}_2(\mathbb{Z}) : \Gamma] = \mu$ and $\text{PSL}_2(\mathbb{Z}) = \bigcup_{i=1}^{\mu} \alpha_i \Gamma$ a coset decomposition

with $\alpha_1 = I$. Let F be the standard fundamental domain for $\mathrm{PSL}_2(\mathbb{Z})$. Then $\bigcup_{i=1}^{\mu} \alpha_i^{-1}F$ is a fundamental domain for Γ . Let

$$E = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

E and V generate $\mathrm{PSL}_2(\mathbb{Z})$, as do L and R . The conversions between them are:

$$\begin{aligned} E &= LR^{-1}L, & V &= R^{-1}L \\ L &= EV^{-1}, & R &= EV^{-2} \end{aligned} \tag{2.1}$$

We have $E^2 = V^3 = 1$. In fact it is well-known that $\mathrm{PSL}_2(\mathbb{Z})$ is isomorphic to the group (cf. [22]):

$$\mathrm{PSL}_2(\mathbb{Z}) \cong \langle e, v : e^2 = v^3 = 1 \rangle. \tag{2.2}$$

For each γ in $\mathrm{PSL}_2(\mathbb{Z})$, left multiplication acts on the left cosets of Γ in $\mathrm{PSL}_2(\mathbb{Z})$ by permutation, i.e. there is a homomorphism $\varphi : \mathrm{PSL}_2(\mathbb{Z}) \rightarrow S_{\mu}$, such that if $\varphi(\gamma) = \sigma_{\gamma}$ then $\gamma\alpha_i\Gamma = \alpha_{\sigma_{\gamma}(i)}\Gamma$. In this way every finite-index subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ is associated with a pair of permutations $e = \varphi(E)$ and $v = \varphi(V)$ with $e^2 = v^3 = 1$ which generate a transitive permutation group (transitivity comes from E and V generating $\mathrm{PSL}_2(\mathbb{Z})$). We call (e, v) a **coset permutation representation** of Γ and (l, r) an **LR-representation** of Γ , where $l = \varphi(L)$ and $r = \varphi(R)$. Each form can be obtained from the other form by the equations (2.1). Note that $\gamma \in \mathrm{PSL}_2(\mathbb{Z})$ is in Γ if and only if $\gamma\Gamma = \Gamma$, i.e. $\sigma_{\gamma}(1) = 1$.

On the other hand, suppose e and v are a pair of permutations on μ letters with $e^2 = v^3 = 1$ that generate a transitive permutation group S (such a permutation we call **valid**). Define a homomorphism $\varphi : \mathrm{PSL}_2(\mathbb{Z}) \rightarrow S$ such that $\varphi(E) = e$ and $\varphi(V) = v$ (This is well-defined because of (2.2)). Let $\Gamma = \{\gamma \in \mathrm{PSL}_2(\mathbb{Z}) : \varphi(\gamma)(1) = 1\}$. Then Γ is an index- μ subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. Thus we have a correlation between valid pairs of permutations and finite-index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$. To test if $A \in \mathrm{PSL}_2(\mathbb{Z})$ is in Γ we write A as a word in L and R (This can be done with a variation of the Euclidean Algorithm) and replace L and R with the permutations l and r . If the resulting permutation fixes 1 then A is in Γ .

If one of the cosets is fixed by e , say $e(i) = i$, it corresponds to an elliptic element in Γ , for $E\alpha_i\Gamma = \alpha_i\Gamma$ means $\alpha_i^{-1}E\alpha_i\Gamma = \Gamma$, meaning $\alpha_i^{-1}E\alpha_i$ (which is order 2) is in Γ . So e_2 , the

number of inequivalent elliptic elements of order 2 in Γ , is equal to the number of elements fixed by e . Similarly, e_3 is the number of elements fixed by v .

The cusp width of Γ at ∞ is the smallest positive integer n such that $L^n \in \Gamma$. Thus the cusp width at infinity is the order of the cycle in $\varphi(L)$ which contains “1”. Likewise, suppose i is in a cycle of length k in $\varphi(L)$, i.e. $L^k \alpha_i \Gamma = \alpha_i \Gamma$, but $L^n \alpha_i \Gamma \neq \alpha_i \Gamma$ for $0 < n < k$. Then $\alpha_i^{-1} L^k \alpha_i \in \Gamma$, but $\alpha_i^{-1} L^n \alpha_i \notin \Gamma$ for $0 < n < k$. If $q = \alpha_i^{-1} \infty$ then $\alpha_i^{-1} L^k \alpha_i q = q$ but $\alpha_i^{-1} L^n \alpha_i q \neq q$ for $0 < n < k$. Thus $\alpha_i^{-1} L^k \alpha_i$ is a generator for the stabilizer of the cusp q , and this cusp has width k .

2.6 Algorithms

2.6.1 Calculating a Farey Symbol

Recall that T , the hyperbolic triangle with vertices ρ , ρ^2 , and ∞ , is the standard fundamental domain for $\mathrm{PSL}_2(\mathbb{Z})$, and let T^* be the hyperbolic triangle with vertices ρ , i and ∞ (So $T = T^* \cup (-\overline{T^*})$). $\mathcal{T} = \{\gamma T : \gamma \in \mathrm{PSL}_2(\mathbb{Z})\}$ is a tessellation of the upper half plane and any finite index subgroup Γ has a fundamental domain which is a simply connected union of \mathcal{T} -tiles. Let $\mathcal{T}^* = \{\gamma T^* : \gamma \in \mathrm{PSL}_2(\mathbb{Z})\} \cup \{\gamma(-\overline{T^*}) : \gamma \in \mathrm{PSL}_2(\mathbb{Z})\}$. \mathcal{T}^* is also a tessellation of the upper half plane, and we will construct a fundamental domain for Γ out of \mathcal{T}^* -tiles. The starting point for our construction will be the six tiles around an odd vertex, the image of ρ under any $\gamma \in \mathrm{PSL}_2(\mathbb{Z})$. The following lemma shows this is a reasonable starting point:

Lemma 11. *Let Γ be a subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ with index ≥ 3 . Then at least one of the stabilizer of $\rho = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ and of $\rho - 1 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ is trivial (i.e. one of these points is not elliptic in Γ).*

Proof. If the two stabilizers are not trivial then they must be $\Gamma_\rho = \{I, A, A^2\}$ and $\Gamma_{\rho-1} = \{I, B, B^2\}$ where $A = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$. But A and B generate an index-2 subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. So Γ is either the (unique) index-2 subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ or $\mathrm{PSL}_2(\mathbb{Z})$ itself. And if the index of Γ in $\mathrm{PSL}_2(\mathbb{Z})$ is bigger than 2, at least one of A and B cannot be in Γ . □

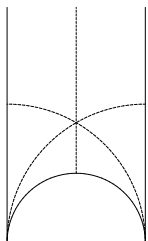


Figure 2.4 A hyperbolic triangle

So if Γ is not $\text{PSL}_2(\mathbb{Z})$ or Γ_2 , the index 2 subgroup of $\text{PSL}_2(\mathbb{Z})$ (cf. [22]), then either the hyperbolic triangle with vertices 0, 1 and ∞ , or the one with vertices -1 , 0 and ∞ (cf. Figure 2.4) is contained in a fundamental domain of Γ . The triangle is made of 6 \mathcal{T}^* -tiles. We will make a polygon P starting with this triangle, then attach \mathcal{T}^* -tiles to P and assign partial pairing information to sides until we get a fundamental domain for Γ (at which point all the pairing information will be filled in). In the algorithm we will say a \mathcal{T}^* -tile T is **adjoinable** to P if T is adjacent to a tile of P and if $P \cup T$ is contained in some fundamental domain of Γ . Note that if T is adjacent to P with adjacency edge e and if e cannot be paired with any other edge of P then T is adjoinable.

Algorithm. 1. If $\Gamma = \text{PSL}_2(\mathbb{Z})$ let P be the special polygon with Farey symbol

$$-\infty \underset{\circ}{\frown} 0 \underset{\bullet}{\smile} \infty$$

or if $\Gamma = \Gamma_2$ let P be the special polygon with Farey symbol

$$-\infty \underset{\bullet}{\frown} 0 \underset{\bullet}{\smile} \infty .$$

In either case return P and terminate.

2. If $\begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$ is not in Γ then let P be the hyperbolic polygon with vertices 0, 1, and ∞ . Otherwise let P be the hyperbolic polygon with edges -1 , 0 and ∞ .
3. If any of the three sides of P map to each other by a $\gamma \in \Gamma$, assign that pairing to the side. (Note that initially all sides are even sides).
4. P is now a polygon where every side is in one of three states:

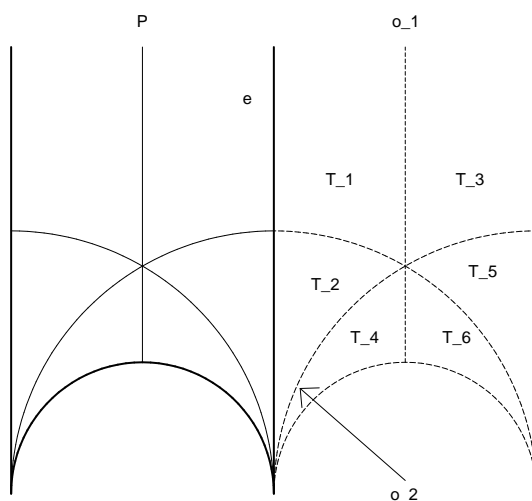


Figure 2.5

- (a) even and already paired.
 - (b) odd and already paired.
 - (c) even and unpaired.
5. Pick an unpaired even side e . Figure 2.5 shows the typical case (The other cases are the same as this case with everything translated by some $\gamma \in \text{PSL}_2(\mathbb{Z})$). Since e is unpaired, T_1 and T_2 must be adjoinable. If o_1 and o_2 are the new odd edges of P after adding T_1 and T_2 to P then either $\gamma o_1 = o_2$ for some $\gamma \in \Gamma$, or there is no such γ . If there is γ pair the two edges and go to Step 3.
 6. If o_1 doesn't pair with o_2 then it doesn't pair with any other side because the only other unpaired sides are odd. So tiles T_3 and likewise T_4 are adjoinable. Each of these tiles has a free edge and the free edges cannot pair with each other (because their common vertex would have an internal angle of $\frac{4\pi}{3}$, and the pairing transformations would make things overlap), so T_5 and T_6 are adjoinable.
 7. We have now added 6 \mathcal{T}^* -tiles to P (One even triangle). If either of the new even edges pair with any of the old unpaired even edges then assign that pairing.
 8. If all the sides of P are paired then we are done. Otherwise go to Step 4.

The output of the algorithm is a special polygon P with $\Gamma_P = \Gamma$. Note that the algorithm must terminate, because a fundamental domain of Γ has hyperbolic area $\frac{\pi}{3}[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma]$ and a single \mathcal{T}^* -tile has area $\frac{\pi}{6}$. So for P to be contained in a fundamental domain of Γ it can have at most $2 \cdot [\mathrm{PSL}_2(\mathbb{Z}) : \Gamma]$ \mathcal{T}^* -tiles.

To effectively implement the algorithm we use Farey symbols. We need only a way to test for group membership. Note that if p_i/q_i and p_{i+1}/q_{i+1} are two adjacent vertices of the fundamental polygon then the hyperbolic triangle added to the edge $[p_i/q_i, p_{i+1}/q_{i+1}]$ in Step 6 is the triangle with vertices p_i/q_i , p_{i+1}/q_{i+1} , and $(p_i + p_{i+1})/(q_i + q_{i+1})$. We can calculate a Farey symbol using the following algorithm:

Algorithm (Calculating a Farey Symbol).

1. If $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ are in Γ then $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, so return

$$-\infty \underset{\circ}{\text{---}} 0 \underset{\bullet}{\text{---}} \infty$$

- and terminate. If $\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$ are in Γ then $\Gamma = \Gamma_2$, so return

$$-\infty \underset{\bullet}{\text{---}} 0 \underset{\bullet}{\text{---}} \infty$$

and terminate.

2. If $\begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} \notin \Gamma$ then let F be the (partial) Farey symbol:

$$-\infty \text{---} \frac{0}{1} \text{---} \frac{1}{1} \text{---} \infty .$$

Otherwise let F be:

$$-\infty \text{---} \frac{-1}{1} \text{---} \frac{0}{1} \text{---} \infty .$$

3. For each i with $0 \leq i \leq n + 1$, if the pairing between x_{i-1} and x_i is not filled in then check if it can be paired with itself (even or odd pairing), or if it can be paired with

another unpaired edge (free pairing). This is done by testing if the appropriate G_i is in Γ . Wherever something can be paired, assign that pairing.

4. If all edges are now paired, return F and terminate.
5. If there is still an unpaired edge, say between p_i/q_i and p_{i+1}/q_{i+1} , make a new vertex $(p_i + p_{i+1})/(q_i + q_{i+1})$ with no pairing information on the edges adjacent to it. Go to Step 3.

The output is a Farey symbol for Γ .

2.6.2 Group Membership

We now describe an algorithm that can test for group membership in the Farey symbol. We will need some lemmas:

Lemma 12. *Distinct even edges with no common vertices do not intersect.*

Proof. Suppose $[A_1, A_2]$ and $[B_1, B_2]$ are even edges with $A_1 < B_1 < A_2 < B_2$ (or if ∞ is a vertex then $A_1 = \infty$ and $B_1 < A_2 < B_2$). There is an element $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma A_1 = \infty$ and $\gamma A_2 = 0$, and since γ is an orientation-preserving homeomorphism on the upper half plane, we have $\gamma B_1 < 0 < \gamma B_2$. But then B_1 and B_2 can't be endpoints of an even edge, because if $B_1 = \frac{x_1}{y_1}$ and $B_2 = \frac{x_2}{y_2}$ with $x_1 \leq -1$ and $x_2 \geq 1, y_1 \geq 1, y_2 \geq 1$, then

$$|x_2 y_1 - x_1 y_2| > 1$$

and so this cannot be an even edge. □

Corollary 13. *Let l be an even line (a hyperbolic line on the upper half plane with rational endpoints x and x' such that $\delta(x, x') = 1$). Let P be a special polygon in \mathbb{H} . Then either $l \subset P$ or $l \cap P = \emptyset$.*

Lemma 14. *If $(a, b) = 1$ then there is $\frac{x'}{y'}$ such that $[\frac{a}{b}, \frac{x'}{y'}]$ is an even edge and the set of even edges with $\frac{a}{b}$ as one endpoint is $[\frac{a}{b}, \frac{x'+an}{y'+bn}]$ for all integers n . Also $(x' + an, y' + bn) = 1$.*

Proof. Let $\frac{x'}{y'}$ be any fraction such that $[\frac{a}{b}, \frac{x'}{y'}]$ is an even edge. Let $A = \begin{bmatrix} a & x' \\ b & y' \end{bmatrix}$. Then $A^{-1}(a/b) = 1/0$, $A^{-1}(x'/y') = 0$, and if $\frac{x}{y}$ is any other vertex where $[\frac{a}{b}, \frac{x}{y}]$ is an even side, then $A^{-1}(x/y) = n$ for some integer n . If $L = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ then

$$AL^{-n}A^{-1}(x/y) = x'/y'$$

We can calculate:

$$AL^{-n}A^{-1} = \begin{bmatrix} 1 - abn & a^2n \\ -b^2n & 1 + abn \end{bmatrix}$$

and

$$x/y = A^{-1}L^nA(x'/y') = \frac{x' + an}{y' + bn}.$$

This fraction is in lowest terms, since if a prime p divides the numerator, $x' \equiv -an \pmod{p}$, and a cannot be $0 \pmod{p}$ since the condition $x'b - y'a = 1$ forces a and x' to be relatively prime. So a is invertible and the same condition gives $bn + y' \equiv a^{-1} \not\equiv 0$, hence p is not a factor of $y' + bn$. A similar argument shows factors of the denominator don't divide the numerator. \square

Lemma 15. Let $[\frac{x_1}{y_1}, \frac{x_2}{y_2}]$ be an even edge and $\frac{a}{b}$ a fraction in lowest terms such that:

$$\frac{x_1}{y_1} < \frac{a}{b} < \frac{x_2}{y_2}.$$

Then $|b| \geq \min\{y_1, y_2\}$

Proof. Suppose $b < y_1$ and $b < y_2$. Using Lemma 14 we can find $\frac{x}{y}$ with $[\frac{a}{b}, \frac{x}{y}]$ an even edge, and $|y| < |b|$. Since even edges don't intersect except possibly at vertices, we must have $\frac{x_1}{y_1} \leq \frac{x}{y} \leq \frac{x_2}{y_2}$. If $\frac{x}{y}$ is not $\frac{x_1}{y_1}$ or $\frac{x_2}{y_2}$ then we rename $\frac{x}{y}$ to $\frac{a}{b}$ and repeat until $\frac{x}{y}$ is an endpoint. If it is the left endpoint, then we have

$$\frac{x_1}{y_1} = \frac{x}{y} < \frac{a}{b} < \frac{x_2}{y_2}$$

where $[\frac{x_1}{y_1}, \frac{a}{b}]$ and $[\frac{x_1}{y_1}, \frac{x_2}{y_2}]$ are even edges, but using Lemma 14 we may show this implies $b > y_2$, contradicting the assumption. The case where $\frac{x}{y}$ is a right endpoint is similar. \square

The following algorithm is adapted from [16]. Given a group Γ with Farey symbol F and special polygon P , and an even edge $[c'_0/d'_0, c_0/d_0]$, the algorithm returns a sequence of generators from the Farey symbol which transform both vertices of the even edge to vertices of P . This allows us to test if an element of $\text{PSL}_2(\mathbb{Z})$ is also an element of Γ .

Let Γ be a finite index subgroup of $\text{PSL}_2(\mathbb{Z})$ and A an element of $\text{PSL}_2(\mathbb{Z})$:

$$A = \begin{bmatrix} c_0 & c'_0 \\ d_0 & d'_0 \end{bmatrix}.$$

A maps the even line $[0, \infty]$ to $l = [c'_0/d'_0, c_0/d_0]$. By Corollary 13, either $l \subset P$ or it is disjoint from P (except possibly at endpoints). If it is disjoint there is an edge (x_i, x_{i+1}) of P which l “lies under”, i.e. $x_i \leq c'_0/d'_0 < c_0/d_0 \leq x_{i+1}$ with one of the endpoints possibly ∞ or $-\infty$. The idea of the algorithm is to repeatedly translate P across this overreaching edge until P intersects $[c'/d', c/d]$, at which point A will be in Γ if and only if l is the image of $(0, \infty)$ or an edge paired with $(0, \infty)$. In the actual algorithm we work in the other direction, translating the even line instead of the special polygon.

Algorithm (LLT). Let $k = 0$ and F be a Farey symbol for Γ with 0 as one of its vertices. Without loss of generality, we can assume $\frac{c'_k}{d'_k} < \frac{c_k}{d_k}$.

1. There are two possibilities: If $\frac{c'_k}{d'_k}$ and $\frac{c_k}{d_k}$ are both vertices of P then terminate. Otherwise we must have $x_i \leq \frac{c'_k}{d'_k} < \frac{c_k}{d_k} \leq x_{i+1}$ with at least one “ \leq ” a strict inequality.
2. Let G_{i+1} be the generator corresponding to the pairing p_{i+1} (recall this is the transformation mapping $[x_i, x_{i+1}]$ to its paired side). If p_{i+1} is a free or even pairing, let $\alpha_k = G_{i+1}$. If p_{i+1} is an odd pairing, let $m = \frac{a_i + a_{i+1}}{b_i + b_{i+1}}$ where $x_i = \frac{a_i}{b_i}$, $x_{i+1} = \frac{b_{i+1}}{b_{i+1}}$. Then the interval $(\frac{c'_k}{d'_k}, \frac{c_k}{d_k})$ must be between either x_i and m or between m and x_{i+1} . If $\frac{c_k}{d_k} \leq m$, let $\alpha_k = G_{i+1}$. Otherwise let $\alpha_k = G_{i+1}^{-1}$.
3. Let $\frac{c_{k+1}}{d_{k+1}} = \alpha_k \cdot \frac{c_k}{d_k}$, $\frac{c'_{k+1}}{d'_{k+1}} = \alpha_k \cdot \frac{c'_k}{d'_k}$. Replace k with $k + 1$ and go to Step 1.

Theorem 16.

1. The algorithm terminates, c'_k/d'_k and c_k/d_k are vertices of P , and the α_i 's satisfy:

$$\frac{c_k}{d_k} = \alpha_{k-1} \cdots \alpha_0 \left(\frac{c_0}{d_0} \right) \quad \text{and} \quad \frac{c'_k}{d'_k} = \alpha_{k-1} \cdots \alpha_0 \left(\frac{c'_0}{d'_0} \right).$$

2. A is in Γ if and only if $\begin{bmatrix} c_k & c'_k \\ d_k & d'_k \end{bmatrix}$ is one of the generators of F , i.e. one of the following is true:

$$(a) \quad \begin{bmatrix} c_k & c'_k \\ d_k & d'_k \end{bmatrix} = \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b) $(\frac{c'_k}{d'_k}, \frac{c_k}{d_k})$ is a free side paired with $(0, \infty)$.

(c) $\begin{bmatrix} c_k & c'_k \\ d_k & d'_k \end{bmatrix} = \pm \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and 0 and ∞ are adjacent vertices with an even pairing between them.

3. If A is in Γ then $A = \alpha_0^{-1} \alpha_1^{-1} \cdots \alpha_{k-1}^{-1} \begin{bmatrix} c_k & c'_k \\ d_k & d'_k \end{bmatrix}$, i.e. the algorithm decomposes A as a word in the generators of Γ .

Proof. To see the algorithm terminates, consider the sequence:

$$P_0 = P, \quad P_k = \alpha_0^{-1} \cdots \alpha_{k-1}^{-1} P$$

This is a sequence of adjacent hyperbolic polygons such that for each k , P_k “lies under” an edge of P_{k-1} (in the sense described above) and the even line $[c'_0/d'_0, c_0/d_0]$ lies under an edge of P_k . Thus the sequence of polygons closes in on the starting edge in a sense. But by Lemma 15, if $d = \max\{d_0, d'_0\}$ then any even edge lying over $[c'_0/d'_0, c_0/d_0]$ must have endpoints with denominators smaller than d . So if (x_i, x_{i+1}) is an edge of P_0 with $[c'_0/d'_0, c_0/d_0]$ lying under it, there are only finitely many rationals with small enough denominators under (x_i, x_{i+1}) and only finitely many translations of P lying below P and above $[c'_0/d'_0, c_0/d_0]$. Thus the algorithm must terminate.

Part 2 is proven in [16] and Part 3 is immediate. \square

Later we will need an adaption of this algorithm which takes a single element x of $\mathbb{P}^1(\mathbb{Q})$ as input and returns a list of generators mapping x to a vertex on P . This can be done with only minor alterations to the algorithm.

2.6.3 Coset Representatives

Let Γ be a group with special polygon P . Let T be the hyperbolic triangle with vertices i , ρ , and ∞ . By the construction of P , T is contained in P . The set of $\gamma \in \Gamma$ such that γT is in P is a set of coset representatives of Γ .

Let a_i/b_i and a_{i+1}/b_{i+1} be adjacent vertices of P , and let $L = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\varphi = \begin{bmatrix} a_i & a_{i+1} \\ b_i & b_{i+1} \end{bmatrix}$. Then $\varphi^{-1}(a_i/b_i) = \infty$ and $\varphi^{-1}(a_{i+1}/b_{i+1}) = 0$. Let w_i be $|a_{i-1}b_{i+1} - a_{i+1}b_{i-1}|$ if the pairing between a_i/b_i and a_{i+1}/b_{i+1} is not an odd pairing and $|a_{i-1}b_{i+1} - a_{i+1}b_{i-1}| + 1$ if it is. Then w_i is the number of \mathcal{T}^* -tiles of the form γT in P . Thus a list of left coset representatives for Γ is $\bigcup_{i=0}^n \{L^{-j}\varphi_i^{-1} : 0 \leq j < w_i\}$.

2.6.4 Congruence Testing

Let Γ be a finite index subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. Lang, Lim and Tan give a test purely in terms of Farey symbols to determine if Γ is a congruence group [16]. Their test relies on Wohlfahrt's Theorem [25] which says that if Γ has level N then Γ is a congruence group if and only if Γ contains $\Gamma(N)$. In Lang, Lim and Tan's test, if Γ has level N one computes a Farey symbol for $\Gamma(N)$, giving a complete set of generators for $\Gamma(N)$. One then checks if each of these generators is contained in Γ using the above algorithm. The difficulty with this algorithm is that the index of $\Gamma(N)$ increases very quickly with N , so if Γ has large level, the calculation of a Farey symbol for $\Gamma(N)$ can be very lengthy, even if Γ has relatively small index.

Another test for congruence was developed by Tim Hsu using Millington's coset permutation representations [8]. If we have an LR-representation of Γ then there is a list of relations that are satisfied if and only if Γ is congruence.

To calculate an LR-representation from a Farey symbol, use the above algorithm to calculate a list of left coset representatives $\alpha_i \in \mathrm{PSL}_2(\mathbb{Z})$ where $\mathrm{PSL}_2(\mathbb{Z}) = \bigcup_{i=1}^{\mu} \alpha_i \Gamma$. To calculate l ,

for instance, recall that l is the permutation such that $L\alpha_i\Gamma = \alpha_{l(i)}\Gamma$. So l sends i to the unique j such that $\alpha_j^{-1}L\alpha_i \in \Gamma$. So we run through every $1 \leq i \leq \mu$ and calculate the permutation. r can be calculated similarly. (Actually, although we need l and r it is easier to calculate e and v , because we know beforehand that they are order 2 and 3 respectively. Then $l = ev^{-1}$ and $r = ev^{-2}$).

Knowing l and r we can directly apply Tim Hsu's congruence algorithm [8]. Depending on the order of l , (i.e. the level of Γ) there are different lists of relations of l and r that are satisfied if and only if Γ is congruence. For example, if N is the order of l and N is odd then Γ is a congruence group if and only if $r^2l^{-\frac{1}{2}}$ is the identity permutation (where $\frac{1}{2}$ is the inverse of 2 modulo N).

2.6.5 Modular Symbols

2.6.5.1 Modular Symbols

We follow the notation of [24]. Recall that $\mathbb{P}^1(\mathbb{Q})$ denotes $\mathbb{Q} \cup \{\infty\}$ and let \mathbb{M}_2 be the free abelian group generated by symbols $\{x, y\}$ with $x, y \in \mathbb{P}^1(\mathbb{Q})$, modulo the relations

$$\begin{aligned} \{x, x\} &= 0 \\ \{x, y\} + \{y, z\} + \{z, x\} &= 0. \end{aligned} \tag{2.3}$$

Combining the relations we conclude:

$$\{x, y\} = -\{y, x\}.$$

The action of $\mathrm{PSL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{Q})$ extends to an action on \mathbb{M}_2 by

$$\gamma\{x, y\} = \{\gamma x, \gamma y\}$$

which is easily seen to be compatible with the relations in (2.3). If Γ is a finite index subgroup of $\mathrm{PSL}_2(\mathbb{Z})$, we define the space of (weight 2) modular symbols $\mathbb{M}_2(\Gamma)$ to be \mathbb{M}_2 modulo the relations:

$$\{x, y\} = \gamma\{x, y\}$$

for all $\gamma \in \Gamma$ and $x, y \in \mathbb{P}^1(\mathbb{Q})$. We think of these being elements of the first homology group of the modular curve X_Γ relative to the cusps:

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

Each symbol $\{x, y\}$ is thought of as a path on the modular curve from x to y . To talk about the regular homology group $H_1(X_\Gamma, \mathbb{Z})$ we want only paths with “boundary zero”. Thus we let $\mathbb{B}_2(\Gamma)$ be the free abelian group generated by the cusps of Γ and define a “boundary map”, $\delta : \mathbb{M}_2(\Gamma) \rightarrow \mathbb{B}_2(\Gamma)$ by

$$\delta(\{x, y\}) = \{y\} - \{x\}.$$

Then the subspace of *weight 2 cuspidal modular symbols* $\mathbb{S}_2(\Gamma)$ is the kernel of δ . As the elements of $\mathbb{S}_2(\Gamma)$ correspond to combinations of paths with no boundary, we can get an isomorphism with the first homology group ([19]):

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

2.6.5.2 Pairing

The reason we are interested in the space of cuspidal modular symbols is that they are dual in a sense to the space of weight 2 cusp forms $S_2(\Gamma)$ of Γ . If $f(z)$ is a weight 2 cusp form and c is an element of $H_1(X_\Gamma, \mathbb{Z})$, we define a pairing $\langle \cdot, \cdot \rangle$ by

$$\langle c, f \rangle = \int_c f(z) dz.$$

Suppose that the genus of Γ (i.e. the genus of its modular curve, X_Γ) is g . Then the space of weight 2 cusp forms, $S_2(\Gamma)$, is a g -dimensional vector space over \mathbb{C} , say with basis f_1, \dots, f_g . On the other hand, $H_1(X_\Gamma, \mathbb{Z})$ is a free rank- $2g$ \mathbb{Z} -module, say generated by c_1, \dots, c_{2g} . We can construct a $2g \times g$ matrix with complex entries $\omega_{ij} = \langle c_i, f_j \rangle$. By separating real and imaginary parts, this matrix can be thought of as a $2g \times 2g$ real valued matrix, and standard Riemann surface theory shows it is nonsingular ([5], Chapter II).

2.6.5.3 Relation to cusp forms and Hecke operators

The vector space of weight k cusp forms has an important family of linear operators called the Hecke operators. The pairing allows us to define an action of the Hecke operators on the cuspidal modular symbols $\mathbb{S}_2(\Gamma)$ which is adjoint to the action on the cusp forms, $S_2(\Gamma)$. Thus if we have an efficient way to compute cuspidal modular symbols, we can compute their Hecke operators and use them to compute spaces of cusp forms for congruence groups.

Since the action of Hecke operators on noncongruence groups is trivial (in a certain sense), for the remainder of this section let Γ denote a congruence group.

Recall how the Hecke operators are defined for congruence groups (See Chapter 3 of [23] or Chapter 5 of [6] for details): Let α be an element of $\mathrm{GL}_2^+(\mathbb{Q})$. Then the double coset $\Gamma\alpha\Gamma$ decomposes into a finite union of right cosets ([6] Lemma 5.1.1 and 5.1.2):

$$\Gamma\alpha\Gamma = \bigcup_{j=1}^n \Gamma\beta_j.$$

(The reason we require Γ to be congruence is to ensure that this is a *finite* union.) If $f(z)$ is a weight k modular form for Γ then we define an action of $\Gamma\alpha\Gamma$ on $f(z)$ by:

$$f(z)[\Gamma\alpha\Gamma]_k = \sum_{j=1}^n f(z)[\beta_j]_k. \quad (2.4)$$

The action defines a linear operator:

$$[\Gamma\alpha\Gamma]_k : M_k(\Gamma) \rightarrow M_k(\Gamma).$$

For all positive integers a and d , define an operator:

$$T(a, d) = [\Gamma\alpha\Gamma]_k \quad \text{where } \alpha = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}. \quad (2.5)$$

Then the n th Hecke operator for Γ is defined to be:

$$T_n = \sum_{ad=n, a|d} T(a, d).$$

Each element of this sum is a double coset operator, so grouping together all the β_j 's in (2.4) from the different terms in the sum (2.5) we get a finite list $B_n \subset \mathrm{GL}_2^+(\mathbb{Q})$ such that

$$T_n(f(z)) = \sum_{\beta \in B_n} f(z)[\beta]_k.$$

Calculating B_n is nontrivial. In the classical example where p is prime and $\Gamma = \Gamma_0(N)$, if p divides N then B_p can be chosen to be:

$$B_p = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & p \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & p \end{bmatrix}, \dots, \begin{bmatrix} 1 & p-1 \\ 0 & p \end{bmatrix} \right\}$$

and if p doesn't divide N then B_p can be chosen to be:

$$B_p = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & p \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & p \end{bmatrix}, \dots, \begin{bmatrix} 1 & p-1 \\ 0 & p \end{bmatrix}, \begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

So the prime Hecke operators for $\Gamma_0(N)$ are:

$$T_p f(z) = \begin{cases} \sum_{j=0}^{p-1} f(z) \begin{bmatrix} 1 & j \\ 0 & p \end{bmatrix}_k & \text{if } p|N \\ \sum_{j=0}^{p-1} f(z) \begin{bmatrix} 1 & j \\ 0 & p \end{bmatrix}_k + f(z) \begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix} & \text{if } p \nmid N \end{cases}$$

Our work can be greatly simplified by the fact that if n and m are relatively prime then:

$$T_{nm} = T_n T_m$$

and for p prime and $k \geq 2$:

$$T_{p^k} = T_p T_{p^{k-1}} - p T(p, p) T_{p^{k-2}}$$

(cf. [23] Theorem 3.24).

We now define an action of the Hecke operators on modular symbols: If x and y are in $\mathbb{P}^1(\mathbb{Q})$ then define:

$$T_n(\{x, y\}) = \sum_{\beta \in B_n} \beta\{x, y\}$$

and extend by linearity to an operator on all of $S_2(\Gamma)$. The Hecke operators on cusp forms are adjoint to those on modular symbols with respect to the pairing defined above. First, we prove a lemma:

Lemma 17. *For any $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$, $c \in H_1(X_\Gamma, \mathbb{Z})$, and $f(z) \in S_2(\Gamma)$:*

$$\langle \gamma c, f \rangle = \langle c, f[\gamma]_2 \rangle$$

Proof. Note that $\frac{d}{dz}(\gamma z) = (cz + d)^{-2}$. Then, using the change of variables $w = \gamma z$:

$$\begin{aligned}\langle \gamma c, f \rangle &= \int_{\gamma c} f(w) dw = \int_c f(\gamma z) d(\gamma z) \\ &= \int_c f(\gamma z) (cz + d)^{-2} dz = \int_c f(z) [\gamma]_2 dz = \langle c, f[\gamma]_2 \rangle.\end{aligned}$$

□

Using this lemma we conclude:

Theorem 18. *Let $f(z)$ be in $S_2(\Gamma)$ and c be in $\mathbb{S}_2(\Gamma)$. Then:*

$$\langle T_n(c), f \rangle = \langle c, T_n(f) \rangle.$$

Proof. Since $\mathbb{S}_2(\Gamma)$ has a basis of modular symbols we can write:

$$c = \sum_{j=1}^N \{x_j, y_j\}.$$

Then:

$$\begin{aligned}\langle T_n(c), f \rangle &= \sum_{j=1}^N \langle T_n(\{x_j, y_j\}), f \rangle = \sum_{j=1}^N \sum_{\beta \in B_n} \langle \beta \{x_j, y_j\}, f \rangle \\ &= \sum_{j=1}^N \sum_{\beta \in B_n} \langle \{x_j, y_j\}, f[\beta]_2 \rangle \\ &= \langle \sum_{j=1}^N \{x_j, y_j\}, \sum_{\beta \in B_n} f[\beta]_2 \rangle \\ &= \langle c, T_n(f) \rangle.\end{aligned}$$

□

2.6.5.4 Modular Symbols and Farey Symbols

Suppose F is a Farey symbol for the group Γ with vertex sequence $\{-\infty, x_0, \dots, x_n, \infty\}$.

Then the set of free pair edges:

$$\{\{x_{i-1}, x_i\} : [x_{i-1}, x_i] \text{ is an edge in a free pairing}\}$$

generates the weight 2 modular symbols of Γ , $\mathbb{M}_2(\Gamma)$. If only one edge is taken from each free pair, we have a basis. This is because the free edges are linearly independent, and there are

$2g + t - 1$ of them (where t is the number of cusps and g is the genus). But by (2.6.5.1), the rank of $\mathbb{M}_2(\Gamma)$ is the rank of $H_1(X_\Gamma, \{\text{cusps}\}, \mathbb{Z})$, which is also $2g + t - 1$. Call this basis $E = \{e_1, \dots, e_n\}$.

Now that we have a basis for $\mathbb{M}_2(\Gamma)$ we need an algorithm to write a general modular symbol in terms of this basis. The main tool for our algorithm will be the algorithm discussed above in Section 2.6.2 where if F is a Farey symbol for Γ , $\{\gamma_1, \dots, \gamma_m\}$ is the set of generators from F , and $a \in \mathbb{P}^1(\mathbb{Q})$ then the algorithm returns a list $\{\gamma_{i_1}, \dots, \gamma_{i_n}\}$ of the generators such that $\gamma_{i_n}^{-1} \cdots \gamma_{i_1}^{-1} a$ is a vertex of F .

Let $\{x, y\}$ be a modular symbol for Γ . We may assume x is a vertex of F , since we can use our algorithm to move it there (Recall that $\gamma\{a, b\} = \{a, b\}$ for every $\gamma \in \Gamma$). Now we apply the algorithm to y and get a list $\{\gamma_{i_1}, \dots, \gamma_{i_n}\}$ of generators in F such that $\gamma_{i_n}^{-1} \cdots \gamma_{i_1}^{-1} y$ is a vertex of F , say $y' = \gamma_{i_n}^{-1} \cdots \gamma_{i_1}^{-1} y$.

Next we use the facts that

$$\{a, b\} = \{a, c\} + \{c, b\}$$

for any $a, b, c \in \mathbb{P}^1(\mathbb{Q})$, and

$$\{a, \gamma a\} = \{b, \gamma b\}$$

for any $a, b \in \mathbb{P}^1(\mathbb{Q})$ and $\gamma \in \text{SL}_2(\mathbb{Z})$ to get:

$$\begin{aligned} \{x, y\} &= \{x, \gamma_{i_1} \cdots \gamma_{i_n} y'\} \\ &= \{x, \gamma_{i_2} \cdots \gamma_{i_n} y'\} + \{\gamma_{i_2} \cdots \gamma_{i_n} y', \gamma_{i_1} \gamma_{i_2} \cdots \gamma_{i_n} y'\} \\ &= \{x, \gamma_{i_2} \cdots \gamma_{i_n} y'\} + \{y', \gamma_{i_1} y'\} \\ &= \dots \\ &= \{x, y'\} + \sum_{j=1}^n \{y', \gamma_{j_i} y'\}. \end{aligned}$$

Modular symbols of the form $\{a, \gamma_i a\}$ can be easily written in our basis. Since this symbol is independent of the choice of a , choose a to be x_i , a vertex on the edge corresponding to γ_i .

Then $\gamma_i x_i$ is another vertex of F (say x_N), so

$$\{x_i, \gamma_i x_i\} = \{x_i, x_N\} = \begin{cases} \sum_{j=i}^{N-1} \{x_j, x_{j+1}\} & \text{if } i < N \\ -\sum_{j=N}^{i-1} \{x_j, x_{j+1}\} & \text{if } i > N \\ 0 & \text{if } i = N \end{cases}$$

Also x and y' are vertices of F so they can be written in terms of the basis in a similar fashion.

Finally, we count how many times each basis element occurs and return that list of numbers.

To summarize:

Theorem 19. *The algorithm described above takes a modular symbol $\{x, y\}$ as input and returns a sequence of numbers a_1, \dots, a_n such that:*

$$\{x, y\} = \sum_{j=1}^n a_j e_j$$

where $\{e_j\}$ is the basis for $\mathbb{M}_2(\Gamma)$ from the Farey symbol.

We can now compute the matrix for the n th Hecke operator on $\mathbb{M}_2(\Gamma)$ in the following way: For a basis element e_i , use the algorithm to reduce βe_i to a linear combination in elements of E for each $\beta \in B_n$. Add up the coefficients to get the expression of $T_n(e_i)$ in the E basis. Repeat for every basis element and put it all together to get a matrix U_n representing T_n .

2.6.5.5 Cusp Forms

If $f(z)$ is a power series in $q = e^{2\pi iz}$, define a linear operator that returns the n th coefficient of f :

$$a_n(f(z)) = a_n \left(\sum_{k=0}^{\infty} c_k q^k \right) = c_n.$$

Then for some groups Γ (at least $\Gamma_1(N)$ and $\Gamma_0(N)$) we have for all n :

$$a_1(T_n(f)) = a_n(f).$$

Again following the notation of [24], let \mathbb{T}' be the ring generated by Hecke operators acting on $S_2(\Gamma)$ and let $\mathbb{T}'_{\mathbb{C}} = \mathbb{T}' \otimes \mathbb{C}$. It can be shown that there is a perfect bilinear pairing of finite

dimensional vector space over \mathbb{C} :

$$\begin{aligned} S_2(\Gamma) \times \mathbb{T}'_{\mathbb{C}} &\rightarrow \mathbb{C} \\ \langle f, t \rangle &\mapsto a_1(t(f)). \end{aligned}$$

Hence there is an isomorphism $\Psi : S_2(\Gamma) \rightarrow \text{Hom}(\mathbb{T}'_{\mathbb{C}}, \mathbb{C})$ mapping a cusp form f to the homomorphism $t \mapsto a_1(t(f))$ where $t \in \mathbb{T}'_{\mathbb{C}}$. Thus if $\varphi \in \text{Hom}(\mathbb{T}'_{\mathbb{C}}, \mathbb{C})$, we get a cusp form $\Psi^{-1}(\varphi)$, which is defined by the condition that $\langle \Psi^{-1}(\varphi), t \rangle = \varphi(t)$ for all $t \in \mathbb{T}'_{\mathbb{C}}$. But this cusp form can be expressed another way:

$$f_{\varphi} = \sum_{k=1}^{\infty} \varphi(T_k) q^k.$$

I claim f_{φ} and $\Psi^{-1}(\varphi)$ are the same, for note that for every n :

$$\langle f_{\varphi}, T_n \rangle = a_1(T_n(f_{\varphi})) = a_n(f_{\varphi}) = \varphi(T_n).$$

So $a_n(f_{\varphi} - g) = \langle f_{\varphi} - g, T_n \rangle = 0$ for all n , so $f_{\varphi}(z) = g(z)$. Thus we have a way of calculating cusp forms from the matrices U_n of the Hecke operators on modular symbols:

Theorem 20. *Fix a basis for $S_2(\Gamma)$ and let U_n be the matrix corresponding to T_n in this basis.*

Let $a_{ij}(U_n)$ be the ij -th entry of U_n . Then the series:

$$f_{ij}(z) = \sum_{n=1}^{\infty} a_{ij}(U_n) q^n$$

is a cusp form for Γ , and the set of all such f_{ij} 's spans the space $S_2(\Gamma)$.

Proof. $\mathbb{T}'_{\mathbb{C}}$ is spanned by T_n , so the function φ defined by setting $\varphi(T_n) = a_{ij}(U_n)$ and extending linearly is on $\text{Hom}(\mathbb{T}'_{\mathbb{C}}, \mathbb{C})$ and as we showed above, $f_{ij}(z) = f_{\varphi}(z) = \Psi^{-1}(\varphi)$ which is on $S_2(\Gamma)$. □

Example 21. Let $\Gamma = \Gamma_0(11)$, which is genus 1. We use the algorithm described above to compute U_n for several values of n :

$$U_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad U_2 = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \quad U_3 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad U_4 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$U_5 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad U_6 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad U_7 = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \quad U_8 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

This sequence gives rise to one cusp form:

$$f_{11} = f_{22} = q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 + \dots$$

This spans the space of weight 2 cusp forms of Γ since $\dim S_2(\Gamma) = \text{genus}(\Gamma)$.

Example 22. Let $\Gamma = \Gamma_0(39)$, which is genus 3. Again we can use the algorithm to compute U_n for n up to 6:

$$U_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad U_2 = \begin{bmatrix} 0 & -1 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & -1 & 2 & -1 \\ 1 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$U_3 = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 1 & 0 \\ -1 & 1 & 0 & -1 & 1 & 0 \end{bmatrix} \quad U_4 = \begin{bmatrix} 0 & 1 & -1 & -2 & 2 & 0 \\ 0 & -1 & 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -3 & 1 \\ -1 & -3 & 1 & 1 & 0 & -2 \\ 1 & -1 & -1 & 0 & 1 & -2 \\ 1 & 1 & 0 & -1 & -1 & 0 \end{bmatrix}$$

$$U_5 = \begin{bmatrix} 1 & -1 & 1 & -2 & 2 & 0 \\ 0 & -2 & 3 & -3 & 1 & -1 \\ 1 & -3 & 4 & -1 & -1 & -1 \\ 1 & -5 & 3 & 0 & 0 & -2 \\ 3 & -3 & 1 & 0 & 0 & -2 \\ 3 & -1 & 0 & 1 & -3 & 1 \end{bmatrix} \quad U_6 = \begin{bmatrix} -1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 2 & -1 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Since $\text{genus}(\Gamma) = 3$ the dimension of $S_2(\Gamma_0(39))$ is also 3 and we just need to pick out three linearly independent cusp forms:

$$\begin{aligned}f_{11} &= q + q^5 - q^6 + \dots \\f_{12} &= -q^2 + q^3 + q^4 - q^5 + \dots \\f_{14} &= q^2 + -2q^4 - 2q^5 + q^6 + \dots\end{aligned}$$

For example in Example 3.1 of [24], Stein shows that

$$f = q + q^2 - q^3 - q^4 + 2q^5 + \dots$$

is a cusp form for $\Gamma_0(39)$. In our notation this is $f = f_{11} - f_{12}$.

CHAPTER 3. Some modular forms for noncongruence groups

3.1 Introduction

3.1.1 Modular Forms for Congruence Groups

Now that we know better how to work with finite-index subgroups of the modular group, let us consider their modular forms. Let Γ be a finite-index subgroup of $\mathrm{SL}_2(\mathbb{Z})$, and recall that a weight k **modular form** of Γ is a function $f(z)$, holomorphic on the upper half plane, satisfying:

$$f(\gamma z) = (cz + d)^k f(z) \quad \text{for every } \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma \quad (3.1)$$

and holomorphic at the cusps (locally as a function on the modular curve X_Γ). We denote the space of weight k modular forms of Γ by $M_k(\Gamma)$. As usual, we define an operator, $[\gamma]_k$, (or just $[\gamma]$ if the weight is clear) by:

$$f[\gamma]_k(z) = (cz + d)^{-k} f(\gamma z)$$

in which case (3.1) becomes:

$$f[\gamma]_k(z) = f(z)$$

for every $\gamma \in \Gamma$. We also define **modular functions** for Γ to be weight 0 functions $f(z)$ satisfying (3.1) but only meromorphic on the upper half plane and at the cusps. The modular functions of Γ form a field which we denote by $\mathbb{C}(\Gamma)$.

Since every finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ contains an element of the form

$$L^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$$

for some n , every modular form is periodic and has a Fourier series in $q_n = e^{2\pi iz/n}$. The coefficients of these Fourier series are often interesting from a number theoretic point of view. In fact, in the next theorem we will show that if Γ is a congruence group, then any of its modular forms with algebraic Fourier coefficients can be multiplied by a constant to get a modular form with coefficients in $\mathcal{O}_{\overline{\mathbb{Q}}}$, the ring of algebraic integers. In this case we say the modular form has **bounded denominators** because if the coefficients are rational then the sequence of denominators of the coefficients is bounded.

Theorem 23. *Let Γ be a congruence group of level N and let $g(z)$ be in $M_k(\Gamma)$ with Fourier coefficients in $\overline{\mathbb{Q}}$. Then there is a nonzero algebraic integer C such that $Cg(z)$ has algebraic integer Fourier coefficients.*

Proof. Suppose Γ is a congruence group of level N , i.e. $\Gamma \supseteq \Gamma(N)$, and $M_k(\Gamma(N)) \supseteq M_k(\Gamma)$. It is well known (see [24] p.172) that $M_k(\Gamma(N))$ has a basis of modular forms with algebraic integer Fourier coefficients, say $\{f_1, \dots, f_n\}$ with

$$f_i = \sum_{j=1}^{\infty} a_{j,i} q^j \quad a_i \in \mathcal{O}_{\overline{\mathbb{Q}}}$$

For each j let v_j be the row vector of q^j -coefficients of the basis elements: $v_j = \begin{bmatrix} a_{j,1} & \dots & a_{j,n} \end{bmatrix}$. By the linear independence of the f_i 's, there is a sequence of distinct integers j_1, \dots, j_n such that $\{v_{j_1}, \dots, v_{j_n}\}$ is linearly independent. Let A be the matrix with rows v_j . A is invertible in $\overline{\mathbb{Q}}$, i.e. $A^{-1} \in \overline{\mathbb{Q}}^{n \times n}$.

Suppose $g(z) = \sum_{j=1}^{\infty} b_j q^j$ is in $M_k(\Gamma)$ and the coefficients b_j are in $\overline{\mathbb{Q}}$. Then there are $c_i \in \mathbb{C}$ such that $g = \sum_{i=1}^n c_i f_i$. Write column vectors:

$$\beta = \begin{bmatrix} b_{j_1} \\ \vdots \\ b_{j_n} \end{bmatrix} \quad \gamma = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Then $\gamma = A^{-1}\beta$, i.e. each c_i is in $\overline{\mathbb{Q}}$.

Let $C \in \mathcal{O}_{\overline{\mathbb{Q}}}$ be such that $Cc_i \in \mathcal{O}_{\overline{\mathbb{Q}}}$ for each i . Then $Cg = \sum_{i=1}^n Cc_i f_i$ has algebraic integer Fourier coefficients. □

We note that this result applies to modular forms, but not generally to modular functions. It does, however, apply to modular functions whose poles occur only at the cusps:

Corollary 24. *Let Γ be a congruence group. If $f(z)$ is a modular function for Γ with algebraic Fourier coefficients, holomorphic on the upper half plane and possibly with poles at the cusps, then there is an algebraic integer C such that $Cg(z)$ has algebraic integer Fourier coefficients.*

Proof. Let $\Delta(z) = \eta^{24}(z)$ where η is the Dedekind eta-function:

$$\eta(z) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n)$$

where $q = e^{2\pi iz}$. Then $\Delta(z)$ is a weight 12 cusp form for $\mathrm{SL}_2(\mathbb{Z})$ with Fourier series:

$$\Delta(z) = q - 24q^2 + 252q^3 + \dots$$

Specifically, $\Delta(z)$ has integer Fourier coefficients, and so does its reciprocal, since:

$$\frac{1}{\Delta} = \frac{1}{q} \cdot \frac{1}{1 + (\Delta/q - 1)} = \frac{1}{q} \left(1 - \left(\frac{\Delta}{q} - 1 \right) + \left(\frac{\Delta}{q} - 1 \right)^2 - \dots \right)$$

Multiplying $f(z)$ by a power of Δ will eliminate any poles at the cusps, i.e. there is n such that $\Delta^n(z)g(z)$ is a modular form for Γ . Hence by the previous theorem there is $C \in \mathcal{O}_{\overline{\mathbb{Q}}}$ such that $C\Delta^n(z)g(z)$ has coefficients in $\mathcal{O}_{\overline{\mathbb{Q}}}$. Then $Cg(z) = (\Delta^{-1}(z))^n C\Delta^n(z)g(z)$ has coefficients in $\mathcal{O}_{\overline{\mathbb{Q}}}$. \square

3.1.2 Modular Forms for Noncongruence Groups

If $f(z)$ is a modular form for a noncongruence group Γ and not a modular form for any larger congruence group containing Γ we say $f(z)$ is a noncongruence modular form.

The first serious study of noncongruence modular forms was Atkin and Swinnerton-Dyer's 1971 paper, "Modular Forms on Noncongruence Groups" [1]. They were interested in the Fourier coefficients of noncongruence modular forms, and they point out that noncongruence modular forms may have unbounded denominators by considering the modular function $\zeta(z) = \eta(z)/\eta(13z)$, where η is the Dedekind eta function. The function $\zeta(z)$ is a generator for the modular functions of a certain congruence group Γ , and its only poles are at the cusps. Its

n th root $\zeta^{1/n}(z)$ is also the generator for the modular functions of a subgroup of Γ , but its Fourier coefficients have unbounded denominators, hence the subgroup must be noncongruence by Corollary 24.

It is conjectured that unbounded denominators are typical of noncongruence groups. Specifically, we say:

Definition 25. A finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$, Γ , is said to satisfy the **unbounded denominator property** (UBD) if every holomorphic integral weight modular form for Γ (which is not a modular form for any larger group containing Γ) with algebraic Fourier coefficients has unbounded denominators.

Then the conjecture can be stated as:

Conjecture 26. *Every noncongruence group satisfies the unbounded denominator property.*

This conjecture seems to be attributed to Atkin and Swinnerton-Dyer (for example in Coste and Gannon [4]), although it does not directly appear in their paper [1]. The earliest written version of this conjecture the author is aware of is in Birch [3], who merely says “It has been conjectured that the coefficients a_n turn out to be algebraic integers if and only if G is a congruence subgroup”.

3.2 Character Groups

The example of Atkin and Swinnerton-Dyer above suggests a general way to construct modular functions with unbounded denominators: taking n th roots of known modular functions with zeros and poles at the cusps. Taking the n th root introduces powers of n to the denominators of the Fourier coefficients, leading often to functions with unbounded denominators. The n th root is not generally a modular function for the original group, but for a subgroup, which must then be noncongruence if the n th root has unbounded denominators. Such subgroups are examples of what we will call character groups.

The unbounded denominator problem for character groups was previously studied by the author and Ling Long in [13] and [14] where the character groups are specialized further into

“Types” (see Definition 28). In [13] it is shown that if there is a prime number p such that every index- p Type II(A) character group of Γ^0 satisfies (UBD) then there is a positive constant c such that for large x :

$$\#\{\text{Type II(A) character groups } \Gamma \text{ of } \Gamma^0: [\Gamma^0 : \Gamma] < x \text{ and } \Gamma \text{ satisfies (UBD)}\} > cx^2$$

And it is shown that the required condition is true for $\Gamma^0(11)$. In [14] it is shown that if M is a positive squarefree integer then every noncongruence Type I(A) character group of $\Gamma_0(M)$ satisfies (UBD).

Let Γ^0 be a finite index congruence subgroup of $\text{SL}_2(\mathbb{Z})$ and let $f(z)$ be a modular function for Γ^0 with no zeros or poles on the upper half plane (for example, a quotient of eta functions). Fix a non-elliptic basepoint z_0 on \mathbb{H} . Let $g(z) = \sqrt[n]{f(z)}$, where the branch is determined by specifying the value of $\sqrt[n]{f(z_0)}$ (since \mathbb{H} is simply connected and contains no zeros or poles of $f(z)$, the analytic continuation is well defined). Let $\mu = e^{2\pi i/n}$. Then for $\gamma \in \Gamma^0$, $f(\gamma z_0) = f(z_0)$, so $g(\gamma z_0) = \mu^m g(z_0)$ for some $m \in \mathbb{Z}/n\mathbb{Z}$. This defines a map $\chi : \Gamma^0 \rightarrow \mathbb{Z}/n\mathbb{Z}$, mapping γ to m . Let $\Gamma = \ker \chi$. Then $[\Gamma^0 : \Gamma]$ divides n and $g(z)$ is a modular function for Γ and not for any G with $\Gamma^0 \geq G \geq \Gamma$.

Let X_Γ and X_{Γ^0} be the modular curves for Γ and Γ^0 with $\pi : X_\Gamma \rightarrow X_{\Gamma^0}$ the natural projection. Then $g^n = f \circ \pi$, i.e. there is a commutative diagram:

$$\begin{array}{ccc} X_\Gamma & \xrightarrow{\pi} & X_{\Gamma^0} \\ & \searrow g^n & \downarrow f \\ & & \mathbb{C}P^1 \end{array}$$

and so:

$$\deg(f) \cdot \deg(\pi) = \deg(g^n) \tag{3.2}$$

Let $\gamma_1, \dots, \gamma_d$ be a set of coset representatives for Γ in Γ^0 . Then the γ_i 's are in one to one correspondence with the elements of $\chi(\Gamma^0)$, i.e. $|\chi(\Gamma^0)| = d$. Let $k = f(z_0)$ and note that if $f(z) = k$ then $g^n(\gamma_i z) = k$ for every i . This is to say that if f maps N points (on X_{Γ^0}) to k then g^n maps dN points (on X_Γ) to k . Since z_0 is not an elliptic point π doesn't ramify there

and we have:

$$\deg(g^n) = |\chi(\Gamma^0)| \deg(f)$$

Combining this with (3.2) we have

$$[\Gamma^0 : \Gamma] = \deg(\pi) = |\chi(\Gamma^0)|$$

If χ is surjective then Γ is an index n subgroup of Γ^0 . More generally:

Definition 27. If Γ^0 is a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and Γ is a normal, finite index subgroup of Γ^0 such that Γ^0/Γ is abelian, we call Γ a **character group** of Γ^0 .

We classify character groups into types using the map of modular curves:

$$\pi : X_\Gamma \rightarrow X_{\Gamma^0}$$

which is a map of Riemann surfaces. The projection induces a map on the homology groups

$$\pi_\# : H_1(X_\Gamma, \mathbb{Z}) \rightarrow H_1(X_{\Gamma^0}, \mathbb{Z})$$

Definition 28. Let Γ be a character group with the notation as above. Γ is said to be **Type I** if $\pi : X_\Gamma \rightarrow X_{\Gamma^0}$ ramifies only at the cusps and elliptic points, and $\pi_\#$ is surjective. Γ is **Type I(A)** if it is Type I but does not ramify at elliptic points. We say Γ is **Type II** if π is ramified only at elliptic points, and **Type II(A)** if π is unramified.

There is another way to think about Type I(A) and Type II(A) character groups, using the relation between Γ^0 and the fundamental group of its modular curve. Let Y_{Γ^0} be the modular curve X_{Γ^0} minus all cusps and elliptic points. There is an isomorphism:

$$\Psi : \Gamma^0 \rightarrow \pi_1(Y_{\Gamma^0})/T \tag{3.3}$$

Where T is the subgroup of $\pi_1(Y_{\Gamma^0})$ generated by every path in Y_{Γ^0} wrapping twice around order two elliptic points and wrapping three times around order three elliptic points. To obtain the map Ψ , fix z_0 , a non-elliptic point on \mathbb{H} and note that for any γ in Γ , γz_0 and z_0 are the same point in Y_{Γ^0} . So pick any path from z_0 to γz_0 in $(\mathbb{H} - \text{elliptic points})$ and define $\Psi(\gamma)$ to be the projection of this path onto Y_{Γ^0} . Note this is well defined, because if C_1 and C_2 are

two different paths from z_0 to γz_0 then the concatenation of C_1 and $-C_2$ is a closed path in \mathbb{H} that possibly encloses elliptic points, but paths around elliptic points in \mathbb{H} project to elements of T .

On the other hand, if c is an element of $\pi_1(Y_{\Gamma^0})/T$, then c lifts to a curve starting at z_0 and ending at γz_0 for some γ (since z_0 is not elliptic, γ is well defined up to sign). The endpoint is well defined because elements of T lift to null-homotopic paths in \mathbb{H} . Thus we have a well defined map from $\pi_1(Y_{\Gamma^0})/T$ to Γ^0 and which is an inverse to Ψ . Thus Ψ is an isomorphism.

It is well known that any genus g orientable compact surface can be obtained by gluing together a polygon with $2g$ sides labeled

$$\overline{A}_1, \overline{B}_1, \overline{A}_1^{-1}, \overline{B}_1^{-1}, \overline{A}_2, \dots, \overline{A}_g, \overline{B}_g, \overline{A}_g^{-1}, \overline{B}_g^{-1}$$

If we do this for X_{Γ^0} choosing paths not passing through elliptic points or cusps, then each \overline{A}_i and \overline{B}_i is a closed path in Y_{Γ^0} , and so we can let $A_i = \Psi^{-1}(\overline{A}_i)$ and $B_i = \Psi^{-1}(\overline{B}_i)$. We can pick a path \overline{a}_i around each cusp and \overline{b}_i around each elliptic point such that the concatenation:

$$\overline{a}_1 \cdots \overline{a}_n \overline{b}_1 \cdots \overline{b}_m \overline{A}_1 \overline{B}_1 \overline{A}_1^{-1} \overline{B}_1^{-1} \cdots \overline{A}_g \overline{B}_g \overline{A}_g^{-1} \overline{B}_g^{-1}$$

is homotopic to 0. Letting $a_i = \Psi^{-1}(\overline{a}_i)$ and $b_i = \Psi^{-1}(\overline{b}_i)$, we can show (cf. [10, Proposition 2.6]):

Theorem 29. *Let Γ^0 be a genus g finite index subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ with cusps $\{c_i\}_{i=1}^n$ and elliptic points $\{d_j\}_{j=1}^m$ of order e_j . Then there is a presentation of Γ^0 :*

$$\Gamma^0 \cong \langle a_1, \dots, a_n, b_1, \dots, b_m, A_1, B_1, \dots, A_g, B_g : R \rangle$$

where each a_i generates the stabilizer of the cusp c_i , each b_j generates the stabilizer of the elliptic point d_j , the paths in \mathbb{H} from z_0 to $A_i z_0$ and from z_0 to $B_i z_0$ projected onto the modular curve are a basis for $H_1(X_{\Gamma^0}, \mathbb{Z})$, and R is the set of relations

$$R = \{b_1^{e_1} = 1, \dots, b_m^{e_m} = 1, a_1 \cdots a_n b_1 \cdots b_m A_1 B_1 A_1^{-1} B_1^{-1} \cdots A_g B_g A_g^{-1} B_g^{-1} = 1\}$$

With this presentation, if Γ is a character group of Γ^0 coming from a character χ :

$$\begin{aligned} \Gamma \text{ is Type I} & \quad \text{if } \chi(A_i) = \chi(B_i) = I \text{ for all } i \\ \Gamma \text{ is Type I(A)} & \quad \text{if } \chi(b_i) = \chi(A_i) = \chi(B_i) = I \text{ for all } i \\ \Gamma \text{ is Type II} & \quad \text{if } \chi(a_i) = I \text{ for all } i \\ \Gamma \text{ is Type II(A)} & \quad \text{if } \chi(a_i) = \chi(b_i) = I \text{ for all } i \end{aligned}$$

The Type I(A) character groups of index p can be easily counted:

Lemma 30. *If Γ^0 has $\nu \geq 2$ inequivalent cusps, then for every prime p , Γ^0 has $\frac{p^{\nu-1}-1}{p-1}$ non-isomorphic index- p Type I(A) character groups.*

Proof. If Γ is an index- p Type I(A) character group of Γ^0 , then it is the kernel of a surjective homomorphism $\varphi : \Gamma^0 \rightarrow G$, for some finite abelian group G of order p . Since Γ is Type I(A), φ factors through the subgroup H of the abelianization of Γ^0 which is generated by the parabolic generators in the presentation of Γ^0 (i.e. φ is well defined on Γ^0/H). So the index- p Type I(A) character groups of Γ are in one-to-one correspondence with the index- p subgroups of H , and since H is isomorphic to $\mathbb{Z}^{\nu-1}$, it has $\frac{p^{\nu-1}-1}{p-1}$ of them. \square

3.3 The Main Theorem

The goal of this chapter is to prove the following theorem:

Theorem 31 (Main Result). *Let Γ^0 be any congruence subgroup with $\nu > 1$ inequivalent cusps. Then for any positive integer n there exists a nonzero integer M such that for all primes $p \nmid M$, Γ^0 has at least n non-isomorphic index- p Type I(A) character groups of Γ^0 satisfying the (UBD) condition.*

To prove this result, we need two results from [13]. The first one gives a condition for a function to have unbounded denominators:

Lemma 32 ([13] Lemma 11). *Let K be a number field, p a prime number, and*

$$f(w) = a_0 + \sum_{m \geq 1} a_m w^m, \quad a_m \in K, a_0 \neq 0$$

such that for every m , a_m is \wp -integral for any prime ideal \wp in \mathcal{O}_K above p . Expand $\sqrt[p]{f(w)} = \sum_{m \geq 0} b_m w^m$ formally (fixing a branch for the p th root of a_0). If there exists at least one b_m such that $\text{ord}_\wp(b_0) - \text{ord}_\wp(b_m) > \frac{\text{ord}_\wp(a_0)}{p}$, then the sequence $\{b_m\}$ has unbounded denominators.

The second lemma (which is modified from Proposition 14 of [13]) gives a condition for a group to satisfy (UBD). For the lemma, let $R_K(\Gamma)$ denote the ring of modular functions on Γ with Fourier coefficients in K and with no poles except possibly at ∞ . In our discussion, K will always be a sufficiently large number field without any further specification.

Lemma 33. *Let Γ^0 be a congruence group with genus larger than 0, and let Γ be a character group of Γ^0 such that the field of meromorphic modular forms for Γ is generated over that of Γ^0 by some $g \in R_K(\Gamma)$ satisfying $g^n = f$ for some $f \in R_K(\Gamma^0)$ which is supported at the cusps of Γ^0 . If g^m has unbounded denominators for all $m \in \{n+1, \dots, 2n-1\}$ then Γ satisfies (UBD).*

3.4 Equation of the Modular Curve

Let Γ^0 be either $\Gamma_0(M)$ or $\Gamma_1(M)$ for some positive integer M . By [26] we can use generalized Dedekind eta functions to find modular functions $x(z)$ and $y(z)$ for Γ^0 with integer Fourier coefficients satisfying an irreducible polynomial of the form:

$$F(x, y) = x^n - y^m - r(x, y) = 0$$

Moreover, x and y have unique poles at infinity of orders m and n respectively, where $(n, m) = 1$ and $n > m$, and $\deg(r) < m$. Also their initial Fourier coefficients are 1. The following lemma shows that the polynomial $F(x, y)$ can be taken to have integer coefficients:

Lemma 34. *Suppose x and y are two power series in q with coefficients in a number field K and $p(z, w)$ is an irreducible polynomial in $\mathbb{C}[z, w]$ such that $p(x, y) = 0$. Suppose also that if S is the set of monomials in p then $\dim(\text{span}_{\mathbb{C}}(S)) = |S| - 1$. Then there is a constant λ such that $\lambda \cdot p(z, w)$ is in $K[z, w]$.*

Proof. Write

$$p(z, w) = \sum_{(i,j) \in S} c_{ij} z^i w^j$$

Each element $x^i y^j$ for $(i, j) \in S$ has a Fourier series, and the coefficients can be treated as an infinite length vector. Since these vectors are linearly dependent, we can do Gaussian elimination on them to get a nontrivial linear combination equal to zero:

$$\sum_{(i,j) \in S} d_{ij} x^i y^j = 0$$

But Gaussian elimination keeps us in the field K , i.e. $d_{ij} \in K$ for all i, j . And since the span of S is one dimension less than $|S|$, all linear combinations summing to zero are multiples of each other, i.e. there is $\lambda \in \mathbb{C}$ such that $d_{ij} = \lambda c_{ij}$. Therefore $\lambda \cdot p(z, w) \in K[z, w]$. \square

The condition that $\dim(\text{span}_{\mathbb{C}}(S)) = |S| - 1$ is satisfied for $F(x, y)$ above because F was constructed as a minimal polynomial in [26]. We thus have that $F(x, y)$ can be chosen to be in $\mathbb{Q}[x, y]$, and hence in $\mathbb{Z}[x, y]$ by appropriate transformations $x \rightarrow p^m x, y \rightarrow p^n y$.

The next theorem is used to show the cusps correspond to algebraic points on the modular curve. The proof is left to the end of the chapter:

Theorem 35. *Let f be a modular function for Γ whose Fourier expansion about ∞ has coefficients in a number field K . Then for every $\gamma \in \text{SL}_2(\mathbb{Z})$, the Fourier expansion of $f[\gamma]$ about ∞ , i.e. the expansion of f at the cusp $\gamma \cdot \infty$, also has coefficients in a number field K' (K' may be larger than K in general).*

Proof. Appendix. \square

The cusp at infinity maps to the unique point at infinity of the curve $F(x, y) = 0$. Suppose $c = \gamma\infty$ is another cusp different from ∞ . Then $x(c) = x(\gamma\infty) = x[\gamma](\infty)$ is algebraic, and so is $y(c)$. The algebraicity of the cusps will be important because of the following theorem ([5] Theorem 2.1.3):

Theorem 36 (Manin-Drinfeld). *Let Γ^0 be a congruence group and let α and β be cusps of Γ^0 . Then the divisor $(\alpha) - (\beta)$ is a torsion element in the Jacobian of the modular curve X_{Γ^0} . (I.e. there is a modular function f for Γ^0 such that $\text{div}(f) = N_{\alpha\beta}(\alpha) - N_{\alpha\beta}(\beta)$ for some integer $N_{\alpha\beta} > 0$.)*

By the Manin-Drinfeld theorem, each cusp is a torsion point in the Jacobian of the modular curve (where the cusp at infinity is the identity element in the Jacobian). This means there is a modular function for Γ^0 with a unique pole of order $N_{\alpha\beta}$ at the cusp β and a unique zero of order $N_{\alpha\beta}$ at the cusp α , for some positive integer $N_{\alpha\beta}$. For the next two sections, we restrict ourselves to the case where $\beta = \infty$. Using suitable quotients of modular functions of this type, one can construct modular functions supported at any two cusps of Γ^0 , however we cannot say anything yet about the algebraicity of their Fourier coefficients. Our present goal will be to construct a function with algebraic Fourier coefficients whose only pole agrees with that of a function coming from Manin-Drinfeld. We will then show the two functions differ only by a constant.

We will need to avoid primes dividing any $N_{\alpha\beta}$, so let $D(\Gamma^0)$ be the least common multiple of all $N_{\alpha\beta}$.

3.5 Constructing the Function

Let P be a cusp of Γ^0 different from ∞ . We continue to assume that x and y have unique poles at infinity of orders n and m respectively where $n > m > 0$, $(n, m) = 1$. By a change of variables we can assume x and y have a zero of order at least 1 at P . Taking functions of the form $x^a y^b$ we get modular functions of Γ^0 with poles of order k at infinity for almost every $k > 0$, specifically those which can be written $k = an + bm$ for some $a \geq 0$ and $b \geq 0$. The number of “missing” functions is given by the following lemma:

Lemma 37. *Let n and m be two relatively prime positive integers. Then there are exactly $(n-1)(m-1)/2$ integers $d \geq 0$ which cannot be written in the form $an + bm = d$ for nonnegative integers a and b .*

Proof. Let S be the set of nonnegative integers which cannot be written in the above form. It is easy to see that any integer greater than nm is not in S , so we count the number of nonnegative integers less than or equal to nm , but not in S . If we associate points (a, b) on a lattice $\mathbb{Z} \times \mathbb{Z}$ with the integers $an + bm$ then this is the same as the counting number of lattice points with $a \geq 0$, $b \geq 0$ and $an + bm \leq nm$ (where $(m, 0)$ and $(0, n)$ count as half points). But

this is half the number of lattice points in the rectangle given by $0 \leq a \leq m$ and $0 \leq b \leq n$, which is $(n+1)(m+1)/2$. Thus

$$|S| = nm + 1 - (n+1)(m+1)/2 = (n-1)(m-1)/2$$

□

The quantity $\frac{1}{2}(n-1)(m-1)$ is also an upper bound for the genus of Γ^0 . The proof of the following lemma is left to the end of the paper:

Lemma 38. *Let C be a projective algebraic curve of the form $x^n = y^m + r(x, y)$ where $n > m > 0$, $(n, m) = 1$, and $\deg(r) \leq m$. Let $g(C)$ be the genus of C . Then:*

$$g(C) \leq \frac{(n-1)(m-1)}{2}$$

with equality if C has no singularities other than at ∞ .

Proof. Appendix. □

Lemma 39. *If $f_1 \dots f_d$ are d linearly independent modular functions with a unique pole at infinity of order at most k and with a zero at P of order at least l , then there is a linear combination of these functions, $g = \sum_{i=1}^d c_i f_i$ such that g has a zero of order at least $d-1+l$ at P and a pole of order at most k at infinity. Moreover if the Fourier coefficients of the f_i 's are in a number field K , then each c_i is also in K .*

Proof. Gaussian elimination. Every function beyond the first allows us to eliminate one term of the power series and increase the order of the zero by one. □

Suppose P is an N -torsion point in the Jacobian of Γ^0 . Then there is a modular function h_1 with divisor $\text{div}(h_1) = N(P) - N(\infty)$. On the other hand, by the construction above, we can construct a function h_2 with a single pole of order at most N at infinity and a zero at P of order at least $N - \frac{1}{2}(n-1)(m-1) = N - g$. I.e.

$$\text{div}(h_2) = -M(\infty) + (N-g)(P) + \text{some positive divisor}$$

for $M \leq N$.

We use a standard result about divisors (See [17] Lemma II.5):

Lemma 40. *Let D be a divisor of degree 0 on a curve C of genus g and let P_1, \dots, P_g be independent generic points. Then there exists exactly one positive divisor on C linearly equivalent to $D + \sum_{i=1}^g (P_i)$.*

From this lemma it follows that principal divisors with too few positive terms are zero:

Lemma 41. *Let C be a curve of genus g and let D be a principal divisor on C (i.e. D is linearly equivalent to 0). Divide D into positive and negative terms: $D = D^+ - D^- = \sum (P_i) - \sum (Q_i)$. If $\deg D^+ \leq g$ then $D = 0$.*

Proof. D is principal so $\deg D = 0$ and there must be equally many terms in the sums $\sum (P_i)$ and $\sum (Q_i)$. By assumption there are at most g terms. If there are n terms with $n < g$ then let $P_i = Q_i = 0$ for $n + 1 < i \leq g$. So

$$D = \sum_{i=1}^g (P_i) - \sum_{i=1}^g (Q_i).$$

Because D is linearly equivalent to 0, $D + \sum (Q_i)$ is linearly equivalent to $\sum (Q_i)$ which is a positive divisor. But $D + \sum (Q_i)$ is itself a positive divisor, and the above lemma says $D + \sum (Q_i)$ is linearly equivalent to exactly one positive divisor. So $D + \sum (Q_i) = \sum (Q_i)$, hence $D = 0$. \square

So since $\operatorname{div}(h_1/h_2) = g(P) - (\text{other stuff})$ has degree $\leq g$, by the Lemma, $\operatorname{div}(h_1/h_2) = 0$. Thus h_2 is a constant multiple of h_1 and $\operatorname{div}(h_2) = N(P) - N(\infty)$. We summarize in a theorem:

Theorem 42. *Let Γ^0 be $\Gamma_0(M)$ or $\Gamma_1(M)$ for some integer $M > 0$ and let P be a cusp other than ∞ . Then there is a smallest positive integer N_P and a modular function $h_P(z)$ for Γ^0 with Fourier coefficients in a number field K such that $\operatorname{div}(h_P) = N_P(P) - N_P(\infty)$.*

3.6 Proof of the Main Theorem

If Γ is an index- p character group of Γ^0 and p is a prime not dividing $D(\Gamma^0)$ then by Lemma 10 of [13], the field of meromorphic functions for Γ is a simple extension of that of Γ^0 by some element $g = \sqrt[p]{f}$ where f is a modular function for Γ^0 . If additionally Γ is a character group

of Γ^0 of Type I(A), then the divisors of f are supported at the cusps of Γ^0 . We will write f in terms of the functions constructed in the previous section.

Let $\infty, c_1, \dots, c_{\nu-1}$ be a set of inequivalent cusp representatives of Γ^0 . Let γ_i be a generator for the stabilizer of c_i for each i , and let h_{c_i} be the modular function constructed in the last section with $\text{div}(h_{c_i}) = N_{c_i}(c_i) - N_{c_i}(\infty)$. If p is a prime not dividing any of the N_{c_i} 's (i.e. $p \nmid D(\Gamma^0)$) then $\sqrt[p]{h_{c_i}}$ is a modular function for the Type I(A) character group $\ker \varphi_i$ where $\varphi_i(\gamma_{c_j}) = 0$ if $i \neq j$ and $\varphi_i(\gamma_{c_j}) = 1$ if $i = j$. This means:

$$\sqrt[p]{h_{c_i}}[\gamma_{c_j}] = \begin{cases} e^{2\pi i/p} \sqrt[p]{h_{c_i}} & \text{if } i = j \\ \sqrt[p]{h_{c_i}} & \text{if } i \neq j \end{cases}$$

Let $\varphi : \Gamma^0 \rightarrow \mathbb{Z}/p\mathbb{Z}$ be such that $\Gamma = \ker \varphi$ is an index- p Type I(A) character group. Recall from the proof of Theorem 30 that φ factors through H , the subgroup of the abelianization of Γ^0 generated by $\{\gamma_{c_i}\}$. Thus φ is completely determined by the values n_j where:

$$\varphi(\gamma_{c_j}) = e^{2\pi i n_j/p}.$$

For each j let e_j be the smallest nonnegative integer congruent to n_j and define

$$g = \left(\prod_i h_{c_i}^{e_i} \right)^{1/p}$$

where the product ranges from 1 to $\nu - 1$. Then g is the generator for the modular functions of Γ since:

$$\begin{aligned} g[\gamma_{c_j}] &= \left(\prod_i h_{c_i}^{e_i} \right)^{1/p} [\gamma_{c_j}] = \prod_i \left(\sqrt[p]{h_{c_i}}[\gamma_{c_j}] \right)^{e_i} \\ &= e^{2\pi i e_j/p} \prod_i \sqrt[p]{h_{c_i}^{e_i}} \\ &= e^{2\pi i n_j/p} g \end{aligned}$$

So $\mathbb{C}(\Gamma)$ is a degree p field extension over $\mathbb{C}(\Gamma^0)$ generated by $(\prod_i h_{c_i}^{e_i})^{1/p}$.

Lemma 43. *Let $h(z)$ be any nonconstant modular function for Γ^0 with a pole at infinity, all other poles and zeros only at the cusps, and with algebraic Fourier coefficients. For almost all primes p , $\sqrt[p]{h(z)}$ generates a degree- p extension of $\mathbb{C}(\Gamma^0)$ corresponding to an index- p Type I(A) noncongruence character group of Γ^0 which satisfies the condition (UBD).*

Proof. Suppose $h(z)$ has a pole of order n at infinity. Then it has a Fourier series

$$h(z) = \sum_{i=-n}^{\infty} a_i q^i.$$

Let p be a prime number such that $\text{ord}_p(a_{-n}) = \text{ord}_p(a_{-n+1}) = 0$ (Note all but finitely many primes satisfy this). Then the formal p th root of $h(z)$ is of the form:

$$\sqrt[p]{h(z)} = q^{-n/p} \left(\sqrt[p]{a_{-n}} + \frac{a_{-n+1}}{p} \cdot q + \dots \right)$$

and $\sqrt[p]{h(z)}$ is a modular function for an index- p character group of Γ^0 . Call this group Γ . Then the modular functions for Γ are generated by those of Γ^0 and $\sqrt[p]{h(z)}$. I.e. $\mathbb{C}(\Gamma) = \mathbb{C}(\Gamma^0)(\sqrt[p]{h(z)})$.

Dividing q^{-n} out of $h(z)$ we can use Lemma 32 with $b_0 = \sqrt[p]{a_{-n}}$ and $b_1 = \frac{a_{-n+1}}{p}$, noting:

$$\text{ord}_p(b_0) - \text{ord}_p(b_1) = \frac{1}{p} \text{ord}_p(a_{-n}) - (\text{ord}_p(a_{-n+1}) - \text{ord}_p(p)) = 1$$

and

$$\frac{\text{ord}_p(a_0)}{p} = 0.$$

So $\text{ord}_p(b_0) - \text{ord}_p(b_1) > \frac{\text{ord}_p(a_0)}{p}$, and by Lemma 32 $\sqrt[p]{h(z)}$ has unbounded denominators.

Moreover, for any m on $\{p+1, \dots, 2p-1\}$

$$\sqrt[p]{(h(z))^m} = q^{-n/p} \left(\sqrt[p]{a_{-n}^m} + \frac{a_{-n+1}m}{p} \cdot q + \dots \right)$$

and $\text{ord}_p(m) = 0$, so Lemma 32 says $\sqrt[p]{(h(z))^m}$ has unbounded denominators. Hence by Lemma 33, Γ satisfies (UBD). \square

Corollary 44. *For each $j = 1, \dots, \nu - 1$, there exists a positive integer M_j such that for every prime $p \nmid M_j$, the index- p noncongruence subgroup $\Gamma_{j,p}$ of Γ^0 whose meromorphic functions are generated over $\mathbb{C}(\Gamma^0)$ by $\sqrt[p]{h_{c_j}}$ satisfies (UBD).*

Our goal was to show that for every positive integer n , there is an integer M such that for every p not dividing M there are at least n non-isomorphic index- p Type I(A) character groups satisfying (UBD). We can now prove this:

Proof of Theorem 31. For any integer $n > 0$, pick n different nonzero vectors of the form $(e_1, \dots, e_{\nu-1}) \in \mathbb{Z}_+^{\nu-1}$ whose first nonzero entry is 1. Then construct n different modular functions $\prod_{j=1}^{\nu-1} (h_{c_j})^{e_j}$. By Lemma 43, for all primes p except for a finite set of bad primes, $\sqrt[p]{\prod_{j=1}^{\nu-1} (h_{c_j})^{e_j}}$ corresponds to an index- p Type I(A) subgroup of Γ^0 satisfying the condition (UBD). If two of the vectors are equal mod p then they generate the same group, so if there are less than n inequivalent vectors for p , we add p to the set of bad primes. We let M be the product of all the bad primes, of which there are finitely many. \square

3.7 Appendix: Proofs of Some Results

3.7.1 Proof of Theorem 35

Let $j(z)$ be the classical modular j -function. By Theorem 1 of [1] there is an irreducible polynomial $G(x, y) \in \mathbb{C}[x, y]$ such that $G(f, j) = 0$. Since both f and j have algebraic coefficients at infinity, Lemma 34 says $G(x, y)$ can be chosen to be in $K'[x, y]$ for some number field K' . Since $j[\gamma] = j$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, $G(f[\gamma], j) = 0$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. The claim of the theorem is equivalent to saying every solution of $G(f, j) = 0$, as a formal power series (in a fractional power of q), has algebraic Fourier coefficients. Since $j(z)$ is a fixed series in q , from now on we will write $g(f, q) = G(f, j)$, a polynomial in f with coefficients in the ring of Laurent series in q .

Lemma 45. *Let M be a nonnegative integer. Then:*

$$\frac{d^M}{dq^M} (g(f(q), q)) = \sum_p \left(c_p \left(\frac{\partial^{n+M-d} g(f, q)}{\partial f^n \partial q^{M-d}} \right) \cdot \prod_{i=1}^n \frac{d^{d_i} f}{dq^{d_i}} \right)$$

where the sum ranges over all partitions

$$p : d_1 + \dots + d_n = d, \quad d_i \geq 1$$

of all d with $0 \leq d \leq M$ (where the partition of 0 is empty), and c_p is a combinatorial constant:

$$c_p = \binom{M}{d} \frac{d!}{d_1! \cdots d_n!} \cdot \frac{1}{\prod_{i=1}^d \#(i \in p)!}$$

Proof. First note that:

$$\frac{d}{dq} \left(\frac{\partial^{a+b} g(f, q)}{\partial f^a \partial q^b} \right) = \frac{\partial^{a+b+1} g}{\partial f^{a+1} \partial q^b} \frac{df}{dq} + \frac{\partial^{a+b+1} g}{\partial f^a \partial q^{b+1}}$$

We claim that each term of $\frac{d^m}{dq^m} (g(f(q), q))$ is of the form:

$$\frac{\partial^{n+m-d} g}{\partial f^n \partial q^{m-d}} \cdot \prod_{i=1}^n \frac{d^{d_i} f}{dq^{d_i}}$$

where $d_1 + \dots + d_n = d \leq m$.

Suppose this is true for m . Then $\frac{d}{dq} \left(\frac{\partial^{n+m-d} g}{\partial f^n \partial q^{m-d}} \cdot \prod_{i=1}^n \frac{d^{d_i} f}{dq^{d_i}} \right)$ has three types of terms, corresponding to new partitions:

$$\begin{aligned} \frac{\partial^{n+m-d+1} g}{\partial f^{n+1} \partial q^{m-d}} \cdot \prod_{i=1}^{n+1} \frac{d^{d_i} f}{dq^{d_i}} & \quad \text{for } d_1 + \dots + d_n + 1 \\ \frac{\partial^{n+m-d+1} g}{\partial f^n \partial q^{m-d+1}} \cdot \prod_{i=1}^n \frac{d^{d_i} f}{dq^{d_i}} & \quad \text{for } d_1 + \dots + d_n \\ \frac{\partial^{n+m-d} g}{\partial f^n \partial q^{m-d}} \cdot \prod_{i=1}^n \frac{d^{d_i} f}{dq^{d_i}} & \quad \text{for } d_1 + \dots + (d_j + 1) + \dots + d_n \end{aligned}$$

where for the last type, there is one for each $1 \leq j \leq n$.

Thus differentiation on a term corresponding to a partition p splits p up into $n+2$ partitions: p itself, p appending “+1”, and all partitions p with 1 added to one of the elements of p .

To get the combinatorial coefficient, we count how many ways to get to a partition p in M steps using the three rules above. If $M > d$ there are steps where p doesn't change, and they can be put in any order, hence the $\binom{M}{d}$ term in c_p . The remaining steps consist of adding +1 to the d_i 's, hence the multinomial coefficient, and the remaining term is to remove any overlap in counting when $d_i = d_j$ for some i and j . \square

Let $g(x, q)$ be a degree N polynomial (in x) with coefficients in $K[[q]]$ for some field K :

$$g(x, q) = \sum_{i=0}^N g_i(q) x^i = \sum_{j=P}^{\infty} h_j(x) q^j.$$

We want to find $f(q)$ such that $g(f(q), q) = 0$. If the order of $f(q)$ at ∞ is Q then $q^{-Q}f(q)$ is holomorphic and non-zero at ∞ , and it satisfies $\bar{g}(q^{-Q}f(q), q) = 0$ where

$$\bar{g}(x, q) = \sum_{i=0}^N (g_i(q) \cdot q^{Qi}) x^i.$$

So, in solving for $f(q)$, we can adjust the h_j polynomials and assume f is holomorphic and non-zero at ∞ . Moreover, we can assume $P = 0$ (and hence each $g_i(q)$ is holomorphic at ∞) since we can multiply powers of q to both sides of $g(f(q), q) = 0$. Let:

$$f(q) = \sum_{i=0}^{\infty} a_i q^i.$$

We will plug these series into the Lemma. Note that:

$$\frac{\partial^{a+b} g(f, q)}{\partial f^a \partial q^b} = \sum_{j=b}^{\infty} h_j^{(a)}(f) j(j-1) \dots (j-b+1) q^{j-b}.$$

So:

$$\left. \frac{\partial^{a+b} g(f, q)}{\partial f^a \partial q^b} \right|_{q=0} = b! h_b^{(a)}(a_0).$$

Similarly:

$$\left. \frac{d^b f}{dq^b} \right|_{q=0} = b! a_b.$$

Now let Q_M be the M -th coefficient of $g(f(q), q)$. Since

$$Q_M = \frac{1}{M!} \left. \frac{d^M}{dq^M} g(f(q), q) \right|_{q=0},$$

putting it all together we have:

$$\begin{aligned} Q_M &= \frac{1}{M!} \sum_p \left(c_p \frac{\partial^{n+M-d} g}{\partial f^n \partial q^{M-d}} \cdot \prod_{i=1}^n \frac{d^{d_i} f}{dq^{d_i}} \right) \Big|_{q=0} \\ &= \frac{1}{M!} \sum_p \left(\binom{M}{d} \frac{d!}{d_1! \dots d_n!} \cdot \frac{1}{\prod_{i=1}^d \#(i \in p)!} \cdot (M-d)! h_{M-d}^{(n)}(a_0) \prod_{i=1}^n d_i! a_{d_i} \right) \\ &= \sum_p \frac{1}{\prod_{i=1}^d \#(i \in p)!} h_{M-d}^{(n)}(a_0) \prod_{i=1}^n a_{d_i}. \end{aligned}$$

For example:

$$\begin{aligned} Q_4 &= a_4 h'_0(a_0) + a_3 a_1 h''_0(a_0) + a_3 h'_1(a_0) + \frac{1}{2} a_2^2 h''_0(a_0) + \frac{1}{2} a_2 a_1^2 h'''_0(a_0) + \\ &\quad + a_2 a_1 h''_1(a_0) + a_2 h'_2(a_0) + \frac{1}{24} a_1^4 h''''_0(a_0) + \frac{1}{6} a_1^3 h'''_1(a_0) + \\ &\quad + \frac{1}{2} a_1^2 h''_2(a_0) + a_1 h'_3(a_0) + h_4(a_0). \end{aligned}$$

We solve $g(f(q), q) = 0$ for the a_i 's. Since $Q_0 = h_0(a_0)$, we pick a_0 to be any non-zero root of h_0 , an (at most) N -th degree polynomial. If a_0 is a simple root we will see that we can successively solve each a_i . Suppose however that $h_0^{(i)}(a_0) = 0$ for all $i \in \{0, 1, \dots, w-1\}$ and $h_0^{(w)}(a_0) \neq 0$. If $w > 1$ then the a_i 's cannot be solved. In this case, instead let $f = \sum a_i q^{i/w}$. Then replace $q^{1/w}$ with q and re-index h_j as h_{jw} and $h_j = 0$ whenever $j \not\equiv 0 \pmod w$. So we have:

$$\begin{aligned} f(q) &= \sum a_i q^i, \\ g(x) &= \sum h_{jw}(x) q^{jw}. \end{aligned}$$

Then $Q_i = 0$ for all i from 1 to $w-1$, because each of their terms contains either $h_0^{(j)}(a_0)$ for $j \in [0, w-1]$ or h_j for $j \in [1, w-1]$. The next non-zero term is:

$$Q_w = h_w(a_0) + a_1^w h_0^{(w)}(a_0).$$

So we solve

$$a_1^w = \frac{-h_w(a_0)}{h_0^{(w)}(a_0)}.$$

There are two cases:

Case 1: If $a_1 \neq 0$ then there are exactly w choices for a_1 and they all differ by an w -th root of unity. Moreover, all subsequent a_i 's are uniquely determined because for example:

$$Q_{w+1} = a_1 h'_w(a_0) + a_1^{w+1} h_0^{(w+1)}(a_0) + a_1^{w-1} a_2 h_0^{(w)}(a_0)$$

can be solved for a_2 . And more generally:

$$Q_{w+c} = a_1^{w-1} a_{c+1} h_0^{(w)}(a_0) + (\text{terms with all } a_i \text{'s having } i \leq c).$$

So we can solve for each a_{c+1} .

(In general, when calculating Q_M , the partitions that give (possibly) non-zero terms are partitions $d_1 + \dots + d_n = d$ such that (1) $d \leq M$, (2) $M \equiv d \pmod N$, and (3) $n \geq N$ if $M = d$.)

Case 2: On the other hand, if $a_1 = 0$, let $\bar{f}(q) = f(q) + q$ and $\bar{g}(x, q) = g(x - q, q)$. The coefficients of \bar{f} and f are the same except the q -term is non-zero, and $\bar{g}(\bar{f}(q), q) = 0$. If we

repeat the above process on \bar{g} and \bar{f} we go into Case 1 and get a sequence for $\bar{f}(q)$, and hence $f(q)$. There is some reindexing involved in this, but note that the “cuspid width” w remains the same after the reindexing, because replacing x with $x - q$ in $g(x, q) = \sum h_j(x)q^j$ does not change the $h_0(x)$ term. That is to say, if

$$\bar{g}(x, q) = \sum_{j=0}^{\infty} \bar{h}_j(x)q^j$$

then $\bar{h}_0(x) = h_0(x)$.

Note that in this recursive solving process, we stay in the field K' . That is to say, $a_i \in K'$ for all i . This proves Theorem 35.

3.7.2 Proof of Lemma 38

Proof of Lemma 38. The genus of C is:

$$g(C) = \frac{(n-1)(n-2)}{2} - \sum_i \frac{r_i(r_i-1)}{2}$$

where the sum ranges over each singularity and infinitely near points to the singularity (see [7, V.3]) and r_i is the singularity multiplicity. Let $R_P(C)$ denote the sum of these contributions from a singularity at P and let ∞ denote the (unique) point at infinity of C . Then I claim that $R_\infty(C) = \frac{(n-1)(n-m-1)}{2}$, from which it will follow that:

$$\begin{aligned} g(C) &\leq \frac{(n-1)(n-2)}{2} - R_\infty(C) \\ &= \frac{(n-1)(m-1)}{2} \end{aligned}$$

Note that the under a change of variables, ∞ becomes the origin of $D : x^n = z^{n-m}(1 + s(x, z))$ where $s(x, z)$ is some polynomial with $s(0, 0) = 0$. I claim that for such a curve, the contribution to the genus formula from $(0, 0)$ is $R_{(0,0)}(D) = \frac{(n-1)(n-m-1)}{2}$. We have

$$R_{(0,0)}(D) = \sum \frac{r_i(r_i-1)}{2}$$

where the sum ranges over all infinitely near points to $(0, 0)$. These are obtained by successively taking the blowing-up about singular points. We work by induction on n and m .

First, if $n = 2$ and $m = 1$ then the formula for $R_{(0,0)}(D)$ is certainly correct because D is non-singular at $(0,0)$. Next, suppose $R_{(0,0)}(D) = \frac{(n-1)(n-m-1)}{2}$ for all curves of the above form with $m < n < N$, and let D be $x^N = z^{N-m}(1 + s(x, z))$ with $s(0,0) = 0$. The blowing-up about $(0,0)$ is obtained by making the substitution $z = tx$ to get a new curve $D' : x^m - t^{N-m}(1 + s(x, xt)) = 0$. By the induction hypothesis, this curve has $R_{(0,0)}(D') = \frac{(m-1)(N-m-1)}{2}$. But also, the singularity multiplicity at $(0,0)$ is m . But then:

$$\begin{aligned} R_{(0,0)}(D) &= R_{(0,0)}(D') + \frac{m(m-1)}{2} \\ &= \frac{(m-1)(N-m-1)}{2} + \frac{m(m-1)}{2} \\ &= \frac{(N-1)(m-1)}{2} \end{aligned}$$

It remains to show the $n = N, m = N - 1$ case. If D is the curve $x^N = z^{N-1}(1 + s(x, z))$, D has singularity multiplicity $N - 1$ at $(0,0)$. The first blowing up is $D' : x = z^{N-1}(1 + s(x, xt))$, and by the induction hypothesis, $R_{(0,0)}(D') = \frac{(1-1)(N-1-1)}{2} = 0$. So:

$$R_{(0,0)}(D) = \frac{(N-1)(N-2)}{2} + R_{(0,0)}(D') = \frac{(N-1)(N-2)}{2}$$

This completes the induction. Then $R_\infty(C) = R_{(0,0)}(D) = \frac{(n-1)(n-m-1)}{2}$, completing the proof. \square

BIBLIOGRAPHY

- [1] Atkin, A. O. L. and Swinnerton-Dyer, H. P. F. (1971). Modular forms on noncongruence subgroups. In *Combinatorics (Proc. Sympos. Pure Math., Vol. XIX, Univ. California, Los Angeles, Calif., 1968)*, pages 1–25. Amer. Math. Soc., Providence, R.I.
- [2] Berger, G. (1994). Hecke operators on noncongruence subgroups. *C. R. Acad. Sci. Paris Sér. I Math.*, 319(9):915–919.
- [3] Birch, B. (1994). Noncongruence subgroups, covers and drawings. In *The Grothendieck theory of dessins d'enfants (Luminy, 1993)*, volume 200 of *London Math. Soc. Lecture Note Ser.*, pages 25–46. Cambridge Univ. Press, Cambridge.
- [4] Coste, A. and Gannon, T. (1999). Congruence groups and rational conformal field theory. *Preprint*.
- [5] Cremona, J. (1997). *Algorithms for Modular Elliptic Curves*. Cambridge University Press.
- [6] Diamond, F. and Shurman, J. (2005). *A first course in modular forms*. Springer.
- [7] Hartshorne, R. (1977). *Algebraic Geometry*. Springer.
- [8] Hsu, T. (1996). Identifying congruence subgroups of the modular subgroup. *Proceedings of the American Mathematical Society*, 124(5):1351–1359.
- [9] Hsu, T. (1997). Permutation techniques for coset representations of modular subgroups. *London Math. Soc. Lecture Note Ser.*, 243:67–77.
- [10] Iwaniec, H. (2002). *Spectral Methods of Automorphic Forms*. American Mathematical Society.

- [11] Koblitz, N. (1993). *Introduction to elliptic curves and modular forms*. Springer-Verlag, New York, second edition.
- [12] Kulkarni, R. S. (1991). An arithmetic geometric method in the study of the subgroups of the modular group. *American Journal of Mathematics*, 113(6):1053–1133.
- [13] Kurth, C. and Long, L. (2008a). On modular forms for some noncongruence subgroups of $SL_2(\mathbb{Z})$. *Journal of Number Theory*, 128:1989–2009.
- [14] Kurth, C. and Long, L. (2008b). On modular forms for some noncongruence subgroups of $SL_2(\mathbb{Z})$ II. *Preprint*.
- [15] Lang, M. L. (2002). Normalisers of subgroups of the modular group. *Journal of Algebra*, 248:202–218.
- [16] Lang, M. L., Lim, C.-H., and Tan, S.-P. (1995). An algorithm for determining if a subgroup of the modular group is congruence. *Journal of the London Mathematical Society*, 51:491–502.
- [17] Lang, S. (1959). *Abelian Varieties*. Interscience Publishers Inc.
- [18] Li, W. (1996). *Number theory with applications*. World Scientific Publishing Co. Inc., River Edge, NJ.
- [19] Manin, Y. (1972). Parabolic points and zeta functions of modular curves. *Izv. Akad. Nauk SSSR Ser. Mat.*, 36:19–66.
- [20] Millington, M. H. (1969a). On cycloidal subgroups of the modular group. *Proc. London Math. Soc.*, 19:164–176.
- [21] Millington, M. H. (1969b). Subgroups of the classical modular group. *J. London Math. Soc.*, 1:351–357.
- [22] Rankin, R. (1977). *Modular forms and functions*. Cambridge University Press.

- [23] Shimura, G. (1971). *Introduction to the arithmetic theory of automorphic functions*. Publications of the Mathematical Society of Japan, No. 11. Iwanami Shoten, Publishers, Tokyo. Kanô Memorial Lectures, No. 1.
- [24] Stein, W. (2007). *Modular Forms, a Computational Approach*. American Mathematical Society.
- [25] Wohlfahrt, K. (1964). An extension of F. Klein's level concept. *Ill. J. Math.*, 8:529–539.
- [26] Yang, Y. (2006). Defining equations of modular curves. *Advances in Mathematics*, 204:481–508.