

**LOW-LYING GEODESICS IN AN ARITHMETIC  
HYPERBOLIC THREE-MANIFOLD**

**BY KATIE LYNN MCKEON**

**A dissertation submitted to the  
School of Graduate Studies  
Rutgers, The State University of New Jersey  
In partial fulfillment of the requirements**

**For the degree of  
Doctor of Philosophy  
Graduate Program in Mathematics**

**Written under the direction of  
Alex Kontorovich**

**And approved by**

---

---

---

---

**New Brunswick, New Jersey**

**October, 2018**

## ABSTRACT OF THE DISSERTATION

# Low-lying Geodesics in an Arithmetic Hyperbolic Three-Manifold

by KATIE LYNN MCKEON

Dissertation Director: Alex Kontorovich

We examine closed geodesics in the quotient of hyperbolic three space by the discrete group of isometries  $SL(2, \mathbb{Z}[i])$ . There is a correspondence between closed geodesics in the manifold, the complex continued fractions originally studied by Hurwitz, and binary quadratic forms over the Gaussian integers. According to this correspondence, a geodesic is called fundamental if the associated binary quadratic form is. Using techniques from sieve theory, symbolic dynamics, and the theory of expander graphs, we show the existence of a compact set in the manifold containing infinitely many fundamental geodesics.

## Acknowledgements

First, I owe many thanks to my adviser Alex Kontorovich for his insight and patient mentorship. This work was made possible by his direction and supervision. Also, I am indebted to Steve Miller for urging me along the path to studying number theory. The number theory faculty and students at Rutgers have taught me so much along the way. I would also like to thank the other members of my thesis committee, Henryk Iwaniec, Lillian Pierce, and Jerrold Tunnell for their time and feedback. Finally, I am grateful for all the support I have received from my friends and family throughout my years as a graduate student.

# Table of Contents

<b>Abstract</b> . . . . .	ii
<b>Acknowledgements</b> . . . . .	iii
<b>List of Tables</b> . . . . .	vi
<b>List of Figures</b> . . . . .	vii
<b>1. Introduction</b> . . . . .	1
1.1. Closed Geodesics on the Modular Surface . . . . .	1
1.2. The Main Theorem . . . . .	2
1.3. Strategy of the Proof . . . . .	3
1.3.1. Thermodynamic Formalism and Renewal Theorems . . . . .	3
1.3.2. A First Attempt via the Affine Sieve . . . . .	5
1.3.3. Beyond Expansion . . . . .	5
<b>2. Closed Geodesics and Dirichlet Forms</b> . . . . .	7
2.0.1. The Picard Group $PSL_2(\mathbb{Z}[i])$ . . . . .	7
2.0.2. Closed Geodesics . . . . .	8
2.0.3. Properties of Closed Geodesics . . . . .	8
2.0.4. Dirichlet Forms . . . . .	10
2.0.5. Fundamental Discriminants for Dirichlet Forms . . . . .	10
<b>3. A Symbolic Encoding of Geodesics</b> . . . . .	11
<b>4. Counting Geodesics via Thermodynamic Formalism</b> . . . . .	18
4.1. Properties of the Spectrum of $\mathcal{L}_s$ . . . . .	20
4.2. A Renewal Theorem for the Counting Function $N(a, X)$ . . . . .	21

4.3. An Asymptotic for Closed Geodesics . . . . .	23
<b>5. Counting Geodesics with Congruence Conditions . . . . .</b>	<b>26</b>
5.1. Bounding $M_z$ in the Supremum Norm . . . . .	27
5.1.1. Expressing $\mu_b$ as a convolution . . . . .	29
5.1.2. Expansion via Selberg's 3/16 Theorem . . . . .	35
5.1.3. Applying Theorem 12 . . . . .	39
5.2. Bounding $M_z$ in Variation . . . . .	40
5.2.1. Applying Theorem 12 Again . . . . .	41
5.3. Fourier-Walsh Decomposition . . . . .	42
5.3.1. Small Imaginary Part . . . . .	43
5.3.2. Large Imaginary Part . . . . .	43
5.4. Fourier Analysis of the Renewal Function . . . . .	44
5.5. The Finite Renewal Function . . . . .	46
<b>6. The Growth Parameter <math>\delta_R</math> . . . . .</b>	<b>48</b>
6.0.1. Perturbation estimates for large alphabets . . . . .	51
<b>7. Construction of the Multilinear Sifting Set . . . . .</b>	<b>53</b>
<b>8. Small Divisors . . . . .</b>	<b>56</b>
<b>9. Large Divisors . . . . .</b>	<b>59</b>
9.1. Spectral Theory of the Laplace Operator . . . . .	59
9.2. Character Sums . . . . .	62
9.3. Large Divisors . . . . .	63
<b>10. Sieve Theorem . . . . .</b>	<b>68</b>
<b>11. Final Estimates . . . . .</b>	<b>72</b>
<b>References . . . . .</b>	<b>75</b>

## List of Tables

3.1. Subshift Rules . . . . .	13
5.1. Admissible Generators . . . . .	31

## List of Figures

1.1. Height Cutoff $t < 2.5$ . . . . .	3
2.1. Fundamental Domain for $\mathrm{SL}_2(\mathbb{Z}[i]) \backslash \mathbb{H}^3$ . . . . .	8
3.1. The Fundamental Region $\mathcal{X}$ . . . . .	11
3.2. The Parts in $\mathcal{P}$ . . . . .	12
3.3. Parts in $P_4$ . . . . .	17
4.1. The Main 8 Regions . . . . .	19

# Chapter 1

## Introduction

### 1.1 Closed Geodesics on the Modular Surface

In [Duk88], Duke showed that closed geodesics on the modular surface equidistribute when grouped by discriminant, see also [CU04]. In [ELM<sup>+</sup>09] Einsiedler, Lindenstrauss, Michel, and Venkatesh gave a modern treatment of Linnik’s approach to this problem using the ergodic method. A key step is ruling out other potential weak-\* limits of closed geodesics. This raises the basic question on what other weak-\* limits could arise. Because the geodesic flow is a shift map (see Chapter 3) this question is trivial without more restrictions. They asked whether there is an infinite collection of closed geodesics having fundamental discriminant and being trapped in a compact subset of the modular surface. That is they don’t visit the cusp, or are “low-lying.” Bourgain and Kontorovich [BK17] showed an abundance of fundamental low-lying geodesics on the modular surface, answering the question above in a quantitative sense.

We will attack the corresponding problem in the Picard 3-manifold  $SL_2(\mathbb{Z}[i])\backslash\mathbb{H}^3$ . However, the solution is not as simple as applying the machinery from [BK17] to a ‘thin semi-group’ with well-established growth properties. For one, continued fractions in the complex plane are much more complicated to work with versus simple continued fractions on the real line. In particular, our symbolic encoding of closed geodesics does not display a semigroup structure because the shift map is restricted. We also have to develop in this setting much of the machinery (Chapters 3-6) which was already available to [BK17] for the modular surface.



## 1.2 The Main Theorem

We must establish some terminology before stating the main result. Consider the upper half space model of hyperbolic three-space:

$$\mathbb{H}^3 = \{z + tj : z \in \mathbb{C}, t \in \mathbb{R}_+\}$$

equipped with the line element

$$ds^2 = (dz_1^2 + dz_2^2 + dt^2)/t^2$$

where an element  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  in the group of isometries  $G = PSL_2\mathbb{C}$  acts by

$$z + tj \mapsto \frac{(az + b)\overline{(cz + d)}}{|cz + d|^2 + |c|^2t^2} + \left( \frac{t}{|cz + d|^2 + |c|^2t^2} \right) j.$$

When  $\mathbb{H}^3$  is embedded as a subset of Hamilton's quaternions, this expression simplifies as  $z + tj \mapsto (a(z + tj) + b)(c(z + tj) + d)^{-1}$ . Since  $G$  extends to a simply transitive action on the frame bundle  $F\mathbb{H}^3$ , we can identify an element in  $G$  with where it moves some representative reference frame. We can also identify  $\mathbb{H}^3 \leftrightarrow G/SU(2)$  and

$$T^1\mathbb{H}^3 \leftrightarrow G/SO(2).$$

Geodesic flow on  $T^1\mathbb{H}^3$  under this identification is represented by right-multiplication by the one parameter group generated by  $a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ . There is a correspondence between conjugacy classes of primitive hyperbolic matrices in  $\Gamma = SL_2(\mathbb{Z}[i])$  and closed geodesics.

To restrict to "low-lying" geodesics, we only consider those in the standard fundamental domain lying in a certain region  $\{z + tj : t < R\}$ . See for example, the figure below which depicts the standard fundamental domain for  $\Gamma \backslash \mathbb{H}^3$  and the cutoff  $t < 2.5$ .

Once we have developed a symbolic encoding of closed geodesics, it is trivial to manufacture infinitely many geodesics which are low-lying by enumerating periodic points (with restricted orbits) of a map analogous to the Gauss map for continued

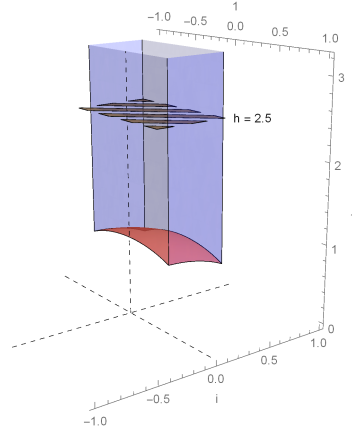


Figure 1.1: Height Cutoff  $t < 2.5$

fractions. Adding the condition that geodesics be fundamental places the requirement that the geodesic  $g$ , as a loxodromic element of  $\mathrm{SL}_2(\mathbb{Z}[i])$ , has  $\mathrm{tr}^2(g) - 4$  which satisfies certain conditions (see Section 2.0.5), a sufficient one being that it is square-free. If  $C_D$  is the set of distinct geodesics with discriminant  $D$  and  $N(\cdot)$  denotes the norm of a Gaussian integer, then we will show the following quantitative result:

**Theorem 1.** *For any  $\epsilon > 0$ , there is a compact region  $Y(\epsilon) \subset \Gamma \backslash \mathbb{H}^3$  and a set  $\mathcal{D}(\epsilon)$  of fundamental discriminants such that*

$$\#\{D \in \mathcal{D}(\epsilon) : N(D) < X\} \gg_{\epsilon} X^{1-\epsilon}, \quad X \rightarrow \infty$$

and for all  $D \in \mathcal{D}(\epsilon)$ ,

$$\#\{\gamma \in C_D : \gamma \subset Y(\epsilon)\} > |C_D|^{1-\epsilon}.$$

Of course, qualitatively, this solves the problem of producing infinitely many fundamental low-lying geodesics in the Picard 3-manifold.

### 1.3 Strategy of the Proof

Our three main tools come from symbolic dynamics, expander graphs, and sieve theory.

#### 1.3.1 Thermodynamic Formalism and Renewal Theorems

After converting the study of closed geodesics to periodic points of an analogue of the ‘Gauss map’, we are led to study the dynamics of a subshift of finite type. In particular,

define

$$\Sigma_R := \{(x_1, x_2, \dots) : x_i \in P_R, A_{x_i, A_{i+1}} = 1 \text{ for all } i\}$$

where  $P_R$  is some finite alphabet (see Chapter 3), and  $A$  is a  $P_R \times P_R$  binary matrix conveying transition rules (i.e.  $A_{x,y} = 1$  if  $x$  can be followed by  $y$  and  $A_{x,y} = 0$  otherwise). The shift map  $\sigma$  on  $\Sigma_R$  is defined as  $\sigma((x_1, x_2, \dots)) = (x_2, x_3, \dots)$ . Choosing the appropriate  $P_R$  and  $A$  following the work of [Pol94], gives an essentially one-to-one correspondence between closed geodesics lying in a compact set (corresponding to the choice of  $R$ ) and periodic points of  $\Sigma_R$  under  $\sigma$ . Denote the set of closed geodesics associated to  $(\Sigma_R, \sigma)$  via this correspondence as  $\Gamma_R \subset \mathrm{SL}_2(\mathbb{Z}[i])$  and

$$B_X := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}[i]) : \|a\|^2 + \|b\|^2 + \|c\|^2 + \|d\|^2 < X^2 \right\}.$$

The methods of Lalley in [Lal89] are straightforward to apply to our situation and lead to the following:

**Theorem 2.** *For fixed  $R \geq 3$ , there is a  $\delta_R \in (0, 2)$  so that*

$$\#(\Gamma_R \cap B_X) \asymp X^{2\delta_R}$$

as  $X \rightarrow \infty$ .

It will be crucial in a later argument that the set above is large.

**Theorem 3.** *The growth parameter given by Theorem 2 satisfies*

$$\lim_{R \rightarrow \infty} \delta_R \rightarrow 2.$$

Finally, we will need local information about closed geodesics. This is where expander graphs are used crucially. We follow the work of [BGS11] together with [BKM]. For  $q \in \mathbb{Z}[i]$ , set  $\mathrm{SL}_2(q) = \mathrm{SL}_2(\mathbb{Z}[i])/(q)$ .

**Theorem 4.** *For each  $R > 8$ , there is some absolute spectral gap  $\Theta_R > 0$  and absolute constants  $c_R, C_R > 0$  such that for all square-free  $q \in \mathbb{Z}[i]$  and  $\omega \in \mathrm{SL}_2(q)$  we have the estimate*

$$\left| \#\{g \in \Gamma_R \cap B_X : g \equiv \omega \pmod{q}\} - \frac{\#(\Gamma_R \cap B_X)}{|\mathrm{SL}_2(q)|} \right| \ll_R \#(\Gamma_R \cap B_X) E(q, X)$$

as  $X \rightarrow \infty$ , where

$$E(q, X) = \begin{cases} e^{-c_R \sqrt{\log X}} & N(q) < C_R \log X \\ N(q)^{C_R} X^{-\Theta_R} & N(q) > C_R \log X \end{cases}.$$

### 1.3.2 A First Attempt via the Affine Sieve

Theorem 4 allows us to follow the affine sieve procedure (see [BGS10], [SGS13], or [Kon14]) up to a level of distribution  $X^\alpha$  where the exponent  $\alpha_R = \Theta_R/C - o(1)$  and show for sufficiently large  $\delta_R$  that

$$\#\{g \in \Gamma_R \cap B_X : p \text{ a Gaussian prime, } p | (\text{tr}^2(g) - 4) \implies N(p) > X^{\alpha_R}\} \gg X^{2\delta_R - o(1)}.$$

Since  $\text{tr}^2(g) - 4$  factors as  $(\text{tr}(g) + 2)(\text{tr}(g) - 2)$ , any Gaussian prime dividing  $\text{tr}^2(g) - 4$  must divide one of its linear factors. This fact allows us to conclude from the “almost prime” estimate above that

$$\#\{g \in \Gamma_R \cap B_X : (\text{tr}^2(g) - 4) \text{ not square-free}\} \ll X^{4 - \alpha_R}.$$

See the proof of Theorem 24 on page 72 for details.

Comparing the two estimates, we would have our main result if only we could show that

$$2\delta_R > 4 - \Theta_R/C.$$

It seems that making  $\delta_R$  near 2 (as in Theorem 3) suffices. However, the shared dependence of  $\delta_R$  and  $\Theta_R$  on  $R$  cannot at this time be decoupled, so this attack fails, and we must use something more than expansion.

### 1.3.3 Beyond Expansion

To have stronger control on the exponent of distribution, we create bilinear (in fact, multilinear) forms, replacing  $\Gamma_R \cap B_X$  by a specially constructed subset  $\Pi$ , see Chapter 7. We analyze the set

$$\#\{\omega \in \Pi : \text{tr}^2(\omega) - 4 \equiv 0 \pmod{q}\}$$

via abelian harmonic analysis (on  $\mathbb{Z}[i]/(q)$ ). The characters of small order up to some intermediate level  $Q_0$  can be handled by expansion. The characters of larger order are

now dealt with by appealing to bilinear forms techniques, namely Cauchy-Schwarz and estimating exponential sums. Our methods allow us to sieve up to the absolute level of distribution  $\alpha = 1/16 - \epsilon$  from which Theorem 1 follows.

## Chapter 2

### Closed Geodesics and Dirichlet Forms

We express the upper half space model of hyperbolic 3-space as a subset of Hamilton's quaternions, i.e.

$$\mathbb{H}^3 = \{z + tj : z \in \mathbb{C}, t \in \mathbb{R}_+\}$$

where  $i^2 = j^2 = -1$  and  $ij = -ji$ . The group  $PSL_2(\mathbb{C})$  acts on  $\mathbb{H}^3$  by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot (z + tj) := (a(z + tj) + b)(c(z + tj) + d)^{-1}.$$

The inverse above should be interpreted as the Hamiltonian inverse. Considering the action of  $PSL_2(\mathbb{C})$  on the boundary  $\hat{\mathbb{C}}$  gives a correspondence with Möbius transformations.

#### 2.0.1 The Picard Group $PSL_2(\mathbb{Z}[i])$

Inside of  $G = PSL_2(\mathbb{C})$ , we have the discrete subgroup  $\Gamma = PSL_2(\mathbb{Z}[i])$ , sometimes referred to as the Picard group. We can express

$$\Gamma = \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \right\rangle$$

and this allows us to write the fundamental Dirichlet domain for the quotient  $\Gamma \backslash \mathbb{H}^3$  as

$$\mathcal{F} := \{(z, t) | \operatorname{Re}(z) \in [-1/2, 1/2], \operatorname{Im}(z) \in [0, 1/2], |z|^2 + t^2 > 1\}.$$

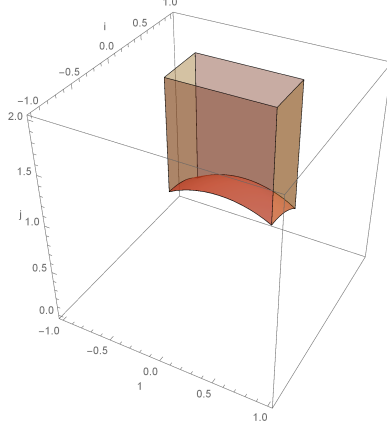


Figure 2.1: Fundamental Domain for  $SL_2(\mathbb{Z}[i]) \setminus \mathbb{H}^3$

### 2.0.2 Closed Geodesics

Write

$$M = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} : \theta \in [0, 2\pi) \right\}.$$

A geodesic which is closed in  $\Gamma \backslash SL_2(\mathbb{C})/M$  is identified by  $(\Gamma g M)a_l = (\Gamma g M)$ . Since  $a_l$  commutes with  $M$ , this translates to

$$g M a_l = A g M$$

for some  $A \in \Gamma$ ,  $g \in G$ . More explicitly, we may write

$$\begin{aligned} A &= g a_l m g^{-1} \\ &= g \begin{pmatrix} e^{l/2} & 0 \\ 0 & e^{-l/2} \end{pmatrix} \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} g^{-1} \end{aligned}$$

for some  $m \in M$ .

### 2.0.3 Properties of Closed Geodesics

- Trace and Eigenvalues:

By the equation above, we infer that  $A$  is diagonalized by  $g$  with eigenvalues  $\exp(\pm(l/2 + i\theta/2))$ . Following Beardon, we call  $A$  hyperbolic if  $\text{tr}^2(A) \in [4, \infty)$  and strictly loxodromic if  $\text{tr}^2(A) \notin [0, \infty)$ . The term loxodromic (referring to

transformations with 2 fixed points in  $\hat{\mathbb{C}}$ ) encompasses both. The trace of  $A$  is

$$\text{tr}(A) = 2 \cosh(l/2 + i\theta/2).$$

On the other hand, we can express the eigenvalue as

$$e^{l/2+i\theta/2} = \frac{\text{tr}(A) + \sqrt{\text{tr}^2(A) - 4}}{2},$$

and this gives us a way to determine the closed geodesic associated to an arbitrary loxodromic transformation in  $PSL_2(\mathbb{Z}[i])$ .

- Length and Primitivity:

The length of the geodesic is given by  $l$ . Note that any powers of  $A$  satisfy  $A^k = g(a_l m)^k g^{-1}$ , which would suggest that their length is  $kl$ . However,  $A^k = (ga_l m g^{-1}) \cdots (ga_l m g^{-1})$  indicates that we are traversing the same closed geodesic  $k$  times. So, we call  $A$  primitive when it is not a power of another element in the Picard group.

- Equivalent geodesics:

Technically, a closed geodesic is an element of the quotient  $SL_2(\mathbb{Z}[i]) \backslash SL_2(\mathbb{C}) / M$  satisfying  $(\Gamma g M) a_l = (\Gamma g M)$ . We chose a particular representative  $A$ , but any conjugation  $BAB^{-1}$  where  $B \in \Gamma$  would give the same geodesic. Hence, geodesics are equivalent if they are in the same  $\Gamma$  conjugacy class.

- Fixed points, visual points:

The loxodromic transformation  $A$  has two fixed points in  $\mathbb{C}$  which can be found by solving

$$z = \frac{az + b}{cz + d}.$$

In other words, the fixed points are roots of the (homogenized) binary quadratic form  $Q_A(1, z) = cz^2 + (d - a)z - b$  with coefficients in  $\mathbb{Z}[i]$ . Solving this, we get

$$\alpha = \frac{(a - d) + \sqrt{\text{tr}^2(A) - 4}}{2c}, \quad \bar{\alpha} = \frac{(a - d) - \sqrt{\text{tr}^2(A) - 4}}{2c}.$$

On the other hand, if we have  $g$  we can calculate

$$\lim_{l \rightarrow \infty} ga_l \cdot j = \alpha.$$



A similar statement (as  $l \rightarrow -\infty$ ) gives the reverse direction of the geodesic. These points  $\alpha, \bar{\alpha}$  are referred to as the visual points of the geodesic.

#### 2.0.4 Dirichlet Forms

Many of the observations above suggest a correspondence between closed geodesics and binary quadratic forms with coefficients in  $\mathbb{Z}[i]$ , also known as Dirichlet forms. We have

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto Q_A(x, y) = cx^2 + (d - a)xy - by^2$$

modulo the greatest common divisor of  $c$ ,  $d - a$ , and  $b$  and up to choice of unit. This correspondence is explained further in [Sar83]. We associate the discriminant of the Dirichlet form  $D_A = \text{tr}^2(A) - 4$  with the closed geodesic corresponding to  $A$ .

#### 2.0.5 Fundamental Discriminants for Dirichlet Forms

Note that a discriminant  $D$  of a Dirichlet form must be a square mod 4 and hence  $D \bmod 4 \in \{0, 1, -1, 2i\}$ . Moreover, each  $D \in \mathbb{Z}[i]$  with a square residue mod 4 is a discriminant of some form. We call a discriminant  $D$  fundamental if it cannot be expressed as  $D = q^2 D_0$  where  $q$  is a non-unit and  $D_0$  is also a discriminant. This is equivalent to another other common definition which states that  $D$  is fundamental if any form  $Q(x, y) = ax^2 + bxy + cy^2$  with discriminant  $D$  must be primitive (i.e.  $(a, b, c) = 1$ .) Note that this also agrees with the work of Hilbert, i.e. that  $D$  is fundamental if and only if  $|D|$  is the relative discriminant of the extension of  $\mathbb{Q}(i, \alpha)$  over  $\mathbb{Q}(i)$ .

A geodesic is fundamental if its associated discriminant is fundamental. We will sieve down to geodesics with square-free discriminant, only catching the  $D = \pm 1 \bmod 4$  case.

## Chapter 3

### A Symbolic Encoding of Geodesics

We now describe the map analyzed by Pollicott in [Pol94]. Denote a circle of radius 1 about a center  $z$  as  $C(z)$ . Consider the region  $\mathcal{X}$  in  $\mathbb{H}^2$  exterior to the three circles  $C(i)$ ,  $C(1)$ , and  $C(-1)$  where we have the removed vertical lines  $\{z : \Im(z) = k/2 : k \in \mathbb{Z}\}$  and horizontal lines  $\{z : \Re z = k/2 : k \in \mathbb{N}\}$ . The region  $\mathcal{X}$  is shown in blue in the figure below.

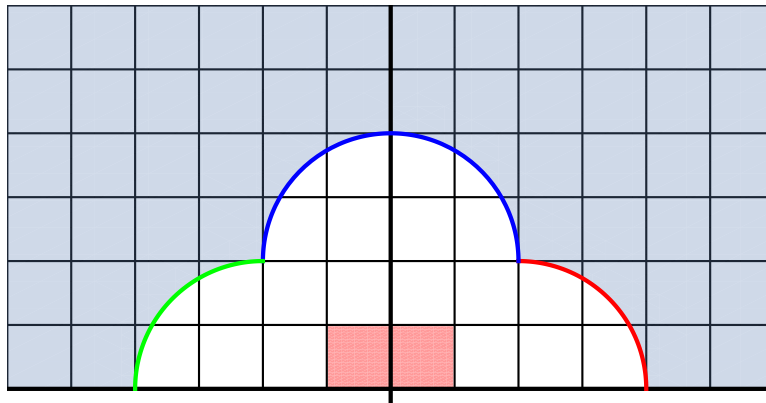


Figure 3.1: The Fundamental Region  $\mathcal{X}$

For each  $z \in \mathcal{X}$  there is a unique closest (in Euclidean distance) Gaussian integer, which we denote  $\lfloor z \rfloor$ . The map  $\hat{f}$  will move  $z$  to the unit square centered about the origin by first subtracting  $\lfloor z \rfloor$ . If  $\Im(z - \lfloor z \rfloor) < 0$ , then  $f$  rotates about the origin by  $\pi$ . Finally, the involution  $\omega \mapsto -1/\omega$  is performed to return to  $\mathcal{X}$ . Formally, we define  $\hat{f}$  as

$$\hat{f}(z) = \begin{cases} \frac{(-1)}{z - \lfloor z \rfloor} & \Im(z) > \Im(\lfloor z \rfloor) \\ \frac{1}{z - \lfloor z \rfloor} & \Im(z) < \Im(\lfloor z \rfloor) \end{cases}.$$

Define  $S : z \mapsto -1/z$  and  $f := S \circ \hat{f} \circ S$  on  $S(\mathcal{X})$ . Note that  $S(\mathcal{X})$ , the image of  $\mathcal{X}$  under  $S$ , is contained within  $\{x + iy : |x| < 1/2, 0 < y < 1/2\}$ . Pollicott proves the

following:

**Theorem 5.** *There is a bijection between closed geodesics and the periodic points of  $f$ .*

*Proof.* Note that each closed geodesic has a unique representative with visual point in  $S(\mathcal{X})$ . The action of  $f$  can be seen as a fractional linear transformation in  $\mathrm{SL}_2(\mathbb{Z}[i])$  extends naturally to the geodesic and gives an  $\mathrm{SL}_2(\mathbb{Z}[i])$  conjugate. If the orbit of the visual point is periodic then the geodesic is  $\mathrm{SL}_2(\mathbb{Z}[i])$ -equivalent to itself.  $\square$

Furthermore, [Pol94] shows that  $(\mathcal{X}, f)$  admits a Markov partition. This will allow us to study the system via a simplified encoding. The naive choice of partition of  $\mathcal{X}$  into connected regions between the grid lines is almost correct. We must introduce two more circles,  $C(1+i)$  and  $C(-1+i)$  and separate each region intersecting the boundary of either of the circles. The figure below illustrates the partition which we denote  $\mathcal{P}$ . Each connected region after removing  $C(1+i)$ ,  $C(-1+i)$  and the  $1/2$ -spaced grid is a part in the partition.

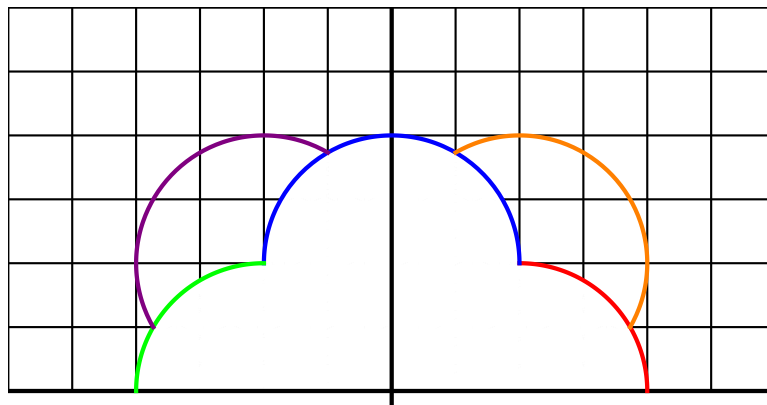


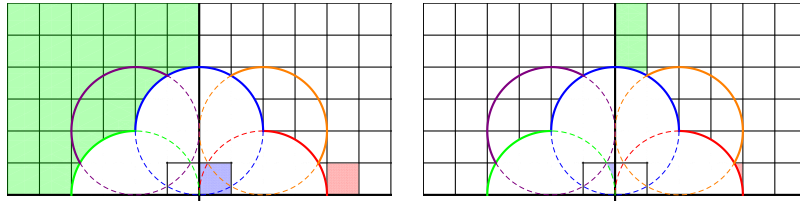
Figure 3.2: The Parts in  $\mathcal{P}$

In Table 3.1, we pictorially reproduce the proof given by Pollicott that  $\mathcal{P}$  is a Markov partition for  $(X, \hat{f})$ . We will divide into cases depending on the closest Gaussian integer. Due to symmetry, we only show Gaussian integers  $z$  with  $\Re(z) \geq 0$ .

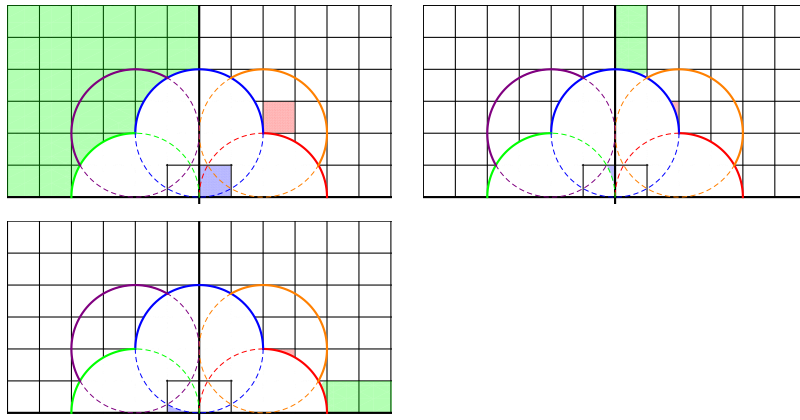
Table 3.1: Subshift Rules

For each part shaded pink, observe that its image shown in green under  $\hat{f}$  is a union of parts in  $\mathcal{P}$ . The blue region within  $S(\mathcal{X})$  is the result of the translation of the pink region before the involution map  $S$  is applied.

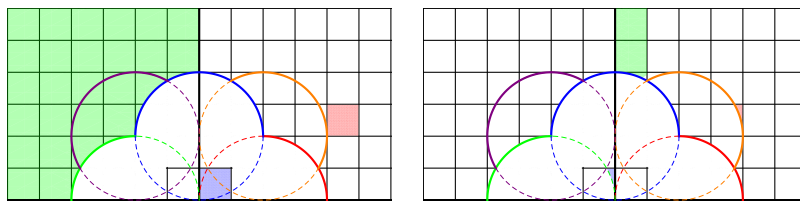
Parts surrounding 2

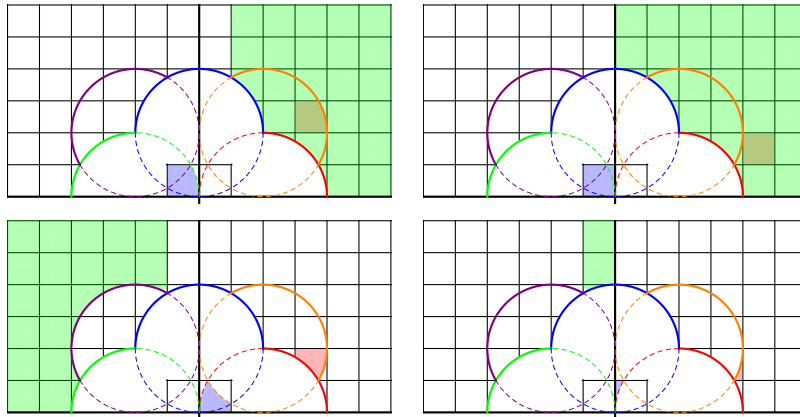
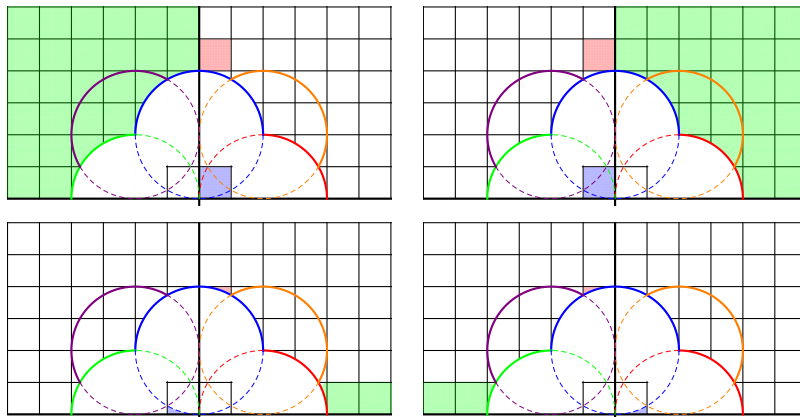
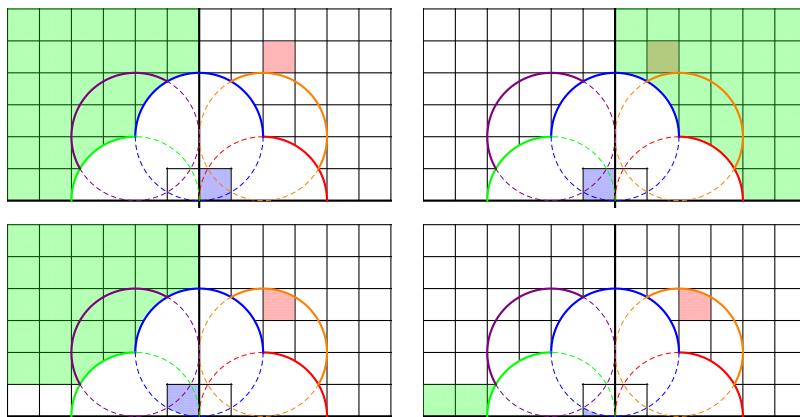


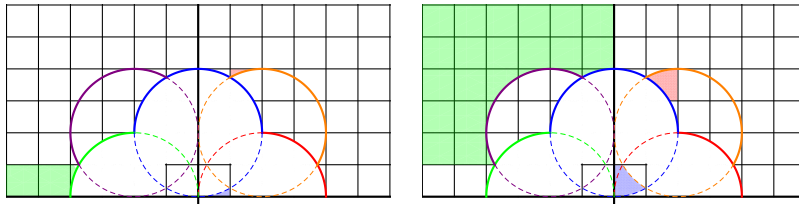
Parts surrounding  $1 + i$



Parts surrounding  $2 + i$

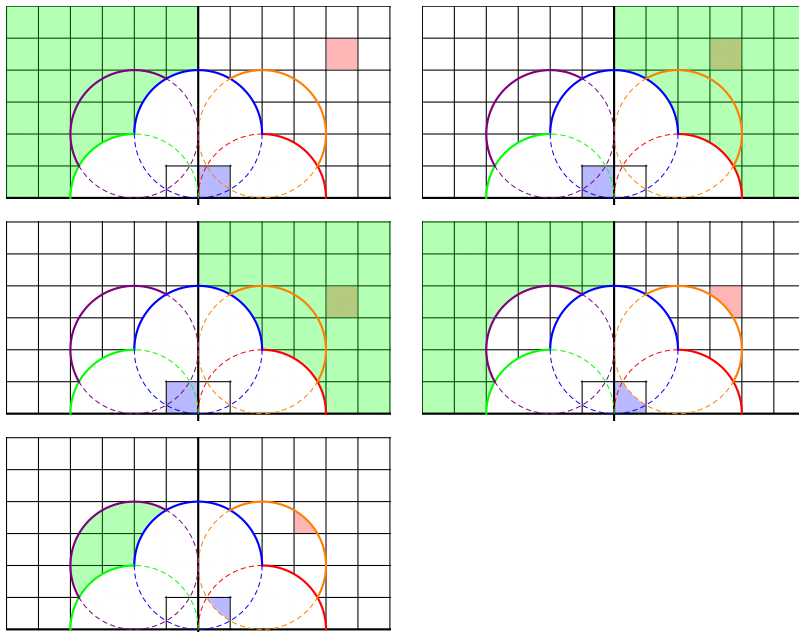


Parts surrounding  $2i$ Parts surrounding  $1 + 2i$ 



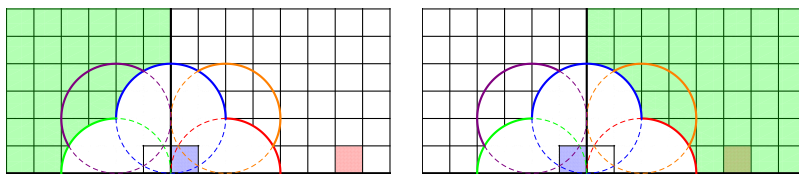
Parts surrounding  $2 + 2i$

---



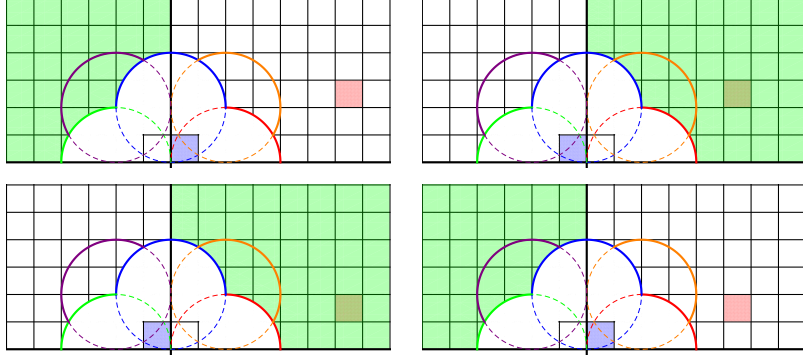
Parts surrounding  $z \in \mathbb{Z}$ , with  $|z| > 2$

---



Parts surrounding  $z \in \mathbb{Z}[i]$  with either  $\Re z$  or  $\Im z$  greater than 2

---



For each part  $\rho \in S(\mathcal{P})$ , we associate a distinct label  $p$ . The set of all labels is the alphabet for our shift map. Since  $S(\mathcal{P})$  is a Markov partition,  $(\Sigma, \sigma)$  gives a symbolic representation of  $(S(\mathcal{X}), f)$  where

$$\Sigma := \{(p_1, p_2, \dots) : p_i \leftrightarrow \rho_i \in S(\mathcal{P}), f(\rho_i) \supset \rho_{i+1}\}$$

and  $\sigma$  is the shift operator, i.e.  $\sigma(p_1, p_2, p_3, \dots) = (p_2, p_3, \dots)$ . More precisely, for any finite admissible word  $a = (p_1 \dots p_n)$  (meaning  $a$  occurs as a subword of some  $\alpha \in \Sigma$ ) we define the cylinder set  $C_a$  as

$$C_a = \bigcap_{k=1}^n f^{-k}(\rho_k).$$

Since  $|f'(z)| \leq \frac{1}{2}$ , we have that the diameter of  $C_a$  is at most  $\frac{1}{2^n}$ . We also have  $C_{(p_1, \dots, p_n)} \subset C_{(p_1, \dots, p_{n-1})}$ . Therefore the map

$$\begin{aligned} \pi : \Sigma &\rightarrow S(\mathcal{X}) \\ (p_1, p_2, \dots) &\mapsto \bigcap_{k=1}^{\infty} C_{(p_1, \dots, p_k)} \end{aligned}$$

is well-defined. Since the interiors of distinct cylinders of length  $n$  (i.e. the cylinder defined on a word of length  $n$ ) are disjoint,  $\pi$  is one-to-one. The image of  $\pi$  is all of  $S(\mathcal{X})$  up to a set of Lebesgue measure 0 (the orbit of the grid under  $f$  must be removed) and the following diagram commutes:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \pi \downarrow & & \downarrow \pi \\ S(\mathcal{X}) & \xrightarrow{f} & S(\mathcal{X}) \end{array}$$

In particular, we have a bijection between the periodic points of  $\sigma$  and the periodic points of  $f$ . This allows us to study closed geodesics by analyzing the system  $(\Sigma, \sigma)$ .

From the correspondence between periodic orbits in  $(S(\mathcal{X}), f)$  and closed geodesics, we can measure how close a geodesic is to the cusp in  $\mathrm{SL}_2(\mathbb{Z}[i]) \backslash \mathrm{SL}_2(\mathbb{C})$  by determining the minimum distance from origin to a point in the orbit of the associated periodic point. So define  $\mathcal{P}_R \subset \mathcal{P}$  by only allowing parts from  $\mathcal{P}$  if they lie completely inside the ball of radius  $R$  centered at the origin. Recall that the parts  $\mathcal{P}$  were defined for the conjugate system  $(\mathcal{X}, \hat{f})$  and so the images of parts in  $\mathcal{P}_R$  under  $S$  fall outside a small ball centered at the origin. The figure below has the parts included in  $\mathcal{P}_4$  shaded.

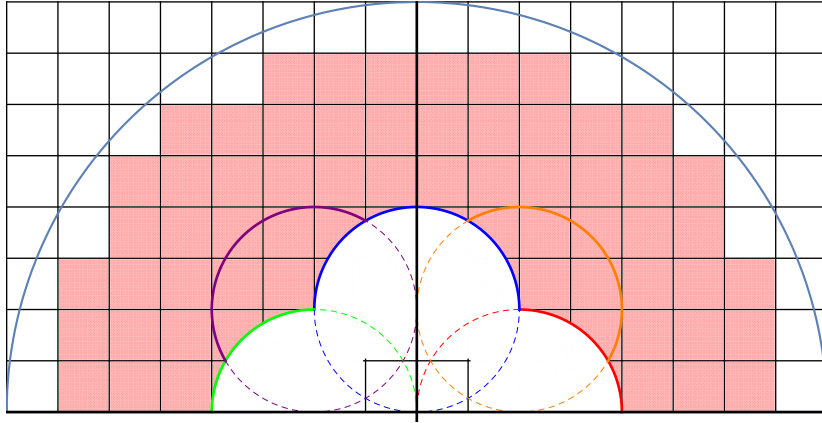


Figure 3.3: Parts in  $\mathcal{P}_4$

Our corresponding symbolic encoding is

$$\Sigma_R := \{(p_1, p_2, \dots) : p_i \leftrightarrow \rho_i \in \mathcal{P}_R, f(\rho_i) \supset \rho_{i+1}\}.$$

The new system  $(\Sigma_R, \sigma)$  is now a subshift of finite type as its alphabet is finite. The problem of counting low-lying geodesics translates into a problem of counting periodic points of  $\sigma$  in  $\Sigma_R$ .



## Chapter 4

### Counting Geodesics via Thermodynamic Formalism

The asymptotic for geodesics derived in this chapter is a straightforward application of the work of Lalley in [Lal89]. We summarize the ideas which give rise to the method in the next section towards an asymptotic for geodesics counted by congruence classes.

For  $a \in \Sigma$ , we define the distortion function  $\tau(a) := \log |f'(\pi(a))|$ . The  $N$ -th Birkhoff sum for the distortion function is then  $S_N \tau(a) := \sum_{k=0}^{N-1} \tau(\sigma^k a)$  where  $S_0(a) = 0$ . If  $\sigma^N(a) = a$  and in particular  $a$  is a visual point corresponding to a closed geodesic, then  $S_N \tau(a) = -l$  where  $l$  is the length of the geodesic. This gives some indication that the following function will be useful in counting geodesics. For  $T > 0$  and  $a \in \Sigma_R$  Lalley's renewal function is defined as

$$N(T, a) := \sum_{k=0}^{\infty} \sum_{b: \sigma^k(b)=a} g(b) \mathbf{1}_{\{S_k \tau(b) \leq T\}}.$$

Partitioning the sum by the preimage  $\sigma^{-1}(a)$ , we arrive at the following recursive relation, known as the renewal equation

$$N(T, a) = g(a) \mathbf{1}_{\{a \leq T\}} + \sum_{b: \sigma(b)=a} N(T - \tau(b), b).$$

In order to analyze the renewal function, we are led to study its Laplace transform and a certain transfer operator.

For a continuous function  $f$  defined on  $\Sigma_R$  and  $0 < \rho < 1$ , we define

$$\begin{aligned} \text{var}_n(f) &:= \sup\{|f(a) - f(b)| : a_i = b_i \text{ for all } 0 \leq i \leq n\}, \\ |f|_\rho &:= \sup_{n \geq 0} \frac{\text{var}_n(f)}{\rho^n}. \end{aligned}$$

Then  $\mathcal{F}_\rho := \{f : |f|_\rho < \infty\}$  is the space of Hölder continuous functions which is a Banach space with norm  $\|\cdot\|_\rho = |\cdot|_\rho + \|\cdot\|_\infty$ . The transfer operator, depending on

$s \in \mathbb{C}$ , acts on  $\mathcal{F}_\rho$  as follows:

$$\mathcal{L}_s f(a) := \sum_{\substack{b \\ \sigma(b)=a}} e^{s\tau(b)} f(b).$$

First note that  $\mathcal{L}_s f(a)$  is a bounded linear operator. Additionally, when  $s \in \mathbb{R}$  the coefficients in the sum are all positive. For real  $s$ , would like to compare the spectrum of  $\mathcal{L}_s$  to that of an positive matrix. In particular, we would like to apply an analogue of the Perron-Frobenius theorem. In order to do so, we must establish a few more properties of the system  $(\Sigma_R, \sigma)$ .

Recall that a matrix  $A$  is irreducible if for each position  $(i, j)$  there is some power of  $A$  such that the  $(i, j)$ -th entry is positive. In analogy, we say that  $(\Sigma_R, \sigma)$  is irreducible if for each two states  $p_1, p_2 \in \mathcal{P}_R$  there is some finite admissible word beginning with  $p_1$  and ending in  $p_2$ .

**Lemma 1.**  $(\Sigma_R, \sigma)$  is irreducible as long as  $R \geq 4$ .

*Proof.* By our definition of  $\Sigma_R$ , a finite subword of the form  $(p_{i_1}, p_{i_2}, \dots, p_{i_n})$  must satisfy  $\hat{f}(p_{i_k}) \supset p_{i_{k+1}}$ . In other words, we must show that for each part in  $p \in \mathcal{P}_R$  that there is some  $k$  such that  $\mathcal{P}_R \subset \hat{f}^n(p)$ .

One may recall from the diagrams in the previous section that for any part  $p \in \mathcal{P}_R$ , the image under  $\hat{f}$  contains at least one of the following regions (1) – (8) : Each region

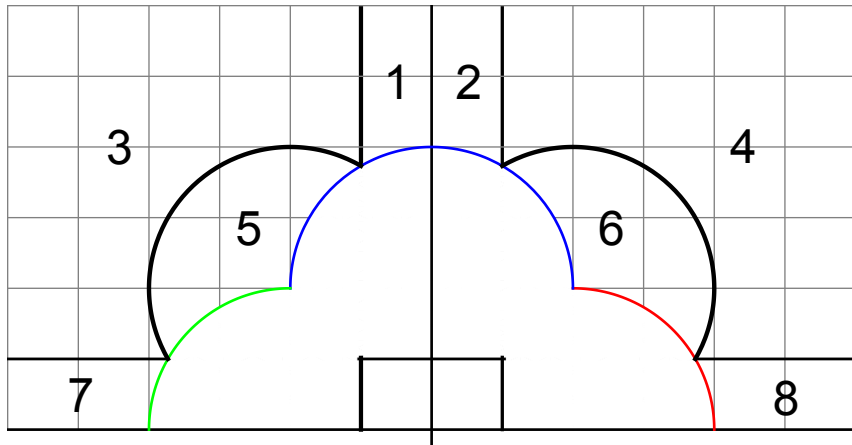


Figure 4.1: The Main 8 Regions

contains at least one square part. The image of a square part under  $\hat{f}$  is either the

union of the even regions or the odd regions. In either case, the next iterate of  $\hat{f}$  is the union of all 8 regions. Hence  $\mathcal{P}_R \subset \hat{f}^3(p)$  for any  $p$ .  $\square$

A state  $p \in \mathcal{P}_R$  is periodic of period  $k$  if any finite admissible word beginning and ending in  $p$  must have length divisible by  $k$ . The period of the system  $(\Sigma_R, \sigma)$  is the greatest common factor of the periods of all of its states and a system is said to be aperiodic if this greatest common factor is 1.

**Lemma 2.**  $(\Sigma_R, \sigma)$  is aperiodic as long as  $R \geq 4$ .

*Proof.* Since we already established irreducibility, we only need to show that one state is aperiodic. Take the square part to the left of 3, i.e.  $p = \{x+iy : x \in (2.5, 3), y \in (0, .5)\}$ . Its image under  $\hat{f}$  is the union of the even regions (refer to the previous figure) and hence  $p \subset \hat{f}(p)$ . Therefore, the orbit of a point in  $p$  may return to  $p$  after any number of iterates of  $\hat{f}$ .  $\square$

#### 4.1 Properties of the Spectrum of $\mathcal{L}_s$

We are now in a position to cite Ruelle's Perron-Frobenius theorem (see [PP90] for proof):

**Theorem 6.** For  $s \in \mathbb{R}$  the spectrum of  $\mathcal{L}_s$  has the following properties:

1.  $\mathcal{L}_s$  has a simple maximal positive eigenvalue  $\lambda_s$  with corresponding eigenfunction  $h_s$  which can be chosen to be positive.
2. The remainder of the spectrum is contained in a disc of radius less than  $\lambda_s$ .
3. There is a unique probability measure  $\mu_s$  on  $\Sigma_R$  such that  $\mathcal{L}_s^* \mu_s = \lambda_s \mu_s$ .
4.  $\frac{1}{\lambda_s^n} \mathcal{L}_s^n v \rightarrow h_s \int v d\mu_s$  uniformly for all continuous  $v$  if  $h_s$  is normalized so that  $\int h_s d\mu_s = 1$ .

Consider  $\mathcal{L}_s$  as a family in  $s \in \mathbb{R}$  and define the pressure functional as  $P(s) = \log \lambda_s$ . The pressure is increasing in  $s$  and there is a unique solution  $s = \delta_R$  to  $P(s) = 0$ . Since  $\delta_R$  is featured in our main asymptotic, it will be necessary to determine the dependence of  $\delta_R$  on  $R$  in a later section. Our next step is to consider the family  $\mathcal{L}_s$  for  $s \in \mathbb{C} - \mathbb{R}$ .

**Theorem 7.** For  $s \in \mathbb{C} - \mathbb{R}$ , the spectrum of  $\mathcal{L}_s$  is contained in a disc centered at zero of radius  $\lambda_{\Re s}$  (the maximal eigenvalue of  $\mathcal{L}_{\Re s}$ ).

*Proof.* We summarize the proof provided by Lalley in section 11 of [Lal89].

There are two cases: either  $\tau$  is lattice in which case the spectral radius of  $\mathcal{L}_s$  is strictly smaller than the radius on the real axis or it is nonlattice in which case  $\mathcal{L}_s$  has spectral radius equal to  $\lambda_{\Re s}$  at regularly spaced intervals along the vertical line  $\Re z = \Re s$ . See Lalley for the precise definition of a lattice function. In order to show that  $\tau$  is non-lattice we relate it to a function known to be non-lattice.

$\tau$  is cohomologous to  $g$ , a height function on suspension of restricted geodesic flow.

It suffices to show that for periodic  $x \in \Sigma$  of period  $n$ , i.e.  $\sigma^n(x) = x$ , we have  $S_n\tau(x) = S_n g(x)$ . If  $k$  is the period of  $x$ , we will show that both  $n$ th Birkhoff sums give  $n/k$  times the length of the closed geodesic associated with  $x$ :

$$\begin{aligned} t^n(M) &= d_H(j, M.j) - d_H(j, \sigma(M).j) \\ &= d_H(j, M.j) - d_H(j, j) \\ &= d_H(j, ga^l g^{-1}.j) \\ &= d_H(g.j, ga^l.j) = l. \end{aligned}$$

On the other hand,  $M$  also corresponds to a mobius function  $m$ . Say  $\alpha$  is the fixed point of  $m$ .

$$m'(\alpha) = \frac{1}{(c\alpha + d)^2} = \frac{1}{(c(e^{l/2} - d)/c + d)^2} = e^{-l}.$$

So  $\log |m'(\alpha)| = \tau^n(\alpha) = l$ .

If  $g$  is lattice then the suspension flow is not mixing for any invariant measure, but [Rud82] proves otherwise. □

## 4.2 A Renewal Theorem for the Counting Function $N(a, X)$

Perturbation estimates ( $|\mathcal{L}_s - \mathcal{L}_z|$  for  $|s - z| < \epsilon$ ) imply that the eigenvalue map  $s \mapsto \lambda_s$ , the lead eigenfunction map  $s \mapsto h_s$  and the invariant measure map  $s \mapsto \mu_s$  are all holomorphic functions in a small neighborhood of  $\delta_R$ . This leads to the local decomposition

$(1 - \mathcal{L}_s)^{-1} = \frac{h_s}{1 - \lambda_s} \mu_s + (1 - \mathcal{L}'_s)^{-1}$  where  $(1 - \mathcal{L}'_s)^{-1}$  is a holomorphic family of bounded operators. The first term in the local decomposition about  $\delta_R$  contributes the main term in Lalley's estimate:

**Theorem 8.** *For any  $a \in \Sigma_R$ ,*

$$N(T, a) = h_{\delta_R}(a)e^{T\delta_R} + o(e^{T\delta_R}).$$

*Proof.* We summarize the proof, ignoring issues of convergence which are addressed in section of 8 of Lalley. First, define the following Laplace transform of the renewal function

$$F(s, a) := \int_{-\infty}^{\infty} e^{Ts} N(T, a) dT$$

and input the renewal equation to get

$$\begin{aligned} F(s, a) &= \int_0^{\infty} g(a)e^{Ts} dT + \sum_{b:\sigma(b)=b} \int_{-\infty}^{\infty} e^{Ts} N(T - \tau(b), b) dT \\ &= \frac{g(a)}{1 - e^s} + \sum_{b:\sigma(b)=b} \int_{-\infty}^{\infty} e^{Ts} e^{-s\tau(b)} N(T, b) dT \\ &= \frac{g(a)}{1 - e^s} + \mathcal{L}_s F(s, a). \end{aligned}$$

So  $I - \mathcal{L}_s$  applied to the Laplace transform of the renewal function gives  $g(a)(1 - e^s)^{-1}$ . Now it is clear how information about the spectrum of  $\mathcal{L}_s$  yields information about the renewal function. In particular, where the resolvent exists we have

$$(1 - e^s)^{-1}(I - \mathcal{L}_s)^{-1}g(a) = F(s, a).$$

Theorems 6 and 7 and the decomposition of  $(1 - \mathcal{L}_s)^{-1}$  in a neighborhood of  $\delta_R$  imply

$$F(s, a) = \frac{C(a)}{z} + G(s, a)$$

where  $G(s, a)$  is holomorphic in  $\Re(s) \geq \delta_R$ . Integrating a smoothed version of  $F(s, a)$  along a vertical line  $\Re s > \delta_R$  and pulling the contour to  $\delta_R$  we get that the pole in  $F(s, a)$  contributes the main term

$$e^{-T\delta} N(T, a) \rightarrow C(a)$$

as  $T \rightarrow \infty$ . □

### 4.3 An Asymptotic for Closed Geodesics

Finally, we relate dynamics on  $(\Sigma_R, \sigma)$  to a geodesic count. Define the set of finite admissible words (plus the empty word  $\emptyset$ ) as

$$\Sigma^* = \{\emptyset\} \cup \{(p_1, \dots, p_n) : p_i \leftrightarrow \rho_i \in \mathcal{P}_R, f(\rho_i) \supset \rho_{i+1}\}.$$

There is a bijection between aperiodic words in  $\Sigma^*$  and closed geodesics. Specifically, to a closed geodesic we associate the element in  $\Sigma_R^*$  that corresponds to the period of an orbit under  $\sigma$ . A periodic word in  $\Sigma^*$  corresponds to the a geodesic traversed multiple times. We define  $\tau^*(a) = d(j, g_a j) - d(j, g_{\sigma(a)} j)$  where  $d(P, P)$  denotes hyperbolic distance between any two points in  $\mathbb{H}^3$  and  $g_a \in \mathrm{SL}_2(\mathbb{Z}[i])$  is the local definition of  $f$  restricted to  $C_a$ . The identity  $2 \cosh d(j, g.j) = \|g\|^2$  (see [EGM13] for a proof specific to  $\mathbb{H}^3$ ) will eventually lead us to the final counts for matrices in a norm ball. The shift operator extends in a natural way to act on  $\Sigma^*$  however we need to resolve the ambiguity of  $\sigma^k$  for words of length less than  $k$ :

$$\begin{aligned} \sigma(p_1, p_2, \dots, p_n) &= (p_2, \dots, p_n) \\ \sigma(p_i) &= \emptyset \\ \sigma(\emptyset) &= \emptyset. \end{aligned}$$

We are now ready to define the finite version of the renewal function

$$N^*(T, a) := \sum_{k=0}^{\infty} \sum_{\substack{b: \sigma^k(b)=a \\ b \neq \emptyset}} g(y) \mathbf{1}_{\{S_n \tau^*(b) \leq T\}}.$$

$N^*(T, a)$  satisfies a renewal equation similar to  $N(T, a)$  and so we are tempted to treat the finite renewal function analogously. However, finding an appropriate Banach space of functions for the transfer operators  $\mathcal{L}_s^*$  to act on is elusive.

In order to model  $\Sigma_R^*$  after  $\Sigma_R$  we introduce a new state 0 and for any  $(p_1, \dots, p_k) \in \Sigma_R^*$  append an infinite tail of 0's to achieve an infinite word. The empty word maps to the infinite string of zeros and the action of  $\sigma$  is well-defined between the finite model of  $\Sigma^*$  and the infinite one. The space  $\mathcal{F}_\rho(\Sigma_R^* \cup \Sigma_R)$  of Hölder continuous functions satisfies the same properties (with the same norm) as previously. However, the addition of the

new ‘0’-state means that the system  $(\Sigma_R^* \cup \Sigma_R, \sigma)$  is no longer irreducible and hence Ruelle’s Perron-Frobenius theorem does not immediately apply to the spectrum of the transfer operators defined on  $\mathcal{F}_\rho(\Sigma_R^* \cup \Sigma_R)$ . Since we are only after an asymptotic for  $N^*(T, a)$  and long words in  $\Sigma_R^*$  may be approximated reasonably well (in the product topology) by words in  $\Sigma_R$ , we can use what we have already shown about the transfer operators on  $\mathcal{F}_\rho(\Sigma_R)$  to prove the following:

**Theorem 9.** *If  $h_s^*$  is the leading eigenfunction of  $\mathcal{L}_s^*$  then*

$$N^*(T, a) = h_{\delta_R}^*(a)e^{T\delta_R} + o(e^{T\delta_R}).$$

*Proof.* Our first claim is that for  $a \in \Sigma_R^*$  and  $b \in \Sigma_R$  which are in the same  $N$ -cylinder in  $\Sigma_R^* \cup \Sigma_R$  (i.e.  $a_i = b_i$  for  $1 \leq i \leq N$ ) and for  $k \ll N$

$$S_k \tau^*(a) = S_k \tau(b) + O(2^{-N+k})$$

Let  $\bar{a} \in \Sigma_R$  be the periodic word with period  $a_1, \dots, a_N$ . We claim that  $S_k \tau^*(a) = S_k \tau(\bar{a})$ . Recall from the proof of Theorem 7 that for any geodesic with period  $c \in \Sigma_R^*$ ,  $S_l \tau^*(c) = S_l \tau(\bar{c})$  where  $l$  is the length of  $c$ . We can also write

$$\begin{aligned} S_k \tau^*(a) &= S_N \tau^*(a) - S_{N-k} \tau^*(\sigma^k(a)) \\ &= S_N \tau(\bar{a}) - S_{N-k} \tau(\overline{\sigma^k(a)}) \\ &= \left( S_k \tau(\bar{a}) + S_{N-k} \tau(\overline{\sigma^k(a)}) \right) - S_{N-k} \tau(\overline{\sigma^k(a)}) \\ &= S_k \tau(\bar{a}) \end{aligned}$$

On the other hand, by bounded distortion (see Lemma 3 which is more easily understood with the notation in the next section)  $S_k \tau(b) = S_k \tau(\bar{a}) + O(2^{-N+k})$ . This proves the first claim.

From the previous claim it follows that for  $a \in \Sigma_R^*$  and  $b \in \Sigma_R$  which are in the same  $N$ -cylinder and  $k \ll N$

$$N(T - 2^{-N+k}, b) \leq N^*(T, a) \leq N(T + 2^{-N+k}, b)$$

for  $g \equiv 1$ .

Then iterating the renewal equation  $N$  times, one has

$$N^*(T, a) = \sum_{\substack{b': \sigma^N(b')=a \\ \sigma^{N-1}(b') \neq \emptyset}} N^*(T - S_N \tau^*(b'), b') \\ + \sum_{k=1}^{N-1} \sum_{\substack{b: \sigma^k(b)=a \\ \sigma^{N-1}(b) \neq \emptyset}} g(b) \mathbf{1}_{\{S_N \tau^*(b) \leq T\}} + g(a) \mathbf{1}_{\{T \geq 0\}}.$$

As  $T \rightarrow \infty$  second line does not change. For each summand in the first line we can find  $b' \in \Sigma_R$  to sandwich between the two terms

$$N(T \pm 2^{-N+k} - S_N \tau^*(b'), b) \asymp h(b) e^{(T - S_N \tau^*(b')) \delta_R} e^{\pm \delta_R 2^{-N+k}}$$

Send  $N \rightarrow \infty$  and use the continuity of  $h$  to get the statement of the theorem.  $\square$

If  $h_s$  is the leading eigenfunction for  $\mathcal{L}_s$  the transfer operator in  $(\Sigma_R, \sigma)$  then the leading eigenfunction  $h_s^*$  of  $\mathcal{L}_s^*$  agrees with  $h_s$  on  $\Sigma_R$ . This is Lemma 6.1 of Lalley.

The theorem follows from a sandwiching argument of the renewal function  $N^*(T, x)$  between  $N(T, x)$  with appropriate parameters. A similar argument will appear in the next section.

Combining the asymptotic for  $N^*(\sqrt{2 \cosh T}, \emptyset)$  with the identity  $2 \cosh d(j, gj) = \|g\|^2$  yields Theorem 2 which we restate here:

**Theorem 10.** *For fixed  $R \geq 3$ , there is a  $\delta_R \in (0, 2)$  so that*

$$\#(\Gamma_R \cap B_X) \asymp X^{2\delta_R}$$

as  $X \rightarrow \infty$ .



## Chapter 5

### Counting Geodesics with Congruence Conditions

Here we combine the work of Bourgain, Gamburd, Sarnak in [BGS11] with the expansion idea of Bourgain, Kontorovich, Magee in [BKM] to estimate

$$\#\{g \in \Gamma_R \cap B_X : g \equiv \omega \pmod{q}\}$$

for some  $\omega \in \mathrm{SL}_2(q)$  (recall that  $\mathrm{SL}_2(q) := \mathrm{SL}_2(\mathbb{Z}[i])/(q)$ ).

In order to detect congruence classes we introduce a new space of functions  $\mathcal{F}_\rho(\Sigma_R \times \mathrm{SL}_2(q))$ . If  $f$ , defined on  $\Sigma_R \times \mathrm{SL}_2(q)$ , is continuous in each variable, we can define

$$\begin{aligned} \|f\|_\infty &:= \sup_x \left( \sum_{g \in \mathrm{SL}_2(q)} |f(x, g)|^2 \right)^{1/2}, \\ \mathrm{var}_n f &:= \sup \left\{ \left( \sum_{g \in \mathrm{SL}_2(q)} |f(a, g) - f(b, g)|^2 \right)^{1/2} : a_i = b_i \text{ for all } 0 \leq i \leq n \right\}, \\ |f|_\rho &:= \sup_n \frac{\mathrm{var}_n f}{\rho^n}. \end{aligned}$$

Then  $\mathcal{F}_\rho(\Sigma_R \times \mathrm{SL}_2(q)) = \{f \in C(\Sigma_R \times \mathrm{SL}_2(q)) : \|f\|_\infty < \infty, |f|_\rho < \infty\}$  with norm  $\|\cdot\|_\rho = \|\cdot\|_\infty + |\cdot|_\rho$  is a Banach space.

Before defining the action of the transfer operators on this space, we must explain how the action of  $(\Sigma_R, \sigma)$  extends to  $\mathrm{SL}_2(q)$ . Recall the definition of  $\hat{f}$  in  $(\mathcal{X}, \hat{f})$ :

$$\hat{f}(z) = \begin{cases} \frac{-1}{z-|z|} & \Im(z) > \Im(|z|) \\ \frac{1}{z-|z|} & \Im(z) < \Im(|z|) \end{cases}.$$

Locally, i.e. when restricting to the interior of a part  $p \in \mathcal{P}$ , we may represent the action of  $\hat{f}$  as a fractional linear transformation:

$$\hat{f}|_p = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}^j \begin{pmatrix} 1 & -|z| \\ 0 & 1 \end{pmatrix}$$

where  $j \in \{0, 1\}$  reflects whether rotation is necessary for the image to be in the upper half-plane. Moreover, the inverse of  $\hat{f}|_p$  is well-defined and represented by a fractional linear transformation in  $\mathrm{SL}_2(\mathbb{Z}[i])$ .

We now introduce some notation and operations on  $\Sigma_R$  in order to describe the preimages of an element under  $\sigma$ . Let

$$\Gamma^n := \{(p_1, \dots, p_n) : p_i \leftrightarrow \rho_i \in \mathcal{P}_R, f(\rho_i) \supset \rho_{i+1}\}$$

be the set of admissible words of length  $n$ . We denote concatenation of two finite words with  $\|$ , i.e. for  $a = (a_1, \dots, a_n) \in \Gamma^n$  and  $b = (b_1, \dots, b_k) \in \Gamma^k$ ,

$$a\|b = (a_1, \dots, a_n, b_1, \dots, b_k)$$

If  $a_n$  and  $b_1$  satisfy the subshift rules, then we say the concatenation  $a\|b \in \Gamma^{n+k}$  is an admissible one. Concatenations of the form  $a\|b$  are also well-defined for  $b \in \Sigma_R$  as long as  $a$  is a finite word. In order to describe finite words which give admissible concatenations we set

$${}_x\Gamma_y^n := \{a \in \Gamma^n : x\|a\|y \text{ is admissible}\},$$

$$\Gamma_y^n := \{a \in \Gamma^n : a\|y \text{ is admissible}\}.$$

While  $y$  may be an infinite word, we must have  $x$  a finite word for the definition above.

For  $a \in \Sigma_R$ , we now write  $\sigma^{-1}(a) = \{b\|a : b \in \Gamma_a^1\}$  and for each  $b\|a \in \sigma^{-1}(a)$  denote the inverse branch of  $f$  at  $b\|a$  as  $g_b$ . In other words,  $g_b \in \mathrm{SL}_2(\mathbb{Z}[i])$  satisfies  $g_b(\pi(a)) = \pi(b\|a)$  and  $f(b\|a) = \pi^{-1} \circ g_b^{-1} \circ \pi(b\|a) = a$ .

We are now ready to describe the congruence transfer operators:

$$M_z f(x, g) = \sum_{a \in \Gamma_{1,x}} e^{z\tau(a\|x)} f(a\|x; g_a g).$$

## 5.1 Bounding $M_z$ in the Supremum Norm

We will exhibit cancellation in the iterates of the transfer operator

$$M_z^N f(x, g) = \sum_{a \in \Gamma_x^N} e^{zS_N \tau(a\|x)} f(a\|x; g_a g)$$

by treating the prefix and suffix of  $a \in \Gamma_x^N$  separately. Let  $N = M + R$  and for any  $b \in \Gamma^M$  define

$$\mu_b := \sum_{a \in_b \Gamma_x^R} e^{z S_N \tau(b||a||x)} \delta_{\pi_q(g_a)} = \sum_{a \in_b \Gamma_x^R} e^{z S_M \tau(b||a||x)} e^{z S_R \tau(a||x)} \delta_{\pi_q(g_a)}$$

where  $\pi_q : \mathrm{SL}_2(\mathbb{Z}[i]) \rightarrow \mathrm{SL}_2(q)$ . In order to decouple  $M$  and  $R$ , we need a lemma which reassures us that the value of  $e^{z S_M \tau(b||a||x)}$  does not change much for fixed  $b$  and varying  $a$ . We first establish a property of the system  $(S(\mathcal{X})_R, f)$  called *bounded distortion*.

**Lemma 3.** *For fixed  $R$  there is some  $C = C(R)$  so that for any  $p \in \mathcal{P}_R$*

$$\sup_{z \in p} \left| \frac{f''}{f'}(z) \right| < C.$$

*Proof.* From the definition

$$f(z) = (-1)^j \left( \frac{1}{z} - [1]z \right)$$

we have that both  $|f'(z)| = \frac{1}{|z|^2}$  and  $|f''(z)| = \frac{1}{|z|^3}$ . Since we have fixed  $R$  all of the parts in  $\mathcal{P}_R$  lie in an annulus (bounded away from 0), the bound follows.  $\square$

Bounded distortion leads to the following estimate for Birkhoff sums

**Lemma 4.** *For any two  $x, x_0 \in \Sigma_R$  and  $b \in \Gamma_x^M \cap \Gamma_{x_0}^M$  we have*

$$S_M \tau(b||x) = S_M \tau(b||x_0)(1 + O(1)).$$

*Proof.* For  $0 \leq k \leq M$  both  $\pi(\sigma^k(b||x))$  and  $\pi(\sigma^k(b||x_0))$  are in the same cylinder  $C_{\sigma^k(b)}$  which has diameter at most  $\frac{1}{2^k}$ .

The Mean Value Theorem combined with bounded distortion yields

$$\begin{aligned} |S_M(\tau(b||x)) - S_M(\tau(b||x_0))| &\leq \sum_{k=0}^M |\tau(\sigma^k(b||x)) - \tau(\sigma^k(b||x_0))| \\ &\leq \sum_{k=0}^M \log |f'(\pi(\sigma^k(b||x)))| - \log |f'(\pi(\sigma^k(b||x_0)))| \\ &\leq \sum_{k=0}^M C |\pi(\sigma^k(b||x)) - \pi(\sigma^k(b||x_0))| \\ &\leq C \sum_{k=0}^{\infty} \frac{1}{2^k} < C'. \end{aligned}$$

$\square$

Returning to our estimation of the measure  $\mu_b$ , we pick an arbitrary  $a_0 \in {}_b\Gamma_x^R$  and the lemma gives

$$\mu_b \leq C e^{z S_M \tau(b \| a_0 \| x)} \sum_{a \in {}_b\Gamma_x^R} e^{z S_R \tau(a \| x)} \delta_{\pi_q(g_a)}.$$

### 5.1.1 Expressing $\mu_b$ as a convolution

Now for a divisor  $L$  of  $R$ , break each  $a \in {}_b\Gamma_x^R$  into subwords of length  $L$ , i.e. write each  $a = a_1 \| \cdots \| a_r$  where  $l(a_i) = L$ . Then

$$\sum_{a \in {}_b\Gamma_x^R} e^{z S_R \tau(a \| x)} \delta_{\pi_q(g_a)} = \sum_{a \in {}_b\Gamma_x^R} e^{z S_L \tau(a_1 \| a_2 \| \cdots \| a_r \| x)} e^{z S_L \tau(a_2 \| a_3 \| \cdots \| a_r \| x)} \cdots e^{z S_L \tau(a_r \| x)} \delta_{\pi_q(g_a)}.$$

For each  $a_i$ , we will further decompose the word as a long prefix (of length  $L-4$ ) and a short suffix. Write  $a_i = a_i^{(L-4)} \| a_i^{(4)}$ . In order to separate dependence on the suffixes, we replace  $e^{z S_L \tau(a_i \| a_{i+1} \| \cdots \| a_r \| x)}$  with  $e^{z S_L \tau(a_i \| a_{i+1}^{(L-4)} \| x_i)}$  where  $x_i \in \Sigma_R$  is some arbitrary admissible choice based on  $a_{i+1}^{(L-4)}$ . Since we will be replacing many of the weights in  $\mu_b$ , we need to sharpen the estimate from the previous lemma. In particular, since  $\pi(\sigma^k(a_i \| a_{i+1}^{(L-4)} \| x_i))$  and  $\pi(\sigma^k(a_i \| a_{i+1} \| \cdots \| x))$  are in the same cylinder  $C_{\sigma^k(a_i \| a_{i+1}^{(L-4)})}$  for  $0 \leq k \leq 2L-4$  we have

$$|S_L \tau(a_i \| a_{i+1} \| \cdots \| a_r \| x) - S_L \tau(a_i \| a_{i+1}^{(L-4)} \| x_i)| < \frac{C}{2^{L-4}}.$$

Hence making the substitution for each of the  $r-1$  subwords (no substitution is necessary for  $a_r$ ) gives

$$\begin{aligned} \mu_b &\leq C e^{z(r-1)2^{-L}} \sum_{a \in {}_b\Gamma_x^R} \left[ \prod_{i=1}^r e^{z S_L \tau(a_i \| a_{i+1}^{(L-4)} \| x_i)} \right] \delta_{\pi_q(g_a)} \\ &= C e^{z(r-1)2^{-L}} \sum_{a \in {}_b\Gamma_x^R} \left[ \prod_{i=1}^r e^{z S_L \tau(a_i \| a_{i+1}^{(L-4)} \| x_i)} \right] [\delta_{\pi_q(g_{a_1})} * \delta_{\pi_q(g_{a_2})} * \cdots * \delta_{\pi_q(g_{a_r})}]. \end{aligned}$$

Instead of decomposing  $a$  into subwords of length  $L$ , we would like to start with subwords and determine which concatenations are admissible. We may choose  $a \in {}_b\Gamma_x^R$  in the following way:

1. select  $a_1^{(L-4)} \in {}_b\Gamma^{L-4}$  and  $a_i^{(L-4)} \in \Gamma^{L-4}$  for the remaining  $i = 2, \dots, r$ .

2. select  $a_r^{(4)} \in {}_{a_r^{(L-4)}}\Gamma_x^4$  and  $a_i^{(4)} \in {}_{a_{i-1}^{(L-4)}}\Gamma_{a_i^{(L-4)}}^4$  for the remaining  $i = 1, \dots, r-1$ .

The effect is to separate the sum into an outer sum depending on the prefixes of length  $L-4$  and an inner sum of the suffixes of length 4. We will also distribute the product  $\left[ \prod_{i=1}^r e^{zS_L\tau(a_i\|a_{i+1}^{(L-4)}\|x_i)} \right]$  into the convolution:

$$\begin{aligned} & \sum_{a_1^{(L-4)}, \dots, a_r^{(L-4)}} \sum_{a_1^{(4)}, \dots, a_r^{(4)}} \left[ e^{zS_L\tau(a_1\|a_2^{(L-4)}\|x_1)} \delta_{\pi_q(g_{a_1})} \right] * \dots * \left[ e^{zS_L\tau(a_r\|x)} \delta_{\pi_q(g_{a_r})} \right] \\ = & \sum_{a_1^{(L-4)}, \dots, a_r^{(L-4)}} \left[ \sum_{a_1^{(4)}} e^{zS_L\tau(a_1\|a_2^{(L-4)}\|x_1)} \delta_{\pi_q(g_{a_1})} \right] * \dots * \left[ \sum_{a_r^{(4)}} e^{zS_L\tau(a_r\|x)} \delta_{\pi_q(g_{a_r})} \right]. \end{aligned}$$

Define

$$\eta_j := \sum_{\alpha_j^{(4)}} e^{zS_L(a_j\|a_{j+1}^{(L-4)}\|x_j)} \delta_{\pi_q(g_{a_j})}$$

as a distribution on  $SL_2(q)$ . Our first observation about the measures  $\eta_j$  is that the ratio of any two coefficients is bounded. For two admissible  $a_j^{(4)}, \tilde{a}_j^{(4)}$

$$\begin{aligned} & \left| S_L\tau(a_j^{(L-4)}\|a_j^{(4)}\|a_{j+1}^{(L-4)}\|x_j) - S_L\tau(a_j^{(L-4)}\|\tilde{a}_j^{(4)}\|a_{j+1}^{(L-4)}\|x_j) \right| \\ & \leq C + \left| S_4\tau(a_j^{(4)}\|a_{j+1}^{(L-4)}\|x_j) - S_4\tau(\tilde{a}_j^{(4)}\|a_{j+1}^{(L-4)}\|x_j) \right|. \end{aligned}$$

The two cylinders  $C_{a_j^{(4)}}, C_{\tilde{a}_j^{(4)}}$  may be disjoint. However, the distance between them is still bounded since they lie in  $S(\mathcal{X}) \subset \{x + iy : |x| < 1/2, 0 < y < 1/2\}$ . So an application of the Mean Value Theorem gives that the second term in the inequality is less than some universal constant. In other words,

$$e^{zS_L\tau(a_j^{(L-4)}\|a_j^{(4)}\|a_{j+1}^{(L-4)}\|x_j)} = e^{zS_L\tau(a_j^{(L-4)}\|\tilde{a}_j^{(4)}\|a_{j+1}^{(L-4)}\|x_j)} (1 + O(1))$$

So the coefficients of the sum defining  $\eta_j$  are nearly flat. In order to establish an expansion result for  $\eta_j$ , we will also need the following:

**Lemma 5.** *For any  $j$ , pairs of admissible suffixes of  $a_j^{(L-4)}$  of length 4 generate all of  $SL_2(\mathbb{Z}[i])$ . Specifically, for any two letters  $i, j \in \mathcal{P}_R$ , we have*

$$\langle \pi(a)\pi(\tilde{a})^{-1} : a, \tilde{a} \in {}_i\Gamma_j^4 \rangle = SL_2(\mathbb{Z}[i])$$

where  $\pi : \Sigma_R \rightarrow SL_2(\mathbb{Z}[i])$ .

*Proof.* Write  $a = \alpha_1\alpha_2\alpha_3\alpha_4$  and  $\tilde{a} = \tilde{\alpha}_1\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_4$  where  $i||a||j$  and  $i||\tilde{a}||j$  are admissible. Note that  $i$  only places a restriction on  $\alpha_1$  and  $\tilde{\alpha}_1$  and  $j$  only places a restriction on  $\alpha_4$  and  $\tilde{\alpha}_4$ . In order to reduce the number of cases we must check, we will select  $\alpha_1 = \tilde{\alpha}_1$  and  $\alpha_4 = \tilde{\alpha}_4$  to maximize the number of options for the middle two letters.

Recall Figure 4.1 represents the image of a part in  $\mathcal{P}$  under  $\hat{f}$ . The image of a single part will be a union of the parts in various regions. For instance a full square part will map to all of the even regions or all of the odds, while a ‘sliver’ part will map to (1), (2), (7), or (8). Each of the eight regions contains at least one full square part. Select the letters  $\alpha_1 = \tilde{\alpha}_1$  and  $\alpha_4 = \tilde{\alpha}_4$  corresponding to that square part. The effect is that we may choose any  $\alpha_2, \alpha_3$  and  $\tilde{\alpha}_2, \tilde{\alpha}_3$  so long as:

1.  $\alpha_2||\alpha_3$  is admissible,
2.  $\tilde{\alpha}_2||\tilde{\alpha}_3$  is admissible,
3.  $\alpha_2$  and  $\tilde{\alpha}_2$  both come from the same quadrant (i.e. the both correspond to parts in the positive reals half-plane or the negative reals half-plane).

Therefore, it suffices to show that the elements  $(\alpha_2\alpha_3)(\tilde{\alpha}_2\tilde{\alpha}_3)^{-1}$  satisfying the above conditions generate  $PSL_2(\mathbb{Z}[i])$ . (Recall that  $\langle S \rangle = \langle \xi_i S \xi_i^{-1} \rangle$  for any generating set  $S$ .)

In the following table, we present a list of matrix equations of the form

$$m = a_1a_2[a_3a_4]^{-1}a_5a_6[a_7a_8]^{-1}.$$

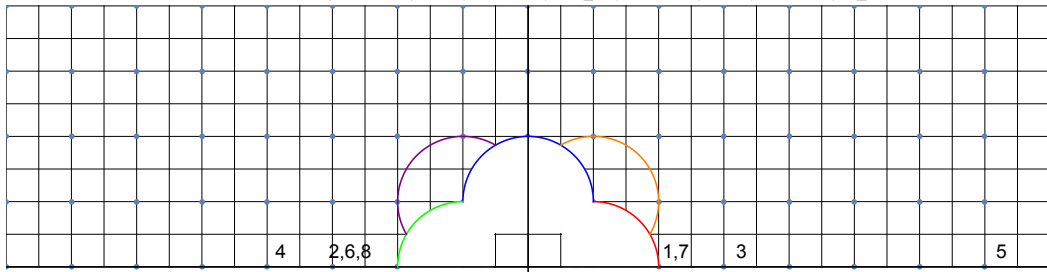
The first four equations all have  $a_1, a_3, a_5, a_7$  corresponding to parts in the right quadrant. The last four equations have them in the left quadrant. Below each equation is a figure showing the part corresponding to each matrix. One can refer to Table 3.1 to ensure that each pairing ((1, 2), (3, 4), (5, 6) and (7, 8)) is admissible.

Table 5.1: Admissible Generators

The Right Quadrant

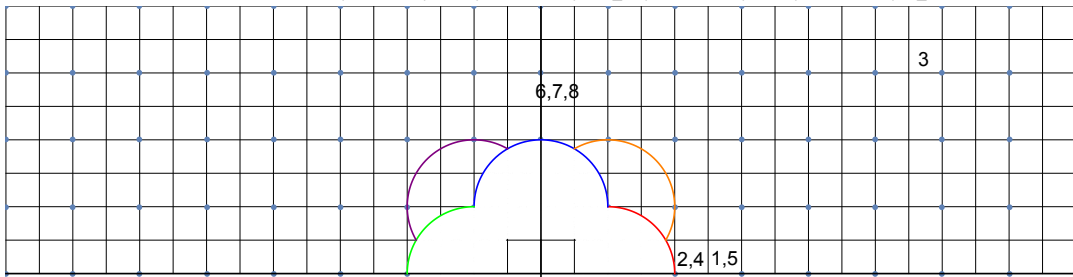
$$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -4 \end{pmatrix} \right]^{-1}$$

$$* \begin{pmatrix} 0 & -1 \\ 1 & 7 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \right]^{-1}$$



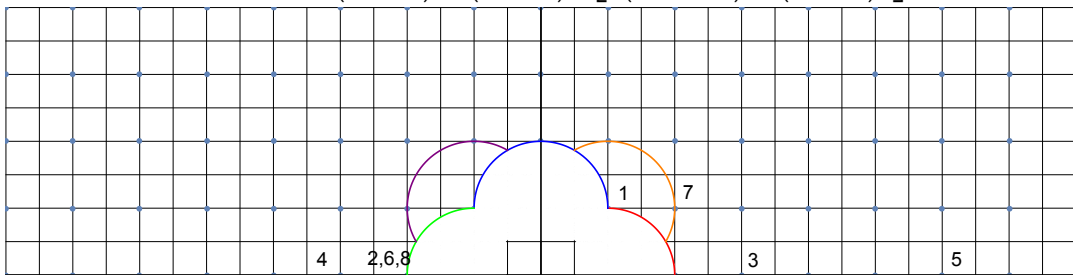
$$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 6+3i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \right]^{-1}$$

$$* \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 3ii \end{pmatrix} \left[ \begin{pmatrix} 0 & -i \\ -i & 3ii \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 3ii \end{pmatrix} \right]^{-1}$$

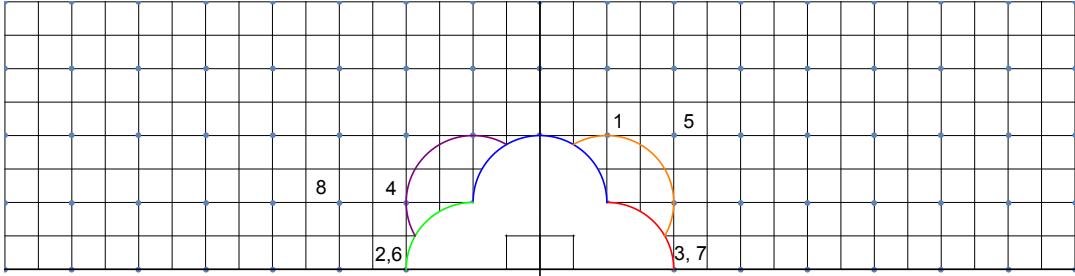


$$\begin{pmatrix} i & 1 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 1+i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \right]^{-1}$$

$$* \begin{pmatrix} 0 & -1 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 2+i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \right]^{-1}$$

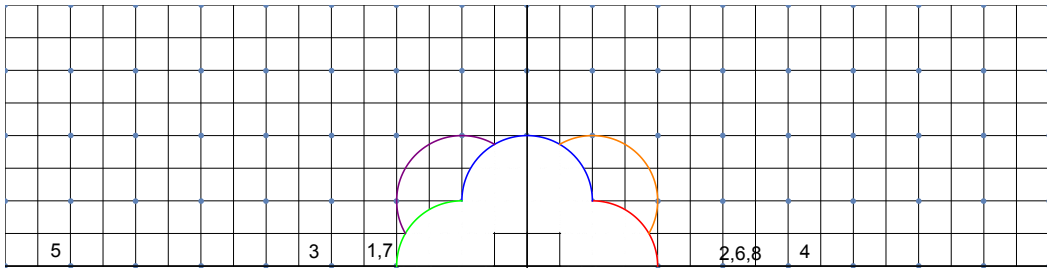


$$\begin{pmatrix} -3 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 1+2i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -2+i \end{pmatrix} \right]^{-1} \\ * \begin{pmatrix} 0 & -1 \\ 1 & 2+2i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -3+i \end{pmatrix} \right]^{-1}$$



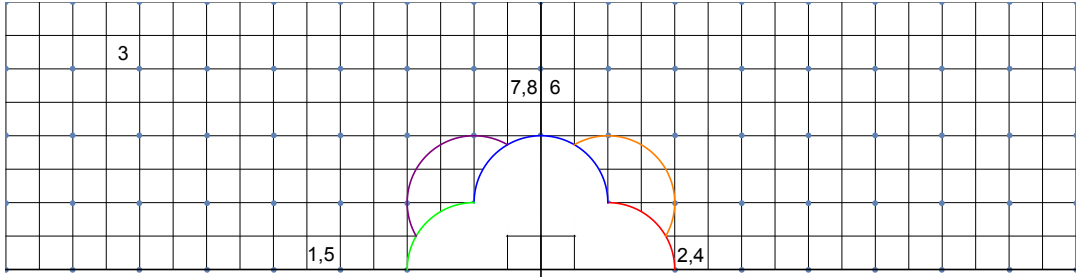
The Left Quadrant

$$\begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 4 \end{pmatrix} \right]^{-1} \\ * \begin{pmatrix} 0 & -1 \\ 1 & -7 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \right]^{-1}$$

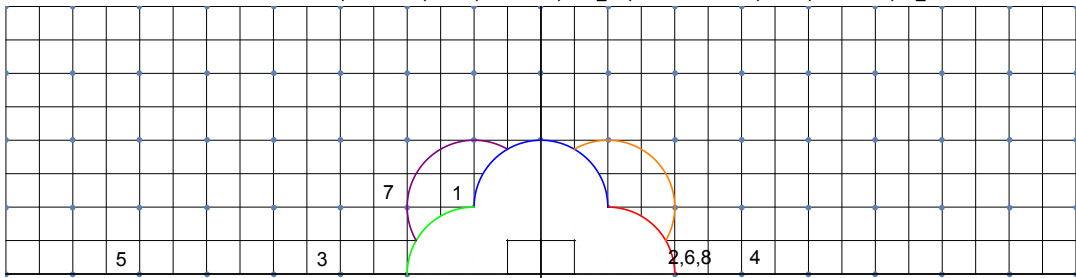


$$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -6+3i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \right]^{-1} \\ * \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 3ii \end{pmatrix} \left[ \begin{pmatrix} 0 & -i \\ -i & 3ii \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 3ii \end{pmatrix} \right]^{-1}$$

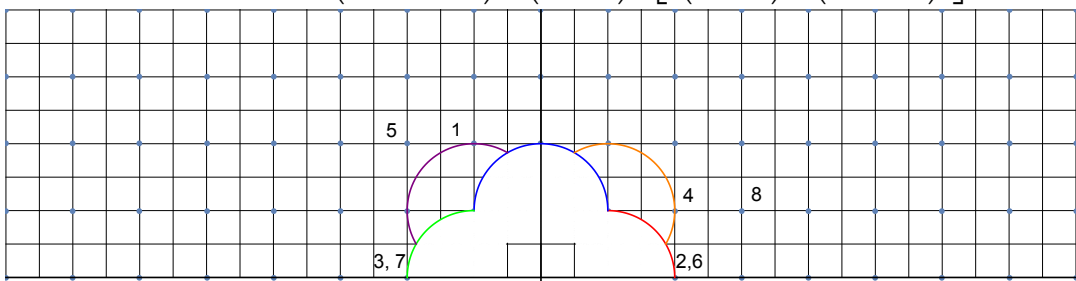




$$\begin{pmatrix} -i & -1 \\ 0 & i \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -1+i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \right]^{-1} \\ * \begin{pmatrix} 0 & -1 \\ 1 & -6 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -2+i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \right]^{-1}$$



$$\begin{pmatrix} -3 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -1+2i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2+i \end{pmatrix} \right]^{-1} \\ * \begin{pmatrix} 0 & -1 \\ 1 & -2+2i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \left[ \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 3+i \end{pmatrix} \right]^{-1}$$



We claim the previous calculations suffice to show that the suffixes generate  $\mathrm{SL}_2(\mathbb{Z}[i])$ .

The canonical generating set for  $PSL_2(\mathbb{Z}[i])$  is

$$\left\{ T_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad T_i = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \quad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}.$$

In both cases, we gave combinations for  $T_1$  and  $Q$ . For  $T_i$ , we appeal to the following identity

$$\begin{pmatrix} i & 1 \\ 0 & -i \end{pmatrix} \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}.$$

Since we have the full group of translation matrices in  $PSL_2(\mathbb{Z}[i])$ , it is enough to find a matrix of the form

$$\begin{pmatrix} * & * \\ -1 & 0 \end{pmatrix}$$

which is the last equation for each case (the left or right quadrant.)  $\square$

### 5.1.2 Expansion via Selberg's 3/16 Theorem

We are now ready to prove the expansion theorem for the  $\eta_j$ 's defined on page 30. For each square-free  $q$ , we have the product representation  $SL_2(q) \cong \prod_{p|q} SL_2(p)$  which gives rise to the following decomposition for functions defined on  $SL_2(q)$ :

$$L_2(SL_2(q)) = \bigoplus_{q'|q} E_{q'}$$

where

$$E_{q'} := \{\varphi : SL_2(q) \rightarrow \mathbb{C} \mid$$

$$\varphi(g) = \varphi(g') \text{ if } g = g'(q'),$$

$$\langle \varphi, \psi \rangle = 0 \text{ for all } \psi \in E_{q''} \text{ such that } q''|q' \text{ and } q'' < q'\}$$

We first treat one  $E_q$  at a time, and then assemble them using Fourier-Walsh decomposition (see Section 5.3 on page 42).

**Theorem 11.** *If  $\varphi \in E_q$ , then*

$$\|\eta_j * \varphi\|_2 \leq (1 - C) \|\eta_j\|_1 \|\varphi\|_2$$

*Proof.* First, we retrace the standard steps to rewrite  $\|\eta_j * \varphi\|_2^2$  in terms of a convolution operator. By definition,

$$\begin{aligned} \|\eta_j * \varphi\|_2^2 &= \langle \eta_j * \varphi, \eta_j * \varphi \rangle \\ &= \left\langle \sum_{a_j^{(4)}} \beta_{a_j^{(4)}} \delta_{\pi_q(a_j)} * \varphi, \sum_{a_j^{(4)}} \beta_{a_j^{(4)}} \delta_{\pi_q(a_j)} * \varphi \right\rangle \end{aligned}$$

where we have

$$\beta_{a_j^{(4)}} := e^{z_{SL} \tau(a_j \|a_{j+1}^{(L-4)}\|x_j)}.$$

Expanding the square gives

$$\begin{aligned} \|\eta_j * \varphi\|_2^2 &= \sum_{k \in G} \left[ \sum_{a_j^{(4)}} \beta_{a_j^{(4)}} \delta_{\pi_q(a_j)} * \varphi(k) \right]^2 \\ &= \sum_k \sum_{a_j^{(4)}, \tilde{a}_j^{(4)}} \beta_{a_j^{(4)}} \beta_{\tilde{a}_j^{(4)}} \varphi(k \pi_q(a_j)^{-1}) \varphi(k \pi_q(\tilde{a}_j)^{-1}). \end{aligned}$$

(Recall that  $a_j = a_j^{(L-4)} \|a_j^{(4)}\|$  and  $\tilde{a}_j = a_j^{(L-4)} \|\tilde{a}_j^{(4)}\|$ .) We reorder the following sums as

$$\begin{aligned} \|\eta_j * \varphi\|_2^2 &= \sum_k \sum_{a_j^{(4)}, \tilde{a}_j^{(4)}} \beta_{a_j^{(4)}} \beta_{\tilde{a}_j^{(4)}} \varphi(k \pi_q(a_j^{(4)})^{-1}) \varphi(k \pi_q(\tilde{a}_j^{(4)})^{-1}) \\ &= \sum_k \sum_{a_j^{(4)}, \tilde{a}_j^{(4)}} \beta_{a_j^{(4)}} \beta_{\tilde{a}_j^{(4)}} \varphi(k \pi_q(a_j^{(4)})^{-1}) \pi_q(\tilde{a}_j^{(4)}) \varphi(k). \end{aligned}$$

Now, from Lemma 5 and the analogue of Selberg's 3/16 theorem for congruence subgroups of  $\mathrm{SL}_2(\mathbb{Z}[i])$  (see [Sar83] or Theorem 6.1 in [EGM13]) we deduce that for any  $\varphi \in l_0^2(\mathrm{SL}_2(q))$  there is some choice of  $a_0, \tilde{a}_0 \in a_j^{(4)}, \tilde{a}_j^{(4)}$  so that

$$\|\pi_q(a_0) \pi_q(\tilde{a}_0)^{-1} * \varphi - \varphi\|_2 > \epsilon \|\varphi\|_2.$$

The law of cosines gives

$$\|\pi_q(a_0) \pi_q(\tilde{a}_0)^{-1} * \varphi - \varphi\|_2^2 = 2 - 2 \langle \pi_q(a_0) \pi_q(\tilde{a}_0)^{-1} * \varphi, \varphi \rangle$$

and so

$$\beta_{a_0} \beta_{\tilde{a}_0} \langle \pi_q(a_0) \pi_q(\tilde{a}_0)^{-1} * \varphi, \varphi \rangle < \beta_{a_0} \beta_{\tilde{a}_0} (1 - \epsilon') \|\varphi\|_2.$$

We separate the  $a_0, \tilde{a}_0$  term from the rest of the sum as follows

$$\|\eta_j * \varphi\|_2^2 \leq \sum_{a_j^{(4)}, \tilde{a}_j^{(4)}} \beta_{a_j^{(4)}} \beta_{\tilde{a}_j^{(4)}} \|\varphi\|_2 + \beta_{a_0} \beta_{\tilde{a}_0} (1 - \epsilon') \|\varphi\|_2$$

Since we established earlier that  $\beta_{\alpha_j} = \beta_{\alpha_j'} (1 + O(1))$ , this gives

$$\|\eta_j * \varphi\|_2^2 \leq (1 - C) \|\eta_j\|_1^2 \|\varphi\|_2^2.$$

□

Apply Theorem 11 to each  $\eta_j$  and we have

**Corollary 1.** For  $\varphi \in E_q$

$$\begin{aligned} & \left\| \sum_{a \in_b \Gamma_x^R} \left[ e^{z S_L \tau(a_1 \| a_2^{(L-4)} \| x_1)} \delta_{\pi_q(g_{a_1})} \right] * \dots * \left[ e^{z S_L \tau(a_r \| x)} \delta_{\pi_q(g_{a_r})} \right] * \varphi \right\|_2 \\ & \leq (1 - C)^r \left( \sum_{a \in_b \Gamma_x^R} \prod_{i=1}^r |e^{z S_L \tau(a_i \| a_{i+1}^{(L-4)} \| x_i)}| \right) \|\varphi\|_2. \end{aligned}$$

Next, we exploit quasi-randomness of  $SL_2(q)$  to get a bound for  $\mu_b$ .

**Theorem 12.** For  $R \asymp \log N(q)$  with  $\mu_b$  as defined previously on page 28, we have

$$\|\mu_b * \varphi\|_2 \leq CN(q)^{-1/4} \|\mu_b\|_1 \|\varphi\|_2$$

for any  $\varphi \in E_q$ .

*Proof.* Recall

$$\mu_b := \sum_{a \in_b \Gamma_x^R} e^{z S_N \tau(b \| a \| x)} \delta_{\pi_q(g_a)}$$

Define

$$\nu := e^{\Re z S_M \tau(b \| a_0 \| x)} \sum_{a \in_b \Gamma_x^R} e^{\Re z S_R \tau(a \| x)} \delta_{\pi_q(g_a)}$$

We established earlier that  $|\mu_b| \leq C\nu$ . Corollary 1 and bounded distortion yield the following bound for  $\nu$ :

$$\left\| \sum_{a \in_b \Gamma_x^R} e^{\Re z S_R \tau(a \| x)} \delta_{\pi_q(g_a)} * \varphi \right\|_2 \leq (1 - C)^r \left( \sum_{a \in_b \Gamma_x^R} e^{\Re z S_R \tau(a \| x)} \right) \|\varphi\|_2$$

and therefore  $\|\nu * \varphi\|_2 \leq (1 - C)^r \|\nu\|_1 \|\varphi\|_2$ .

Define  $A$  as the convolution operator  $\varphi \mapsto \mu_b * \varphi$ . First note that  $A$  acts on  $E_q$  since it's a linear combination of convolutions with delta functions. Since  $A^*A$  is self adjoint, we have  $\text{tr}(A^*A) = \lambda_1^2 + \dots + \lambda_{|G|-1}^2$ .

$$\begin{aligned} \lambda^2 \text{mult}(\lambda) &\leq \text{tr}(A^*A) \\ &= \sum_{g \in G} \langle (A^*A)^2 \delta_g, \delta_g \rangle \\ &= \sum_{g \in G} \|\tilde{\mu}_b * \mu_b * \delta_g\|_2^2 \\ &= |G| \|\tilde{\mu}_b * \mu_b\|_2^2 \\ &\leq C |G| \|\tilde{\nu} * \nu\|_2^2. \end{aligned}$$

where  $\tilde{\mu}_b(g) = \overline{\mu_b}(g^{-1})$  and a similar definition applies to  $\nu$ . The multiplicity of  $\lambda$  equal to the dimension of the eigenspace is at least  $\frac{N(q)-1}{2}$  by the Fobenius lemma. So

$$\|A\|_{op} = \max \lambda^{1/2} \leq C \left( \frac{|G| \|\tilde{\nu} * \nu\|_2^2}{N(q)} \right)^{1/4}.$$

We bound  $\|\tilde{\nu} * \nu\|_2$  by introducing  $\psi = \delta_e - \frac{1}{|G|} \mathbf{1}_G$ . Observe  $\psi \in l_0^2$  and  $\|\psi\|_2 < 1$ .

$$\begin{aligned} \|\tilde{\nu} * \nu\|_2 &= \|\tilde{\nu} * \nu * \delta_e\|_2 \\ &\leq \|\tilde{\nu} * \nu * \frac{\mathbf{1}_G}{|G|}\|_2 + \|\tilde{\nu} * \nu * \psi\|_2 \\ &\leq \frac{\|\nu\|_1^2}{|G|^{1/2}} + \|\nu\|_1 \|\nu * \psi\|_2. \end{aligned}$$

Since  $\|\nu * \psi\|_2 < (1 - C)^R \|\nu\|_1$ , we can choose  $R = C' \log q$  to get  $\|\tilde{\nu} * \nu\|_2 < \frac{\|\nu\|_1}{G^{1/2}}$ .  $\square$

### 5.1.3 Applying Theorem 12

Now we would like to use the previous bound on our congruence transfer operators defined on page 27. Rewrite  $M_z^n$  as

$$\begin{aligned} M_s^n f(x, g) &= \sum_{a \in \Gamma_x^n} e^{sS_n \tau(a||x)} f(a||x; g_a g) \\ &= \sum_{a \in \Gamma_x^M} \sum_{b \in \Gamma_a^R} e^{sS_n \tau(b||a||x)} f(b||a||x; g_b g_a g) \\ &= \sum_{a \in \Gamma_x^M} \sum_{b \in \Gamma_a^R} e^{sS_n \tau(b||a||x)} f(b||x_b; g_b g_a g) + O \left( |f|_\rho \rho^M \sum_{a \in \Gamma_x^n} e^{sS_n \tau(a||x)} \right). \end{aligned}$$

where  $x_b$  is arbitrarily chosen as long as  $b||x_b$  is admissible. We will frequently use the fact that for  $s \leq \delta_R$  and the transfer operators defined in the non-congruence setting we have

$$|\mathcal{L}_s^n 1(x)| = \left| \sum_{a \in \Gamma_x^n} e^{z\tau(a||x)} \right| < \lambda_{\Re s}^n.$$

Filling this in for the second term in our bound for the congruence transfer operator gives

$$\begin{aligned} M_z^n f(x, g) &= \sum_{b \in \Gamma^R} \sum_{a \in_b \Gamma_x^M} e^{\tau^n(b||a||x)} f(b||x_b; g_b g_a g) + O(|f|_\rho \rho^M \lambda_{\Re s}^n) \\ &= \sum_{b \in \Gamma^R} [\mu_b * f(b||x_b; g_b \cdot)](g) + O(|f|_\rho \rho^M \lambda_{\Re s}^n). \end{aligned}$$

If  $f(b||x_b, g_b \cdot) \in E_q$ , then we are in position to use the bound for  $\mu_b$ :

$$\|\mu_b * f(b||x_b; g_b \cdot)\|_2 < CN(q)^{-1/4} \|\mu_b\|_1 \|f(b||x_b; g_b \cdot)\|_{l_2(G)} < CN(q)^{1/4} \|\mu_b\|_1 \|f\|_\rho.$$

For any  $b \in \Gamma^R$ ,

$$\begin{aligned} \|\mu_b\|_1 &= \left\| \sum_{a \in_b \Gamma_x^R} e^{sS_n \tau(b||a||x)} \delta_{g_a} \right\|_1 \\ &\leq \sum_{a \in_b \Gamma_x^R} \left| e^{sS_R \tau(b||a||x)} e^{sS_M \tau(a||x)} \right| \\ &\leq C |e^{sS_R \tau(b||a_0||x)}| \sum_{a \in_b \Gamma_x^R} |e^{sS_M \tau(a||x)}| \\ &\leq C e^{\Re s S_R \tau(b||a_0||x)} \mathcal{L}_{\Re s}^R 1(x). \end{aligned}$$

Apply this bound for each summand in  $M_s^n$ ,

$$\begin{aligned}
\|M_s^n f\|_\infty &\leq CN(q)^{-1/4} \sum_{b \in \Gamma^R} e^{\Re s S_R \tau(b\|a_0\|x)} \|\mathcal{L}_{\Re s}^R 1(x)\|_\infty \|f\|_\infty + O(|f|_\rho \rho^M \lambda_{\Re s}^n) \\
&\leq CN(q)^{-1/4} \|\mathcal{L}_{\Re s}^n 1(x)\|_\infty \|f\|_\infty + O(|f|_\rho \rho^M \lambda_{\Re s}^n) \\
&\leq CN(q)^{-1/4} \lambda_{\Re s}^n \|f\|_\infty + O(|f|_\rho \rho^M \lambda_{\Re s}^n).
\end{aligned}$$

## 5.2 Bounding $M_z$ in Variation

Now, we need to bound  $|M_z^n f|_\rho$ . Suppose  $x, y \in \Sigma$  and  $w \in \Gamma_x^k \cap \Gamma_y^k$ ,

$$\begin{aligned}
|M_z^n f(w\|x; g) - M_z^n f(w\|y; g)| &= \left| \sum_{a \in \Gamma_w^n} e^{z S_n \tau(a\|w\|x)} f(a\|w\|x, g_a g) - e^{z S_n \tau(a\|w\|y)} f(a\|w\|y, g_a g) \right| \\
&\leq \left| \sum_{a \in \Gamma_w^n} e^{z S_n \tau(a\|w\|x)} (f(a\|w\|x; g_a g) - f(a\|w\|y; g_a g)) \right| \\
&\quad + \left| \sum_{a \in \Gamma_w^n} (e^{z S_n \tau(a\|w\|x)} - e^{z S_n \tau(a\|w\|y)}) f(a\|w\|y, g_a g) \right|.
\end{aligned}$$

For the first term, we note that  $a\|w\|x$  and  $a\|w\|y$  agree in the first  $n+k$  letters so

$$\begin{aligned}
\left| \sum_{a \in \Gamma_w^n} e^{z S_n \tau(a\|w\|x)} (f(a\|w\|x; g_a g) - f(a\|w\|y; g_a g)) \right| &\leq |f|_\rho \rho^{n+k} \sum_{a \in \Gamma_w^n} |e^{z S_n \tau(a\|w\|x)}| \\
&\leq |f|_\rho \rho^{n+k} \mathcal{L}_{\Re z}^n 1(w\|x) \\
&\leq |f|_\rho \rho^{n+k} \lambda_{\Re z}^n.
\end{aligned}$$

For the second term, we will use a similar approach as before. Decompose  $\Gamma^n$  into  $\Gamma^R \otimes \Gamma^M$  and decouple:

$$\begin{aligned}
&\sum_{a \in \Gamma_w^n} (e^{z S_n \tau(a\|w\|x)} - e^{z S_n \tau(a\|w\|y)}) f(a\|w\|y, g_a g) \\
&= \sum_{a \in \Gamma_w^M} \sum_{b \in \Gamma_a^R} (e^{z S_n \tau(b\|a\|w\|x)} - e^{z S_n \tau(b\|a\|w\|y)}) f(b\|a\|w\|y; g_b g_a g) \\
&= \sum_{a \in \Gamma_w^M} \sum_{b \in \Gamma_a^R} (e^{z S_n \tau(b\|a\|w\|x)} - e^{z S_n \tau(b\|a\|w\|y)}) f(b\|x\|b; g_b g_a g) \\
&\quad + O\left(|f|_\rho \rho^M \sum_{a \in \Gamma_x^n} (e^{z S_n \tau(a\|w\|x)} - e^{z S_n \tau(a\|w\|y)})\right).
\end{aligned}$$

For the error term, we estimate

$$\sum_{a \in \Gamma_x^n} (e^{zS_n \tau(a\|w\|x)} - e^{zS_n \tau(a\|w\|y)}) \leq \sum_{a \in \Gamma_x^n} |e^{zS_n \tau(a\|w\|x)}| |1 - e^{z(S_n \tau(a\|w\|x) - S_n(a\|w\|y))}|.$$

For the first term, we use the eigenvalue bound. For the second, we use the fact that  $\pi(a\|w\|x)$  and  $\pi(a\|w\|y)$  are in the same  $n+k$  cylinder combined with bounded distortion (i.e.  $|(S_n \tau(a\|w\|x) - S_n(a\|w\|y))| < C2^{-k}$ .) This gives

$$\sum_{a \in \Gamma_x^n} (e^{zS_n \tau(a\|w\|x)} - e^{zS_n \tau(a\|w\|y)}) < C\lambda_{\mathfrak{R}z}^n (1 + |\Im z|)2^{-k}.$$

### 5.2.1 Applying Theorem 12 Again

For each  $b \in \Gamma_R$  define

$$\mu_b := \sum_{a \in_b \Gamma_w^M} (e^{\tau^n(b\|a\|w\|x)} - e^{\tau^n(b\|a\|w\|y)}) \delta_{\pi_q(g_a)}.$$

The same proof as before follows through for  $\mu_b$ , as long as  $R \asymp \log N(q)$ , to yield the following for  $\varphi \in E_q$ :

$$\|\mu_b * \varphi\|_2 < CN(q)^{-1/4} \|\mu_b\|_1 \|\varphi\|_2.$$

Referring to the proof, we note that the  $\mu_b$ 's have slightly different coefficients in the corresponding  $\eta_j$ 's:

$$\eta_j := \sum_{a_j^{(4)}} (e^{zS_L(a_j\|a_{j+1}^{(L-4)}\|x_j)} - e^{zS_L(a_j\|a_{j+1}^{(L-4)}\|y_j)}) \delta_{\pi_q(a_j)}.$$

However, the important property of 'nearly flat coefficients' (i.e. that the constant for each summand varies by at most a constant ratio) is preserved. Returning to the bound for  $M_s^n$ , we have

$$|M_s^n f|_\rho \leq CN(q)^{-1/4} \sum_{b \in \Gamma^R} \|\mu_b\|_1 \|f\|_\infty + O(|f|_\rho \rho^M (\lambda_{\mathfrak{R}s}^n (1 + |\Im z|)2^{-k} + \lambda_{\mathfrak{R}s}^{n+k})).$$

For each  $\mu_b$  we have

$$\|\mu_b\|_1 \leq \sum_{a \in \Gamma_x^M} |e^{zS_n \tau(a\|w\|x)}| \cdot |1 - e^{z(S_n \tau(a\|w\|x) - S_n(a\|w\|y))}| < C\lambda_{\mathfrak{R}z}^M (1 + |\Im z|)2^{-k} |e^{zS_R(b\|x_b)}|.$$

So finally we have

$$|M_s^n f|_\rho \leq CN(q)^{-1/4} \lambda_{\mathfrak{R}s}^n \|f\|_\infty (1 + |\Im z|)2^{-k} + |f|_\rho \rho^M \lambda_{\mathfrak{R}s}^n (1 + |\Im z|)2^{-k} + |f|_\rho \rho^M \lambda_{\mathfrak{R}s}^{n+k}.$$



Recall also our bound for  $\|M_s^n f\|_\infty$ :

$$\|M_s^n f\|_\infty \leq CN(q)^{-1/4} \lambda_{\mathfrak{R}s}^n \|f\|_\infty + |f|_\rho \rho^R \lambda_{\mathfrak{R}s}^n.$$

We assume  $n > \log N(q)$  because we needed  $R = c \log N(q)$  and  $n = M + R$ . Thus, we have

$$\begin{aligned} |M_s^n f|_\rho &\leq C \lambda_{\mathfrak{R}s}^n N(q)^{-1/4} (1 + |\Im z|) (\|f\|_\infty + \rho^{n/2} |f|_\rho) \\ \|M_s^n f\|_\infty &\leq C \lambda_{\mathfrak{R}s}^n N(q)^{-1/4} (\|f\|_\infty + \rho^{n/2} |f|_\rho). \end{aligned}$$

Further, take  $n \sim \log q + C \log(1 + |\Im z|)$  and we have

$$\|M_s^n f\|_\infty + \rho^{n/2} |M_s^n f|_\rho < \lambda_{\mathfrak{R}s}^n N(q)^{-1/4} (\|f\|_\infty + \rho^{n/2} |f|_\rho).$$

Iterating the inequality yields

$$\|M_s^{mn} f\|_\rho < \lambda_{\mathfrak{R}s}^{mn} N(q)^{-m/4} N(q) (1 + |\Im s|) \|f\|_\rho.$$

### 5.3 Fourier-Walsh Decomposition

We would like to extrapolate from the previous bound (valid for  $f \in E_q \subset L^2(\mathrm{SL}_2(q))$ ) a bound for any of the non-constant level subspaces. In particular, recall that  $L_2(\mathrm{SL}_2(q)) = \bigoplus_{q'|q} E_{q'}$ . We temporarily denote  $E_{q'}$  inside of  $L_2(\mathrm{SL}_2(q))$  as  $E_{q'}(q)$  in order to compare  $E_{q'}(q)$  with  $E_{q'}(q')$ . This decomposition extends to one for  $C(\Sigma_R \times \mathrm{SL}_2(q))$ , namely

$$C(\Sigma_R \times \mathrm{SL}_2(q)) = \bigoplus_{q'|q} C(\Sigma_R, E_{q'}).$$

1.  $M_{z,q}$  preserves the subspaces  $C(\Sigma_R, E_{q'})$  because

$$M_{z,q} f(x, \cdot) = \sum_{a \in \Gamma_x^1} e^{z\tau(a\|x)} f(a\|x, g_a \cdot)$$

and  $g \mapsto g_a g$  is an automorphism of  $\mathrm{SL}_2(q)$  for each  $a$ . So the right hand side is a linear combination of functions in  $C(\Sigma_R, E_{q'})$ .

2. The natural projection from  $\mathrm{SL}_2(q)$  to  $\mathrm{SL}_2(q')$  extends to the subspaces  $E_{q'}(q)$  and  $E_{q'}(q')$ . In particular,  $f \in E_{q'}(q)$  and the corresponding  $\hat{f} \in E_{q'}(q')$  satisfy

$$\|f\|_{L^2(q)} = \sqrt{\frac{|\mathrm{SL}_2(q)|}{|\mathrm{SL}_2(q')|}} \|\hat{f}\|_{L^2(q')}$$

and if we denote the norm  $\|\cdot\|_{\rho,q'}$  on  $\mathcal{F}_\rho(\Sigma_R, E'_q)$  we have

$$\|f\|_{\rho,q} \leq \sum_{q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} \|\hat{f}\|_{\rho,q'}.$$

3.  $M_{z,q}(\widehat{1 \otimes f}) = M_{z,q'}(\widehat{1 \otimes f})$ , i.e. the projection is equivariant under the appropriate congruence transfer operators.

These three properties allow us to decompose a function  $f = \sum_{q'|q} f_{q'}$  where  $f_{q'} \in C(\Sigma_R \times E_{q'})$  and apply our bound as if  $f_{q'}$  is in  $\mathcal{F}_\rho(\Sigma, E_{q'}(q'))$ . Assume  $f_1 = 0$ , i.e.  $f$  is orthogonal to the constant function.

### 5.3.1 Small Imaginary Part

For small imaginary part ( $|\Im s| < N(q)$ )

$$\begin{aligned} \sum_{1 \neq q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} \|M_{z,q'}^m \hat{f}_{q'}\|_{\rho,q'} &\leq \sum_{1 \neq q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} \lambda_{\Re s}^m N(q)^{-m/(4n)} N(q') (1 + |\Im s|) \|\hat{f}_{q'}\|_{q'} \\ &\leq \sum_{1 \neq q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} N(q)^C e^{-\epsilon n} \lambda_{\Re z}^n \|\hat{f}_{q'}\|_{q'} \\ &\leq N(q)^C e^{-\epsilon n} \lambda_{\Re z}^n \|f\|_q \end{aligned}$$

where we used that  $\|\hat{f}_{q'}\|_{q'} \leq \|f\|_q$  and that the number of divisors of  $q$  is at most  $N(q)^{\epsilon'}$ .

### 5.3.2 Large Imaginary Part

For large imaginary part ( $|\Im s| > N(q)$ ), we have

$$\begin{aligned} \sum_{1 \neq q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} \|M_{z,q'}^m \hat{f}_{q'}\|_{\rho,q'} &\leq \sum_{1 \neq q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} \lambda_{\Re s}^m N(q)^{-m/(4n)} N(q') (1 + |\Im s|) \|\hat{f}_{q'}\|_{q'} \\ &\leq \sum_{1 \neq q'|q} \sqrt{\frac{|SL_2(q)|}{|SL_2(q')|}} |\Im z|^C e^{-\epsilon n \frac{\log N(q')}{\log |\Im z|}} \lambda_{\Re z}^n \|\hat{f}_{q'}\|_{q'} \\ &\leq \|f\|_q |\Im z|^C \lambda_{\Re z}^n \sum_{1 < N(q') < |\Im z|} e^{-\epsilon n \frac{\log N(q')}{\log |\Im z|}}. \end{aligned}$$

To bound the sum, observe

$$\begin{aligned}
\sum_{1 < N(q') < |\Im z|} e^{-\epsilon n \frac{\log N(q)}{\log |\Im z|}} &< \prod_{p|q} \left( 1 + \exp \left( \epsilon n \frac{-\log N(p)}{\log |\Im z|} \right) \right) - 1 \\
&\leq \exp \left( \sum_p e^{-\epsilon n \frac{\log N(p)}{\log |\Im z|}} \right) - 1 \\
&< C \exp \left( \sum_{s=2}^{\infty} e^{-\epsilon n \frac{\log s}{\log |\Im z|}} \right) - 1 \\
&< e^{-\frac{\epsilon n}{\log |\Im z|}}.
\end{aligned}$$

We have shown the following:

**Theorem 13.** *For  $f$  orthogonal to the constant functions, there is some  $\epsilon$  so that*

$$\|M_s^n f\|_{\rho} \leq \begin{cases} N(q)^C e^{-\epsilon n} \lambda_{\Re s}^n \|f\|_{\rho} & |\Im z| \leq N(q) \\ |\Im s|^C e^{\frac{-\epsilon n}{\log |\Im s|}} \lambda_{\Re s}^n \|f\|_{\rho} & |\Im z| \gg N(q) \end{cases}.$$

Now, in order to find the region of analyticity of  $(1 - M_z)^{-1}$ , we recall that  $z \mapsto \lambda_z$  is holomorphic in a small neighborhood of  $\delta_R$  and  $\lambda_{\delta_R} = 1$ . In particular, this means there is some  $\epsilon_2$  such that  $\lambda_{\Re z} < e^{\epsilon/2}$  for all  $\Re z \in [-\delta, -\delta + \epsilon_2]$ . We can also find  $\epsilon_3$  so that for  $z$  satisfying  $\Re z \in [-\delta, -\delta + \epsilon_3 \frac{1}{\log |\Im z|}]$  we have

$$\lambda_{\Re z} < e^{\epsilon/2 \frac{1}{\log |\Im z|}}.$$

For  $z$  in both regions, we have that  $(1 - M_z)^{-1}$  (restricted to the space orthogonal to constant functions) is holomorphic and bounded by

$$\|(1 - M_z)^{-1}\|_{\rho} < (N(q) + |\Im z|)^C.$$

## 5.4 Fourier Analysis of the Renewal Function

Similar to the previous chapter, we introduce a counting function which satisfies a functional equation relating it to the resolvent of the congruence transfer operator  $M_z$  on  $\mathcal{F}_{\rho}(\Sigma_R \times \mathrm{SL}_2(q))$ . To incorporate the congruence aspect, we define

$$N_q(X, a, g, f) := \sum_{n=0}^{\infty} \sum_{b \in \Gamma_a^n} g(b|a) f(\pi_q(b)) \mathbf{1}_{\{S_n \tau(b|a) \leq X\}}.$$

where  $g$  is a function on  $\Sigma_R$ ,  $f$  is a function on  $\mathrm{SL}_2(q)$ ,  $a \in \Sigma_R$ , and  $X > 0$ . The renewal equation is

$$N_q(X, a, g, f) = \sum_{b \in \Gamma_a^1} N_q(X - \tau(b||a), b||a, g, \rho(\pi_q(b))f) + g(a)f(1)\mathbf{1}_{\{0 \leq X\}}$$

where  $\rho$  denotes the right regular representation of  $\mathrm{SL}_2(q)$ . In particular

$$\rho(\pi_q(b))f(z) = f(z\pi_q(b)).$$

The Laplace transform

$$F_q(s, a, g, f) := \int_{-\infty}^{\infty} e^{-st} N_q(t, a, g, f) dt$$

satisfies

$$sF_q(s, a, g, f) = (1 - M_{s,q})^{-1}[g \otimes f](a).$$

Observe that  $N_q(X, a, g, f)$  is linear in  $f$  and so is its Laplace transform. The main contribution to  $F_q(s, a, g, f)$  comes from the constant term and this analysis is a straightforward extension of the previous chapter since

$$N_q(t, a, g, \mathbf{1}) = N(t, a, g).$$

The contribution from functions orthogonal to constants is bounded using Theorem 13.

As in [BGS11], we can choose a smooth nonnegative function  $k$  on  $\mathbb{R}$  such that  $\int k = 1$ ,  $\mathrm{supp}(k) \subset [1, 1]$  with the following bound for its Fourier transform

$$|\hat{k}(\xi)| \leq C \exp(-|\xi|^{1/2}).$$

For some small parameter  $\gamma$ , we define

$$k_\gamma(t) = \frac{1}{\gamma} k\left(\frac{t}{\gamma}\right).$$

Note that  $|\hat{k}_\gamma(\xi)| \leq C \exp(-|\gamma\xi|^{1/2})$ .

Inserting the smoothing function gives

$$\int_{-\infty}^{\infty} k_\gamma(t) N_q(X + t, a, f) dt = \frac{1}{2\pi i} \int_{\delta_R + i\mathbb{R}} e^{Xs} \hat{k}_\gamma(s) (I - M_s)^{-1} ds.$$

Theorem 13 allows us to shift the contour by

$$-\delta_R + it \mapsto -\delta_R + \frac{\epsilon}{2} \min\left(1, \frac{1}{\log(1 + |t|)}\right) + it$$

and gives

**Theorem 14.** For  $\varphi \in L^2(SL_2(q))$  with  $\langle \varphi, 1 \rangle = 0$ ,

$$\left| \int_{-\gamma/2}^{\gamma/2} k_\gamma(t) N(T+t, \varphi, a) dt \right| < N(q)^C \gamma^{-C} \exp\left(\min\left(1, \frac{1}{\log \frac{T}{\gamma}}\right)\right) e^{T\delta} \|\varphi\|_2.$$

## 5.5 The Finite Renewal Function

As in the previous chapter, the geodesic count comes from an analysis of a lattice point counting function which is close to our renewal function  $N_q$  for large sequences. The strategy is analogous to [Lal89] or [MOW17].

We define

$$N_q^*(X, a, g, f) := \sum_{n=0}^{\infty} \sum_{b \in \Gamma_a^n} g(b|a) f(\pi_q(b)) \mathbf{1}_{\{S_n \tau^*(b|a) \leq X\}}$$

where  $a \in \Sigma_R^* \cup \Sigma$  (see Section 4.3 for the definition of the space  $(\Sigma^* \cup \Sigma, \sigma)$  and  $\tau^*$ ). To count geodesics with congruence conditions, it suffices to provide an asymptotic for

$$\begin{aligned} N^*(T, \emptyset, \mathbf{1}, \delta_{\pi_q(g)}) &= \#\{b \in \Sigma_R^* : g_b \equiv g \pmod{q}, d(j, g_b j) < T\} \\ &= \#\{b \in \Sigma_R^* : g_b \equiv g \pmod{q}, \|g_b\| < \sqrt{2 \cosh T}\}. \end{aligned}$$

Iterating the finite renewal equation yields

$$\begin{aligned} N^*(X, a, g, f) &= \sum_{b \in \Gamma_a^n} N^*(X - S_n \tau^*(b), b, g, \rho(\pi_q(b)) f) \\ &\quad + \sum_{k=1}^{n-1} \sum_{b \in \Gamma_a^k} g(b|a) f(\pi_q(b)) \mathbf{1}_{\{S_k \tau^*(b) \leq X\}} + g(a) f(1) \mathbf{1}_{\{X \geq 0\}}. \end{aligned}$$

As  $X \rightarrow \infty$ , the second line does not change. For each summand corresponding to  $b \in \Gamma_a^n$  in the first line we can find  $b' \in \Sigma_R$  to sandwich between the two terms

$$N_q(T \pm 2^{-n+k} - S_n \tau^*(b'), b', g, \rho(\pi_q(b)) f) \asymp N_q^*(T - S_n \tau^*(b), b, g, \rho(\pi_q(b)) f).$$

So, it suffices to analyze  $N_q(T \pm 2^{-n+k} - S_n \tau^*(b'), b', g, \rho(\pi_q(b)) \mathbf{1}_{\{\pi_q(g)\}})$  (sending  $n \rightarrow \infty$  to get the final theorem). Let  $\varphi = \delta_{\pi_q(g)}$  and write

$$\varphi = \frac{\langle \varphi, 1 \rangle}{|\mathrm{SL}_2(q)|} \mathbf{1} + \varphi'$$

where  $\langle \varphi', 1 \rangle = 0$ . Then we have

$$\begin{aligned} \int_{-\infty}^{\infty} k_{\lambda}(t)N(T+t, a, \varphi)dT &= \frac{\langle \varphi, 1 \rangle}{|SL_2(q)|} \frac{1}{2\pi i} \int_{\delta+i\mathbb{R}} e^{as} F(s, a, 1) \hat{k}\lambda(s) ds \\ &\quad + \frac{1}{2\pi i} \int_{\delta+i\mathbb{R}} e^{as} F(s, a, \varphi') \hat{k}\lambda(s) ds. \end{aligned}$$

Observe  $N_q(T, a, 1) = N(T, a)$  where  $N(T, a)$  is the renewal function defined in the previous section. We established in the previous section that

$$N(T - O(2^{-N}), b) \leq N^*(T, a) \leq N(T + O(2^{-N}), b)$$

when  $a$  and  $b$  are in the same  $N$ -cylinder. So the first term is  $N^*(T, a) + O(e^{T(\delta-c)})$ .

Since  $N_q(T, a, g, f)$  is increasing in  $T$ , we have

$$N_q(T - \gamma, a, g, f) \leq \int_{-\gamma}^{\gamma} k_{\gamma}(t)N_q(T+t, a, g, f)dt \leq N(T + \gamma, a, g, f).$$

Sending  $\gamma \rightarrow 0$  and appealing to Theorem 14 gives

$$N_q(T, a, g, \varphi') \ll N(q)^C (e^{-T\epsilon_3} + e^{\frac{1}{\log T}}) e^{T\delta_R}.$$

After renaming constants, we have shown Theorem 4 from the introduction:

**Theorem 15.** *For each  $R > 8$ , there is some absolute spectral gap  $\Theta_R > 0$  and absolute constants  $c_R, C_R > 0$  such that for all square-free  $q \in \mathbb{Z}[i]$  and  $\omega \in SL_2(q)$  we have the estimate*

$$\left| \#\{g \in \Gamma_R \cap B_X : g \equiv \omega \pmod{q}\} - \frac{\#(\Gamma_R \cap B_X)}{|SL_2(q)|} \right| \ll_R \#(\Gamma_R \cap B_X) E(q, X)$$

as  $X \rightarrow \infty$ , where

$$E(q, X) = \begin{cases} e^{-c_R \sqrt{\log X}} & N(q) < C_R \log X \\ N(q)^{C_R} X^{-\Theta_R} & N(q) > C_R \log X \end{cases}.$$

## Chapter 6

### The Growth Parameter $\delta_R$

Recall from Chapter 4 that we proved the following asymptotic

$$\#\Gamma_R \cap B_X \asymp X^{2\delta_R}.$$

The exponent  $\delta_R$  was the unique solution to the pressure equation

$$P(s) = \log \lambda_s = 0.$$

In other words, the function  $s \mapsto \lambda_s$  (where  $\lambda_s$  is the maximal eigenvalue of  $\mathcal{L}_s$  on  $s \in \mathbb{S}$ ) is strictly decreasing in  $s$  and  $\lambda_{\delta_R} = 1$ .

We will make use of the fact that  $\delta_R$  is arbitrarily close to 2 as  $R \rightarrow \infty$  in a later section. Therefore, we now show that  $\delta = \delta(R) \rightarrow 2$  as  $R \rightarrow \infty$ . In order to proceed, we need to consider the action of the transfer operators on a different space where it becomes easier to compare  $\mathcal{L}_{s,R}$  for varying  $R$ .

Previously, we have considered transfer operators on the subshift of finite type  $(\Sigma_R, \sigma)$ . In order to show that  $\delta_R \rightarrow 2$  as  $R \rightarrow \infty$ , we will need to compare the dynamics in  $(\Sigma_R, \sigma)$  to that of the subshift on a countable alphabet  $(\Sigma, \sigma)$ .

Recall the set of admissible words in the countably infinite alphabet  $\mathcal{P}$  is

$$\Sigma := \{(p_1, p_2, \dots) : p_i \in \mathcal{P}, f(p_i) \supset p_{i+1}\}.$$

We work in the space of Hölder continuous functions  $\mathcal{F}_\rho(\Sigma)$  with the norm

$$\|\cdot\|_\rho = \|\cdot\|_\infty + |\cdot|_\rho.$$

Also let  $\mathcal{F}_\infty(\Sigma)$  be the space of continuous functions endowed with the sup norm.

Consider the infinite transfer operator

$$\mathcal{L}_{s,\infty}g(x) := \sum_{y \in \Sigma: \sigma(y)=x} e^{-s\tau(y)}g(y)$$

where the distortion function  $\tau$  is the same as before:  $\tau(y) = \log |f'(\pi(y))|$ . In order to establish that  $\mathcal{L}_{s,\infty}$  is a bounded linear operator, we need that  $s\tau$  is summable, i.e.

$$\sum_{y \in E^1} \left( \sup_{x \in \Sigma} e^{-s\tau(y||x)} \right) < \infty.$$

In the interior of each one-cylinder,  $f$  is defined as  $f(z) = (-1)^k(1/z - \lfloor 1/z \rfloor)$  and so  $|f'(z)| = 1/|z|^2$ . This bounds our weights by

$$|e^{-s\tau(y||x)}| \leq |\pi(y||x)|^{2\Re s}.$$

For each lattice point in  $\mathbb{Z}[i]$  falling in  $\mathcal{X}$ , we have between two and six adjacent one-cylinders. This gives

$$\sum_{y \in E^1} \left( \sup_{x \in \Sigma} e^{-s\tau(y||x)} \right) < 6 \sum_{\substack{z \in \mathbb{Z}[i] \\ |z| \geq 2, \Im z \geq 0}} \sup_{\substack{w \\ |w-z| < 1/\sqrt{2}}} |w|^{-2\Re s} < C \sum_{z \in \mathbb{Z}[i] \setminus \{0\}} |z|^{-2\Re s}.$$

So, for  $s\tau$  to be summable, it suffices that  $\Re s > 1$ .

In place of the irreducible and periodic properties of the finite subshifts  $(\Sigma_R, \sigma)$ , we must now have that  $(\Sigma, \sigma)$  is finitely primitive. In other words, there exists some  $k$  and finite subset  $W \subset E^k$  such that for each  $i, j \in E^1$  there is some  $\omega \in W$  such that  $i\omega j$  is admissible. This is clear from the proof we provided for the irreducibility of  $(\Sigma_R, \sigma)$ . We divided the one-cylinders into eight regions, each containing a full square. The image of each under  $f$  contained two squares which in turn maps to the full region. Thus, we can choose  $W \subset E^2$  of size 16.

With these properties, we may apply the Perron-Frobenius theory for subshifts on a countable alphabet. See [MU01] or [MU03] for proof of the following theorem

**Theorem 16.** *For the infinite transfer operator  $\mathcal{L}_{s,\infty}$ , as long as  $\Re s > 1$ , we have*

1. *The spectral radius of  $\mathcal{L}_s$  acting on either  $\mathcal{F}_\infty$  or  $\mathcal{F}_\rho$  is  $\lambda_{s,\infty} = e^{P(s\tau)}$*
2.  *$\lambda_{s,\infty}$  is a simple eigenvalue and has a corresponding eigenfunction  $h_{s,\infty} \in \mathcal{F}_\rho \subset \mathcal{F}_\infty$  which is positive.*
3. *The remainder of the spectrum on  $\mathcal{F}_\rho$  is in a disc centered at 0 with radius strictly smaller than  $\lambda_{s,\infty}$ .*



where  $P(s\tau)$  denotes the pressure function:

$$P(s\tau) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\omega \in E^n} \left( \sup_{x \in \Sigma} e^{sS_n\tau(\omega\|x)} \right).$$

Note that the topological pressure  $P(s\tau)$  is increasing in  $s$  and there is a unique 0. Combining [Sul84] with [Ser81], we have that this value is the critical exponent  $\delta(\Gamma)$  which is equal to the Hausdorff dimension of the limit set of  $\Gamma$  which is 2 (since the limit set has non-zero Lebesgue measure). Thus,  $\lambda_{2,\infty} = 1$ .

Information about  $\lambda_{s,R}$  will follow from Keller-Liverani Perturbation Theorem (see Appendix A of [PU18] or [LV98] for proof.)

We will work in the setting of the two norms  $\|\cdot\|_\infty \leq \|\cdot\|_\rho$  in the Banach space  $\mathcal{F}_\rho(\Sigma)$ . Define

$$|||Q||| := \sup\{\|Qf\|_\infty : \|f\|_\rho \leq 1\}$$

and consider a family of operators  $P_k$  compared to some ‘limit operator’  $P$ . The next theorem will require the four following conditions:

1. There are  $C, M > 0$  such that for all  $k, n \in \mathbb{N}$ ,

$$\|P_k^n\|_\infty \leq CM^n$$

.

2. There are  $C_1, C_2, \alpha > 0$  such that for all  $n, k \in \mathbb{N}$ ,

$$\|P_k^n f\|_\rho \leq C_2 \alpha^n \|f\|_\rho + C_2 M^n \|f\|_\infty$$

.

3. If  $z \in \sigma(P_n) \cap \overline{B}^c(0, \alpha)$ , then  $z$  is not in the residual spectrum of  $P_n$ .

4.  $|||P_k - P||| \rightarrow 0$  as  $k \rightarrow \infty$ .

Although the full Keller-Liverani Perturbation Theorem provides more refined information about the spectrum of  $P_k$  and  $P$ , we only need the following

**Theorem 17.** *Assume the family  $\{P_k\}_{k \in \mathbb{N}}$  and  $P$  satisfy conditions (1)-(4) above. If  $\lambda$  is a simple, isolated eigenvalue of  $P$ , then for every sufficiently large  $k$ , there exists a unique simple eigenvalue  $\lambda_k$  of  $p_k$  such that*

$$\lim_{k \rightarrow \infty} \lambda_k = \lambda.$$

We will apply the theorem to the following family:  $\mathcal{L}_{s,R}f := \mathcal{L}_{s,R}[\mathbf{1}_{E_R^1}f]$ . In other words, we'll only sum over the one-cylinders in  $\Sigma_R$ . Conditions (1) and (2) follow from  $\|L_{s,R}\|_\infty \leq \|L_{s,\infty}\|_\infty$  and  $\|L_{s,R}\|_\infty \leq \|L_{s,\infty}\|_\infty$ . Condition (3) is automatically satisfied by our choice of  $\|\cdot\|_\infty$  in  $\mathcal{F}_\rho(\Sigma)$ . In particular, the unit ball in  $\mathcal{F}_\rho(\Sigma)$  is  $\|\cdot\|_\infty$ -compact by Ascoli's theorem (see [PP90].) Therefore, we must now only establish property (4).

### 6.0.1 Perturbation estimates for large alphabets

Assuming  $\Re s > 1$ , we have

$$\begin{aligned} \|(L_{s,\infty} - L_{s,R})f(x)\|_\infty &= \left\| \sum_{y \in E^1 \setminus E_R^1} e^{s\tau(y\|x)} f(y\|x) \right\|_\infty \\ &< C \|f\|_\infty \sum_{z \in \mathbb{Z}[i], |z| > R-1} |z|^{-2\Re s} \\ &< C \frac{\|f\|_\infty}{R}. \end{aligned}$$

So the Keller-Liverani perturbation theorem implies  $\lambda_{s,R} \rightarrow \lambda_{s,\infty}$  as  $R \rightarrow \infty$  when  $\Re s > 1$ . Since  $\lambda_{2,\infty} = 1$  and  $\lambda_{s,R}$  is analytic in  $s$  for fixed  $R$  we have that the solution to  $\lambda_{s,R} = 1$  lies in an epsilon neighborhood of 2 for large enough  $R$ .

Finally, we must relate the eigenvalues of  $\mathcal{L}_{s,R}$  on  $(\Sigma, \sigma)$  to those of  $\mathcal{L}_s$  on  $(\Sigma_R, \sigma)$ . It suffices to show that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\omega \in E_R^n} \left( \sup_{x \in \Sigma} e^{sS_n \tau(\omega\|x)} \right) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\omega \in E_R^n} \left( \sup_{x \in \Sigma_R} e^{sS_n \tau(\omega\|x)} \right) \\ (P_1) &= (P_2) \end{aligned}$$

since the left hand side is the log of the lead eigenvalue for  $L_{s,R}$  on  $(\Sigma, \sigma)$  and the right hand side is the log of the lead eigenvalue of  $L_s$  on  $(\Sigma_R, \sigma)$ .

Combining the fact that the diameter of an  $n$  cylinder is at most  $2^{-n}$  with bounded distortion, we have

$$\sup_{x \in \Sigma} e^{sS_n \tau(w \| x)} = \sup_{x \in \Sigma_R} e^{sS_n \tau(w \| x)} + O_R(2^{-n})$$

where the bound is uniform over  $w \in E_R^n$  for fixed  $s$ . Inserting this into  $(P_1)$  gives

$$(P_1) = (P_2) + \lim_{n \rightarrow \infty} \frac{1}{n} \log \left( 1 + \frac{\sum_{w \in E_R^n} \sup_{x \in \Sigma_R} e^{sS_n \tau(w \| x)}}{C \sum_{w \in E_R^n} 2^{-n}} \right)$$

For  $w \in E_R^n$  and  $x \in \Sigma_R$ , bounded distortion yields  $|e^{sS_n \tau(w \| x)}| \in [(1/R)^{2\Re s}, (1/2)^{2\Re s}]$ .

So,

$$(P_1) = (P_2) + \lim_{n \rightarrow \infty} \frac{1}{n} \log \left( 1 + O \left( \left( \frac{2}{R} \right)^n \right) \right) = (P_2).$$

## Chapter 7

### Construction of the Multilinear Sifting Set

Recall from Chapter 3, that the set of finite admissible words

$$\Sigma_R^* = \{(p_1, \dots, p_n) : p_i \leftrightarrow \rho_i \in \mathcal{P}_R, f(\rho_i) \supset \rho_{i+1}\}$$

represents closed geodesics in  $\mathrm{SL}_2(\mathbb{Z}[i]) \backslash \mathbb{H}^3$ . Specifically, when we restrict to aperiodic words in  $\Sigma_R^*$  we have an exact correspondence. In order to detect whether a geodesic is fundamental, we will need to use the correspondence between closed geodesics and hyperbolic matrices in  $\mathrm{SL}_2(\mathbb{Z}[i])$ . We define

$$\begin{aligned} \pi : \Sigma_R^* &\rightarrow \mathrm{SL}_2(\mathbb{Z}[i]) \\ a &\mapsto f|_{C_a} \end{aligned}$$

where  $C_a$  is the cylinder  $\rho_{a_1} \cap f^{-1}(\rho_{a_2}) \cap \dots \cap f^{-k-1}(\rho_{a_k})$  and  $f$ , the Pollicott map, acts as a fractional linear transformation locally on that cylinder. Therefore, we may express the correspondence between  $\Sigma_R^*$  and closed geodesics (as primitive hyperbolic matrices in  $\mathrm{SL}_2(\mathbb{Z}[i])$ ) as

$$\Gamma_R := \{\pi(a) : a \in \Sigma_R^*, a \text{ is aperiodic}\}.$$

In order to develop an asymptotic estimate for

$$\#\{\omega \in \Gamma_R : \mathrm{tr}^2(\omega) - 4 \text{ square-free}, \|\omega\| < X\}$$

we will use the multilinear structure coming from  $\Sigma_R^*$ .

In order to construct a large set of geodesics with the desired expansion properties, we specify the general expansion result from a previous chapter, so that we have a fixed radius  $R$  and corresponding spectral gap  $\Theta$ :

**Theorem 18.** *There is an absolute  $\Theta > 0$  (fixed throughout the remaining sections) and  $c, C > 0$  such that for any square-free  $q \in \mathbb{Z}[i]$*

$$\#\{\gamma \in \Gamma_8 \cap B_Y : \gamma \equiv \alpha \pmod{q}\} = \frac{\#(\Gamma_8 \cap B_Y)}{|SL_2(q)|} + O(\#(\Gamma_8 \cap B_Y)E(Y, q))$$

as  $Y \rightarrow \infty$  where

$$E(Y, q) = \begin{cases} e^{-c\sqrt{\log Y}} & N(q) \leq C \log Y \\ N(q)^C Y^{-\Theta} & N(q) > C \log Y \end{cases}.$$

We also have from Chapter 4 that  $\#(\Gamma_8 \cap B_Y) \asymp X^{2\delta_8}$ . In order to increase the size of the sifting set, we embed  $\Gamma_8$  inside a larger subset of geodesics. In particular, define

$$\Xi_0 := \{a \in \Sigma_R^* : \pi(a) \in B_X\}, \quad \Omega_0 := \{a \in \Sigma_R^* : \pi(a) \in B_Z\}$$

while

$$\aleph_0 := \{a \in \Sigma_8^* : \pi(a) \in B_Y\}$$

We would like to construct a set which is the product of  $\Xi_0$ ,  $\aleph_0$ , and  $\Omega_0$  such that any combination of three elements from the respective sets gives a unique geodesic. To this end, we recall the notation from page 27. Specifically, if  $a||b$  denotes the admissible concatenation of two finite words. There are two possible issues we must address:

1. For each  $\xi \in \Xi_0$ ,  $a \in \aleph_0$ ,  $\omega \in \Omega_0$ , the concatenation  $\xi||a||\omega$  may not be admissible.
2. For each  $\xi \in \Xi_0$ ,  $a \in \aleph_0$ ,  $\omega \in \Omega_0$ , the concatenation  $\xi||a||\omega$  may not be unique.

Since the length of words in  $\Xi_0, \aleph, \Omega_0$  may vary it is possible that the same product may occur at different concatenation spots.

To avoid (2), we establish a uniform length of the words in  $\Omega_0$  and  $\Xi_0$ . For (1), we will add a universal transition between  $\xi$  and  $a$  as well as  $a$  and  $\omega$ .

In  $\Gamma_R$ , wordlength is commensurate with the log of the norm. Specifically, for  $a \in \Sigma_R^*$  we have  $l(a) \asymp \|\pi(a)\|$ . Since

$$\#\Xi_0 = \#(\Gamma_R \cap B_X) \asymp X^{2\delta_R}$$

there is some  $l_x$  so that

$$\Xi := \{a \in \Sigma_R^* | a \in \Xi_0, l(a) = l_x\} \subset \Xi_0$$

has size  $\gg X^{2\delta_R}/\log X$ . Similarly, we can find  $l_z$  and

$$\Omega := \{a : a \in \Sigma_R^* | a \in \Omega_0, l(a) = l_z\} \subset \Omega_0$$

of size  $\gg Z^{2\delta_R}/\log Z$ .

Finally, we address the issue that  $\xi||a||\omega$  may not be an admissible concatenation in  $\Sigma_R^*$ . Recall from Chapter 3 that  $(\Sigma_R, \sigma)$  is irreducible and aperiodic as long as  $R > 3$ . In the course of proving these properties, we showed that  $\hat{f}^3(p_1) \supset p_2$  for any two states in  $p_1, p_2 \in \mathcal{P}_R$ . Therefore for any two words  $a, b \in \Sigma_R^*$  there is a word  $\iota$  of length 3 such that  $a||\iota||b \in \Sigma_R^*$ . Note that  $\iota$  only depends on the final letter of  $a$  and the first letter of  $b$ . So we can arbitrarily choose a dictionary of three letter words  $\iota$  such that  $a||\iota||b$  is admissible and abbreviate the new word as  $a||b$ . Thus, we define our sifting set as

$$\Pi = \{\xi||a||\omega : \xi \in \Xi, a \in \mathfrak{N}, \omega \in \Omega\}.$$

Since  $\pi(\Xi) \subset B_X, \pi(\mathfrak{N}) \subset B_Y, \pi(\Omega) \subset B_Z$  there is some universal  $C$  such that  $\pi(\Pi) \subset B_{CN}$  where  $N = XYZ$ .

Write

$$\begin{aligned} |\mathcal{U}_{\mathfrak{q}}| &= \sum_{\varpi \in \Pi} \mathbf{1}_{\{\text{tr}^2(\pi(\varpi)) - 4 \equiv 0(\mathfrak{q})\}} = \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\varpi \in \Pi} \mathbf{1}_{\{\text{tr}(\pi(\varpi)) \equiv t(\mathfrak{q})\}} \\ &= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\varpi \in \Pi} \frac{1}{N(\mathfrak{q})} \sum_{q|N(\mathfrak{q})} \sum_{\chi(q)} \chi(\text{tr}(\pi(\varpi)) - t) \end{aligned}$$

where  $\sum_{q|N(\mathfrak{q})}$  is a sum over real  $q$  and  $\sum_{\chi(q)}$  ranges over all characters of  $\mathbb{Z}[i]/(\mathfrak{q})$  of order  $q$ . We will now separate the sum into characters of small and large order.

## Chapter 8

### Small Divisors

We estimate

$$\mathcal{M}_{\mathfrak{q}} := \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\varpi \in \Pi} \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \leq Q_0}} \sum_{\chi(q)} \chi(\mathrm{tr}(\pi(\varpi)) - t)$$

Expand from the definition of the sifting set  $\Pi$  and partition  $\aleph$  into residues classes in  $\mathrm{SL}_2(q)$ :

$$\begin{aligned} \mathcal{M}_{\mathfrak{q}} &= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\substack{\xi a \omega \in \Pi \\ \omega \in \Omega}} \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \leq Q_0}} \sum_{\chi(q)} \chi(\mathrm{tr}(\pi(\xi a \omega)) - t) \\ &= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\substack{\xi \in \Sigma \\ \omega \in \Omega}} \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \leq Q_0}} \sum_{\chi(q)} \sum_{a_0 \in \mathrm{SL}_2(q)} \chi(\mathrm{tr}(\pi(\xi a_0 \omega)) - t) \left[ \sum_{\substack{a \in \aleph \\ a \equiv a_0(q)}} 1 \right]. \end{aligned}$$

Now, using Theorem 18 from page 54 we write  $\mathcal{M}_{\mathfrak{q}} = \mathcal{M}_{\mathfrak{q}}^{(1)} + r^{(1)}(\mathfrak{q})$  where

$$\begin{aligned} \mathcal{M}_{\mathfrak{q}}^{(1)} &:= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\substack{\xi \in \Sigma \\ \omega \in \Omega}} \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \leq Q_0}} \sum_{\chi(q)} \frac{|\aleph|}{|\mathrm{SL}_2(\mathfrak{q})|} \sum_{a_0 \in \mathrm{SL}_2(q)} \chi(\mathrm{tr}(\pi(\xi) a_0 \pi(\omega)) - t), \\ r^{(1)}(\mathfrak{q}) &:= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \sum_{\substack{\xi \in \Sigma \\ \omega \in \Omega}} \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \leq Q_0}} \sum_{\chi(q)} \sum_{a_0 \in \mathrm{SL}_2(q)} \chi(\mathrm{tr}(\pi(\xi) a_0 \pi(\omega)) - t) |\aleph| E(\mathfrak{q}, Y). \end{aligned}$$

First, we address  $|r^{(1)}(\mathfrak{q})|$ :

$$\sum_{N(\mathfrak{q}) < \mathcal{Q}} |r^{(1)}(\mathfrak{q})| \ll \sum_{N(\mathfrak{q}) < \mathcal{Q}} \tau(\mathfrak{q}) \frac{|\Pi|}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \leq Q_0}} q^4 q^C Y^{-\Theta} \ll |\Pi| (\log \mathcal{Q})^2 Q_0^C Y^{-\Theta}.$$

Back to  $\mathcal{M}_{\mathfrak{q}}^{(1)}$ , we want to sum over all divisors of  $\mathfrak{q}$ , not only the small ones. So, we reintroduce the large divisors by writing  $\mathcal{M}_{\mathfrak{q}}^{(2)} = \mathcal{M}_{\mathfrak{q}}^{(1)} + r^{(2)}(\mathfrak{q})$  where

$$\begin{aligned}\mathcal{M}_{\mathfrak{q}}^{(2)} &:= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \frac{|\Pi|}{N(\mathfrak{q})} \sum_{q|N(\mathfrak{q})} \sum_{\chi(q)} \frac{1}{|\mathrm{SL}_2(\mathfrak{q})|} \sum_{a_0 \in \mathrm{SL}_2(q)} \chi(\mathrm{tr}(a_0) - t), \\ r^{(2)}(\mathfrak{q}) &:= \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \frac{|\Pi|}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q > Q_0}} \sum_{\chi(q)} \frac{1}{|\mathrm{SL}_2(\mathfrak{q})|} \sum_{a_0 \in \mathrm{SL}_2(q)} \chi(\mathrm{tr}(a_0) - t).\end{aligned}$$

Observe that we collapsed the sums over  $\Sigma$  and  $\Omega$  since they only reindex the sum over  $\mathrm{SL}_2(q)$ .

Now, define

$$\begin{aligned}\rho_t(\mathfrak{p}) &:= \frac{1}{|\mathrm{SL}_2(\mathfrak{p})|} \sum_{\gamma \in \mathrm{SL}_2(\mathfrak{p})} \sum_{\chi(\mathfrak{p})} \chi(\mathrm{tr}(\pi(\gamma)) - t) \\ &= \frac{1}{|\mathrm{SL}_2(\mathfrak{p})|} \sum_{\gamma \in \mathrm{SL}_2(\mathfrak{p})} (N(\mathfrak{p}) - 1) \mathbf{1}_{\{\mathrm{tr}(\pi(\gamma)) = t(\mathfrak{p})\}} + (-1) \mathbf{1}_{\{\mathrm{tr}(\pi(\gamma)) \neq t(\mathfrak{p})\}} \\ &= \frac{N(\mathfrak{p}) (\#\{\gamma \in \mathrm{SL}_2(\mathfrak{p}) : \mathrm{tr}(\pi(\gamma)) = t(\mathfrak{p})\}) - |\mathrm{SL}_2(\mathfrak{p})|}{|\mathrm{SL}_2(\mathfrak{p})|}.\end{aligned}$$

**Lemma 6.** For  $t = \pm 2 \bmod \mathfrak{p}$ ,

$$\#\{\gamma \in \mathrm{SL}_2(\mathfrak{p}) : \mathrm{tr}(\pi(\gamma)) = t(\mathfrak{p})\} = N(\mathfrak{p})^2$$

*Proof.* We may as well assume  $t = 2$ , the case when  $t = -2$  is similar. We partition

$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  into two cases according to whether or not  $c = 0$ :

1.  $c = 0$ : In this case, we must have  $ad = 1$  hence  $a + d = a + a^{-1} = 2$ . This only happens when  $x^2 - 2x + 1$  has a root in  $\mathbb{F}_{\mathfrak{p}}$ . Examining the discriminant, we see that there is exactly one solution. So, there's only one choice for  $a$  and  $d$ . There are  $N(\mathfrak{p})$  choices for  $b$ .
2.  $c \neq 0$ : There are  $N(\mathfrak{p}) - 1$  nonzero choices of  $c$ . In this case, there are  $N(\mathfrak{p})$  choices for  $a$  after which  $d$  is determined. The determinant equation implies  $b = c^{-1}(ad - 1)$ .

Combining the two cases, we have  $(N(\mathfrak{p}) - 1)N(\mathfrak{p}) + N(\mathfrak{p}) = N(\mathfrak{p})^2$  total matrices in  $\mathrm{SL}_2(\mathfrak{p})$  with trace 2.  $\square$



Therefore, we have

$$\rho_t(\mathfrak{p}) = \frac{N(\mathfrak{p})^3 - (N(\mathfrak{p})^3 - N(\mathfrak{p}))}{N(\mathfrak{p})^3 - N(\mathfrak{p})} = \frac{1}{N(\mathfrak{p})^2 - 1}.$$

We can rewrite

$$\mathcal{M}_{\mathfrak{q}}^{(2)} = |\Pi| \sum_{\substack{t \bmod \mathfrak{q} \\ t^2 \equiv 4(\mathfrak{q})}} \prod_{\mathfrak{p}|\mathfrak{q}} \frac{(1 + \rho_t(\mathfrak{p}))}{N(\mathfrak{p})} = |\Pi| \prod_{\mathfrak{p}|\mathfrak{q}} (1 + \mathbf{1}_{\{p \neq (1+i)\}}) \frac{(1 + \rho_t(\mathfrak{p}))}{N(\mathfrak{p})}.$$

Above, we used the Chinese Remainder Theorem to count  $t \bmod \mathfrak{q}$  satisfying  $t^2 \equiv 4(\mathfrak{q})$ .

As long as  $p \neq (1+i)$  we will have  $\pm 2$  are distinct which gives two solutions to  $t^2 \equiv 4(\mathfrak{p})$ .

If  $p = (1+i)$ , then we only have one solution.

Finally, we bound  $\sum_{\mathfrak{q} < \mathcal{Q}} |r^{(2)}(\mathfrak{q})|$ . Note that  $|\rho_t(\mathfrak{p})| \leq \frac{1}{N(\mathfrak{p})}$  so

$$|r^{(2)}(\mathfrak{q})| \ll \tau(\mathfrak{q}) |\Pi| \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|\mathfrak{q} \\ q \geq Q_0}} \frac{1}{N(q)} \ll |\Pi| \frac{N(\mathfrak{q})^\epsilon}{N(\mathfrak{q})} \frac{1}{Q_0}.$$

Then  $\sum_{N(\mathfrak{q}) < \mathcal{Q}} |r^{(2)}(\mathfrak{q})| \ll |\Pi| \frac{\mathcal{Q}^\epsilon}{Q_0}$ .

## Chapter 9

### Large Divisors

We first need to establish the existence of a smooth cutoff function which will allow us to extend sums over  $\Xi$  and  $\Omega$  to all of  $\mathrm{SL}_2(\mathbb{Z}[i])$ .

#### 9.1 Spectral Theory of the Laplace Operator

The Laplacian  $\Delta = z^2(\partial_{xx} + \partial_{yy} + \partial_{zz}) - z\partial_z$  acts on  $L^2(\mathrm{SL}_2(\mathbb{Z}[i])\backslash\mathbb{H}^3)$ . There are a finite number of discrete eigenvalues in  $[0, 1)$ :

$$0 = \lambda_0 < \lambda_1 \leq \cdots \lambda_K < 1.$$

In  $[1, \infty)$ , there is the continuous spectrum as well as the remaining part of the discrete spectrum (see [Sel56] or [Sar83].) For the congruence subgroup  $\Gamma(q)$  we have the analogue of Selberg's 3/16-Theorem:  $\lambda_1(q) = \delta(2 - \delta) \geq 3/4$  (see [Sar83] or Theorem 6.1 of [EGM13] for proof.) In other words,  $\delta \leq 3/2$ .

Denote  $G = \mathrm{SL}_2(\mathbb{C})$ ,  $K = \mathrm{SU}(2)$ ,  $V := L^2(\Gamma(q)\backslash G)$

$$V = \bigoplus_{\lambda_j < \theta(2-\theta)} V_{\lambda_j} \oplus V^\perp$$

where  $V_{\lambda_j}$  is a complementary series representation of parameter  $s_j$  and  $V^\perp$  does not weakly contain any complementary series representation of parameter  $s > \theta$ .

The following is standard, see [BK14] or [KO11].

**Theorem 19.** *Let  $\Theta$  and  $(\pi, V)$  be a unitary representation of  $G$  which does not weakly contain any complementary series representation with parameter  $s > \Theta$ . Then for any right  $K$ -invariant vectors  $\Psi_1, \Psi_2 \in V$*

$$|\langle \pi(g) \cdot \Psi_1, \Psi_2 \rangle| \ll \|g\|^{-2(2-\Theta)} \|\Psi_1\| \|\Psi_2\|$$

as  $\|g\| \rightarrow \infty$ .

We pick a nonnegative smooth bump function  $\psi$  satisfying  $\int_{G/K} \psi = 1$ . The support of  $\psi$  is in a ball about the identity of (tiny) radius  $\eta$ . For large  $X$ , we define our indicator function on  $\mathrm{SL}_2(\mathbb{C})$  as

$$\varphi_X(g) := \int_{G/K} \int_{G/K} \mathbf{1}_{\{\|h_1^{-1}gh_2\| < X\}} \psi(h_1)\psi(h_2)dh_1dh_2.$$

With  $\eta$  small enough, we have

$$\varphi_X(g) = \begin{cases} 1 & \text{if } \|g\| < (1 - \epsilon)X \\ 0 & \text{if } \|g\| > (1 + \epsilon)X \\ \in [0, 1] & \text{otherwise} \end{cases}.$$

Since the support of  $\varphi_X$  is within a ball of radius  $(1 + \epsilon)X$ , we have the following

$$\sum_{\gamma \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\gamma) \ll \#\{\gamma \in \mathrm{SL}_2(\mathbb{Z}[i]) : \|\gamma\| < C_1X\} \ll X^4$$

Now, we wish to establish that  $\varphi_X$  assigns roughly equal weights to residue classes in  $\mathrm{SL}_2(p)$ .

**Proposition 1.** *For squarefree  $q$  and  $\gamma_0 \in \mathrm{SL}_2(q)$ , we have*

$$\sum_{\substack{\xi \in \mathrm{SL}_2(\mathbb{Z}[i]) \\ \xi \equiv \gamma_0(q)}} \varphi_X(\xi) = \frac{1}{|\mathrm{SL}_2(q)|} \sum_{\xi \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\xi) + O(X^3).$$

We follow along with the proof found in [BK14].

*Proof.*

$$\sum_{\substack{\xi \in \mathrm{SL}_2(\mathbb{Z}[i]) \\ \xi \equiv \gamma_0(q)}} \varphi_X(\xi) = \sum_{\gamma \in \Gamma(q)} \varphi_X(\gamma\gamma_0) = \int_{G/K} \int_{G/K} \sum_{\gamma \in \Gamma(q)} \mathbf{1}_{\{\|h_1^{-1}\gamma\gamma_0h_2\| < X\}} \psi(h_1)\psi(h_2)dh_1dh_2.$$

Define

$$\mathcal{F}_q(h, g) = \sum_{\gamma \in \Gamma(q)} \mathbf{1}_{\{\|h^{-1}\gamma g\| < X\}},$$

$$\Psi_q(g) = \sum_{\gamma \in \Gamma(q)} \psi(\gamma g),$$

$$\Psi_{q, \gamma_0}(g) = \sum_{\gamma \in \Gamma(q)} \psi(\gamma_0^{-1}\gamma g)$$

so that we have the following identity (via a standard folding and unfolding argument):

$$\begin{aligned} \sum_{\substack{\xi \in \mathrm{SL}_2(\mathbb{Z}[i]) \\ \xi \equiv \gamma_0(q)}} \varphi_X(\xi) &= \langle \mathcal{F}_q, \Psi_q \otimes \Psi_{q, \varphi_0} \rangle_q \\ &= \int_{G/K} \mathbf{1}_{\{\|h\| < X\}} \langle \pi(h) \cdot \Psi_q, \Psi_{q, \gamma_0} \rangle_q dh \end{aligned}$$

Where  $\langle \cdot, \cdot \rangle_q$  denotes the inner product on  $L^2(\Gamma(q) \backslash G/K)$ . Now, decompose the matrix coefficient into the following

$$= \int_{G/K} \mathbf{1}_{\{\|h\| < X\}} \langle \pi(h) \cdot \Psi_q|_V, \Psi_{q, \gamma_0}|_V \rangle_q dh + \int_{G/K} \mathbf{1}_{\{\|h\| < X\}} \langle \pi(h) \cdot \Psi_q|_{V^\perp}, \Psi_{q, \gamma_0}|_{V^\perp} \rangle_q dh$$

where  $|_V$  denotes projection onto the old forms and  $V^\perp$  is orthogonal.

$$\text{Claim 1: } \langle \pi(h) \cdot \Psi_q|_V, \Psi_{q, \gamma_0}|_V \rangle_q = \frac{1}{[\Gamma(q) : \Gamma]} \langle \pi(h) \cdot \Psi_1|_V, \Psi_{1, \gamma_0}|_V \rangle_1$$

If  $\varphi_1^{(q)}, \dots, \varphi_l^{(q)}$  denote the oldforms for the spectrum below  $\Theta(2 - \Theta)$ , then we can rewrite them as normalizations of the eigenfunctions at level one:

$$\varphi_i^{(q)} = \frac{1}{\sqrt{[\Gamma(q) : \Gamma]}} \varphi_i^{(1)}$$

The rest follows from folding and unfolding.

$$\text{Claim 2: } \int_{G/K} \mathbf{1}_{\{\|h\| < X\}} \langle \pi(h) \cdot \Psi_q|_{V^\perp}, \Psi_{q, \gamma_0}|_{V^\perp} \rangle_q dh \ll X^3$$

Since  $\eta$  is fixed and  $\psi$  has 'bounded' support, we have  $\mathcal{S}\Psi_q|_{V^\perp} \leq \mathcal{S}\Psi \ll 1$  (and similar for  $\Psi_{q, \gamma_0}$ ). So, Theorem 19 gives

$$\int_{G/K} \mathbf{1}_{\{\|h\| < X\}} \|h\|^{2\delta-4} dh \ll X^3.$$

Thus, we have shown

$$\sum_{\substack{\xi \in \mathrm{SL}_2(\mathbb{Z}[i]) \\ \xi \equiv \gamma_0(q)}} \varphi_X(\xi) = \frac{1}{[\Gamma(q) : \Gamma]} \langle \pi(h) \cdot \Psi_1|_V, \Psi_{1, \gamma_0}|_V \rangle_1 + O(X^3).$$

When we use the same argument, setting  $q = 1$ , the theorem follows from

$$\sum_{\xi \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\xi) = \langle \pi(h) \cdot \Psi_1|_V, \Psi_{1, \gamma_0}|_V \rangle_1 + O(X^3)$$

□

## 9.2 Character Sums

**Proposition 2.** *Suppose  $q$  is a rational prime and  $\mathfrak{q}$  lies over  $q$ ,  $\chi : \mathbb{Z}[i] \rightarrow S^1$  has order  $q$  and  $\xi \in \mathbb{Z}[i]^4$  satisfies  $(\xi, \mathfrak{q}) = 1$  (component-wise.) We have*

$$\left| \sum_{s \in \text{SL}_2(\mathfrak{q})} \chi_q(s \cdot \xi) \right| \ll N(\mathfrak{q})^{3/2}.$$

*Proof.* We may as well assume  $y \neq 0(\mathfrak{q})$ , otherwise we could alter the following argument on  $s$ .

We now partition all  $s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  into two cases, either  $c = 0$  or  $c \neq 0$ . If  $c \equiv 0(\mathfrak{q})$ , then  $ad - bc = 1$  implies  $d = a^{-1}(\mathfrak{q})$  and  $b$  is anything. Since  $b$  ranges over all  $\mathbb{Z}[i]/(\mathfrak{q})$ , we have

$$\sum_{\substack{s \in \text{SL}_2(\mathfrak{q}) \\ c=0(\mathfrak{q})}} \chi_q(s \cdot \xi) = \sum_{a(\mathfrak{q})} \sum_{b(\mathfrak{q})} \chi_q(ax + by + a^{-1}w) = \sum_{a(\mathfrak{q})} \chi_q(ax + a^{-1}w) \sum_{b(\mathfrak{q})} \chi_q(by) = 0.$$

On the other hand if  $c \neq 0$ , then we can pick any  $a, d$  which imply  $b = c^{-1}(ad - 1)$  so

$$\begin{aligned} \sum_{\substack{s \in \text{SL}_2(\mathfrak{q}) \\ c \neq 0(\mathfrak{q})}} \chi_q(s \cdot \xi) &= \sum_{c(\mathfrak{q})} \sum_{a, d(\mathfrak{q})} \chi_q(ax + c^{-1}(ad - 1)y + cz + dw) \\ &= \sum_{c(\mathfrak{q})} \chi_q(cz - c^{-1}y) \sum_{a(\mathfrak{q})} \chi_q(ax) \sum_{d(\mathfrak{q})} \chi_q(d(c^{-1}ay + w)). \end{aligned}$$

Now, note that

$$\sum_{d(\mathfrak{q})} \chi_q(d(c^{-1}ay + w)) = \begin{cases} 0 & a \not\equiv -cy^{-1}w(\mathfrak{q}) \\ N(\mathfrak{q}) & a \equiv -cy^{-1}w(\mathfrak{q}) \end{cases}.$$

Hence

$$\begin{aligned} \sum_{\substack{s \in \text{SL}_2(\mathfrak{q}) \\ c \neq 0(\mathfrak{q})}} \chi_q(s \cdot \xi) &= N(\mathfrak{q}) \sum_{c(\mathfrak{q})} \chi_q(cz - c^{-1}y) \chi_q(-cy^{-1}wx) \\ &= N(\mathfrak{q}) \sum_{c(\mathfrak{q})} \chi_q(c(z - y^{-1}wx) - c^{-1}y). \end{aligned}$$

Now, since  $y \neq 0$ , we have a nontrivial Kloosterman sum (or perhaps a Ramanujan sum if  $z - y^{-1}wx = 0$ .) Regardless of whether  $q$  is split or inert, we have

$$|K(\chi_q; a, b)| \leq 2N(q)^{1/2}.$$

When  $q$  splits, this is just the Weil bound. When  $q$  is inert, see Theorem 5.45 of [LN97].  $\square$

From Proposition 2, we have the following

**Corollary 2.** *For square-free  $\mathfrak{q}$ ,  $\chi$  a character of  $\mathbb{Z}[i]/(\mathfrak{q})$  of order  $N(\mathfrak{q})$  and  $(\xi, \mathfrak{q}) = 1$ , we have*

$$\left| \sum_{s \in \mathrm{SL}_2(\mathfrak{q})} \chi_q(s \cdot \xi) \right| \ll N(\mathfrak{q})^{3/2}.$$

*Proof.* Pontryagin duality allows us to express  $\chi_q$  as a product of characters of prime order. Apply Proposition 2 to each term.  $\square$

Now, we combine this bound with our indicator function  $\varphi$ :

**Proposition 3.** *For square-free  $q$  and a  $\chi_q$  character of order  $N(q)$ , we have*

$$\left| \sum_{\xi \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\xi) \chi_q(s \cdot \xi) \right| \ll N(\mathfrak{q})^{-3/2+\epsilon} X^4 + N(q)^3 X^3.$$

*Proof.* Partition over residue classes over  $\mathfrak{q}$  and apply Proposition 1:

$$\begin{aligned} \left| \sum_{\xi \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\xi) \chi_q(s \cdot \xi) \right| &= \left| \sum_{\gamma \in \mathrm{SL}_2(\mathfrak{q})} \chi_q(\gamma \cdot s) \sum_{\substack{\xi \in \mathrm{SL}_2(\mathbb{Z}[i]) \\ \xi \equiv \gamma(\mathfrak{q})}} \varphi_X(\xi) \right| \\ &= \left| \sum_{\gamma \in \mathrm{SL}_2(\mathfrak{q})} \chi_q(\gamma \cdot s) \right| \frac{X^4}{|\mathrm{SL}_2(\mathfrak{q})|} + O(N(\mathfrak{q})^3 X^3) \\ &\ll \frac{N(\mathfrak{q})^{3/2+\epsilon} X^4}{N(\mathfrak{q})^3} + N(\mathfrak{q})^3 X^3. \end{aligned}$$

$\square$

### 9.3 Large Divisors

Define

$$r(\mathfrak{q}) := \sum_{t^2 \equiv 4(\mathfrak{q})} \sum_{\varpi \in \Pi} \frac{1}{N(\mathfrak{q})} \sum_{\substack{q|N(\mathfrak{q}) \\ q \geq Q_0}} \sum_{\chi(q)} \chi(\mathrm{tr}(\pi(\varpi)) - t).$$

Our goal is to bound

$$\mathcal{E} := \sum_{N(\mathfrak{q}) < Q} |r(\mathfrak{q})| = \sum_{N(\mathfrak{q}) < Q} r(\mathfrak{q})\zeta(\mathfrak{q}),$$

where the sum ranges over square-free Gaussian ideals up to norm  $Q$  and  $\zeta(\mathfrak{q})$  captures the argument of  $r(\mathfrak{q})$ . In order to interchange the order of  $q$  and  $\mathfrak{q}$  in  $\mathcal{E}$  we introduce

$$\zeta_1(q, r) := q \sum_{\substack{N(\mathfrak{q}) < Q \\ N(\mathfrak{q}) \equiv 0(q)}} \frac{\zeta(\mathfrak{q})}{N(\mathfrak{q})} \sum_{t^2 \equiv 4(\mathfrak{q})} \chi_q^r(-t) \mathbf{1}_{\{\chi_q^r \in \widehat{\mathbb{Z}[i]/(\mathfrak{q})}\}}$$

for real  $q$  and  $r \in (\mathbb{Z}/q\mathbb{Z})^*$ . Insert  $\zeta_1$  into  $\mathcal{E}$ :

$$\mathcal{E} = \sum_{Q_0 \leq q < Q} \frac{1}{q} \sum'_{r(q)} \sum_{\varpi \in \Pi} \chi_q^r(\text{tr}(\pi(\varpi))) \zeta_1(q, r)$$

where the sum is over all square-free  $q \in [Q_0, Q]$ .

We now restrict our attention to short intervals in  $q$  and fixed  $a \in \mathbb{N}$ , i.e.

$$\begin{aligned} \mathcal{E}_1(Q, a) &:= \sum_{q \asymp Q} \left| \sum'_{r(q)} \zeta_1(q, r) \sum_{\xi \in \Sigma} \sum_{\omega \in \Omega} \chi_q^r(\text{tr}(\pi(\xi a \omega))) \right| \\ &= \sum_{q \asymp Q} \zeta_2(q) \sum'_{r(q)} \zeta_1(q, r) \sum_{\xi \in \Sigma} \sum_{\omega \in \Omega} \chi_q^r(\text{tr}(\pi(\xi a \omega))) \end{aligned}$$

where we introduced  $\zeta_2(q)$  to capture the absolute value of each term. Now apply Cauchy-Schwarz in the  $\xi$  parameter:

$$|\mathcal{E}_1|^2 \ll |\Sigma| \sum_{\xi \in \Sigma} \left| \sum_{q \asymp Q} \zeta_2(q) \sum'_{r(q)} \zeta_1(q, r) \sum_{\omega \in \Omega} \chi(\text{tr}(\pi(\xi a \omega))) \right|^2.$$

The support of  $\pi(\Sigma)$  is within  $B_X$ , so we replace sequences in  $\Sigma$  with matrices in  $B_X$  as follows

$$|\mathcal{E}_1|^2 \ll |\Sigma| \sum_{\gamma \in \text{SL}_2(\mathbb{Z}[i])} \varphi_X(\gamma) \left| \sum_{q \asymp Q} \zeta_2(q) \sum'_{r(q)} \zeta_1(q, r) \sum_{\omega \in \Omega} \chi_q^r(\gamma \cdot \pi(a\omega)) \right|^2$$

where we have replaced the trace of  $\pi(\xi a \omega)$  with the dot product  $\gamma \cdot \pi(a\omega)$  since

$$\text{tr}(\pi(\xi a \omega)) = \sum_{i,j} \pi(\xi)_{i,j} \pi(a\omega)_{i,j}.$$

Opening the square, we have

$$\begin{aligned}
|\mathcal{E}_1|^2 &\ll |\Sigma| \sum_{q,q'} \sum_{\substack{r(q) \\ r'(q')}} \zeta_2(q) \zeta_1(q, r) \overline{\zeta_2(q') \zeta_1(q', r')} \\
&\quad \sum_{\omega, \omega'} \sum_{\gamma \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\gamma) \chi_q^r(\gamma \cdot \pi(a\omega)) \chi_{q'}^{r'}(-\gamma \cdot \pi(a\omega')) \\
&\ll \mathcal{Q}^\epsilon |\Sigma| \sum_{q,q'} \sum_{\substack{r(q) \\ r'(q')}} \sum_{\omega, \omega'} \left| \sum_{\gamma \in \mathrm{SL}_2(\mathbb{Z}[i])} \varphi_X(\gamma) \chi_q^r(\gamma \cdot \pi(a\omega)) \chi_{q'}^{r'}(-\gamma \cdot \pi(a\omega')) \right|.
\end{aligned}$$

We would like to combine  $\chi_q$  and  $\chi_{q'}$ . Let  $\hat{q}$  be the least common multiple  $[q, q']$  and  $q_1 = \hat{q}/q$ ,  $q'_1 = \hat{q}/q'$  be the primes distinct to  $q$  and  $q'$  respectively. If  $b \equiv (q_1)^{-1}(q')$  and  $b' \equiv (q'_1)^{-1}(q)$  we have

$$\chi_q^r(\gamma \cdot \pi(a\omega)) \chi_{q'}^{r'}(-\gamma \cdot \pi(a\omega')) = \chi_{\hat{q}}(\xi \cdot (rq'_1 b' \pi(a\omega) - r'q_1 b \pi(a\omega'))).$$

If  $s := rq'_1 b' \pi(a\omega) - r'q_1 b \pi(a\omega')$ , then we may not have  $(s, \hat{q}) = 1$ . So, remove common factors to get  $(s', q_0) = 1$ . If  $\hat{q}_0 = \hat{q}/q_0$  represents all the factors we removed from  $\hat{q}$ , then notice

$$rq'_1 b' \pi(a\omega) \equiv r'q_1 b \pi(a\omega') \pmod{\hat{q}_0}.$$

Multiply (on the left) by  $\pi(a)^{-1} \in \mathrm{SL}_2(\hat{q}_0)$  to remove  $\pi(a)$ . Take determinants of both sides and since  $\det(\pi(\omega)) = \det(\pi(\omega')) = 1$ , we get

$$(rq'_1 b')^2 \equiv (r'q_1 b)^2 \pmod{\hat{q}_0}.$$

Now,  $\hat{q}_0 | (q, q')$  so  $(q_1, \hat{q}_0) = (r', \hat{q}_0) = (b, \hat{q}_0) = 1$ . Similarly,  $(q'_1 b' r, \hat{q}_0) = 1$ . So, we can find some  $u$  such that

$$q'_1 b' r \equiv u q_1 b r' \pmod{\hat{q}_0}$$

which satisfies  $u^2 \equiv 1(\hat{q}_0)$  by the squared congruence relation above - there are only  $N^\epsilon$  such  $u$ . We also have  $\pi(\omega) \equiv u \pi(\omega') \pmod{\hat{q}_0}$ .

Back to our estimate of  $\mathcal{E}_1$ , partition the sum over  $q$  and  $q'$  via their least common



multiple as follows:

$$|\mathcal{E}_1(Q, a)|^2 \ll |\Sigma| \sum_{\hat{q} \asymp Q} \sum_{q_1 q'_1 \hat{q}_0 \hat{q}_0 = \hat{q}} \sum_{\substack{u(\hat{q}_0) \\ u^2 \equiv 1(\hat{q}_0)}} \sum'_{r(q)} \sum_{\substack{r'(q') \\ q'_1 b' r \equiv u q_1 b r'(\hat{q}_0)}} \sum_{\omega' \in \Omega} \sum_{\substack{\omega \in \Omega \\ \pi(\omega) \equiv u \pi(\omega')(\hat{q}_0)}} \left| \sum_{\xi \in \text{SL}_2(\mathbb{Z}[i])} \varphi_X(\xi) \chi_{q_0}(\xi \cdot s) \right|.$$

Since the support of  $\pi(\Omega)$  is in  $B_Z$ , we replace the second sum over  $\Omega$  with one over  $\text{SL}_2(\mathbb{Z}[i])$  via  $\varphi_Z$ :

$$|\mathcal{E}_1(Q, a)|^2 \ll |\Sigma| \sum_{\hat{q} \asymp Q} \sum_{q_1 q'_1 \hat{q}_0 \hat{q}_0 = \hat{q}} \sum_{\substack{u(\hat{q}_0) \\ u^2 \equiv 1(\hat{q}_0)}} \sum'_{r(q)} \sum_{\substack{r'(q') \\ q'_1 b' r \equiv u q_1 b r'(\hat{q}_0)}} \sum_{\omega' \in \Omega} \sum_{\substack{\beta \in \text{SL}_2(\mathbb{Z}[i]) \\ \beta \equiv u \pi(\omega')(\hat{q}_0)}} \varphi_Z(\beta) \left| \sum_{\xi \in \text{SL}_2(\mathbb{Z}[i])} \varphi_X(\xi) \chi_{q_0}(\xi \cdot s) \right|.$$

Now, we apply our bounds for the smoothing function  $\varphi_X$ :

$$|\mathcal{E}_1(Q, a)|^2 \ll |\Sigma| \sum_{\hat{q} \asymp Q} \sum_{q_1 q'_1 \hat{q}_0 \hat{q}_0 = \hat{q}} \sum_{\substack{u(\hat{q}_0) \\ u^2 \equiv 1(\hat{q}_0)}} \sum'_{r(q)} \sum_{\substack{r'(q') \\ q'_1 b' r \equiv u q_1 b r'(\hat{q}_0)}} |\Omega| \left[ \frac{Z^4}{N(\hat{q}_0)^3} + Z^3 \right] \cdot \left[ N(q_0)^{-3/2+\epsilon} X^4 + N(q_0)^3 X^3 \right].$$

There are at most  $N(q)$  choices for  $r$  and then  $\frac{N(q')}{N(\hat{q}_0)}$  choices for  $r'$ . Therefore,

$$|\mathcal{E}_1(Q, a)|^2 \ll |\Sigma| \sum_{\hat{q} \asymp Q} N^\epsilon \sum_{q_0 \hat{q}_0 = \hat{q}} \frac{Q^2 N(q_0)}{N(\hat{q})} |\Omega| \left[ \frac{Z^4}{N(\hat{q}_0)^3} + Z^3 \right] \left[ N(q_0)^{-3/2+\epsilon} X^4 + N(q_0)^3 X^3 \right].$$

Now, we use the fact that  $|\Omega| \asymp \frac{Z^{2\delta}}{\log Z}$  and similarly  $|\Sigma| \asymp \frac{X^{2\delta}}{\log X}$  to insert  $|\Omega| Z^{-2\delta}$  and  $|\Sigma| X^{-2\delta}$ :

$$|\mathcal{E}_1(Q, a)|^2 \ll N^\epsilon Q^2 |\Sigma|^2 |\Omega|^2 (XZ)^{2(2-\delta)} \sum_{Q \ll N(\hat{q}) \ll Q^2} \frac{1}{N(\hat{q})} \left[ \frac{1}{N(\hat{q})^{1/2}} + \frac{1}{Z} + \frac{Q^8}{X} \right].$$

So, we get the following

**Theorem 20.**

$$|\mathcal{E}_q(Q, a)| \ll N^\epsilon Q |\Sigma| |\Omega| (XZ)^{2-\delta} \left[ \frac{1}{Q^{1/4}} + \frac{1}{Z^{1/2}} + \frac{Q^4}{X^{1/2}} \right].$$

Summing over  $a$  and  $Q$  gives

**Theorem 21.**

$$\mathcal{E} \ll N^\epsilon |\Pi| (XZ)^{2-\delta} \left[ \frac{1}{Q_0^{1/4}} + \frac{1}{Z^{1/2}} + \frac{Q^4}{X^{1/2}} \right].$$

## Chapter 10

### Sieve Theorem

We have

$$|U_q| = \beta(q)|\Pi| + r^{(1)}(q) + r^{(2)}(q) + r(q)$$

where

$$\begin{aligned} \sum_{N(q) < Q} |r^{(1)}(q)| &\ll |\Pi| (\log Q)^2 \left( e^{-c\sqrt{\log Y}} + Q_0^C Y^{-\Theta} \right), \\ \sum_{N(q) < Q} |r^{(2)}(q)| &\ll |\Pi| \frac{Q^\epsilon}{Q_0}, \\ \sum_{N(q) < Q} |r(q)| &\ll N^\epsilon |\Pi| (XZ)^{2-\delta} \left( \frac{1}{Q_0^{1/4}} + \frac{1}{Z^{1/2}} + \frac{Q^4}{X^{1/2}} \right). \end{aligned}$$

Recall that  $c$ ,  $C$  and  $\Theta$  are fixed constants coming from Theorem 18 page 54. By our construction of the sifting set (in which  $\Xi$  and  $\Omega$  depend on  $R$  but  $\aleph$  does not), as we send  $R \rightarrow \infty$  we can get  $\delta_R$  arbitrarily close to 2 while  $c$ ,  $C$ , and  $\Theta$  remain constant.

For  $r^{(1)}(q)$  and  $r^{(2)}(q)$  we need  $y > 0$ ,  $\alpha_0 > 0$  and  $\alpha_0 C < y\Theta$ , where  $x + y + z = 1$ .

For  $r(q)$  we need

$$\begin{aligned} \alpha_0/4 &> (x+z)(2-\delta), \\ z/2 &> (x+z)(2-\delta), \\ x/2 &> (x+z)(2-\delta) + 8\alpha. \end{aligned}$$

Observe from the last inequality that taking  $x$  near 1 and  $\delta_R$  near 2, we must have  $\alpha < 1/16$ . In order to achieve this level of distribution, we set  $\alpha = 1/16 - \eta$  and assume for now that  $\delta > 2 - \eta$ . In order to have  $|\Pi| \gg N^{2\delta-\eta}$  we set  $x = 1 - \eta$ .

For the remaining parameters, we set

$$z = \frac{\eta}{1 + C/\Theta}, \quad y = \frac{zC}{\Theta}, \quad \alpha_0 = \frac{5z}{6}$$

so that  $\alpha_0 C < y\Theta$ . Moreover if  $\delta > 2 - z/5 = 2 - \frac{\eta}{5(1+C/\Theta)}$  we have

$$z/2 > \alpha_0/4 = \frac{5z}{24} > \frac{z}{5} > 2 - \delta > (x+z)(2-\delta).$$

Hence the three inequalities for  $r(q)$  are satisfied. We get a power savings in each error term except  $r^{(1)}$  where we save an arbitrary power of  $\log$ . This proves the level of distribution for the following

**Theorem 22.** *For sufficiently small  $\eta$ , there is a large enough  $R$  so that  $U$  has level of distribution  $Q = T^{1/16-\eta}$ . In other words, there exists a multiplicative function  $\beta : \mathbb{Z}[i] \rightarrow \mathbb{R}$  satisfying*

$$\prod_{\substack{p \\ w \leq N(p) < z}} (1 - \beta(p))^{-1} \leq C \left( \frac{\log z}{\log w} \right)^2$$

for any  $2 \leq w < z$  and a decomposition

$$|U_q| = |\Pi|\beta(q) + r(q)$$

so that for all  $K$

$$\sum_{\substack{\text{square-free } q \\ N(q) < Q}} |r(q)| \ll_K \frac{|\Pi|}{\log^K N}.$$

Moreover, when

$$X = N^{1-\eta}$$

we have

$$|\Pi| \gg N^{2\delta-\eta}.$$

*Proof.* We must show that the sieve dimension is 2. The following summation formulas for primes in arithmetic progressions are consequences of Mertens work in [Mer74]:

$$\sum_{\substack{p \leq n \\ p \equiv 1(4)}} \frac{1}{p} = \frac{1}{2} \log \log n + B_1 + O\left(\frac{1}{\log n}\right),$$

$$\sum_{\substack{p \leq n \\ p \equiv 3(4)}} \frac{1}{p} = \frac{1}{2} \log \log n + B_3 + O\left(\frac{1}{\log n}\right).$$

For Gaussian primes, we only need the first equality:

$$\sum_{N(\mathfrak{p}) \leq n} \frac{1}{N(\mathfrak{p})} = \frac{1}{2} + \sum_{\substack{p \leq n \\ p \equiv 1(4)}} \frac{2}{p} + \sum_{\substack{p \leq \sqrt{n} \\ p \equiv 3(4)}} \frac{1}{p^2} = \log \log n + O(1).$$

Since  $\beta(\mathfrak{p}) \in (0, 1)$  for all  $p$ , we have

$$\begin{aligned} \prod_{w \leq N(\mathfrak{p}) \leq z} (1 - \beta(\mathfrak{p}))^{-1} &= \exp \left( - \sum_{w \leq N(\mathfrak{p}) \leq z} \log(1 - \beta(\mathfrak{p})) \right) \\ &= \exp \left( \sum_{w \leq N(\mathfrak{p}) \leq z} \sum_{k=1}^{\infty} \frac{\beta(\mathfrak{p})^k}{k} \right) \\ &= \exp \left( \sum_{w \leq N(\mathfrak{p}) \leq z} \beta(\mathfrak{p}) \right) \exp \left( \sum_{w \leq N(\mathfrak{p}) \leq z} \sum_{k>1} \frac{\beta(\mathfrak{p})^k}{k} \right). \end{aligned}$$

The second exponential is negligible since

$$\begin{aligned} S &= \sum_{\mathfrak{p}} \sum_{k>1} \frac{\beta(\mathfrak{p})^k}{k} \leq \frac{1}{2} \sum_{\mathfrak{p}} \beta(\mathfrak{p})^2 \sum_{k=0}^{\infty} \beta(\mathfrak{p})^k \\ &< \frac{1}{2} \sum_{\mathfrak{p}} \frac{\beta(\mathfrak{p})^2}{(1 - \beta(\mathfrak{p}))} \\ &\ll \sum_{p \equiv 1(4)} 2 \frac{b(p)^2}{1 - b(p)} + \sum_{p \equiv 3(4)} \frac{b(p^2)^2}{1 - b(p^2)}, \end{aligned}$$

where we have defined

$$b(x) = \frac{2}{x} \left( 1 + \frac{1}{x^2 - 1} \right).$$

Since  $\sum_n b(n)^2$  converges and  $(1 - \beta(n))^{-1} \rightarrow 1$  partial summation gives  $S < \infty$ . Hence

$$\begin{aligned} \prod_{w \leq N(\mathfrak{p}) \leq z} (1 - \beta(\mathfrak{p}))^{-1} &\leq C \exp \left( \sum_{w \leq N(\mathfrak{p}) \leq z} \beta(\mathfrak{p}) \right) \\ &< C \exp \left( \sum_{w \leq N(\mathfrak{p}) \leq z} \frac{2}{N(\mathfrak{p})} \right) \\ &< C \exp \left( 2 \log \left( \frac{\log z}{\log w} \right) \right). \end{aligned}$$

□

From Theorem 22 and the Fundamental Lemma of sieve theory (see Lemma 6.3 of [IK04]) we have the following:

**Theorem 23.** *Define*

$$\Pi_{AP} := \{\varpi \in \Pi : p | (tr^2(\pi(\varpi)) - 4) \implies N(p) > N^\alpha\}.$$

*We have*

$$\#\Pi_{AP} > N^{2\delta-\eta}.$$

## Chapter 11

### Final Estimates

We have the lower bound

$$\#\Pi_{AP} = \#\{\varpi \in \Pi : p | (\text{tr}^2(\pi(\varpi)) - 4) \implies N(p) > N^\alpha\} > N^{2\delta-\eta}.$$

On the other hand,

$$\begin{aligned} & \#\{\gamma \in \Gamma_R \cap B_N : (\text{tr}^2(\pi(\gamma)) - 4) \text{ is square-free}\} \\ & \geq \#\{\gamma \in \Pi_{AP} : (\text{tr}^2(\pi(\gamma)) - 4) \text{ is square-free}\} \\ & > N^{2\delta-\eta} - \#\{\gamma \in \Pi_{AP} : (\text{tr}^2(\pi(\gamma)) - 4) \text{ is not square-free}\}. \end{aligned}$$

We will examine the last term

$$\Pi_{AP}^\square := \{\gamma \in \Pi_{AP} : (\text{tr}^2(\pi(\gamma)) - 4) \text{ is not square-free}\}$$

more closely. If  $\gamma \in \Pi_{AP}^\square$ , then we can find a prime  $p$  with  $p^2 | (\text{tr}^2(\pi(\gamma)) - 4)$ . On one hand,  $\text{tr}^2(\pi(\gamma)) - 4$  factors as  $\text{tr}(\pi(\gamma)) \pm 2$  so  $N(p) \ll N^{1/2}$ . On the other hand,  $\Pi_{AP}^\square \subset \Pi_{AP}$  implies  $N(p) > N^\alpha$ . Therefore,

$$\begin{aligned} \#\Pi_{AP}^\square & \leq \sum_{N^\alpha < N(p) \ll N^{1/2}} \sum_{\substack{N(t) < N \\ t^2 - 4 \equiv 0(p^2)}} \#\{\gamma \in \Gamma_R \cap B_N : \text{tr}(\pi(\gamma)) = t\} \\ & \ll \sum_{N^\alpha < N(p) \ll N^{1/2}} \frac{N^2}{N(p)^2} N^{2+\epsilon} \ll N^{4-\alpha+\epsilon}, \end{aligned}$$

where we trivially bounded the trace multiplicity, i.e.

$$\#\{\gamma \in \Gamma_R \cap B_N : \text{tr}(\pi(\gamma)) = t\} < \#\{s \in B_N : \text{tr}(s) = t\} \ll N^{2+\epsilon}.$$

So, as long as  $2\delta - \eta > 4 - \alpha + \epsilon$ , we have:

**Theorem 24.** *There is an  $R$  such that as  $N \rightarrow \infty$  we have*

$$\#\{\gamma \in \Gamma_R \cap B_N : \text{tr}^2(\pi(\gamma)) - 4 \text{ is square-free}\} > N^{2\delta_R-\eta}.$$

We now select discriminants which contribute the most to the count above. Define

$$\mathcal{T} := \{t \in \mathbb{Z}[i] : t^2 - 4 \text{ is square-free}\}$$

and

$$\mathcal{M}_R(t) := \#\{\gamma \in \Gamma_R : \text{tr}(\pi(\gamma)) = t\}.$$

**Proposition 4.** *For any  $\eta$  there is a large enough  $R$  so that*

$$\sum_{\substack{t \in \mathcal{T} \\ N(t) \leq N}} \mathbf{1}_{\{\mathcal{M}_R(t) \geq N(t)^{2\delta_R - 2 - 2\eta}\}} > N^{2\delta_R - 2 - \eta}.$$

*Proof.* The previous theorem gives

$$N^{2\delta_R - \eta} < \sum_{\substack{t \in \mathcal{T} \\ N(t) \leq N}} \mathcal{M}_{R,N}(t) = \sum_{\substack{t \in \mathcal{T} \\ N(t) \leq N}} \mathcal{M}_{R,N}(t) \left( \mathbf{1}_{\{\mathcal{M}_{R,N}(t) \geq W\}} + \mathbf{1}_{\{\mathcal{M}_{R,N}(t) < W\}} \right)$$

where  $W$  is a parameter of our choice. Trivially, we have  $\mathcal{M}_{R,N}(t) \ll N^{1+\epsilon}$ . So,

$$N^{2\delta_R - \eta} \ll N^{2+\epsilon} \sum_{\substack{t \in \mathcal{T} \\ N(t) \leq N}} \mathbf{1}_{\{\mathcal{M}_{R,N}(t) \geq W\}} + N^2 W.$$

Now, set  $W = N^{2\delta_R - 2 - 2\eta}$  and the claim follows.  $\square$

For any  $\epsilon > 0$ , we can find  $\eta$  small and  $R$  large so that

$$2\delta_R - 2 - \eta > 2 - \epsilon.$$

The choice of  $R$  gives our compact region (geodesics do not visit the cusp when their symbolic encodings have small entries.) We define

$$\mathcal{D} := \{D = t^2 - 4 : t \in \mathcal{T}, \mathcal{M}_R(t) > N(t)^{2\delta_R - 2 - 2\eta}\}$$

which gives a subset of all fundamental discriminants. The previous claim gives us a lower bound for the number of these discriminants:

$$\begin{aligned} \#\{d \in \mathcal{D} : N(d) \leq T\} &\geq \#\{t \in \mathcal{T} : N(t) \leq T^{1/2}, \mathcal{M}_R(t) > N(t)^{2\delta_R - 2 - 2\eta}\} \\ &> (T^{1/2})^{2\delta_R - 2 - \eta} \\ &> T^{1-\epsilon}. \end{aligned}$$



Moreover, for each discriminant  $d \in \mathcal{D}$ , if  $d = t^2 - 4$  then

$$\mathcal{M}_R(t) > N(t)^{2-\epsilon} > \sqrt{N(D)}^{2-\epsilon} \gg |C_D|^{1-\epsilon'}$$

Thus, after renaming constants we have proved Theorem 1, which we restate here:

**Theorem 25.** *For any  $\epsilon > 0$ , there is a compact region  $Y(\epsilon) \subset \Gamma \backslash \mathbb{H}^3$  and a set  $\mathcal{D}(\epsilon)$  of fundamental discriminants such that*

$$\#\{D \in \mathcal{D}(\epsilon) : N(D) < X\} \gg_{\epsilon} X^{1-\epsilon}, \quad X \rightarrow \infty$$

and for all  $D \in \mathcal{D}(\epsilon)$ ,

$$\#\{\gamma \in C_D : \gamma \subset Y(\epsilon)\} > |C_D|^{1-\epsilon}.$$

## References

- [BGS10] Jean Bourgain, Alex Gamburd, and Peter Sarnak. Affine linear sieve, expanders, and sum-product. *Inventiones mathematicae*, 179(3):559–644, 2010.
- [BGS11] Jean Bourgain, Alex Gamburd, and Peter Sarnak. Generalization of Selbergs 3/16-theorem and affine sieve. *Acta mathematica*, 207(2):255–290, 2011.
- [BK14] Jean Bourgain and Alex Kontorovich. The affine sieve beyond expansion I: Thin hypotenuses. *International Mathematics Research Notices*, 2015(19):9175–9205, 2014.
- [BK17] Jean Bourgain and Alex Kontorovich. Beyond expansion II: low-lying fundamental geodesics. *Journal of the European Mathematical Society*, 19(5):1331–1359, 2017.
- [BKM] J Bourgain, A Kontorovich, and M Magee. Thermodynamic expansion to arbitrary moduli, 2015. *arXiv preprint arXiv:1507.07993*, 9.
- [CU04] Laurent Clozel and Emmanuel Ullmo. Equidistribution des points de Hecke, contributions to automorphic forms, geometry, and number theory, 193254, 2004.
- [Duk88] William Duke. Hyperbolic distribution problems and half-integral weight Maass forms. *Inventiones mathematicae*, 92(1):73–90, 1988.
- [EGM13] Jürgen Elstrodt, Fritz Grunewald, and Jens Mennicke. *Groups acting on hyperbolic space: Harmonic analysis and number theory*. Springer Science & Business Media, 2013.
- [ELM<sup>+</sup>09] Manfred Einsiedler, Elon Lindenstrauss, Philippe Michel, Akshay Venkatesh, et al. Distribution of periodic torus orbits on homogeneous spaces. *Duke Mathematical Journal*, 148(1):119–174, 2009.
- [IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53. American Mathematical Soc., 2004.
- [KO11] Alex Kontorovich and Hee Oh. Apollonian circle packings and closed horospheres on hyperbolic 3-manifolds. *Journal of the American Mathematical Society*, 24(3):603–648, 2011.
- [Kon14] Alex Kontorovich. The orbital circle method. In *Thin Groups and Superstrong Approximation*, volume 61, pages 93–106. Cambridge University Press, 2014.

- [Lal89] Steven P Lalley. Renewal theorems in symbolic dynamics, with applications to geodesic flows, noneuclidean tessellations and their fractal limits. *Acta mathematica*, 163(1):1–55, 1989.
- [LN97] Rudolf Lidl and Harald Niederreiter. *Finite fields*, volume 20. Cambridge university press, 1997.
- [LV98] Carlangelo Liverani and Tor Vergata. Stability of the spectrum for transfer operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.(4) vol.*, 28:141–152, 1998.
- [Mer74] Franz Mertens. Ein beitrage zur analytischen zahlentheorie. *Journal für die reine und angewandte Mathematik*, 78:46–62, 1874.
- [MOW17] Michael Magee, Hee Oh, and Dale Winter. Uniform congruence counting for schottky semigroups in  $sl_2(\mathbb{Z})$ . *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2017.
- [MU01] R Daniel Mauldin and Mariusz Urbański. Gibbs states on the symbolic space over an infinite alphabet. *Israel Journal of Mathematics*, 125(1):93–130, 2001.
- [MU03] R Daniel Mauldin and Mariusz Urbanski. *Graph directed Markov systems: geometry and dynamics of limit sets*, volume 148. Cambridge University Press, 2003.
- [Pol94] Mark Pollicott. The Picard group, closed geodesics and zeta functions. *Transactions of the American Mathematical Society*, 344(2):857–872, 1994.
- [PP90] William Parry and Mark Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics*. Société mathématique de France, 1990.
- [PU18] Mark Pollicott and Mariusz Urbański. Open conformal systems and perturbations of transfer operators. *Lecture Notes in Mathematics*, 2206, 2018.
- [Rud82] Daniel J Rudolph. Ergodic behaviour of Sullivan’s geometric measure on a geometrically finite hyperbolic manifold. *Ergodic Theory and Dynamical Systems*, 2(3-4):491–512, 1982.
- [Sar83] P. Sarnak. The arithmetic and geometry of some hyperbolic three manifolds. *Acta Math.*, 151:253–295, 1983.
- [Sel56] Atle Selberg. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *Journal of the Indian Mathematical Society*, 20(1-3):47–87, 1956.
- [Ser81] Caroline Series. The infinite word problem and limit sets in Fuchsian groups. *Ergodic theory and dynamical systems*, 1(3):337–360, 1981.
- [SGS13] Alireza Salehi Golsefidy and Peter Sarnak. The affine sieve. *Journal of the American Mathematical Society*, 26(4):1085–1105, 2013.
- [Sul84] Dennis Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Mathematica*, 153(1):259–277, 1984.