Groups Geom. Dyn. 16 (2022), 683–703
DOI 10.4171/GGD/667

Combinatorial growth in the modular group

Ara Basmajian and Robert Suzzi Valli

Abstract. We consider an exhaustion of the modular orbifold by compact subsurfaces and show that the growth rate, in terms of word length, of the reciprocal geodesics on such subsurfaces (so named low lying reciprocal geodesics) converges to the growth rate of the full set of reciprocal geodesics on the modular orbifold. We derive a similar result for the low lying geodesics and their growth rate convergence to the growth rate of the full set of closed geodesics.

1. Introduction

Consider the modular surface; that is, the $(2,3,\infty)$ triangle orbifold, $S = \mathbb{H}/\operatorname{PSL}(2,\mathbb{Z})$. A *reciprocal geodesic* on the modular surface is a closed geodesic that begins and ends at the order-two cone point, traversing its image twice. Its lift is the conjugacy class in $\operatorname{PSL}(2,\mathbb{Z})$ of a hyperbolic element with axis passing through an order-two fixed point. Specifying the geometric length of a geodesic is equivalent to specifying the absolute trace of such a hyperbolic element by way of the formula, $T_\gamma = 2\cosh\frac{\ell(\gamma)}{2}$. Sarnak [27] showed

$$|\{\gamma \text{ a primitive reciprocal geodesic with } T_\gamma \le T\}| \sim \frac{3}{8}T.$$

Let $\mathcal{C} \subset S$ be the cusp with its natural horocycle boundary of length one. The *depth* of a point in $\mathcal{C}$ is its distance to the natural horocycle of length one. For a positive integer m, we define the *m-thick part* of S, denoted by S_m, to be S with the points a depth larger than $\log\frac{m+1}{2}$ deleted. Thus the m-thick parts form a compact exhaustion of S. We are interested in the reciprocal geodesics that lie in the m-thick part (so called m-low lying reciprocal geodesics). See Figure 1. Bourgain and Kontorovich [3] showed (in our terminology) that for any $\varepsilon > 0$, there is an $m > 0$ so that the number of fundamental reciprocal geodesics in the m-thick part having absolute trace $\le T$ has growth rate at least $T^{1-\varepsilon}$ (fundamental geodesics correspond to certain classes of binary quadratic forms, see [2, 3, 18] for the definition). In particular, as ε goes to zero, m goes to infinity and we have a nested, increasing set of compact subsets that converge (in an appropriate sense) to the modular orbifold S. Combined with the Sarnak result this shows that the growth rates of the low lying reciprocal geodesics converge (for each T, as $m \to \infty$) to the growth rate of the reciprocal geodesics. Using the fact that $\mathbb{Z}_2 * \mathbb{Z}_3$ is isomorphic to the modular group we

2020 Mathematics Subject Classification. Primary 20F69; Secondary 32G15, 57K20, 20H10, 53C22.
Keywords. Asymptotic growth, binary words, closed geodesics, modular orbifold.

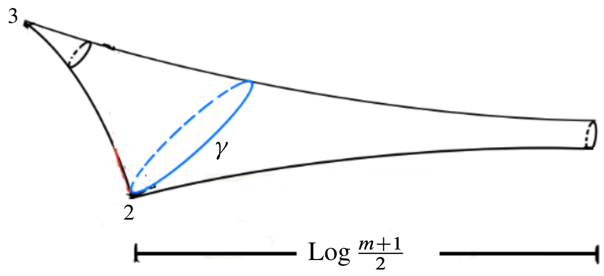

Figure 1. A reciprocal geodesic γ on S_m

give a combinatorial analogue of the above results using word length, instead of absolute trace, with respect to the generators of the factors in $\mathbb{Z}_2 * \mathbb{Z}_3$. For γ a closed geodesic on S we define its word length, denoted by $|\gamma|$, to be the minimal word length in the conjugacy class of a lift in $\mathrm{PSL}(2,\mathbb{Z})$, which is necessarily even.

Our main results are given in the following theorem.

Theorem 1.1. *The following hold:*

(1) $|\{\gamma \text{ a primitive reciprocal geodesic with } |\gamma| \leq 2t\}| \sim 2^{\lfloor \frac{t}{2} \rfloor}$.

(2) $|\{\gamma \text{ a primitive reciprocal geodesic in } S_{m\geq 2} \text{ with } |\gamma| \leq 2t\}|$
$\sim \left(\frac{\alpha_m}{2+(m+1)(\alpha_m-2)}\right)\alpha_m^{\lfloor \frac{t}{2} \rfloor}$.

(3) $|\{\gamma \text{ a primitive closed geodesic with } |\gamma| \leq 2t\}| \sim \frac{2^{t+1}}{t}$.

(4) $|\{\gamma \text{ a primitive closed geodesic in } S_{m\geq 3} \text{ with } |\gamma| \leq 2t\}| \gtrsim \frac{2^{t(1-1/m)}}{t}$.

We use the notation $f \sim g$ to denote that the ratio of $f(t)$ and $g(t)$ approaches one, as t goes to infinity. We use the symbol $f \gtrsim g$ to mean that there exist a constant C and a t_0 so that $f(t) \geq Cg(t)$ for all $t \geq t_0$.

The constant α_m in item (2) of Theorem 1.1 is the unique positive root of the polynomial

$$z^m - z^{m-1} - z^{m-2} - \cdots - z - 1.$$

The $\{\alpha_m\}_{m=2}^{\infty}$ are increasing in m, satisfy $2(1-\frac{1}{2^m}) \leq \alpha_m \leq 2$, and go to 2 as $m \to \infty$. See [6] for the details and the proofs of these algebraic properties. Using these properties on the functions in Theorem 1.1, we get the following result.

Corollary 1.2. *The asymptotic growth rate of the primitive reciprocal geodesics in the m-thick part, S_m, converges to the asymptotic growth rate of the primitive reciprocal geodesics on the modular orbifold, as $m \to \infty$. Similarly, the asymptotic growth rate of the primitive closed geodesics in S_m converges, up to a multiplicative constant, to the asymptotic growth rate of the primitive closed geodesics on the modular orbifold, as $m \to \infty$.*

Geodesic set	**Bijection to**	**Cardinality**
Geodesics of length $2t$	Lyndon binary words (primitive and non-primitive) of length t	$\frac{1}{t}\sum_{j=1}^{t} 2^{\gcd(j,t)} - 2$
Reciprocal geodesics of length $4t$	Compositions of t	2^{t-1}
Geodesics in S_m of length $2t$	m-Lyndon binary words of length t	$\geq \frac{2^{t-\frac{t}{m}-1}}{t}$
Reciprocal geodesics in S_m of length $4t$	Compositions of t with parts bounded by m	$\left\lfloor \frac{1}{2} + \frac{\alpha_m - 1}{2+(m+1)(\alpha_m-2)}\alpha_m^t \right\rfloor$

Table 1. Cardinality of geodesic classes

Remark 1.3. We derive the exact size of the full set of the low lying reciprocal geodesics, reciprocal geodesics, and closed geodesics of word length exactly $2t$ (see Table 1), allowing us to achieve tight coarse bounds, and hence in the limit the asymptotic growth rates for word length $\leq 2t$; thus proving the results of Theorem 1.1. As these are asymptotic growth rates, rough approximations such as quasi-isometries between growth rates involving geometric length and combinatorial length of geodesics are not applicable. In particular, our results do not follow in any obvious way from the results of Sarnak [27].

The study of asymptotic growth rates of geometric lengths of various classes of closed geodesics has a long and storied history beginning with Huber's result for all closed geodesics, to Mirzakhani's growth rate of the simple closed geodesics, to more general results for non-simple closed geodesics and reciprocal geodesics [1–4, 8, 9, 21, 23, 27]. Concurrently there is the study of such geodesics in terms of word length or equivalently primitive conjugacy classes and their word length growth rates leading to more abstract, algebraic investigations of groups such as surface groups or free groups [5, 7, 16, 22, 25, 26, 29]. Papers involving normal forms, enumeration schemes for curves, Farey arithmetic, and generating elements in a non-abelian free group include [10–15, 28].

Our focus in this paper is on so called low lying and reciprocal words in the modular group. These are the lifts of the low lying and reciprocal geodesics, respectively. Although neither of these sets forms a group, they have the minimal properties needed to consider the growth rate of their primitive conjugacy classes. Namely, these subsets of the modular group are comprised of infinite order elements, are conjugacy invariant, are closed under taking powers, and the unique positive power primitive element in the modular group is also a member of the subset.

We take, for the most part, a combinatorial approach to determine the growth rate of the primitive conjugacy classes. Typically in such arguments a convenient normal form for the conjugacy classes is used and then counted. Of course, one needs to determine when two elements in normal form represent the same conjugacy class. In the case of reciprocal words we prove a crucial lemma (Lemma 3.6) showing exactly two elements in normal form are conjugate and identifying these two elements. The fact that there are two primitive conjugates was first proven by Sarnak [27] using different methods.

This normal form using the isomorphism from $\mathbb{Z}_2 * \mathbb{Z}_3$ to $\mathrm{PSL}(2,\mathbb{Z})$ allows us to represent a closed geodesic as a product of parabolic elements. How deep a geodesic wanders into the cusp is directly related to the exponents of these parabolics. See Lemma 7.1 for a precise statement. For the connection between geodesic excursions into the cusp and number theoretic quantities see [17] and the references therein.

The paper is organized as follows. In Section 2, we derive some of the elementary but key lemmas as well as set up notation. In Section 3, we talk about the normal form of a reciprocal word and prove a crucial lemma which identifies when two such words in normal form are conjugate. In Sections 4 and 5, we determine the size of the conjugacy classes of elements in $\mathbb{Z}_2 * \mathbb{Z}_3$ of length $2t$ and reciprocal words of length $4t$ as well as the primitive classes of these sets. In Section 6, we identify the low lying conjugacy classes of length $2t$ with so called m-Lyndon words of length t, and derive an effective lower bound for the growth of such words. We next construct a bijection between the conjugacy classes of m-low lying reciprocal words and compositions with parts bounded by m allowing us to count these classes. Section 7 relates the cusp geometry of the modular orbifold with geodesic excursions into the cusp. Finally, in Section 8, we put the work of the previous sections together to prove our main theorem.

2. Basics and notation

We use the notation $f \sim g$ to mean asymptotic to and the symbol $f \gtrsim g$ to mean that there exist a constant C and a t_0 so that $f(t) \geq Cg(t)$ for all $t \geq t_0$.

Consider the group $G = \mathbb{Z}_2 * \mathbb{Z}_3$. Assume the generator of $\mathbb{Z}_2$ is a and the generator of $\mathbb{Z}_3$ is b. An element $g \in G$ is *primitive* if it is not a non-trivial power of another element of G. The *word length* of g, denoted by $\|g\|$, is the minimum length among all words representing g using the symmetric set of generators $\{a, b, b^{-1}\}$. Set $\mathcal{W} = \{$reduced words in the generators of $G\}$, that is, where the exponent of a is always $+1$ and the exponent of b is always ± 1. The conjugacy class of $g \in G$ is denoted by $[g]$. For a positive integer s, since conjugation commutes with taking powers, we may define $[g]^s := [g^s]$. The *length of a conjugacy class* $[g]$ is given by $\|[g]\| = \min\{\|h\| : h \in [g]\}$. A word in $\mathcal{W}$ is *cyclically reduced* if any cyclic permutation of it is a reduced word. Though cyclically reduced words in a conjugacy class are not unique they do realize the minimum length in the conjugacy class. In fact, all conjugates of a cyclically reduced word are cyclic permutations of each other. For the basics on combinatorial group theory see [19, 20].

We call a reduced word that begins with a and ends with b or b^{-1} an (ab)-word. Similarly, we have (aa)-, (bb)-, (ba)-words. We remark the obvious but important fact that an (ab)- or (ba)-word is cyclically reduced but an (aa)- or (bb)-word is not.

We have the following fundamental lemma.

Lemma 2.1. *Let $x \in \mathcal{W}$ where x is not conjugate to one of the generators, that is, not conjugate to a, b, or b^{-1}. Then:*

(1) *x is conjugate to an (ab)-word y with $\|x\| \geq \|y\|$.*

(2) *The only conjugates of the word $ab^{\varepsilon_1} \dots ab^{\varepsilon_t}$, $\varepsilon_i = \pm 1$, that are (ab)-words are its even cyclic permutations. That is, $ab^{\varepsilon_t} ab^{\varepsilon_1} \dots ab^{\varepsilon_{t-1}}$ and so on.*

(3) *If y is an (ab)-word and $x^s = y$ for a positive integer s, then x is an (ab)-word and $s\|x\| = \|y\|$.*

(4) *If $[x]^s = [y]$, then $s\|[x]\| = \|[y]\|$.*

Proof. Items (1)–(3) follow immediately. To prove item (4) we may assume, by conjugating if necessary, that y is an (ab)-word. Now, by assumption there exists an x so that $x^s = y$, and hence x is an (ab)-word by item (3). Moreover, we have $\|[y]\| = \|y\| = s\|x\| = s\|[x]\|$, where the second and third equalities also follow from item (3). ∎

Remark 2.2. In the group $G = \mathbb{Z}_2 * \mathbb{Z}_3$ each infinite order element is a positive power of a unique, primitive element. Although this can be proven using purely combinatorial methods, the easiest way to see this is by the fact that G has a discrete, faithful representation into $\mathrm{PSL}(2, \mathbb{Z})$. In particular, every infinite order element is contained in a maximal cyclic subgroup of G, and hence, there is a corresponding maximal primitive root.

Throughout this work we consider subsets of G that have the following minimal properties.

Definition 2.3. A set $\mathcal{A} \subset G$ made up of infinite order elements is said to satisfy condition $(*)$ if the following properties are satisfied.

- All positive and negative powers of any element in $\mathcal{A}$ are also in $\mathcal{A}$.
- For any element in $\mathcal{A}$, the corresponding unique primitive root in G of which it is a positive power is also in $\mathcal{A}$.
- $\mathcal{A}$ is conjugacy invariant.

In the sequel the subsets $\mathcal{A}$ will denote either infinite order words in G, reciprocal words (to be defined later), or low lying words (to be defined later). For now we proceed abstractly with any set $\mathcal{A}$ satisfying condition $(*)$, we fix notation, and derive some basic facts.

As in Section 2, let $\mathcal{W}$ be the set of reduced words in the generators of G. Setting $\mathcal{A}^p = \{\text{primitive elements of } \mathcal{A}\}$, we have $\mathcal{A}^p \subseteq \mathcal{A} \subseteq \mathcal{W}$. Since each of these subsets is closed under conjugation by elements of G, we define the conjugacy classes of these subsets by capitals: A^p, A, and W, respectively. Note that W is the full set of conjugacy classes in G. We denote the *non-primitive conjugacy classes* in A by A^{np}.

For various choices of $\mathcal{A}$ we are interested in the growth rate of primitive conjugacy classes in $\mathcal{A}$. Here growth is measured by word length in terms of the generators.

For a positive integer t we use t as a subscript to denote the elements in that set of length t. Similarly, we use $\leq t$ as a subscript to denote the elements in the set of length $\leq t$. For example, A_t denotes the conjugacy classes in A of length t, and $A_{\leq t}$ denotes the conjugacy classes in A of length less than or equal to t. The growth function for the set A is denoted by $|A_{\leq t}|$. A *proper divisor* of t is a positive integer that divides t but is not 1 or t.

We next define a map from primitive conjugacy classes to non-primitive conjugacy classes given by a power map.

Lemma 2.4. *The map* $\iota : \dot{\bigcup}_{s|t} A_s^p \to A_t^{np}$ *given by* $[x] \mapsto [x^{t/s}]$ *is well-defined and a bijection. That is, the non-primitive conjugacy classes in* A_t *are in one-to-one correspondence with elements of* $\dot{\bigcup}_{s|t} A_s^p$, *where the union is over all proper divisors, s, of t.*

Proof. ι is well-defined since powers commute with conjugation. To prove surjectivity, suppose $[y] \in A_t^{np}$ and hence there exists $[x] \in A_s^p$ so that $[x]^n = [y]$, where n is a positive integer greater than 1. By item (4) of Lemma 2.1, $n\|[x]\| = \|[y]\|$ and hence s divides t. If $s = 1$, then $n = t$ and $\|[x]\| = 1$. So x is conjugate to a or $b^{\pm 1}$, which implies x has finite order. Thus s properly divides t.

Injectivity follows from establishing the following two items, which we leave to the reader.

(1) If $[x_1] \neq [x_2]$ in A_s^p, then $\iota([x_1])$ is not conjugate to $\iota([x_2])$.

(2) If s_1 and s_2 divide t, $s_1 \neq s_2$, then $\iota(A_{s_1}^p) \cap \iota(A_{s_2}^p) = \emptyset$. ∎

We have established the following result.

Proposition 2.5. *Suppose* $\mathcal{A} \subset G$ *satisfies condition* $(*)$. *Then*

$$A_t^p = A_t - \dot{\bigcup_{s|t}} \iota(A_s^p) \quad \text{and} \quad |A_t^p| = |A_t| - \sum_{s|t} |A_s^p|$$

where the union and sum are over all proper divisors, s, of t.

Our goal in the next few sections is to compute the asymptotics as $t \to \infty$ of the functions $|A_{2t}^p|$ and $|A_{\leq 2t}^p|$ for various choices of $\mathcal{A}$.

Set

$$\mathcal{R} = \{xy : x, y \text{ are distinct order-two elements in } G\}.$$

Denoting the commutator of x and y by $[x, y]$, and noting that there is one conjugacy class of order-two elements, we may write

$$\begin{aligned} \mathcal{R} &= \left\{[xax^{-1}, xyx^{-1}] : x, y \in G, y \neq a\right\} \\ &= \left\{x[a, y]x^{-1} : x, y \in G, y \neq a\right\}. \end{aligned}$$

We call the elements of $\mathcal{R}$ *reciprocal words*. We remark that $\mathcal{R}$ is closed under taking powers. Moreover, if an element of $\mathcal{R}$ is a power of the unique, primitive $y \in G$, then y

is also in $\mathcal{R}$. Thus $\mathcal{R}$ satisfies condition $(*)$. Denote the primitive conjugacy classes in R by R^p. The conjugacy classes of $\mathcal{R}$ correspond exactly to the set of reciprocal geodesics on the modular orbifold. This is because a lift of a reciprocal geodesic is the product of a pair of order-two elements in $\mathrm{PSL}(2,\mathbb{Z})$.

Remark 2.6. Any reciprocal word conjugated to an (ab)-word has the form

$$w = [a,\gamma] = ab^{\varepsilon_1}\dots ab^{\varepsilon_t}ab^{-\varepsilon_t}\dots ab^{-\varepsilon_1},$$

where $\varepsilon_i = \pm 1$ and γ is a (bb)-word. With this in mind we define the normal form of a reciprocal word to be $[a,\gamma]$ where γ is a (bb)-word. The full set of normal forms is denoted by

$$\mathcal{N} = \big\{[a,\gamma] : \gamma \text{ a } (bb)\text{-word}\big\}.$$

Note that non-primitive elements of $\mathcal{N}$ are powers of elements of the same form. That is, $[a,\gamma] = [a,\beta]^n$, where $[a,\beta]$ is primitive.

3. Binary words and the normal form for a reciprocal word

Our interest is in counting words in G. To make our computations less cumbersome we identify (ab)-words in G with binary words. Namely, we identify the (ab)-word $ab^{\varepsilon_0}ab^{\varepsilon_1}\dots ab^{\varepsilon_{t-1}}$ with the binary word $(\varepsilon_0,\varepsilon_1,\dots,\varepsilon_{t-1})$, where $\varepsilon_i = \pm 1$. Denote the set of all binary words of length t by X_t. We define a cyclic permutation map

$$\alpha : X_t \to X_t, \quad (\varepsilon_0,\dots,\varepsilon_{t-1}) \mapsto (\varepsilon_{t-1},\varepsilon_0,\dots,\varepsilon_{t-2}).$$

Focusing on reciprocal words, the length of a reciprocal word in normal form, $\mathcal{N}$, is a multiple of 4. Hence we identify $\mathcal{N}$ with the subset

$$Y_{2t} = \big\{(\varepsilon_0,\dots,\varepsilon_{2t-1}) : \varepsilon_j = -\varepsilon_{2t-j-1} \text{ for all } j = 0,\dots,2t-1\big\} \subset X_{2t}.$$

Conjugate words of an (ab)-word in the same (ab)-form are cyclic permutations of even order. With this in mind, using the bijection, we use the cyclic action $\langle\alpha\rangle$ on X_{2t} given by $\alpha^k(\varepsilon_0,\dots,\varepsilon_{2t-1}) = (\varepsilon'_0,\dots,\varepsilon'_{2t-1})$, where $\varepsilon'_j = \varepsilon_{j-k}$ for all $j = 0,\dots,2t-1$. Here we use the convention that the subscripts are modulo $2t$.

Lemma 3.1. *Fix k. For any $(\varepsilon_0,\dots,\varepsilon_{2t-1}) \in X_{2t}$ we have:*

- $\alpha^k(\varepsilon_0,\dots,\varepsilon_{2t-1}) = (\varepsilon_0,\dots,\varepsilon_{2t-1})$ *if and only if* $\varepsilon_j = \varepsilon_{j-k}$ *for all* j.
- $\alpha^k(\varepsilon_0,\dots,\varepsilon_{2t-1}) \in Y_{2t}$ *if and only if* $\varepsilon_{j-k} = -\varepsilon_{2t-j-1-k}$ *for all* j.
- *If* $(\varepsilon_0,\dots,\varepsilon_{2t-1}) \in Y_{2t}$, *then* $\alpha^t(\varepsilon_0,\dots,\varepsilon_{2t-1}) \in Y_{2t}$.
  In general, $\alpha^k(\varepsilon_0,\dots,\varepsilon_{2t-1}) \in Y_{2t}$ *if and only if* $\varepsilon_j = \varepsilon_{j-2k}$ *for all* j.

The proof of Lemma 3.1 is straightforward, we leave it to the reader.

With an eye toward applications to elements in the group G we define a notion of primitivity for a binary word.

Definition 3.2. For any positive integer t, an element $x \in X_t$ is called non-primitive if there exists s which properly divides t so that x is the juxtaposition of s-subtuples where each subtuple is the same. That is, there exists $z \in X_{t/s}$ so that $x = z^s$. Otherwise, we say x is primitive.

An element in X_t is the positive integer power of some primitive subtuple. Moreover, if $x \in Y_{2t} \subset X_{2t}$ is non-primitive then $x = y^s$, where $y \in Y_{2t/s}$.

Let $\mathcal{O} = \mathcal{O}(x)$ be the orbit in X_{2t} of an element $x \in Y_{2t} \subset X_{2t}$ under the action of α. We are interested in $\mathcal{O} \cap Y_{2t}$. We set k_0 to be the *smallest power of* α for which an element of $\mathcal{O} \cap Y_{2t}$ maps back to $\mathcal{O} \cap Y_{2t}$. Note that k_0 is an invariant of the orbit and $1 \le k_0 \le t$.

Lemma 3.3. *The following hold:*

(1) *For any integer n, $\alpha^{nk_0}(\mathcal{O} \cap Y_{2t}) = \mathcal{O} \cap Y_{2t}$.*

(2) *$\alpha^l(\mathcal{O} \cap Y_{2t}) \cap (\mathcal{O} \cap Y_{2t}) = \emptyset$ for l not a multiple of k_0.*

(3) *$\alpha^{k_0}(x) \neq x$ for any $x \in \mathcal{O} \cap Y_{2t}$.*

Proof. Set $x = (\varepsilon_0, \ldots, \varepsilon_{2t-1}) \in \mathcal{O} \cap Y_{2t}$ throughout this proof.

To prove item (1), it is enough to show that $\alpha^{nk_0}(\mathcal{O} \cap Y_{2t}) \subset \mathcal{O} \cap Y_{2t}$ for any integer n. Noting that $\alpha^{k_0}(x) \in \mathcal{O} \cap Y_{2t}$, we have, by Lemma 3.1, $\varepsilon_j = \varepsilon_{j-2k_0} = \cdots = \varepsilon_{j-2nk_0}$ for all j. Therefore, again by Lemma 3.1, $\alpha^{nk_0}(x) \in \mathcal{O} \cap Y_{2t}$.

For item (2), suppose for contradiction $\alpha^l(x) \in \mathcal{O} \cap Y_{2t}$ and $nk_0 < l < (n+1)k_0$. Then $\alpha^{l-nk_0}(\alpha^{nk_0}(x)) = \alpha^l(x) \in \mathcal{O} \cap Y_{2t}$. On the other hand by item (1), $\alpha^{nk_0}(x) \in \mathcal{O} \cap Y_{2t}$ and so $\alpha^{l-nk_0}(\mathcal{O} \cap Y_{2t}) \cap (\mathcal{O} \cap Y_{2t}) \neq \emptyset$, contradicting our assumption that k_0 is minimal.

To prove item (3), assume $\alpha^{k_0}(x) = x$ and hence, by Lemma 3.1, $\varepsilon_j = \varepsilon_{j-k_0}$ for all j, and there is a repeating subtuple y of length k_0 which fills out x. Moreover, this subtuple must lie in Y_{k_0} and since the elements of Y_{k_0} have even length, k_0 must be even. On the other hand, setting $k = \frac{k_0}{2}$ we have $\varepsilon_{j-2k} = \varepsilon_{j-k_0} = \varepsilon_j$ for all j, where the last equality follows from Lemma 3.1. This contradicts the minimality of k_0 and thus $\alpha^{k_0}(x) \neq x$. ∎

Proposition 3.4. *Assume $x \in Y_{2t}$. Then the following hold:*

(1) *$\alpha^{nk_0}(x) = x$ for n even, and $\alpha^{nk_0}(x) = \alpha^{k_0}(x)$ for n odd. Namely, $\mathcal{O} \cap Y_{2t}$ consists of two distinct elements, $\{x, \alpha^{k_0}(x)\}$.*

(2) *The element x is primitive if and only if $k_0 = t$.*

Proof. Set $x = (\varepsilon_0, \ldots, \varepsilon_{2t-1}) \in Y_{2t}$. To prove (1), we note by items (1) and (2) of Lemma 3.3, that all the points in $\mathcal{O} \cap Y_{2t}$ are of the form $\alpha^{nk_0}(x)$. Since $\alpha^{k_0}(x) \in \mathcal{O} \cap Y_{2t}$ and assuming n is even, we have $\varepsilon_j = \varepsilon_{j-2k_0} = \varepsilon_{j-4k_0} = \cdots = \varepsilon_{j-nk_0}$ by Lemma 3.1, and hence $\alpha^{nk_0}(x) = x$. For odd $n = 2m+1$, we have that

$$\alpha^{nk_0}(x) = \alpha^{(2m+1)k_0}(x) = \alpha^{k_0}(\alpha^{2mk_0}(x)) = \alpha^{k_0}(x).$$

However, $\alpha^{k_0}(x) \neq x$ by item (3) of Lemma 3.3.

To prove (2), we first remark that x is non-primitive if and only if there exists a minimal subtuple $y \in Y_{2s}$ repeated $\frac{t}{s}$-times, where s properly divides t, giving x. We have $k_0 < 2s \leq t$, where the left inequality follows from item (3) of Lemma 3.3. Thus $k_0 < t$. On the other hand, if x is primitive then there is no proper s and thus $k_0 = t$. ■

Remark 3.5. A schematic picture emerges. We picture Y_{2t} as the diagonal in X_{2t} and the $\langle\alpha\rangle$-orbit of a point in Y_{2t} as intersecting Y_{2t} in exactly two distinct points, and the number of all orbit points in X_{2t} being $2k_0$.

We now return to the main objective of this section to consider the normal form of reciprocal words. We remind the reader of the bijection

$$\mathcal{N}_{4t} = \{[a, \gamma] : \gamma \text{ a } (bb)\text{-word of length } 2t - 1\} \to Y_{2t}$$

given by

$$ab^{\varepsilon_0} \dots ab^{\varepsilon_{t-1}} ab^{-\varepsilon_{t-1}} \dots ab^{-\varepsilon_0} \mapsto (\varepsilon_0, \dots, \varepsilon_{t-1}, -\varepsilon_{t-1}, \dots, -\varepsilon_0).$$

Using the bijection with Proposition 3.4, and noting that when $[a, \beta]$ is primitive, β cannot be of order two, we have proven the following result.

Lemma 3.6. *Each conjugacy class of an element of $\mathcal{R}$ has exactly two representatives in the normal form $\mathcal{N}$. Namely, the two conjugates in $\mathcal{N}$ are $[a, \beta]^n$ and $[a, \beta^{-1}]^n$, where $[a, \beta]$ is primitive, n is a unique positive integer, and β is a unique (bb)-word not of order two.*

Remark 3.7. In proving Lemma 3.6 we largely took a combinatorial point of view. However, using different methods in [27], this lemma is proven for primitive reciprocals by considering conjugacy classes of infinite maximal dihedral subgroups of $\mathrm{PSL}(2, \mathbb{Z})$.

4. Counting conjugacy classes in $\mathbb{Z}_2 * \mathbb{Z}_3$

The goal of this section is to investigate the growth rate of primitive conjugacy classes in G. Fix a positive integer t, recall that $X_t = \{(\varepsilon_0, \dots, \varepsilon_{t-1}) : \varepsilon_i = \pm 1\}$, and consider the cyclic permutation map, $\alpha : X_t \to X_t$, given by $(\varepsilon_0, \dots, \varepsilon_{t-1}) \mapsto (\varepsilon_{t-1}, \varepsilon_0, \dots, \varepsilon_{t-2})$. Appealing to the Burnside lemma we have

$$|X_t / \langle\alpha\rangle| = \frac{1}{t} \sum_{j=1}^{t} 2^{\gcd(j,t)}. \tag{4.1}$$

Denote the set of all words in G of (ab)-form with length $2t$ by

$$\mathcal{W}_{2t}(ab) = \{ab^{\varepsilon_0} \dots ab^{\varepsilon_{t-1}} : \varepsilon_i = \pm 1\}.$$

Consider the action

$$\beta : \mathcal{W}_{2t}(ab) \to \mathcal{W}_{2t}(ab), \quad ab^{\varepsilon_0} \dots ab^{\varepsilon_{t-1}} \mapsto ab^{\varepsilon_{t-1}} ab^{\varepsilon_0} \dots ab^{\varepsilon_{t-2}}.$$

The group $\langle \beta \rangle$ is cyclic of order t.

Lemma 4.1. *We have*

$$|\mathcal{W}_{2t}(ab)/\langle \beta \rangle| = \frac{1}{t} \sum_{j=1}^{t} 2^{\gcd(j,t)}.$$

Proof. Note that $(\varepsilon_0, \dots, \varepsilon_{t-1}) \mapsto ab^{\varepsilon_0} \dots ab^{\varepsilon_{t-1}}$ is an equivariant bijection between the $\langle \alpha \rangle$ action on X_t and the $\langle \beta \rangle$ action on $\mathcal{W}_{2t}(ab)$. The result follows using (4.1). ■

Recall that W_{2t} is the full set of conjugacy classes in G of length $2t$.

Theorem 4.2. *The following hold:*

(1) $|W_{2t}| = \frac{1}{t} \sum_{j=1}^{t} 2^{\gcd(j,t)}$.

(2) $|W_{2t}| \sim \frac{2^t}{t}$, *as* $t \to \infty$.

(3) $|W_{\leq 2t}| \sim \frac{2^{t+1}}{t}$, *as* $t \to \infty$.

Proof. A proof of item (1) appears in [29], however for completeness we supply a proof. Suppose $[w] \in W_{2t}$. Then in the conjugacy class of w there is a representative in $\mathcal{W}_{2t}(ab)$. Now the only other conjugates in $\mathcal{W}_{2t}(ab)$ are the ones equivalent under the action of $\langle \beta \rangle$. There is a one-to-one correspondence between the set of conjugacy classes W_{2t} and $\mathcal{W}_{2t}(ab)/\langle \beta \rangle$, and hence by Lemma 4.1 the result follows.

To prove item (2) we begin by noting that $\gcd(t,t) = t$ and $\gcd(j,t) \leq \frac{t}{2}$ for $j < t$. It follows that

$$\frac{2^t}{t} \leq |W_{2t}| = \frac{1}{t} \left(\sum_{j=1}^{t-1} 2^{gcd(j,t)} + 2^t \right) \leq \left(\frac{t-1}{t} \right) 2^{t/2} + \frac{2^t}{t}.$$

The claimed asymptotic follows since

$$\frac{\left(\frac{t-1}{t}\right) 2^{t/2}}{\frac{2^t}{t}} \to 0.$$

For item (3) note that

$$|W_{\leq 2t}| = 3 + \sum_{n=1}^{t} |W_{2n}| = 3 + \sum_{n=1}^{t} \frac{1}{n} \sum_{j=1}^{n} 2^{\gcd(j,n)}.$$

We remark that the term 3 appears above since there are 3 length-one conjugacy classes. Using the same reasoning from item (2), we have

$$\sum_{n=1}^{t} \frac{2^n}{n} \leq |W_{\leq 2t}| \leq 3 + \sum_{n=1}^{t} \frac{2^n}{n} + \sum_{n=1}^{t} 2^{n/2}.$$

Since
$$\frac{\sum_{n=1}^{t} 2^{n/2}}{\sum_{n=1}^{t} \frac{2^n}{n}} \to 0,$$
we have that $|W_{\leq 2t}| \sim \sum_{n=1}^{t} \frac{2^n}{n}$. Finally, an application of the Stolz–Cesaro theorem [24] yields
$$\sum_{n=1}^{t} \frac{2^n}{n} \sim \frac{2^{t+1}}{t+1}. \qquad \blacksquare$$

Recall that W_{2t}^{np} is the full set of conjugacy classes of non-primitive elements in G of length $2t$.

Lemma 4.3. *The following hold:*

(1) $|W_{2t}^{np}| \leq \frac{1}{2}t2^{t/2}$.

(2) $|W_{\leq 2t}^{np}| \leq \frac{1}{2}t^2 2^{t/2}$.

Proof. For item (1), using Proposition 2.5, we have
$$\begin{aligned}|W_{2t}^{np}| &= \sum_{s|t} |W_{2s}^{p}| \leq \sum_{s|t} |W_{2s}| = \sum_{s|t} \frac{1}{s} \sum_{j=1}^{s} 2^{\gcd(j,s)} \\ &\leq \sum_{s|t} \frac{1}{s} \sum_{j=1}^{s} 2^s = \sum_{s|t} 2^s \leq \frac{t}{2} 2^{t/2},\end{aligned}$$
where the last inequality follows from the fact that the largest proper divisor of t is $\frac{t}{2}$ and there are at most $\frac{t}{2}$ divisors. For item (2) we apply item (1):
$$|W_{\leq 2t}^{np}| = \sum_{n=1}^{t} |W_{2n}^{np}| \leq \sum_{n=1}^{t} \frac{1}{2} n 2^{n/2} \leq \frac{1}{2} t 2^{t/2} \sum_{n=1}^{t} 1 = \frac{1}{2} t^2 2^{t/2}. \qquad \blacksquare$$

Recall that W_{2t}^{p} is the full set of conjugacy classes of primitive elements in G of length $2t$.

Theorem 4.4. *The following hold:*

(1) $|W_{2t}^{p}| \sim |W_{2t}| \sim \frac{2^t}{t}$, *as* $t \to \infty$.

(2) $|W_{\leq 2t}^{p}| \sim |W_{\leq 2t}| \sim \frac{2^{t+1}}{t}$, *as* $t \to \infty$.

Proof. Applying Lemma 4.3, we have
$$|W_{2t}| - \frac{1}{2} t 2^{t/2} \leq |W_{2t}^{p}| \leq |W_{2t}|.$$
Dividing by $|W_{2t}|$ and noting that $|W_{2t}| \geq \frac{2^t}{t}$, we get
$$1 - \frac{\frac{1}{2} t 2^{t/2}}{\frac{2^t}{t}} \leq \frac{|W_{2t}^{p}|}{|W_{2t}|} \leq 1,$$
which yields the desired asymptotic.

For part (2), we use an analogous argument. From Lemma 4.3 we deduce

$$|W_{\leq 2t}| - \frac{1}{2}t^2 2^{t/2} \leq |W^p_{\leq 2t}| \leq |W_{\leq 2t}|.$$

Dividing by $|W_{\leq 2t}|$ and using that $|W_{\leq 2t}| \geq \sum_{n=1}^{t} \frac{2^n}{n} \geq \frac{2^t}{t}$, we have

$$1 - \frac{\frac{1}{2}t^2 2^{t/2}}{\frac{2^t}{t}} \leq \frac{|W^p_{\leq 2t}|}{|W_{\leq 2t}|} \leq 1,$$

giving the desired asymptotic. ∎

5. Counting conjugacy classes of reciprocal words

In this section, we compute the growth rate of conjugacy classes of reciprocal words R in $G = \mathbb{Z}_2 * \mathbb{Z}_3$. The conjugacy classes of reciprocal words have length a multiple of 4. For this reason we use the parameter $4t$ for ease of computation.

Lemma 5.1. *The following hold:*

(1) $|R_{4t}| = 2^{t-1}$.

(2) $|R_{\leq 4t}| = 2^t - 1$.

Proof. Let $[w] \in R_{4t}$. We pick as representative a cyclically reduced word of the form

$$w = ab^{\varepsilon_0} \ldots ab^{\varepsilon_{t-1}} ab^{-\varepsilon_{t-1}} \ldots ab^{-\varepsilon_0},$$

where $\varepsilon_i = \pm 1$. The result follows since there are exactly two conjugates of this form (Lemma 3.6) and there are 2^t words of this form.

For the second claim,

$$|R_{\leq 4t}| = \sum_{n=1}^{t} |R_{4n}| = \frac{1}{2}\sum_{n=1}^{t} 2^n = \frac{1}{2}\left[2\left(\frac{1-2^t}{1-2}\right)\right] = 2^t - 1.$$ ∎

Lemma 5.2. *The following hold:*

(1) $|R^{np}_{4t}| \leq \frac{1}{4}t2^{t/2}$.

(2) $|R^{np}_{\leq 4t}| \leq \frac{1}{4}t^2 2^{t/2}$.

The proof of the lemma follows in an analogous way to the proof of Lemma 4.3. We leave the details to the reader.

Theorem 5.3. *The following hold:*

(1) $|R^p_{4t}| \sim |R_{4t}| = 2^{t-1}$, *as* $t \to \infty$.

(2) $|R^p_{\leq 4t}| \sim |R_{\leq 4t}| = 2^t - 1$, *as* $t \to \infty$.

Proof. For part (1), we apply Lemma 5.2 to get

$$|R_{4t}| - \frac{1}{4}t2^{t/2} \le |R^p_{4t}| \le |R_{4t}|.$$

Dividing by $|R_{4t}| = \frac{1}{2}2^t$ and letting $t \to \infty$ yields the claimed asymptotic.

For part (2), we again apply Lemma 5.2 to get

$$|R_{\le 4t}| - \frac{1}{4}t^2 2^{t/2} \le |R^p_{\le 4t}| \le |R_{\le 4t}|.$$

Dividing by $|R_{\le 4t}| = 2^t - 1$ and letting $t \to \infty$ yields the claimed asymptotic. ∎

6. Lying low

In this section, we would like to count the low lying geodesics. In other words, we consider the growth rate of conjugacy classes of low lying words as well as low lying reciprocal words. For a positive integer m we say that a word in $\mathcal{W}$ is an *m-low lying word* if, when conjugated to an (ab)-word $ab^{\varepsilon_1}ab^{\varepsilon_2}\dots ab^{\varepsilon_t}$ with $\varepsilon_i = \pm 1$, no $(m+1)$ consecutive ε_i considered cyclically have the same sign. Put another way, the highest exponent of ab or ab^{-1} considered cyclically in an m-low lying word is at most m. When the m is understood we simply say that the word is *low lying*. We denote the conjugacy classes of m-low lying words of word length $2t$ by $L_{2t,m}$, and the m-low lying primitive conjugacy classes of words by $L^p_{2t,m}$. We note that the property of being m-low lying is preserved under conjugation, taking powers, and taking roots; it hence satisfies condition $(*)$.

6.1. Low lying words

We fix a positive integer $m \ge 2$. We now consider the growth rate of the conjugacy classes of all m-low lying words. Let $\mathcal{L}_{2t,m}(ab)$ be the set of normalized m-low lying words of length $2t$. That is, a word of the form $w = ab^{\varepsilon_1}ab^{\varepsilon_2}\dots ab^{\varepsilon_t}$, where $\varepsilon_i = \pm 1$, and no $(m+1)$ consecutive ε_i considered cyclically have the same sign.

As before we identify such a word w with the t-tuple of ± 1's, $(\varepsilon_1, \dots, \varepsilon_t)$, and of course via this identification we have the notion of a primitive and non-primitive t-tuple. The cyclic action on the word w induces a cyclic action on this t-tuple. With this in mind, we consider the t-tuple of ± 1's on a circle oriented counterclockwise. Within the cyclic equivalence class we identify a distinguished element. Namely, using the lexicographical ordering (-1 precedes 1), among cyclic permutations of $(\varepsilon_1, \dots, \varepsilon_t)$ choose the smallest and call such a word an *m-Lyndon binary word*. An m-Lyndon binary word is primitive if and only if none of its non-trivial cyclic permutations are equal to it. Note that an m-Lyndon binary word selects a representative in the conjugacy class of an m-low lying word. For example, the Lyndon binary word $(-1,-1,-1,1,-1,-1,-1,1,-1,-1,1,1)$ selects the low lying word $[w] = [(ab^{-1})^3(ab)(ab^{-1})^3(ab)(ab^{-1})^2(ab)^2] \in L_{24}$. We have established the following.

Proposition 6.1. *Fix a positive integer $m \geq 2$. Then $L_{2t,m}$ is bijectively equivalent to the m-Lyndon binary words of length t.*

Remark 6.2. In the case that $m \geq t$ and we restrict to primitive binary words, such words are known as *Lyndon words* in the literature.

Our next goal is to derive an effective lower bound on the number of m-Lyndon words or equivalently the m-low lying conjugacy classes of length $2t$. We consider the normal form of the m-low lying words $ab^{\varepsilon_1}ab^{\varepsilon_2}\dots ab^{\varepsilon_t}$ or equivalently $(\varepsilon_1,\dots,\varepsilon_t)$ with cyclic runs of length at most m. In order to achieve a lower bound we construct a subset of m-low lying words of length $2t$ or equivalently m-low lying t-tuples. To that end, we picture t ordered slots and we group them into the first m, second m, and so on. There are exactly $\lceil \frac{t}{m} \rceil$ groups where the last grouping has less than or equal to m slots. We color all these slots black except the following which are colored red: the first one in the second group of m slots, the first one in the third group of m slots, and so on. If the t-th slot (that is, the last slot) is not red, it should also be colored red. Let $\mathcal{B}_{2t,m}$ be the subset of m-low lying words where any black slot can be $+1$ or -1, and the red slots are determined to insure there are no runs of length greater than m. Hence there are at least $t - \lfloor \frac{t}{m} \rfloor - 1$ black slots, and since the worst case up to cyclic conjugacy is that all t of the cyclic conjugates are distinct and in $\mathcal{B}_{2t,m}$, we have that the conjugacy classes of these elements satisfy

$$|B_{2t,m}| \geq \frac{2^{t-\lfloor \frac{t}{m} \rfloor -1}}{t}.$$

Theorem 6.3. *The following hold:*

1. $|L_{2t,m}| \geq |B_{2t,m}| \geq \frac{2^{t-\frac{t}{m}-1}}{t}$, *for* $m \geq 2$.
2. *For* $m \geq 3$, $|L^p_{2t,m}| \sim |L_{2t,m}|$, *as* $t \to \infty$.
3. *There exists* $t_0 > 0$ *so that* $|L^p_{2t,m}| \geq \frac{1}{2}\left(\frac{2^{t-\frac{t}{m}-1}}{t}\right)$, *for* $t \geq t_0$ *and* $m \geq 3$.
4. *There exists* $t_0 > 0$ *so that* $|L^p_{\leq 2t,m}| \geq \frac{1}{4}\sum_{s=t_0}^{t} \frac{2^{s-\frac{s}{m}}}{s}$, *for* $t \geq t_0$ *and* $m \geq 3$.

Proof. Item (1) was proven in the discussion before the theorem. To prove item (2), we first use Lemma 4.3 to bound the non-primitive low lying growth rate:

$$|L^{np}_{2t,m}| \leq |W^{np}_{2t}| \leq \frac{1}{2} t 2^{t/2}.$$

Hence,

$$1 \geq \frac{|L^p_{2t,m}|}{|L_{2t,m}|} = 1 - \frac{|L^{np}_{2t,m}|}{|L_{2t,m}|} \geq 1 - \frac{\frac{1}{2}t2^{t/2}}{\frac{2^{t-\frac{t}{m}-1}}{t}} \geq 1 - \frac{\frac{1}{2}t2^{t/2}}{\frac{2^{t-\frac{t}{3}-1}}{t}}, \tag{6.1}$$

where we have used $m \geq 3$ in the right-hand inequality.

Noting that the lower bound in expression (6.1) does not depend on m, item (3) follows from item (1) and by choosing t_0 large enough so that

$$\frac{|L^p_{2t,m}|}{|L_{2t,m}|} \geq \frac{1}{2}.$$

Finally, to prove item (4),

$$
\begin{aligned}
|L^p_{\leq 2t,m}| = \sum_{s=1}^{t} |L^p_{2s,m}| &= \sum_{s=1}^{t_0} |L^p_{2s,m}| + \sum_{s=t_0}^{t} |L^p_{2s,m}| \\
&\geq \sum_{s=1}^{t_0} |L^p_{2s,m}| + \frac{1}{2} \sum_{s=t_0}^{t} \Big(\frac{2^{s-\frac{s}{m}-1}}{s} \Big) \\
&\geq \frac{1}{2} \sum_{s=t_0}^{t} \Big(\frac{2^{s-\frac{s}{m}-1}}{s} \Big) = \frac{1}{4} \sum_{s=t_0}^{t} \frac{2^{s-\frac{s}{m}}}{s}. \qquad \blacksquare
\end{aligned}
$$

Remark 6.4. Since our eventual goal is to prove Theorem 1.1 and Corollary 1.2, it is critical that t_0 in the above theorem does not depend on m.

6.2. Low lying reciprocal words

In this section, we count the low lying reciprocal words. Recall that

$$\mathcal{N}_{4t} = \big\{ [a, \gamma] : \gamma \text{ a } (bb)\text{-word of length } 2t - 1 \big\},$$

where $[a, \gamma]$ is the group commutator of a and γ. We define $\pi : \mathcal{N}_{4t} \to R_{4t}$ to be the map taking elements in $\mathcal{N}_{4t}$ to its conjugacy class. From Section 3 we know that there are exactly two conjugacy class representatives in normal form for reciprocal words, hence π is a surjective 2-1 mapping.

Definition 6.5. Fix an integer $t > 0$. A composition of t is an ordered sequence of positive integers $(k_1, \ldots, k_l)$ which sums to t. The k_i's are called the parts of the composition. The set of all compositions of t is denoted by C_t. Compositions of t having parts bounded by a fixed positive integer m are denoted by $C_{t,m}$.

Next, define

$$g : \mathcal{N}_{4t} \to C_t, \quad [a, b^{\varepsilon_1} a b^{\varepsilon_2} \ldots a b^{\varepsilon_t}] \mapsto (k_1, \ldots, k_l).$$

Here $(k_1, \ldots, k_l)$ is the ordered sequence of lengths of $(+1)$ and (-1)-runs starting from the left in the ε_i's. For example, if $\omega = [a, b^{-1} a b^{-1} a b^{-1} a b^{1} a b^{1} a b^{-1} a b^{1}] \in \mathcal{N}_{28}$, then $g(\omega) = (3, 2, 1, 1) \in C_7$. We remark here that g is a surjective 2-1 mapping. Namely, suppose $\omega = [a, b^{\varepsilon_1} a b^{\varepsilon_2} \ldots a b^{\varepsilon_t}] \in \mathcal{N}_{4t}$ so that $g(\omega) = (k_1, \ldots, k_l) \in C_t$. Then there exists exactly one other element in $\mathcal{N}_{4t}$ whose image under g is $(k_1, \ldots, k_l)$, namely, $[a, b^{-\varepsilon_1} a b^{-\varepsilon_2} \ldots a b^{-\varepsilon_t}]$.

$$
\begin{array}{ccc}
\mathcal{N}_{4t} & \xrightarrow{\;g\;} & C_t \\
{\scriptstyle \pi}\big\downarrow & & \\
R_{4t} & &
\end{array}
\qquad\qquad
\begin{array}{ccc}
\mathcal{L}_{4t,m} \cap \mathcal{N}_{4t} & \xrightarrow{\;g_m\;} & C_{t,m} \\
{\scriptstyle \pi_m}\big\downarrow & \nearrow{\scriptstyle \Phi_m} & \\
L_{4t,m} \cap R_{4t} & &
\end{array}
\tag{6.2}
$$

The set of reciprocal words is filtered by low lying reciprocal words. This is because the set of all words $\mathcal{W}$ has filtration,

$$\mathcal{L}_{4t,1} \subset \mathcal{L}_{4t,2} \subset \cdots \subset \mathcal{L}_{4t,m} \subset \cdots .$$

This induces a filtration of each of the spaces in the left diagram of (6.2) to yield restricted mappings in the right diagram. Noting that the maps g_m and π_m are surjective 2-1 maps, even though the preimage of a point in $L_{4t,m} \cap R_{4t}$ is generically not the same as the preimage of a point in $C_{t,m}$, we have

$$2|L_{4t,m} \cap R_{4t}| = |\mathcal{L}_{4t,m} \cap \mathcal{N}_{4t}| = 2|C_{t,m}|.$$

Thus we have the following result.

Theorem 6.6. *For each positive integer $m \geq 2$, there is a bijection Φ_m from $L_{4t,m} \cap R_{4t}$ to $C_{t,m}$, given by the right diagram of* (6.2).

We have now reduced the problem to counting $C_{t,m}$, whose computation involves using the recursion relation: $|C_{t,m}| = \sum_{i=1}^{m} |C_{t-i,m}|$. The combinatorial analysis solving this problem follows from [6, Theorem 2], where $C_{t,m}$ is denoted by $F_{t+1}^{(m)}$. We have

$$|C_{t,m}| = \operatorname{rnd}\Big(\frac{\alpha_m - 1}{2 + (m+1)(\alpha_m - 2)}\alpha_m^t\Big), \tag{6.3}$$

where $\operatorname{rnd}(x) = \lfloor x + \frac{1}{2} \rfloor$ and α_m is the unique positive root of $z^m - z^{m-1} - \cdots - 1 = 0$. We remark that the coefficient $\frac{\alpha_m - 1}{2+(m+1)(\alpha_m-2)}$ in equation (6.3) only depends on m and thus we denote it by d_m. We note that $2(1 - 2^{-m}) \leq \alpha_m < 2$ and the α_m are increasing as m increases. For the details see [6].

We have proven the following theorem.

Theorem 6.7. *The following hold:*

(1) $d_m\alpha_m^t - \frac{1}{2} \leq |L_{4t,m} \cap R_{4t}| \leq d_m\alpha_m^t + \frac{1}{2}$ *for all* $t \geq 1$.

(2) $|L_{4t,m} \cap R_{4t}| \sim d_m\alpha_m^t$, *as* $t \to \infty$.

Corollary 6.8. *We have*

$$|L_{\leq 4t,m} \cap R_{\leq 4t}| \sim \Big(\frac{\alpha_m}{2 + (m+1)(\alpha_m - 2)}\Big)\alpha_m^t, \quad \text{as } t \to \infty.$$

Proof. First, note that

$$|L_{\leq 4t,m} \cap R_{\leq 4t}| = \sum_{n=1}^{t} |L_{4n,m} \cap R_{4n}|.$$

Applying item (1) of Theorem 6.7 gives us

$$\sum_{n=1}^{t}\Big(d_m\alpha_m^n - \frac{1}{2}\Big) \leq |L_{\leq 4t,m} \cap R_{\leq 4t}| \leq \sum_{n=1}^{t}\Big(d_m\alpha_m^n + \frac{1}{2}\Big).$$

Simplifying, we have

$$\Big(\frac{\alpha_m}{2+(m+1)(\alpha_m-2)}\Big)(\alpha_m^t-1)-\frac{t}{2} \\ \le |L_{\le 4t,m}\cap R_{\le 4t}| \le \Big(\frac{\alpha_m}{2+(m+1)(\alpha_m-2)}\Big)(\alpha_m^t-1)+\frac{t}{2}.$$

The result follows by dividing by $(\frac{\alpha_m}{2+(m+1)(\alpha_m-2)})\alpha_m^t$ and letting $t\to\infty$. ∎

7. Representation and geodesic excursion into the cusp

Consider the group $G=\mathbb{Z}_2 * \mathbb{Z}_3$ with generators a and b of the first and second factors, respectively. This group is isomorphic to the modular group. We consider the following representation of G: $a\mapsto A$ and $b\mapsto B$ where,

$$A=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \quad\text{and}\quad B=\begin{pmatrix}1 & -1\\ 1 & 0\end{pmatrix}.$$

This is a discrete, faithful representation with image $\mathrm{PSL}(2,\mathbb{Z})$. Let $S=\mathbb{H}/\mathrm{PSL}(2,\mathbb{Z})$ be the associated orbifold surface. It follows that S is a generalized pair of pants. In particular, S has zero genus and signature $(2,3,\infty)$.

It is well known that if an orbifold surface has a cusp then it has an embedded cusp of area one and boundary a horocycle segment of length one. A closed geodesic that wanders into (and hence out of a cusp) has a maximal depth in which it enters. More precisely, let γ be a closed geodesic on the modular orbifold and let $\mathcal{C}$ be the cusp of area one. The closed geodesic may wander in and out of the cusp a number of times, and each time it enters and exits the cusp we call this an *excursion* of γ. The *depth* of an excursion is the furthest distance into the cusp the excursion goes.

Lemma 7.1. *Let γ be a closed geodesic on S.*

(1) *An excursion of γ winds $k\ge 2$ times around the cusp if and only if the depth of the excursion is strictly between* $\log\frac{k}{2}$ *and* $\log\frac{k+1}{2}$.

(2) *γ is contained in the m-thick part of S if and only if some, and hence any, representative $g\in\mathrm{PSL}(2,\mathbb{Z})$ of γ is an m-low lying word.*

Proof. Let $\tilde\gamma$ be a lift of γ normalized so that its endpoints at infinity are $-r$ and r and the parabolic associated to the cusp normalized to be $f(z)=z+1$. The depth for this excursion into the cusp is $\log r$. Now if the excursion winds around the cusp k-times then $f^k(\tilde\gamma)\cap\tilde\gamma\ne\emptyset$ and $f^{k+1}(\tilde\gamma)\cap\tilde\gamma=\emptyset$. That is, $\frac{k}{2}<r<\frac{k+1}{2}$. See Figure 2. Note that equality is not included as that would violate the fact that two hyperbolic elements in a Fuchsian group can not share a unique fixed point. Equivalently, $\log\frac{k}{2}<\log r<\log\frac{k+1}{2}$. These steps are reversible. Hence we have proven item (1).

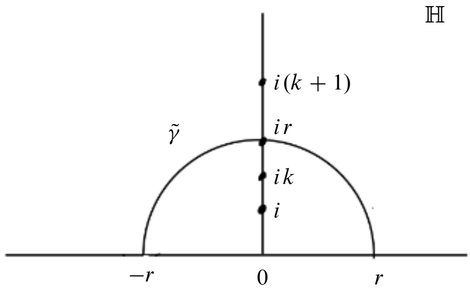

Figure 2. The hyperbolic plane

To prove item (2), the word g written as a product of the generators in normal form is the product of the inverse conjugate parabolic elements AB and AB^{-1}. Suppose g is m-low lying. Then the longest run of AB or AB^{-1} (considered cyclically) is at most m. Now, a run in the word g, say k, corresponds to γ winding around the cusp k times. By item (1), we know that the depth of this excursion is at most $\log \frac{k+1}{2} \leq \log \frac{m+1}{2}$. Thus γ is contained in the m-thick part of S. For the converse, if γ is in the m-thick part, item (1) again guarantees that there is no run of AB or AB^{-1} longer than m. Therefore g is an m-low lying word. ■

8. All together now: The proof of Theorem 1.1

In this section, we put together the work of the previous sections to prove Theorem 1.1. The closed geodesics on S correspond to conjugacy classes of hyperbolic elements in $\mathrm{PSL}(2, \mathbb{Z})$. Similarly, reciprocal closed geodesics correspond to hyperbolic elements whose axes pass through an order-two fixed point. Finally, low lying closed geodesics correspond to conjugacy classes of low lying words, as in Lemma 7.1. Using the notation from the previous sections, we have the following correspondence between geodesics and conjugacy classes of words in the group:

$$\begin{aligned}
&\big\{\gamma \text{ a primitive reciprocal geodesic with } |\gamma| \leq 2t\big\} \\
&\longleftrightarrow R^{p}_{\leq 2t}, \\
&\big\{\gamma \text{ a primitive reciprocal geodesic in } S_{m\geq 2} \text{ with } |\gamma| \leq 2t\big\} \\
&\longleftrightarrow L^{p}_{\leq 2t,m} \cap R^{p}_{\leq 2t}, \\
&\big\{\gamma \text{ a primitive closed geodesic with } |\gamma| \leq 2t\big\} \\
&\longleftrightarrow \big\{[w] \in W^{p}_{\leq 2t} : \|[w]\| > 2\big\},
\end{aligned}$$

and

$$\begin{aligned}&\{\gamma \text{ a primitive closed geodesic in } S_{m\geq 3} \text{ with } |\gamma| \leq 2t\}\\ &\longleftrightarrow \{[w] \in L^{p}_{\leq 2t,m} : \|[w]\| > 2\}.\end{aligned}$$

Proof of Theorem 1.1. To prove item (1), we first note that

$$|R_{\leq 2t}| = \sum_{n=1}^{\lfloor \frac{t}{2} \rfloor} |R_{4n}| = 2^{\lfloor \frac{t}{2} \rfloor} - 1,$$

where the last equality follows as in the proof of the second part of Lemma 5.1. Using the fact that $|R_{2t}| \leq \frac{1}{2} 2^{\frac{t}{2}}$, we can establish $|R^{p}_{\leq 2t}| \sim |R_{\leq 2t}|$ in a similar way to what was done for length $4t$.

To prove item (2), note that

$$|L_{\leq 2t,m} \cap R_{\leq 2t}| = \sum_{n=1}^{\lfloor \frac{t}{2} \rfloor} |L_{4n,m} \cap R_{4n}|.$$

As in the proof of Corollary 6.8, applying item (1) from Theorem 6.7 yields

$$|L_{\leq 2t,m} \cap R_{\leq 2t}| \sim \Big(\frac{\alpha_m}{2 + (m+1)(\alpha_m - 2)}\Big) \alpha_m^{\lfloor \frac{t}{2} \rfloor}.$$

Using the fact that $|L_{2t,m} \cap R_{2t}| \leq \operatorname{rnd}(d_m \alpha_m^{\frac{t}{2}})$, it is not difficult to show that

$$|L^{p}_{\leq 2t,m} \cap R^{p}_{\leq 2t}| \sim |L_{\leq 2t,m} \cap R_{\leq 2t}|.$$

Item (3) follows from Theorem 4.4 and the fact that the primitive hyperbolic conjugacy classes have the same growth as all primitive conjugacy classes.

Lastly, item (4) follows from Theorem 6.3, item (4), and an application of the Stolz–Cesaro theorem to show

$$\sum_{s=t_0}^{t} \frac{2^{s-\frac{s}{m}}}{s} \sim \frac{2}{2 - 2^{\frac{1}{m}}} \Big(\frac{2^{t-\frac{t}{m}}}{t}\Big). \qquad \blacksquare$$

Acknowledgments. The authors would like to thank Hugo Parlier and Ser Peow Tan for informative conversations as well as the referees for their comments, which helped to improve the overall readability of the paper.

Funding. The first author was supported by a PSC-CUNY grant and a grant from the Simons foundation (359956, A.B.). The second author was supported in part by the Faculty Development Plan through the School of Science at Manhattan College.

References

[1] F. P. Boca, V. Paşol, A. A. Popa, and A. Zaharescu, Pair correlation of angles between reciprocal geodesics on the modular surface. *Algebra Number Theory* **8** (2014), no. 4, 999–1035 Zbl 1304.11073 MR 3248992

[2] J. Bourgain and A. Kontorovich, Beyond expansion II: low-lying fundamental geodesics. *J. Eur. Math. Soc. (JEMS)* **19** (2017), no. 5, 1331–1359 Zbl 1403.11054 MR 3635355

[3] J. Bourgain and A. Kontorovich, Beyond expansion, III: Reciprocal geodesics. *Duke Math. J.* **168** (2019), no. 18, 3413–3435 Zbl 1447.11079 MR 4034890

[4] P. Buser, *Geometry and spectra of compact Riemann surfaces*. Progr. Math. 106, Birkhäuser, Boston, MA, 1992 Zbl 0770.53001 MR 1183224

[5] D. Calegari and J. Louwsma, Immersed surfaces in the modular orbifold. *Proc. Amer. Math. Soc.* **139** (2011), no. 7, 2295–2308 Zbl 1276.20049 MR 2784794

[6] G. P. B. Dresden and Z. Du, A simplified Binet formula for k-generalized Fibonacci numbers. *J. Integer Seq.* **17** (2014), no. 4, Article 14.4.7 Zbl 1360.11031 MR 3181762

[7] V. Erlandsson, A remark on the word length in surface groups. *Trans. Amer. Math. Soc.* **372** (2019), no. 1, 441–455 Zbl 1429.30034 MR 3968775

[8] V. Erlandsson, H. Parlier, and J. Souto, Counting curves, and the stable length of currents. *J. Eur. Math. Soc. (JEMS)* **22** (2020), no. 6, 1675–1702 Zbl 1439.53040 MR 4092896

[9] V. Erlandsson and J. Souto, Counting curves in hyperbolic surfaces. *Geom. Funct. Anal.* **26** (2016), no. 3, 729–777 Zbl 1350.57017 MR 3540452

[10] J. Gilman, Informative words and discreteness. In *Combinatorial group theory, discrete groups, and number theory*, pp. 147–155, Contemp. Math. 421, American Mathematical Society, Providence, RI, 2006 Zbl 1116.30030 MR 2303833

[11] J. Gilman, Primitive curve lengths on pairs of pants. In *Infinite group theory*, pp. 141–155, World Scientific, Hackensack, NJ, 2018 MR 3586884

[12] J. Gilman and L. Keen, Word sequences and intersection numbers. In *Complex manifolds and hyperbolic geometry* (Guanajuato, 2001), pp. 231–249, Contemp. Math. 311, American Mathematical Society, Providence, RI, 2002 Zbl 1015.30021 MR 1940172

[13] J. Gilman and L. Keen, Cutting sequences and palindromes. In *Geometry of Riemann surfaces*, pp. 194–216, London Math. Soc. Lecture Note Ser. 368, Cambridge University Press, Cambridge, 2010 Zbl 1194.20026 MR 2665010

[14] J. Gilman and L. Keen, Enumerating palindromes and primitives in rank two free groups. *J. Algebra* **332** (2011), 1–13 Zbl 1237.20023 MR 2774675

[15] J. Gilman and L. Keen, Lifting free subgroups of $PSL(2,\mathbb{R})$ to free groups. In *Quasiconformal mappings, Riemann surfaces, and Teichmüller spaces*, pp. 109–122, Contemp. Math. 575, American Mathematical Society, Providence, RI, 2012 Zbl 1256.30027 MR 2933896

[16] V. Guba and M. Sapir, On the conjugacy growth functions of groups. *Illinois J. Math.* **54** (2010), no. 1, 301–313 Zbl 1234.20041 MR 2776997

[17] A. Haas, Diophantine approximation on hyperbolic Riemann surfaces. *Acta Math.* **156** (1986), no. 1-2, 33–82 Zbl 0593.10028 MR 822330

[18] A. Kontorovich, Applications of thin orbits. In *Dynamics and analytic number theory*, pp. 289–317, London Math. Soc. Lecture Note Ser. 437, Cambridge University Press, Cambridge, 2016 Zbl 1406.11068 MR 3618792

[19] R. C. Lyndon and P. E. Schupp, *Combinatorial group theory*. Class. Math., Springer, Berlin, 2001 Zbl 0997.20037 MR 1812024

[20] W. Magnus, A. Karrass, and D. Solitar, *Combinatorial group theory*. 2nd edn., Dover Publications, Inc., New York, 1976 Zbl 0362.20023 MR 0422434

[21] G. McShane and I. Rivin, Simple curves on hyperbolic tori. *C. R. Acad. Sci. Paris Sér. I Math.* **320** (1995), no. 12, 1523–1528 Zbl 0835.53050 MR 1340065

[22] B. T. Marmolejo, *Growth of conjugacy classes of reciprocal words in triangle groups.* Ph.D. thesis, City University of New York, 2020 MR 4187474

[23] M. Mirzakhani, Growth of the number of simple closed geodesics on hyperbolic surfaces. *Ann. of Math. (2)* **168** (2008), no. 1, 97–125 Zbl 1177.37036 MR 2415399

[24] M. Mureşan, *A concrete approach to classical analysis*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2009 Zbl 1163.26001 MR 2456018

[25] P. S. Park, Conjugacy growth of commutators. *J. Algebra* **526** (2019), 423–458 Zbl 1447.20014 MR 3920909

[26] I. Rivin, Growth in free groups (and other stories)—twelve years later. *Illinois J. Math.* **54** (2010), no. 1, 327–370 Zbl 1225.05128 MR 2776999

[27] P. Sarnak, Reciprocal geodesics. In *Analytic number theory*, pp. 217–237, Clay Math. Proc. 7, American Mathematical Society, Providence, RI, 2007 Zbl 1198.11039 MR 2362203

[28] C. Series, The geometry of Markoff numbers. *Math. Intelligencer* **7** (1985), no. 3, 20–29 Zbl 0566.10024 MR 795536

[29] C. Traina, The conjugacy problem of the modular group and the class number of real quadratic number fields. *J. Number Theory* **21** (1985), no. 2, 176–184 Zbl 0567.10018 MR 808285

Received 20 August 2020.

Ara Basmajian
The Graduate Center, CUNY, 365 Fifth Ave., New York, NY 10016; and Hunter College, CUNY, 695 Park Ave., New York, NY 10065, USA; abasmajian@gc.cuny.edu

Robert Suzzi Valli
Manhattan College, 4513 Manhattan College Pkwy., Riverdale, NY 10471, USA; robert.suzzivalli@manhattan.edu