FIBERED COMMENSURABILITY AND ARITHMETICITY OF RANDOM MAPPING TORI

HIDETOSHI MASAI

Abstract. We consider a random walk on the mapping class group of a surface of finite type. We assume that the random walk is determined by a probability measure whose support is finite and generates a non-elementary subgroup $H$. We further assume that $H$ is not consisting only of lifts with respect to any one covering. Then we prove that the probability that the random walk gives a non-minimal mapping class in its fibered commensurability class decays exponentially. As an application of the minimality, we prove that for the case where a surface has at least one puncture, the probability that the random walk gives a mapping class with an arithmetic mapping torus decays exponentially.

1. Introduction

Let $S$ be an orientable surface of finite type $(g, n)$, where $g$ is the genus and $n$ is the number of punctures. We consider random walks on the mapping class group $G := \text{Mod}(S)$ which are determined by probability measures on $G$ each of whose support generates a non-elementary subgroup. It has been shown that such a random walk gives rise to pseudo-Anosov elements with asymptotic probability one [15, 17, 18, 23]. Let $\mu$ be a probability measure on $G$. A subset $A \subset G$ is said to be exponentially small (with respect to $\mu$) if the probability that the random walk determined by $\mu$ visits $A$ decays exponentially with the number of steps. A subset is called exponentially large (with respect to $\mu$) if its complement is exponentially small. Then the work of Maher [18] can be stated as “the set of pseudo-Anosov elements is exponentially large”. In this paper, we consider fibered commensurability, a notion introduced by Calegari-Sun-Wang [6], of random mapping classes. Roughly, a mapping class $\phi$ is said to cover another mapping class $\varphi$ if $\phi$ is a power of some lift of $\varphi$ with respect to some finite covering of underlying surfaces. The commensurability with respect to this covering relation is called fibered commensurability. Each commensurability class enjoys an order by the covering relation. It has been shown [6, 20] that for pseudo-Anosov case, each commensurability class contains a unique minimal (orbifold) element (see Theorem 2.3). Our aim is to prove that the set of minimal elements is exponentially large with respect to any measure which satisfies a suitable condition (Condition 1.2). As an application of the minimality, we also show a result on arithmeticity of random mapping tori. By using random walks on $G$, we may generate randomly 3-manifolds by taking mapping tori. The work of Thurston [24] together with [18] shows that the set of mapping classes with hyperbolic mapping tori is exponentially large. A cusped hyperbolic 3-manifold is called arithmetic if it is commensurable to a Bianchi orbifold (see [5]).

2010 Mathematics Subject Classification. 57M50, 60G50, 20F65.
Several distinguished hyperbolic 3-manifolds, for example the complement of the figure eight knot or the Whitehead link, are known to be arithmetic. However, a “generic” hyperbolic 3-manifold is believed to be non-arithmetic. The minimality of random mapping classes together with the work by Bowditch-Maclachlan-Reid \cite{4} enables us to prove that the set of mapping classes with arithmetic mapping tori is exponentially small if $S$ has at least one puncture.

The paper is organized as follows. In §2 we prepare several definitions and facts about random walks on groups and mapping class groups. Note that to prove that a given mapping class $\phi$ is minimal, it suffices to show that $\phi$ is primitive and not a lift. In §3 we prove the primitivity of random mapping classes.

**Theorem 1.1.** Let $\mu$ be a probability measure on $G$ whose support is finite and generates a non-elementary subgroup. Then the set of primitive elements in $G$ is exponentially large with respect to $\mu$.

Next, we prove that random mapping classes are not lifts in §4. We need further assumption for the measure $\mu$ to avoid the case that there is some finite covering $\pi : S \to S'$ such that every element in the support of $\mu$ is a lift of a mapping class on $S'$. Let $\mathcal{PMF}(S)$ denote the set of projective measured foliations on $S$, where in case of orbifolds, we consider the one for the surface we get by puncturing the orbifold points. Each covering $\pi : S \to S'$ determines a map $\Pi : \mathcal{PMF}(S') \to \mathcal{PMF}(S)$ so that $a \in \Pi(\mathcal{PMF}(S'))$ if and only if $\pi(a) \in \mathcal{PMF}(S')$. Let $\text{sgr}(\mu)$ denote the semigroup generated by the support of $\mu$. Then the condition for the measure $\mu$ which we need is the following.

**Condition 1.2.** The support is finite, and generates a non-elementary subgroup of $G$. Moreover, for any (possibly orbifold) covering $\pi : S \to S'$, $\text{sgr}(\mu)$ contains a pseudo-Anosov element whose fixed points are not in $\Pi(\mathcal{PMF}(S'))$.

In §4 we prove:

**Theorem 1.3.** Let $\mu$ be a probability measure on $G$ which satisfies Condition 1.2. Then the set of elements which is a lift with respect to some (possibly orbifold) covering $\pi : S \to S'$ is exponentially small with respect to $\mu$.

Putting Theorem 1.1 and 1.3 together, we have:

**Theorem 1.4.** Let $\mu$ be a probability measure on $G$ which satisfies Condition 1.2. Then the set of minimal elements in $G$ is exponentially large with respect to $\mu$.

Finally in §5 we prove the following theorem.

**Theorem 1.5.** Suppose that $S$ has at least one puncture. Let $\mu$ be a probability measure on $G$ which satisfies Condition 1.2. Then the set of mapping classes with arithmetic mapping tori is exponentially small with respect to $\mu$.

2. Preliminary

In this section, we summarize several definitions and facts that we use throughout the paper. Interested readers may refer to several papers regarding to random walks on the mapping class groups (for example \cite{13,17}) in which there are detailed expositions of basic theory of both random walks and mapping class groups.
2.1. Random walks on groups. We recall the definitions and terminologies of random walks. See [26] for more details about random walks on groups. Let $G$ be a countable group. A (possibly infinite) matrix $P = (p_{g,h})_{g,h \in G}$ is called stochastic if every element is non-negative and
\[
\sum_{h \in G} p_{g,h} = 1
\]
for all $g \in G$. For a given probability measure $\mu$ on $G$, by putting $p_{g,h} = \mu(hg^{-1})$, we have a stochastic matrix $P_\mu = (p_{g,h})_{g,h \in G}$. Let $P_n$ denote the probability measure on $(G^n, 2^{G^n})$ defined by
\[
P_n(A) = \sum_{(g_1, \ldots, g_n) \in A} p_{g_1,g_2} \cdots p_{g_{n-1},g_n} \quad \text{for } A \subseteq 2^{G^n}.
\]
Note that by the definition, we have $P_n(A \times G) = P_n(A)$ for any $A \subseteq 2^{G^n}$. Let $\mathcal{B}(G^n)$ denote the $\sigma$-algebra generated by cylinder sets, where a cylinder set is a subset defined as
\[
\{ \omega = (\omega_n)_{n \in \mathbb{N}} \in G^n \mid (\omega_1, \ldots, \omega_n) \in A \}
\]
for some $A \subset G^n$. Then by the Kolmogorov extension theorem, there exists a unique measure $P$ on $(G^n, \mathcal{B}(G^n))$ which satisfies
\[
P(A \times G^n) = P_n(A) \quad \text{for all } n \in \mathbb{N}, \text{ and } A \subseteq 2^{G^n}.
\]
For $\omega = (\omega_n) \in G^n$, we define $G$-valued random variables $X_n$ on $(G^n, \mathcal{B}(G^n))$ by $X_n(\omega) = \omega_n$. Thus we have a stochastic process $\{X_n\}_{n \in \mathbb{N}}$ which is a Markov chain with the transition matrix $P_\mu$. We call this Markov chain $\{X_n\}_{n \in \mathbb{N}}$ the random walk determined by $\mu$.

Let us fix a probability measure $\mu$ and the random walk determined by $\mu$. Each element $(\omega_n)_{n \in \mathbb{N}} \in G^n$ is called a sample path. Let $A \subseteq G$. By abbreviation of notations, we write $P(\omega_n \in A)$ to mean $P(G^{n-1} \times A \times G^n)$. A subset $A \subset G$ is called exponentially small (with respect to $\mu$) if there exist $c < 1, K > 0$ which depend only on $\mu$ such that $P(\omega_n \in A) < Ke^n$. A subset is called exponentially large (with respect to $\mu$) if its complement is exponentially small. Let $Q$ be a property for elements in $G$. We say that the random walks determined by $\mu$ has property $Q$ with exponentially high probability if $S_Q := \{ g \in G \mid g \text{ is } Q \}$ is exponentially large. It can be readily seen that if $A, B \subset G$ are both exponentially small (resp. large), then so is $A \cup B$ (resp. $A \cap B$).

2.2. Mapping class groups and curve graphs. For more details about topics in this subsection, one may refer to the books [3, 7]. Let $S := S_{g,n}$ be an orientable surface of finite type $(g,n)$ where $g$ is the genus and $n$ is the number of punctures.

In this paper, we always suppose $3g - 3 + n > 0$ unless otherwise stated. The mapping class group $\text{Mod}(S)$ is the group of isotopy classes of orientation preserving automorphisms on $S$. A mapping class is called pseudo-Anosov if it is aperiodic and has no fixed 1-dimensional submanifold of $S$. Thurston [25] showed that each pseudo-Anosov mapping class has exactly two fixed points $F_s, F_u$ in the space $\mathcal{PMF}(S)$ of projective measured foliations. A subgroup of $\text{Mod}(S)$ is called non-elementary if it contains two pseudo-Anosov mapping classes with distinct fixed points in $\mathcal{PMF}(S)$.

The curve graph $\mathcal{C}(S)$ of $S$ is a graph whose vertices consist of isotopy classes of simple closed curves, and two vertices are connected by an edge if the corresponding
curves can be disjointly represented on $S$. By giving length 1 to every edge, the curve graph enjoys a metric $d_{c(S)}(\cdot, \cdot)$.

Let $(X, d_X)$ be a metric space. For a fixed point $p \in X$, the Gromov product $(x, x')_p$ of two points $x, x' \in X$ is defined by

$$(x \cdot x')_p = \frac{1}{2}(d_X(x, p) + d_X(x', p) - d_X(x, x')).$$

Then for $r > 0$, a shadow $S_p(x, r) \subset X$ is defined by

$$S_p(x, r) := \{ y \in X \mid (x \cdot y)_p \geq r \}.$$ 

If we have another metric space $(Y, d_Y)$, a map $f : X \to Y$ is said to be $Q$-quasi-isometric if for any $x, x' \in X$,

$$d_X(x, x')/Q - Q \leq d_Y(f(x), f(x')) \leq Qd_X(x, x') + Q.$$ 

Such $f$ is called $Q$-quasi-isometry if it further satisfies that for any $y \in Y$, there exists $x \in X$ such that $d_Y(y, f(x)) < Q$. Two metric spaces are said to be quasi-isometric if there is a $Q$-quasi-isometry between the two. Suppose further that $X$ is a geodesic space. Then $X$ is called $\delta$-hyperbolic if every geodesic triangle is $\delta$-thin; one side of a geodesic triangle is contained in the $\delta$-neighborhood of the other two sides. $X$ is called hyperbolic if it is $\delta$-hyperbolic for some $\delta \geq 0$. Two geodesics in $X$ are said to be asymptotic if they are finite Hausdorff distance apart. We may define the Gromov boundary as the set of asymptotic classes of geodesics. Hyperbolicity is invariant under quasi-isometries, and a quasi-isometry induces a homeomorphism of the Gromov boundaries. For two points $x, x'$ in a geodesic space $X$, we denote by $[x, x']$ a geodesic connecting $x$ and $x'$. Note that there can be many such geodesics, and $[x, x']$ is an arbitrarily chosen one. We suppose that if $a, b \in [x, x']$, then $[a, b] \subset [x, x']$.

**Remark 2.1.** It is well known that if $X$ is $\delta$-hyperbolic, the Gromov product $(x, x')_p$ is equal to the distance from $p$ to $[x, x']$ up to additive constant $K$ which depends only on $\delta$ (c.f. Lemma [1.6]). By this fact, a shadow $S_p(x, r)$ for $x \in X$ and $r > 0$ can be (coarse equivalently) regarded as the set of $x' \in X$ such that every geodesic connecting $p$ and $x'$ passes through a point in the $(d_X(x, p) - r + C)$-neighborhood of $x$ for some $C$ depending only on $\delta$.

In [21], Masur-Minsky proved that the curve graph $C(S)$ is hyperbolic. The mapping class group $G := \text{Mod}(S)$ acts isometrically on $C(S)$. Using this action, by fixing a base point $p \in C(S)$, $G$ admits a $\delta$-hyperbolic (improper) metric by

$$d(g, h) = d_{C(S)}(gp, hp).$$

2.3. **Commensurability of mapping classes.** In [6], Calegari-Sun-Wang defined commensurability of mapping classes on possibly distinct surfaces as follows.

**Definition 2.2 ([6]).** Let $S_1$ and $S_2$ be orientable surfaces of finite type. A mapping class $\phi_1 \in \text{Mod}(S_1)$ covers $\phi_2 \in \text{Mod}(S_2)$ if there exists a finite covering $\pi : S_1 \to S_2$ and $k \in \mathbb{Z} \setminus \{0\}$ such that a lift $\varphi$ of $\phi_2$ with respect to $\pi$ satisfies $\varphi^k = \phi_1$. Two mapping classes are said to be commensurable if there exists a mapping class that covers both.

Since this gives commensurability of the monodromies of fibers on orientable surface bundles over the circle, this notion is also called fibered commensurability. Fibered commensurability gives rise to an equivalence relation by taking transitive
Theorem 2.3 (\[6, 20\]). If \( \phi \in \text{Mod}(S) \) is pseudo-Anosov, then the commensurability class of \( \phi \) contains a unique minimal (orbifold) element.

Note that a mapping class \( \phi \) is minimal if it is primitive (i.e. if \( \varphi^k = \phi \), then \( k = 1 \) and \( \phi = \varphi \), or \( k = -1 \) and \( \phi = \varphi^{-1} \)) and it is not a lift of any orbifold automorphism.

3. Random mapping classes are primitive

Throughout this section, let us fix an orientable surface \( S \) of finite type and denote by \( G \) the mapping class group \( \text{Mod}(S) \). To prove the primitivity, we consider the action of \( G \) on the curve graph \( \mathcal{C}(S) \). We shall fix a base point \( p \in \mathcal{C}(S) \). For \( g \in G \), the translate \( gp \in \mathcal{C}(S) \) is also denoted by \( g \) by abuse of notation. We abbreviate the distance on \( \mathcal{C}(S) \) to \( d_{\mathcal{C}}(\cdot, \cdot) \). In this section, unless otherwise stated, we consider the random walk determined by a probability measure \( \mu \) on \( G \) with finite support which generates a non-elementary subgroup.

3.1. Random mapping classes do not (anti-)align. We first recall the work of Calegari-Maher [5].

**Definition 3.1.** Let \( p_0, \ldots, p_n \) be points in \( \mathcal{C}(S) \) and \( \gamma = [p_0, p_n] \). A point \( y \in \gamma \) is \( D \)-proximal (with respect to \( p_0, \ldots, p_n \)) if \( d_{\mathcal{C}}(y, p_i) < D \) for some \( 0 \leq i \leq n \). Let \( \gamma_D \) denote the subset of \( D \)-proximal points of \( \gamma \).

Let \( \omega = (\omega_n) \) be a sample path in \( G^\mathbb{N} \), then for large enough \( n \), Calegari-Maher proved that most part of \([\omega_0, \omega_n]\) should be \( D \)-proximal with exponentially high probability.

**Lemma 3.2** ([5, Lemma 5.14]). There are constants \( C_1, K > 0 \) and \( c < 1 \) so that for any \( \epsilon > 0 \), there is a further constant \( D \) depending on \( C_1 \) and \( \epsilon \) with the following property. Let \( \gamma := [\omega_0, \omega_n] \) and \( \gamma_D \) denote the set of \( D \)-proximal points on \( \gamma \) with respect to \( \omega_0, \ldots, \omega_n \in \mathcal{C}(S) \). Then

\[
P((\text{length}(\gamma) \geq C_1 n) \land (\text{length}(\gamma_D)/\text{length}(\gamma) \geq 1 - \epsilon)) \geq 1 - Ke^n.
\]

Lemma 3.2 shows that coarsely, a random walk fellow travels with a geodesic connecting the endpoints with exponentially high probability.

We also recall the work of Maher which shows that each shadow is exponentially small.

**Lemma 3.3** ([13]). There are constants \( K > 0 \) and \( c < 1 \) such that for any \( q \in \mathcal{C}(S) \) and any \( r \),

\[
P(\omega_n \in S_1(q, r)) < Ke^r.
\]

In what follows, we suppose that a path in \( \mathcal{C}(S) \) is a continuous map \([0, 1] \rightarrow \mathcal{C}(S) \). Hence for a given path \( \gamma \), \( \gamma(0) \) denotes the initial point and \( \gamma(1) \) denotes the terminal point. Two paths \( \gamma_1 \) and \( \gamma_2 \) are said to be \( D \)-aligned (resp. \( D \)-anti-aligned) if there exists \( h \in G \) such that \( d_\mathcal{C}(h\gamma_1(0), \gamma_2(0)) < D \) and \( d_\mathcal{C}(h\gamma_1(1), \gamma_2(1)) < D \).
(resp. $d_C(h\gamma_1(1), \gamma_2(0)) < D$ and $d_C(h\gamma_1(0), \gamma_2(1)) < D$). Lemma 3.3 below looks quite similar to [5, Lemma 5.26] showing the probability that a random walk has two anti-aligned subpaths decays polynomially. Lemma 3.4 shows the probability that a random walk has aligned subpaths decays exponentially. The order of the decay is exponential since we consider the case that a random walk has aligned subpaths of length of linear order (see property (1) of Lemma 3.4) while in [5], the order was of logarithm. Although one can prove Lemma 3.4 by almost the same argument as in [5], we include a proof for completeness. Recall that by the work of Bowditch [2], the action of $G$ on $C(S)$ is acylindrical; for any $C_1 > 0$, there are constants $C_2, C_3$ such that for $a, b \in C(S)$ with $d_C(a, b) \geq C_2$, there are at most $C_3$ elements $h \in G$ with $d_C(a, ha) \leq C_1$ and $d_C(b, hb) \leq C_1$.

**Lemma 3.4** (c.f. [5, Lemma 5.26]). Fix $D, M > 0$. Then there is a constant $c_1 < 1$, $K > 0$ such that the following holds. Consider the collection of indices $a < a' < b < c < c' < d$ for which there are geodesics $\alpha \in [\omega_a, \omega_b]$ and $\beta \in [\omega_c, \omega_d]$ with the following properties:

1. $\text{length}(\alpha) \geq Mn$ and similarly for $\beta$;
2. there is $t \in [0, 1, 0.2]$ so that $d_C(\omega_{a'}, \alpha(t)) \leq D$ and $d_C(\omega_{c'}, \beta(t)) \leq D$;
3. there is some $h \in G$ so that $d_C(h\alpha(0), \beta(0)) \leq D$, and $d_C(h\alpha(1), \beta(1)) \leq D$.

The probability that this collection of indices is non-empty is at most $Kc_1^3$. 

**Proof.** We first fix $a < a' < b < c < c'$. To satisfy conditions (2) and (3), we need to have $h \in G$ such that $d_C(h\omega_a, \omega_b) \leq C_1$ and $d_C(h\omega_a, \omega_b) \leq C_1$ for some constant $C_1$ depending only on $D$ and the hyperbolicity constant $\delta$. Hence, the acylindricity of the action of $G$ on $C(S)$ implies that if $\alpha = [\omega_a, \omega_b]$ is long enough, there is a set $A \subset C(S)$ of at most $C_3$ points so that $\omega_d$ should be in $D$ neighborhood of $x \in A$ where $C_3$ depends only on $D$ and $\delta$. As we noted in Remark 2.1 it follows that $\omega_d \in S_{\omega_c}(x, d_C(\omega_c, x) - C)$ for some $C$ depending only on $\delta$. Then by Lemma 3.3, the probability that a random walk from $\omega_d$ is in $S_{\omega_c}(x, d_C(\omega_c, x) - C)$ decays exponentially since $d_C(\omega_c, x) = \text{at least } 8Mn/10$ by the conditions (1) and (2). Since the number of elements of $A$ is universally bounded, the probability that $a < a' < b < c < c'$ satisfies (1)-(3) is less than $Kc_2^3$ for some $K' > 0$ and $c_2 < 1$ which depend only on $D, \delta$ and $M$ but on $n$ and $a < a' < b < c < c'$. The number of all possible choices of $a < a' < b < c < c'$ is of order $n^3$. We may find some $K > 0$ and $c_1 < 1$ such that $n^3Kc_2^3 < Kc_1^3$. Thus we complete the proof. 

**Remark 3.5.** As shown in [5], almost the same argument shows anti-aligned version of Lemma 3.4. Namely, we may replace the conditions (2) and (3) of Lemma 3.4 with

(2’) there is $t \in [0, 1, 0.2]$ so that $d_C(\omega_{a'}, \alpha(1 - t)) \leq D$ and $d_C(\omega_{c'}, \beta(t)) \leq D$,

(3’) there is some $h \in G$ so that $d_C(h\alpha(0), \beta(1)) \leq D$, and $d_C(h\alpha(1), \beta(0)) \leq D$,

to have the probability that we have indices satisfying (1), (2’) and (3’) decays exponentially.

### 3.2. Proof of Theorem 1.1

For $g \in G$, let $\tau(g)$ denote the translation length

$$
\tau(g) := \lim_{n \to \infty} \frac{d_C(g^n(p), p)}{n}
$$

of $g$ on the curve graph $C(S)$. Maher-Tiozzo proved that the translation length grows linearly [19].
Lemma 3.6 ([19]). There exists $L > 0$, $K > 0$ and $c < 1$ which only depends on $S$ and $\mu$ such that

$$\mathbb{P}(\tau(\omega_n) < Ln) < Ke^n.$$

We first prepare an elementary observation for an action of a group on a $\delta$-hyperbolic space.

Proposition 3.7 (c.f. [17] Lemma 3.3]). Let $H$ be a group acting isometrically on a $\delta$-hyperbolic space $(Y, d_Y)$ with a base point $x$. Fix $h \in H$. Suppose that $h$ has a geodesic axis $\alpha$, i.e. a geodesic satisfying $h^n(\alpha) \subset N_{2\delta}(\alpha)$ for all $n \in \mathbb{Z}$ where $N_{2\delta}(\alpha)$ denotes the $2\delta$ neighborhood of $\alpha$. Let $q$ be a nearest point projection of $x$ to $\alpha$. If $d_Y(q, hq) > 28\delta$, the following holds. There exist $D_1, D_2 \geq 0$ which depend only on $\delta$ such that the geodesic $\gamma = [x, hx]$ can be decomposed into three subsegments $\gamma = \gamma_1\gamma_2\gamma_3$ so that

- The distance $d_c(\gamma_1(1), q) \leq D_1$ and $d_c(\gamma_3(0), hq) \leq D_1$, and
- $\gamma_2 \subset N_{D_2}(\alpha)$ and $\text{length}(\gamma_2) \geq d_Y(q, hq) - 28\delta$.

Proof. Any side of a geodesic quadrilateral in a $\delta$-hyperbolic space is in the $2\delta$ neighborhood of the other three sides. We consider a geodesic quadrilateral whose vertices are $x, q, hq, hx$. Since $q, hq$ are nearest point projections, if a point $s \in [q, hq]$ is at least $4\delta$ apart from $q$ and $hq$, then $d_Y(s, \gamma) \leq 2\delta$. This is because if $d_Y(s, \gamma) > 2\delta$, then there must be $s' \in [x, q] \cup [hq, hx]$ such that $d_Y(s, s') \leq 2\delta$, which contradicts the fact that $q$ and $hq$ are nearest point projections to $\alpha$. Let $x'_1$ (resp. $x'_2$) be a nearest point projection to $\gamma$ of the point $q_1$ (resp. $q_2$) on $[q, hq]$ which is exactly $4\delta$ apart from $q$ (resp. $hq$). Then $d_Y(x'_i, q) \leq 6\delta$ for $i = 1, 2$. By $\delta$-hyperbolicity, if a point $a \in [x'_1, x'_2]$ is at least $4\delta$ away from both $x'_1$ and $x'_2$, then $d_Y(a, [q, hq]) \leq 2\delta$. Let $x_1$ (resp. $x_2$) denote the point on $[x'_1, x'_2]$ exactly $4\delta$ away from $x'_1$ (resp. $x'_2$). Put $\gamma_1 := [x, x_1]$, $\gamma_2 := [x_1, x_2]$ and $\gamma_3 := [x_2, hx]$. Note that $d_Y(x_1, q) \leq 10\delta$ for $i = 1, 2$, so we put $D_1 := 10\delta$. By $\delta$-hyperbolicity, except for the $3\delta$ neighborhood of $hq$, points on $[q, hq]$ is in the $\delta$ neighborhood of $\alpha$. Hence by putting $D_2 := 3\delta$, we have $\gamma_2 \subset N_{D_2}(\alpha)$. Let $q'_1, q'_2$ be nearest point projections of $x_1, x_2$ to $[q, hq]$ respectively. Then $d_Y(q, q'_1) \leq d_Y(q, q_1) + d_Y(q_1, x'_1) + d_Y(x'_1, x_1) + 2\delta \leq 12\delta$. By symmetry we have $d_Y(q'_2, hq) \leq 12\delta$. By triangle inequality, we have we have $\text{length}(\gamma_2) \geq d_Y(q'_1, q'_2) - d_Y(x_1, q'_1) - d_Y(x_2, q'_2) \geq d_Y(q, hq) - 28\delta$. Thus we have a required decomposition. \qed

We are now in a position to prove Theorem 1.1.

Proof of Theorem 1.1. Suppose $\omega_n = \phi^k$ for some $\phi \in G$ and $k > 1$. Let $\eta$ be a geodesic axis of $\phi$, and $\gamma = [\omega_0, \omega_n]$. By Lemma 3.6 $\gamma_2$ of the decomposition of $\gamma = \gamma_1\gamma_2\gamma_3$ from Proposition 3.7 has length at least $Ln$ for some $L > 0$ with exponentially high probability. Let $L' := \text{length}(\gamma_2)$. Then by applying Lemma 5.2 for small enough $\epsilon$, say $1/100$, we may find $D' > 0$ such that $\text{length}(\gamma_3')/\text{length}(\gamma) \geq 1 - \epsilon$ with exponentially high probability. Then we can find a $D'$-proximal point $q_0 \in \gamma_2$ such that $d_c(q_0, \gamma_2(0)) \leq L'\epsilon$. Let $a$ denote the index that $d_c(\omega_a, q_0) \leq D'$. Similarly we may find a point $q_b \in \gamma_2$ such that

- $L' \leq d_c(q_a, q_b) \leq L' + L'\epsilon$,
- $q_b$ is $D'$-proximal so that $d_c(\omega_a, q_b) \leq D'$ for $a < b$. 


We consider translating \([q_a, q_b] \subset \gamma_2\) by \(\varphi := \phi^{[k/2]}\) where \([k/2]\) is the largest integer among all integers smaller than \(k/2\). Note that
\[
\frac{\tau(\omega_n)}{2} \leq \tau(\varphi) \leq \frac{\tau(\omega_n)}{3}.
\]

By perturbing at most \(L'\) if necessary, we may assume that both \(\varphi(\omega_a)\) and \(\varphi(\omega_b)\) are within at most \(2D_2 + 2\delta\) distance from \(D'\)-proximal points \(q_c, q_d \in \gamma_2\) respectively. The constant \(D_2\) is from Proposition \([5,7]\). Hence there exist indices \(c, d\) with \(a < b < c < d\) such that \(d_C(\omega_i, q_i) \leq D' + 2D_2 + 2\delta\) for \(i \in \{a, b, c, d\}\). Let \(\alpha := [\omega_a, \omega_b]\) and \(\beta := [\omega_c, \omega_d]\). By \(\delta\)-hyperbolicity, we may decompose \(\alpha = \alpha_1\alpha_2\alpha_3\) so that \(\text{length}(\alpha_1), \text{length}(\alpha_3) < D' + 2D_2 + 4\delta\) and \(\alpha_2 \subset \mathcal{N}_{2\delta}(\gamma)\). Hence if \(n\) is large enough, then for some \(t \in [0.1, 0.2]\) we can find a \(D'\)-proximal point \(q_{a'} \in \gamma_2\) with \(d_C(q_{a'}, \alpha(t)) \leq 2\delta\). Similarly, we can also find a \(D'\)-proximal point \(q_{c'}\) such that \(d_C(q_{c'}, \beta(t)) \leq 2\delta\). Thus we have indices \(a'\) and \(c'\) such that \(d_C(\omega_{a'}, \alpha(t)) \leq D' + 2\delta\) and \(d_C(\omega_{c'}, \beta(t)) < D' + 2\delta\). Thus if \(\omega_n\) is not primitive we may find indices satisfying conditions (1)-(3) of Lemma \([3,4]\) for \(M = L/4\) and \(D = D' + 2D_2 + 2\delta\). Therefore the probability that \(\omega_n\) is not primitive decays exponentially. \(\square\)

4. Random Mapping Classes are not Lifts

Let \(\mu\) be a probability measure on the mapping class group \(G\) of surface \(S\) of finite type. In this section, we suppose that \(\mu\) satisfies Condition 1.2. A measure \(\nu\) on \(\mathcal{PMF}(S)\) is called \(\mu\)-stationary if
\[
\nu(X) = \sum_{g \in G} \mu(g) \nu(g^{-1}X)
\]
for any measurable subset \(X \subset \mathcal{PMF}(S)\). We first recall the work of Kaimanovich-Masur. Recall that a projective measured foliation is said to be uniquely ergodic if its supporting foliation admits only one transverse measure up to scale. We denote by \(\mathcal{UE}(S) \subset \mathcal{PMF}(S)\) the space of uniquely ergodic foliations with unique projective measures.

**Theorem 4.1** (\([13\text{ Theorem 2.2.4(1)]}\)). There exists a unique \(\mu\)-stationary probability measure \(\nu\) on \(\mathcal{PMF}(S)\). The measure \(\nu\) is non-atomic and concentrated on the uniquely ergodic foliations \(\mathcal{UE}(S)\).

Then in the following subsection, we will measure by \(\nu\) the lifts of projective measured foliations.

4.1. Lifts of Projective Measured Foliations Have \(\mu\)-Stationary Measure Zero. We first recall the Teichmüller space of \(S\). The Teichmüller space \(\mathcal{T}(S)\) is the space of conformal structures on \(S\). In this paper we consider the Teichmüller metric on \(\mathcal{T}(S)\):
\[
d_T(X, Y) = \frac{1}{2} \log \inf_h K(h), \quad (X, Y \in \mathcal{T}(S)),
\]
where the infimum is taken over all quasi-conformal maps \(h : X \to Y\) homotopic to the identity, and \(K(h)\) is the maximal dilatation of \(h\). Thurston (c.f.\([7]\)) showed that \(\mathcal{PMF}(S)\) compactifies \(\mathcal{T}(S)\) so that the action of \(G := \text{Mod}(S)\) extends continuously. This compactification is called the Thurston compactification. Let \(\overline{\mathcal{T}}(S) := \mathcal{T}(S) \cup \mathcal{PMF}(S)\).
Let $\pi : S \to S'$ be a (possibly orbifold) covering. If $S'$ is an orbifold, $\mathcal{PMF}(S')$ and $\mathcal{T}(S')$ are defined to be the ones on the surface that we get by puncturing the orbifold points of $S'$. The covering $\pi$ determines $\Pi : \bar{\mathcal{T}}(S') \to \bar{\mathcal{T}}(S)$ so that $X \in \Pi(\mathcal{T}(S'))$ if $\pi(X) \in \mathcal{T}(S)$, and $\lambda \in \Pi(\mathcal{PMF}(S'))$ if $\pi(\lambda) \in \mathcal{PMF}(S')$. As pointed out in [22] Section 7, $\Pi$ is an isometric embedding of $\mathcal{T}(S')$. We may also extend the $\mu$-stationary measure $\nu$ in Theorem 4.1 to $\bar{\mathcal{T}}(S)$ by $\nu(A) = \nu(A \cap \mathcal{PMF}(S))$ for each subset $A \subset \bar{\mathcal{T}}(S)$. Our goal in this subsection is the following lemma.

Lemma 4.2. Let $\mu$ be a probability measure on $G$ which satisfies Condition 1.2 and $\nu$ the $\mu$-stationary measure on $\bar{\mathcal{T}}(S)$ from Theorem 4.1. Then for any finite covering $\pi : S \to S'$, we have for all $g \in G$,

$$\nu(g\Pi(\bar{\mathcal{T}}(S'))) = 0.$$

Recall that $\mathcal{PMF}(S)$ is homeomorphic to the sphere $\mathbb{S}^{6g-7+2n}$. Although the image $\Pi(\mathcal{PMF}(S'))$ is a sphere of lower dimension, Lemma 4.2 is non-trivial. This is because by the work of Gadre [8], the $\mu$-stationary measure $\nu$ is singular to the standard Lebesgue measure on the sphere.

First, we give a sufficient condition for a subset of $\mathcal{PMF}(S)$ to have $\nu$ measure zero.

Proposition 4.3 (c.f. [13] Lemma 2.2). If $E$ is a subset of $\mathcal{PMF}(S)$ and there exist infinitely many distinct translations by elements in $\text{sgr}(\mu)$. Suppose further that

$$(*) \quad \nu(g_1E \cap g_2E) = 0 \text{ or } \nu(g_1E) = \nu(g_2E) \text{ for all } g_1, g_2 \in G.$$  

Then $\nu(E) = 0$.

Proof. By $(*)$, we see that there is some $h \in G$ such that $E' := hE$ satisfies $\nu(E') \geq \nu(E)$ for all $g \in G$. Then since $\nu$ is $\mu$-stationary, we have

$$\nu(E') = \sum_{g \in G} \mu(g)\nu(g^{-1}E') \leq \sum_{g \in G} \mu(g)\nu(E') = \nu(E').$$

Thus we see that $\nu(gE') = \nu(E')$ for every $g$ in the support of $\mu$. By discussing the $n$-convolution $\mu^n$ of $\mu$, we see that $\nu(gE') = \nu(E')$ for every $g \in \text{sgr}(\mu)$. Since we have infinitely many distinct translates of $E'$ by elements of $\text{sgr}(\mu)$, we see that $\nu(E') = \nu(E) = 0$. \qed

To prove Lemma 4.2 we recall Teichmüller geodesics on the Teichmüller space, see for example [7, 9, 13] for more details. Recall that $S$ is a surface of finite type $(g, n)$. Let $X \in \mathcal{T}(S)$, then $X$ can be written as $\mathbb{H}/\Gamma$ for some $\Gamma \subset \text{PSL}(2, \mathbb{R})$, where $\mathbb{H}$ is the upper half subspace of $\mathbb{C}$. Then a holomorphic quadratic differential is a holomorphic map $\varphi : X \to \mathbb{C}$ induced from $\tilde{\varphi} : \mathbb{H} \to \mathbb{C}$ satisfying for all $\gamma \in \Gamma$

$$\tilde{\varphi}(\gamma(z))\gamma'(z)^2 = \tilde{\varphi}(z).$$

Teichmüller showed that for any given point $X \in \mathcal{T}(S)$, a holomorphic quadratic differential determines a geodesic with respect to Teichmüller metric. For $X \in \mathcal{T}(S)$, let $\text{QD}(X)$ denote the Banach space of holomorphic quadratic differentials on $X$ with $\|\varphi\| := \int_X |\varphi|$. Each $\varphi \in \text{QD}(X)$ determines two measured foliation, called the horizontal foliation and the vertical foliation. By Riemann-Roch theorem, $\text{QD}(X)$ has complex dimension $3g - 3 + n$. Let $Q_0 \subset \text{QD}(X)$ denote the unit sphere. This $Q_0$ compactifies $T(X)$ which is called the Teichmüller compactification.

By the work of Hubbard-Masur (compact) and Gardiner (finite type), we see:
Lemma 4.4 ([12], [9 Chapter 11]). For any $X \in \mathcal{T}(S)$ and $F \in \mathcal{PMF}(S)$, there is a unique $\varphi \in \text{QD}(X)$ whose horizontal foliation is $F$ up to scale.

Proof of Lemma 4.4. The proof goes by induction. Let $E := g \Pi(\mathcal{T}(S'))$ and $d$ the complex dimension of $\text{QD}(X')$ for any $X' \in \mathcal{T}(S')$. We consider intersection $E' := g_1E \cap g_2E \cap \cdots \cap g_nE$. We first define $d(E') \in \mathbb{N}$. If $E' \cap \mathcal{PMF}(S)$ contains at most one uniquely ergodic foliation, then we define $d(E') = 0$. In this case we also have $\nu(E) = 0$ since $\nu$ is non-atomic. If $E' \cap \mathcal{PMF}(S)$ contains at least two uniquely ergodic foliations $E_1, E_2$, then there is a unique Teichmüller geodesic $\gamma$ connecting $E_1$ and $E_2$ by [10]. Since covering maps induce isometric embeddings of Teichmüller spaces [22 Section 7], any point of $\gamma$ is in $E'$. In particular $E' \cap T(S)$ is non-empty. For any $X \in E' \cap T(S)$, each $g_iE$ determines a subspace of $S_i(X) \subset \text{QD}(X)$ which consists of the lifts of holomorphic quadratic differentials with respect to the covering $\pi \circ g_i^{-1}$. Let $S(X) := \cap_{i=1,...,n} S_i(X)$ and $d(X) := \dim S(X)$. Since $d(X) \in \mathbb{N}$, there exists $X' \in E' \cap T(S)$ such that $d(X') \geq d(X)$ for any $X \in E' \cap T(S)$. We define $d(E') := d(X')$. Then we explain how the induction works by using a style of inductive algorithm, see Algorithm 1 which is named MVIT. By MVIT($E, d$), we have $\nu(E) = 0$. Note that although the depth of Algorithm 1 is finite, the width is infinite.

Algorithm 1 MVIT(Measure by $\nu$ the Intersection of Translates)

Input: $(E' := g_1E \cap g_2E \cap \cdots \cap g_nE, d(E'))$

Ensure: $\nu(g_1E \cap g_2E \cap \cdots \cap g_nE) = 0$.

if $d(E') = 0$ then
    By the definition of $d(E')$ and Lemma 4.4 we have $\nu(E') = 0$.
end if

for $h_1, h_2 \in G$ do
    Let $E'_1 := h_1E'$ and $E'_2 := h_2E'$. Note that since each $g \in G$ induces a vector isomorphism between $\text{QD}(X)$ and $\text{QD}(gX)$, we have $d(E'_1) = d(E'_2) = d(E')$.
    if $d(E'_1 \cap E'_2) = d(E')$ then
        We see that $\nu(E'_1) = \nu(E'_2) = \nu(E')$ by Lemma 4.4.
    else
        In this case we have $d(E') > d(E'_1 \cap E'_2)$. Then we apply
        MVIT($E'_1 \cap E'_2, d(E'_1 \cap E'_2)$), which proves $\nu(E'_1 \cap E'_2) = 0$.
    end if
end for

We have seen that the condition $(*)$ of Proposition 4.3 is satisfied. By Condition 1.2 and the north-south dynamics of pseudo-Anosov maps (see [25]), we see that there are infinitely many translates of $E'$. Thus by Proposition 4.3 we have $\nu(E') = 0$.

$\square$

4.2. Proof of Theorem 1.3 We turn to consider the curve graph $\mathcal{C}(S)$ again. By the work of Klarreich [14] (see also Hamenstädt [11]), the Gromov boundary $\partial \mathcal{C}(S)$ of $\mathcal{C}(S)$ is identified with the space $\mathcal{F}_{\text{min}}(S)$ of minimal foliations. There is a natural measure forgetting map from $\mathcal{U} \mathcal{E}(S)$ to $\mathcal{F}_{\text{min}}(S)$. Hence we may consider the push forward of $\nu$ to $\mathcal{F}_{\text{min}}(S)$, which we again write as $\nu$ by abuse of notation. This $\nu$ extends to $\bar{\mathcal{C}} := \mathcal{C}(S) \cup \partial \mathcal{C}(S)$ by $\nu(A) = \nu(A \cap \partial \mathcal{C}(S))$ for $A \subset \bar{\mathcal{C}}(S)$. 
We fix a (possibly orbifold) covering \( \pi : S \to S' \). Here if \( S' \) is an orbifold, we define \( C(S') \) as the curve graph of the surface that we get by puncturing every orbifold point of \( S' \). We may define one to finite relation \( \Pi : C(S') \to C(S) \) as follows. A curve \( b \in C(S) \) is in \( \Pi(a) \) for some \( a \in C(S') \) if \( \pi(a) = b \) as isotopy classes of simple closed curves. In [22], Rafi-Schleimer showed that \( \Pi \) is quasi-isometric. Hence the map \( \Pi \) extends continuously to the Gromov boundary \( \partial C(S') \). This implies that \( \Pi(\bar{C}(S')) \) has the same \( \nu \) measure as \( \Pi(\bar{T}(S')) \). Let

\[
F := \bigcup_{g \in G} g(\Pi(\bar{C}(S'))).
\]

By Lemma 4.2, we have \( \nu(F) = 0 \) since \( G \) is a countable group.

For a subset \( A \subset \bar{C}(S) \), we define the shadow \( S_p(A, r) \) for \( r > 0 \) and \( p \in \bar{C}(S) \) by

\[
S_p(A, r) := \bigcup_{a \in A} S_p(a, r).
\]

We first prove the following lemma, which is a key step for showing Theorem 1.3.

**Lemma 4.5** (c.f. [18, Lemma 2.10]). There is a constants \( K > 0 \) and \( c < 1 \), such that for any \( r \),

\[
\nu(S_1(F, r)) < c^r, \quad P(\omega_n \in (S_1(F, r))) < Kc^r,
\]

and the constants \( K \) and \( c \) depend on \( \mu \) and \( \delta \), but not on \( r \) and \( n \).

We prove Lemma 4.5 by borrowing several arguments from the proof of [18, Lemma 2.10]. In [18], Maher uses several lemmas from [5], which are applications of Lemma 4.6 below. Instead of using those lemmas, we only use Lemma 4.6 since the proof of each lemma in [5] that we need is short and elementary.

**Lemma 4.6** (see for example [3, Proposition 6.7]). Let \( (X, d_X) \) be a \( \delta \)-hyperbolic space. Then there is a constant \( K_1 \) which depends only on \( \delta \) with the following property. For any four points \( x_1, x_2, x_3, x_4 \in X \), there is an embedded tree \( T \) connecting the four point such that

\[
(1a) \quad d_T(x_i, x_j) \leq d_X(x, y) - K_1
\]

\[
(1b) \quad (x_i \cdot x_j)_{x_k} - 2K_1 \leq (x_i \cdot x_j)_x \leq (x_i \cdot x_j)_{x_k} + K_1
\]

for \( 1 \leq i, j \leq 4 \). Where \( d_T \) denotes the distance in \( T \), and for \( a, b, c \in T, (a, b)_c^T \) denotes the Gromov product with respect to \( d_T \).

Note that the only combinatorial type of the tree up to reindexing is as depicted in Figure 1.

![Figure 1. Approximate tree.](image)

We will use the following lemma in [18].

**Lemma 4.7** ([18, Proposition 2.12].) For any \( \epsilon > 0 \), there is a constant \( K_2 \) which depends on \( \epsilon \) and \( \mu \), such that if \( r \geq K_2 \), then \( \nu(S_1(x, r)) < \epsilon \).

For the proof of Lemma 4.5, we also prepare the following lemma.
Lemma 4.8. For any $\epsilon > 0$, there is a constant $K_1$ which depends on $\epsilon$ and $\mu$, such that if $r \geq K_1$, then $\nu(S_1(F, r)) < \epsilon$.

Proof. Note that if $r > s$, $S_1(F, r) \subset S_1(F, s)$. Recall that $\nu(F) = 0$ by Lemma 4.9. Since
\[
\bigcap_{r \in \mathbb{R}^+} S_1(F, r) = F \cap \partial \mathcal{C}(S),
\]
by the monotone convergence theorem
\[
\lim_{r \to \infty} \nu(S_1(F, r)) = 0.
\]
\[\square\]

We recall the following lemma due to Maher.

Lemma 4.9 (LS Lemma 2.11). Let $X_0 \subset X_1 \subset X_2 \subset \ldots$ be a sequence of nested closed subsets of $\mathcal{C}(S)$ with the following properties:

1. $1 \not\in X_0$
2. $(\mathcal{C}(S) \setminus X_n) \cap X_{n+1} = \emptyset$
3. $d_{\mathcal{C}}(\mathcal{C}(S) \setminus X_n, X_{n+1}) \leq D$

Furthermore, suppose there is a constant $0 < \epsilon < 1$ such that, for any $x \in X_n \setminus X_{n+1}$ which is the translate of the base point $p$ by $x \in G$,

4. $\nu_x(X_{n+2}) \leq \epsilon$
5. $\nu_x(\mathcal{C}(S) \setminus X_{n-1}) \leq \epsilon$

where $\nu_x(A) := \nu(x^{-1}A)$ for any $A \subset \mathcal{C}(S)$. Then there are constants $c < 1$ and $K$, which depend only on $\epsilon$ and $\mu$, such that $\nu(X_n) < c^n$ and $\mu_n(X_n) < K c^n$.

Then, to prove Lemma 4.5, it suffices to prove

Lemma 4.10. For any $D > 0$, there exists $L$ which depend on $\mu$, $\delta$ and $D$ with the following property. The sets $X_n := S_1(F, L(n+1))$ for all $n \in \mathbb{N}$ form a sequence of nested sets which satisfies (1)-(5) in Lemma 4.9.

Proof. The proof goes in a similar way to LS Lemma 2.13. We use the constants $K_1, \ldots, K_3$ from Lemma 4.6. Let $L := 4K_1 + \max\{D, K_2, K_3, 2\delta\}$.

1. The Gromov product $(1 \cdot a)_1 = 0$ for all $a \in \mathcal{C}(S)$. For all $y \in X_0$, there is $e_y \in F$ such that $(e_y \cdot y)_1 \geq L > 0$, hence $1 \not\in X_0$.
2. If $y_i \to y \in \partial G$, then by the property of the Gromov product (see for example LS IIIH 3.17(5)), $\liminf (x \cdot y_i)_1 \geq (x \cdot y)_1 - 2\delta$. This implies if $y \in X_{n+1}$, then for any sequence $y_i \to y$, all but finitely many $y_i$’s are in $X_n = S_1(F, L(n+1))$ since $L > 2\delta$. Thus we have $X_{n+1} \cap (\mathcal{C}(S) \setminus X_n) = \emptyset$.
3. Let $a \in X_{n+1}$, then there exists $e_a \in F$ such that $a \in S_1(e_a, L(n+2))$. Let $b \in \mathcal{C}(S) \setminus X_n$, then for all $e \in F$, we have $b \not\in S_1(e, L(n+1))$. In particular $b \not\in S_1(e_a, L(n+1))$. Then we consider a tree $T_1$ from Lemma 4.6 that connects $\{1, b, a, e_a\}$. Since $(a \cdot e_a)_1 \geq L(n+2)$ and $(b \cdot e_a)_1 < L(n+1)$,

\[\hspace{1cm}\]

\[\text{A correction for the proof in LS can be found in Maher’s webpage.}\]

\[\text{http://www.math.csi.cuny.edu/maher/research}\]
Proof of Theorem 1.3. We first fix a (possibly orbifold) covering \( \pi : S \to S' \). Since every argument in this section is independent of the choice of the base point \( p \), we may suppose that \( p \in \Pi(\hat{C}(S')) \). Then if \( \omega_n \) is a lift with respect to covering map \( \pi \circ g \) for some \( g \in G \), we have \( \omega_n p \in F \). By Lemma 4.6 we may suppose for some \( L' > 0 \), \( \tau(\omega_n) \geq L'n \) with exponentially high probability. This implies that if \( \omega_n \in F \), then \( \omega_n \) must be in \( X_{L'/L'n-1} \), where \( X_i := S_i(F, L(i + 1)) \) as in Lemma 4.10. Therefore by Lemma 4.10 we have \[
P(\omega_n \in F) \leq P(\omega_n \in X_{L'/L'n-1}) + P(d_C(\omega_n, \omega_n) < L'n) \leq Ke^n \]
for some \( K > 0 \) and \( c < 1 \). Thus for a fixed covering map \( \pi : S \to S' \), we have proved that the probability that a random walk is a lift of some covering that corresponds to \( \pi_1(S) < \pi_1(S') \) decays exponentially.

The number of possible types of orbifolds which may be covered by \( S \) is finite. Furthermore, for each such an orbifold, there are only finitely many possible covering maps up to conjugacy. This is because the number of subgroups of bounded index in a finitely generated group is finite. Hence the argument above suffices to complete the proof. \( \square \)
5. Non-compact random mapping tori are non-arithmetic

First, we recall the definition of non-compact arithmetic 3-manifold, see [16] for more details and properties of arithmetic 3-manifolds. Let \( d \) be a positive square-free integer and \( \mathcal{O}_d \) denote the ring of integers of \( \mathbb{Q}(\sqrt{-d}) \). A Bianchi group is a subgroup of \( \text{PSL}(2, \mathbb{C}) \) which is of the form \( \text{PSL}(2, \mathcal{O}_d) \). One can show that every Bianchi group is a lattice. The quotient \( \mathbb{H}^3 / \text{PSL}(2, \mathcal{O}_d) \) is called a Bianchi orbifold, where \( \mathbb{H}^3 \) is the hyperbolic 3-space. A non-compact hyperbolic 3-manifold \( M = \mathbb{H}^3 / \Gamma \) of finite volume is arithmetic if a conjugate of \( \Gamma \) in \( \text{PSL}(2, \mathbb{C}) \) is commensurable to some Bianchi group \( \text{PSL}(2, \mathcal{O}_d) \). Recall that two subgroups of \( \text{PSL}(2, \mathbb{C}) \) are said to be commensurable if their intersection is a finite index subgroup in both. Let \( S \) be an orientable surface of finite type with at least one puncture. For \( \phi \in \text{Mod}(S) \), the mapping torus \( M(S, \phi) \) is defined by

\[
M(S, \phi) = S \times [0, 1]/(x, 1) \sim (\phi(x), 0).
\]

Two mapping tori \( M(S, \phi_1) \) and \( M(S, \phi_2) \) are said to be cyclic commensurable if there exists \( k_1, k_2 \in \mathbb{Z} \setminus \{0\} \) such that \( M(S, \phi_1^{k_1}) = M(S, \phi_2^{k_2}) \). Bowditch-Maclachlan-Reid proved the following theorem.

**Theorem 5.1** ([4, Theorem 4.2]). Let \( S \) be an orientable surface of finite type with at least one puncture. There are at most finitely many cyclic commensurability classes of arithmetic mapping tori with fiber \( S \).

Now we are in a position to prove Theorem 1.5.

**Proof of Theorem 1.5**. Note that if two mapping classes give rise to cyclic commensurable mapping tori, then they are fibred commensurable. By Theorem 1.4 it suffices to discuss minimal mapping classes in their commensurability classes. The uniqueness of the minimal element (Theorem 2.3) implies that two minimal mapping classes give rise to cyclic commensurable mapping tori if and only if they are conjugate. Hence there are at most finitely many conjugacy classes of minimal elements that give arithmetic mapping tori by Theorem 5.1. Hence there is an upper bound of the translation length for minimal mapping classes to have arithmetic mapping tori. Then Lemma 3.6 applies to complete the proof.

**Remark 5.2.** For \( S \) closed, one can prove similar statement as Theorem 5.1 with upper bound for the degree of the invariant trace fields, see [4, Corollary 4.4.]. For \( S \) closed, we do not know if the set of a random mapping classes with arithmetic mapping tori is exponentially small or not.

**Acknowledgments**

The author would like to thank Joseph Maher and Giulio Tiozzo for helpful conversations and sharing the preprint [19]. He started this work when he was in ICERM, Brown University. Thanks also goes to ICERM and JSPS for supporting the visit. He would also like to thank Brian Bowditch for bringing the paper [4] in his attention.

**References**

[1] M. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Vol. 319. Springer, 1999.

[2] B. Bowditch, *Tight geodesics in the curve complex*, Invent. Math. 171(2) (2008), 281-300.
[3] B. Bowditch, A course on geometric group theory, MSJ Memoirs 16, 2006.

[4] B. Bowditch, C. Maclachlan and A. Reid, Arithmetic hyperbolic surface bundles. Math. Ann. 302 (1995), no. 1, 31-60.

[5] D. Calegari and J. Maher, Statistics and compression of scl, Ergodic Theory and Dynamical Systems (2010): 1-47. [arXiv:1008.4952]

[6] D. Calegari, H. Sun and S. Wang, On fibered commensurability. Pacific J. Math. 250 (2011), no. 2, 287-317.

[7] B. Farb and D. Margalit, A Primer on Mapping Class Groups (PMS-49), Princeton University Press, 2011.

[8] V. Gadre, Harmonic measures for distributions with finite support on the mapping class group are singular, to appear in Duke Math. J. [arXiv:0911.2891]

[9] F. Gardiner, Teichm"uller theory and quadratic differentials, Wiley-Interscience, 1987.

[10] F. Gardiner and H. Masur, Extremal length geometry of Teichm"uller space, Complex Variables and Elliptic Equations 16.2-3 (1991): 209-237.

[11] U. Hamenst"adt, Train tracks and the Gromov boundary of the complex of curves, Spaces of Kleinian groups, London Math. Soc. Lecture Note Ser., vol. 329, Cambridge Univ. Press, Cambridge, 2006, pp. 187-207.

[12] J. Hubbard and H. Masur, Quadratic differentials and foliations, Acta. Math. 142 (1979), 221-274.

[13] V. Kaimanovich and H. Masur, The Poisson boundary of the mapping class group, Invent. Math. 125 (1996) 221-264.

[14] E. Klarreich, The boundary at infinity of the curve complex and the relative Teichmüller space, preprint, available at http://www.ericaklarreich.com/curvecomplex.pdf.

[15] E. Kowalski, The large sieve and its applications, Arithmetic geometry, random walks and discrete groups, Cambridge Tracts in Mathematics 175 (Cambridge University Press, Cambridge, 2008).

[16] C. Maclachlan and A. Reid. The arithmetic of hyperbolic 3-manifolds, Vol. 219. Springer, 2003.

[17] J. Maher, Random walks on the mapping class group, Duke Math. J. 156 (2011), no. 3, 429-468.

[18] J. Maher, Exponential decay in the mapping class group, J. Lond. Math. Soc. (2) 86 (2012), no. 2, 366-386.

[19] J. Maher and G. Tiozzo, Random walks on weakly hyperbolic groups, preprint.

[20] H. Masai, On commensurability of fibrations on a hyperbolic 3-manifold, Pacific J. Math. 266 (2013), No. 2, 313-327.

[21] H. Masur and Y. Minsky, Geometry of the complex of curves. I. Hyperbolicity, Invent. Math., 138(1):103-149, 1999.

[22] K. Rafi and S. Schleimer, Covers and the curve complex, Geom. Topol. 13 (2009), no. 4, 2141-2162.

[23] I. Rivin, Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms, Duke Math. J. 142 (2008) 353-379.

[24] W. Thurston, Hyperbolic Structures on 3-manifolds, II: Surface groups and 3-manifolds which fiber over the circle, preprint. [arXiv:math/9801045]

[25] W. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces, Bulletin of the American mathematical society 19, (1988) 417-431.

[26] W. Woess, Random walks on infinite graphs and groups, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000.