1

LINEAR GROWTH OF TRANSLATION LENGTHS OF RANDOM ISOMETRIES ON GROMOV HYPERBOLIC SPACES AND TEICHMÜLLER SPACES

HYUNGRYUL BAIK [1], INHYEOK CHOI[2] AND DONGRYUL M. KIM[3]

[1]*Department of Mathematical Sciences, KAIST, 291 Daehak-ro Yuseong-gu, Daejeon, 34141, South Korea*
(hrbaik@kaist.ac.kr)

[2]*Department of Mathematical Sciences, KAIST, 291 Daehak-ro Yuseong-gu, Daejeon, 34141, South Korea*
(inhyeokchoi48@gmail.com)

[3]*Department of Mathematics, Yale University, 219 Prospect Street, New Haven, CT 06511, USA*
(dongryul.kim@yale.edu)

Abstract We investigate the translation lengths of group elements that arise in random walks on the isometry groups of Gromov hyperbolic spaces. In particular, without any moment condition, we prove that non-elementary random walks exhibit at least linear growth of translation lengths. As a corollary, almost every random walk on mapping class groups eventually becomes pseudo-Anosov, and almost every random walk on $\mathrm{Out}(F_n)$ eventually becomes fully irreducible. If the underlying measure further has finite first moment, then the growth rate of translation lengths is equal to the drift, the escape rate of the random walk.

We then apply our technique to investigate the random walks induced by the action of mapping class groups on Teichmüller spaces. In particular, we prove the spectral theorem under finite first moment condition, generalizing a result of Dahmani and Horbez.

1. Introduction

Throughout, X denotes either a separable geodesic Gromov hyperbolic space or the Teichmüller space of a closed orientable surface of genus at least two, which is equipped with the Teichmüller metric. G denotes a countable subgroup of $\mathrm{Isom}(X)$ containing

© The Author(s), 2023. Published by Cambridge University Press.

a pair of independent hyperbolic isometries.[1] When X is a Gromov hyperbolic space, following Maher and Tiozzo [MT18], we call G a *weakly hyperbolic group*. μ denotes a non-elementary probability measure on G, i.e., a measure on G such that the semigroup generated by the support of μ contains independent hyperbolic isometries. $\mathbb{P}$ stands for the probability measure for random walks induced by μ. See section 2.3 for details.

Random walks on the isometry group of Gromov hyperbolic spaces or Teichmüller spaces have been studied in depth for several decades. For example, Kaimanovich investigated random walks on hyperbolic groups and semi-simple Lie groups in [Kai00], identifying the Poisson boundary with other natural boundaries such as the Gromov boundary or the Furstenberg boundary, under some moment and entropy conditions on the measure. See also [Led01] for the behavior of random walks on free groups with general conditions. A generalization to relatively hyperbolic groups was considered by Gautero and Mathéus in [GM12]. This was further generalized to weakly hyperbolic groups by Maher and Tiozzo in [MT18].

In the course of characterizing the Poisson boundary, Maher and Tiozzo observed the following phenomenon.

Theorem 1.1 [MT18, Theorem 1.2, 1.4]. *Let ω be the random walk generated by μ. Then there exists a constant $L > 0$ such that*

$$\liminf_{n \to \infty} \frac{d_X(x_0, \omega_n x_0)}{n} \geq L$$

for $\mathbb{P}$-almost every (a.e. in short) sample path (ω_n). Moreover, the translation length $\tau(\omega_n)$ of ω_n grows at least linearly; that is,

$$\mathbb{P}(\tau(\omega_n) \leq L'n) \to 0 \quad as \quad n \to \infty$$

for some constant $L' > 0$.

Maher and Tiozzo proved these results for weakly hyperbolic groups, which implies the results for Teichmüller space thanks to the coarsely Lipschitz systole map from Teichmüller space to the curve complex. One can also refer to the earlier work of Ledrappier [Led01] that proves similar results with considerations on harmonic measures in the case of free groups.

If μ further has finite first moment, Kingman's subadditive ergodic theorem provides a constant λ such that $\lim_{n \to \infty} d_X(x_0, \omega_n x_0)/n = \lambda$ for $\mathbb{P}$-a.e. (ω_n). Here, λ is called the *escape rate* or the *drift* of the random walk. One consequence of Theorem 1.1 is that the drift of a non-elementary random walk on weakly hyperbolic groups is strictly positive.

As presented in Theorem 1.1, due to the lack of subadditivity, translation length is more difficult to investigate than displacement. Hence, although the growth of displacements is given *almost surely*, the growth of translation lengths is given *in probability*. Maher and Tiozzo also proved in [MT18] *almost sure linear growth*; that is,

$$\liminf_{n \to \infty} \tau(\omega_n) \geq L'n \quad \text{for} \quad \mathbb{P}\text{-a.e.}(\omega_n)$$

---

[1] A pair of independent hyperbolic isometries means two hyperbolic isometries (pseudo-Anosovs on the surface when X is the Teichmüller space) with a disjoint set of fixed points on ∂X.

when the support of μ is bounded. This relies on the exponential decay of shadows explained in [Mah12].

Amongst weakly hyperbolic groups are mapping class groups and $\mathrm{Out}(F_n)$ that act on the corresponding curve complexes or the complexes of free factors, respectively. We remark that before the work [MT18] of Maher and Tiozzo, Maher proved in [Mah11] that random element of the mapping class group becomes pseudo-Anosov *in probability*.

Previously known results about the growth of translation lengths rely on moment conditions that require μ to have bounded support, finite exponential moment or finite second moment. One of our purposes is to study the growth of translation lengths without any moment condition. With the finite first moment condition (so that the drift is available), one can describe this growth with greater precision. Our first main theorem is as follows.

Theorem A. *Let G be a weakly hyperbolic group acting on a Gromov hyperbolic space X and let μ be a non-elementary probability measure on G. Then $\mathbb{P}$-almost every sample path shows at least linear growth of translation lengths. More precisely, there exists a constant $\mathcal{L} > 0$ such that for $\mathbb{P}$-a.e. (ω_n), $\tau(\omega_n) \geq \mathcal{L}n$ for sufficiently large n.*

Moreover, if μ further has finite first moment, then for $\mathbb{P}$-a.e. (ω_n), we have

$$\lim_{n\to\infty} \frac{\tau(\omega_n)}{n} = \lambda,$$

where λ is the drift of the random walk.

As a corollary, *almost every* sample path of non-elementary random walks on mapping class groups or $\mathrm{Out}(F_n)$ becomes pseudo-Anosov or fully irreducible, respectively. Here, mapping class groups are acting on the curve complexes, and $\mathrm{Out}(F_n)$ are acting on the complexes of free factors; see [MM99], [MM00], [BF14]. In addition to full irreducibility, another notion that captures the loxodromic property of an outer automorphism is that of atoroidality. Using the action of $\mathrm{Out}(F_n)$ on hyperbolic spaces such as Dowdall-Taylor's co-surface graph [DT17] or Brian Mann's intersection graph [Man14], one can even deduce the genericity of fully irreducible atoroidal outer automorphisms in $\mathrm{Out}(F_n)$.

Corollary 1.2 (Eventually pseudo-Anosov behavior). *Let S be a closed orientable surface of genus at least 2 and $\mathrm{Mod}(S)$ be its mapping class group. Let μ be a non-elementary measure on $\mathrm{Mod}(S)$. Then for $\mathbb{P}$-a.e. sample path (ω_n) of the random walk generated by μ, there exists $N > 0$ such that ω_n is pseudo-Anosov for all $n > N$.*

Corollary 1.3 (Spectral theorem for $\mathrm{Mod}(S)$ on the curve complex). *Let S be a closed orientable surface of genus at least 2 and $\mathrm{Mod}(S)$ be its mapping class group. Let μ be a non-elementary measure on $\mathrm{Mod}(S)$ having finite first moment on the curve complex and let $\lambda > 0$ be its drift. Then for $\mathbb{P}$-a.e. (ω_n), we have*

$$\lim_{n\to\infty} \frac{1}{n}\tau(\omega_n) = \lambda.$$

We record a relevant work [EST20] regarding the genericity of pseudo-Anosov elements in $\mathrm{Mod}(S)$.

Recently, there has been a remarkable progress on the study of the large deviation principle for random walks. In [BMSS22], Boulanger, Mathieu, Sert and Sisto showed the large deviation property of the displacement and translation length of non-elementary random walks on weakly hyperbolic groups under finite exponential moment condition. Also investigated are the properties of the deviation function. Their result is related to the large deviation principle of spectral radii of some random matrix products; see [AS21] for this perspective.

Recently, Gouëzel proved the exponential error bounds for non-elementary random walks on Gromov hyperbolic spaces without moment condition in [Gou22]. His strategy shares the philosophy of pivoting described in the present article, but his aim differs from ours. A similar technique appears in [HMM18], where the authors aim to deal with surface groups and central limit theorems.

Meanwhile, mapping class groups also act on Teichmüller spaces equipped with the Teichmüller metric, which are not Gromov hyperbolic in general ([MW95], [Len08], [Mas75], [MP89], [Iva02]). Thus, random walks on Teichmüller spaces are of independent interest. Kaimanovich and Masur [KM96] proposed a way to investigate random walks on Teichmüller spaces and related them with their Poisson boundary. Meanwhile, Dahmani and Horbez proved in [DH18] the spectral theorem for mapping class groups with respect to the Teichmüller metric under finite second moment condition. Their strategy is to lift the deviation of random paths on the curve complex to Teichmüller setting. In [BCK22], the authors applied these techniques to establish the spectral theorems for free subgroups of $\mathrm{Mod}(S)$ generated by two multitwists under finite second moment condition.

Our next main theorem generalizes the result of Dahmani and Horbez in [DH18].

Theorem B. *Let S be a closed orientable surface of genus at least 2. Let $G = \mathrm{Mod}(S)$ be its mapping class group, $X = \mathcal{T}(S)$ be its Teichmüller space equipped with the Teichmüller metric, and μ be a non-elementary probability measure on $\mathrm{Mod}(S)$. Then $\mathbb{P}$-a.e. sample path shows at least linear growth of translation lengths. More precisely, there exists a constant $\mathcal{L} > 0$ such that for $\mathbb{P}$-a.e. (ω_n), $\tau(\omega_n) \geq \mathcal{L}n$ for sufficiently large n.*

Moreover, if μ further has finite first moment with respect to the Teichmüller metric, then for $\mathbb{P}$-a.e. (ω_n), we have

$$\lim_{n \to \infty} \frac{\tau(\omega_n)}{n} = \lambda,$$

where $\lambda > 0$ is the drift.

We remark that Theorem B implies the spectral theorem for Teichmüller spaces equipped with Thurston's asymmetric Lipschitz metric. This is due to the result of Choi and Rafi in [CR07]: among the (marked) surfaces of the same injectivity radius, the distances with respect to the Teichmüller metric and the Thurston metric differ by a uniformly bounded amount.

Our methods are influenced by [MT18], where Maher and Tiozzo defined the notion of persistent joints that records the permanent depart of the random walk from the origin. We also make use of the boundary convergence of the random walk and the non-atomness of the limiting measure, which are established in [MT18]. Another approach to

study the deviation of random walks was suggested by [MS20]. These traditional methods require finite second moment condition to deduce a deviation inequality that leads to the summable decay of shadows. A related result is the central limit theorem on hyperbolic groups established by Benoist and Quint in [BQ16].

In contrast, apart from the boundary convergence, pivoting and probability estimation appearing in our methods rely purely on the elementary properties of Gromov products and fellow-traveling geodesics. As long as pivots are present in the sample path, the pivoting method works regardless of deviation from the escape rate. (See [Gou22] for a similar idea.) In order to implement this idea on Teichmüller spaces, we utilize the fellow-traveling phenomena of certain Teichmüller geodesics due to Rafi [Raf14]. We remark that Duchin also investigated "thin triangles" in Teichmüller spaces in [Duc05] to study the dynamics of random walks on Teichmüller spaces.

2. Preliminaries

2.1. Geometry of a Gromov hyperbolic space

Let (X,d) be a metric space. Throughout, we fix a basepoint x_0 in X. The following notion is crucial to defining Gromov hyperbolicity.

Definition 2.1 (Gromov product). For $x,y,z \in X$, the *Gromov product* $(x,y)_z$ is defined by

$$(x,y)_z := \frac{1}{2} \left[d(x,z) + d(y,z) - d(x,y) \right].$$

Definition 2.2 (Gromov hyperbolic space). A metric space (X,d) is said to be *Gromov hyperbolic* if it satisfies the following property for some $\delta > 0$:

Property 2.3. For any $x,y,z,w \in X$, we have

$$(x,y)_w \geq \min\{(x,z)_w, (y,z)_w\} - \delta.$$

For details on the properties of Gromov hyperbolic spaces, see [Gro87] and [BH99]. In sections 2 and 3, we assume that (X,d) is a separable, geodesic, Gromov hyperbolic space.

We now consider the *Gromov boundary* ∂X of X in terms of the Gromov product. We regard that a sequence $(x_n)_{n \in \mathbb{N}}$ in X is converging to a point at infinity if $(x_n, x_m)_{x_0} \to \infty$ as $\min\{m,n\} \to \infty$. Furthermore, two such sequences $(x_n), (y_n)$ will be considered as converging to the same point at infinity if $(x_n, y_n)_{x_0} \to \infty$ as $n \to \infty$. In this case, we regard (x_n) and (y_n) to be equivalent and denote $(x_n) \sim (y_n)$. In this point of view, the *Gromov boundary* is defined as follows:

$$\partial X := \left\{ (x_n) \in X^{\mathbb{N}} : \lim_{\min\{m,n\} \to \infty} (x_n, x_m)_{x_0} \to \infty \right\} / \sim .$$

The Gromov product defined above can be extended to the Gromov boundary by setting

$$(x,y)_{x_0} := \sup_{(x_n)=x, (y_n)=y} \liminf_{m,n \to \infty} (x_m, y_n)_{x_0}.$$

One can interpret the Gromov product $(x,y)_z$ as a crude distance from z to the geodesic connecting x and y. Having this in mind, we have the following useful lemma.

Lemma 2.4. *Let $n \geq 1$ and $x_0, \ldots, x_n$ be points in X. Suppose that*

$$(x_{i-1}, x_{i+1})_{x_i} + (x_i, x_{i+2})_{x_{i+1}} < d(x_i, x_{i+1}) - 3\delta \tag{2.1}$$

for $i = 1, \ldots, n-2$. Then

(1) $|(x_i, x_k)_{x_j} - (x_{j-1}, x_{j+1})_{x_j}| \leq 2\delta$ *for $0 \leq i < j < k \leq n$, and*

(2)

$$\left| \left(\sum_{i=0}^{n-1} d(x_i, x_{i+1}) - 2 \sum_{i=1}^{n-1} (x_{i-1}, x_{i+1})_{x_i} \right) - d(x_0, x_n) \right| \leq 2(n-1)\delta.$$

Proof. We prove (1) and (2) by induction on n. For $n = 1$, (1) is void and (2) holds automatically.

Let us now assume (1) and (2) for $n = m \geq 1$ and prove them for $n = m+1$. We claim that

$$|(x_i, x_k)_{x_j} - (x_{j-1}, x_k)_{x_j}| \leq \delta, \tag{2.2}$$

$$|(x_i, x_k)_{x_j} - (x_i, x_{j+1})_{x_j}| \leq \delta \tag{2.3}$$

for $0 \leq i < j < k \leq m+1$ concludes (1) for $n = m+1$. Indeed, given a triple (i,j,k), Inequality 2.2 and Inequality 2.3 (with i replaced with $j-1$) lead to the inequality in (1).

The nontrivial case for Inequality 2.2 is when $i < j-1$. Since $0 \leq i < j-1 < j \leq m$ in this case, Inequality 2.2 for $n = m$ implies

$$|(x_i, x_j)_{x_{j-1}} - (x_{j-2}, x_j)_{x_{j-1}}| \leq \delta. \tag{2.4}$$

Moreover, since $1 \leq j-1 < j < k \leq m+1$, Inequality 2.3 for $n = m$ implies

$$|(x_{j-1}, x_k)_{x_j} - (x_{j-1}, x_{j+1})_{x_j}| \leq \delta. \tag{2.5}$$

Combining them, we have

$$
\begin{aligned}
(x_i, x_{j-1})_{x_j} &= d(x_{j-1}, x_j) - (x_i, x_j)_{x_{j-1}} \\
&\geq d(x_{j-1}, x_j) - (x_{j-2}, x_j)_{x_{j-1}} - \delta &&(\because \text{Inequality } 2.4) \\
&> (x_{j-1}, x_{j+1})_{x_j} + 2\delta &&(\because \text{Inequality } 2.1) \\
&\geq (x_{j-1}, x_k)_{x_j} + \delta. &&(\because \text{Inequality } 2.5)
\end{aligned}
\tag{2.6}
$$

Now Property 2.3 reads

$$(x_i, x_k)_{x_j} \geq \min\{(x_i, x_{j-1})_{x_j}, (x_{j-1}, x_k)_{x_j}\} - \delta \geq (x_{j-1}, x_k)_{x_j} - \delta \tag{2.7}$$

and

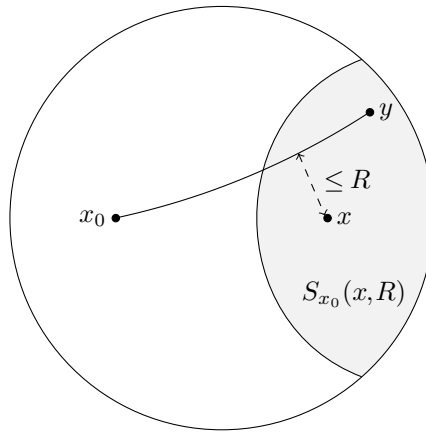$$(x_{j-1}, x_k)_{x_j} \geq \min\{(x_{j-1}, x_i)_{x_j}, (x_i, x_k)_{x_j}\} - \delta. \tag{2.8}$$

Figure 1. Shadow of x with respect to x_0.

If $\min\{(x_{j-1},x_i)_{x_j}, (x_i,x_k)_{x_j}\} = (x_{j-1},x_i)_{x_j}$ then Inequality 2.6 and 2.8 imply

$$(x_i,x_{j-1})_{x_j} > (x_{j-1},x_k)_{x_j} + \delta \geq (x_{j-1},x_i)_{x_j} + 1.5\delta - \delta,$$

a contradiction. Hence, $\min\{(x_{j-1},x_i)_{x_j}, (x_i,x_j)_{x_j}\} = (x_i,x_k)_{x_j}$, and Inequality 2.8 reads

$$(x_{j-1},x_k)_{x_j} \geq (x_i,x_k)_{x_j} - \delta. \tag{2.9}$$

Inequality 2.7 and 2.9 lead to Inequality 2.2. Inequality 2.3 is deduced in a similar manner.
For (2), we have

$$\left| \left(\sum_{i=0}^{m} d(x_i,x_{i+1}) - 2\sum_{i=1}^{m}(x_{i-1},x_{i+1})_{x_i} \right) - d(x_0,x_{m+1}) \right|$$

$$\leq \left| \left(\sum_{i=0}^{m-1} d(x_i,x_{i+1}) - 2\sum_{i=1}^{m-1}(x_{i-1},x_{i+1})_{x_i} \right) - d(x_0,x_m) \right|$$

$$+ |d(x_0,x_m) + d(x_m,x_{m+1}) - d(x_0,x_{m+1}) - 2(x_{m-1},x_{m+1})_{x_m}|$$

$$\leq 2(m-1)\delta + 2|(x_0,x_{m+1})_{x_m} - (x_{m-1},x_{m+1})_{x_m}| \leq 2m\delta. \qquad \square$$

We now define the shadows of a point in terms of the Gromov product.

Definition 2.5 (Shadow). For $x_0, x \in X$ and $R > 0$, the *shadow* $S_{x_0}(x,R)$ is defined by

$$S_{x_0}(x,R) := \{y \in X : (x_0,y)_x \leq R\}.$$

Intuitively, the shadow is a set of points $y \in X$ that the geodesic segment connecting x_0 and y is of distance at most R from x up to an additive constant. See Figure 1.

2.2. Isometries of a Gromov hyperbolic space

We mainly consider the group $\mathrm{Isom}(X)$ of isometries $X \to X$. Isometries of Gromov hyperbolic spaces are classified into the following categories.

Proposition 2.6 (Classification of isometries). *For $g \in \mathrm{Isom}(X)$, one of the following holds.*

(1) *g is elliptic (i.e., $\{g^n x : n \in \mathbb{Z}\}$ is bounded for any $x \in X$);*

(2) *g is parabolic (i.e., g is not elliptic and has exactly one fixed point in ∂X); or*

(3) *g is hyperbolic (i.e., g is not elliptic and has exactly two fixed points in ∂X).*

In particular, when g is hyperbolic, its action shows a source-sink dynamics; one fixed point in ∂X is the attracting point while the other is the repelling point.

One quantity representing the dynamics of an isometry is its *translation length*.

Definition 2.7 (Translation length). *For $g \in \mathrm{Isom}(X)$, its translation length $\tau(g)$ is*

$$\tau(g) := \liminf_{n \to \infty} \frac{d(x, g^n x)}{n}$$

for any $x \in X$.

Note that Lemma 2.4 gives the following corollary. (cf. [MT18, Proposition 5.8])

Corollary 2.8. *For $g \in \mathrm{Isom}(X)$, if g and $x_0 \in X$ satisfy*

$$d(x_0, g x_0) > 2(g x_0, g^{-1} x_0)_{x_0} + 3\delta,$$

then we have

$$\left| \tau(g) - \left(d(x_0,\ g x_0) - 2(g^{-1} x_0,\ g x_0)_{x_0} \right) \right| \leq 2\delta.$$

2.3. Random walks

We define a bi-infinite random walk on a group by adopting the convention in [MT18].

For a countable group G, let $\mu : G \to [0,1]$ be a probability measure on G. Then the product space $(G^{\mathbb{Z}}, \mu^{\mathbb{Z}})$ forms the step space consisting of bi-infinite step paths. To obtain random walks, we consider the map $G^{\mathbb{Z}} \to G^{\mathbb{Z}}$, $(g_n) \mapsto (\omega_n)$ so that ω_0 is the identity and $\omega_{n-1}^{-1} \omega_n = g_n$ for all n. In other words, we set

$$\omega_n = \begin{cases} g_1 \cdots g_n & , n > 0 \\ id & , n = 0 \\ g_0^{-1} g_{-1}^{-1} \cdots g_{n+1}^{-1} & , n < 0. \end{cases}$$

Via this map, $(G^{\mathbb{Z}}, \mu^{\mathbb{Z}})$ induces the probability space $(\Omega, \mathbb{P})$, where Ω denotes the space of sample paths for random walks.

We often need to estimate the distance or the Gromov product among the translates of x_0 by isometries. In this situation, bringing a particular point to the basepoint can ease notation and computation. For example, consider isometries w, g and h. Then $(wg x_0, wh x_0)_{wx_0}$ can also be computed by $(g x_0, h x_0)_{x_0}$, which does not depend on w. In particular, when w are words at step n and g, h are the next steps, then this reduction helps unify the cases.

In this philosophy, we keep the following notations throughout the paper: for a path $\vec{w} = (w_1, \ldots, w_n)$, we write

$$x_n := w_n x_0, \quad x_{n \to m} := w_n^{-1} w_m x_0.$$

Morally, this amounts to shifting the basepoint to x_n and observing the phenomena relative to step n. One can readily observe the following equalities:

(1) $x_{0 \to n} = x_n = w_n x_0$, $x_{n \to 0} = w_n^{-1} x_0$, $x_{n \to n} = x_0$.
(2) $d(x_n, x_m) = d(x_{k \to n}, x_{k \to m}) = d(x_0, x_{n \to m}) = d(x_0, x_{m \to n})$.
(3) $(x_n, x_m)_{x_k} = (x_{k \to n}, x_{k \to m})_{x_0}$.

3. Random walks on hyperbolic spaces

3.1. Ingredients for persistent joints

In this section, (X, d) is a separable geodesic Gromov hyperbolic space, $G \leq \mathrm{Isom}(X)$ is a weakly hyperbolic group, and $\mu : G \to [0,1]$ is a non-elementary probability measure. As in section 2.3, μ induces the probability space $(\Omega, \mathbb{P})$ for random walks.

We recall some facts proved in [MT18]. First of all, Maher and Tiozzo proved in [MT18, Theorem 1.1] that almost every sample path $\omega_n x_0$ converges to a point ω_+ in ∂X. This induces the hitting measure ν on ∂X by

$$\nu(S) := \mathbb{P}(\omega : \omega_n o \to \xi \in S \text{ as } n \to \infty)$$

for Borel sets $S \subseteq \partial X$. Maher and Tiozzo also showed that ν is the unique μ-stationary probability measure on ∂X and is non-atomic. We also borrow the notation

$$Sh(x_0, r) := \{S_{x_0}(g x_0, R) : g \in G, d_X(x_0, g x_0) - R \geq r\}.$$

Proposition 3.1. [MT18, Proposition 5.1]

$$\lim_{r \to \infty} \sup_{S \in Sh(x_0, r)} \nu(\bar{S}) = 0$$

For each subset U of X, we also define

$$H_x^+(U) := \mathbb{P}(\omega_n x \in U \text{ for some } n \geq 0).$$

Similarly, we define $H_x^-(U) := \mathbb{P}(\omega_n x \in U \text{ for some } n \leq 0)$. Then the above proposition holds analogously.

Proposition 3.2 [MT18, Proposition 5.2]. *We have*

$$\lim_{r \to \infty} \sup_{S \in Sh(x_0, r)} H_{x_0}^{\pm}(S) = 0.$$

Let $R_1 > 0$ be such that $\sup_{S \in Sh(x_0, R_1)} H_{x_0}^{\pm}(S) < 0.01$.

Meanwhile, since μ is assumed to be non-elementary, there exist two independent hyperbolic isometries w_+, w_- in $\langle\langle \mathrm{supp}\, \mu \rangle\rangle$, the subsemigroup generated by the support
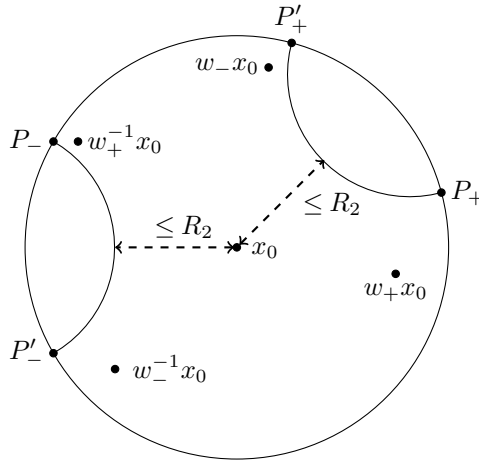
Figure 2. Choice of R and $w_\pm$. Here, P_+ and P'_+ are attracting points and P_- and P'_- are repelling points of w and w', respectively.

of μ. Their independence implies the following: $\{(w_1^n x_0, w_2^m x_0)_{x_0}\}_{m,n>0}$ is bounded, say by $R_2 > 0$, for

$$(w_1, w_2) \in \{(w_+, w_+^{-1}), (w_-, w_-^{-1}), (w_+, w_-), (w_+, w_-^{-1}), (w_+^{-1}, w_-), (w_+^{-1}, w_-^{-1})\}.$$

See Figure 2. We now fix $R = 1000(R_1 + R_2 + \delta + 1)$. Note that the bound R_2 for $\{(w_1^n x_0, w_2^m x_0)_{x_0}\}_{m,n>0}$ still works if we replace w_+ and w_- with their positive powers. By taking suitable powers of $w_\pm$ if necessary, we can assume that:

(1) $w_+, w_- \in \operatorname{supp} \mu^L$ for the same power L (i.e., $w_+ = a_1 \cdots a_L$ and $w_- = b_1 \cdots b_L$ for some isometries $a_i, b_i \in \operatorname{supp} \mu$), and

(2) $d(x_0, w_+ x_0), d(x_0, w_- x_0) > 100R$.

For convenience, we fix the following notations:

$$p_+ = \mu(a_1) \cdots \mu(a_L),$$
$$p_- = \mu(b_1) \cdots \mu(b_L),$$
$$P = \max\left(\frac{p_+}{p_+ + p_-}, \frac{p_-}{p_+ + p_-}\right).$$

3.2. Persistent joints

We define a random variable $\chi_k(\omega)$ that witnesses *persistent joints* at position $3kL$, which is a slight variation of the one defined by Maher and Tiozzo in [MT18, Section 5.2]. See also Figure 3.

Definition 3.3 (Persistent joint). For a sample path $\omega = (\omega_n)$ with the step sequence $(g_n) = (\omega_{n-1}^{-1} \omega_n)$, we define a random variable $\chi_k(\omega)$ as follows.

$\chi_k(\omega) = 1$ if

(1)

$$(g_{3(k-1)L+1},\ldots,g_{3kL}) = \left\{ \begin{array}{c} (b_1,\ldots,b_L,a_1,\ldots,a_L,b_1,\ldots,b_L) \\ \text{or} \\ (b_1,\ldots,b_L,b_1,\ldots,b_L,b_1,\ldots,b_L) \end{array} \right. ,$$

(2) $x_n \in S_{x_{(3k-2)L}}(x_{3(k-1)L}, 0.9R)$ for all integers $n \leq 3(k-1)L$, and

(3) $x_n \in S_{x_{(3k-1)L}}(x_{3kL}, 0.9R)$ for all integers $n \geq 3kL$

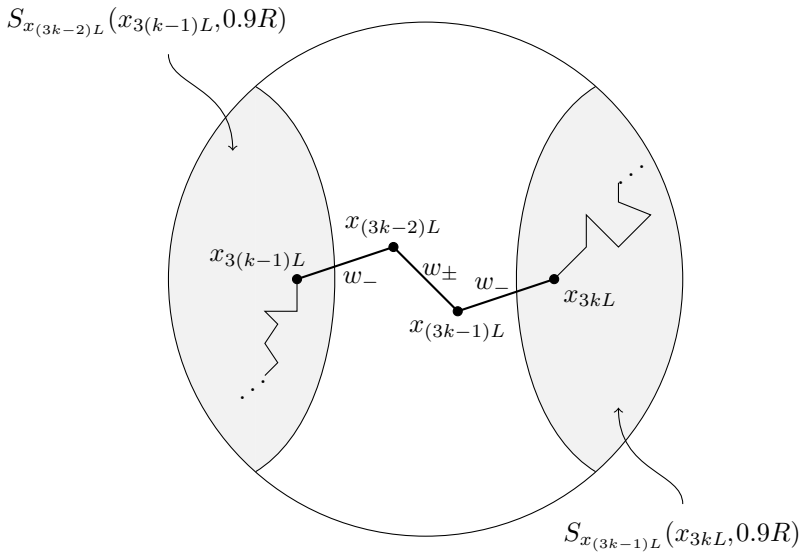where $x_n = \omega_n x_0$. Otherwise, $\chi_k(\omega) = 0$.

Figure 3. Description of a persistent joint.

We first observe that $\mathbb{E}(\chi_1(\omega)) > 0$. The probability for condition (1) is $(p_+ + p_-)p_-^2 \neq 0$. Given (1) as the prior condition, (2) and (3) become independent events. (2) holds if the shifted random walk $T^{3(k-1)L}\omega$ does not hit $S_{x_0}(w_- x_0, d(x_0, w_- x_0) - 0.9R)$ in negative time. Indeed, we have

$$g^{-1}_{3(k-1)L} \cdots g^{-1}_{n+1} x_0 \notin S_{x_0}(w_- x_0, d(x_0, w_- x_0) - 0.9R)$$

$$\Leftrightarrow \omega_n x_0 \notin S_{\omega_{3(k-1)L} x_0}(\omega_{3(k-1)L} w_- x_0, d(x_0, w_- x_0) - 0.9R)$$

$$\Leftrightarrow \omega_n x_0 \in S_{\omega_{(3k-2)L} x_0}(\omega_{3(k-1)L} x_0, 0.9R)$$

for $n < 3(k-1)L$. Here, the last inequality follows from the equality

$$(p, y)_z + (p, z)_y = d(y, z)$$

for $p \in X$ and $y = \omega_{3(k-1)L} x_0$, $z = \omega_{3(k-1)L} w_- x_0 = \omega_{(3k-2)L} x_0$. Thus, the probability for condition (2) is at least

$$1 - H_{x_0}^-(S_{x_0}(w_- x_0, d(x_0, w_- x_0) - R)) \geq 1 - H_{x_0}^-(S_{x_0}(w_- x_0, 99R)) \geq 0.99.$$

Similarly, the probability for condition (3) is at least

$$1 - H_{x_0}^+(S_{x_0}(w_-^{-1} x_0, d(x_0, w_- x_0) - R)) \geq 0.99.$$

Overall, we have

$$\eta := \mathbb{P}(\chi_1(\omega) = 1) \geq (p_+ + p_-)p_-^2 \cdot (0.99)^2 > 0.$$

Note also that $\chi_k(\omega) = \chi_1(T^{3(k-1)L} \omega)$. We invoke a variant of Kingman's subadditive ergodic theorem.

Theorem 3.4 [Woe00, (8.10) Theorem]. *Let $(\Omega, \mathbb{P})$ be a probability space and $U : \Omega \to \Omega$ be a measure-preserving transformation. If W_n is a non-negative real-valued random variable on Ω satisfying the subadditivity $W_{n+m} \leq W_n + W_m \circ U^n$ for all $m, n \in \mathbb{N}$, and W_1 has finite first moment, then there is a U-invariant random variable W_∞ such that*

$$\lim_{n \to \infty} \frac{1}{n} W_n = W_\infty$$

almost surely and in $L^1(\Omega, \mathbb{P})$. If U is ergodic in addition, then W_∞ is constant a.e.

We define $W_n = \sum_{k=1}^n \chi_k(\omega)$. Then $W_{n+m} = W_n + W_m \circ T^{3Ln}$ holds. Since W_1 is bounded, it has finite first moment. Applying Theorem 3.4, we get almost everywhere convergence of $\frac{1}{n} W_n$ to an a.e. constant W_∞. Since $\mathbb{E}(W_1) = \eta > 0$, we have $W_\infty = \eta$ a.e.

We now consider a modified version of W_n. Given positive integers $m \leq n$, we say that $\mathcal{N} = \{n_1 < \cdots < n_k\} \subseteq 3L\mathbb{Z}$ is an (m,n)-*set of pivots* for a finite path $\vec{w} = (w_1, \ldots, w_n)$ if the following hold:

(1) $\mathcal{N} \subseteq \{1, \ldots, m\}$;

(2) for each $i = 1, \ldots, k$,

$$(g_{n_i - 3L + 1}, \ldots, g_{n_i}) = \begin{cases} (b_1, \ldots, b_L, a_1, \ldots, a_L, b_1, \ldots, b_L) \\ \text{or} \\ (b_1, \ldots, b_L, b_1, \ldots, b_L, b_1, \ldots, b_L) \end{cases},$$

(3) for each $i = 1, \ldots, k$, $x_j \in S_{x_{n_i - 2L}}(x_{n_i - 3L}, 0.9R)$ for $n_{i-1} - L \leq j \leq n_i - 3L$, and

(4) for each $i = 1, \ldots, k$, $x_j \in S_{x_{n_i - L}}(x_{n_i}, 0.9R)$ for $n_i \leq j \leq n_{i+1} - 2L$.

(For convenience, we set $n_0 = L$ and $n_{k+1} = n + 2L$.) Note that if $\mathcal{N}$ and $\mathcal{N}'$ are (m,n)-sets of pivots for $\vec{w} = (w_1, \ldots, w_n)$, then so is their union. Thus, we can associate each finite path $\vec{w} = (w_1, \ldots, w_n)$ with its maximal (m,n)-set of pivots $\mathcal{N}(\vec{w}) = \mathcal{N}_{m,n}(\vec{w})$. We also define

$$F_n := \left\{ \vec{w} = (w_1, \ldots, w_n) : \#\mathcal{N}_{n,n}(\vec{w}) \geq \frac{\eta n}{6L} + 1 \right\}.$$

Note that $\#\mathcal{N}_{n,n}(\omega_1,\ldots,\omega_n) \geq W_{\lfloor n/3L \rfloor}(\omega_1,\ldots,\omega_n)$ for each n and $W_{\lfloor n/3L \rfloor}(\omega) \geq \eta \lfloor n/3L \rfloor \geq \frac{\eta n}{6L} + 1$ eventually holds for a.e. sample path ω. Consequently, $(\omega_1,\ldots,\omega_n) \in F_n$ eventually holds for a.e. ω.

Let us now fix a finite path $\vec{w} = (w_1,\ldots,w_n)$ with $\mathcal{N}_{n,n}(\vec{w}) = \{n_1 < \ldots < n_k\}$. For convenience, we define the following for $i = 1,\ldots,k$:

$$\begin{aligned}
A_i'(\vec{w}) &:= n_i(\vec{w}) - 3L, \\
\alpha_i'(\vec{w}) &:= n_i(\vec{w}) - 2L, \\
\beta_i'(\vec{w}) &:= n_i(\vec{w}) - L, \\
B_i'(\vec{w}) &:= n_i(\vec{w}).
\end{aligned} \tag{3.1}$$

We also let $\beta_0'(\vec{w}) = B_0'(\vec{w}) := 0$ and $\alpha_{k+1}'(\vec{w}) = A_{k+1}'(\vec{w}) := n$. The following lemma allows us to calculate the distances among x_j.

Lemma 3.5. *We have the following:*

(1) $d(x_{\beta_{i-1}'}, x_{\alpha_i'}) > 99R$ *for each* $i = 1,\ldots,k+1$;
(2) $(x_{\alpha_i' \to \beta_{i-1}'}, w_\pm x_0)_{x_0} < 0.6R$ *for each* $i = 1,\ldots,k$, *and*
(3) $(x_{\beta_i' \to \alpha_{i+1}'}, w_\pm^{-1} x_0)_{x_0} < 0.6R$ *for each* $i = 1,\ldots,k$.

Proof. Let us discuss (1). For $2 \leq i \leq k+1$, we have

$$(x_{\alpha_i'}, x_{B_{i-1}'})_{x_{\beta_{i-1}'}} = d(x_{\beta_{i-1}'}, x_{B_{i-1}'}) - (x_{\alpha_i'}, x_{\beta_{i-1}'})_{x_{B_{i-1}'}} \geq 100R - 0.9R > 99R.$$

Since $(x_{\alpha_i'}, x_{B_{i-1}'})_{x_{\beta_{i-1}'}} \leq d(x_{\alpha_i'}, x_{\beta_{i-1}'})$, we deduce the desired conclusion. Similar discussion on $(x_{\beta_{i-1}'}, x_{A_i'})_{x_{\alpha_i'}}$ for $1 \leq i \leq k$ handles the remaining case:

$$(x_{\beta_{i-1}'}, x_{A_i'})_{x_{\alpha_i'}} = d(x_{\alpha_i'}, x_{A_i'}) - (x_{\beta_{i-1}'}, x_{\alpha_i'})_{x_{A_i'}} \geq 100R - 0.9R > 99R. \tag{3.2}$$

For (2), we invoke Property 2.3:

$$\min\left\{(w_-^{-1} x_0, x_{\alpha_i' \to \beta_{i-1}'})_{x_0}, (x_{\alpha_i' \to \beta_{i-1}'}, w_\pm x_0)_{x_0}\right\} - \delta \leq (w_-^{-1} x_0, w_\pm x_0)_{x_0} < 0.5R.$$

Here, $(w_-^{-1} x_0, x_{\alpha_i' \to \beta_{i-1}'})_{x_0} = (x_{\alpha_i' \to A_i'}, x_{\alpha_i' \to \beta_{i-1}'})_{x_0} > 99R$ was proven in Inequality 3.2. Hence, we deduce $(x_{\alpha_i' \to \beta_{i-1}'}, w_\pm x_0)_{x_0} < 0.5R + \delta < 0.6R$. (3) is argued similarly. □

Corollary 3.6. *Let*

$$(y_{2i-1}, y_{2i}) := (x_{\alpha_i'}, x_{\beta_i'})$$

for $i = 1,\ldots,k$ *and* $y_0 = x_0$, $y_{2k+1} = x_n$. *Then we have*

$$(y_i, y_l)_{y_j} \leq 0.8R, \quad d(y_i, y_j) \leq d(y_i, y_l) - 95(l-j)R$$

for all $0 \leq i \leq j \leq l \leq 2k+1$.

Proof. The first item of Lemma 3.5 tells us that $d(y_{2i}, y_{2i+1}) > 99R$ for $i = 0,\ldots,k$. Moreover, $d(y_{2i-1}, y_{2i}) = d(x_0, w_\pm x_0) \geq 100R$ for $i = 1,\ldots,k$. Finally, (2) and (3) of

Lemma 3.5 read that $(y_{i-1}, y_{i+1})_{y_i} < 0.6R$ for $i = 1, \ldots, 2k$ and

$$(y_{i-1}, y_{i+1})_{y_i} + (y_i, y_{i+2})_{y_{i+1}} < 1.2R < 99R - 3\delta \leq d(y_i, y_{i+1}) - 3\delta$$

for $i = 1, \ldots, 2k-1$. We can then apply Lemma 2.4 and conclude

$$(y_i, y_l)_{y_j} \leq (y_{j-1}, y_{j+1})_{y_j} + 2\delta \leq 0.8R$$

for $0 \leq i < j < l \leq 2k+1$. Moreover, this implies

$$d(y_i, y_j) \leq d(y_i, y_{j+1}) - d(y_j, y_{j+1}) + 2(y_i, y_{j+1})_{y_j} \leq d(y_i, y_{j+1}) - 99R + 1.6R$$

for $0 \leq i \leq j < 2k+1$, which leads to the second conclusion. □

We now define constants D, M and set G_n such that

$$D := R\eta/L,$$
$$M > 1 + 10R + 2d(x_0, w_+ x_0) + 2d(x_0, w_- x_0) + 4L,$$
$$G_n := \left\{ \vec{w} \in F_n : \tau(w_n) \leq \left(2D - \frac{2\eta}{M} \right) n \right\}.$$

Furthermore, for $\vec{w} \in F_n$ and $Q > 0$, we define

$$\mathcal{N}_f(\vec{w}; Q) = \left\{ n_i \in \mathcal{N}(\vec{w}) : d(x_0, x_{\beta_i'}) \leq \frac{1}{2} d(x_0, x_n) - (D + Q\eta/M)n \right\},$$
$$\mathcal{N}_b(\vec{w}; Q) = \left\{ n_i \in \mathcal{N}(\vec{w}) : d(x_n, x_{\alpha_i'}) \leq \frac{1}{2} d(x_0, x_n) - (D + Q\eta/M)n \right\},$$
$$\mathcal{N}_0(\vec{w}; Q) = \mathcal{N}(\vec{w}) \setminus [\mathcal{N}_f(\vec{w}; Q) \cup \mathcal{N}_b(\vec{w}; Q)].$$

We observe the following estimation of $\mathcal{N}_f$ and $\mathcal{N}_b$.

Lemma 3.7. *For $n > 10L/\eta$, if $\vec{w} \in F_n$ and $Q \leq 1$, then*

$$|\mathcal{N}_f(\vec{w}; Q)| \geq \frac{\eta n}{20LM^2} \quad or \quad |\mathcal{N}_b(\vec{w}; Q)| \geq \frac{\eta n}{20LM^2}$$

holds.

Proof. Suppose not. Then $\mathcal{N}_0(\vec{w}; Q)$ contains at least $\frac{\eta n}{15L} + 1$ indices. By setting $i = 1$ and increasing j, Corollary 3.6 implies that $d(x_0, x_{\beta_j'})$ increases as j increases. Similarly, by setting $j = N + 1$ and decreasing i, we realize that $d(x_{\alpha_i}, x_n)$ increases as i decreases. Consequently, we may take t, t' such that

$$\mathcal{N}_0(\vec{w}; Q) = \{n_t, n_{t+1}, \ldots, n_{t'}\}.$$

Again, using Corollary 3.6, we deduce that

$$d(x_{\beta_t'}, x_{\alpha_{t'}'}) \geq 90R \cdot \frac{\eta n}{15L} \geq 6R\eta n/L.$$

Recall that $R > 1000 > Q$, $M > L$, $D = \frac{R\eta}{L}$ and $\frac{2\eta n}{L} \geq 20$. Then Corollary 3.6 implies that

$$d(x_0, x_n) \geq d(x_0, x_{\beta'_t}) + d(x_{\beta'_t}, x_{\alpha'_{t'}}) + d(x_{\alpha'_{t'}}, x_n) - 4.4R$$

$$\geq d(x_0, x_n) - 2Dn - \frac{2Q\eta n}{M} + \frac{6R\eta n}{L} - 4.4R$$

$$\geq d(x_0, x_n) - \frac{2R\eta n}{L} - \frac{2R\eta n}{L} + \frac{6R\eta n}{L} - 4.4R > d(x_0, x_n),$$

wich is a contradiction. $\qquad\square$

From this lemma, it follows that

$$F_{n,f}(Q) := \left\{ \vec{w} \in F_n : |\mathcal{N}_f(\vec{w}; Q)| \geq \frac{\eta n}{20LM^2} \right\},$$

$$F_{n,b}(Q) := \left\{ \vec{w} \in F_n : |\mathcal{N}_b(\vec{w}; Q)| \geq \frac{\eta n}{20LM^2} \right\}$$

cover entire F_n for $Q \leq 1$ and large enough n. We also define

$$G_{n,f}(Q) := F_{n,f}(Q) \cap G_n \quad \text{and} \quad G_{n,b}(Q) := F_{n,b}(Q) \cap G_n.$$

For each $\vec{w} \in F_{n,f}(Q)$ ($\vec{w} \in F_{n,b}(Q)$, resp.), we fix an integer $N = N(\vec{w})$ between $\frac{\eta n}{50M^2 L}$ and $\frac{\eta n}{20M^2 L}$. Then we pick *pivot indices* $p_1(\vec{w}) < \ldots < p_N(\vec{w})$ from $\mathcal{N}_f$ ($\mathcal{N}_b$, resp.). We introduce a notation:

$$A_i(\vec{w}) := p_i(\vec{w}) - 3L,$$
$$\alpha_i(\vec{w}) := p_i(\vec{w}) - 2L,$$
$$\beta_i(\vec{w}) := p_i(\vec{w}) - L,$$
$$B_i(\vec{w}) := p_i(\vec{w}).$$

For convenience, we also let $\beta_0(\vec{w}) = B_0(\vec{w}) := 0$ and $\alpha_{N+1}(\vec{w}) = A_{N+1}(\vec{w}) := n$. Recalling Figure 3, A_i, α_i, β_i, B_i are described as in Figure 4

For each choice of $\sigma \in \{0,1\}^N$, we now define the pivoted word $\vec{w}^\sigma$ by declaring the steps $\{g_i^\sigma\}$ as follows. In plain words, we modify the type of joints that are marked by σ only. For $\sigma(i) = 1$, we set

$$(g_{\alpha_i+1}^\sigma, \ldots, g_{\beta_i}^\sigma) := \begin{cases} (a_1, \ldots, a_L) & \text{if } (g_{\alpha_i+1}, \ldots, g_{\beta_i}) = (b_1, \ldots, b_L) \\ (b_1, \ldots, b_L) & \text{if } (g_{\alpha_i+1}, \ldots, g_{\beta_i}) = (a_1, \ldots, a_L). \end{cases}$$

Other steps remain unchanged. For $\vec{w}^\sigma = (w_i^\sigma)$, we similarly denote $w_n^\sigma x_0$ by x_n^σ and $(w_n^\sigma)^{-1} w_m^\sigma x_0$ by $x_{n \to m}^\sigma$. Then we observe

$$x_{\alpha_i \to \beta_{i-1}} = x_{\alpha_i \to \beta_{i-1}}^\sigma, \quad x_{\beta_{i-1} \to \alpha_i} = x_{\beta_{i-1} \to \alpha_i}^\sigma.$$

By definition, we have

$$x_{\alpha_i \to \beta_i}^\sigma \in \{w_+, w_-\}, \quad x_{\alpha_i \to \beta_i}^\sigma = x_{\alpha_i \to \beta_i} \quad \text{iff} \quad \sigma(i) = 0.$$

Clearly, $\vec{w}^\sigma \neq \vec{w}^{\sigma'}$ for $\sigma \neq \sigma'$. Finally, note that $\mathcal{N}(\vec{w}) = \mathcal{N}(\vec{w}^\sigma)$ for $\sigma \in \{0,1\}^N$, while *a priori* $\mathcal{N}_f(\vec{w})$ and $\mathcal{N}_f(\vec{w}^\sigma)$ may not coincide.

The proof of Lemma 3.5 and Corollary 3.6 also apply here, so we omit the proof.
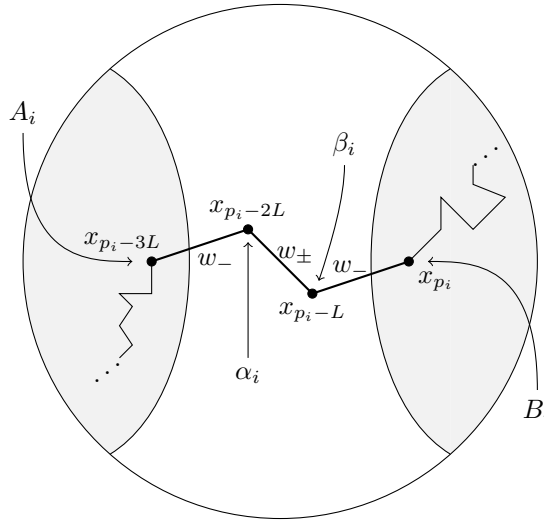
Figure 4. Choice of A_i, α_i, β_i, B_i.

Lemma 3.8. *We have the following.*

(1) $d(x^\sigma_{\beta'_{i-1}}, x^\sigma_{\alpha'_i}) > 99R$ *for each* $1 \le i \le k+1$.

(2) $(x^\sigma_{\alpha'_i \to \beta'_{i-1}}, w_\pm x_0)_{x_0} < 0.6R$ *for each* i.

(3) $(x^\sigma_{\beta'_i \to \alpha'_{i+1}}, w_\pm^{-1} x_0)_{x_0} < 0.6R$ *for each* i.

Corollary 3.9. *Let*

$$(y^\sigma_{2i-1}, y^\sigma_{2i}) := (x^\sigma_{\alpha_i}, x^\sigma_{\beta_i})$$

for $i = 1, \ldots, k$ *and* $y_0 = x_0$, $y_{2k+1} = x_n$. *Then* $(y_i, y_l)_{y_j} \le 0.8R$ *and* $d(y_i, y_j) \le d(y_i, y_l) - 95(l-j)R$ *for all* $0 \le i < j < l \le 2N+1$.

Now we have all ingredients ready.

Lemma 3.10. *Let* $Q \ge 0.9$ *and* $n \ge 40MR/\eta$. *For each* $\vec{w} \in F_{n,f}(Q)$ *and* $\sigma \ne \kappa \in \{0,1\}^N$, *if* $\tau(w_n^\kappa) \le (2D - 2\eta/M)n$, *then* $\tau(w_n^\sigma) \ge (2D + \eta/M)n$.

Proof. We first establish a bound on $d(x_0, x_n^\sigma)$. Note that this part does not require any assumption on $\tau(w_n^\kappa)$.

Claim 3.11. We have

$$|d(x_0, x_n^\kappa) - d(x_0, x_n^\sigma)| \le \frac{\eta n}{40M},$$

and for each $i = 1, \cdots, N$,

$$|d(x_n^\kappa, x_{\alpha_i'}^\kappa) - d(x_n^\sigma, x_{\alpha_i'}^\sigma)| \le \frac{\eta n}{40M},$$
$$|d(x_0, x_{\beta_i'}^\kappa) - d(x_0, x_{\beta_i'}^\sigma)| \le \frac{\eta n}{40M}.$$

Proof of the Claim. By Lemma 2.4, we have

$$
\begin{aligned}
d(x_0, x_n^\sigma) &\ge \sum_{i=1}^{N+1} d(x_{\beta_{i-1}}^\sigma, x_{\alpha_i}^\sigma) + \sum_{i=1}^{N} d(x_{\alpha_i}^\sigma, x_{\beta_i}^\sigma) \\
&\quad - 2\sum_{i=1}^{N} (x_{\beta_i}^\sigma, x_{\beta_{i-1}}^\sigma)_{x_{\alpha_i}^\sigma} - 2\sum_{i=1}^{N} (x_{\alpha_i}^\sigma, x_{\alpha_{i+1}}^\sigma)_{x_{\beta_i}^\sigma} - 2 \cdot 2N \cdot \delta \\
&\ge \sum_{i=1}^{N+1} d(x_{\beta_{i-1}}^\sigma, x_{\alpha_i}^\sigma) = \sum_{i=1}^{N+1} d(x_{\beta_{i-1}}^\kappa, x_{\alpha_i}^\kappa).
\end{aligned}
$$

Recall also that $d(x_{\alpha_i}^\kappa, x_{\beta_i}^\kappa)$ is either $d(x_0, w_+ x_0)$ or $d(x_0, w_- x_0)$, which are both smaller than $M/2$. Since N is chosen to be less than $\frac{\eta n}{20M^2 L}$, we have

$$
\begin{aligned}
d(x_0, x_n^\sigma) &\ge \sum_{i=1}^{N+1} d(x_{\beta_{i-1}}^\kappa, x_{\alpha_i}^\kappa) + \sum_{i=1}^{N} d(x_{\alpha_i}^\kappa, x_{\beta_i}^\kappa) - 0.5M \cdot \frac{\eta n}{20M^2 L} \\
&\ge d(x_0, x_n^\kappa) - \frac{\eta n}{40ML} \\
&\ge d(x_0, x_n^\kappa) - \frac{\eta n}{40M}.
\end{aligned}
$$

A symmetric argument shows that $d(x_0, x_n^\sigma) \le d(x_0, x_n^\kappa) + \eta n/40M$. Similar arguments involving partial sums give the remaining results. $\square$

Now suppose $\tau(w_n^\kappa) \le (2D - 2\eta/M)n$. We observe the following:

Claim 3.12.

$$(x_n^\kappa, x_{n\to 0}^\kappa)_{x_0} \ge \frac{1}{2} d(x_0, x_n) - Dn. \tag{3.3}$$

Proof of the Claim. Recall that $M > L$, $R > 1 + 3\delta$ and $n > 40MR/\eta$. This implies $1/40M < R/L$, and consequently,

$$3\delta < R < \frac{\eta n}{40M} < Dn. \tag{3.4}$$

Now suppose that Inequality 3.3 does not hold. This implies

$$
\begin{aligned}
d(x_0, x_n^\kappa) - 2(x_n^\kappa, x_{n\to 0}^\kappa)_{x_0} &\ge \left[d(x_0, x_n) - \frac{\eta n}{40M} \right] - 2(x_n^\kappa, x_{n\to 0}^\kappa)_{x_0} \quad (\because \text{ Claim 3.11}) \\
&> 2Dn - \frac{\eta n}{40M} > \frac{\eta n}{40M} > 3\delta.
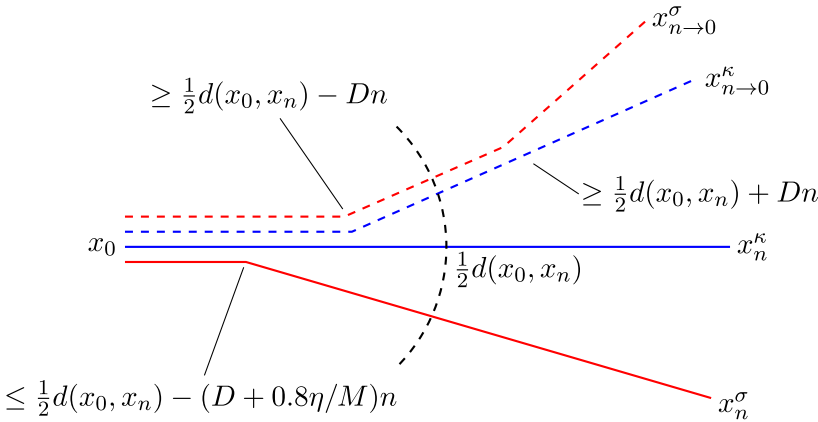\end{aligned}
$$

Figure 5. Four segments in the pivoting process and the Gromov products.

Then from Corollary 2.8, we have

$$\tau(w_n^\kappa) \geq d(x_0, x_n^\kappa) - 2(x_n^\kappa, x_{n \to 0}^\kappa)_{x_0} - 2\delta \geq \left(2D - \frac{\eta}{40M} - \frac{\eta}{40M}\right)n. \tag{3.5}$$

This contradicts the assumption that $\tau(w_n^\kappa) \leq (2D - 2\eta/M)n$. $\qquad\square$

Fixing $\sigma \in \{0,1\}^N$ other than κ, we now estimate the Gromov products among four points $x_n^\kappa, x_n^\sigma, x_{n \to 0}^\kappa$, and $x_{n \to 0}^\sigma$ based at x_0. See Figure 5.

Claim 3.13. We have

$$(x_{n \to 0}^\kappa, x_{n \to 0}^\sigma)_{x_0} \geq \frac{1}{2}d(x_0, x_n) + Dn. \tag{3.6}$$

Proof of the Claim. Since $\vec{w} \in F_{n,f}$, $B_N = p_N$ belongs to $\mathcal{N}_f(\vec{w})$. Note that the steps $(g_i)_i$ of $\vec{w}$ after β_N are not altered by the pivoting; thus, we have

$$x_{n \to \beta_N}^\kappa = x_{n \to \beta_N}^\sigma = x_{n \to \beta_N}.$$

Moreover, Claim 3.11 and $p_N \in \mathcal{N}_f(\vec{w})$ imply

$$d(x_{n \to 0}^\kappa, x_{n \to \beta_N}^\kappa) = d(x_0^\kappa, x_{\beta_N}^\kappa) \leq d(x_0, x_{\beta_N}) + \frac{\eta n}{40M}$$
$$\leq \frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n + \frac{\eta n}{40M}.$$

The same upper bound also applies to $d(x_{n \to 0}^\sigma, x_{n \to \beta_N}^\sigma)$. Hence,

$$d(x_{n \to 0}^\kappa, x_{n \to 0}^\sigma) \leq d(x_{n \to 0}^\kappa, x_{n \to \beta_N}^\kappa) + d(x_{n \to \beta_N}^\kappa, x_{n \to \beta_N}^\sigma) + d(x_{n \to \beta_N}^\sigma, x_{n \to 0}^\sigma)$$
$$\leq 2 \cdot \left(\frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n + \frac{\eta n}{40M}\right).$$

This inequality and Claim 3.11 together yield

$$(x_{n\to 0}^{\kappa}, x_{n\to 0}^{\sigma})_{x_0} \geq \frac{1}{2}[d(x_0, x_n^{\kappa}) + d(x_0, x_n^{\sigma})] - \left(\frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n + \frac{\eta n}{40M}\right)$$

$$\geq d(x_0, x_n) - \frac{\eta n}{40M} - \left(\frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n + \frac{\eta n}{40M}\right)$$

$$\geq \frac{1}{2}d(x_0, x_n) + Dn + \left(\frac{0.9\eta}{M} - \frac{\eta}{20M}\right)n. \qquad \square$$

In contrast, we get the following inequality.

Claim 3.14. We have

$$(x_n^{\kappa}, x_n^{\sigma})_{x_0} \leq \frac{1}{2}d(x_0, x_n) - (D + 0.8\eta/M)n.$$

Proof of the Claim. To see this, let l be the minimum among $\{1, \ldots, N\}$ such that $\sigma(l) \neq \kappa(l)$. We let

$$y_0 = x_n^{\kappa}, \quad y_1 = x_{\beta_l}^{\kappa}, \quad y_2 = x_{\alpha_l}^{\kappa} = x_{\alpha_l}^{\sigma}, \quad y_3 = x_{\beta_l}^{\sigma}, \quad y_4 = x_n^{\sigma}.$$

Then Corollary 3.9 tells us that $d(y_{i-1}, y_i) > 99R$ for $i = 1, \ldots, 4$ and $(y_0, y_2)_{y_1}, (y_2, y_4)_{y_3} < 0.8R$. We also have $(y_1, y_3)_{y_2} = (w_+ x_0, w_- x_0)_{x_0} \leq R_1 < 0.5R$. Hence, we have

$$(y_{i-1}, y_{i+1})_{y_i} + (y_i, y_{i+2})_{y_{i+1}} \leq 1.3R < 99R - 3\delta \leq d(y_i, y_{i+1})$$

for $i = 1, 2$. We then apply Lemma 2.4 and deduce $(y_0, y_4)_{y_2} < 0.5R + 2\delta < 0.6R$. This implies

$$\begin{aligned}
2(x_n^{\kappa}, x_n^{\sigma})_{x_0} &= d(x_0, y_0) + d(x_0, y_4) - d(y_0, y_4) \\
&\leq \big(d(x_0, y_2) + d(y_2, y_0)\big) + \big(d(x_0, y_2) + d(y_2, y_4)\big) \\
&\quad - \big(d(y_0, y_2) + d(y_2, y_4) - 2(y_0, y_4)_{y_2}\big) \\
&\leq 2d(x_0, y_2) + 2 \cdot 0.6R.
\end{aligned} \qquad (3.7)$$

Meanwhile, the distance from x_0 and $y_2 = x_{\alpha_l}^{\kappa}$ is estimated as follows:

$$d(x_0, x_{\alpha_l}^{\kappa}) \leq d(x_0, x_{\beta_N}^{\kappa}) \qquad (\because \text{ Corollary } 3.9)$$

$$\leq d(x_0, x_{\beta_N}) + \frac{\eta n}{40M} \qquad (\because \text{ Claim } 3.11)$$

$$\leq \frac{1}{2}d(x_0, x_n) - \left(D + \frac{Q\eta}{M}\right)n + \frac{\eta n}{40M} \qquad (p_N \in \mathcal{N}_f(\vec{w})).$$

By plugging this into Inequality 3.7, we obtain

$$(x_n^{\kappa}, x_n^{\sigma})_{x_0} \leq \frac{1}{2}d(x_0, x_n) - \left(D + \frac{Q\eta}{M}\right)n + \frac{\eta n}{40M} + 0.6R$$

$$\leq \frac{1}{2}d(x_0, x_n) - \left(D + \frac{0.9\eta}{M}\right)n + \frac{0.05\eta n}{M} + \frac{0.02 n\eta}{M}$$

$$\leq \frac{1}{2}d(x_0, x_n) - \left(D + \frac{0.8\eta}{M}\right)n$$

for $n > 40RM/\eta$, as desired. $\qquad \square$

Now let us finish the proof of Lemma 3.10. If we have

$$(x_n^\sigma, x_{n\to0}^\sigma)_{x_0} \geq \frac{1}{2}d(x_0, x_n) - \left(D + \frac{0.7\eta}{M}\right)n, \tag{3.8}$$

then we have

$$
\begin{aligned}
(x_n^\kappa, x_n^\sigma)_{x_0} &\geq \min\{(x_n^\kappa, x_{n\to0}^\kappa)_{x_0}, (x_{n\to0}^\kappa, x_n^\sigma)_{x_0}\} - \delta && (\because \text{ Property } 2.3) \\
&\geq \min\left\{ \begin{array}{c} (x_n^\kappa, x_{n\to0}^\kappa)_{x_0}, (x_{n\to0}^\kappa, x_{n\to0}^\sigma)_{x_0}, \\ (x_{n\to0}^\sigma, x_n^\sigma)_{x_0} \end{array} \right\} - 2\delta && (\because \text{ Property } 2.3) \\
&\geq \frac{1}{2}d(x_0, x_n) - \left(D + \frac{0.7\eta}{M}\right)n - 2\delta && \left(\because \begin{array}{c} \text{Claim } 3.12, \\ \text{Claim } 3.13 \end{array} \right) \\
&> \frac{1}{2}d(x_0, x_n) - \left(D + \frac{0.8\eta}{M}\right)n. && (\because \text{ Inequality } 3.4)
\end{aligned}
$$

This contradicts Claim 3.14. Therefore, Inequality 3.8 does not hold and

$$
\begin{aligned}
d(x_0, x_n^\sigma) - 2(x_n^\sigma, x_{n\to0}^\sigma)_{x_0} &\geq \left[d(x_0, x_n) - \frac{\eta n}{40M}\right] - 2(x_n^\sigma, x_{n\to0}^\sigma)_{x_0} \\
&\geq 2\left(D + \frac{0.7\eta}{M}\right)n - \frac{\eta n}{40M} \\
&\geq (2D + \eta/M)n + 3\delta > 3\delta. && (\because \text{ Inequality } 3.4)
\end{aligned}
$$

Then Corollary 2.8 tells us that $\tau(\omega_n^\sigma) \geq (2D + \eta/M)n$. □

In particular, for $\vec{w} \in G_{n,f}(Q)$, $\vec{w}^\sigma \notin G_n$ for any nontrivial σ. Similar discussion holds for $F_{n,b}$ and $G_{n,b}$.

In the following crucial observation, we finally set the optimal value for Q and $N(\vec{w})$.

Lemma 3.15. *Let $Q = 1$ and $n \geq 40RM/\eta$. Suppose that $\vec{w}, \vec{w}' \in G_{n,f}(Q=1)$ and the numbers of pivots $N(\vec{w})$, $N(\vec{w}')$ are $\lfloor \frac{\eta n}{40M^2 L} \rfloor$. Then for $\sigma, \sigma' \in \{0,1\}^N$, $\vec{w}^\sigma = \vec{w}'^{\sigma'}$ if and only if $\vec{w} = \vec{w}'$ and $\sigma = \sigma'$.*

Proof. Let $\vec{v} = \vec{w}^\sigma = \vec{w}'^{\sigma'}$. The key idea here is to change the roles of $\vec{w}$, $\vec{w}'$ and $\vec{v}$. First note that

$$\mathcal{N}_f(\vec{w}; Q=1) \subseteq \mathcal{N}(\vec{w}) \subseteq \mathcal{N}(\vec{v} = \vec{w}^\sigma).$$

Moreover, for each $n_i \in \mathcal{N}_f(\vec{w}; Q=1)$, we have

$$
\begin{aligned}
d(x_0, x_{\beta_i}^\sigma) &\leq d(x_0, x_{\beta_i}) + \frac{\eta n}{40M} && (\because \text{ Claim } 3.11) \\
&\leq \frac{1}{2}d(x_0, x_n) + \frac{\eta n}{40M} - \left(D + \frac{\eta}{M}\right)n && (\because p_i \in \mathcal{N}_f(\vec{w}; Q=1)) \\
&\leq \frac{1}{2}d(x_0, x_n^\sigma) + \frac{3\eta n}{80M} - \left(D + \frac{\eta}{M}\right)n && (\because \text{Claim } 3.11) \\
&\leq \frac{1}{2}d(x_0, x_n^\sigma) - \left(D + \frac{0.9\eta}{M}\right)n.
\end{aligned}
$$

This implies $n_i \in \mathcal{N}_f(\vec{v}; Q = 0.9)$. It follows that

$$\mathcal{N}_f(\vec{w}; Q = 1) \subseteq \mathcal{N}_f(\vec{v}; Q = 0.9) \quad \text{and} \quad \vec{v} \in F_{n,f}(Q = 0.9).$$

Similarly, $\mathcal{N}_f(\vec{w}'; Q = 1) \subseteq \mathcal{N}_f(\vec{v}; Q = 0.9)$.

Thus, we are able to pick forward pivots $p_i(\vec{w})$ and $p_i'(\vec{w}')$ of $\vec{w}$ and $\vec{w}'$ altogether for $\vec{v}$. (This will give $N(\vec{v}) \leq \frac{\eta n}{20M^2L}$, which is legitimate.) Then Lemma 3.10 applied to $\vec{v} \in F_{n,f}(Q = 0.9)$ yields a contradiction with $\vec{w}, \vec{w}' \in G_n$ unless $\sigma = \sigma'$. $\square$

Similar discussion also holds for $G_{n,b}$.

3.3. Translation lengths of random isometries

We now prove the first main theorem:

Theorem A. *Let G be a weakly hyperbolic group acting on a Gromov hyperbolic space X and let μ be a non-elementary probability measure on G. Then $\mathbb{P}$-almost every sample path shows at least linear growth of translation lengths. More precisely, there exists a constant $\mathcal{L} > 0$ such that for $\mathbb{P}$-a.e. (ω_n), $\tau(\omega_n) \geq \mathcal{L}n$ for sufficiently large n.*

Moreover, if μ further has finite first moment, then for $\mathbb{P}$-a.e. (ω_n), we have

$$\lim_{n \to \infty} \frac{\tau(\omega_n)}{n} = \lambda,$$

where λ is the drift of the random walk.

Proof. Let $n \geq 40RM/\eta$. By Lemma 3.7, $F_{n,f}(Q = 1)$ and $F_{n,b}(Q = 1)$ cover entire F_n. Consequently, $G_{n,f}(Q = 1)$ and $G_{n,b}(Q = 1)$ cover G_n.

As in Lemma 3.15, we let $N(\vec{w}) = \lfloor \frac{\eta n}{40M^2L} \rfloor$ for each $\vec{w} \in G_{n,f}(Q = 1)$. Let $\mathcal{C}(\vec{w}) := \{\vec{w}^\sigma : \sigma \in \{0,1\}^N\}$. Lemma 3.15 asserts that $\mathcal{C}(\vec{w})$ and $\mathcal{C}(\vec{w}')$ are disjoint for distinct elements $\vec{w}, \vec{w}' \in G_{n,f}(Q = 1)$. Moreover, for each $\vec{w} \in G_{n,f}(Q = 1)$, the conditional probability of $\vec{w}$ in $\mathcal{C}(\vec{w})$ is bounded by P^N. Indeed, elements $\vec{v} = (v_i)_{i=1}^n$ in $\mathcal{C}(\vec{w})$ are determined by N independent choices at pivotal times, with probability p_+ for $(v_{\alpha_i+(j-1)}^{-1} v_{\alpha_i+j})_{j=1}^L = (a_1, \ldots, a_n)$ and with probability p_- for $(v_{\alpha_i+(j-1)}^{-1} v_{\alpha_i+j})_{j=1}^L = (b_1, \ldots, b_n)$. Since $\vec{w}$ corresponds to a single outcome, its probability is at most P^N.

This implies that

$$\mathbb{P}((\omega_i)_{i=1}^n \in G_{n,f}(Q = 1)) = \sum_{w \in G_{n,f}(Q=1)} \mathbb{P}((\omega_i)_{i=1}^n = \vec{w})$$

$$\leq P^N \sum_{w \in G_{n,f}(Q=1)} \mathbb{P}((\omega_i)_{i=1}^n \in \mathcal{C}(\vec{w}))$$

$$\leq P^N,$$

where the disjointness of $\mathcal{C}(\vec{w})$ was used at the end. Similarly, we have

$$\mathbb{P}((\omega_i)_{i=1}^n \in G_{n,b}(Q = 1)) \leq P^N.$$

Then the Borel-Cantelli lemma guarantees that $\mathbb{P}$-a.e. sample path ω avoids G_n eventually. Pictorially, we have Figure 6.
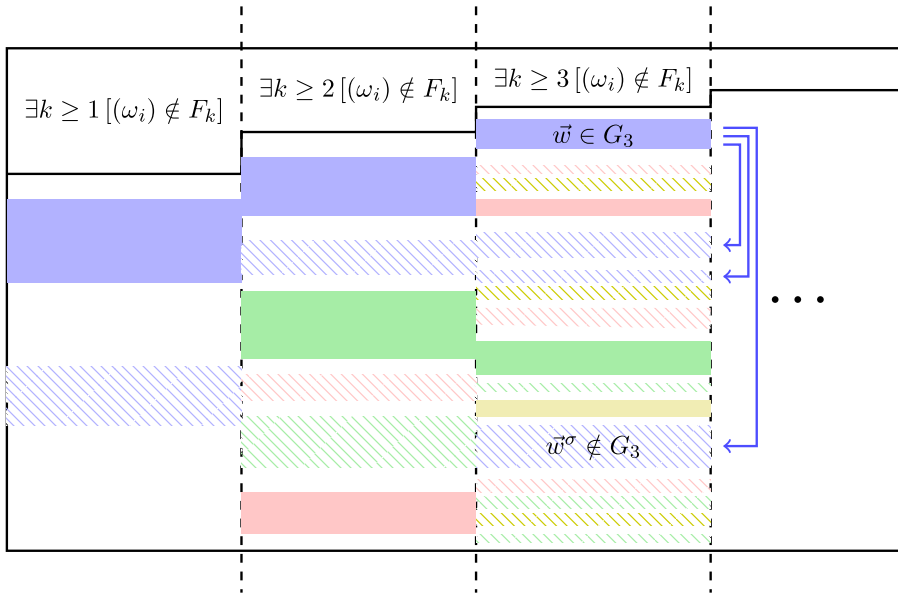
Figure 6. Schematic for F_n and G_n. Each colored region corresponds to $\{\omega : (\omega_i)_{i=1}^k \in F_k$ for $k \geq n$, $(\omega_i)_{i=1}^n = \vec{w} \in G_n\}$ and is copied inside $\{\omega : (\omega_i)_{i=1}^k \in F_k$ for $k \geq n\}$ by pivoting (hatched regions). Copies of distinct words in G_n (colored in distinct colors) are disjoint. The sum of the measures of colored regions is bounded. Note that $\{\omega : (\omega_i)_{i=1}^k \notin F_k$ for some $k \geq n\}$ decreases to a measure zero set.

For those paths, if $\tau(\omega_n) \leq \frac{1.5R\eta n}{L} < (2D - 2\eta/M)n$ happens infinitely often, then $(\omega_1, \ldots, \omega_n)$ should avoid F_n infinitely often. However, as mentioned before, almost every sample path avoids F_n eventually. Consequently, $\tau(\omega_n) \geq \frac{1.5R\eta n}{L}$ eventually holds for $\mathbb{P}$-a.e. (ω_n).

Now let us prove the second assertion. When μ has finite first moment, there exists a constant $0 < \lambda < +\infty$, called *drift*, satisfying that $\lambda = \lim_{k\to\infty} \frac{1}{k} d(x_0, \omega_k x_0)$ for $\mathbb{P}$-a.e. (ω_k). Set a threshold $0 < \epsilon < 1$. We slightly modify the definition of F_n and G_n as follows:

$$F_n := \left\{ \vec{w} : \frac{\#\mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}(\vec{w}) \geq \frac{\eta \epsilon n}{24L}, \, d(x_0, w_n x_0) \geq \left(1 - \frac{\epsilon}{1000}\right)\lambda n,}{d(x_0, w_{\lfloor \epsilon n/7 \rfloor} x_0) \leq \frac{\epsilon \lambda n}{6}} \right\}$$

$$G_n := \left\{ \vec{w} \in F_n : \tau(w_n) \leq \left(2D - \frac{2\eta}{M}\right)n \right\}.$$

We again invoke the subadditive ergodic theorem. Theorem 3.4 asserts that for a.e. path ω, there exists some $t(\omega)$ such that for all $i > t(\omega)$,

(1) $W_i \geq 0.999\eta i$,

(2) $d(x_0, \omega_i x_0) \geq (1 - \epsilon/1000)\lambda i$,

(3) $d(x_0, \omega_i x_0) \leq 1.001\lambda i$.

Note also that $\#N_{i,j}(\omega) \geq W_{\lfloor i/3L \rfloor}$ for any $i \leq j$. Considering this, such path ω will be contained in F_n for $n \geq 1000(L+1)t(\omega)/\epsilon$.

This time, we set

$$D := 0.5(1 - \epsilon/2),$$

$$M > 1 + 10R + 2d(x_0, w_+ x_0) + 2d(x_0, w_- x_0) + 4L + \frac{2000\eta}{\lambda\epsilon} + \frac{2000\eta}{\epsilon} + \frac{2000}{\epsilon}.$$

Note that Inequality 3.4 still holds. Considering this, Lemma 3.8, Lemma 3.10 and Lemma 3.15 also hold with the same proofs. The only missing step is Lemma 3.7. We now claim that $F_{n,f}(D, Q = 1) = F_n$. If $\vec{w} \in F_n$, then

$$\begin{aligned}
d(x_0, w_{\lfloor \epsilon n/7 \rfloor} x_0) &\leq \frac{\epsilon \lambda n}{6} \\
&\leq \frac{1}{2}\left(1 - \frac{\epsilon}{1000}\right)\lambda n - \left(\frac{1}{2}(1 - \epsilon/2)\lambda + \eta \cdot \frac{\epsilon \lambda}{2000\eta}\right)n \quad (3.9) \\
&\leq \frac{1}{2}d(x_0, w_n x_0) - \left(D + \frac{\eta}{M}\right)n.
\end{aligned}$$

Let us denote $\mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}(\vec{w}) = \{n_1 < \ldots < n_k\}$. We have $k \geq \eta \epsilon n / 24L$ (since $\vec{w} \in F_n$) and $\mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}(\vec{w}) \subseteq \mathcal{N}_{\lfloor \epsilon n/7 \rfloor, \lfloor \epsilon n/7 \rfloor}(\vec{w})$. Let

$$y_{2i-1} = x_{n_i - 2L}, y_{2i} = x_{n_i - L}$$

for $i = 1, \ldots, k$ and $y_{2k+1} = x_{\lfloor \epsilon n/7 \rfloor}$. Then, Corollary 3.6 tells us that

$$d(x_0, x_{n_i - L}) \leq d(x_0, w_{\lfloor \epsilon n/7 \rfloor} x_0) \leq \frac{1}{2}d(x_0, w_n x_0) - \left(D + \frac{\eta}{M}\right)n$$

for each $n_i \in \mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}$. Hence, these joints contribute to $\mathcal{N}_f(\vec{w})$. We then have $\vec{w} \in F_{n,f}$ because

$$\frac{\eta n}{20LM^2} \leq \frac{\eta n}{20LM} \leq \frac{\eta n}{20L} \cdot \frac{\epsilon}{2000} \leq \frac{\eta \epsilon n}{24L}.$$

Given that the lemmata still holds true, the previous argument to deduce that almost every path ω does not fall into G_n infinitely often. Note also that

$$2D - \frac{2\eta}{M} \geq (1 - \epsilon)\lambda.$$

Together with the observation that a.e. path does not avoid F_n infinitely often, we deduce that $\liminf \frac{1}{n}\tau(\omega_n) \geq (1 - \epsilon)\lambda$ almost surely. By setting $\epsilon = 1/k$ and taking intersection of those events for $k \in \mathbb{Z}_{>1}$, we deduce that $\liminf \frac{1}{n}\tau(\omega_n) \geq \lambda$ almost surely. Since $\frac{1}{n}\tau(\omega_n) \leq \frac{1}{n}d(x_0, \omega_n x_0)$ and the latter one converges to λ a.s., we finally conclude that $\frac{1}{n}\tau(\omega_n) \to \lambda$ a.s. □

4. Teichmüller space and its geometry

4.1. Basic notions

In this section, we investigate the geometry of the Teichmüller space (X, d) equipped with the Teichmüller metric. Let S be a closed orientable surface of genus at least

2. Its *Teichmüller space* is the space of equivalence classes $[(f, \Sigma)]$ of an orientation preserving homeomorphism $f : S \to \Sigma$ to a hyperbolic surface Σ, where (f, Σ) and (g, Σ') are equivalent if there is an isometry $i : \Sigma \to \Sigma'$ so that $i \circ f$ is homotopic to g. The Teichmüller space admits a metric d called the *Teichmüller metric* defined by $d([(f, \Sigma)], [(g, \Sigma')]) = \frac{1}{2} \inf_\varphi \log K_\varphi$, where the infimum is taken over quasiconformal homeomorphisms φ with $f = \varphi \circ g$ up to homotopy and K_φ is its quasiconformal constant. Then the mapping class group $\mathrm{Mod}(S)$ of the surface acts isometrically on its Teichmüller space $X = \mathcal{T}(S)$. We denote $G = \mathrm{Mod}(S)$ for the rest of the paper. For details on the Teichmüller geometry, see [Hub06], [Pap07], [IT92].

We recall the notion of the extremal length of a curve.

Definition 4.1 (Extremal length). For a point $x \in X$ in the Teichmüller space and an isotopy class α of simple closed curves on the underlying surface S, the *extremal length* is defined as

$$Ext_x(\alpha) := \sup_{\sigma \in [x]} \frac{l_\sigma^2(\alpha)}{\mathrm{area}(\sigma)}.$$

Here, σ is a Riemannian metric in the conformal class x and $l_\sigma(\alpha)$ is the length of α measured by σ.

Kerckhoff [Ker80] characterized the Teichmüller metric in terms of the extremal length:

$$d(x, y) = \frac{1}{2} \log \sup_{\alpha \in \mathcal{C}_0(S)} \frac{Ext_y(\alpha)}{Ext_x(\alpha)},$$

where $\mathcal{C}_0(S)$ is a space of isotopy classes of essential simple closed curves on S.

We denote the ϵ-thick part of the Teichmüller space X by $X_{\geq \epsilon}$. That is, $X_{\geq \epsilon}$ consists of surfaces whose shortest extremal length is at least ϵ. From Kerckhoff's formula, $x \in X_{\geq \epsilon}$ implies $y \in X_{\geq \epsilon'}$ where $\epsilon' = \epsilon e^{-2d(x,y)}$.

A *geodesic segment* on X refers to an isometric embedding $\Gamma : I = [a, b] \to X$ of a closed interval $[a, b]$ into X (i.e., $d_X(\Gamma(t_1), \Gamma(t_2)) = |t_1 - t_2|$ for all $t_1, t_2 \in I$). We make an abuse of notation: the image $\Gamma([a, b])$ of Γ is also called the geodesic segment connecting $\Gamma(a)$ and $\Gamma(b)$ and denoted by $[\Gamma(a), \Gamma(b)]$. Note that for each $x, y \in X$, $[x, y]$ uniquely exists by Teichmüller's theorem. We say that a segment $\Gamma' : J = [c, d] \to X$ is a *subsegment* of $\Gamma : I = [a, b] \to X$ if $\Gamma|_J = \Gamma'$. In this situation, we also say that $[\Gamma'(c), \Gamma'(d)]$ is a subsegment of $[\Gamma(a), \Gamma(b)]$. For $\epsilon > 0$, we also call a segment ϵ-thick if it is contained in the ϵ-thick part.

For subsets $A, B \subseteq X$ and $R > 0$, we define

$$N_R(A) := \{y \in X : \text{there exists } x \in A \text{ such that } d(x, y) < R\},$$
$$d(A, B) := \inf\{d(x, y) : x \in A, y \in B\}, \text{ and}$$
$$d_H(A, B) := \inf\{R > 0 : A \subseteq N_R(B) \text{ and } B \subseteq N_R(A)\}.$$

4.2. Curve complex

We temporarily digress to the relationship between the Teichmüller geometry and the curve complex. The payback for this is Theorem 4.3, which guarantees the existence of persistent joints.

The *curve complex* $Y = Y(S)$ of a surface S is a simplicial complex whose vertices are isotopy classes of essential simple closed curves and each $(k+1)$-simplex corresponds to $k+1$ vertices represented by disjoint simple closed curves. The curve complex was first introduced by Harvey [Har81], and it endows with a natural metric that each edge has a length 1.

There exists a projection $\pi : X \to Y$ from Teichmüller space to the curve complex that associates each surface $x \in X$ with the curve $\pi(x)$ with the shortest extremal length. Masur and Minsky [MM99] showed that π is coarsely $\mathrm{Mod}(S)$-equivariant and coarsely Lipschitz. That is, there exists a constant $A > 0$ such that

$$d_{Y(S)}(g\pi(x), \pi(gx)) \leq A, \quad d_{Y(S)}(\pi(x), \pi(y)) \leq A d_X(x,y) + A$$

for each $x, y \in X$ and $g \in \mathrm{Mod}(S)$. Furthermore, there exists a constant $K > 0$ such that every geodesic γ in the Teichmüller space X descends to a K-quasi-geodesic $\pi \circ \gamma$ in the curve complex $Y(S)$ up to reparametrization. Finally, since $Y(S)$ is Gromov hyperbolic, the following assertions hold for some $A' > 0$.

(1) If γ, η are K-quasi-geodesics with the same endpoints, then $d_H(\gamma, \eta) < A'$.

(2) Given $x, y, z \in Y(S)$, $d(y, [x,z]) \geq d(x,y) - (y,z)_x - A'$.

Thurston [Thu88] compactified the Teichmüller space by introducing $\mathcal{PMF}$, the projective space of measured foliations, as the boundary of the Teichmüller space. Pseudo-Anosov maps exhibit source-sink dynamics on this compactified Teichmüller space as follows. Let $\mathcal{UE} \subseteq \mathcal{PMF}$ be the set of uniquely ergodic measured foliations. Then for each pseudo-Anosov ϕ, there exists distinct foliations $\phi^{+\infty}, \phi^{-\infty} \in \mathcal{UE}$ such that $\lim_{n \to +\infty} \phi^{\pm n} x = \phi^{\pm \infty}$ for any $x \in X$.

Moreover, Kaimanovich and Masur [KM96] proved the boundary convergence of random walks on the Teichmüller space. More precisely, for non-elementary random walks on mapping class groups, almost every orbit on the Teichmüller space converges to $\mathcal{UE} \subset \mathcal{PMF}$, the space of uniquely ergodic measured foliations.

Klarreich [Kla22] and Hamenstädt [Ham06] also proved that the boundary of the curve complex can be identified with $\widetilde{\mathcal{MIN}}$, the space of equivalence classes of minimal topological foliations. As such, we can define a map $\pi_\infty : \mathcal{UE} \to \widetilde{\mathcal{MIN}}$ which forgets the measure structure and quotients by topological equivalence. This map π_∞ is injective, and further, if a sequence $x_n \in X$ converges to a point $x_\infty \in \mathcal{UE}$, then the corresponding $\pi(x_n)$ converges to a point $\pi_\infty(x_\infty)$.

We now observe that two sequences on Teichmüller space diverge from each other if they are heading to distinct uniquely ergodic foliations.

Proposition 4.2. *Let ϕ be a pseudo-Anosov mapping class and $\{z_k\}_{k \geq 0}$ be a sequence in X that converges to a uniquely ergodic foliation ξ other than $\phi^{+\infty}$. Then for each $C > 0$, there exists $F = F(C, \phi, \{z_k\}_k) > 0$ such that $d(\phi^m x_0, [x_0, z_k]) > C$ for all $k \geq 0$ and $m \geq F$.*

Proof. Since $\{z_k\}_k$ and $\{\phi^n x_0\}_n$ are both heading to points in $\mathcal{UE}$, we can descend them to the curve complex; $\pi(z_k) \to \pi_\infty(\xi)$ and $\pi(\phi^n x_0) \to \pi_\infty(\phi^{+\infty})$. Since ξ and $\phi^{+\infty}$ are distinct, we have

$$P := \sup_{k,n}\{(\pi(\phi^n x_0), \pi(z_k))_{\pi(x_0)}\} < \infty.$$

Take F such that

$$d(\pi(x_0), \pi(\phi^m x_0)) > P + 2A' + A(C+1)$$

for any $m \geq F$. Suppose to the contrary that $d(\phi^m x_0, [x_0, z_k]) \leq C$ for some $m \geq F$ and $k \geq 0$. Let $x^\dagger \in [x_0, z_k]$ be the closest point to $\phi^m x_0$.

Note that $[\pi(x_0), \pi(z_k)]$ and $\pi([x_0, z_k])$ are (unparametrized) K-quasi-geodesics on $Y(S)$ with the same endpoints. Hence, we have

$$d_H([\pi(x_0), \pi(z_k)], \pi([x_0, z_k])) < A'.$$

Moreover, since the projection π is coarsely Lipschitz, we have

$$d(\pi(x^\dagger), \pi(\phi^m x_0)) \leq Ad(x^\dagger, \phi^m x_0) + A \leq A(C+1).$$

Combining these yields

$$d([\pi(x_0), \pi(z_k)], \pi(\phi^m x_0)) \leq d_H([\pi(x_0), \pi(z_k)], \pi([x_0, z_k])) + d(\pi(x^\dagger), \pi(\phi^m x_0))$$
$$\leq A' + A(C+1).$$

Meanwhile, we also have

$$d([\pi(x_0), \pi(z_k)], \pi(\phi^m x_0)) \geq d(\pi(x_0), \pi(\phi^m x_0)) - (\pi(\phi^m x_0), \pi(z_k))_{\pi(x_0)} - A'$$
$$> (P + 2A' + A(C+1)) - P - A'$$
$$= A' + A(C+1),$$

a contradiction. $\qquad\qquad\square$

Proposition 4.2 leads to the following probabilistic observation. Recall that ω is a non-elementary random walk on the mapping class group.

Theorem 4.3. *Let ϕ be a pseudo-Anosov mapping class. Then for each $C > 0$, there exists a constant $F = F(C, \phi) > 0$ such that*

$$\mathbb{P}\left((\omega_n)_{n\in\mathbb{Z}} : \begin{array}{c} \exists m, n \in \mathbb{Z} \text{ such that } |m| \geq F \text{ and} \\ d(\phi^m x_0, [x_0, \omega_n x_0]) < C \end{array}\right) < 1.$$

Proof. Suppose not. Then for each N, for almost every sample path $\omega = (\omega_i)_i$ there exists $m, n \in \mathbb{Z}$ such that $|m| \geq N$ and $d(\phi^m x_0, [x_0, \omega_n x_0]) < C$. Taking intersection of these events, we deduce that

$$\mathbb{P}\left(\mathcal{E} := \left\{(\omega_n)_{n\in\mathbb{Z}} : \begin{array}{c} \exists (n_i), (m_i) \text{ such that } \lim_{i\to\infty} |m_i| = +\infty \text{ and} \\ d(\phi^{m_i} x_0, [x_0, \omega_{n_i} x_0]) < C \end{array}\right\}\right) = 1. \quad (4.1)$$

Meanwhile, recall the result of Kaimanovich and Masur that a.e. path ω have boundary points $\xi^+(\omega), \xi^-(\omega) \in \mathcal{UE}$ such that $\lim_{n \to \pm\infty} \omega_n x_0 = \xi^\pm(\omega)$. Since the forward and the backward stationary measure does not have atoms, we have

$$\mathbb{P}\left(\mathcal{E}' := \left\{\omega : \{\xi^+(\omega), \xi^-(\omega)\} \cap \{\phi^{+\infty}(\omega), \phi^{-\infty}(\omega)\} = \emptyset\right\}\right) = 1.$$

Finally, Proposition 4.2 tells us that $\mathcal{E} \cap \mathcal{E}' = \emptyset$. This implies that $\mathbb{P}(\mathcal{E} \cup \mathcal{E}') = \mathbb{P}(\mathcal{E}) + \mathbb{P}(\mathcal{E}') = 2 > 1$, a contradiction. $\square$

4.3. Fellow traveling

Let $\gamma : [0, L] \to X$ and $\gamma' : [0, L'] \to X$ be paths on X. We say that γ and γ' ϵ-*fellow travel* if $d(\gamma(kL), \gamma'(kL')) < \epsilon$ for each $0 \leq k \leq 1$. We remark that we always stick to the arclength parametrization when discussing geodesics on X. The following is a direct observation.

Lemma 4.4. *Let $\gamma^{(i)} : [0, L_i] \to X$ be arcs on X for $i = 1, 2, 3$. Suppose that $\gamma^{(1)}$ and $\gamma^{(2)}$ ϵ-fellow travel, and $\gamma^{(2)}$ and $\gamma^{(3)}$ ϵ'-fellow travel. Then*

(1) *$\gamma^{(1)}|_{[kL_1, k'L_1]}$ and $\gamma^{(2)}|_{[kL_2, k'L_2]}$ ϵ-fellow travel for each $0 \leq k \leq k' \leq 1$. In particular, any initial (terminal, resp.) subarc of γ_1 ϵ-fellow travel with an initial (terminal, resp.) subarc of γ_2.*

(2) *γ_1 and γ_3 $(\epsilon + \epsilon')$-fellow travel.*

In Gromov hyperbolic spaces, two geodesics fellow travel if their endpoints are pairwise near. In contrast, such geodesics in Teichmüller spaces need not fellow travel since Teichmüller spaces are not Gromov hyperbolic in general. Indeed, Rafi [Raf14] presented examples of pairs of geodesics whose endpoints are pairwise near while they are not fellow traveling with a uniform constant. Nonetheless, Rafi also proved the fellow traveling phenomenon of geodesics with near endpoints, given that the endpoints are lying on an ϵ-thick part.

Theorem 4.5 [Raf14, Theorem 7.1]. *There exists a constant $\mathscr{B}_0(\epsilon)$ satisfying the following. For $x, x', y, y' \in X_{\geq \epsilon}$ such that*

$$d(x, x') \leq 1 \quad and \quad d(y, y') \leq 1,$$

$[x, y]$ and $[x', y']$ $\mathscr{B}_0(\epsilon)$-fellow travel.

From this, one can also see the following.

Corollary 4.6. *For each $C > 0$, there exists a constant $\mathscr{B}(\epsilon, C) > C$ satisfying the following. For all $x, y \in X_{\geq \epsilon}$ and all $x', y' \in X$ such that*

$$d(x, x') \leq C \quad and \quad d(y, y') \leq C,$$

$[x, y]$ and $[x', y']$ $\mathscr{B}(\epsilon, C)$-fellow travel.

Proof. Let $\{x_t\}$ ($\{y_t\}$, resp.) be a segment connecting x and x' (y and y', resp.) with speed less than 1 and $x_0 = x$, $x_C = x'$, $y_0 = y$ and $y_C = y'$. Then each of x_t, y_t is ϵe^{-2t}-thick. Thus, $[x_t, y_t]$ and $[x_{t+1}, y_{t+1}]$ $\mathscr{B}_0(\epsilon e^{-2(t+1)})$-fellow travel. This implies that $[x, y]$ and $[x', y']$ $(\sum_{t=1}^{\lceil C \rceil} \mathscr{B}_0(\epsilon e^{-2t}))$-fellow travel. $\square$

Rafi also proved that geodesic triangles in a thick part of the Teichmüller space are thin.

Theorem 4.7 [Raf14, Theorem 8.1]. *There exist constants $\mathscr{C}_0(\epsilon)$ and $\mathscr{D}_0(\epsilon)$ such that the following holds. Let $x, y, z \in X_{\geq \epsilon}$ and suppose that the geodesic $[x,y]$ contains a segment $\gamma \subseteq X_{\geq \epsilon}$ of length at least $\mathscr{C}_0(\epsilon)$. Then there exists a point $w \in \gamma$ such that*

$$\min \{d(w, [x,z]), d(w, [z,y])\} < \mathscr{D}_0(\epsilon).$$

From now on, we fix a point $x_0 \in X$ as the basepoint. Since the random walk ω is generated by a non-elementary probability measure μ, there exist two independent pseudo-Anosovs ϕ_+, ϕ_- in $\langle\langle \operatorname{supp} \mu \rangle\rangle$, the subsemigroup generated by the support of μ. By taking suitable powers, we may assume that they are made of equal numbers of elements in $\operatorname{supp} \mu$.

Let $\Gamma(\phi_\pm)$ be the invariant axis of $\phi_\pm$ on X, respectively. We fix points $y_+ \in \Gamma(\phi_+)$ and $y_- \in \Gamma(\phi_-)$. We also let $\tau_\pm$ be the translation length of $\phi_\pm$ on X, respectively.

Lemma 4.8. *There exists $\mathscr{M}, \epsilon > 0$ such that the following holds. Let $\phi \in \{\phi_+, \phi_-\}$ and $n \leq m$. Then*

  (1) $[\phi^n x_0, \phi^m x_0]$ *is ϵ-thick, and*

  (2) $\{\phi^i x_0 : n \leq i \leq m\}$ *is contained in the $\mathscr{M}$-neighborhood of $[\phi^n x_0, \phi^m x_0]$.*

Proof. Let us discuss the case for $\phi = \phi_+$. Let $\epsilon_0 > 0$ be such that x_0, y_+ are ϵ_0-thick.

First, $\{\phi_+^i x_0 : n \leq i \leq m\}$ and $\{\phi_+^i y_+ : n \leq i \leq m\}$ are within Hausdorff distance $d(x_0, y_+)$. Moreover, $\{\phi_+^i y_+ : n \leq i \leq m\}$ periodically appear on the Teichmüller geodesic $[\phi_+^n y_+, \phi_+^m y_+]$ with period τ_+. Hence, they are within Hausdorff distance τ_+, and $[\phi_+^n y_+, \phi_+^m y_+]$ is $\epsilon_0 e^{-2\tau_+}$-thick.

Finally, note that $\phi_+^n y_+, \phi_+^m y_+$ are ϵ_0-thick and $d(\phi_+^n y_+, \phi_+^n x_0), d(\phi_+^m y_+, \phi_+^m x_0) < d(x_0, y_+)$. Corollary 4.6 then tells us that the Hausdorff distance between $[\phi_+^n y_+, \phi_+^m y_+]$ and $[\phi_+^n x_0, \phi_+^m x_0]$ is bounded by $\mathscr{B}(\epsilon_0, d(x_0, y_+))$. $\qquad\square$

Remark 4.9. We have actually proved something stronger, namely, that the Hausdorff distance between $[\phi^n x_0, \phi^m x_0]$ and $\{\phi^i x_0 : n \leq i \leq m\}$ is bounded by $\mathscr{M}$. We stick to the weaker conclusion as in Lemma 4.8 since it still holds true after replacing $\phi_\pm$ with their powers.

Lemma 4.10. *For each $C > 0$, there exists a constant $\mathscr{G}(C) > 0$ such that $d(\phi_+^k y_+, \phi_-^m y_-) \leq C$ implies $\max(|k|, |m|) \leq \mathscr{G}(C)$.*

Proof. Suppose not. Without loss of generality, for each $i \in \mathbb{Z}_{>0}$, let $k_i, m_i \in \mathbb{Z}$ be such that $|k_i| > i$ and $d(\phi_+^{k_i} y_+, \phi_-^{m_i} y_-) \leq C$. Note that

$$d(x_0, \phi_+^{k_i} y_+) \geq d(y_+, \phi_+^{k_i} y_+) - d(y_+, x_0) \geq |k_i| \tau_+ - d(y_+, x_0)$$

tends to infinity. Since $d(\phi_+^{k_i} y_+, \phi_-^{m_i} y_-) < C$, $d(x_0, \phi_-^{m_i} y_-)$ also tends to infinity. Possibly after passing to subsequences, this implies that $\phi_+^{k_i} y_+$ approaches an endpoint of $\Gamma(\phi_+)$ and $\phi_-^{m_i} y_-$ approaches an endpoint of $\Gamma(\phi_-)$. Since $d(\phi_+^{k_i} y_+, \phi_-^{m_i} y_-)$ is bounded, this

implies that those endpoints are identical ([KM96, Lemma 1.4.2]); this contradicts the independence of ϕ_+ and ϕ_-. $\qquad\square$

Lemma 4.11. *There exists $C_{Grom} > 0$ such that the following holds. For any pair of distinct elements ϕ, ψ of $\{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ and $m, n \geq 0$, we have $(\phi^n x_0, \psi^m x_0)_{x_0} < C_{Grom}$.*

Proof. Let $\mathscr{M}, \epsilon > 0$ be as in Lemma 4.8 and let $\mathscr{C}_0(\epsilon), \mathscr{D}_0(\epsilon)$ be as in Theorem 4.7. Finally, let $\mathscr{G} = \mathscr{G}(\mathscr{D}_0(\epsilon) + 2\mathscr{M})$ be as in Lemma 4.10.

When $\phi = \psi^{-1}$, the conclusion follows from the fact that ϕ_+, ϕ_- have positive translation lengths. In particular, we have

$$|d(\phi_\pm^m x_0, \phi_\pm^n x_0) - \tau_\pm |m - n|| \leq 2d(y_\pm, x_0)$$

for any integers $m, n \in \mathbb{Z}$. This leads to

$$|d(\phi_\pm^m x_0, \phi_\pm^{-n} x_0) - d(\phi_\pm^m x_0, x_0) - d(\phi_\pm^{-n} x_0, x_0)| \leq 6d(y_\pm, x_0)$$

for $m, n \geq 0$.

Let us now fix $m, n \in \mathbb{Z}$ and estimate $(\phi_+^m x_0, \phi_-^n x_0)_{x_0}$. If $d(x_0, \phi_+^m x_0)$ is smaller than $\mathscr{G}\tau_+ + 2d(x_0, y_+) + \mathscr{M} + 1 + \mathscr{C}_0(\epsilon)$, so is the Gromov product. If not, we take the subsegment η of $[x_0, \phi_+^m x_0]$ with length $\mathscr{C}_0(\epsilon)$ such that $d(x_0, \eta) = \mathscr{G}\tau_+ + 2d(x_0, y_+) + \mathscr{M} + 1$. By Theorem 4.7, either $[x_0, \phi_-^m x_0]$ or $[\phi_+^m x_0, \phi_-^n x_0]$ intersects the $\mathscr{D}_0(\epsilon)$-neighborhood of η.

In the former case, let $p \in [x_0, \phi_-^n x_0]$ and $q \in \eta$ be such that $d(p, q) \leq \mathscr{D}_0(\epsilon)$. By Lemma 4.8, there exists m', n' such that $\phi_+^{m'} x_0 \in N_{\mathscr{M}}(p)$ and $\phi_-^{n'} x_0 \in N_{\mathscr{M}}(q)$. Then we have $d(\phi_+^{m'} x_0, \phi_-^{n'} x_0) \leq \mathscr{D}_0(\epsilon) + 2\mathscr{M}$ for some $m', n' \in \mathbb{Z}$ such that $d(x_0, \phi_+^{m'} x_0) \geq d(x_0, \eta) - \mathscr{M} > \mathscr{G}\tau_+ + 2d(x_0, y_+)$. This contradicts Lemma 4.10.

Hence, the latter case holds: there exists $p \in [\phi_+^m x_0, \phi_-^n x_0]$ that belongs to the $\mathscr{D}_0(\epsilon)$-neighborhood of η. This implies that

$$2(\phi_+^m x_0, \phi_-^n x_0)_{x_0} = [d(x_0, \phi_+^m x_0) - d(\phi_+^m x_0, p)] + [d(x_0, \phi_-^n x_0) - d(p, \phi_-^n x_0)]$$
$$\leq 2d(x_0, p) \leq 2d(x_0, \eta) + 2\operatorname{diam}(\eta) + 2\mathscr{D}_0(\epsilon).$$

In conclusion, we have $(\phi_+^m x_0, \phi_-^n x_0)_{x_0} \leq \mathscr{G}\tau_+ + 2d(x_0, y_+) + \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M} + 1$. $\quad\square$

4.4. Witnessing

Recall that G stands for the mapping class group of the underlying surface.

Definition 4.12. Let $D > 0$, γ, γ' be paths on X. We say that γ is *D-witnessed by γ'* if there exists a subsegment η of γ that D-fellow travels with γ'. Here, if γ and γ' share the beginning point (ending point, resp.), we additionally require η to begin at (end at, resp.) that shared point.

For $w, w' \in G$, we say that

- w is D-witnessed by γ' if $[x_0, wx_0]$ is D-witnessed by γ';
- γ is D-witnessed by w' if γ is D-witnessed by $[x_0, w'x_0]$, and
- w is D-witnessed by w' if $[x_0, wx_0]$ is D-witnessed by $[x_0, w'x_0]$.

We also define

$$\mathcal{C}_D(\phi \to 0) := \{w \in G : w \text{ is } D\text{-witnessed by } \phi\},$$
$$\mathcal{C}_D(0 \to \phi) := \{w \in G : w^{-1} \text{ is } D\text{-witnessed by } \phi^{-1}\},$$
$$\mathcal{C}_D(\phi \to \psi) := \mathcal{C}_D(\phi \to 0) \cap \mathcal{C}_D(0 \to \psi).$$

The following lemma will tell us that sample paths are witnessed by $\phi_\pm$ for a definite probability.

Lemma 4.13. *For each $F > 0$, there exists $\mathscr{F} = \mathscr{F}(F)$ such that the following holds. Let $\phi \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ and $w \in G$. If $d(\phi^F x_0, [x_0, w x_0]) > \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M}$, then $\phi^{-m} w$ is $\mathscr{F}$-witnessed by ϕ^{-m} for every $m \geq 0$.*

Proof. Let $K_1 > 0$ be such that $d(x_0, \phi^i x_0) > \mathscr{M} + \mathscr{C}_0(\epsilon)$ for each $\phi \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ and $i \geq K_1$. Then, let $K_2 > 0$ be such that $d(x_0, \phi^i x_0) \leq K_2$ for each $\phi \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ and $0 \leq i \leq F + K_1$. We then take

$$\mathscr{F} = K_1 + K_2 + \mathscr{B}(\epsilon, \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M} + K_2) + F.$$

Note that for any $x, y, z \in X$, $[x, y]$ is $d(x, z)$-witnessed by $[x, z]$. Since $\mathscr{F} \geq K_2$, $\phi^{-m} w$ is $\mathscr{F}$-witnessed by ϕ^{-m} for $0 \leq m \leq F + K_1$.

Let us now consider the case of $m > F + K_1$. Let y be a point on $[\phi^{-m} x_0, x_0]$ that lies in the $\mathscr{M}$-neighborhood of $\phi^{F-m} x_0$. Note that $d(y, x_0) \geq d(\phi^{m-F} x_0, x_0) - \mathscr{M} \geq \mathscr{C}_0(\epsilon)$. Hence, we can take a subsegment $\eta = [y, y']$ of $[\phi^{-m} x_0, x_0]$ such that $d(y, y') = \mathscr{C}_0(\epsilon)$.

By Theorem 4.7, either $[\phi^{-m} w x_0, x_0]$ or $[\phi^{-m} w x_0, \phi^{-m} x_0]$ intersects the $\mathscr{D}_0(\epsilon)$-neighborhood of η. If $d([\phi^{-m} w x_0, \phi^{-m} x_0], \eta) < \mathscr{D}_0(\epsilon)$, then we deduce $d(\phi^{F-m} x_0, [\phi^{-m} w x_0, \phi^{-m} x_0]) < \mathscr{D}_0(\epsilon) + \mathscr{M} + \mathscr{C}_0(\epsilon)$. This is equivalent to $d(\phi^F x_0, [x_0, w x_0]) < \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M}$, which contradicts the assumption.

Hence, $[\phi^{-m} w x_0, x_0]$ contains a point p that lies in the $\mathscr{D}_0(\epsilon)$-neighborhood of η. Note that

$$d(p, \phi^{-m} x_0) \leq d(p, \eta) + \operatorname{diam}(\eta) + \mathscr{M} + d(\phi^{F-m} x_0, \phi^{-m} x_0)$$
$$\leq \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M} + K_2.$$

Then $[p, x_0]$ and $[\phi^{-1} x_0, x_0]$ $\mathscr{B}(\epsilon, \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M} + K_2)$-fellow travel as desired. $\square$

Corollary 4.14. *There exists $\mathscr{F} > 0$ such that for $\phi \in \{\phi_+, \phi_-\}$, we have*

$$\mathbb{P}((\omega_n)_{n \in \mathbb{Z}} : \phi^m \omega_n \text{ is } \mathscr{F}\text{-witnessed by } \phi^m \text{ for all } m, n \in \mathbb{Z}) > 0.$$

Proof. Let $F = F(\mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M}, \phi)$ be as in Theorem 4.3 and $\mathscr{F} = \mathscr{F}(F)$ be as in Lemma 4.13. We consider the event $\mathcal{E}$ of sample paths $\omega = (\omega_n)_n$ such that

$$d(\phi^F x_0, [x_0, \omega_n x_0]), d(\phi^{-F} x_0, [x_0, \omega_n x_0]) > \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M}$$

for all $n \in \mathbb{Z}$. Then $\mathbb{P}(\mathcal{E}) > 0$ thanks to Theorem 4.3. Moreover, Lemma 4.13 implies that for $\omega \in \mathcal{E}$, $\phi^m \omega_n$ is $\mathscr{F}$-witnessed by ϕ^m for each $m \in \mathbb{Z}$, as desired. $\square$

We now fix some constants. As the base case, we set

$$\mathscr{D}_1 := \mathscr{D}_0(\epsilon) + \mathscr{B}(\epsilon, \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M}) + \mathscr{F},$$
$$\epsilon_1 := \epsilon e^{-2\mathscr{D}_1}.$$

Now given $\mathscr{D}_j$ and ϵ_j, we define

$$\mathscr{D}_{j+1} := \mathscr{D}_j + \mathscr{B}(\epsilon, 4\mathscr{D}_j + 2C_{Grom} + 5\mathscr{C}_0(\epsilon_j) + 5\mathscr{D}_0(\epsilon_j) + 2),$$
$$\epsilon_{j+1} := \epsilon e^{-2\mathscr{D}_{j+1}}. \tag{4.2}$$

Note that the constants $\epsilon, \mathscr{M}, \mathscr{F}, C_{Grom}$ remain the same even if we replace ϕ_+, ϕ_- with their powers. Hence, by employing sufficiently large powers of $\phi_\pm$ if necessary, we may assume

$$d(x_0, \phi_\pm x_0) \geq 6\mathscr{D}_j + 2C_{Grom} + 2\mathscr{C}_0(\epsilon_j) + 2\mathscr{D}_0(\epsilon_j) + 2 \tag{4.3}$$

for $j = 1, \ldots, 6$. Finally, we set

$$Z = \mathscr{C}_0(\epsilon_3) + \mathscr{D}_0(\epsilon_3) + 2[\mathscr{D}_6 + d(x_0, \phi_+ x_0) + d(x_0, \phi_- x_0)] + 1. \tag{4.4}$$

Lemma 4.15. *For any $\phi \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ and $n > 0$, ϕ^n belongs to $\mathcal{C}_{\mathscr{D}_1}(\phi \to \phi)$.*

Proof. By Lemma 4.8, we have $p \in [x_0, \phi^n x_0]$ such that

$$d(p, \phi x_0) < \mathscr{M} < \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M}.$$

Since x_0 and ϕx_0 are ϵ-thick, Corollary 4.6 tells us that $[x_0, p]$ and $[x_0, \phi x_0]$ $\mathscr{B}(\epsilon, \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M})$-fellow travel. By a similar reason, there exists $q \in [x_0, \phi^n x_0]$ such that $[q, \phi^n x_0]$ and $[\phi^{n-1} x_0, \phi^n x_0]$ $\mathscr{B}(\epsilon, \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + \mathscr{M})$-fellow travel. $\square$

Remark 4.16. We will observe that if $\phi, \psi \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ are distinct, then $\mathcal{C}_{\mathscr{D}_5}(\phi \to 0) \cap \mathcal{C}_{\mathscr{D}_5}(\psi \to 0) = \emptyset$.

We can now discuss the concatenation of witnessed mapping classes.

Definition 4.17. Let $\phi_i, \psi_i \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$. We say that the sequences $(\phi_i)_i$ and $(\psi_i)_i$ are *repulsive* if $\phi_{i+1} \neq \psi_i^{-1}$ for each i.

Given repulsive sequences $(\phi_i)_i$ and $(\psi_i)_i$, we say that a sequence $(g_i)_{i=1}^n \subseteq G$ is *D-marked with* $(\phi_i)_{i=2}^n$ and $(\psi_i)_{i=1}^{n-1}$ if

(1) $g_1 \in \mathcal{C}_D(0 \to \psi_1)$,
(2) $g_i \in \mathcal{C}_D(\phi_i \to \psi_i)$ for $i = 2, \ldots, n-1$,
(3) $g_n \in \mathcal{C}_D(\phi_n \to 0)$.

If (g_i) additionally satisfies that $g_1 \in \mathcal{C}_D(\phi_1 \to \psi_1)$, then we say that $(g_i)_{i=1}^n$ is *D-strongly marked with* $(\phi_i)_{i=1}^n$ and $(\psi_i)_{i=1}^{n-1}$.

Lemma 4.18. *For each $j = 1, \ldots, 5$, we have the following.*

Let $(\phi_i)_i$, $(\psi_i)_i$ be repulsive sequences and $(g_i)_{i=1}^n$ be a sequence that is $\mathscr{D}_j$-marked with $(\phi_i)_{i=2}^n, (\psi_i)_{i=1}^{n-1}$. Let also $w := g_1 \cdots g_n$. Then w is $\mathscr{D}_{j+1}$-witnessed by $[g_1 \psi_1^{-1} x_0, g_1 x_0]$. If, moreover, $(g_i)_i$ is $\mathscr{D}_j$-strongly marked with $(\phi_i)_{i=1}^n$, $(\psi_i)_{i=1}^{n-1}$, then w is $\mathscr{D}_{j+1}$-witnessed by ϕ_1.

Proof. Let C_{Grom} be the constant obtained from Lemma 4.11 and let $K = \mathscr{C}_0(\epsilon_j) + \mathscr{D}_0(\epsilon_j) + 2\mathscr{D}_j + C_{Grom} + 1$. Recall that we have assumed in Condition 4.3 that $d(x_0, \phi^{\pm} x_0) \geq 2K + 2\mathscr{D}_{j+1}$.

We induct on the number of segments. For $n = 1$, the conclusion follows from Lemma 4.13 since $\mathscr{D}_j \leq \mathscr{D}_{j+1}$.

Now suppose that the theorem holds for $(g_2, \ldots, g_n)$. By induction hypothesis, there exists $q_2 \in [g_1 x_0, w x_0]$ such that $[g_1 x_0, q_2]$ and $[g_1 x_0, g_1 \phi_2 x_0]$ $\mathscr{D}_{j+1}$-fellow travel. Note that the length of $[g_1 x_0, q_2]$ is at least $d(x_0, \phi_2 x_0) - 2\mathscr{D}_{j+1} \geq 2K$.

There exists $q_1 \in [x_0, g_1 x_0]$ such that $[g_1 x_0, q_1]$ and $[g_1 x_0, g_1 \psi_1^{-1} x_0]$ $\mathscr{D}_j$-fellow travel. Then $[g_1 x_0, q_1]$ is ϵ_j-thick and its length is at least $d(x_0, \psi_1^{-1} x_0) - 2\mathscr{D}_j \geq 2K$.

Let $\eta = [y, y']$ be the subsegment of $[q_1, g_1 x_0] \subseteq [x_0, g_1 x_0]$ such that $d(y, y') = \mathscr{C}_0(\epsilon_j)$ and $d(y', g_1 x_0) = K$. By Theorem 4.7, at least one of $[x_0, w x_0]$ and $[g_1 x_0, w x_0]$ intersects the $\mathscr{D}_0(\epsilon_j)$-neighborhood of η.

Suppose that there exists a point $p \in [g_1 x_0, w x_0]$ that belongs to the $\mathscr{D}_0(\epsilon_j)$-neighborhood of η. We then have

$$
\begin{aligned}
d(g_1 x_0, p) &\leq d(g_1 x_0, \eta) + \operatorname{diam}(\eta) + \mathscr{D}_0(\epsilon_j) \\
&= K + \mathscr{C}_0(\epsilon_j) + \mathscr{D}_0(\epsilon_j) \\
&< 2K \leq d(g_1 x_0, q_2).
\end{aligned}
$$

Hence, $d(g_1 \psi_1^{-1} x_0, g_1 \phi_2 x_0)$ is at most

$$
\begin{aligned}
&\leq d(g_1 \psi_1^{-1} x_0, q_1) + d(q_1, y) + d(y, p) + d(p, q_2) + d(q_2, g_1 \phi_2 x_0) \\
&= d(g_1 \psi_1^{-1} x_0, q_1) + [d(q_1, g_1 x_0) - d(y, g_1 x_0)] + d(y, p) \\
&\quad + [d(g_1 x_0, q_2) - d(g_1 x_0, p)] + d(q_2, g_1 \phi_2 x_0) \\
&\leq d(g_1 \psi_1^{-1} x_0, q_1) + [d(q_1, g_1 \psi_1^{-1} x_0) + d(g_1 \phi_1^{-1} x_0, g_1 x_0)] - d(y, g_1 x_0) + d(y, p) \\
&\quad + [d(g_1 x_0, g_1 \phi_2 x_0) + d(g_1 \phi_2 x_0, q_2)] - d(g_1 x_0, p) + d(q_2, g_1 \phi_2 x_0) \\
&\leq d(g_1 \phi_1^{-1} x_0, g_1 x_0) + d(g_1 x_0, g_1 \phi_2 x_0) - 2d(y, g_1 x_0) \\
&\quad + 2d(y, p) + 2d(q_1, g_1 \psi_1^{-1} x_0) + 2d(q_2, g_1 \phi_2 x_0) \\
&\leq d(g_1 \phi_1^{-1} x_0, g_1 x_0) + d(g_1 x_0, g_1 \phi_2 x_0) - 2[K - \mathscr{D}_0(\epsilon_j) - \mathscr{C}_0(\epsilon_j) - 2\mathscr{D}_j].
\end{aligned}
\tag{4.5}
$$

This implies that $(\phi_2 x_0, \psi_1^{-1} x_0)_{x_0} > C_{Grom}$, a contradiction.

Hence, we instead obtain a point $p \in [x_0, w x_0]$ that belongs to the $\mathscr{D}_0(\epsilon_j)$-neighborhood of η. Then p is within distance $\mathscr{D}_0(\epsilon_j) + \mathscr{C}_0(\epsilon_j) + K$ from $g_1 x_0$. Then Corollary 4.6 tells us that $[x_0, g_1 x_0]$ and $[x_0, p]$ $\mathscr{B}(\epsilon, 2K)$-fellow travel. Since $[g_1 \psi_1^{-1} x_0, g_1 x_0]$ $\mathscr{D}_j$-fellow travel with a terminal subsegment of $[x_0, g_1 x_0]$, we conclude that $[g_1 \psi_1^{-1} x_0, g_1 x_0]$ and a terminal subsegment of $[x_0, p]$ $\mathscr{D}_{j+1}$-fellow travel. This establishes the first item.

If $[x_0, \phi_1 x_0]$ $\mathscr{D}_j$-fellow travel with an initial subsegment of $[x_0, g_1 x_0]$, we also conclude that an initial subsegment of $[x_0, p]$ and $[x_0, \phi_1 x_0]$ $\mathscr{D}_{j+1}$-fellow travel. This establishes the second item. $\qquad\square$

Several corollaries of Lemma 4.18 follow.

Corollary 4.19. *Let $\phi_1, \phi_2 \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ and $v \in G$. Suppose v is $\mathscr{D}_5$-witnessed by ϕ_1 and ϕ_2. Then $\phi_1 = \phi_2$.*

Proof. Suppose not (i.e., $\phi_1 \neq \phi_2$). This implies that the sequence (v^{-1}, v) is $\mathscr{D}_5$-marked with ϕ_1, ϕ_2, which are repulsive. Lemma 4.18 then implies that $[x_0, x_0] = \{x_0\}$ is $\mathscr{D}_6$-witnessed by $[v^{-1}\phi_2^{-1}x_0, v^{-1}x_0]$, which is impossible since $d(x_0, \phi_2 x_0) > 2\mathscr{D}_6$. $\qquad\square$

Corollary 4.20. *Let $1 \leq j \leq 4$, (ϕ_i), (ψ_i) be repulsive sequences and $(g_i)_{i=1}^n$ be a sequence that is $\mathscr{D}_j$-marked with $(\phi_i)_{i=2}^n$ and $(\psi_i)_{i=1}^{n-1}$. Let $w = g_1 \cdots g_n$ and $i \in \{1, \ldots, n-1\}$. Then*

(1) *$[x_0, wx_0]$ is $\mathscr{D}_{j+2}$-witnessed by $[g_1 \cdots g_i x_0, g_1 \cdots g_i \phi_{i+1} x_0]$;*

(2) *$[x_0, wx_0]$ is $\mathscr{D}_{j+2}$-witnessed by $[g_1 \cdots g_i \psi_i^{-1} x_0, g_1 \cdots g_i x_0]$; and*

(3) *$d(x_0, wx_0) \geq d(x_0, g_1 \cdots g_i x_0) + d(x_0, g_{i+1} \cdots g_n x_0) - \mathscr{D}_{j+2}$.*

Proof. Note that $(g_{i+1}, \ldots, g_n)$ is $\mathscr{D}_j$-strongly marked with $(\phi_{i+j})_{j=1}^{n-i}$, $(\psi_{i+j})_{j=1}^{n-i-1}$ and $(g_i^{-1}, \ldots, g_1^{-1})$ is $\mathscr{D}_j$-strongly marked with $(\psi_{i+1-j}^{-1})_{j=1}^{i}$, $(\phi_{i+1-j}^{-1})_{j=1}^{i-1}$. Thus, by Lemma 4.18, $g_{i+1} \cdots g_n$ is $\mathscr{D}_{j+1}$-witnessed by ϕ_{i+1} and $g_i^{-1} \cdots g_1^{-1}$ is $\mathscr{D}_{j+1}$-witnessed by ψ_i^{-1}. In other words, $(g_1 \cdots g_i, g_{i+1} \cdots g_n)$ is $\mathscr{D}_{j+1}$-marked with ϕ_{i+1}, ψ_i. We then apply Lemma 4.18 again to deduce the conclusion. $\qquad\square$

The next corollary plays the role of Lemma 2.4 on the Teichmüller space.

Corollary 4.21. *Let $1 \leq j \leq 4$, $(\phi_i)_{i=2}^n$, $(\psi_i)_{i=1}^{n-1}$ be repulsive sequences, $(g_i)_{i=1}^n$ be a step sequence $\mathscr{D}_j$-marked with (ϕ_i), (ψ_i), and $1 = t_0 < t_1 < \cdots < t_k = n$. Then*

$$\left| d(x_0, g_1 \cdots g_n x_0) - \sum_{i=1}^{k} d(x_0, g_{t_{i-1}+1} \cdots g_{t_i} x_0) \right| \leq (k-1)\mathscr{D}_{j+2}$$

holds.

Proof. Note that a subsequence of a step sequence marked with repulsive sequences is again marked by the corresponding repulsive subsequences. Thus, using induction on k, it suffices to prove the result for $k = 2$. Then Corollary 4.20 applies. $\qquad\square$

4.5. Translation lengths of mapping classes

So far, we observed that the directions witnessed by repulsive sequences of mapping classes are persistent in the final geodesic. Our next aim is to relate these recorded directions with the translation length of w and analyze the effect of pivoting.

Lemma 4.22. *Let $h_1, h_2 \in G$ and $\phi, \psi_1, \psi_2 \in \{\phi_+, \phi_+^{-1}, \phi_-, \phi_-^{-1}\}$ be such that the following hold:*

- *$d(x_0, h_1 x_0) \geq d(x_0, h_2 x_0) + Z$,*
- *h_1^{-1} is $\mathscr{D}_3$-witnessed by ϕ, and*
- *$\psi_1 \neq \psi_2$.*

Then the following hold:

(1) $h_1^{-1} h_2 \psi_i$ *is* $\mathscr{D}_4$-*witnessed by* ϕ *for each* $i = 1, 2$.
(2) *If* $\psi_1^{-1} h_2^{-1} h_1$ *is not* $\mathscr{D}_2$-*witnessed by* ψ_1^{-1}, *then* $\psi_2^{-1} h_2^{-1} h_1$ *is* $\mathscr{D}_2$-*witnessed by* ψ_2^{-1}.
(3) *If* $\psi_2^{-1} h_2^{-1} h_1$ *is not* $\mathscr{D}_2$-*witnessed by* ψ_2^{-1}, *then* $\psi_1^{-1} h_2^{-1} h_1$ *is* $\mathscr{D}_2$-*witnessed by* ψ_1^{-1}.

Proof. Let us first establish (1). The assumption tells us that $[h_1 x_0, x_0]$ contains an initial subsegment $[h_1 x_0, q_1]$ that $\mathscr{D}_3$-fellow travels with $[h_1 x_0, h_1 \phi x_0]$. This implies that $[h_1 x_0, q_1]$ is ϵ_3-thick. Moreover, the length of $[h_1 x_0, q_1]$ is at least $d(x_0, \phi x_0) - \mathscr{D}_3 \geq \mathscr{C}_0(\epsilon_3)$ and at most $d(x_0, \phi x_0) + \mathscr{D}_3$.

Let $\eta = [y, q_1]$ be the subsegment of $[h_1 x_0, q_1] \subseteq [h_1 x_0, x_0]$ such that $d(y, q_1) = \mathscr{C}_0(\epsilon_3)$. By Theorem 4.7, at least one of $[h_1 x_0, h_2 \psi_i x_0]$ and $[h_2 \psi_i x_0, x_0]$ intersects the $\mathscr{D}_0(\epsilon_3)$-neighborhood of η.

If there exists a point $p \in [h_2 \psi_i x_0, x_0]$ that belongs to the $\mathscr{D}_0(\epsilon_3)$-neighborhood of η, we have

$$
\begin{aligned}
d(x_0, h_2 x_0) &\geq d(x_0, h_2 \psi_i x_0) - d(h_2 x_0, h_2 \psi_i x_0) \\
&\geq d(x_0, p) - d(x_0, \psi_i x_0) \\
&\geq d(x_0, \eta) - \mathrm{diam}(\eta) - \mathscr{D}_0(\epsilon_3) - d(x_0, \psi_i x_0) \\
&\geq d(x_0, h_1 x_0) - [d(x_0, \phi x_0) + \mathscr{D}_3] - \mathscr{C}_0(\epsilon_3) - \mathscr{D}_0(\epsilon_3) - d(x_0, \psi_i x_0).
\end{aligned}
$$

This contradicts the assumption that $d(x_0, h_1 x_0) \geq d(x_0, h_2 x_0) + Z$.

Hence, $[h_1 x_0, h_2 \psi_i x_0]$ contains a point p that belongs to the $\mathscr{D}_0(\epsilon_3)$-neighborhood of η. Then

$$
\begin{aligned}
d(p, h_1 \phi x_0) &\leq d(p, \eta) + \mathrm{diam}(\eta) + d(q_1, h_1 \phi x_0) \\
&\leq \mathscr{D}_0(\epsilon_3) + \mathscr{C}_0(\epsilon_3) + \mathscr{D}_3
\end{aligned}
$$

and $[h_1 x_0, p]$ $\mathscr{D}_4$-fellow travels with $[h_1 x_0, h_1 \phi x_0]$.

Let us now establish (2). Let $K = 2\mathscr{C}_0(\epsilon) + 2\mathscr{D}_0(\epsilon) + C_{Grom}$. Recall that $[h_2 x_0, h_2 \psi_1 x_0]$ is an ϵ-thick geodesic whose length is at least $\mathscr{C}_0(\epsilon) + K$.

Let $\eta = [y, y']$ be the subsegment of $[h_2 x_0, h_2 \psi_1 x_0]$ with $d(y, y') = \mathscr{C}_0(\epsilon)$ and $d(h_2 x_0, y) = K$. By Theorem 4.7, either $[h_1 x_0, h_2 \psi_1 x_0]$ or $[h_1 x_0, h_2 x_0]$ intersects the $\mathscr{D}_0(\epsilon)$-neighborhood of η. If there exists a point $p \in [h_1 x_0, h_2 \psi_1 x_0]$ that belongs to $N_{\mathscr{D}_0(\epsilon)}(\eta)$, then $d(p, h_2 x_0) \leq \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) + K$. This implies that $[p, h_2 \psi_1 x_0]$ and $[h_2 x_0, h_2 \psi_1 x_0]$ $\mathscr{D}_2$-fellow travel, which contradicts the assumption. Hence, we instead have a point $p \in [h_1 x_0, h_2 x_0]$ that belongs to $N_{\mathscr{D}_0(\epsilon)}(\eta)$.

We now consider a subsegment $\eta' = [z, z']$ of $[h_2 x_0, h_2 \psi_2 x_0]$ with $d(z, z') = \mathscr{C}_0(\epsilon)$ and $d(h_2 x_0, z) = K$. By Theorem 4.7, either $[h_1 x_0, h_2 \psi_2 x_0]$ or $[h_1 x_0, h_2 x_0]$ intersects the $\mathscr{D}_0(\epsilon)$-neighborhood of η.

Suppose that there exists a point $q \in [h_1 x_0, h_2 x_0]$ that belongs to $N_{\mathscr{D}_0(\epsilon)}(\eta')$. Then from the inequalities

$$
\begin{aligned}
|d(h_2 x_0, p) - d(h_2 x_0, y)| &\leq \mathrm{diam}(\eta) + \mathscr{D}_0(\epsilon), \\
|d(h_2 x_0, q) - d(h_2 x_0, z)| &\leq \mathrm{diam}(\eta') + \mathscr{D}_0(\epsilon)
\end{aligned}
$$

and the fact that h_2x_0, p, q are on the same geodesic, we have

$$d(p,q) = |d(h_2x_0, p) - d(h_2x_0, q)|$$
$$\leq \operatorname{diam}(\eta) + \operatorname{diam}(\eta') + 2\mathscr{D}_0(\epsilon).$$

This in turn implies

$$d(h_2\psi_1x_0, h_2\psi_2x_0) \leq d(h_2\psi_1x_0, y) + d(y,p) + d(p,q) + d(q,z) + d(z, h_2\psi_2x_0)$$
$$\leq [d(h_2x_0, h_2\psi_1x_0) - K] + [\mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon)] + [2\mathscr{C}_0(\epsilon) + 2\mathscr{D}_0(\epsilon)]$$
$$+ [\mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon)] + [d(h_2x_0, h_2\psi_2x_0) - K]$$
$$\leq d(h_2x_0, h_2\psi_1x_0) + d(h_2x_0, h_2\psi_2x_0) - 2C_{Grom}.$$

This contradicts the fact that $(\psi_1x_0, \psi_2x_0)_{x_0} < C_{Grom}$.

Hence, we instead have a point $q \in [h_1x_0, h_2\psi_2x_0]$ that is within distance $\mathscr{D}_0(\epsilon)$ from η. Then we have

$$d(q, h_2x_0) \leq K + \mathscr{C}_0(\epsilon) + \mathscr{D}_0(\epsilon) \leq 3\mathscr{C}_0(\epsilon) + 3\mathscr{D}_0(\epsilon) + C_{Grom}.$$

Corollary 4.6 then tells us that $[q, h_2\psi_2x_0]$ and $[h_2x_0, h_2\psi_2x_0]$ $\mathscr{D}_2$-fellow travel as desired.

(3) is deduced from a similar argument. □

5. Random walks on Teichmüller space

In this section, we adapt the proof of Theorem A to deal with Teichmüller space. Before delving into details, we briefly sketch our plan. We will define persistent joints in sample paths and pivot the path at those joints as in section 3. (See Figure 5.) The basic philosophy of section 3 was using Property 2.3 for hyperbolic spaces. More precisely, we applied Property 2.3 to the Gromov products

$$(x^\sigma_{n\to0}, x^\kappa_{n\to0})_{x_0} >> (x^\kappa_{n\to0}, x^\kappa_n)_{x_0} >> (x^\kappa_n, x^\sigma_n)_{x_0}$$

to conclude that $(x^\sigma_n, x^\sigma_{n\to0})_{x_0}$ is small enough. This led to the lower bound on $\tau(w^\sigma_n)$.

Our aim is to copy this phenomenon to the Teichmüller space X. However, several issues arise due to the fact that X is not Gromov hyperbolic:

- Property 2.3 among the Gromov products may not hold;
- small $(x^\sigma_n, x^\sigma_{n\to0})_{x_0}$ may not lead to large translation length of w^σ_n.

In order to overcome these difficulties, we copy the following property of Gromov hyperbolicity: for a geodesic triangle with vertices x, y and z, the edge $[x,y]$ fellow travels with either $[x,z]$ or $[z,x]$. Thanks to Rafi's theorems (Corollary 4.6 and Theorem 4.7), we can partially guarantee such a phenomenon among certain Teichmüller geodesics witnessed by pseudo-Anosov mapping classes.

We now begin our discussion. Recall that μ is a non-elementary probability measure on the mapping class group G. We have fixed two independent pseudo-Anosov mapping classes ϕ_+, ϕ_- in $\operatorname{supp}\mu^L$ for some $L > 0$. Here, $\phi_\pm$ are associated with constants $\epsilon, \mathscr{M}, C_{Grom}$ and $\mathscr{F}$ that satisfy Lemma 4.8, Lemma 4.11 and Corollary 4.14. We have also defined constants $\mathscr{D}_j, \epsilon_j$ as in Display 4.2 and assumed Inequality 4.3.

Let $a_i, b_i \in \operatorname{supp}\mu$ be the letters for $\phi_\pm$ satisfying $\phi_+ = b_1 \cdots b_L$ and $\phi_- = a_1 \cdots a_L$. As before, we fix the following notations:

$$p_+ = \mu(a_1) \cdots \mu(a_L),$$
$$p_- = \mu(b_1) \cdots \mu(b_L),$$
$$P = \max\left(\frac{p_+}{p_+ + p_-}, \frac{p_-}{p_+ + p_-}\right).$$

This time, we define $\chi_k(\omega)$ as follows. $\chi_k(\omega) = 1$ if

(1)

$$(g_{3(k-1)L+1}, \ldots, g_{3kL}) = \left\{ \begin{array}{c} (b_1, \ldots, b_L, a_1, \ldots, a_L, b_1, \ldots, b_L) \\ \text{or} \\ (b_1, \ldots, b_L, b_1, \ldots, b_L, b_1, \ldots, b_L) \end{array} \right. ,$$

(2) $[x_0, x_{(3k-2)L \to n}]$ is $\mathscr{D}_1$-witnessed by ϕ_+^{-1} for $n \le 3(k-1)L$, and

(3) $[x_0, x_{(3k-1)L \to n}]$ is $\mathscr{D}_1$-witnessed by ϕ_+ for $n \ge 3kL$,

and $\chi_k(\omega) = 0$ otherwise.

We first observe that $\mathbb{E}(\chi_1(\omega)) > 0$. The probability for condition (1) is $(p_+ + p_-)p_-^2 \ne 0$. Given (1) as the prior condition, (2) and (3) become independent events. Moreover, since $\mathscr{F} \le \mathscr{D}_1$, Corollary 4.14 tells us that (2) and (3) hold for nonzero probability. Hence, we conclude that $\eta := \mathbb{P}(\chi_1(\omega) = 1) \ne 0$.

Note that $\chi_k(\omega) = \chi_1(T^{3(k-1)L}\omega)$. We define $W_n = \sum_{k=1}^{n} \chi_k(\omega)$. Then $W_{n+m} = W_n + W_m \circ T^{3Ln}$ holds. Since W_1 is bounded, it has finite first moment. Applying Theorem 3.4, we get almost everywhere convergence of $\frac{1}{n}W_n$ to an a.e. constant variable W_∞. Since $\mathbb{E}(\frac{1}{n}W_n) = \eta > 0$, we have $W_\infty = \eta$ a.e.

We also consider a modified version of W_n as in section 3. Given positive integers $n \le m$, we say that $\mathcal{N}_{m,n} = \{n_1 < \cdots < n_k\} \subseteq 3L\mathbb{Z}$ is an (m,n)-set of pivots for $\vec{w} = (w_1, \cdots, w_m)$ if the following holds:

(1) $\mathcal{N}_{m,n} \subseteq \{1, \ldots, n\}$;

(2) for each i,

$$(g_{n_i - 3L+1}, \ldots, g_{n_i}) = \left\{ \begin{array}{c} (b_1, \ldots, b_L, a_1, \ldots, a_L, b_1, \ldots, b_L) \\ \text{or,} \\ (b_1, \ldots, b_L, b_1, \ldots, b_L, b_1, \ldots, b_L) \end{array} \right. ,$$

(3) for each i, $[x_0, x_{n_i - 2L \to j}]$ is $\mathscr{D}_1$-witnessed by ϕ_+^{-1} for $n_{i-1} - L \le j \le n_i - 3L$, and

(4) for each i, $[x_0, x_{n_i - L \to j}]$ is $\mathscr{D}_1$-witnessed by ϕ_+ for $n_i \le j \le n_{i+1} - 2L$.

For convenience, we set $n_0 = L$ and $n_{k+1} = m + 2L$.

As before, we denote the maximal (m,n)-set of pivots of $\vec{w}$ by $\mathcal{N}(\vec{w}) = \mathcal{N}_{m,n}(\vec{w})$, whose cardinality is denoted by W_n^m. Note that $W_n^n \ge W_{\lfloor n/3L \rfloor}$ always holds.

Now given a finite path $\vec{w} = (w_1, \ldots, w_n)$ with $\mathcal{N}_{n,n}(\vec{w}) = \{n_1 < \ldots < n_k\}$, we define the following:

$$A'_i(\vec{w}) := n_i(\vec{w}) - 3L,$$
$$\alpha'_i(\vec{w}) := n_i(\vec{w}) - 2L,$$
$$\beta'_i(\vec{w}) := n_i(\vec{w}) - L,$$
$$B'_i(\vec{w}) = n_i(\vec{w}).$$

Next, we set

$$(h'_{2i-1}, h'_{2i}) := (w_{\alpha'_i(\vec{w})}, w_{\beta'_i(\vec{w})})$$

for $i = 1, \ldots, k$ and $h'_0 = id$, $h'_{2k+1} = w_n$. We can now discuss an analogue of Lemma 3.5:

Lemma 5.1. *The sequence $(h'^{-1}_{i-1} h'_i)^{2k+1}_{i=1}$ is $\mathcal{D}_1$-marked with repulsive sequences*

$$(\phi_+, \phi'_1, \phi_+, \phi'_2, \ldots, \phi'_k), (\phi'_1, \phi_+, \phi'_2, \phi_+, \ldots, \phi_+),$$

where $\phi'_i := h'^{-1}_{2i-1} h'_{2i}$ is either ϕ_+ or ϕ_-. We also have $h'^{-1}_i h'_j \in \mathcal{C}_{\mathcal{D}_2}(\psi \to \psi')$ for $0 \leq i < j \leq 2k+1$, where

$$\psi = \begin{cases} \phi_+ & i \text{ is odd} \\ \phi'_{i/2} & i \text{ is even} \end{cases}, \quad \psi' = \begin{cases} \phi'_{(i+1)/2} & i \text{ is odd} \\ \phi_+ & i \text{ is even} \end{cases}.$$

Here, we set $\phi'_0 = \phi'_{k+1} = id$ for convenience.

Moreover, we have

$$(h'_i x_0, h'_l x_0)_{h'_j x_0} \leq \mathcal{D}_3, \tag{5.1}$$

$$d(h'_i x_0, h'_l x_0) \geq d(h'_i x_0, h'_j x_0) + \mathcal{D}_3(l - j) \tag{5.2}$$

for each $0 \leq i \leq j \leq l \leq 2k+1$.

Proof. For each $i = 1, \ldots, k$, $h'^{-1}_{2i-1} h'_{2i} = w^{-1}_{\alpha'_i(\vec{w})} w_{\beta'_i(\vec{w})} = \phi'_i$ for some $\phi'_i \in \{\phi_+, \phi_-\}$. Lemma 4.15 then tells us that $h'^{-1}_{2i-1} h'_{2i} \in \mathcal{C}_{\mathcal{D}_1}(\phi'_i \to \phi'_i)$.

Moreover, for each $i = 2, \ldots, k$, Condition (3) and (4) for an (m,n)-set of pivots tell us that

(1) $[x_0, x_{n_i - 2L \to n_{i-1} - L}] = [x_0, x_{\alpha'_i(\vec{w}) \to \beta'_{i-1}(\vec{w})}]$ is $\mathcal{D}_1$-witnessed by ϕ_+^{-1}, and

(2) $[x_0, x_{n_{i-1} - L \to n_i - 2L}] = [x_0, x_{\beta'_{i-1}(\vec{w}) \to \alpha'_i(\vec{w})}]$ is $\mathcal{D}_1$-witnessed by ϕ_+.

This means that $h'^{-1}_{2i-2} h'_{2i-1} = w^{-1}_{\beta'_{i-1}(\vec{w})} w_{\alpha'_i(\vec{w})}$ belongs to $\mathcal{C}_{\mathcal{D}_1}(\phi_+ \to \phi_+)$. Similarly, we observe $h'_1 \in \mathcal{C}_{\mathcal{D}_1}(0 \to \phi_+)$ and $h'_{2k+1} \in \mathcal{C}_{\mathcal{D}_1}(\phi_+ \to 0)$. Since ϕ'_i is either ϕ_+ or ϕ_-, ϕ'_i and ϕ_+^{-1} are distinct for each i. This concludes that $(h'^{-1}_{i-1} h'_i)^{2k+1}_{i=1}$ is $\mathcal{D}_1$-marked with repulsive sequences

$$(\phi_+, \phi'_1, \phi_+, \phi'_2, \ldots, \phi_k), (\phi'_1, \phi_+, \phi'_2, \phi_+, \ldots, \phi_+).$$

Lemma 4.18 then tells us that $h'^{-1}_i h_j \in \mathcal{C}_{\mathcal{D}_2}(\psi \to \psi')$ for each pair (i,j) and the corresponding ψ, ψ'. Moreover, Corollary 4.20 implies Inequality 5.1 for each triple (i,j,l).

Finally, for each $i = 2, \ldots, 2k+1$, $h_{i-1}'^{-1}h_i'$ is $\mathscr{D}_1$-witnessed by either ϕ_+ or ϕ_-. Moreover, $h_1'^{-1}$ is $\mathscr{D}_1$-witnessed by ϕ_+^{-1}. Hence, for $i = 1, \ldots, 2k+1$, we have

$$d(h_i'x_0, h_{i+1}'x_0) \geq d(x_0, \phi_\pm x_0) - 2\mathscr{D}_1 \geq 4\mathscr{D}_3.$$

Combining this with Inequality 5.1 yields Inequality 5.2. $\qquad\square$

We now set

$$M > 1000(1 + L + L/\mathscr{D}_3 + \mathscr{D}_3 + Z),$$
$$D = 10\eta/M, \quad 0.9 \leq Q \leq 1,$$

and

$$F_n := \left\{ \vec{w} = (w_1, \cdots, w_n) : W_n^n \geq \frac{\eta n}{6L} + 1 \right\},$$
$$G_n := \left\{ \vec{w} \in F_n : \tau(w_n) \leq \left(2D - \frac{2\eta}{M} \right) n \right\}.$$

As before, for $\vec{w} \in F_n$ and $Q > 0$, we also define

$$\mathcal{N}_f(\vec{w}; Q) := \left\{ s \in \mathcal{N}(\vec{w}) : d(x_0, x_{s-L}) \leq \frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n \right\},$$
$$\mathcal{N}_b(\vec{w}; Q) = \left\{ s \in \mathcal{N}(\vec{w}) : d(x_n, x_{s-2L}) \leq \frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n \right\},$$
$$\mathcal{N}_0(\vec{w}; Q) := \mathcal{N}(\vec{w}) \setminus (\mathcal{N}_f(\vec{w}; Q) \cup \mathcal{N}_b(\vec{w}; Q)).$$

Lemma 5.2. *For sufficiently large n, if $\vec{w} \in F_n$ and $Q \leq 1$, then*

$$|\mathcal{N}_f(\vec{w}; Q)| \geq \frac{\eta n}{20L} \quad or \quad |\mathcal{N}_b(\vec{w}; Q)| \geq \frac{\eta n}{20L}$$

holds.

Proof. If not, we have $|\mathcal{N}_0(\vec{w}; Q)| \geq \frac{\eta n}{15L} + 1$. As in the proof of Lemma 3.7, we can take t, t' such that $\mathcal{N}_0(\vec{w}; Q) = \{n_t, n_{t+1}, \ldots, n_{t'}\}$. Using Inequality 5.1 and 5.2, we obtain

$$d(x_{\beta_t'}, x_{\alpha_{t'}'}) \geq \mathscr{D}_3(t' - t + 1) \geq \frac{1000L}{M} \cdot \frac{\eta n}{15L} \geq \frac{60\eta n}{M} + 4\mathscr{D}_3$$

for large enough n. Moreover, since $n_t \notin \mathcal{N}_f(\vec{w}; Q)$ and $n_{t'} \notin \mathcal{N}_b(\vec{w}; Q)$, we have

$$d(x_0, x_{\beta_t'}), d(x_{\alpha_{t'}'}, x_n) > \frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n.$$

Combining these inequalities, we obtain

$$d(x_0, x_n) \geq d(x_0, x_{\beta_t'}) + d(x_{\beta_t'}, x_{\alpha_{t'}'}) + d(x_{\alpha_{t'}'}, x_n) - 4\mathscr{D}_3$$
$$\geq 2\left(\frac{1}{2}d(x_0, x_n) - (D + Q\eta/M)n \right) + \frac{60\eta n}{M}$$
$$\geq d(x_0, x_n) + \frac{\eta n}{M},$$

a contradiction. $\qquad\square$

Thus, when $Q \leq 1$,

$$F_{n,f}(Q) := \left\{ \vec{w} \in F_n : |\mathcal{N}_f(\vec{w};Q)| \geq \frac{\eta n}{20L} \right\},$$
$$F_{n,b}(Q) := \left\{ \vec{w} \in F_n : |\mathcal{N}_b(\vec{w};Q)| \geq \frac{\eta n}{20L} \right\}$$

cover entire F_n. We also define

$$G_{n,f}(Q) := F_{n,f}(Q) \cap G_n \quad \text{and} \quad G_{n,b}(Q) := F_{n,b}(Q) \cap G_n.$$

From now on, we will focus on $\vec{w} \in F_{n,f}(Q)$; the argument for $\vec{w} \in F_{n,b}(Q)$ is analogous.

For each $\vec{w} \in F_{n,f}(Q)$, we fix an integer $N = N(\vec{w})$ between $\frac{\eta n}{50M^2 L}$ and $\frac{\eta n}{20M^2 L}$ as before and pick pivot indices $p_1(\vec{w}), \ldots, p_N(\vec{w})$ from $\mathcal{N}_f(\vec{w};Q)$. We also bring the previous notation

$$A_i(\vec{w}) := p_i(\vec{w}) - 3L,$$
$$\alpha_i(\vec{w}) := p_i(\vec{w}) - 2L,$$
$$\beta_i(\vec{w}) := p_i(\vec{w}) - L,$$
$$B_i(\vec{w}) := p_i(\vec{w}).$$

We also let $\beta_0(\vec{w}) = B_0(\vec{w}) := 0$, $\alpha_{N+1}(\vec{w}) = A_{N+1}(\vec{w}) := n$,

$$(h_{2i-1}, h_{2i}) := (w_{\alpha_i(\vec{w})}, w_{\beta_i(\vec{w})})$$

for $i = 1, \ldots, N$ and $h_0 = id$, $h_{2k+1} = w_n$. Since $(\alpha_i)_i$, $(\beta_i)_i$ are subsequences of $(\alpha'_i)_i$ and $(\beta'_i)_i$, respectively, the marking information from Lemma 5.1 and Lemma 4.18 tells us the following.

Lemma 5.3. *The sequence $(h_{i-1}^{-1} h_i)_{i=1}^{2k+1}$ is $\mathscr{D}_2$-marked with repulsive sequences*

$$(\phi_+, \phi_1, \phi_+, \phi_2, \ldots, \phi_k), \ (\phi_1, \phi_+, \phi_2, \phi_+, \ldots, \phi_+),$$

where $\phi_i = h_{2i-1}^{-1} h_{2i}$ is either ϕ_+ or ϕ_-. Moreover, we have

$$(h_i x_0, h_l x_0)_{h_j x_0} \leq \mathscr{D}_3, \tag{5.3}$$

$$d(h_i x_0, h_l x_0) \geq d(h_i, h_j x_0) + \mathscr{D}_3(l - j) \tag{5.4}$$

for each $0 \leq i \leq j \leq l \leq 2k+1$.

For each choice $\sigma \in \{0,1\}^N$, we define the pivoted $\vec{w}^\sigma$ as before: we modify the type of joints that are marked by σ only. Precisely speaking, the step sequences $(g_i)_i$ and $(g_i^\sigma)_i$ for $\vec{w}$ and $\vec{w}^\sigma$ coincide except at $\alpha_j + 1 \leq i \leq \beta_j$ for some j such that $\sigma(j) = 1$. For $\sigma(j) = 1$, we set

$$(g_{\alpha_j+1}^\sigma, \ldots, g_{\beta_j}^\sigma) := \begin{cases} (a_1, \ldots, a_L) & \text{if } (g_{\alpha_j+1}, \ldots, g_{\beta_j}) = (b_1, \ldots, b_L) \\ (b_1, \ldots, b_L) & \text{if } (g_{\alpha_j+1}, \ldots, g_{\beta_j}) = (a_1, \ldots, a_L). \end{cases}$$

Other steps remain unchanged. Now, for

$$(h_{2i-1}^\sigma, h_{2i}^\sigma) := (w_{\alpha_i(\vec{w})}^\sigma, w_{\beta_i(\vec{w})}^\sigma),$$

we have the following observation:

Lemma 5.4. *The sequence $((h_{i-1}^\sigma)^{-1} h_i^\sigma)_{i=1}^{2k+1}$ is $\mathscr{D}_2$-marked with repulsive sequences*

$$(\phi_+, \phi_1^\sigma, \phi_+, \phi_2^\sigma, \ldots, \phi_k), \ (\phi_1^\sigma, \phi_+, \phi_2^\sigma, \phi_+, \ldots, \phi_+),$$

where $\phi_i^\sigma \in \{\phi_+, \phi_-\}$ and $\phi_i^\sigma = \phi_i$ if and only if $\sigma(i) = 0$. Moreover, we have

$$(h_i^\sigma x_0, h_l^\sigma x_0)_{h_j^\sigma x_0} \le \mathscr{D}_3, \tag{5.5}$$

$$d(h_i^\sigma x_0, h_l^\sigma x_0) \ge d(h_i^\sigma, h_j^\sigma x_0) + \mathscr{D}_3(l - j) \tag{5.6}$$

for each $0 \le i \le j \le l \le 2k + 1$.

Proof. Since $g_i = g_i^\sigma$ for all i except for $\alpha_{j+1} \le i \le \beta_j$ where $\sigma(j) = 1$. In particular, we have $w_{\beta'_{i-1}(\vec{w}) \to \alpha'_i(\vec{w})}^\sigma \in \mathcal{C}_{\mathscr{D}_1}(\phi_+ \to \phi_+)$ for various i. Moreover, $w_{\alpha'_i(\vec{w}) \to \beta'_i(\vec{w})}^\sigma = (\phi_i'^\sigma)^{F+1} \in \mathcal{C}_{\mathscr{D}_1}(\phi_i' \to \phi_i')$ for some $\phi_i'^\sigma \in \{\phi_+, \phi_-\}$. Then we have the marking information as in Lemma 5.1 (with constant $\mathscr{D}_1$), and the marking information as in Lemma 5.3 (with constant $\mathscr{D}_2$) by taking subsequences. Finally, $\phi_i'^\sigma \ne \phi_i'$ if and only if p_i' is chosen as a pivot index p_j such that $\sigma(j) = 1$. This leads to the condition when ϕ_i^σ equals ϕ_i. $\square$

We now prove an analogy of Lemma 3.10 for sufficiently large n in a different way, in the case of the Teichmüller space.

Lemma 5.5. *Suppose that $n \ge (Z + 2\mathscr{D}_6)M/\eta$. If $\kappa \ne \sigma \in \{0, 1\}^N$ and $\tau(w_n^\kappa) \le (2D - 2\eta/M)n$, then*

$$\tau(w_n^\sigma) \ge (2D + \eta/M)n.$$

Proof. We first need the result of Claim 3.11. Lemma 5.1 tells us that

$$\left| d(x_0, x_n^\sigma) - \left(\sum_{i=1}^{N+1} d(x_{\beta_{i-1}}^\sigma, x_{\alpha_i}^\sigma) + \sum_{i=1}^{N} d(x_{\alpha_i}^\sigma, x_{\beta_i}^\sigma) \right) \right| \le 2N\mathscr{D}_3,$$

$$\left| d(x_0, x_n^\kappa) - \left(\sum_{i=1}^{N+1} d(x_{\beta_{i-1}}^\kappa, x_{\alpha_i}^\kappa) + \sum_{i=1}^{N} d(x_{\alpha_i}^\kappa, x_{\beta_i}^\kappa) \right) \right| \le 2N\mathscr{D}_3.$$

Recall that

  (1) $d(x_{\beta_{i-1}}^\kappa, x_{\alpha_i}^\kappa) = d(x_{\beta_{i-1}}^\sigma, x_{\alpha_i}^\sigma)$ for all i,
  (2) $d(x_{\alpha_i}^\kappa, x_{\beta_i}^\kappa) = d(x_{\alpha_i}^\sigma, x_{\beta_i}^\sigma)$ for i such that $\sigma(i) = 0$, and
  (3) $|d(x_{\alpha_i}^\kappa, x_{\beta_i}^\kappa) - d(x_{\alpha_i}^\sigma, x_{\beta_i}^\sigma)| \le Z$.

From these, we deduce that

$$|d(x_0, x_n^\sigma) - d(x_0, x_n^\kappa)| \le 4N(\mathscr{D}_3 + Z) \le \frac{\eta n}{5M^2 L}(\mathscr{D}_3 + Z) \le \frac{\eta n}{40M}.$$

For similar reasons, we also get

$$|d(x_n, x_{\alpha_i'}^\kappa) - d(x_n, x_{\alpha_i'}^\sigma)| \le \frac{\eta n}{40M},$$

$$|d(x_0, x_{\beta_i'}^\kappa) - d(x_0, x_{\beta_i'}^\sigma)| \le \frac{\eta n}{40M}$$

for $i = 1, \cdots, N$ by considering partial sums.

Among the indices at which σ and κ differ, let i and j be the first and the last ones, respectively. We then have

$$
\begin{aligned}
d(x_0, x^\kappa_{\alpha(i)}) + d(x_0, x^\kappa_{\beta(j)}) &\leq d(x_0, x_{\alpha(i)}) + d(x_0, x_{\beta(j)}) + \frac{\eta n}{20M} \\
&\leq d(x_0, x_n) - 2(D + Q\eta/M)n + \frac{\eta n}{20M} \\
&\leq d(x_0, x^\kappa_n) - 2(D + 0.9\eta/M)n \\
&\leq d(x_0, x^\kappa_n) - Z
\end{aligned}
\tag{5.7}
$$

for large enough n. We similarly have

$$
2d(x_0, x^\kappa_{\beta(j)}) \leq d(x_0, x^\kappa_n) - 2(D + 0.95\eta/M)n. \tag{5.8}
$$

Now Inequality 5.7 implies

$$
\begin{aligned}
d(x_0, x^\kappa_{n \to \beta(j)}) &\geq d(x_0, x^\kappa_n) - d(x_0, x^\kappa_{\beta(j)}) \\
&\geq d(x_0, x^\kappa_{\alpha(i)}) + Z.
\end{aligned}
$$

Moreover, $w^\kappa_{\beta(j) \to n}$ is $\mathscr{D}_2$-witnessed by ϕ_+. Lemma 4.22 then tells us that $w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \alpha(i)} \phi_\pm$ is $\mathscr{D}_4$-witnessed by ϕ_+. Also note that $w^\kappa_{\beta(j) \to n} = w^\sigma_{\beta(j) \to n}$ and $w^\kappa_{0 \to \alpha(i)} = w^\sigma_{0 \to \alpha(i)}$.

At this moment, if $(\phi^\kappa_i)^{-1} w^\kappa_{\alpha(i) \to 0} w^\kappa_{n \to \beta(j)}$ is $\mathscr{D}_2$-witnessed by $(\phi^\kappa_i)^{-1}$, then

$$
\begin{aligned}
&(w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \beta(i)} \quad, w^\kappa_{\beta(i) \to \beta(j)}, w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \beta(i)} \quad, \omega^\kappa_{\beta(i) \to \beta(j)}, \dots) \\
&= (w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \alpha(i)} \phi^\kappa_i, w^\kappa_{\beta(i) \to \beta(j)}, w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \alpha(i)} \phi^\kappa_i, w^\kappa_{\beta(i) \to \beta(j)}, \dots)
\end{aligned}
$$

is $\mathscr{D}_4$-marked with repulsive sequences

$$
(\phi_+, \phi_+, \dots), (\phi^\kappa_i, \phi_+, \phi^\kappa_i, \phi_+, \dots).
$$

Corollary 4.21 then tells us that $(v^i x_0, v^k x_0)_{v^j x_0} \leq \mathscr{D}_6$ for $v = w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \beta(j)}$ and $i < j < k$. This implies

$$
\begin{aligned}
\tau(w^\kappa_n) &= \tau((w^\kappa_{\beta(j)})^{-1} w^\kappa_n w^\kappa_{\beta(j)}) \\
&= \lim_{m \to \infty} \frac{1}{m} d(x_0, v^m x_0) \geq d(x_0, v x_0) - 2\mathscr{D}_6 \\
&\geq d(x_0, w^\kappa_{\beta(j) \to n} x_0) - d(w^\kappa_{\beta(j) \to n} w^\kappa_{0 \to \beta(j)} x_0, w^\kappa_{\beta(j) \to n} x_0) - 2\mathscr{D}_6 \\
&\geq d(x_0, x^\kappa_n) - 2d(x_0, x^\kappa_{\beta(j)}) - 2\mathscr{D}_6 \\
&\geq 2(D + 0.95\eta/M)n - 2\mathscr{D}_6,
\end{aligned}
$$

which contradicts the fact that $\tau(w^\kappa_n) \leq (2D - 2\eta/M)n$.

Hence, $(\phi^\kappa_i)^{-1} w^\kappa_{\alpha(i) \to 0} w^\kappa_{n \to \beta(j)}$ is not $\mathscr{D}_2$-witnessed by $(\phi^\kappa_i)^{-1}$. Instead, by Lemma 4.22, $(\phi^\sigma_i)^{-1} w^\kappa_{\alpha(i) \to 0} w^\kappa_{n \to \beta(j)}$ is $\mathscr{D}_2$-witnessed by $(\phi^\sigma_i)^{-1}$. Then the above argument tells us that $\tau(w^\sigma_n) \geq (2D + 0.9\eta/M)n$. $\qquad \square$

In particular, for $\vec{w} \in G_{n,f}(Q)$, $\vec{w}^\sigma \notin G_n$ for any nontrivial σ. We now observe an analogue of Lemma 3.15.

Lemma 5.6. *Let $Q = 1$ and n be a sufficiently large integer. Suppose that $\vec{w}, \vec{w}' \in G_{n,f}(Q=1)$ and the number of pivots $N(\vec{w})$, $N(\vec{w}')$ are $\lfloor \frac{\eta n}{40 M^2 L} \rfloor$. Then for $\sigma, \sigma' \in \{0,1\}^N$, $\vec{w}^\sigma = \vec{w}'^{\sigma'}$ if and only if $\vec{w} = \vec{w}'$ and $\sigma = \sigma'$.*

Proof. Let $\vec{v} = \vec{w}^\sigma = \vec{w}'^{\sigma'}$. As before, note that

$$\mathcal{N}_f(\vec{w}; Q=1) \subseteq \mathcal{N}(\vec{w}) \subseteq \mathcal{N}(\vec{v} = \vec{w}^\sigma).$$

Moreover, for each $n_i \in \mathcal{N}_f(\vec{w}; Q=1)$, we have

$$
\begin{aligned}
d(x_0, x_{\beta_i'}^\sigma) &\leq d(x_0, x_{\beta_i'}) + \frac{\eta n}{40 M} \\
&\leq \frac{1}{2} d(x_0, x_n) + \frac{\eta n}{40 M} - \left(D + \frac{\eta}{M} \right) n \\
&\leq \frac{1}{2} d(x_0, x_n^\sigma) + \frac{3\eta n}{80 M} - \left(D + \frac{\eta}{M} \right) n \\
&\leq \frac{1}{2} d(x_0, x_n^\sigma) - \left(D + \frac{0.9\eta}{M} \right) n
\end{aligned}
$$

from Claim 3.11. Thus, $n_i \in \mathcal{N}_f(\vec{v}; Q=0.9)$. It follows that

$$\mathcal{N}_f(\vec{w}; Q=1) \subseteq \mathcal{N}_f(\vec{v}; Q=0.9) \quad \text{and} \quad \vec{v} \in F_{n,f}(Q=0.9).$$

Similarly, $\mathcal{N}_f(\vec{w}'; Q=1) \subseteq \mathcal{N}_f(\vec{v}; Q=0.9)$.

Thus, we are able to pick forward pivots $p_i(\vec{w})$ and $p_i'(\vec{w}')$ of $\vec{w}$ and $\vec{w}'$ altogether for $\vec{v}$. (This will give $N(\vec{v}) \leq \frac{\eta n}{20 M^2 L}$, so we are safe.) Then Lemma 5.5 applied to $\vec{v} \in F_{n,f}(Q = 0.9)$ yields a contradiction with $\vec{w}, \vec{w}' \in G_n$ unless $\sigma = \sigma'$. □

We are now ready to prove the second main theorem, Theorem B:

Theorem B. *Let S be a closed orientable surface of genus at least 2. Let $G = \mathrm{Mod}(S)$ be its mapping class group, $X = \mathcal{T}(S)$ be its Teichmüller space equipped with the Teichmüller metric, and μ be a non-elementary probability measure on $\mathrm{Mod}(S)$. Then $\mathbb{P}$-a.e. sample path shows at least linear growth of translation lengths. More precisely, there exists a constant $\mathcal{L} > 0$ such that for $\mathbb{P}$-a.e. (ω_n), $\tau(\omega_n) \geq \mathcal{L}n$ for sufficiently large n.*

Moreover, if μ further has finite first moment with respect to the Teichmüller metric, then for $\mathbb{P}$-a.e. (ω_n), we have

$$\lim_{n \to \infty} \frac{\tau(\omega_n)}{n} = \lambda,$$

where $\lambda > 0$ is the drift.

Proof. Lemma 5.6 tells us that $\mathbb{P}((\omega_i)_{i=1}^n \in G_{n,f}(Q=1)) \leq P^n$ for sufficiently large n. Similarly, we obtain $\mathbb{P}((\omega_i)_{i=1}^n \in G_{n,b}(Q=1)) \leq P^n$. Since $G_{n,f}(Q=1) \cup G_{n,b}(Q=1) = G_n$, we deduce that $\mathbb{P}((\omega_i)_{i=1}^n \in G_n) \leq 2P^n$ for sufficiently large n. Thus, by the Borel-Cantelli lemma, almost every path avoids G_n eventually. If ω avoids G_n eventually but $\tau(\omega_n) \leq (2D - 2\eta/M)n$ infinitely often, then ω_n avoids F_n infinitely often. However, since $W_n^n \geq W_{\lfloor n/3L \rfloor}$, the subadditive ergodic theorem implies that such path constitutes a measure zero set.

We now investigate the second assertion of Theorem B. When μ has finite first moment, the drift λ of the random walk ω is finite and strictly positive. Here, the strict positivity of λ follows from the nontriviality of the Poisson boundary of non-elementary random walks on Teichmüller space (cf. [KM96, Theorem 2.3.2]). As in the proof of the second assertion of Theorem A, we define the following set for $0 < \epsilon < 1$ and $D = 0.5(1 - \epsilon/2)\lambda$:

$$F_n := \left\{ \vec{w} : \begin{matrix} \#\mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}(\vec{w}) \geq \frac{\eta \epsilon n}{24L}, \, d(x_0, w_n x_0) \geq \left(1 - \frac{\epsilon}{1000}\right)\lambda n, \\ d(x_0, w_{\lfloor \epsilon n/7 \rfloor} x_0) \leq \frac{\epsilon \lambda n}{6} \end{matrix} \right\}$$

$$G_n := \left\{ \vec{w} \in F_n : \tau(w_n) \leq \left(2D - \frac{2\eta}{M}\right)n \right\}.$$

By the subadditive ergodic theorem, a.e. sample path ω belongs to F_n for all sufficiently large n. We then require $M > \frac{2000\eta}{\lambda \epsilon} + \frac{2000\eta}{\epsilon} + \frac{2000}{\epsilon} + 1$ and $0.9 \leq Q \leq 1$. Given $\vec{w} \in F_n$, we observe as before that

$$d(x_0, x_{n_i - L}) \leq d(x_0, w_{\lfloor \epsilon n/7 \rfloor} x_0) \leq \frac{1}{2} d(x_0, w_n x_0) - \left(D + \frac{\eta}{M}\right)n$$

for each $n_i \in \mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}(\vec{w})$. This implies that $\mathcal{N}_{\lfloor \epsilon n/7 \rfloor, n}(\vec{w}) \subseteq \mathcal{N}_f(\vec{w})$ and $\vec{w} \in F_{n,f}$. In other words, $F_{n,f}$ already covers F_n, which is stronger than what we hope in Lemma 5.2.

Given this, Lemma 5.3, Lemma 5.5 and Lemma 5.6 still hold true. Using them, we deduce that almost every path ω does not fall into $G_n(Q = 1)$ infinitely often. Note also that

$$2D - \frac{2\eta}{M} \geq (1 - \epsilon)\lambda.$$

Together with the observation that a.e. path does not avoid F_n infinitely often, we deduce that $\liminf \frac{1}{n} \tau(\omega_n) \geq (1 - \epsilon)\lambda$ almost surely, as desired. $\qquad \square$

Competing Interests. The authors have no competing interest to declare.

References

[AS21]   R. Aoun and C. Sert, Law of large numbers for the spectral radius of random matrix products, *Amer. J. Math.* **143**(3) (2021), 995–1010.

[BCK22] H. Baik, I. Choi and D. M. Kim, Topological entropy of pseudo-Anosov maps from a typical Thurston's construction, *Int. Math. Res. Not. IMRN* **24** (2022), 19862–19904, 2022.

[BF14]   M. Bestvina and M. Feighn, Hyperbolicity of the complex of free factors, *Adv. Math.* **256** (2014), 104–155.

[BH99]   M. R. Bridson and A. Haefliger, *Metric Spaces of Non-Positive Curvature* (Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]) vol. 319 (Springer-Verlag, Berlin, 1999).

[BMSS22] A. BOUNLANGER, P. MATHIEU, Ç. SERT AND A. SISTO, Large deviations for random walks on hyperbolic spaces, *Ann. Sci. Éc. Norm. Supér. (4)* (2022).

[BQ16] Y. BENOIST AND J.-F. QUINT, Central limit theorem on hyperbolic groups, *Izv. Ross. Akad. Nauk Ser. Mat.* **80**(1) (2016), 5–26.

[CR07] Y.-E. CHOI AND K. RAFI, Comparison between Teichmüller and Lipschitz metrics, *J. Lond. Math. Soc. (2)* **76**(3) (2007), 739–756.

[DH18] F. DAHMANI AND C. HORBEZ, Spectral theorems for random walks on mapping class groups and Out(F_N), *Int. Math. Res. Not. IMRN* **9** (2018), 2693–2744.

[DT17] S. DOWDALL AND S. J. TAYLOR, The co-surface graph and the geometry of hyperbolic free group extensions, *J. Topol.* **10**(2) (2017), 447–482.

[Duc05] M. DUCHIN, Thin triangles and a multiplicative ergodic theorem for Teichmüller geometry, Preprint, 2005, arXiv:math/0508046.

[EST20] V. ERLANDSSON, J. SOUTO AND J. TAO, Genericity of pseudo-Anosov mapping classes, when seen as mapping classes, *L'Enseingement Mathématique* **66**(3/4) (2020), 419–439.

[GM12] F. GAUTERO AND F. MATHÉUS, Poisson boundary of groups acting on $\mathbb{R}$-trees, *Israel J. Math.* **191**(2) (2012), 585–646.

[Gou22] S. GOUËZEL, Exponential bounds for random walks on hyperbolic spaces without moment conditions, *Tunis. J. Math.* **4**(4) (2022), 635–671.

[Gro87] M. GROMOV, Hyperbolic groups, in *Essays in Group Theory* (Math. Sci. Res. Inst. Publ.) vol. **8** (Springer, New York, 1987), 75–263.

[Ham06] U. HAMENSTÄDT, Train tracks and the Gromov boundary of the complex of curves, in *Spaces of Kleinian Groups* (London Math. Soc. Lecture Note Ser.) vol. 329 (Cambridge Univ. Press, Cambridge, 2006), 187–207.

[Har81] W. J. HARVEY, Boundary structure of the modular group, in *Riemann Surfaces and Related Topics: Proceedings of the 1978 Stony Brook Conference* (State Univ. New York, Stony Brook, N.Y., 1978) (Ann. of Math. Stud.) vol. 97 (Princeton Univ. Press, Princeton, N.J., 1981), 245–251.

[HMM18] P. HAÏSSINSKY, P. MATHIEU AND S. MÜLLER, Renewal theory for random walks on surface groups, *Ergodic Theory Dynam. Systems* **38**(1) (2018), 155–179.

[Hub06] J. H. HUBBARD, *Teichmüller Theory and Applications to Geometry, Topology, and Dynamics* vol. 1 (Matrix Editions, Ithaca, NY, 2006). With contributions by Adrien Douady, William Dunbar, Roland Roeder, Sylvain Bonnot, David Brown, Allen Hatcher, Chris Hruska and Sudeb Mitra. With forewords by William Thurston and Clifford Earle.

[IT92] Y. IMAYOSHI AND M. TANIGUCHI, *An Introduction to Teichmüller Spaces* (Springer-Verlag, Tokyo, 1992). Translated and revised from the Japanese by the authors.

[Iva02] N. V. IVANOV. A short proof of non-Gromov hyperbolicity of Teichmüller spaces, *Ann. Acad. Sci. Fenn. Math.* **27**(1) (2002), 3–5.

[Kai00] V. A. KAIMANOVICH, The Poisson formula for groups with hyperbolic properties, *Ann. of Math. (2)* **152**(3) (2000), 659–692.

[Ker80] S. P. KERCKHOFF, The asymptotic geometry of Teichmüller space, *Topology* **19**(1) (1980), 23–41.

[Kla22] E. KLARREICH, The boundary at infinity of the curve complex and the relative Teichmüller space, *Groups Geom. Dyn.* **16**(2) (2022), 705–723.

[KM96] V. A. KAIMANOVICH AND H. MASUR, The Poisson boundary of the mapping class group, *Invent. Math.* **125**(2) (1996), 221–264.

[Led01] F. LEDRAPPIER, Some asymptotic properties of random walks on free groups, in *Topics in Probability and Lie Groups: Boundary Theory* (CRM Proc. Lecture Notes) vol. **28**. (Amer. Math. Soc., Providence, RI, 2001), 117–152.

[Len08]   A. LENZHEN, Teichmüller geodesics that do not have a limit in PMF, *Geom. Topol.* **12**(1) (2008), 177–197.

[Mah11]   J. MAHER, Random walks on the mapping class group, *Duke Math. J.* **156**(3) (2011), 429–468.

[Mah12]   J. MAHER, Exponential decay in the mapping class group, *J. Lond. Math. Soc. (2)* **86**(2) (2012), 366–386.

[Man14]   B. MANN, Some hyperbolic Out(FN)-graphs and nonunique ergodicity of very small FN-trees, Ph.D. Thesis, 2014, The University of Utah (ProQuest LLC, Ann Arbor, MI).

[Mas75]   H. MASUR, On a class of geodesics in Teichmüller space, *Ann. of Math. (2)* **102**(2) (1975), 205–221.

[MM99]   H. A. MASUR AND Y. N. MINSKY, Geometry of the complex of curves. I. Hyperbolicity, *Invent. Math.* **138**(1) (1999), 103–149.

[MM00]   H. A. MASUR AND Y. N. MINSKY, Geometry of the complex of curves. II. Hierarchical structure, *Geom. Funct. Anal.* **10**(4) (2000), 902–974.

[MP89]   J. MCCARTHY AND A. PAPADOPOULOS, Dynamics on Thurston's sphere of projective measured foliations, *Comment. Math. Helv.* **64**(1) (1989), 133–166.

[MS20]   P. MATHIEU AND A. SISTO, Deviation inequalities for random walks, *Duke Math. J.* **169**(5) (2020), 961–1036.

[MT18]   J. MAHER AND G. TIOZZO, Random walks on weakly hyperbolic groups, *J. Reine Angew. Math.* **742** (2018), 87–239.

[MW95]   H. A. MASUR AND M. WOLF, Teichmüller space is not Gromov hyperbolic, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **20**(2) (1995), 259–267.

[Pap07]   A. PAPADOPOULOS (ed.), *Handbook of Teichmüller Theory. Vol. I* (IRMA Lectures in Mathematics and Theoretical Physics) vol. 11 (European Mathematical Society (EMS), Zürich, 2007).

[Raf14]   K. RAFI, Hyperbolicity in Teichmüller space, *Geom. Topol.* **18**(5) (2014), 3025–3053.

[Thu88]   W. P. THURSTON, On the geometry and dynamics of diffeomorphisms of surfaces, *Bull. Amer. Math. Soc. (N.S.)* **19**(2) (1988), 417–431.

[Woe00]   W. WOESS, *Random Walks on Infinite Graphs and Groups* (Cambridge Tracts in Mathematics) vol. **138** (Cambridge University Press, Cambridge, 2000).