WORD LENGTH STATISTICS FOR TEICHMÜLLER GEODESICS AND SINGULARITY OF HARMONIC MEASURE

VAIBHAV GADRE, JOSEPH MAHER, AND GIULIO TIOZZO

Abstract. Given a measure on the Thurston boundary of Teichmüller space, one can pick a geodesic ray joining some basepoint to a randomly chosen point on the boundary. Different choices of measures may yield typical geodesics with different geometric properties. In particular, we consider two families of measures: the ones which belong to the Lebesgue or visual measure class, and harmonic measures for random walks on the mapping class group generated by a distribution with finite first moment in the word metric.

We consider the word length of approximating mapping class group elements along a geodesic ray, and prove that this quantity grows superlinearly in time along almost all geodesics with respect to Lebesgue measure, while along almost all geodesics with respect to harmonic measure the growth is linear. As a corollary, the harmonic and Lebesgue measures are mutually singular. We also prove a similar result for the ratio between the word metric and the relative metric (i.e. the induced metric on the curve complex).

1. Introduction

Let $G = Mod(S)$ be the mapping class group of an orientable surface S of finite type, which acts on the Teichmüller space $\mathcal{T}(S)$ of marked hyperbolic metrics on S. Following Thurston, a boundary of Teichmüller space is given by the space $\mathcal{PMF}$ of projective measured foliations, which carries several measures:

- on the one hand, there is a natural *Lebesgue measure class* Leb on $\mathcal{PMF}$ given by pulling back Lebesgue measure from the charts defined using train track coordinates;
- on the other hand, Kaimanovich and Masur [18] showed that if μ is a probability distribution on G, whose support generates a non-elementary subgroup, then the image of a random walk on G under the orbit map $g \mapsto gX_0$ converges to a point in $\mathcal{PMF}$ almost surely. We let ν be the corresponding *hitting measure* (also known as *harmonic measure*).

In this paper, we analyze geometric properties of typical geodesics with respect to these measures. To define what we mean by typical geodesics, let us fix a basepoint $X_0 \in \mathcal{T}(S)$. Then there is a map from $\mathcal{PMF}$ to the set of Teichmüller geodesic rays based at X_0. In fact, we can associate to each measured foliation F the unique unit area quadratic differential at X_0 which has vertical foliation proportional to F, and this quadratic differential determines a geodesic ray based at X_0. Thus, we can think of the above measures as measures on the set of geodesic rays from X_0, and we can talk about the behavior of geodesics which are typical with respect to either measure.

1.1. **The word length ratio.**

Let us denote as $\|g\|_G$ the *word length* of a group element g with respect to some fixed generating set. As different choices of generators lead to quasi-isometric metrics, our results will be independent of the particular choice.

Let γ be a Teichmüller geodesic ray based at X_0. For each time t, we denote as γ_t the point at distance t from the basepoint along γ, and we let g_t be a group element such that g_tX_0 is a closest element of the G-orbit of X_0 to γ_t. This is illustrated schematically in Figure 1 below. We then define the *word length ratio* as the quantity

$$\rho(\gamma) := \lim_{\substack{t\to\infty \\ \gamma_t \notin \mathcal{T}_\epsilon}} \frac{\|g_t\|_G}{t}$$

(for technical reasons, we restrict to taking limits over points γ_t which do not lie in the ϵ-thin part $\mathcal{T}_\epsilon$; however, since with respect to both measures generic geodesics are recurrent to the thick part, the limit makes sense almost surely). Our first result establishes that the word metric grows superlinearly along geodesics which are typical with respect to Lebesgue measure:

Theorem 1.1. *Let X_0 be a point in Teichmüller space. Then for Lebesgue-almost every geodesic ray γ based at X_0, the word length ratio is infinite:*

$$\rho(\gamma) = \infty.$$

We now state a corresponding result for random walks. Recall that a measure μ on the mapping class group G has *finite first moment* in the word metric if $\int_G \|g\|_G \, d\mu(g) < \infty$. Moreover, we say that μ is *non-elementary* if the support of μ generates a non-elementary subgroup of G as a semigroup. The word length ratio is almost surely finite along typical geodesics with respect to harmonic measure:

Theorem 1.2. *Let μ be a non-elementary probability measure on $Mod(S)$ with finite first moment in the word metric, let ν be the harmonic measure determined by the corresponding random walk, and let $X_0 \in \mathcal{T}(S)$ a basepoint. Then there is a constant $c > 0$ such that*

$$\rho(\gamma) = c$$

for ν-almost every geodesic ray γ based at X_0.

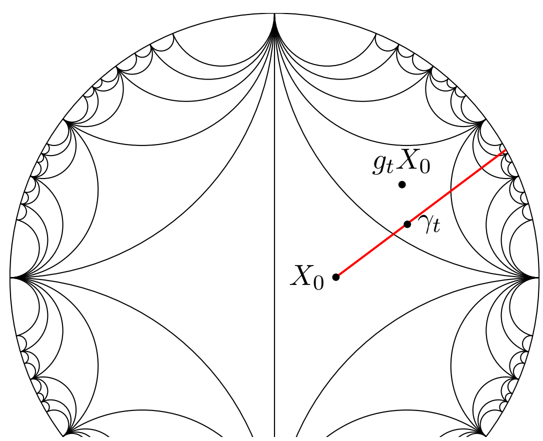

FIGURE 1. Our definition of g_t. The point γ_t lies on the geodesic γ, and $g_t X_0$ is a closest orbit point to γ_t. In the case of genus 1, this corresponds to taking g_t to be the element in the orbit of X_0 which lies in the same tile as γ_t of the Farey tessellation.

1.2. **The relative metric.** A way to interpret this result is in term of different metrics on the mapping class group. If G acts isometrically on a metric space (X, d), one can pick a basepoint $x_0 \in X$ and consider the distance on G given by $d(g, h) := d(gx_0, hx_0)$. In particular, the mapping class group acts on the following three different metric spaces, and this can be used to define three metrics on G:

(1) the *word metric* $\|\cdot\|_G$ (or d_G) previously mentioned arises from the action of G on its Cayley graph;
(2) the mapping class group acts on the Teichmüller space equipped with the *Teichmüller metric*: we will denote this metric as $\|\cdot\|_{\mathcal{T}}$ or $d_{\mathcal{T}}$;
(3) the mapping class group acts on the curve complex $\mathcal{C}(S)$: the resulting metric is called the *relative metric* $\|\cdot\|_{rel}$ (or d_{rel}). The curve complex is a locally infinite, hyperbolic, simplicial complex. The relative metric on G can also be seen as the word metric with respect to an infinite generating set, constructed by adding to a finite generating set the stabilizers of simple closed curves α_i, where the α_i are a set of representatives for orbits of simple closed curves under G [27].

It is well known that $\|\cdot\|_{rel} \lesssim \|\cdot\|_{\mathcal{T}} \lesssim \|\cdot\|_G$, and the three metrics are not quasi-isometric to each other. Moreover, except in some cases of low complexity, neither the word metric nor the Teichmüller metric are hyperbolic in the sense of Gromov; the curve complex, on the other hand, is known to be hyperbolic [27], but this comes at the price of not being locally compact.

However, if one restricts to geodesics that lie completely in some thick part, then the three metrics are indeed quasi-isometric. Thus, one can interpret the results of Theorem 1.1 and Theorem 1.2 by saying that typical geodesics with respect to Lebesgue measure have larger excursions in the thin part than typical geodesics with respect to harmonic measure.

We can also modify the word length ratio and define the *relative word length ratio* $\rho_{rel}(\gamma)$ as

$$\rho_{rel}(\gamma) := \lim_{\substack{t\to\infty \\ \gamma_t \notin \mathcal{T}_\epsilon}} \frac{\|g_t\|_G}{\|g_t\|_{rel}}.$$

For the relative word length ratio we have a similar result.

Theorem 1.3. *Let X_0 be a point in Teichmüller space. Then for Lebesgue-almost every geodesic ray γ based at X_0, we have*

$$\rho_{rel}(\gamma) = \infty.$$

Moreover, let μ be a non-elementary probability measure on $Mod(S)$ with finite first moment in the word metric, and let ν be the harmonic measure determined by the corresponding random walk. Then there exists a constant $c > 0$ such that

$$\rho_{rel}(\gamma) = c$$

for ν-almost every geodesic ray γ based at X_0.

Let us remark that in the special case of $G = SL_2(\mathbb{Z})$ (mapping class group of the torus), Theorem 1.3 is equivalent to the following classical statement in terms of continued fractions. For each $r \in \mathbb{R}$, let us denote as $a_n(r)$ the n^{th} coefficient in the continued fraction expansion of r. Then for Lebesgue-almost every r, we have

$$\lim_{n\to\infty} \frac{a_1(r) + \cdots + a_n(r)}{n} = +\infty,$$

while for almost every r with respect to harmonic measure, we have

$$\lim_{n\to\infty} \frac{a_1(r) + \cdots + a_n(r)}{n} = c < \infty.$$

Note that the statement of Theorem 1.3 for the Lebesgue measure follows immediately from Theorem 1.1, as the Teichmüller metric is a coarse upper bound for the relative metric. For the same reason, the statement of Theorem 1.3 implies that for the harmonic measure the word length ratio $\rho(\gamma)$ is almost surely bounded above and below between two positive constants: this is almost the statement of Theorem 1.2, except for the existence of the limit.

1.3. **Singularity of harmonic measure.** The theorem has the following corollary for the harmonic measure:

Theorem 1.4. *Let μ be a measure on the mapping class group with finite first moment in the word metric, and such that the semigroup generated by its support is a non-elementary subgroup of $Mod(S)$. Then the corresponding harmonic measure ν on $\mathcal{PMF}$ is singular with respect to Lebesgue measure.*

The singularity of harmonic measure for random walks on the mapping class group was conjectured by Kaimanovich and Masur [18]. The origin of this conjecture lies in the following analogy with Riemannian manifolds. If $\widetilde{M}$ is the universal cover of a compact surface M of negative curvature, the harmonic measure on the ideal boundary of $\widetilde{M}$ given by the Brownian motion is singular with respect to the visual measure unless the surface has constant curvature (see Katok [16] and Ledrappier [20]). In this light, the singularity of Theorem 1.4 is to be expected as Teichmüller space is inhomogeneous (e.g., its isometry group is discrete). Note that the Teichmüller metric is not Riemannian, so it is not clear how to define a Brownian motion, but one can use the random walk as a discrete analog.

Note that the finite first moment assumption is essential; indeed, it is conjectured that there exists a measure μ on $Mod(S)$ such that the hitting measure of the corresponding random walk is absolutely continuous on $\mathcal{PMF}$: as a consequence of our result, such a measure must have infinite first moment.

In [10], Gadre proved singularity of the harmonic measure for random walks on $Mod(S)$ generated by measures with finite support, while here we consider arbitrary measures of finite first moment. His proof uses train track splittings on $\mathcal{PMF}$, and relies on the exponential decay of measures of shadow sets, which are not known for measures of finite first moment.

In this paper, we get the Lebesgue measure statistics by using the ergodicity of the Teichmüller geodesic flow, combined with estimates on the volume of the thin part of the space of quadratic differentials. In particular, we define the following function $L : \mathcal{QM} \to \mathbb{R}$ on the moduli space of quadratic differentials:

$$L(q) := \sum_{\alpha \in C_q(\delta,\epsilon)} \frac{1}{\ell_q^2(\alpha)},$$

where $\ell_q(\alpha)$ is the length of the curve α in the flat metric q, and $C_q(\delta, \epsilon)$ the set of cylinders which have core length $\leq \sqrt{\epsilon}$ and area $\geq \delta$. We then prove, using results of Eskin-Masur [7], that the ergodic average of L is infinite along orbits of the Teichmüller flow; Theorem 1.1 follows since the time integral of L is a lower bound for the word metric $\| \cdot \|_G$.

The statistics for harmonic measure follows from linear progress in the relative metric, combined with sublinear tracking between geodesics and sample paths. In particular, in [32] it is proven that random walks on the mapping class group track Teichmüller geodesics sublinearly (in the Teichmüller metric). In order to transfer the information about the ratio between the word and the relative metric along sample paths, we will prove a tracking result in the word metric.

1.4. **Background and remarks.** For random walks on general groups, the question of singularity of harmonic measure has a long history (see also the introduction of Kaimanovich and Le Prince [17]). In the context of lattices in Lie groups, Furstenberg [8, 9] first constructed random walks on discrete groups whose hitting measure is absolutely continuous on the boundary. For non-uniform lattices of rank 1, these random walks have finite first moment in the Riemannian metric on the Lie group, but do not have finite first moment in the word metric on the discrete subgroup.

For non-uniform lattices Γ in $SL(2, \mathbb{R})$, Guivarc'h and Le Jan [13, 14] proved the singularity of harmonic measures by studying the asymptotic winding around the cusp of the geodesic flow on $\Gamma \backslash \mathbb{H}^2$. Other approaches are given by Deroin, Kleptsyn and Navas [4] and by Blachère, Haïssinsky and Mathieu [2]. Our approach via the word length ratio can be also applied to non-uniform lattices in $SL(2,\mathbb{R})$ [12]. On the other hand, for finitely supported measures on uniform lattices in $SL(2,\mathbb{R})$, harmonic measure is expected to be singular; however, the question appears to be still open.

Several authors have considered cusp excursions of Lebesgue-typical geodesics; in particular, Sullivan [31] showed that on a non-compact hyperbolic manifold a generic geodesic ray ventures into the cusps infinitely often with maximum depth in the cusps of about $\log t$, where t is the time along the geodesic ray. The same approach has been then adapted to the Teichmüller geodesic flow by Masur [26].

Our method uses essentially only the geometry of the cusp, so it is natural to expect it to apply to other group actions for which the orbit space is a non-compact manifold of finite volume and the geodesic flow is ergodic, e.g. for fundamental groups of higher-dimensional hyperbolic manifolds with cusps.

1.5. **Outline of the paper.** In Section 2 we present background material on Teichmüller theory; in particular, we review the curve complex and marking complex, define the concept of excursion and use results of Rafi in order to prove the coarse monotonicity in the word metric of the approximating group elements along Teichmüller geodesics. In Section 3 we prove the asymptotic result for the Lebesgue measure, i.e. Theorem 1.1. This is done by considering the ergodic average with respect to the Teichmüller flow of an appropriate function defined on the moduli space of quadratic differentials (Theorem 3.3) and then relate the average to the growth rate of the word metric along typical geodesics. In Section 4 we prove Theorem 1.2, namely the asymptotics for harmonic measure.

1.6. **Notation.** We shall find it convenient to occasionally use *big O* notation. We say that $f(x) = O(g(x))$ if there are constants A and B such that $|f(x)| \leqslant A\,|g(x)|$ for all $x \geqslant B$. In particular, $f(x) = O(1)$ means that the function $f(x)$ is bounded. We will also write $f(x) \lesssim g(x)$ to mean that the inequality holds up to additive and multiplicative constants, i.e. there are constants K and c such that

$$f(x) \leqslant K g(x) + c,$$

and similarly $f(x) \asymp g(x)$ will mean that there exist constants K, c such that

$$\frac{1}{K} g(x) - c \leqslant f(x) \leqslant K g(x) + c.$$

2. Preliminaries from Teichmüller theory

In Sections 2.1–2.3 we review some background material on quadratic differentials, subsurface projections and short markings. In Sections 2.4 and 2.5 we review in detail some results of Rafi [29] which relate subsurface projection distance first to the twist parameter along a Teichmüller geodesic, and then to the excursion distance along the geodesic. In Section 2.6, we use results of Rafi [30] to show that word length grows coarsely monotonically along Teichmüller geodesics, and finally in Section 2.7, we show that a similar result holds for the nearest lattice points to the geodesic, if they lie in the thick part of Teichmüller space.

2.1. **Quadratic differentials and Teichmüller discs.** Let S be a hyperbolic surface of finite type, i.e. a surface of finite area which may have boundary components or punctures. We say such a surface S is *sporadic* if it is a sphere with at most four punctures or boundary components, or a torus with at most one puncture or boundary component. We shall primarily be interested in non-sporadic surfaces, as in the sporadic cases the Teichmüller spaces are either trivial, or isometric to $\mathbb{H}^2$, and covered by the case of $SL(2,\mathbb{Z})$ (see [12]).

Let S be a non-sporadic surface with no boundary components, but which may have punctures. We will write $\mathcal{T}(S)$ for the Teichmüller space of a surface S, or just $\mathcal{T}$ if we do not need to explicitly refer to the surface. We shall consider $\mathcal{T}$ together with the Teichmüller metric

$$d_{\mathcal{T}}(x, y) = \tfrac{1}{2} \inf_{f} \log K(f),$$

where the infimum is taken over all quasiconformal maps $f\colon x \to y$, and $K(f)$ is the quasiconformal constant for the map f. The mapping class group $G = \mathrm{Mod}(S)$ of the surface acts by isometries on $\mathcal{T}$, and we shall write $\mathcal{T}_\epsilon$ for the *thin part* of Teichmüller space, i.e. all surfaces which contain a curve of hyperbolic length at most ϵ. We shall write $\mathcal{M}$ for the quotient $G\backslash\mathcal{T}$, which is known as moduli space. The thin part of Teichmüller space is mapping class group invariant, and we shall write $\mathcal{M}_\epsilon$ for the subset of moduli space given by $G\backslash\mathcal{T}_\epsilon$.

Let $\mathcal{Q}$ be the space of unit area quadratic differentials, which may be identified with the unit cotangent bundle to Teichmüller space [15]. We shall write π for the projection $\pi : \mathcal{Q} \to \mathcal{T}$ which sends a quadratic differential to its underlying Riemann surface, and we shall write μ_{hol} for the Masur-Veech measure, also known as the holonomy measure, as it may be defined in terms of holonomy coordinates. The measure μ_{hol} is mapping class group invariant, and so gives a measure on the moduli space of unit area quadratic differentials $\mathcal{MQ} = G\backslash\mathcal{Q}$, which has finite volume [25, 33].

A quadratic differential q determines a flat structure on the surface, which may be thought of as a union of polygons glued together along parallel sides, where the vertices of the polygons may correspond to points of cone angle $n\pi$, for $n \geqslant 1$. If $n \geqslant 2$, then the vertex corresponds to a zero of order $n - 2$ for the quadratic differential q, and for $n = 1$ the vertices correspond to cone points of angle π which are simple poles for the quadratic differential, and correspond to the punctures of the surface. There is an affine action of $SL(2, \mathbb{R})$ on the flat surface, which gives rise to a new quadratic differential. The orbits of quadratic differentials under the action of $SL(2, \mathbb{R})$ give a foliation of $\mathcal{Q}$ by copies of $SL(2, \mathbb{R})$, and we shall write $\widetilde{D}_q$ for the orbit of the quadratic differential q. We shall write D_q for the image of $\widetilde{D}_q$ in $\mathcal{T}$, and this is called a Teichmüller disc, which is geodesically embedded in $\mathcal{T}$. With the metric induced from the Teichmüller metric, D_q is isometric to the hyperbolic plane of constant curvature -4, and it will be convenient for us to use coordinates coming from the disc model of hyperbolic plane, with the initial quadratic differential q corresponding to the origin.

The group of rotations of $\mathbb{R}^2$ acts on flat surfaces, and hence on $\mathcal{Q}$. In terms of quadratic differentials, rotation by angle θ in $\mathbb{R}^2$ sends $q \mapsto e^{-2i\theta}q$, and this action is trivial on Teichmüller space $\mathcal{T}$. It follows from the definition that holonomy measure is invariant under rotation, i.e. $\mu_{\mathrm{hol}}(U) = \mu_{\mathrm{hol}}(e^{i\theta}U)$, for all θ, for any subset $U \subset \mathcal{Q}$. In particular, this means that if we consider the conditional measure from μ_{hol} on the image of a point $q \in \mathcal{Q}$ under rotation, i.e. $\{e^{i\theta}q : \theta \in [0, 2\pi]\}$, then this is precisely the invariant Haar or Lebesgue measure on the circle.

Finally, given $X \in \mathcal{T}$, the space $\mathcal{Q}(X)$ of unit area quadratic differentials on X is the unit cotangent space at X, and we can denote by s_X the conditional measure induced by the holonomy measure on $\mathcal{Q}(X)$. The map $\mathcal{Q}(X) \to \mathcal{PMF}$ which associates to each quadratic differential on X the projective class of its vertical foliation pushes forward the measure s_X to a measure in the Lebesgue measure class, so we can indifferently use s_X and Lebesgue measure on $\mathcal{PMF}$ when discussing sets of full measure. For a thorough review of the different measures on $\mathcal{T}(S)$, $\mathcal{PMF}$ and related spaces, we refer the reader to Athreya, Bufetov, Eskin and Mirzakhani [1, Section 2] and Dowdall, Duchin and Masur [6, Section 3].

2.2. **Curve complex and subsurface projections.** In this section we review the properties we will use of two combinatorial objects associated with a surface, namely the *curve complex* and the *marking complex*.

We say a simple closed curve on a surface S is *essential* if it does not bound a disc, and is not parallel to a puncture or boundary component. The *curve complex* $\mathcal{C}(S)$ is a finite dimensional but locally infinite simplicial complex whose vertices are isotopy classes of essential simple closed curves on S, and whose simplices consist of collections of curves which can be realised disjointly on the surface. For the non-sporadic surfaces, the curve complex is a non-empty connected simplicial complex. In the case of a torus with one puncture or boundary component, or a sphere with four punctures or boundary components, the definition above gives a complex with no edges, so we alter the definition to connect two vertices if their corresponding curves can be realised by curves which intersect at most once (in the case of the once punctured torus) or at most twice (in the case of the four punctured sphere). In the case of the annulus the curve complex is defined to be the infinite graph with vertices consisting of arcs connecting the two boundary components of the annulus modulo isotopy fixing the endpoints with edges between two arcs if they can be realized disjointly. The curve complex of the annulus is quasi-isometric to $\mathbb{Z}$ with a quasi-isometry given by the algebraic intersection number. We define the curve complex to be empty for the remaining sporadic surfaces.

We say a subsurface $Y \subseteq S$ is *essential* if each boundary component is an essential simple closed curve in S. Given an essential subsurface $Y \subseteq S$, which is not a disc or a three-punctured sphere, one can also consider $\mathcal{C}(Y)$, the complex of curves of Y. There is a coarsely well-defined *subsurface projection* $\pi_Y \colon \mathcal{C}(S) \to \mathcal{C}(Y)$ which we now describe. Choose an element in the isotopy class of the curve γ which has the minimal possible number of intersections with Y, and then take a regular neighbourhood of the union of the boundary of Y with the intersection of the curve γ with Y, i.e. $N(\partial Y \cup (\gamma \cap Y))$. Choose a component of the boundary of this regular neighbourhood to be $\pi_Y(\gamma)$. This is coarsely well defined.

To define the annular projection $\pi(\gamma)$ of a curve γ with essential intersection with an annulus A essentially one passes to the annulus cover $\widetilde{S}$ of S given by the core curve α of A and chooses $\pi_A(\gamma)$ to be a component of the lift of γ that is an arc running from one boundary component of $\widetilde{S}$ to the other. The set of components of the lift of γ that satisfy this property form a finite diameter set in the curve complex of the annulus $\widetilde{S}$ and so the projection is coarsely well-defined. Finally the map π_A has the property that if D_α denotes the Dehn twist about α, then

$$d_{\mathcal{C}(A)}(\pi_A(D_\alpha^n(\gamma)), \pi_A(\gamma)) = 2 + |n|. \tag{1}$$

Thus, defining the projection this way achieves the desired property of recording the twisting around α. There is a natural $\mathbb{Z}$ action on $\mathcal{C}(A)$ by Dehn twisting around the core curve of the annular cover $\widehat{S}$. The group $\mathbb{Z}$ also has an inclusion into the mapping class group of S as Dehn twists around α, and so it acts on $\mathcal{C}(A)$ through this inclusion. The projection map π_A is coarsely equivariant with respect to the two $\mathbb{Z}$ actions. we will often abuse notation and write π_α to mean the subsurface projection to an annulus whose core curve is α.

A *marking* consists of a collection of simple closed curves α_i forming a maximal simplex in the curve complex, or equivalently, a pants decomposition of the surface, together with a *transverse curve* τ_i for each pants curve α_i, which is an element of the annular curve complex corresponding to α_i. The curves α_i are known as the *base curves* of the marking. We remark that the definition we give here corresponds to the definition of a *complete* marking from [28]. They consider more general markings, in which the set of base curves does not need to form a maximal simplex in $\mathcal{C}(S)$, and all base curves are not required to have a transversal. However, complete markings suffice for our purposes.

If α is a simple closed curve in S, then a *clean transverse curve* for α is a simple closed curve β, such that a regular neighbourhood of $\alpha \cup \beta$, isotoped to have minimal intersection, is either a sphere with four boundary components, or a torus with a single boundary component. A *clean marking* is a marking (α_i, τ_i), such that each transverse curve τ_i is of the form $\pi_{\alpha_i}(\beta_i)$, for some clean transverse curve β_i, which is disjoint from the union of the other base base curves $\cup\alpha_j$, for $j \neq i$. A clean marking $m' = (\alpha_i, \beta_i)$ is compatible with a marking $m = (\alpha_i, \tau_i)$, if the base of m is equal to the base of m', and for each simple closed curve α_i in the base, $d_{\alpha_i}(\tau_i, \pi_{\alpha_i}\beta_i)$ is minimal. There are only finitely many clean markings m' compatible with a given marking m.

The *marking complex* $M(S)$ is a graph whose points are clean markings, and whose edges are given by *elementary moves* as defined by Masur and Minsky [28]. These moves are called twists and flips. In a *twist*, a transverse curve β_i is replaced by the image of the transverse curve under a Dehn twist along its corresponding pants curve $D_{\alpha_i}(\beta_i)$. In a *flip*, a transverse curve β_i and its corresponding base curve α_i are interchanged, i.e. a new clean marking is chosen which is compatible with the marking formed by replacing $(\alpha_i, \pi_{\alpha_i}(\beta_i))$ with $(\beta_i, \pi_{\beta_i}(\alpha))$. The mapping class group acts on the marking complex and the space of orbits is finite. We will write d_M for the induced metric on the marking complex obtained by setting the length of each edge equal to one.

The mapping class group is finitely generated, so a choice of generating set gives rise to a word metric, in which the length of a group element is the shortest length of any product of generators representing the group element. Different generating sets give rise to quasi-isometric metrics. We shall assumed we have fixed a generating set, and we shall write $\| \cdot \|_G$, or d_G, for the word metric distance in the mapping class group. Masur and Minsky showed that the distance d_M in the marking complex is quasi-isometric to the word metric in the mapping class group.

Proposition 2.1. [28, Theorems 6.10 and 7.1] *Fix a complete clean marking m_0 and a system of generators for* $\mathrm{Mod}(S)$*. Then there exist constants C_1, C_2 such that for each $g \in \mathrm{Mod}(S)$*

$$C_1^{-1}\|g\|_G - C_2 \leqslant d_M(m_0, gm_0) \leqslant C_1\|g\|_G + C_2.$$

There is a coarsely well-defined map from the marking complex $M(S)$ to the curve complex $\mathcal{C}(S)$, which takes a marking to one of the short curves in the marking. In particular, for any essential subsurface $Y \subseteq S$, this gives us a map from the marking space $M(S)$ to $\mathcal{C}(Y)$, given by composing π and π_Y. Given markings m and n, denote by $d_Y(m, n)$ the diameter in $\mathcal{C}(Y)$ of the union of the projections of m and n. If α is a simple closed curve, then d_α will denote the distance in the curve complex of the annulus with core curve α.

Masur and Minsky [28, Theorem 6.12] proved a distance formula expressing the distance in the marking complex $M(S)$, and hence by Proposition 2.1, the distance in $\mathrm{Mod}(S)$ in the word metric, in terms of subsurface projections. We now describe their formula, using the cutoff function $\lfloor x \rfloor_A$, defined by

$$\lfloor x \rfloor_A = \begin{cases} x & \text{if } x \geqslant A \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Theorem 2.2 ([28] Quasi-distance formula)**.** *There exists a constant $A_0 > 0$, which depends only on the topology of the surface S, such that for any $A \geqslant A_0$, there are constants C_1 and C_2, which depend only on A and the topology of S, such that for any pair of clean markings m and m' in $M(S)$,*

$$C_1^{-1} d_M(m, m') - C_2 \leqslant \sum_{Y \subseteq S} \lfloor d_Y(m, m') \rfloor_A \leqslant C_1 d_M(m, m') + C_2$$

where the sum runs over all subsurfaces Y of S, including S.

2.3. **Short curves and short markings.** Given a hyperbolic surface X, there is a *systole* map from Teichmüller space $\mathcal{T}(S)$ to the curve complex $\mathcal{C}(S)$ given by sending X to a shortest curve on X. This map is coarsely well defined: there may be multiple shortest curves, but there are only finitely many choices, and they are a bounded distance apart in the curve complex, where these bounds depend only on the topology of S. This follows from the fact that by Bers' Lemma, for any surface S there is a constant L depending on S such that any hyperbolic metric on S contains a simple closed curve of length at most L, and the collar lemma says that for any simple closed curve γ of length L, there is an $\epsilon > 0$, depending on L, such that an ϵ-neighbourhood of γ is embedded, and so this bounds the number of intersections of any pair of curves of length L. In particular, for any Teichmüller geodesic γ_t, this gives a sequence of simple closed curves α_t.

A *reparameterization* of $\mathbb{R}$ is a continuous, monotonically increasing function $\phi \colon \mathbb{R} \to \mathbb{R}$, which need not be onto. We say a function f from $\mathbb{R}$ to a metric space is an *unparameterized* (K, c)-*quasigeodesic* if there is a reparameterization ϕ such that $f \circ \phi$ is a (K, c)-quasigeodesic, which may be of finite length.

Masur and Minsky [28] showed that the image of a Teichmüller geodesic under the shortest curve map is an unparameterized quasigeodesic in the curve complex. Rafi [30] showed that the composition of subsurface projection with the shortest curve map gives an unparameterized quasigeodesic in the curve complex of the subsurface.

Theorem 2.3 ([30, Theorem B])**.** *There are constants K and c, which only depend on the surface S, such that for any Teichmüller geodesic γ, and any subsurface $Y \subseteq S$, the sequence of curves $\pi_Y(\alpha_t)$ arising from the projection of the shortest curves α_t to $\mathcal{C}(Y)$ is an unparameterized (K, c)-quasigeodesic in $\mathcal{C}(Y)$.*

Given a hyperbolic surface X, let us define the *shortest marking* $m(X)$ in the following way. First, choose a pants decomposition by picking the shortest simple closed curves in the hyperbolic metric, using the greedy algorithm. To be precise, start by choosing one of the shortest curves on the surface, then choose one of the shortest curves on the complementary surface, and continue until you have a pants decomposition of the original surface. Then, for each curve α_i of the pants decomposition choose a transverse curve τ_i which is perpendicular to α_i in the hyperbolic metric. If there are multiple shortest curves, then the shortest marking may not be unique, but there are only a finite number of choices, with a bound depending on the topology of the surface. This gives a map from $\mathcal{T}(S)$ to $M(S)$, which is coarsely well-defined, and we shall write m_t for the image of a point on a Teichmüller geodesic γ_t under this map.

2.4. **Projections and isolated intervals.** Rafi [29] shows that for any Teichmüller geodesic γ, and any subsurface Y, there is a (possibly empty) interval I_Y during which Y is *isolated*, i.e. the boundary components of Y are short in the hyperbolic metric. In order to make this statement precise, let us pick a constant $\epsilon_0 > 0$ which is smaller than the Margulis constant. Given a Teichmüller geodesic $\gamma(t)$, and a simple closed curve α, we shall write $L_t(\alpha)$ for the length of α in the hyperbolic metric $\gamma(t)$.

Proposition 2.4 ([29, Corollary 3.3])**.** *Let $\epsilon_0 > 0$ be sufficiently small. Then there exists $\epsilon_1 \leqslant \epsilon_0$ such that, for any geodesic in the Teichmüller space and any curve α in S, there exists a connected (perhaps empty) interval I_α such that*

1. *for $t \in I_\alpha$, $L_t(\alpha) \leqslant \epsilon_0$;*
2. *for $t \notin I_\alpha$, $L_t(\alpha) \geqslant \epsilon_1$.*

Outside the active interval I_α, the map from the Teichmüller geodesic to the curve complex of the annulus corresponding to α is coarsely constant.

Proposition 2.5 ([29, Proposition 3.7])**.** *There is a constant K, depending only on the topology of the surface S, and the choices for the constants ϵ_0 and ϵ_1, such that if $[r, s] \cap I_\alpha = \emptyset$, then*

$$d_\alpha(m_r, m_s) \leqslant K.$$

In the next section we show that the length of the active interval for an annulus is roughly log of the projection distance of the endpoints of the geodesic into the subsurface.

2.5. **Excursions and twist parameter.** The material in this section is due to Rafi [29, 30]. However, we need versions of his results in terms of the excursion parameter, and we use some of the contents of the proofs, not just the main stated results, so we write out all of the details for the convenience of the reader.

A horoball H in the hyperbolic plane is a subset of the plane which in the Poincaré disc model corresponds to a Euclidean disc whose boundary circle is tangent to the boundary at infinity. Given a horoball H and a geodesic γ which spends a finite amount of time in H, let us define the *excursion* $E(\gamma, H)$ of γ in H as the "relative visual size" of the set of rays which go deeper than γ inside H. Namely, consider a basepoint X_0 on the Teichmüller disc in $\mathcal{T}$, and let γ_H be the geodesic through X_0 which tends to the cusp of H, and γ_T a geodesic through X_0 which is tangent to H. Let ϕ_0 be the angle between γ and γ_H, and ϕ_{max} be the angle between γ_H and γ_T (see Figure 2). Then

Definition 2.6. *The* excursion *of the geodesic γ in the horoball H is defined as*

$$E(\gamma, H) := \frac{\phi_{max}}{\phi_0}. \tag{3}$$

It turns out that $E(\gamma, H)$ is, up to an additive error, also the hyperbolic length of the projection of $\gamma \cap H$ to the complement of H.

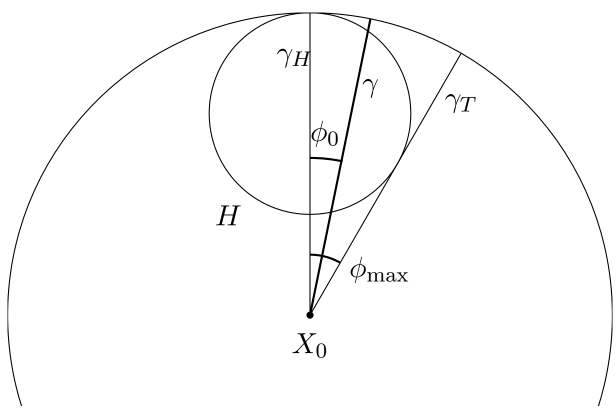

FIGURE 2. Excursion in the horoball H.

Let (X, q) be a quadratic differential on X, and α a simple closed curve on X. The choice of q determines a Teichmüller geodesic γ and a pair (F^+, F^-) of contracting and expanding foliations. Each t determines a new quadratic differential q_t and hence a flat metric on X, which we will call the q_t-metric.

For a given t, α is realized by a family of parallel flat geodesics, and we will denote as β_t the perpendicular to α in the q_t-metric. The *twist parameter* $tw_t^+(\alpha)$ is the highest intersection number between a leaf of F^+ and the transversal β_t, and similarly we define $tw_t^-(\alpha)$.

Given a simple closed curve α corresponding to metric cylinder, there is a unique rotation $e^{i\theta_\alpha}$ which takes the metric cylinder to a vertical metric cylinder. The endpoint of the geodesic ray corresponding to the quadratic differential $e^{i\theta_\alpha} q$ determines a point ξ_α on the boundary at infinity of the Teichmüller disc $\mathbb{D}$. We shall write $H_\epsilon(\alpha)$ as the set of points in the disc for which α is short in the flat metric:

$$H_\epsilon(\alpha) := \{q \in \mathbb{D} \ : \ell_q^2(\alpha) \leqslant \epsilon\}.$$

As seen in the disc, this set is a horoball tangent to the boundary at infinity at ξ_α. The fundamental estimate is the following:

Proposition 2.7. *Let $H = H_\epsilon(\alpha)$ as above, and let t_1 and t_2 respectively be the entry time and exit time from H (i.e. $t_1 \leqslant t_2$) along the Teichmüller geodesic γ. Let moreover A be the area of the maximal flat cylinder in (X, q_0) with core curve α. Then we have, up to universal multiplicative and additive constants,*

$$tw^-_{t_2}(\alpha) - tw^-_{t_1}(\alpha) \asymp \frac{A}{\epsilon} E(\gamma, H).$$

Proof. Consider the universal cover of the flat cylinder corresponding to α at time t, in the flat metric q_t. We shall assume that the contracting foliation is vertical, and the expanding foliation is horizontal. Let ℓ_t be the length of α at time t, and let θ_t be the angle α_t makes with the vertical contracting foliation, as illustrated below in Figure 3.

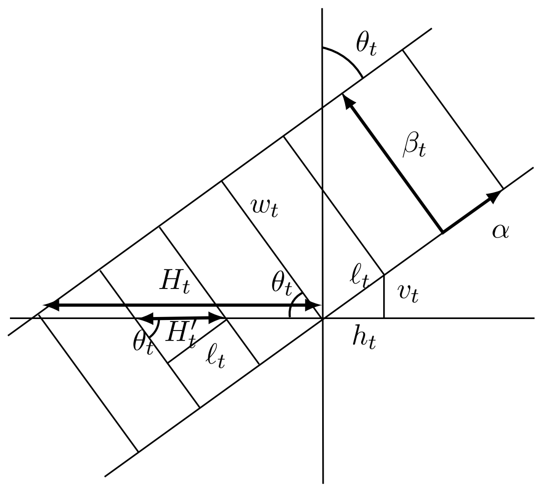

FIGURE 3. Estimating intersections in the flat annulus.

Let h_t and v_t be the horizontal and vertical lengths of ℓ_t in the q_t metric, i.e.

$$h_t = h_0 e^t = \ell_0 \sin\theta_0 e^t$$

$$v_t = v_0 e^{-t} = \ell_0 \cos\theta_0 e^{-t}.$$

Let w_t be the length of β_t, which is the width of the flat annulus. Let H_t be the length of the intersection of a leaf of the horizontal foliation with the universal cover of the flat annulus, and let H'_t be the length of the intersection of the horizontal leaf with two adjacent translates of β_t.

Up to constant additive error, $tw^-_t(\alpha)$, which is the maximum number of intersections between the horizontal leaf of the foliation and β, is given by H_t/H'_t. Therefore

$$tw^-_t(\alpha) = \frac{H_t}{H'_t} + O(1) = \frac{w_t \sin\theta_t}{\ell_t \cos\theta_t} + O(1).$$

The area of the annulus is $A = w_t \ell_t$, and $\tan\theta_t = \tan\theta_0 e^{2t}$, so this implies that

$$tw^-_t(\alpha) = \frac{A}{\ell_t^2} \tan\theta_0 e^{2t} + O(1). \tag{4}$$

The total length of α is given by

$$\ell_t^2 = h_t^2 + v_t^2 = \ell_0^2(\sin^2\theta_0 e^{2t} + \cos^2\theta_0 e^{-2t}), \tag{5}$$

and recall that we choose t_i such that $\ell_{t_i}^2 = \epsilon$, which by (4) implies

$$tw_{t_2}^-(\alpha) - tw_{t_1}^-(\alpha) = \frac{A}{\epsilon}\tan\theta_0(e^{2t_2} - e^{2t_1}) + O(1). \tag{6}$$

Note that by definition the t_i are solutions to the equation

$$\ell_{t_i}^2 = \ell_0^2(\sin^2\theta_0 e^{2t_i} + \cos^2\theta_0 e^{-2t_i}) = \epsilon \qquad i = 1, 2$$

If we set $X_i := e^{2t_i}$, then X_i are the solutions to

$$X^2 - \frac{\epsilon^2}{\ell_0^2\sin^2\theta_0}X + \frac{1}{\tan^2\theta_0} = 0 \tag{7}$$

hence

$$e^{2t_2} - e^{2t_1} = X_2 - X_1 = \sqrt{\frac{\epsilon^2}{\ell_0^4\sin^4\theta_0} - \frac{4}{\tan^2\theta_0}} \tag{8}$$

and putting (6) and (8) together

$$tw_{t_2}^-(\alpha) - tw_{t_1}^-(\alpha) = \frac{A\tan\theta_0}{\epsilon}(e^{2t_2} - e^{2t_1}) + O(1) = \frac{A}{\epsilon}\sqrt{\frac{\epsilon^2}{\ell_0^4\sin^2\theta_0\cos^2\theta_0} - 4} + O(1). \tag{9}$$

Let us now relate this quantity to the excursion in the horoball H.

Lemma 2.8. *Let ϕ_{max} be the angle between a geodesic γ_T tangent to H and the geodesic γ_H which goes straight into the cusp of H. Then*

$$\sin\phi_{max} = \frac{\epsilon}{\ell_0^2},$$

where ℓ_0 is the length of α at time $t = 0$.

Proof. When $\theta_0 = \theta_{max}$ then the geodesic is tangent to the horoball H, hence $t_1 = t_2$ in equation (8), so

$$\frac{\epsilon^2}{\ell_0^4\sin^4\theta_{max}} = \frac{4}{\tan^2\theta_{max}}.$$

The claim follows by recalling that a rotation of angle θ in the flat metric picture corresponds to multiplying the quadratic differential by $e^{2i\theta}$, hence $\phi_{max} = 2\theta_{max}$. □

The proposition now follows easily from the lemma, equation (9) and the fact that $\phi_0 = 2\theta_0$:

$$tw_{t_2}^-(\alpha) - tw_{t_1}^-(\alpha) = \frac{2A}{\epsilon}\frac{\sin\phi_{max}}{\sin\phi_0}\sqrt{1 - \frac{\sin^2\phi_0}{\sin^2\phi_{max}}} + O(1) \asymp \frac{A}{\epsilon}E(\gamma, H)$$

where in the last equality we used equation (3) and the fact that $\sin\phi \asymp \phi$ (note that we can assume $\sin\phi_0 \leq \frac{1}{2}\sin\phi_{max}$, otherwise the claim is trivially verified).

□

Remark. *Note that one can also relate the twist parameter to the time spent by the geodesic inside the horoball, namely*

$$tw_{t_2}^-(\alpha) - tw_{t_1}^-(\alpha) \asymp \frac{A}{\epsilon}(e^{t_2 - t_1} - e^{t_1 - t_2}).$$

The distance between the projections from the marking complex to the complex of the annulus can be compared to the excursion in the horoball:

Proposition 2.9. *Let $\epsilon > 0$ sufficiently small, and (X_0, q) a unit area quadratic differential, which determines the geodesic ray γ_t. Let A be the q-area of the maximal flat cylinder with core curve α, and suppose that α is not short in the q-metric (i.e. $\ell_q^2(\alpha) \geq \epsilon$). If the geodesic γ crosses the horoball $H = H_\epsilon(\alpha)$ and t is larger than the exit time of γ from H, then*

$$d_\alpha(m_0, m_t) \asymp \frac{A}{\epsilon} E(\gamma, H)$$

where m_t is the shortest marking on γ_t, and the quasi-isometry constants depend only on X_0, ϵ and the topology of S.

Before proving the proposition, let us recall the definition of extremal length:

$$\mathrm{Ext}_\sigma(\alpha) := \sup_\rho \frac{(\ell_\rho(\alpha))^2}{A(\rho)}$$

where the sup is taken over all metrics ρ in the same conformal class as σ. For any quadratic differential q with area 1 and any curve α,

$$(\ell_q(\alpha))^2 \leqslant \mathrm{Ext}_\sigma(\alpha) \leqslant \frac{L_\sigma(\alpha)}{2} e^{L_\sigma(\alpha)/2}$$

where the left-hand side is by definition, while the right-hand side is due to Maskit [23], and $L_\sigma(\alpha)$ is the length of α in the hyperbolic metric corresponding to the conformal structure σ.

Recall $tw_q^\pm$ denotes the twist parameter in the flat metric associated to q, as defined in the previous section. Analogously, given a hyperbolic metric σ on S and a simple closed curve α, we can define a twist parameter $tw_\sigma^\pm(\alpha)$ with respect to the hyperbolic metric by taking a curve β perpendicular to α with respect to the hyperbolic metric and letting

$$tw_\sigma^\pm(\alpha) := i(F^\pm, \beta).$$

The following proposition of Rafi relates the two twist parameters:

Proposition 2.10 ([29, Theorem 4.3])**.** *The two twist parameters are the same up to an additive error comparable to $1/L_\sigma(\alpha)$. That is,*

$$tw_\sigma^\pm(\alpha) = tw_q^\pm(\alpha) + O\left(\frac{1}{L_\sigma(\alpha)}\right).$$

Proof of Proposition 2.9. Let us choose ϵ_0 in such a way that if $\ell_q^2(\alpha) = \epsilon$, then $L_\sigma(\alpha) \geqslant \epsilon_0$. Let t_1 and t_2 be the times when the q_t-length of α is exactly ϵ, and let $t > t_2$. By the previous choice, $L_{\sigma_{t_i}}(\alpha) \geqslant \epsilon_0$ for $i = 1, 2$, so by Proposition 2.4 $[0, t_1]$ and $[t_2, t]$ are disjoint from I_α, hence by Proposition 2.5

$$d_\alpha(m_0, m_t) = d_\alpha(m_{t_1}, m_{t_2}) + O(1).$$

On the other hand, the progress in subsurface projection across the horoball is comparable to the progress in the twist parameter for the hyperbolic metric,

$$d_\alpha(m_{t_1}, m_{t_2}) = |i_\alpha(\beta_{t_1}, F^+) - i_\alpha(\beta_{t_2}, F^+)| + O(1) = |tw_{\sigma_1}^+(\alpha) - tw_{\sigma_2}^+(\alpha)| + O(1)$$

and by Proposition 2.10 it is also comparable to the progress in the twist parameter defined via the flat metric:

$$|tw_{\sigma_1}^+(\alpha) - tw_{\sigma_2}^+(\alpha)| = |tw_{q_1}^+(\alpha) - tw_{q_2}^+(\alpha)| + O\left(\frac{1}{L_{\sigma_1}(\alpha)}\right) + O\left(\frac{1}{L_{\sigma_2}(\alpha)}\right)$$

and since $L_{\sigma_i}(\alpha) \geqslant \epsilon_0$

$$d_\alpha(m_0, m_t) = |tw_{q_1}^+(\alpha) - tw_{q_2}^+(\alpha)| + O(1).$$

Finally, by Proposition 2.7 the twist is comparable to the excursion, thus

$$d_\alpha(m_0, m_t) \asymp \frac{A}{\epsilon} E(\gamma, H).$$

□

2.6. **Coarse monotonicity for the word metric.** In [30], Rafi shows the following non-backtracking or reverse triangle inequality for subsurface projections along a Teichmüller geodesic. Recall that given a Teichmüller geodesic γ_t we write m_t for the shortest marking at γ_t, and we write $d_Y(m_s, m_t)$ to mean the distance in the curve complex $\mathcal{C}(Y)$ between the images of m_s and m_t under subsurface projection to Y.

Theorem 2.11 ([30, Theorem 6.1])**.** *There exists a constant C, only depending on the topology of S, such that for every Teichmüller geodesic γ, and every subsurface Y,*

$$d_Y(m_r, m_s) + d_Y(m_s, m_t) \leqslant d_Y(m_r, m_t) + C, \tag{10}$$

for all constants $r \leqslant s \leqslant t$.

The above theorem along with the Masur-Minsky quasi-distance formula (2.2) implies that the distance in the marking complex is coarsely monotonic along a Teichmüller ray.

Proposition 2.12. *There exists constants $C_1 > 0$ and C_2 that depend only on S such that along a Teichmüller geodesic γ_t, for $0 < s < t$ the distance in the marking complex satisfies*

$$d_M(m_0, m_s) \leqslant C_1 d_M(m_0, m_t) + C_2.$$

Proof. Let C be the constant in Rafi's reverse triangle inequality, Theorem 2.11. Assume $0 < s < t$, then (10) implies

$$d_Y(m_0, m_t) \geqslant d_Y(m_0, m_s) - C \tag{11}$$

for all subsurfaces $Y \subseteq S$. The Masur-Minsky quasi-distance formula (Theorem 2.2) holds for all floor constants sufficiently large, though the quasi-isometry constants depend on A. Choose a floor constant $A > 2C$, and let K_1 and K_2 be the associated quasi-isometry constants. By the definition of the floor function, if $\lfloor x \rfloor_A$ is non zero, then $x \geqslant A$. This implies that $x - A/2 \geqslant x/2$, and as the floor function is monotonic,

$$\lfloor x - A/2 \rfloor_A \geqslant \lfloor x/2 \rfloor_A. \tag{12}$$

As we have chosen $A > 2C$, combining (11) and (12) implies

$$\lfloor d_Y(m_0, m_t) \rfloor_A \geqslant \lfloor \tfrac{1}{2} d_Y(m_0, m_s) \rfloor_A, \tag{13}$$

again for all subsurfaces $Y \subseteq S$. Now summing (13) over all subsurfaces $Y \subseteq S$, the quasi-distance formula implies

$$d_M(m_0, m_t) \geqslant \frac{1}{K_1} \left(\sum \lfloor \tfrac{1}{2} d_Y(m_0, m_s) \rfloor_A - K_2 \right).$$

By definition of the floor function, $\lfloor \frac{1}{2} x \rfloor_A = \frac{1}{2} \lfloor x \rfloor_{2A}$, so

$$d_M(m_0, m_t) \geqslant \frac{1}{2K_1} \left(\sum \lfloor d_Y(m_0, m_s) \rfloor_{2A} - 2K_2 \right).$$

The quasi-distance formula holds for all A sufficiently large, so in particular holds for $2A$, though with different quasi-isometry constants, which we shall denote K_3 and K_4. This implies that

$$d_M(m_0, m_t) \geqslant \frac{1}{2K_1 K_3} \left(d_M(m_0, m_s) - K_3 K_4 - K_2 \right)$$

whence the result. □

2.7. **Projection to closest Teichmüller lattice point.** Let q be a quadratic differential, let q_t be the image of q under the Teichmüller geodesic flow after time t, and let X_t be the image of q_t in $\mathcal{T}$. The orbit of X_0 under the mapping class group is called a *Teichmüller lattice*, and let $g_t X_0$ be a choice of closest lattice point in $\mathcal{T}$ to X_t, i.e. such that

$$d_{\mathcal{T}}(g_t X_0, X_t) \leqslant d_{\mathcal{T}}(g X_0, X_t) \text{ for all } g \in \mathrm{Mod}(S).$$

For any given point X_t, there are at most finitely many closest lattice points, however it is possible that the number of closest lattice points increases as you choose points deeper in the thin part. Let m_t be a shortest marking on X_t, and $\|\cdot\|_G$ the word metric on the mapping class group with respect to some choice of generators.

Lemma 2.13. *If X_0 and X_t both lie in the thick part $\mathcal{T} \setminus \mathcal{T}_\epsilon$, then*

$$\|g_t\|_G \asymp d_M(m_0, m_t)$$

where the quasi-isometry constants only depend on X_0, the choice of ϵ and the generating set for the mapping class group.

Proof. Let K_1 be the diameter of the thick part $\mathcal{T} \setminus \mathcal{T}_\epsilon$ in moduli space; then, by definition there exists a group element g such that in Teichmüller space $d_{\mathcal{T}}(gX_0, X_t) \leqslant K_1$, so by definition of g_t

$$d_{\mathcal{T}}(g_t X_0, X_t) \leqslant K_1.$$

Hence by group invariance

$$d_{\mathcal{T}}(X_0, g_t^{-1} X_t) \leqslant K_1.$$

In the Teichmüller ball of radius K_1 only finitely many markings appear as short markings, hence there exists K_2, depending only on K_1, and the surface S, such that the distance in the marking complex is bounded:

$$d_M(m_0, g_t^{-1} m_t) \leqslant K_2.$$

As a consequence,

$$|d_M(m_0, m_t) - d_M(m_0, g_t m_0)| \leqslant d_M(g_t m_0, m_t) = d_M(m_0, g_t^{-1} m_t) \leqslant K_2.$$

Finally, the distance in the word metric $\|g_t\|_G$ is quasi-isometric to the distance $d_M(m_0, g_t m_0)$ in the marking complex by Proposition 2.1. □

By combining the previous lemma with the coarse monotonicity statement of Proposition 2.12, we get that the word length of the closest point projection to the Teichmüller lattice is coarsely monotone along the thick part of a Teichmüller ray:

Proposition 2.14. *There exists constants $C_1 > 0$ and C_2, that depend only on X_0 and ϵ_0 and the choice of generators, such that along a Teichmüller geodesic γ_t, for $0 < s < t$ the word metric satisfies*

$$\|g_s\|_G \leqslant C_1 \|g_t\|_G + C_2$$

whenever γ_0, γ_s and γ_t all lie in the thick part $\mathcal{T} \setminus \mathcal{T}_{\epsilon_0}$.

3. Lebesgue measure sampling

The goal of this section is to study the asymptotic behaviour of typical Teichmüller geodesics with respect to Lebesgue measure, proving Theorem 1.1. More precisely, we want to keep track of short curves in the flat metric as the metric changes under the action of Teichmüller flow, and prove an asymptotic result, Theorem 3.3. In Section 3.1 we recall results of Masur [26] and Eskin and Masur [7] which show that the growth rate of the number of metric cylinders with area bounded below is quadratic. In Section 3.2 we consider the function given by the sum of the squares of the reciprocals of the short curves, and show that the average value of this function tends to infinity along almost every Teichmüller geodesic with respect to Lebesgue measure. Then in Section 3.3 we show that this function gives a lower bound for the average of the sums of the excursions along the geodesic. Finally in Section 3.4 we show that the sum of the excursions is a lower bound for the word metric along the Teichmüller geodesic, and so the word metric along the geodesic has faster than linear growth, which completes the proof of the Theorem 1.1.

3.1. **Metric cylinders with bounded area.** Let q be a quadratic differential of unit area. A *metric cylinder* for q is a cylinder in the flat metric associated to q which is the union of freely homotopic closed trajectories of q. We shall label each metric cylinder by the homotopy class α of the corresponding closed trajectory.

Let us now fix some $0 < \delta < 1$, and let $C_q(\delta)$ be the set of metric cylinders for the q-metric with area bounded below by δ. Moreover, let us denote by $C_q(\delta, \epsilon)$ the set of cylinders whose area is bounded below by δ and whose core curve has length shorter than the square root of ϵ:

$$C_q(\delta, \epsilon) := \{\alpha \in C_q(\delta) \; : \; \ell_q^2(\alpha) \leqslant \epsilon\}.$$

Lemma 3.1. *Suppose $\epsilon < \delta$. Then any two distinct elements of $C_q(\delta, \epsilon)$ are disjoint on q. As a corollary, the cardinality of $C_q(\delta, \epsilon)$ is bounded above by a constant which depends only on the topology of S.*

Proof. We follow the argument in [26, Lemma 2.2]. Denote by α the core curve of some cylinder which belongs to $C_q(\delta, \epsilon)$. Since the metric cylinder of α has area $A(\alpha) \geqslant \delta$, any curve τ which crosses α is such that $\delta \leqslant \ell_q(\alpha)\ell_q(\tau) \leqslant \ell_q(\tau)\sqrt{\epsilon}$, hence $\ell_q(\tau) > \sqrt{\epsilon}$, so τ cannot belong to $C_q(\delta, \epsilon)$. □

Given the quadratic differential q, let us denote as $N_q(\delta, T)$ the number of cylinders in the q-metric which have area bounded below by δ and length smaller than T. As Eskin and Masur showed, $N_q(\delta, T)$ grows quadratically as a function of T:

Theorem 3.2. *There exists $0 < \delta < 1$ and a constant $c_\delta > 0$ such that, for almost every quadratic differential q of unit area, we have*

$$\lim_{T\to\infty} \frac{N_q(\delta, T)}{T^2} = c_\delta.$$

Proof. Let $0 < \delta < 1$. By the general counting argument of Eskin-Masur [7, Theorem 2.1] applied to the set of metric cylinders with area bounded below by δ, we get the existence of the limit c_δ almost everywhere. On the other hand, by [26, Proposition 2.5], for every quadratic differential there exists some $\delta > 0$ such that $\liminf_{T\to\infty} \frac{N_q(\delta,T)}{T^2} > 0$, so the constant c_δ must be positive for some δ. □

A finer statement, at least in the case of translation surfaces, is due to Vorobets [34].

3.2. **Asymptotic length of short curves.** Let us now quantify the idea of keeping track of short curves in the flat metric. For the rest of the paper, we will fix some $\delta > 0$ for which Theorem 3.2 holds, and some $\epsilon < \delta$. Let us define the function $L : \mathcal{QM} \to \mathbb{R}$ as

$$L(q) := \sum_{\alpha \in C_q(\delta,\epsilon)} \frac{1}{\ell_q^2(\alpha)}.$$

Note that by Lemma 3.1 the number of terms in the sum is always finite, so the function is well-defined. Let us fix denote by q_t the image of the quadratic differential q under the Teichmüller geodesic flow after time t. Our goal is to prove that the ergodic average of L is infinite:

Theorem 3.3. *For μ_{hol}-a.e. quadratic differential q of unit area, we have*

$$\lim_{T\to\infty} \frac{\int_0^T L(q_t)\ dt}{T} = \infty.$$

In the proof of Theorem 3.3, we will make use of the following relations between metric cylinders and the geometry of Teichmüller discs. Let us fix a base point q_0 in the space of quadratic differentials, and call D_{q_0} the Teichmüller disc given by the $SL_2(\mathbb{R})$-orbit of q_0. For every metric cylinder α on q_0, there is an angle θ_α such that α is vertical in the quadratic differential $e^{i\theta_\alpha}q_0$. The angle θ_α determines a point in the circle at infinity of D_{q_0}. For each metric cylinder on q_0 with core curve α, let us define the set

$$H_\epsilon(\alpha) := \{q \in D_{q_0} \ : \ \ell_q^2(\alpha) \leqslant \epsilon\},$$

of points in the Teichmüller disc for which the length of α is less than the square root of ϵ. Recall the metric induced on D_{q_0} by the Teichmüller metric is the hyperbolic metric of constant curvature -4, and $H_\epsilon(\alpha)$ is a horoball for that metric.

Lemma 3.4. *The Euclidean diameter s of the horoball $H_\epsilon(\alpha)$ is*

$$s = \frac{2\epsilon}{\epsilon + \ell_{q_0}^2(\alpha)}$$

where $\ell_{q_0}(\alpha)$ is the length of α in the flat metric associated to the quadratic differential q_0.

Proof. By integrating the hyperbolic metric of curvature -4 we have

$$d(q_0, H_\epsilon(\alpha)) = \int_0^{1-s} \frac{dx}{1-x^2} = \frac{1}{2}\log\frac{2-s}{s}$$

and, since the Teichmüller map exponentially shrinks the curve α,

$$e^{-2d(q_0,H_\epsilon(\alpha))}\ell_{q_0}^2(\alpha) = \epsilon$$

hence the claim. □

We will need the following estimate from elementary Euclidean geometry:

Lemma 3.5. *In the unit disc, let $\theta(r,R)$ be the angle at the center of the disc corresponding to the intersection of the circle of radius $R \geqslant \frac{1}{2}$ centered at the origin, with a circle of radius $r \leqslant \frac{1}{2}$ tangent to the boundary, with $R + 2r - 1 \geqslant 0$. Then there is a constant K such that*

$$\frac{1}{K}\sqrt{(1-R)(R+2r-1)} \leqslant \theta(r,R) \leqslant K\sqrt{(1-R)(R+2r-1)}.$$

Proof. By the law of cosines, $r^2 = (1-r)^2 + R^2 - 2R(1-r)\cos(\theta/2)$. The claim follows by standard algebraic manipulation and approximation. □

Let $q_{t,\theta}$ denote the quadratic differential given by flowing the quadratic differential $e^{i\theta}q_0$ for time t.

Lemma 3.6. *For almost every quadratic differential q_0 there exists a constant $c > 0$, such that for each $\epsilon > 0$ there exists a time t_ϵ such that*

$$\sum_{\alpha\in C_{q_0}(\delta)} \mathrm{Leb}(\{\theta \in [0, 2\pi] \ : \ q_{t,\theta} \in H_\epsilon(\alpha)\}) \geqslant c\epsilon \qquad \forall\ t \geqslant t_\epsilon,$$

where Leb *denotes Lebesgue measure on the circle.*

Proof. Let $\frac{1}{2} < R < 1$, and consider the set of horoballs of the collection $H_\epsilon(\alpha)$ with $\alpha \in C_{q_0}(\delta)$ and Euclidean diameter $s \geqslant \frac{3}{2}(1-R)$. By Lemma 3.4, these horoballs correspond precisely to metric cylinders with core curve α such that

$$\ell^2_{q_0}(\alpha) \leqslant \frac{3R+1}{3(1-R)}\epsilon.$$

By Theorem 3.2, the number of such cylinders is, for R large, at least $\frac{c_\delta}{2}\frac{3R+1}{3(1-R)}\epsilon$. By Lemma 3.5, every corresponding horoball intersects the circle of Euclidean radius R centered at the origin in an arc of visual angle

$$\theta \geqslant \frac{1}{K\sqrt{2}}(1-R)$$

and by Lemma 3.1 every quadratic differential belongs to at most a universally bounded number M of horoballs, hence the total visual angle is at least $\frac{c_\delta}{6KM\sqrt{2}}\epsilon$. □

In order to prove Theorem 3.3, let us first define a discrete version of L. Namely, for each n and α we denote as $H_n(\alpha)$ the horoball

$$H_n(\alpha) := \{q \in D_{q_0} \ : \ell^2_q(\alpha) \leqslant 2^{-n}\epsilon\}.$$

Now, the function $\Psi : \mathcal{QM} \to \mathbb{R}$ is defined as

$$\Psi(q) := \sum_{\alpha \in C_q(\delta)} \sum_{n=1}^{\infty} 2^n \chi_{H_n(\alpha)}.$$

It is easy to see that Ψ is bounded above by a multiple of L:

Lemma 3.7. *For each quadratic differential q, we have*

$$\Psi(q) \leqslant 4\epsilon L(q).$$

Proof. Let $\alpha \in C_q(\delta)$ be a short curve on q: then there exists a positive integer M such that

$$2^{-M}\epsilon \leqslant \ell^2_q(\alpha) \leqslant 2^{-M+1}\epsilon.$$

Now, since q lies in $H_1(\alpha) \cup \cdots \cup H_M(\alpha)$,

$$\sum_{n=1}^{\infty} 2^n \chi_{H_n(\alpha)} \leqslant 1 + 2 + \cdots + 2^M \leqslant 2 \cdot 2^M \leqslant \frac{4\epsilon}{\ell^2_q(\alpha)}$$

and summing over α yields the claim. □

Proof of Theorem 3.3. By Lemma 3.7, it is enough to prove the statement for Ψ. Let us now truncate the function Ψ by defining, for each N,

$$\Psi_N(q) := \sum_{\alpha \in C_q(\delta)} \sum_{n=1}^{N} 2^n \chi_{H_n(\alpha)}.$$

Let us now fix N. By Lemma 3.1, Ψ_N is bounded on the moduli space $\mathcal{MQ}$ of unit area quadratic differentials, hence μ_{hol}-integrable; by ergodicity of the geodesic flow, for a generic Teichmüller disc for almost all radial directions θ the ergodic average of Ψ_N along the flow tends to its integral:

$$\lim_{T\to\infty} \int_0^T \frac{\Psi_N(q_{t,\theta})\ dt}{T} = \int_{\mathcal{MQ}} \Psi_N(q)\ d\mu_{hol} \qquad \text{for a.e. } \theta.$$

Then, if we integrate both sides w.r.t. to the angular measure $d\theta$ and apply the dominated convergence theorem,

$$\lim_{T\to\infty}\int_{S^1} d\theta \int_0^T \frac{\Psi_N(q_{t,\theta})\ dt}{T} = \int_{\mathcal{MQ}} \Psi_N(q)\ d\mu_{hol}$$

and by Fubini

$$\lim_{T\to\infty}\frac{\int_0^T dt \int_{S^1} \Psi_N(q_{t,\theta})\ d\theta}{T} = \int_{\mathcal{MQ}} \Psi_N(q)\ d\mu_{hol}.$$

Now, by Lemma 3.6, for each $t > T_{2^{-N}}$

$$\int_{S^1} \Psi_N(q_{t,\theta})\ d\theta \geqslant \sum_{n=1}^{N} 2^n \cdot c2^{-n} = cN$$

hence

$$\int_{\mathcal{MQ}} \Psi_N(q)\ d\mu_{hol} = \lim_{T\to\infty}\frac{\int_0^T dt \int_{S^1} \Psi_N(q_{t,\theta})\ d\theta}{T} \geqslant cN.$$

Since the previous estimate works for all N, then also

$$\int_{\mathcal{MQ}} L(q)\ d\mu_{hol} = \infty$$

hence the ergodic average tends to infinity almost everywhere:

$$\lim_{T\to\infty}\int_0^T \frac{L(q_t)\ dt}{T} = \int_{\mathcal{MQ}} L(q)\ d\mu_{hol} = \infty \qquad \text{for a.e. } q \in \mathcal{MQ}.$$

□

3.3. **Average excursion.** Let us now turn the asymptotic estimate of the previous section into an asymptotic about excursions. If q is a quadratic differential, let us denote as γ_q the corresponding Teichmüller geodesic ray. We now define the concept of *total excursion* traveled by the geodesic γ_q inside the horoballs up to time T:

Definition 3.8. *Given a quadratic differential q, the* total excursion $E(q,T)$ *is the sum of all excursions in all horoballs crossed by the geodesic ray γ_q up to time T:*

$$E(q,T) := \sum_{\gamma_q([0,T])\cap H_\epsilon(\alpha)\neq\emptyset} E(\gamma_q, H_\epsilon(\alpha)).$$

Our goal is to prove that also the average total excursion is infinite.

Theorem 3.9. *For μ_{hol}-almost every quadratic differential q of unit area, we have*

$$\lim_{T\to\infty}\frac{E(q,T)}{T} = \infty.$$

Theorem 3.9 follows from Theorem 3.3 and the following

Proposition 3.10. *Let q be a quadratic differential with geodesic ray γ_q, and let $T > 0$ be such that both q and $\gamma_q(T)$ lie outside all horoballs of the type $H_\epsilon(\alpha)$. Then*

$$\int_0^T L(q_t)\ dt \leqslant \frac{C}{\epsilon} E(q,T)$$

for some universal constant C.

Proof. Let $\alpha \in C_q(\delta)$ be a curve which has become short before time T, i.e. such that $\gamma_q([0,T]) \cap H_\epsilon(\alpha)$ is non-empty. Let T_1 be the time the geodesic enters $H_\epsilon(\alpha)$, and T_2 the time the geodesic exits. Moreover, let N be the maximum integer k such that the geodesic enters $H_k(\alpha)$. Note that there is a universal constant C_1 such that for each $n \geqslant 1$ and each α

$$\mathrm{Leb}(\{t \in [0,T] \ : \ q_t \in H_{n+1}(\alpha) \setminus H_n(\alpha)\}) \leqslant C_1.$$

Then

$$\int_{T_1}^{T_2} \frac{1}{\ell^2_{q_t}(\alpha)} \, dt \leqslant \sum_{n=1}^{N} \frac{2^n}{\epsilon} \mathrm{Leb}(\{t \in [0,T] \ : \ q_t \in H_{n+1}(\alpha) \setminus H_n(\alpha)\}) \leqslant \frac{C_1 \cdot 2^{N+1}}{\epsilon}.$$

In order to compare the right hand side with the excursion, let us denote by $\tilde{\epsilon}$ the smallest value of $\ell^2_{q_t}(\alpha)$ along the geodesic ray γ_q. By the definition of N, we have $\tilde{\epsilon} \asymp 2^{-N}\epsilon$. Now, by the definition of excursion and Lemma 2.8,

$$E(\gamma_q, H_\epsilon(\alpha)) = \frac{\phi_{max}}{\phi_0} \asymp \frac{\sin \phi_{max}}{\sin \phi_0} = \frac{\epsilon}{\tilde{\epsilon}} \asymp 2^N$$

(where all the approximate equalities hold up to multiplicative constants), hence the claim follows. □

Remark. *A precise analysis of how $E(q,T)$ grows along Leb-typical geodesics is carried out in* [11]. *It culminates in a strong law analogous to the one established by Diamond and Vaaler for continued fractions* [5].

3.4. **The word metric.** Let us complete the proof of Theorem 1.1 and Theorem 1.3 for the Lebesgue measure by proving that the word metric is bounded below by the total excursion. Let us pick ϵ_0 to define the thick part as in section 2.4, and let us choose δ so that Theorem 3.2 holds. Finally, we choose ϵ so that if X belongs to the thick part $\mathcal{T} \setminus \mathcal{T}_{\epsilon_0}$ and α is the core curve of a metric cylinder of q-area larger than δ on X, then $\ell^2_q(\alpha) \geqslant \epsilon$.

Let X_0 lie in the thick part $\mathcal{T} \setminus \mathcal{T}_{\epsilon_0}$, and let γ be a Teichmüller geodesic with $\gamma(0) = X_0$. Recall for each time t, g_t is a closest point projection of $\gamma(t)$ to the Teichmüller lattice.

Proposition 3.11. *If $\gamma(T)$ lies in the thick part $\mathcal{T} \setminus \mathcal{T}_{\epsilon_0}$, then*

$$\|g_T\|_G \geqslant C_1 E(\gamma, T) - C_2$$

where the constants depend only on X_0, the choice of ϵ_0 and the choice of generating set for $Mod(S)$.

Proof. Since $\gamma(0)$ and $\gamma(T)$ lie in the thick part, by Lemma 2.13

$$\|g_T\|_G \asymp d_M(m_0, m_T).$$

By the Masur-Minsky quasi-distance formula (Theorem 2.2), for any B large enough

$$d_M(m_0, m_T) \asymp \sum_{Y \subseteq S} \lfloor d_Y(m_0, m_T) \rfloor_B \geqslant \sum_{\gamma([0,T]) \cap H_\epsilon(\alpha) \neq \emptyset} \lfloor d_\alpha(m_0, m_T) \rfloor_B$$

where on the right-hand side we only consider projections to annuli of area bounded below, and whose core curve becomes short before time T. Now by Proposition 2.9, for some constants K_1 and K_2,

$$d_\alpha(m_0, m_T) \geqslant K_1 E(\gamma, H_\epsilon(\alpha)) - K_2$$

so if $B \geqslant K_2$

$$\lfloor d_\alpha(m_0, m_T) \rfloor_B \geqslant \lfloor K_1 E(\gamma, H_\epsilon(\alpha)) - K_2 \rfloor_B \geqslant \frac{K_1}{2} \lfloor E(\gamma, H_\epsilon(\alpha)) \rfloor_{\frac{2B}{K_1}}$$

and we can choose $\tilde{\epsilon}$ a bit smaller than ϵ so that $\lfloor E(\gamma, H_\epsilon(\alpha) \rfloor_{\frac{2B}{K_1}} \geqslant E(\gamma, H_{\tilde{\epsilon}}(\alpha))$, hence

$$\sum_{\gamma([0,T]) \cap H_\epsilon(\alpha) \neq \emptyset} \lfloor d_\alpha(m_0, m_T) \rfloor_B \geqslant \sum_{\gamma([0,T]) \cap H_{\tilde{\epsilon}}(\alpha) \neq \emptyset} E(\gamma, H_{\tilde{\epsilon}}(\alpha)) = E(q, T).$$

□

Proofs of Theorem 1.1 and Theorem 1.3 (Lebesgue measure). Theorem 1.1 follows directly from Theorem 3.9 and Proposition 3.11. Moreover, by Theorem 2.3, the relative metric is a lower bound for Teichmüller distance, i.e. there exist constants K, c, depending only on the topology of S, such that

$$\|g_T\|_{rel} \leqslant KT + c.$$

This completes the proof of the first part of Theorem 1.3. □

4. Hitting measure sampling

In Section 4.1 we review some background material from the theory of random walks, and recall some previous results which show that the ratio between the word metric and the relative metric along the locations $w_n X_0$ of a sample path of the random walk remains bounded for almost all sample paths. This means that if a location $w_n X_0$ of the sample path is close to the geodesic, then this ratio is also bounded for points on the geodesic close to $w_n X_0$. However, the results of the previous section apply to all points along the geodesic which lie in the thick part of Teichmüller space, so we need to extend the bounds to these other points. In Section 4.2 we use some results of [32] to show that the distance between the locations of the sample path and the corresponding geodesic grows sublinearly, and then in Section 4.3 we use the coarse monotonicity of word length along the geodesic to show that this also bounds the ratio between word length and relative length for all points along the geodesic which lie in the thick part.

4.1. **Random walks.** Let μ be a measure on the mapping class group $G = \mathrm{Mod}(S)$. We say that μ has *finite first moment* with respect to the word metric on G if

$$\int_G \|g\|_G \, d\mu(g) < \infty$$

where $\| \cdot \|_G$ is a word metric on G with respect to a choice of finite set of generators (note that the finiteness does not depend on this choice). The *step space* is the infinite product $G^{\mathbb{N}}$ with the product measure $\mathbb{P} := \mu^{\mathbb{N}}$. Let $w_n = g_1 g_2 \ldots g_n$ be the location of the random walk after n steps. The *path space* is $G^{\mathbb{N}}$, with the pushforward of the product measure under the map

$$(g_1, g_2, g_3, \ldots) \mapsto (w_1, w_2, w_3, \ldots).$$

It will also be convenient to consider *bi-infinite* sample paths. In this case the step space is the set $G^{\mathbb{Z}}$ of bi-infinite sequences of group elements with the product measure. The location of the random walk is given by $w_0 = 1$, and $w_n = g_1 g_2 \ldots g_n$ if n, if n is positive, and $w_n = g_0^{-1} g_{-1}^{-1} \ldots g_{n-1}^{-1}$, if n is negative. The path space is $G^{\mathbb{Z}}$, as a set, but with measure coming from the pushforward of $\mathbb{P}$ under the map

$$(\ldots, g_{-1}, g_0, g_1, g_2, \ldots) \mapsto (\ldots w_{-1}, w_0, w_1, w_2, \ldots).$$

Let us fix a base point $X_0 \in \mathcal{T}$, and consider the image of the sample paths $w_n X_0$ in $\mathcal{T}$. Kaimanovich and Masur showed that almost every sample path converges to a uniquely ergodic foliation in the space $\mathcal{PMF}$ of projective measured foliations, Thurston's boundary for Teichmüller space. Recall that the *harmonic measure* ν on $\mathcal{PMF}$ is defined as the hitting measure of the random walk, i.e. for any measurable subset $A \subseteq \mathcal{PMF}$,

$$\nu(A) := \mathbb{P}(w_n \ : \lim_{n \to \infty} w_n X_0 \in A).$$

Theorem 4.1 (Kaimanovich and Masur [18]). *Let μ be a probability distribution on the mapping class group whose support generates a non-elementary subgroup. Then almost every sample path $(w_n)_{n\in\mathbb{N}}$ converges to a uniquely ergodic foliation in $\mathcal{PMF}$, and the resulting hitting measure ν is the unique non-atomic μ-stationary measure on $\mathcal{PMF}$.*

Since the mapping class group is non-amenable, the random walk makes linear progress in the word metric $\|\cdot\|_G$, or indeed in any metric quasi-isometric to the word metric.

Theorem 4.2 (Kesten [19], Day [3]). *Let μ be a probability distribution on a group, whose support generates a non-amenable subgroup. Then there exists a constant $c_1 > 0$ such that for almost all sample paths*

$$\lim_{n\to\infty} \frac{\|w_n\|_G}{n} = c_1. \tag{14}$$

Even though the relative metric is smaller than the word metric, more recent results prove that the growth rate is still linear in the number of steps.

Theorem 4.3 (Maher [21], Maher-Tiozzo [22]). *Let μ be a probability distribution on the mapping class group which has finite first moment in the word metric, and such that the semigroup generated by its support is a non-elementary subgroup. Then there is a constant $c_2 > 0$ such that for almost all sample paths*

$$\lim_{n\to\infty} \frac{\|w_n\|_{rel}}{n} = c_2.$$

Note that in [21], the result is proven under the additional condition that the support of μ is bounded in the relative metric, while such condition is not needed in [22].

From these results it follows that the quotient between the word metric and the relative metric converges to c_1/c_2 for almost every sample path, i.e.

$$\lim_{n\to\infty} \frac{\|w_n\|_G}{\|w_n\|_{rel}} = \frac{c_1}{c_2}$$

for almost all sample paths. The limit above is a limit taken along the locations $(w_n)_{n\in\mathbb{N}}$ of the random walk. In order to compare this to the previous statistic we need to relate locations of the random walk to points on a Teichmüller geodesic.

By the work of Kaimanovich and Masur [18], for almost every bi-infinite sample path $w \in G^{\mathbb{Z}}$, there are well-defined maps

$$F^{\pm} : G^{\mathbb{Z}} \to \mathcal{PMF}$$

given by

$$F^+(w) := \lim_{n\to\infty} w_n X_0$$

and

$$F^-(w) := \lim_{n\to\infty} w_{-n} X_0.$$

Furthermore, the two foliations $F^+(w)$ and $F^-(w)$ are almost surely uniquely ergodic and distinct, so there is a unique oriented Teichmüller geodesic γ_w whose forward limit point in $\mathcal{PMF}$ is $F^+(w)$ and whose backward limit point is $F^-(w)$. There is also a unique geodesic ray ρ_w starting at the basepoint X_0 whose forward limit point is F^+. We shall always parameterize ρ_w as unit speed geodesic with $\rho_w(0) = X_0$. As F^+ is uniquely ergodic, the distance between γ_w and ρ_w tends to zero, by Masur [24], and we shall parameterize γ_w such that $d_{\mathcal{T}}(\rho_w(t), \gamma_w(t)) \to 0$.

For each bi-infinite sample path we can define the function

$$D : G^{\mathbb{Z}} \to \mathbb{R}$$

given by

$$D(w) := d_{\mathcal{T}}(X_0, \gamma_w)$$

which represents the Teichmüller distance between the base point X_0 and the geodesic γ_w. This is well-defined and measurable, by Lemma 1.4.4 of [18]. In particular, this implies that for any $\epsilon > 0$ there is a constant M such that the probability that $D(w) \leqslant M$ is at least $1 - \epsilon$.

The shift map σ maps the step space to itself by incrementing the index of each step by one, i.e.

$$\sigma \colon (g_n)_{n \in \mathbb{Z}} \mapsto (g_{n+1})_{n \in \mathbb{Z}}.$$

This is a measure preserving ergodic transformation on the step space, and the induced action of σ on the path space is given by

$$\sigma \colon (w_n)_{n \in \mathbb{Z}} \mapsto (w_1^{-1} w_{n+1})_{n \in \mathbb{Z}}.$$

4.2. **Distance between geodesic and sample path.** The geodesic γ_w is determined by its endpoints $F^+(w)$ and $F^-(w)$, and the distribution of these pairs is given by harmonic measure ν and reflected harmonic measure $\check{\nu}$ respectively.

The distance from a location w_n to the corresponding geodesic γ_w is given by

$$d_{\mathcal{T}}(w_n X_0, \gamma_w) = d_{\mathcal{T}}(X_0, w_n^{-1} \gamma_w)$$

since the mapping class group acts on $\mathcal{T}$ by isometries, and by the definition of the shift map,

$$d_{\mathcal{T}}(w_n X_0, \gamma_w) = d_{\mathcal{T}}(X_0, \gamma_{\sigma^n w}).$$

As already noted in [18], if ϵ is sufficiently small, almost every geodesic with respect to harmonic measure returns to the ϵ-thick part $\mathcal{T} \setminus \mathcal{T}_\epsilon$ infinitely often.

Our goal is to show that every step of the random walk lies within sublinear distance in the word metric from some point in the thick part of the limit geodesic.

In [32], sublinear tracking is proven in the Teichmüller metric: we will adapt the argument to the word metric. The fundamental argument for sublinear tracking in [32] is the following lemma.

Lemma 4.4 (Tiozzo [32])**.** *Let $T : \Omega \to \Omega$ a measure-preserving, ergodic transformation of the probability measure space (Ω, λ), and let $f : \Omega \to \mathbb{R}^{\geqslant 0}$ any measurable, non-negative function. If the function*

$$g(\omega) := f(T\omega) - f(\omega)$$

belongs to $L^1(\Omega, \lambda)$, then for λ-almost every $\omega \in \Omega$ one has

$$\lim_{n \to \infty} \frac{f(T^n \omega)}{n} = 0.$$

We now explain how to apply the lemma above in the current setting. Given a point $X \in \mathcal{T}$, let us denote as $\mathrm{proj}(X)$ the set of lattice points at minimal distance from X:

$$\mathrm{proj}(X) := \{h \in G : d_{\mathcal{T}}(hX_0, X) \text{ is minimal}\}.$$

Such a projection may possibly vary wildly if X lies in the thin part, but it is controlled in the thick part: namely, given $\epsilon > 0$ there is a constant $K(\epsilon)$ such that

$$d_{\mathcal{T}}(X, hX_0) \leqslant K(\epsilon), \qquad \forall X \notin \mathcal{T}_\epsilon \quad \forall h \in \mathrm{proj}(X).$$

We now associate to almost every sample path w a subset $P(w)$ of the mapping class group, which we now describe. Almost every bi-infinite sample path $w \in G^{\mathbb{Z}}$ determines two uniquely ergodic foliations, $F^\pm(w)$. Let γ_w be the bi-infinite Teichmüller geodesic joining them. Now, let us define

$P(w)$ as the set of mapping class group elements $h \in G$ such that hX_0 is the closest projection from some point X in $\gamma_w \setminus \mathcal{T}_\epsilon$, i.e.

$$P(w) := \bigcup_{X \in \gamma_w \setminus \mathcal{T}_\epsilon} \mathrm{proj}(X).$$

This is illustrated in Figure 4.

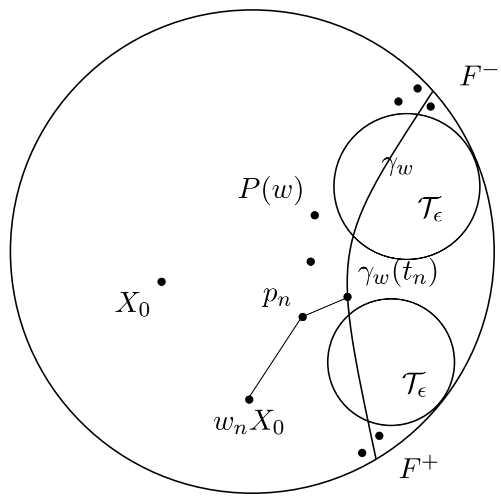

FIGURE 4. Sample path locations and basepoint orbits close to the geodesic.

The key result is the following:

Proposition 4.5. *Fix $\epsilon > 0$, sufficiently small. Then for almost every sample path $(w_n)_{n \in \mathbb{N}}$, with corresponding Teichmüller ray ρ_w, there exists a sequence of times $t_n \to \infty$ with $\rho_w(t_n) \in \rho_w \setminus \mathcal{T}_\epsilon$, such that*

$$\lim_{n \to \infty} \frac{d_G(w_n, h_n)}{n} = 0$$

for any $h_n \in \mathrm{proj}(\rho_w(t_n))$.

Proof. Let us fix $\epsilon > 0$ sufficiently small. Recall that $P(w)$ is the collection of group elements corresponding to closest lattice points to points on the geodesic γ_w which lie in the thick part of Teichmüller space. Note that, since the mapping class group acts by isometries with respect to both the Teichmüller and word metrics, then P is equivariant, in the sense that

$$P(\sigma^n w) = w_n^{-1} P(w).$$

Let us now define the function $\varphi : G^{\mathbb{Z}} \to \mathbb{R}$ on the space of bi-infinite sample paths as

$$\varphi(w) := d_G(1, P(w))$$

i.e. the minimal word-metric distance between the base point X_0 and the set of closest projections from the thick part of the geodesic γ_w. The shift map $\sigma : G^{\mathbb{Z}} \to G^{\mathbb{Z}}$ acts on the space of sequences, ergodically with respect to the product measure $\mu^{\mathbb{Z}}$. By the equivariance of P, we have for each n the equality

$$\varphi(\sigma^n w) = d_G(w_n, P(w)). \tag{15}$$

We shall now apply Lemma 4.4, setting $(\Omega, \lambda) = (G^{\mathbb{Z}}, \mu^{\mathbb{Z}})$, $T = \sigma$, and $f = \varphi$. The only condition to be checked is the L^1-condition on the function $g(\omega) = f(T\omega) - f(\omega)$, which in this case becomes

$$g(\omega) = \varphi(\sigma w) - \varphi(w) = d_G(1, P(\sigma w)) - d_G(1, P(w)).$$

Now, using (15) we have

$$|d_G(1, P(\sigma w)) - d_G(1, P(w))| = |d_G(w_1, P(w)) - d_G(1, P(w))| \leq d_G(1, w_1)$$

which has finite integral precisely by the finite first moment assumption. Thus, it follows from Lemma 4.4 that for almost all bi-infinite paths w one gets

$$\lim_{n\to\infty} \frac{d_G(w_n, P(w))}{n} = 0.$$

By definition of $P(w)$, there exists a sequence of times t_n, such that $\gamma_w(t_n)$ lies in $\gamma_w \setminus \mathcal{T}_\epsilon$, the ϵ-thick part of the geodesic γ_w, and group elements $p_n \in G$ such that $p_n \in \text{proj}(\gamma_w(t_n))$, and furthermore

$$\lim_{n\to\infty} \frac{d_G(w_n, p_n)}{n} = 0. \tag{16}$$

Now let F^+ be the terminal foliation of the geodesic γ_w, and denote as ρ_w the geodesic ray through X_0 with terminal foliation F^+. We have obtained a sequence of points lying in the intersection of the geodesic γ_w with the thick part $\mathcal{T} \setminus \mathcal{T}_\epsilon$, and we now show how to obtain a sequence of points lying in the intersection of the geodesic ρ_w with the thick part $\mathcal{T} \setminus \mathcal{T}_\epsilon$.

Recall that since γ_w and ρ_w have the same terminal foliation F^+, and F^+ is almost surely uniquely ergodic, then the distance between the positive ray ρ_w and the geodesic γ_w tends to zero, and we have chosen parameterizations such that $d_\mathcal{T}(\gamma_w(t), \rho_w(t)) \to 0$. In particular, after discarding finitely many initial values, we may assume

$$d_\mathcal{T}(\gamma_w(t_n), \rho_w(t_n)) \leqslant \frac{\log 2}{2},$$

for all n. Now for each n sufficiently large consider the sequence take $\rho_w(t_n)$. Then:

(1) By Wolpert's lemma, $\rho_w(t_n)$ lies in the $\frac{\epsilon}{2}$-thick part;
(2) if $h_n \in \text{proj}(\rho_w(t_n))$, then $d_\mathcal{T}(\rho_w(t_n), h_n X_0) \leqslant K(\epsilon/2)$ so

$$\begin{aligned} d_\mathcal{T}(h_n X_0, p_n X_0) &\leqslant d_\mathcal{T}(h_n X_0, \rho_w(t_n)) + d_\mathcal{T}(\rho_w(t_n), \gamma_w(t_n)) + d_\mathcal{T}(\gamma_w(t_n), p_n X_0) \\ &\leqslant 1 + 2K(\epsilon/2), \end{aligned}$$

hence $d_G(h_n, p_n) \leqslant K'$, so by equation (16) we have also

$$\lim_{n} \frac{d_G(w_n, h_n)}{n} \to 0.$$

This completes the proof of Proposition 4.5. □

4.3. **Intermediate times.** So far, we have shown that every step of the sample path is close enough to some point on the thick part of the geodesic, hence the closest projection to the lattice will behave like the sample path. However, we still need to deal with the case in which there are points in the thick part of the Teichmüller geodesic which are not close to the sample path.

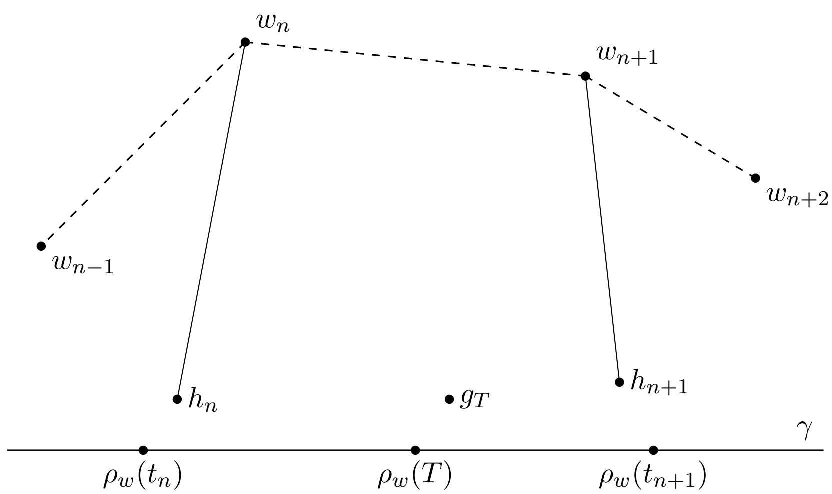

FIGURE 5. Intermediate times.

Proof of Theorem 1.2 and Theorem 1.3 (harmonic measure). Given a sample path w, let ρ_w be the geodesic ray joining the base point X_0 to the limit foliation $F^+(w)$, and let t_n be the sequence of times given by Proposition 4.5. Let now $T > 0$ be a time for which the geodesic $\rho_w(T)$ lies in the thick part, and let $g_T X_0$ be a projection of $\rho_w(T)$ to the Teichmüller lattice. Since $t_n \to \infty$, there exists an index $n = n(T)$ such that $t_n \leqslant T \leqslant t_{n+1}$. By Proposition 2.14, there exist constants $C_1 > 0$, C_2 such that

$$d_G(h_n, g_T) \leqslant C_1 d_G(h_n, h_{n+1}) + C_2.$$

Moreover, by Proposition 4.5 and triangle inequality,

$$\lim_{n\to\infty} \frac{d_G(h_n, h_{n+1})}{n} \leqslant \lim_{n\to\infty} \frac{d_G(h_n, w_n) + d_G(w_n, w_{n+1}) + d_G(w_{n+1}, h_{n+1})}{n} = 0$$

(where we used the finite first moment condition to ensure $d_G(w_n, w_{n+1})/n \to 0$). Thus, we also have

$$\lim_{n\to\infty} \frac{d_G(h_n, g_T)}{n} = 0$$

and again by Proposition 4.5

$$\lim_{n\to\infty} \frac{d_G(w_n, g_T)}{n} \leqslant \lim_{n\to\infty} \frac{d_G(w_n, h_n) + d_G(h_n, g_T)}{n} = 0.$$

Similarly, since the relative metric is bounded above by the word metric,

$$\lim_{n\to\infty} \frac{d_{rel}(w_n, g_T)}{n} = 0.$$

Finally, by computing the ratio between the word and relative metric,

$$\lim_{\substack{T\to\infty \\ \rho_w(T)\notin\mathcal{T}_\epsilon}} \frac{\|g_T\|_G}{\|g_T\|_{rel}} = \lim_{\substack{T\to\infty \\ \rho_w(T)\notin\mathcal{T}_\epsilon}} \frac{\frac{d_G(1,g_T)}{n(T)}}{\frac{d_{rel}(1,g_T)}{n(T)}} = \lim_{n\to\infty} \frac{\frac{d_G(1,w_n)}{n}}{\frac{d_{rel}(1,w_n)}{n}} = \frac{c_1}{c_2} > 0.$$

This completes the proof of Theorem 1.3. The proof of Theorem 1.2 is exactly the same, replacing d_{rel} with $d_{\mathcal{T}}$ and noting that, since the Teichmüller metric is a coarse upper bound for the relative metric, Theorem 4.3 implies also positive drift in the Teichmüller metric. □

References

[1] Jayadev Athreya, Alexander Bufetov, Alex Eskin, and Maryam Mirzakhani, *Lattice point asymptotics and volume growth on Teichmüller space*, Duke Math. J. **161** (2012), no. 6, 1055–1111.

[2] Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu, *Harmonic measures versus quasiconformal measures for hyperbolic groups*, Ann. Sci. Éc. Norm. Supér. (4) **44** (2011), no. 4, 683–721.

[3] Mahlon Marsh Day, *Convolutions, means, and spectra*, Illinois J. Math. **8** (1964), 100–111.

[4] Bertrand Deroin, Victor Kleptsyn, and Andrés Navas, *On the question of ergodicity for minimal group actions on the circle*, Mosc. Math. J. **9** (2009), no. 2, 263-303.

[5] Harold Diamond and Jeffrey Vaaler, *Estimates for partial sums of continued fraction partial quotients*, Pacific J. Math. **122** (1986), no. 1, 73-82.

[6] Spencer Dowdall, Moon Duchin, and Howard A. Masur, *Statistical hyperbolicity in Teichmüller space*, Geom. Funct. Anal **24** (2014), 748-795.

[7] Alex Eskin and Howard A. Masur, *Asymptotic formulas on flat surfaces*, Ergodic Theory Dynam. Systems **21** (2001), no. 2, 443–478.

[8] Harry Furstenberg, *Random walks and discrete subgroups of Lie groups*, Dekker, New York, 1971.

[9] Harry Furstenberg, *Boundary theory and stochastic processes on homogeneous spaces*, Amer. Math. Soc., Providence, R.I., 1973.

[10] Vaibhav Gadre, *Harmonic measures for distributions with finite support on the mapping class group are singular*, Duke Math. J. **163** (2014), no. 2, 309-368.

[11] Vaibhav Gadre, *Partial sums of excursions along random geodesics and volume asymptotics for thin parts of moduli spaces of quadratic differentials* (2014), available at `arXiv:1408.5812`.

[12] Vaibhav Gadre, Joseph Maher, and Giulio Tiozzo, *Word length statistics and Lyapunov exponents for Fuchsian groups with cusps* (2014).

[13] Yves Guivarc'h and Yves Le Jan, *Sur l'enroulement du flot géodésique*, C. R. Acad. Sci. Paris Sér. I Math. **311** (1990), no. 10, 645–648.

[14] Yves Guivarc'h and Yves Le Jan, *Asymptotic winding of the geodesic flow on modular surfaces and continued fractions*, Ann. Sci. École Norm. Sup. (4) **26** (1993), no. 1, 23–50.

[15] John Hubbard and Howard Masur, *Quadratic differentials and foliations*, Acta Math. **142** (1979), no. 3-4, 221–274.

[16] Anatole Katok, *Four applications of conformal equivalence to geometry and dynamics*, Ergodic Theory Dynam. Systems **8*** (1988), 139–152.

[17] Vadim A. Kaimanovich and Vincent Le Prince, *Matrix random products with singular harmonic measure*, Geom. Dedicata **150** (2011), 257–279.

[18] Vadim A. Kaimanovich and Howard A. Masur, *The Poisson boundary of the mapping class group*, Invent. Math. **125** (1996), no. 2, 221–264.

[19] Harry Kesten, *Symmetric random walks on groups*, Trans. Amer. Math. Soc. **92** (1959), 336–354.

[20] François Ledrappier, *Applications of dynamics to compact manifolds of negative curvature*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), 1995, pp. 1195–1202.

[21] Joseph Maher, *Linear progress in the complex of curves*, Trans. Amer. Math. Soc. **362** (2010), no. 6, 2963–2991.

[22] Joseph Maher and Giulio Tiozzo, *Random walks on weakly hyperbolic groups* (2014), available at `arXiv:1410.4173`.

[23] Bernard Maskit, *Comparison of hyperbolic and extremal lengths*, Ann. Acad. Sci. Fenn. Ser. A I Math. **10** (1985), 381–386.

[24] Howard A. Masur, *Uniquely ergodic quadratic differentials*, Comment. Math. Helv. **55** (1980), no. 2, 255–266.

[25] Howard A. Masur, *Interval exchange transformations and measured foliations*, Ann. of Math. (2) **115** (1982), no. 1, 169–200.

[26] Howard A. Masur, *Logarithmic law for geodesics in moduli space*, (Göttingen, 1991/Seattle, WA, 1991), Contemp. Math., vol. 150, Amer. Math. Soc., Providence, RI, 1993, pp. 229–245.

[27] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149.

[28] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974.

[29] Kasra Rafi, *A combinatorial model for the Teichmüller metric*, Geom. Funct. Anal. **17** (2007), no. 3, 936–959.

[30] Kasra Rafi, *Hyperbolicity in Teichmüller space* (2010), available at `arXiv:1011.6004`. to appear in Geom. Topol.

[31] Dennis Sullivan, *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics*, Acta Math. **3-4** (1982), 215–237.

[32] Giulio Tiozzo, *Sublinear deviation between geodesics and sample paths*, Duke Math. J. **164** (2015), no. 3, 511–539.

[33] William A. Veech, *The Teichmüller geodesic flow*, Ann. of Math. (2) **124** (1986), no. 3, 441–530.
[34] Yaroslav Vorobets, *Periodic geodesics on generic translation surfaces*, Contemp. Math., vol. 385, Amer. Math. Soc., Providence, RI, 2005.

Vaibhav Gadre
Warwick Mathematics Institute
Zeeman Building, University of Warwick, Coventry CV4 7AL, UK
v.gadre@warwick.ac.uk;gadre.vaibhav@gmail.com

Joseph Maher
CUNY College of Staten Island and CUNY Graduate Center
2800 Victory Boulevard, Staten Island NY 10314 USA
joseph.maher@csi.cuny.edu

Giulio Tiozzo
Yale University
10 Hillhouse Avenue, New Haven CT 06511 USA
giulio.tiozzo@yale.edu