STATISTICAL HYPERBOLICITY IN TEICHMÜLLER SPACE

SPENCER DOWDALL, MOON DUCHIN, AND HOWARD MASUR

ABSTRACT. In this paper we explore the idea that Teichmüller space with the Teichmüller metric is hyperbolic "on average." We consider several different measures on Teichmüller space and show that with respect to each one, the average distance between points in a ball of radius r is asymptotic to $2r$, which is as large as possible.

1. INTRODUCTION

Let S be a closed surface of genus $g > 1$. In this paper we continue the study of Teichmüller space $\mathcal{T}(S)$, which is the parameter space for several types of geometric structures on S. It is known that Teichmüller space equipped with the Teichmüller metric $d_{\mathcal{T}}$ is a complete metric space homeomorphic to $\mathbb{R}^{6g-6}$. It is not δ–hyperbolic [15], and several kinds of obstructions to hyperbolicity are known: for instance, pairs of geodesic rays through the same point may fellow-travel arbitrarily far apart [9], and there are large "thin parts" of the space which, up to bounded additive error, are isometric to product spaces equipped with sup metrics [16]. These exceptions to the negative-curvature phenomena seem to come from rare occurrences, so one might expect properties that are characteristic of hyperbolicity to hold on average.

One way to make this precise is to consider the average distance between points on metric spheres in a metric space (X, d). Writing $\mathcal{S}_r(x)$ for the sphere of radius r based at x, we define a geometric statistic for the large spheres as follows. Given a family of probability measures μ_r on the spheres $\mathcal{S}_r(x)$, let $E(X) = E(X, x, d, \{\mu_r\})$ be the average distance between points on large spheres:

$$E(X) := \lim_{r\to\infty} \frac{1}{r} \int_{\mathcal{S}_r(x)\times\mathcal{S}_r(x)} d(y, z) \quad d\mu_r(y)d\mu_r(z),$$

if the limit exists. It is shown in [5] that non-elementary hyperbolic groups all have $E(G, S) = 2$ for any finite generating set S; this is also the case in the hyperbolic space $\mathbb{H}^n$ of any dimension endowed with the natural measure on spheres. By contrast, it is shown that $E(\mathbb{R}^n) < 2$ (increasing over the range $[4/\pi, \sqrt{2})$ as the dimension goes from 2 to infinity), and that $E(\mathbb{Z}^n, S) < 2$ for all n and S, with nontrivial dependence on S. (See [5] for more examples.)

In the case of Teichmüller space, the sphere $\mathcal{S}_r(x)$ can be identified with the unit sphere $\mathcal{Q}^1(x)$ in the vector space of quadratic differentials on x via the Teichmüller map. The latter has various natural measures, and corresponding measures on $\mathcal{S}_r(x)$ will be called *visual measures*.

Date: September 12, 2011.

The second author was partially supported by NSF DMS-0906086. The third author was partially supported by NSF DMS-0905907.

On the other hand, as a metric space $\mathcal{T}(S)$ carries a natural $6g-6$ dimensional *Hausdorff measure* η. Using this measure, we also consider the average distance between points in the ball $\mathcal{B}_r(x)$ of radius r centered at x. Our main theorem is:

Theorem 1. *For every point $x \in \mathcal{T}(S)$,*

$$\lim_{r\to\infty} \frac{1}{r}\frac{1}{\eta(\mathcal{B}_r(x))^2}\int_{\mathcal{B}_r(x)\times\mathcal{B}_r(x)} d_{\mathcal{T}}(y,z) \quad d\eta(y)d\eta(z) = 2.$$

In other words, the average distance between points in $\mathcal{B}_r(x)$ is asymptotic to $2r$, which, in light of the triangle inequality, is the maximum possible distance. In addition to Hausdorff measure, we will also prove this theorem for averages with respect to various other measures; the complete statement is given in §5 at the end of the paper.

Our approach to Theorem 1 will be via certain visual measures on spheres; these will be discussed in §3. In this regard we have the following result.

Theorem 2. *$E(\mathcal{T}(S)) = 2$ with respect to Teichmüller distance and various visual measures.*

We will actually prove a slightly stronger statement that allows for the pair of points to lie on spheres of different radii that go to infinity and we conclude that the average distance is asymptotically the sum of those radii.

Before proceeding to the distance estimates needed for these results, in §3 we will establish comparisons among a number of a priori different measures that are natural from various points of view, such as the metric structure (Hausdorff measure), the Finsler structure (Busemann measure and Holmes-Thompson measure), the quadratic differentials (holonomy, or Masur-Veech, measure), and the symplectic structure. We find that all of these are absolutely continuous with respect to each other and cite a theorem of [2] to conclude that they are absolutely continuous with respect to certain visual measures as well. These comparisons may be of independent interest.

We sketch here the main ideas in the proof of Theorem 2. The first step is to show (Proposition 23) that on average, pairs of geodesics separate from each other in the Teichmüller metric after a threshold time. Then one would hope that, as in a hyperbolic space, the geodesic joining their endpoints would follow the first geodesic back to approximately where they separate before following the other so that its length is roughly the sum of the lengths of the two geodesics, as on the left in Figure 1. The Minsky product regions theorem [16] says that this in fact may not happen. If the pair of geodesics separate because they enter thin parts that are disjoint, then the geodesic joining their endpoints travels through those disjoint thin regions simultaneously and its length is then smaller than the sum, as on the right in Figure 1. Our goal is to show that this phenomenon does not happen on average. The mechanism for showing this is the coarsely contracting map from $\mathcal{T}(S)$ to the curve complex. We show (Theorem 35) that, generically, pairs of Teichmüller geodesics separate in the curve complex after a bounded time. We then show (Theorem 36) that for generic pairs, a geodesic connecting the two rays must pass through the region where the rays separated, which gives the needed distance estimate.

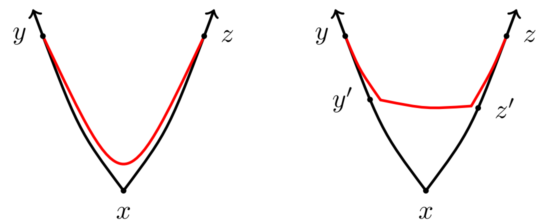

FIGURE 1. We will show that the geodesic between points on generic rays "dips" back near the basepoint. This requires an analysis of the time spent in thick and thin parts: if $[x, y']$ and $[x, z']$ lie in disjoint thin parts, then Minsky's product regions theorem shows that the connecting geodesic can take a "shortcut." We show that this effect is rare.

1.1. **Acknowledgments.** We would like to thank Alex Eskin, Benson Farb, Curtis McMullen, and especially Kasra Rafi for numerous helpful comments and explanations.

2. BACKGROUND MATERIAL

2.1. **Teichmüller space and quadratic differentials.** Recall that Teichmüller space $\mathcal{T}(S)$ is the space of marked Riemann surfaces X that are homeomorphic to the topological surface S. More precisely, it consists of pairs (X, f), where $f : S \to X$ is a homeomorphism, up to the equivalence relation that $(X_1, f_1) \sim (X_2, f_2)$ when there exists a conformal map $F : X_1 \to X_2$ such that $F \circ f_1$ is isotopic to f_2. Alternately, we may define $\mathcal{T}(S)$ as the space of marked hyperbolic surfaces (ρ, f); namely, maps $f : S \to \rho$ with $(\rho_1, f_1) \sim (\rho_2, f_2)$ when there exists an isometry $F : \rho_1 \to \rho_2$ s.t. $F \circ f_1$ is isotopic to f_2.

Using the first definition of $\mathcal{T}(S)$, the Teichmüller distance is given by

$$d_{\mathcal{T}}((X_1, f_1), (X_2, f_2)) := \inf_{F \sim f_2 \circ f_1^{-1}} \frac{1}{2} \log K(F),$$

where the minimum is taken over all quasiconformal maps F and $K(F)$ is the maximal dilatation of F. The space $\mathcal{T}(S)$ is homeomorphic to the ball $\mathbb{R}^{6g-6}$, and from now on we will use $h = 6g - 6$ to designate this dimension. In this paper, we will denote a point of $\mathcal{T}(S)$ by x, regarding it either as a Riemann surface or a hyperbolic surface, and suppressing the marking f.

For $x, y \in \mathcal{T}(S)$, the Teichmüller geodesic segment joining x to y will usually be denoted $[x, y]$. We will also use $\gamma(t)$ to denote a geodesic ray or segment when the time parameter is important.

A *quadratic differential* on a Riemann surface X is an integrable holomorphic 2-tensor $q = \phi(z)dz^2$ on X. The space of all quadratic differentials on all Riemann surfaces homeomorphic to S is denoted $\mathcal{Q}(S)$. A point of $\mathcal{Q}(S)$ will be denoted q, with the underlying complex structure implicit in the notation. The real dimension of $\mathcal{Q}(S)$ is $12g - 12 = 2h$. Reading off the Riemann surface, we obtain a projection to the Teichmüller space $\pi : \mathcal{Q}(S) \to \mathcal{T}(S)$. Under this projection, $\mathcal{Q}(S)$ forms vector bundle over $\mathcal{T}(S)$ which is canonically identified with the cotangent bundle of $\mathcal{T}(S)$. Each fiber $\mathcal{Q}(X)$ is equipped with a norm given by the total area of q; namely $\|q\| = \int_X |\phi(z)dz^2|$. Recall that $d_{\mathcal{T}}$ is not a Riemannian metric on $\mathcal{T}(S)$, but rather a Finsler metric; it comes from dualizing the norm on $\mathcal{Q}$ to give a norm on each tangent space of $\mathcal{T}(S)$ that is not induced by any inner product.

It is the famous theorem of Teichmüller that the infimum in the definition of $d_{\mathcal{T}}$ is realized uniquely by a *Teichmüller map* from X_1 to X_2. A Teichmüller map is determined by an initial quadratic differential $q = \phi(z)dz^2$ on X_1 and the number K. The Teichmüller map expands along the horizontal trajectories of q by a factor of $K^{1/2}$ and contracts along the vertical trajectories by the same factor to obtain a terminal quadratic differential q' on the image surface X_2. If we fix q and let $K = e^{2t}$ vary over $t \in [0, \infty)$ we get a Teichmüller geodesic ray.

Recall that the *mapping class group* of S, defined by

$$\mathrm{Mod}(S) := \mathrm{Diff}^+(S)/\mathrm{Diff}_0(S),$$

is the discrete group of orientation-preserving diffeomorphisms of S, up to isotopy. This group acts isometrically on $\mathcal{T}(S)$ by changing the marking: $\phi \cdot (X, f) = (X, f \circ \phi^{-1})$. In fact, by a result of Royden [20], $\mathrm{Mod}(S)$ is the full group of (orientation-preserving) isometries of $(T(S), d_{\mathcal{T}})$.

2.2. **Curve graph.** When we speak of a *curve* on S, this will mean an isotopy class of essential simple closed curves. Given $x \in \mathcal{T}(S)$, the *length* $l_x(\alpha)$ of a curve α is the infimal length achieved by the isotopy class in the hyperbolic metric x.

We recall the definition of the *curve complex* (or curve graph) $\mathcal{C}(S)$ of S. The vertices of $\mathcal{C}(S)$ are curves on S. Two vertices are joined by an edge if the corresponding curves can be realized disjointly. Assigning edges to have length 1 we have a metric graph. Properly speaking, $\mathcal{C}(S)$ is the flag complex associated to this curve graph, but since we are working coarsely, we can identify $\mathcal{C}(S)$ with the graph.

It is known that the curve graph is hyperbolic [14]. That is, there exists a constant $\delta > 0$ such that every geodesic triangle in $\mathcal{C}(S)$ is δ–thin: each side of the triangle is contained in the union of the δ–neighborhoods of the other two sides. It follows that every geodesic quadrilateral in $\mathcal{C}(S)$ is 2δ–thin (each side is within 2δ of the union of the other three sides). Furthermore, in any δ–hyperbolic metric space and for any quasi-isometry constants (K, C), there exists a constant τ, depending only on δ, K, C, such that any two (K, C)–quasi-geodesic segments with the same endpoints remain within τ of each other. Since actual geodesics are $(1, 0)$–quasi-geodesics, this implies that every (K, C)–quasi-geodesic quadrilateral is $2(\delta + \tau)$–thin.

2.3. **Thick parts and subsurface projections.** For any given ϵ_0, we say a curve is ϵ_0–short if its hyperbolic length is less than ϵ_0. Then define $\mathcal{T}_{\epsilon_0}$, the ϵ_0*–thick part* of Teichmüller space, to be the subset of $x \in \mathcal{T}(S)$ on which no curve is ϵ_0–short.

For each $x \in \mathcal{T}(S)$ there is associated a *Bers marking* μ_x. To construct μ_x, greedily choose a shortest *pants decomposition* of the surface (a collection of $3g - 3$ disjoint simple geodesics). Then for each pants curve β, choose a shortest geodesic crossing β minimally (either once or twice depending on the topology) that is disjoint from all other pants curves. The total collection of $6g - 6$ curves is called a Bers marking and is defined up to finitely many choices.

Throughout, a *proper subsurface of* S will mean a compact, properly embedded subsurface $V \subset S$ which is not equal to S and for which the induced map on fundamental groups is injective. Subsurfaces which are isotopic to each other will not be considered distinct. The proper subsurfaces of S fall into two categories, annuli and non-annuli, which behave somewhat differently. Nevertheless, we will strive to develop intuitive notation under which these two possibilities may be dealt with on equal footing.

Every proper subsurface V has a nonempty boundary ∂V consisting of a disjoint union of curves on S. We say that two subsurfaces V and W *transversely intersect*, denoted $V \pitchfork W$, if they are neither (isotopically) disjoint nor nested. In this case, ∂V necessarily intersects W, and ∂W intersects V.

Consider a non-annular subsurface V, possibly equal to S. The *subsurface projection* $\pi_V(\beta)$ of a simple closed curve $\beta \subset S$ to V is defined as follows: Realize β and ∂V as geodesics (in any hyperbolic metric on S). If $\beta \subset V$, then $\pi_V(\beta)$ is defined to be β. If β is disjoint from V, then $\pi_V(\beta)$ is undefined. Otherwise, $\beta \cap V$ is a disjoint union of finitely many homotopy classes of arcs with endpoints on ∂V, and we obtain $\pi_V(\beta)$ by choosing any arc and performing a surgery along ∂V to create a simple closed curve contained in V. The subsurface projection of a point $x \in \mathcal{T}(S)$ is then defined to be the collection

$$\pi_V(x) := \{\pi_V(\beta)\}_{\beta \in \mu_x}$$

of curves obtained by varying β in the Bers marking at x. This is a non-empty subset of the curve complex $\mathcal{C}(V)$ with uniformly bounded diameter.

Definition 3 (Non-annular projection distance)**.** For a non-annular subsurface $V \subseteq S$, the *projection distance in* V of a pair of points $x, y \in \mathcal{T}(S)$ is defined to be

$$d_V(x,y) := \operatorname{diam}_{\mathcal{C}(V)}(\pi_V(x) \cup \pi_V(y)).$$

In particular, $d_S(x,y)$ denotes the curve complex distance. When convenient, we will also denote this distance by $d_{\mathcal{C}(V)} := d_V$.

For an annular subsurface $A \subset S$ with *core curve* $\alpha = \partial A$, there are two kinds of projection distances: one that measures twisting about α and is analogous to the definition above, and a second which also incorporates the length of α. Any simple closed curve β that crosses α may be realized by a geodesic and then lifted to a geodesic $\tilde{\beta}$ in the annular cover $\tilde{A}$, that is, the quotient of $\mathbb{H}^2$ by the deck transformation corresponding to α, with the Gromov compactification. For a pair β, γ of such curves, we may then consider the intersection number $\mathrm{i}(\tilde{\beta}, \tilde{\gamma})$ in $\tilde{A}$. The *twisting distance* in A of a pair of points $x, y \in \mathcal{T}(S)$ is then defined as

$$d_{\mathcal{C}(A)}(x,y) := \sup_{\beta \in \mu_x, \gamma \in \mu_y} \mathrm{i}_{\tilde{A}}(\tilde{\beta}, \tilde{\gamma}).$$

We additionally define a hyperbolic projection distance as follows.

Definition 4 (Annular projection distance)**.** For an annular subsurface $A \subset S$ with core curve $\alpha = \partial A$, denote by $\mathbb{H}_\alpha$ a copy of the standard horoball $\{\mathrm{Im}(z) \geq 1\} \subset \mathbb{H}^2$. Given $x, y \in \mathcal{T}(S)$, we consider the points $(0, 1/l_x(\alpha))$ and $(d_{\mathcal{C}(A)}(x,y), 1/l_y(\alpha)) \in \mathbb{H}^2$ and denote their closest point projections to the horoball $\mathbb{H}_\alpha$ by

$$\pi_\alpha(x) = \left(0, \max\left\{1, \frac{1}{l_x(\alpha)}\right\}\right), \quad \pi_\alpha(y) = \left(d_{\mathcal{C}(A)}(x,y), \max\left\{1, \frac{1}{l_y(\alpha)}\right\}\right).$$

The *projection distance* in A (or *hyperbolic distance* $d_{\mathbb{H}_\alpha}$) between x and y is then defined to be

$$d_A(x,y) := d_{\mathbb{H}^2}\left(\pi_\alpha(x), \pi_\alpha(y)\right).$$

2.4. **Distance formula.** For functions f, g and constants $K \geq 1$, $C \geq 0$, we will use the notation $f(x) \overset{K,C}{\asymp} g(x)$ if the inequalities $\frac{1}{K}g(x) - C \leq f(x) \leq K \cdot g(x) + C$ hold for all x. As usual we denote $f = O(g)$ if the second inequality holds, and $f = o(g)$ if $f(x)/g(x) \to 0$ as $x \to \infty$.

The following distance formula due to Rafi relates the Teichmüller distance between two points x and y to the combinatorics of the corresponding Bers markings μ_x and μ_y.

Theorem 5 (Distance formula, Rafi [17])**.** *Fix a small $\epsilon_0 > 0$. For any sufficiently large threshold M_0, there exist quasi-isometry constants $K \geq 1$ and $C \geq 0$ depending only on M_0 and the topology of S such that, for all $x, y \in \mathcal{T}(S)$ we have*

$$\begin{aligned} d_{\mathcal{T}}(x,y) \quad &\overset{K,C}{\asymp} d_S(x,y) + \sum_V \left[d_V(x,y)\right]_{M_0} + \max_{\alpha \in \Gamma_{xy}} d_{\mathbb{H}_\alpha}(x,y) \\ &+ \sum_{A\,:\,\partial A \notin \Gamma_{xy}} \log_+ \left[d_{\mathcal{C}(A)}(x,y)\right]_{M_0} + \max_{\alpha \in \Gamma_x} \log_+ \left(\frac{1}{l_x(\alpha)}\right) + \max_{\alpha \in \Gamma_y} \log_+ \left(\frac{1}{l_y(\alpha)}\right), \end{aligned}$$

where Γ_{xy} is the set of ϵ_0–short curves in both x and y, Γ_x is the set of curves that are ϵ_0–short in x but not in y, and Γ_y is defined similarly. Here and throughout, $\log_+$ is a modified logarithm so that $\log_+ a = 0$ for $a \in [0,1]$, and $[\cdot]_{M_0}$ is a threshold function for which $[N]_{M_0} := N$ when $N \geq M_0$ and $[N]_{M_0} := 0$ otherwise.

Remark 6. The definition of $d_{\mathbb{H}_\alpha} = d_A$ given above is technically different than that used by Rafi in [17]; however, the two definitions agree up to bounded additive error.

In order to simplify notation and streamline our arguments, it will be beneficial to repackage this distance formula in a way that treats annular and non-annular subsurfaces on equal footing, which has the effect of simply expressing the Teichmüller distance as the sum of large subsurface projections. For simplicity and without loss of generality, below we suppose that ϵ_0 is fixed small enough that $\log_+(1/\epsilon_0) \geq 100$, say. We begin with a straightforward reformulation.

Lemma 7. *Given any sufficiently large threshold M_0, there exist $K \geq 1$ and $C \geq 0$ such that for all $x, y \in \mathcal{T}(S)$ we have:*

$$\begin{aligned} d_{\mathcal{T}}(x,y) &\overset{K,C}{\asymp} d_S(x,y) + \sum_V \left[d_V(x,y)\right]_{M_0} + \sum_{A\,:\,\partial A \in \Gamma_{xy}} \left[d_A(x,y)\right]_{M_0} + \\ &\sum_{A\,:\,\partial A \notin \Gamma_{xy}} \left[\max\left\{\log_+\left(d_{\mathcal{C}(A)}(x,y)\right), \log_+\left(\frac{1}{l_x(\partial A)}\right), \log_+\left(\frac{1}{l_y(\partial A)}\right)\right\}\right]_{\log M_0} \end{aligned}$$

Proof. Since Γ_{xy}, Γ_x and Γ_y each contain at most $3g-3$ curves, each max over these sets is within bounded multiplicative error of the corresponding sum, and applying a threshold only creates bounded additive error, so the first three terms of the lemma are established. By the definition of Γ_x we have

$$\sum_{\alpha \in \Gamma_x} \log_+ \left(\frac{1}{l_x(\alpha)}\right) = \sum_{\alpha \notin \Gamma_{xy}} \log_+ \left[\frac{1}{l_x(\alpha)}\right]_{1/\epsilon_0}.$$

Since this is a sum with at most $3g-3$ nonzero terms, we can increase the threshold to any number $M_0 \geq 1/\epsilon_0$ with bounded additive error. Finally, for functions f, g, h, we have

$$\log_+[f]_{M_0} + \log_+[g]_{M_0} + \log_+[h]_{M_0} \overset{3,0}{\asymp} \left[\max\{\log_+ f, \log_+ g, \log_+ h\}\right]_{\log M_0}. \quad \square$$

We now show that each term in the last summand is bilipschitz equivalent to the corresponding hyperbolic distance $d_A(x,y)$.

Lemma 8. *Consider an annular subsurface $A \subset S$ with core curve $\partial A = \alpha$. For each pair of points $x, y \in \mathcal{T}(S)$, set*

$$H_A(x,y) := \max\left\{\log_+\left(d_{\mathcal{C}(A)}(x,y)\right), \log_+\left(\frac{1}{l_x(\alpha)}\right), \log_+\left(\frac{1}{l_y(\alpha)}\right)\right\}.$$

If $x, y \in \mathcal{T}(S)$ are such that $\alpha \notin \Gamma_{xy}$ and either $d_A(x,y)$ or $H_A(x,y)$ is greater than $36\log_+(1/\epsilon_0)$, then $6^{-1}d_A(x,y) \le H_A(x,y) \le 6d_A(x,y)$.

Proof. Choose points $x, y \in \mathcal{T}(S)$ that satisfy the hypotheses. To fix notation, set $\pi'_\alpha(x) = (0,1)$ and $\pi'_\alpha(y) = (d_{\mathcal{C}(A)}(x,y), 1)$. These are the closest-point projections of $\pi_\alpha(x)$ and $\pi_\alpha(y)$ to the horocycle bounding $\mathbb{H}_\alpha$, and their distances from these points are exactly given by $\log_+(1/l_x(\alpha))$ and $\log_+(1/l_y(\alpha))$. Let

$$B = d_{\mathbb{H}^2}(\pi'_\alpha(x), \pi'_\alpha(y)) = \operatorname{arccosh}\left(1 + \frac{d_{\mathcal{C}(A)}(x,y)^2}{2}\right)$$

denote the hyperbolic distance between these projections. Using this formula, one may easily check that the inequalities

$$\log_+ d_{\mathcal{C}(A)}(x,y) \le B \le 4\log_+ d_{\mathcal{C}(A)}(x,y) \tag{1}$$

hold provided that either $B \ge 3$ or $d_{\mathcal{C}(A)}(x,y) \ge 3$.

Applying the triangle inequality with the points $\pi'_\alpha(x)$ and $\pi'_\alpha(y)$ implies that

$$d_A(x,y) \le \log_+\left(\frac{1}{l_x(\alpha)}\right) + B + \log_+\left(\frac{1}{l_y(\alpha)}\right). \tag{2}$$

Then (1), (2), and the definition of H_A imply that $d_A(x,y) \le 6H_A(x,y)$ in the case that $B \ge 3$. If $B < 3$, we claim that the hypotheses of the Lemma ensure that B cannot be the largest term on the right-hand side and therefore that $d_A(x,y) \le 3L \le 3H_A(x,y)$, where L denotes the larger of the other two terms. Indeed, if B were the largest term and $B < 3$, then (2) would imply $d_A(x,y) < 9$, and (1) would necessitate $\log_+ d_{\mathcal{C}(A)}(x,y) < 3$ so that $H_A(x,y) < 3$. But then both d_A and H_A are less than 9, contradicting the hypothesis.

By the above, the assumption $d_A(x,y) \ge 36\log_+(1/\epsilon_0)$ implies that $H_A(x,y) \ge 6\log_+(1/\epsilon_0)$; therefore all cases will be covered by proving that this in turn implies $H_A(x,y) \le 6d_A(x,y)$. Without loss of generality, we may assume that $l_x(\alpha) \le l_y(\alpha)$; since $\alpha \notin \Gamma_{xy}$ this guarantees $l_y(\alpha) \ge \epsilon_0$. First suppose that $\log_+ d_{\mathcal{C}(A)}(x,y) \ge 3\log_+(1/l_x(\alpha))$, in which case we have $\log_+ d_{\mathcal{C}(A)}(x,y) = H_A(x,y) \ge 6\log_+(1/\epsilon_0)$. In particular we certainly have $d_{\mathcal{C}(A)}(x,y) \ge 3$; thus (1) and the triangle inequality give

$$\log_+ d_{\mathcal{C}(A)}(x,y) \le B \le \log_+\left(\frac{1}{l_x(\alpha)}\right) + d_A(x,y) + \log_+\left(\frac{1}{l_y(\alpha)}\right).$$

Therefore $H_A(x,y) = \log_+ d_{\mathcal{C}(A)}(x,y) \le 3d_A(x,y)$ in this case. The remaining possibility $\log_+ d_{\mathcal{C}(A)}(x,y) \le 3\log_+(1/l_x(\alpha))$ necessitates $3\log_+(1/l_x(\alpha)) \ge H_A(x,y)$.

Recall that $\pi'_\alpha(x)$ is the *closest* point projection of $\pi_\alpha(x)$ to the horocycle bounding $\mathbb{H}_\alpha$; since $\pi'_\alpha(y)$ is also on this horocycle we have

$$\log_+\left(\frac{1}{l_x(\alpha)}\right) \le d_{\mathbb{H}^2}(\pi_\alpha(x), \pi'_\alpha(y)) \le d_A(x,y) + \log_+\left(\frac{1}{l_y(\alpha)}\right).$$

The assumptions $3\log_+(1/l_x(\alpha)) \ge H_A(x,y) \ge 6\log_+(1/\epsilon_0)$ and $l_y(\alpha) \ge \epsilon_0$ now ensure that $H_A(x,y) \le 6 d_A(x,y)$. □

Corollary 9. *Let $H_A(x,y)$ be defined as in Lemma 8. Then for any threshold $M_0 \ge 36\log_+(1/\epsilon_0)$ and any $x, y \in \mathcal{T}(S)$ we have*

$$\sum_{\partial A \notin \Gamma_{xy}} 6^{-1}\left[d_A(x,y)\right]_{6M_0} \le \sum_{\partial A \notin \Gamma_{xy}} \left[H_A(x,y)\right]_{M_0} \le \sum_{\partial A \notin \Gamma_{xy}} 6\left[d_A(x,y)\right]_{M_0/6}$$

With these estimates, the distance formula now takes a particularly simple form.

Proposition 10 (Repackaged distance formula)**.** *For any sufficiently large threshold M_0, there exist quasi-isometry constants $K \ge 1$ and $C \ge 0$ depending only on M_0 and the topology of S such that, for all $x, y \in \mathcal{T}(S)$ we have:*

$$d_{\mathcal{T}}(x,y) \overset{K,C}{\asymp} d_S(x,y) + \sum_Y \left[d_Y(x,y)\right]_{M_0} \tag{3}$$

Here, the sum is over all (annular and non-annular) proper subsurfaces.

Proof. Fix a small $\epsilon_0 > 0$ and choose any sufficiently large threshold M_0 such that Lemma 7 holds for both e^{6M_0} and $M_0/6$ and such that $M_0/6 \ge 36\log_+(1/\epsilon_0)$. Let $K \ge 1$ and $C \ge 0$ denote the larger of the quasi-isometry constants given by Lemma 7 for the thresholds e^{6M_0} and $M_0/6$.

Notice that, in any sum of the form $\sum [f]_M$, raising the threshold can only decrease the value of the sum, and lowering the threshold can only increase its value. Therefore, combining Lemma 7 and Corollary 9 we find that for any $x, y \in \mathcal{T}(S)$ the various distances satisfy

$$\begin{aligned} d_{\mathcal{T}} &\le K\left(d_S + \sum_V [d_V]_{e^{6M_0}} + \sum_{\partial A \in \Gamma_{xy}} [d_A]_{e^{6M_0}} + \sum_{\partial A \notin \Gamma_{xy}} [H_A]_{6M_0}\right) + C \\ &\le 6K\left(d_S + \sum_V [d_V]_{M_0} + \sum_{\partial A \in \Gamma_{xy}} [d_A]_{M_0} + \sum_{\partial A \notin \Gamma_{xy}} [d_A]_{M_0}\right) + C, \end{aligned}$$

where we have suppressed the x and y in the notation. The lower bound on $d_{\mathcal{T}}(x,y)$ is similar. □

2.5. **Thinness and time-ordering.** We will use some results from Rafi's work [17, Prop 3.7] combinatorializing the Teichmüller metric. For every Teichmüller geodesic and every proper subsurface V, there is a (possibly empty) interval along the geodesic where ∂V is short. Outside of this interval, the projections d_V move by at most a bounded amount. In the form that we will use below: there is a global constant M and constants $\epsilon_0 < \epsilon_1$ such that for any pair of points $x, y \in \mathcal{T}(S)$ there is a possibly nonempty connected interval T_V along the geodesic segment $[x,y]$ such that

- for $a \in \mathrm{T}_V$ the length of ∂V on a is at most ϵ_1;
- for $a \in [x,y] \setminus \mathrm{T}_V$ the length of ∂V on a is at least ϵ_0;

- for a, b in the same component of $[x, y] \setminus \mathrm{T}_V$, we have $d_V(a, b) < \mathsf{M}$; and
- if $V \pitchfork W$ then $\mathrm{T}_V \cap \mathrm{T}_W = \emptyset$.

We should note that the interval T_V is not uniquely defined. We also note that the second condition says that on the complement of the union of thin intervals the point lies in the ϵ_0–thick part of $\mathcal{T}(S)$. If $\mathrm{T}_V \neq \emptyset$ we will say that V is *thin* along T_V. In particular if $d_V(x, y) \geq \mathsf{M}$, then the interval $\mathrm{T}_V \neq \emptyset$.

We will write $\mathrm{T}_V < \mathrm{T}_W$ along $[x, y]$ if both endpoints of T_V occur before both endpoints of T_W when traveling from x to y. This T_V is called the *thin interval* for V (or the *active interval* in some papers). Note that for us thin intervals are segments in Teichmüller space, whereas Rafi works with the corresponding time intervals $I_V \subset \mathbb{R}$. The geodesic $[x, y]$ is suppressed in the notation T_V, and so the same notation re-occurs when there are multiple segments in an argument; the geodesic with respect to which the interval is defined should be clear from context. We will take M to be large enough to be a valid threshold in the distance formula (3).

The properties of thin intervals imply the following *time-ordering principle.*

Lemma 11 (Time ordering [18])**.** *Choose any constant $M_0 \geq \mathsf{M}$. Consider a pair of geodesic segments in $\mathcal{T}$ with a common basepoint x which end in y, y', respectively. Suppose $V \pitchfork W$ are transversely intersecting subsurfaces of S.*

(1) *If $d_V(x, y), d_V(x, y'), d_W(x, y), d_W(x, y') \geq 3M_0$, then the thin intervals $\mathrm{T}_V, \mathrm{T}_W$ appear in the same order along both geodesics, as they are traced out from x.*
(2) *If $d_V(x, y), d_W(x, y), d_W(x, y') \geq 3M_0$ and T_V appears before T_W along $[x, y]$, then V determines a thin interval along $[x, y']$ which appears before T_W.*

Proof. Assume the pairs of endpoints have large projection to V and W as in the hypothesis of the first statement, and suppose T_V appears before T_W along $[x, y]$. Since the endpoints of T_V contain ∂V in their markings, we have $d_W(x, \partial V) \leq M_0$. If the intervals appear in the opposite order along $[x, y']$, then letting z be the endpoint of T_V closest to x, since z contains ∂V in its marking, we use the triangle inequality to get

$$d_W(x, \partial V) = d_W(x, z) \geq 3M_0 - M_0 = 2M_0,$$

a contradiction. This proves the first statement. Turning to the second statement, the assumption on $[x, y]$ gives us that

$$d_V(x, \partial W) \geq 2M_0.$$

Let z' be the endpoint of T_W along $[x, y']$ closest to x. It contains ∂W in its marking. We therefore have

$$d_V(x, z') \geq 2M_0 \geq \mathsf{M},$$

and so T_V must appear between x and z' along $[x, y']$. □

2.6. **Reverse triangle inequality.** We will repeatedly use the fact that the projection of a Teichmüller geodesic to the curve complex of any subsurface forms an unparameterized quasi-geodesic that, in particular, does not backtrack. This phenomenon is captured by the following "reverse triangle inequality," which was proved first in the case of the curve complex of the whole surface by Masur–Minsky [14] and then for general subsurfaces by Rafi [18, Thm B].

Lemma 12 (Reverse triangle inequality)**.** *There exists $B > 0$ such that for any nonannular subsurface V (including S itself) and for any geodesic interval $[x, y]$ and any subinterval $[a, b] \subset [x, y]$ we have*

$$d_V(x,a) + d_V(a,y) \le d_V(x,y) + B, \text{ and} \quad (4)$$
$$d_V(a,b) \le d_V(x,y) + B.$$

For an annulus A, these inequalities hold with the twisting distance $d_{\mathcal{C}(A)}$, but not necessarily with the projection distance d_A.

In the exceptional annulus case, we have the following theorem from Rafi [17].

Theorem 13 (R.T.I. exception)**.** *For any sufficiently large M_0, there exists a constant $B' > 0$ with the following property. For any geodesic segment $[x, y]$ and any annulus A, if $a \in [x, y]$ is such that $d_A(x, a) + d_A(a, y) - d_A(x, y) \ge B'$ (i.e., the reverse triangle inequality fails), then there exists a proper subsurface $V \ne A$ containing a family of subsurfaces $W_i \subset V$ such that*

- *$\partial V = \partial A$,*
- *the W_i fill V,*
- *for each W_i the reverse triangle inequality (4) holds along $[x, y]$,*
- *$d_{W_i}(a, y) \ge M_0$ for each i,*
- *$d_A(a, y) \le \sum d_{W_i}(a, y)$.*

We remark that this is not exactly how the result in [17] is stated. Rafi finds an annulus about the short curve which with respect to the quadratic differential is a disjoint union of a *flat* annulus and an *expanding* annulus. Each is foliated by equidistant lines. In the flat annulus case, the lines are geodesics of the quadratic differential and have 0 curvature. In the latter case they have negative curvature. Rafi measures the path traveled in $\mathbb{H}^2$ defined by the length and twist coordinates by computing the modulus of these annuli. He shows that the distance traveled in $\mathbb{H}^2$ due to the expanding annulus is much smaller than the Teichmüller distance and if the reverse triangle inequality fails, it is due to the presence of an expanding annulus whose modulus is much bigger than the modulus of the flat annulus. The fact that path length in $\mathbb{H}^2$ is much smaller than Teichmüller length forces, by his distance formula, the presence of the domains W_i as in the statement of the theorem.

Going forward, we fix once and for all a constant M large enough to satisfy the quantitative parts of the thinness statements, the distance formula, and these reverse triangle inequality statements.

3. Comparing measures

To address averaging questions, one of course needs to consider a measure. In the present context of metric geometry, it is perhaps most natural to consider Hausdorff measure of the appropriate dimension.

Definition 14 (Hausdorff measure)**.** The n-dimensional Hausdorff measure on a metric space will be denoted by η. It is defined by

$$\eta(E) := \lim_{\delta \to 0} \left[\inf \sum \operatorname{diam}(U_i)^n \right],$$

where the infimum is over countable covers $\{U_i\}$ of E with $\operatorname{diam} U_i < \delta \ \forall i$.

For the Teichmüller metric, there is a nontrivial h-dimensional Hausdorff measure. As we shall see, in order to understand average distances with respect to this measure, it will be necessary to compare with other measures, defined below, which are also natural to consider in their own right.

3.1. **Measures on Finsler manifolds.** The Teichmüller space carries several natural volume forms coming from its structure as a Finsler manifold. Let us discuss these general constructions first before returning to the case of $M = \mathcal{T}(S)$. The treatment closely follows the survey by Álvarez and Thompson [1].

Recall that a Finsler metric on M is a continuous function $F\colon T(M) \to \mathbb{R}$ that restricts to a norm on each tangent space $T_x(M)$. There is a dual norm on each cotangent space $T^*_x(M)$. For a point $x \in M$, let $B_x \subset T_x(M)$ and $B^*_x \subset T^*_x(M)$ denote the unit balls for these two norms. A local coordinate system $(x_1, \ldots, x_n)$ on M induces a pair of isomorphisms

$$\phi\colon T_x(M) \to \mathbb{R}^n \qquad \text{and} \qquad \psi\colon T^*_x(M) \to \mathbb{R}^n \tag{5}$$

defined by writing vectors and covectors with respect to the dual bases $\{\partial_{x_1}, \ldots, \partial_{x_n}\}$ and $\{dx_1, \ldots, dx_n\}$. By definition of the dual norm, the pairing $T_x(M) \times T^*_x(M) \to \mathbb{R}$ is sent to the standard inner product on $\mathbb{R}^n$ under these isomorphisms. In the local coordinate chart we may now define two functions

$$f(x) = \frac{\varepsilon_n}{\lambda\left(\phi(B_x)\right)} \qquad \text{and} \qquad g(x) = \frac{\lambda\left(\psi(B^*_x)\right)}{\varepsilon_n},$$

where λ is Lebesgue measure and $\varepsilon_n := \lambda(\mathrm{Ball}^n)$ is the Lebesgue measure of the standard unit ball in $\mathbb{R}^n$. While these functions clearly depend on the choice of coordinates $(x_1, \ldots, x_n)$, one may easily check that the n–forms

$$f(x)\, dx_1 \wedge \cdots \wedge dx_n \qquad \text{and} \qquad g(x)\, dx_1 \wedge \cdots \wedge dx_n$$

are independent of the coordinate system and therefore define global volume forms on M. The former is called the *Busemann volume* on the Finsler manifold and the latter is the *Holmes–Thompson volume*; see [1] for more details. These both define measures on M.

A third measure to consider is the one induced by the canonical symplectic form ω on the cotangent bundle, defined as follows. Consider local coordinates $(x_1, \ldots, x_n)$ defined in a neighborhood $U \subset M$. The 1–forms $dx_1, \ldots, dx_n$ then give a trivialization of $T^*(M)$ over U, and we have a local coordinate system on $T^*(M)$ given by

$$(x_1, y_1, \ldots, x_n, y_n) \mapsto \left((x_1, \ldots, x_n), \sum_{i=1}^{n} y_i\, dx_i \right). \tag{6}$$

In these coordinates the canonical symplectic form may be written simply as $\omega = \sum dx_i \wedge dy_i$. Taking exterior powers then yields a volume form $\mu_{\mathrm{sp}} = \omega^n/n!$ on $T^*(M)$. By restricting to the unit disk bundle $T^{*,\leq 1}(M)$ and pushing forward by the projection $\pi : T^*(M) \to M$, we obtain a *symplectic measure* $\mathbf{n}$ on M.

Finally, a Finsler metric on a smooth manifold M^n induces a path metric d in the usual way, and this in turn gives rise to a Hausdorff measure in any dimension.

Recall that a centrally symmetric convex body $\Omega \subset \mathbb{R}^n$ determines a *polar body* $\Omega^\circ \subset (\mathbb{R}^n)^* = \mathbb{R}^n$ via

$$\Omega^\circ := \{\xi \in \mathbb{R}^n \mid \xi \cdot v \leq 1\ \forall v \in \Omega\}.$$

The *Mahler volume* of Ω is then defined to be the product $M(\Omega) := \lambda(\Omega)\cdot\lambda(\Omega^\circ)$ of the Lebesgue volumes of Ω and Ω°. For any centrally symmetric convex body Ω, it is known that

$$\frac{\varepsilon_n^2}{n^{n/2}} \leq M(\Omega) \leq \varepsilon_n^2 = M(\mathrm{Ball}^n). \tag{7}$$

The first inequality was established by John [8], and the latter, which gives an equality if and only if the norm is Euclidean, is known as the Blaschke–Santaló inequality [3].

Theorem 15 (Assembling facts on Finsler measures)**.** *Suppose that M^n is a continuous Finsler manifold. Then*

- *the Busemann measure μ_{B} and the n-dimensional Hausdorff measure η are equal;*
- *the Holmes–Thompson measure μ_{HT} and the symplectic measure $\mathbf{n}$ are scalar multiples: $\mu_{\mathrm{HT}} = \frac{1}{\epsilon_n}\mathbf{n}$;*
- *$\mu_{\mathrm{HT}} \leq \mu_{\mathrm{B}} \leq (n^{n/2})\,\mu_{\mathrm{HT}}$,with equality of measures if and only if the metric is Riemannian.*

Note that it is still possible for μ_{HT} and μ_{B} to be scalar multiples of each other in the non-Riemannian case, for instance on a vector space with a Finsler norm.

Proof. The first statement was originally shown by Busemann in the 1940s in [4] and is stated in modern language in [1, Thm 3.23].

The second statement is straightforward and we include a proof for completeness. Working in the local coordinates and applying the Fubini theorem, we see that the Holmes–Thompson volume of a subset $E \subset M$ is given by:

$$\begin{aligned}\int_E g(x)\,dx_1 \wedge \cdots \wedge dx_n &= \int_E \left(\int_{\psi(B_x^*)} \frac{1}{\varepsilon_n} d\lambda\right) dx_1 \wedge \cdots \wedge dx_n \\ &= \frac{1}{\varepsilon_n}\int_{\pi^{-1}(E)\cap T^{*,\leq 1}(M)} dy_1 \wedge \cdots \wedge dy_n \wedge dx_1 \wedge \cdots \wedge dx_n \\ &= \frac{1}{\varepsilon_n}\,\mathbf{n}(E).\end{aligned}$$

For the third statement, recall that the measures are defined by

$$\mu_{\mathrm{B}}(E) = \int_E f(x)\,dx_1 \wedge \cdots \wedge dx_n \quad \text{and} \quad \mu_{\mathrm{HT}}(E) = \int_E g(x)\,dx_1 \wedge \cdots \wedge dx_n.$$

For each $x \in M$, the unit ball $B_x \subset T_x(M)$ is sent to a centrally symmetric convex body $\phi(B_x) \subset \mathbb{R}^n$ under the isomorphism ϕ defined in (5). The polar body is exactly given by $\phi(B_x)^\circ = \psi(B_x^*)$. Therefore, the Mahler volume of $\phi(B_x)$ is

$$M(\phi(B_x)) = \lambda(\phi(B_x))\cdot\lambda(\psi(B_x^*)) = \varepsilon_n^2\frac{g(x)}{f(x)}.$$

Combining with (7) now implies that $n^{-n/2}f(x) \leq g(x) \leq f(x)$ for all $x \in M$. We conclude that $\mu_{\mathrm{HT}}(E) \leq \mu_{\mathrm{B}}(E) \leq n^{n/2}\mu_{\mathrm{HT}}(E)$ for all $E \subset M$. Finally, since Blaschke–Santaló can only give equality for a Euclidean norm, it follows that μ_{B} and μ_{HT} can only be equal for a Riemannian metric. $\square$

3.2. **Measures coming from quadratic differentials.** Recall that quadratic differential space $\mathcal{Q}(S)$ is naturally identified with the cotangent bundle $T^*(\mathcal{T}(S))$ of Teichmüller space, and that each quadratic differential $q \in \mathcal{Q}(S)$ has a norm $\|q\|$ given by the area of the flat structure on S induced by q. The unit disk bundle for this norm will be denoted by

$$\mathcal{Q}^{\leq 1}(S) = \{q \in \mathcal{Q}(S) : \|q\| \leq 1\}.$$

Using this disk bundle, the natural symplectic measure $\mu_{\rm sp}$ on $\mathcal{Q}(S)$ descends to a measure $\mathbf{n}$ on $\mathcal{T}(S)$ exactly as above. We note that ω and therefore $\mu_{\rm sp}$ and $\mathbf{n}$ are invariant under the action of the mapping class group.

The space $\mathcal{Q}(S)$ also carries a natural $\operatorname{Mod}(S)$–invariant measure $\mu_{\rm hol}$ that is defined in terms of holonomy coordinates and which we will refer to as *holonomy measure*; it is also sometimes called Masur–Veech measure in the literature (see [11] for details). This measure has been studied extensively, for instance to establish ergodicity results for the geodesic flow. The measure $\mu_{\rm hol}$ is also related to the "Thurston measure" $\mu_{\rm TH}$ on the space of measured foliations $\mathcal{MF}$ induced by the piecewise-linear structure of $\mathcal{MF}$ [7]. Indeed, as seen in [11], $\mu_{\rm hol}$ is equal to the pullback of $\mu_{\rm TH} \times \mu_{\rm TH}$ under the $\operatorname{Mod}(S)$-invariant map $\mathcal{Q}(S) \to \mathcal{MF} \times \mathcal{MF}$ that sends a quadratic differential to its vertical and horizontal foliations.

Just as $\mu_{\rm sp}$ induces $\mathbf{n}$, the holonomy measure $\mu_{\rm hol}$ descends to a measure $\mathbf{m}$ on $\mathcal{T}(S)$. Explicitly, the $\mathbf{m}$–measure of a set $E \subset \mathcal{T}(S)$ is given by

$$\mathbf{m}(E) := \mu_{\rm hol}\left(\pi^{-1}(E) \cap \mathcal{Q}^{\leq 1}(S)\right).$$

This measure $\mathbf{m}$ has been studied previously in [2] and [6].

Proposition 16. [13, p.3746] *There is a scalar $k > 0$ such that $\mu_{\rm sp} = k \cdot \mu_{\rm hol}$.*

We recall the outlines of the argument here. In [13], it was shown that the Teichmüller flow on $\mathcal{Q}(S)$ is a Hamiltonian flow for the function

$$H(q) = \frac{\|q\|^2}{2}.$$

As such, the Teichmüller flow preserves the symplectic form ω and the corresponding measure $\mu_{\rm sp}$. The measures $\mu_{\rm sp}$ and $\mu_{\rm hol}$ both descend to the quotient space $\mathcal{Q}(S)/\operatorname{Mod}(S)$; furthermore, the latter defines an ergodic measure for the Teichmüller flow on $\mathcal{Q}(S)/\operatorname{Mod}(S)$ [11]. Since $\mu_{\rm sp}$ is absolutely continuous with respect to $\mu_{\rm hol}$, the result follows.

We therefore also have $\mathbf{n} = k\mathbf{m}$, and combining Proposition 16 with Theorem 15 we get:

Corollary 17. *There are scalars $k_2 > k_1 > 0$ such that*

$$k_1 \mathbf{m} \leq \eta \leq k_2 \mathbf{m}.$$

3.3. **Visual measures.** The unit sphere subbundle of $\mathcal{Q}(S)$ will be denoted by

$$\mathcal{Q}^1(S) = \{q \in \mathcal{Q}(S) \colon \|q\| = 1\}.$$

For each $x \in \mathcal{T}(S)$, the fiber $\mathcal{Q}^1(x)$ is identified with the "space of directions" at x, and the Teichmüller geodesic flow $\varphi_t \colon \mathcal{Q}(S) \to \mathcal{Q}(S)$ gives rise to a homeomorphism

$$\begin{array}{cccc} \Psi_x \colon & \mathcal{Q}^1(x) \times (0,\infty) & \to & \mathcal{T}(S) \setminus \{x\} \\ & (q,r) & \mapsto & \pi(\varphi_r(q)) \ , \end{array}$$

which serves as "polar coordinates" centered at x. Furthermore, this conjugates φ_t to a *radial flow* based at x given by

$$\hat{\varphi}_t(\pi(\varphi_r(q))) := \pi(\varphi_{r+t}(q)).$$

We will consider measures on $\mathcal{T}(S)$ that are compatible with these polar coordinates and with the radial flow.

Definition 18 (Visual measure)**.** Given any measure κ_x on the unit sphere $\mathcal{Q}^1(x) \cong S^{h-1}$, we define the corresponding *visual measures* on $\mathcal{S}_r(x)$ and $\mathcal{T}(S)$ as follows. Firstly, the visual measure $\mathrm{Vis}_r(\kappa_x)$ on the sphere $\mathcal{S}_r(x)$ of radius r is just the push-forward of $e^{hr}\kappa_x$ under the homeomorphism $\mathcal{Q}^1(x)\times\{r\} \cong \mathcal{S}_r(x)$. Integrating these over $(0,\infty)$ then gives a visual measure on $\mathcal{T}(S)$ defined by

$$\mathrm{Vis}(\kappa_x)(E) := \int_{(q,r)\in E\subset \mathcal{Q}^1(S)\times(0,\infty)} e^{hr}d\kappa_x(q)d\lambda(r).$$

Said differently, $\mathrm{Vis}(\kappa_x)$ is equal to the push-forward of $\kappa_x \times \lambda_0$ under the homeomorphism Ψ_x, where λ_0 is the weighted Lebesgue measure on $(0,\infty)$ given by $\lambda_0([a,b]) = \int_a^b e^{hr}d\lambda(r) = (e^{hb}-e^{ha})/h$. (We have scaled things in this way so that the visual measure of the ball of radius R grows like e^{hR}.)

The essential feature of visual measures is that they enjoy the following "normalized invariance" under the radial flow: For any $t\geq 0$ and measurable $E\subset \mathcal{S}_r(x)$ we have

$$\frac{\mathrm{Vis}_{r+t}(\kappa_x)(\hat{\varphi}_t(E))}{\mathrm{Vis}_{r+t}(\kappa_x)(\mathcal{S}_{r+t}(x))} = \frac{\mathrm{Vis}_r(\kappa_x)(E)}{\mathrm{Vis}_r(\kappa_x)(\mathcal{S}_r(x))}.$$

The same invariance holds for $\mathrm{Vis}(\kappa_x)$ when we normalize with respect to annular shells $\mathcal{B}_b(x)\setminus\mathcal{B}_a(x)$ instead of spheres.

There are two visual measures that specifically interest us. Firstly, the normed vector space $\mathcal{Q}(x)$ carries a unique translation-invariant measure ν_x normalized so that $\nu_x(B_x^*) = 1$; recall that the unit ball B_x^* is just the intersection $\mathcal{Q}^{\leq 1}(S)\cap\mathcal{Q}(x)$. This induces a measure (also denoted ν_x) on the unit sphere $\mathcal{Q}^1(x)$ via the usual method of coning off: $\nu_x(E) := \nu_x([0,1]\times E)$ for $E\subset\mathcal{Q}^1(x)$.

Secondly, since $\mathcal{Q}(S)$ has the structure of a fiber bundle over $\mathcal{T}(S)$, we can define a conditional measure s_x on $\mathcal{Q}(x)$ by disintegration from μ_{hol}. More precisely, s_x is the unique measure on $\mathcal{Q}(x)$ such that the μ_{hol}–measure of $E\subset\mathcal{Q}(S)$ is given by

$$\mu_{\mathrm{hol}}(E) = \int_{\mathcal{T}(S)} s_x(E\cap\mathcal{Q}(x))\,d\mathbf{m}(x).$$

Via the process of coning off, we again think of s_x as a measure on $\mathcal{Q}^1(x)$. The corresponding visual measure $\mathrm{Vis}(s_x)$ is crucial to our argument because of the following comparison theorem of Athreya, Bufetov, Eskin and Mirzakhani.

Theorem 19 (Holonomy vs. visual measure [2, Prop 2.5])**.** *For any $x\in\mathcal{T}(S)$, there exists a constant c_1 such that*

$$\mathbf{m} \leq c_1\cdot\mathrm{Vis}(s_x).$$

The space $\mathcal{Q}(S)$ of quadratic differentials is a complex vector bundle; as such, there is a natural circle action $S^1 \curvearrowright \mathcal{Q}(S)$ that preserves each fiber $\mathcal{Q}(x)$ and unit sphere $\mathcal{Q}^1(x)$. We say that a visual measure $\mathrm{Vis}(\kappa_x)$ is *rotation-invariant* if the corresponding measure κ_x on $\mathcal{Q}^1(x)$ is invariant under this action of S^1. The visual measure $\mathrm{Vis}(\nu_x)$ is rotation-invariant because S^1 preserves the unit ball B_x. Similarly, $\mathrm{Vis}(s_x)$ is rotation-invariant because S^1 preserves μ_{hol}.

3.4. **Summary.** The measures on $\mathcal{T}(S)$ considered above are $\mathbf{n}$ and $\mathbf{m}$ (induced by the symplectic and holonomy measures on $\mathcal{Q}(S)$, respectively, via the covering map), Hausdorff measure η, the visual measures $\mathrm{Vis}(\kappa_x)$ created by radially flowing measures on the sphere of directions $\mathcal{Q}^1(x)$, and the measures μ_{B} and μ_{HT} coming from the Finsler structure.

We found that $\mathbf{n}$, $\mathbf{m}$, and μ_{HT} are scalar multiples of each other, Hausdorff measure and Busemann measure coincide, and all of these are mutually comparable in the sense of being bounded above and below by scalar multiples of each other. Thus, we get from Theorem 19 that all are absolutely continuous with respect to the visual measure $\mathrm{Vis}(s_x)$. In the following section we will get results about the structure of generic geodesic rays with respect to any rotation-invariant visual measure and deduce results for the other measures considered here.

4. Statistics for the geometry of Teichmüller rays

4.1. **Separation and thickness of rays.** We need an estimate for how likely it is for two geodesics to fellow-travel past some radius R_0. The next theorem gives the appropriate sort of estimate for pairs of geodesics that are nearby at a given radius. Proposition 23 then extends this to geodesics that are near each other at any time during an interval $[R_0, R]$. After that, we use the ergodicity of the geodesic flow to conclude that most rays are thick for a definite proportion of the time.

Definition 20. For any $r > 1$, let $\mathcal{A}_r(x) = \mathcal{B}_r(x) \setminus \mathcal{B}_{r-1}(x)$ be the annular shell between radius r and $r-1$. We say a measure μ satisfies an *exponential decay estimate* if given any $M_0 > 0$ there exist C, R_0 such that for all t, r with $R_0 \le t \le r$ and any $x_1 \neq x$ we have

$$\mu\big\{x_2 \in \mathcal{A}_r(x) :\ d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0\big\} < Ce^{-t} \cdot \mu(\mathcal{A}_r(x)),$$

where γ_i is the geodesic ray based at x and passing through x_i.

Lemma 21. *Every rotation-invariant visual measure $\mu_x = \mathrm{Vis}(\kappa_x)$ on $\mathcal{T}(S)$ satisfies an exponential decay estimate.*

Proof. Choose $0 < t \le r$, fix a point $x_1 \neq x$, and let $E = \{x_2 \in \mathcal{A}_r(x) : d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0\}$. Looking instead in the sphere $\mathcal{S}_t(x)$, we have the set $E' = \{z \in \mathcal{S}_t(x) : d_{\mathcal{T}}(\gamma_1(t), z) < M_0\}$. Notice that, by definition,

$$E = \bigcup_{s \in [r-t-1, r-t]} \hat{\varphi}_s(E').$$

Therefore, by the normalized invariance, we have

$$\begin{aligned}
\mu(E) &= \int_{r-1}^{r} \mathrm{Vis}_s(\kappa_x)(\hat{\varphi}_{s-t}(E'))\, d\lambda(s) \\
&= \int_{r-1}^{r} \mathrm{Vis}_t(\kappa_x)(E') \frac{\mathrm{Vis}_s(\kappa_x)(\mathcal{S}_s(x))}{\mathrm{Vis}_t(\kappa_x)(\mathcal{S}_t(x))}\, d\lambda(s) \\
&= \frac{\kappa_x(E')}{\kappa_x(\mathcal{Q}^1(x))} \mu_x(\mathcal{A}_r(x)),
\end{aligned}$$

where, in the last line, we have identified E' with its image in $\mathcal{Q}^1(x) \cong \mathcal{S}_t(x)$.

It remains to show that $\kappa_x(E')/\kappa_x(\mathcal{Q}^1(x)) \leq Ce^{-t}$ when t is large. Recall that S^1 acts freely on $\mathcal{Q}^1(x)$ by rotations. Choosing orbit representatives, we may realize $\mathcal{Q}^1(x)$ as a setwise product $(\mathcal{Q}^1(x)/S^1) \times S^1$. The measure κ_x pushes forward to a measure on $\mathcal{Q}^1(x)/S^1$. By disintegration, we then obtain a measure on each fiber S^1 which, by the rotation-invariance of κ_x, must agree with Lebesgue measure up to a scalar. For any two points $z, z' \in E'$, the triangle inequality gives $d_{\mathcal{T}}(z, z') \leq 2M_0$. Now suppose that z and z' lie in the same Teichmüller disk, meaning that the unit quadratic differentials associated to the geodesics $[x, z]$ and $[x, z']$ lie in the same S^1–orbit. Each Teichmüller disk is an isometrically embedded copy of the hyperbolic plane. Thus, when t is large compared to M_0, hyperbolic geometry implies that the fraction of each S^1–orbit contained in E' is at most Ce^{-t} for some constant C. Using the product structure and integrating over the $\mathcal{Q}^1(x)/S^1$ factor, Fubini's theorem then implies that $\kappa_x(E')/\kappa_x(\mathcal{Q}^1(x)) \leq Ce^{-t}$ as well. □

Theorem 22 (Exponential decay of fellow-travelers)**.** *The visual measures* $\mathrm{Vis}(\nu_x)$ *and* $\mathrm{Vis}(s_x)$*, the holonomy measure* $\mathbf{m}$*, and the Hausdorff measure* η *all satisfy exponential decay estimates.*

Proof. The visual measures $\mathrm{Vis}(\nu_x)$ and $\mathrm{Vis}(s_x)$ are rotation-invariant, so this follows from Lemma 21 above. Since η and $\mathbf{m}$ are bounded in terms of each other (Cor 17), and $\mathbf{m}$ is bounded above by the visual measure $\mathrm{Vis}(s_x)$ (Thm 19), we need only verify that for either η or $\mathbf{m}$ the measure of $\mathcal{A}_r(x)$ is bounded below by $Ce^{hr} \sim \mathrm{Vis}(s_x)(\mathcal{A}_r(x))$ for some constant C. For η this follows from the observation that there are e^{hr} orbit points of the mapping class group in this shell; alternately, this was proved for $\mathbf{m}$ in [2]. □

We also have

Proposition 23 (Separation is forever)**.** *Suppose a measure* μ *satisfies an exponential decay estimate. Then for any* $M_0 > 0$ *and for sufficiently large* R_0 *there exists a* C *such that for all* $r \geq R_0$ *and any* $x_1 \neq x$*,*

$$\mu\big\{x_2 \in \mathcal{A}_r(x) : \; d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0 \textit{ for some } t \in [R_0, r]\big\} < Ce^{-R_0}\mu(\mathcal{A}_r(x)),$$

where γ_i *is the geodesic ray based at* x *and passing through* x_i*.*

Proof. If x_2 is such a point, then there is some $k \in \mathbb{N}$, $k \leq r - R_0$, such that $d_{\mathcal{T}}(\gamma_1(R_0 + k), \gamma_2(R_0 + k)) < M_0 + 2$. Thus our set of points is contained in the union of the exceptional sets corresponding to the radii $R_0, R_0+1, \ldots, R_0+\lfloor r-R_0 \rfloor$. Using the exponential decay estimate, we see that our set has measure at most

$$\left(e^{-R_0} + e^{-R_0-1} + \cdots + e^{-R_0-\lfloor r-R_0 \rfloor}\right) C\mu(\mathcal{A}_r(x)) \leq \left(\frac{e}{e-1}\right) Ce^{-R_0}\mu(\mathcal{A}_r(x)). \quad \square$$

Thus we can conclude that after throwing out a subset of $\mathcal{A}_r(x) \times \mathcal{A}_r(x)$ of measure which is an arbitrarily small proportion, all pairs of geodesics stay separated by an arbitrarily chosen distance in Teichmüller space after a threshold time has elapsed. Later we will show that after waiting even longer we may also assume that every pair of geodesics has big curve complex distance. This relies both on the large Teichmüller distance established above and the fact that most geodesics spend a definite fraction of their time in the thick part.

Fix a measure κ_x (either s_x or ν_x) on $\mathcal{Q}^1(x)$ and an ϵ_0–thick part $\mathcal{T}_{\epsilon_0}$.

Proposition 24 (Thickness of geodesics)**.** *There is a constant $\delta > 0$ such that for all $\epsilon > 0$ there is a threshold R_0 and a set $E \subset \mathcal{Q}^1(x)$ with $\kappa_x(E^c) \leq \rho$ so that each geodesic $[x, \gamma(r)]$ corresponding to $q \in E, r \geq R_0$ spends at least time δr in $\mathcal{T}_{\epsilon_0}$.*

Proof. By the ergodicity of the geodesic flow [11] there is $\delta > 0$ such that the geodesic determined by almost every $q \in \mathcal{Q}(S)$ spends proportion δ of its time in $\mathcal{T}_{\epsilon_0}$, asymptotically. The vertical foliation of each such q is uniquely ergodic [12]. If two quadratic differentials have the same vertical uniquely ergodic measured foliation then they are forwards asymptotic [10]. We conclude that almost every measured foliation $F \in \mathcal{MF}$ (with respect to Thurston measure μ_{TH}) has the property that for any quadratic differential with vertical foliation F, the corresponding geodesic spends at least δ proportion of its time in the thick part, asymptotically.

The map $\mathcal{Q}(x) \to \mathcal{MF}$ which assigns to q its vertical foliation is a smooth map off the multiple zero locus, so it is smooth on a set of full measure. Thus it is absolutely continuous with respect to the measures κ_x and μ_{TH}. Thus the property of asymptotically spending proportion δ in the thick part is true of almost every $q \in \mathcal{Q}(x)$. Thus for every ϵ, except for a set of κ_x measure at most ϵ, long enough geodesics are thick for proportion δ of their length. ☐

4.2. **Thick-thin structure and progress in the curve complex.** The goal of this section is to prove Proposition 33, which says that any geodesic that spends a definite fraction of its time in the thick part must move a definite amount in the curve complex. The idea is that long subintervals contained in $\mathcal{T}_{\epsilon_0}$ contribute to progress in $\mathcal{C}(S)$; alternately, one could consider intervals in the complement of $\bigcup_V \mathrm{T}_V$. For this analysis, we would like to bound the number of connected components of $\bigcup_V \mathrm{T}_V$ in terms of $d_S(x, y)$. One bound is given by the number of nonempty thin intervals. While there may be arbitrarily many such T_V, some of these will be redundant in the sense that $\mathrm{T}_V \subset \mathrm{T}_W$ for some other subsurface W.

Definition 25 (Thin-significance)**.** Recall the choice of global constant M. A proper subsurface $V \subsetneq S$ is said to be *thin-significant* for the geodesic segment $[x, y]$ if $d_{\mathcal{C}(V)}(x, y) \geq 3\mathsf{M}$ and for every other proper subsurface $Z \subsetneq S$ with $d_{\mathcal{C}(Z)}(x, y) \geq 3\mathsf{M}$ we have $\mathrm{T}_V \not\subset \mathrm{T}_Z$.

Remark. In this subsection we will focus on the curve complex distance $d_{\mathcal{C}(V)}$ for a subsurface V. Recall that this agrees with the usual projection distance d_V in the case that V is non-annular, but that $d_{\mathcal{C}(A)}$ and d_A differ for annuli. We will take care to handle exceptional annuli carefully.

Our next goal is to bound the number of thin-significant subsurfaces along an arbitrary geodesic. For this, we will use the work of Rafi–Schleimer [19] bounding the size of an antichain in the poset of subsurfaces of S.

Definition 26 (Antichain)**.** Given a subsurface $\Sigma \subset S$ a pair of points $x, y \in \mathcal{T}(S)$ and constants $T_1 \geq T_0 > 0$, a collection Ω of proper subsurfaces of Σ is an *antichain* for (Σ, x, y, T_0, T_1) if the following hold:

- if $Y, Y' \in \Omega$, then Y is not a proper subsurface of Y';
- if $Y \in \Omega$, then $d_{\mathcal{C}(Y)}(x, y) \geq T_0$; and
- if $Z \subsetneq \Sigma$ and $d_{\mathcal{C}(Z)}(x, y) \geq T_1$, then $Z \subset Y$ for some $Y \in \Omega$.

Lemma 27 (Antichain bound [19, Lem 7.1])**.** *For every $\Sigma \subset S$ and sufficiently large $T_1 \geq T_0 > 0$, there is a constant $A = A(\Sigma, T_0, T_1)$ so that if Ω is an antichain*

for (Σ, x, y, T_0, T_1) *then*
$$|\Omega| \leq A \cdot d_{\mathcal{C}(\Sigma)}(x,y).$$

We now prove a proposition showing that if there are a large enough number of thin-significant subsurfaces along a geodesic, then the image of the geodesic makes definite progress in the curve complex. The following notation will be used in the proof.

Definition 28. Consider a geodesic segment $[x,y] \subset \mathcal{T}(S)$ and a collection Ω of proper subsurfaces of S. We will consider three partial orders on the set Ω:

(1) $V \leq_1 W \iff V \subset W$,
(2) $V \leq_2 W \iff \mathrm{T}_V \subset \mathrm{T}_W$, and
(3) $V \leq_3 W \iff V \subset W$ and $\mathrm{T}_V \subset \mathrm{T}_W$.

The subcollection of Ω consisting of maximal elements with respect to $\leq_*$ will be denoted $(\Omega)_*$; notice that these sets are related by $(\Omega)_1 \subset (\Omega)_3 \supset (\Omega)_2$. Elements of $(\Omega)_1$ are said to be *topologically maximal* with respect to Ω.

Proposition 29 (Progress from thin-significant subsurfaces)**.** *For any* t_0*, there is a constant* N *such that if* $d_{\mathcal{C}(S)}(x,y) \leq t_0$*, then the number of thin-significant subsurfaces* Y *along* $[x,y]$ *is at most* N*.*

Proof. Let $\Omega = \{V \subsetneq S : d_{\mathcal{C}(V)}(x,y) \geq 3\mathsf{M}\}$ be the collection of proper subsurfaces which have a large projection. By definition, the set of thin-significant subsurfaces is exactly given by $(\Omega)_2$. On the other hand, the subcollection $(\Omega)_1$ of topologically maximal subsurfaces clearly forms an antichain for $(S, x, y, 3\mathsf{M}, 3\mathsf{M})$. By Lemma 27, we therefore have $|(\Omega)_1| \leq At_0$ for some constant A. We will extend this to a bound on the cardinality of the larger set $(\Omega)_3$; this will imply the proposition because $(\Omega)_2 \subset (\Omega)_3$.

Fix a proper subsurface $W \in \Omega$ and consider the set $\mathcal{U}_W = \{V \in (\Omega)_3 : V \subsetneq W\}$. We claim that $|\mathcal{U}_W|$ is bounded by a constant depending only on the complexity of W. By the above, this will suffice because each $V \in (\Omega)_3$ is either equal to or properly contained in some topologically maximal proper subsurface $W \in (\Omega)_1$.

First consider those $V \in \mathcal{U}_W$ for which $\mathrm{T}_V \cap \mathrm{T}_W \neq \emptyset$. The definition of $\leq_3$ implies that $\mathrm{T}_V \not\subset \mathrm{T}_W$; therefore T_V must overlap with at least one endpoint of T_W. If T_{V_1} and T_{V_2} both contain the initial endpoint of T_W, then $\mathrm{T}_{V_1} \cap \mathrm{T}_{V_2} \neq \emptyset$ and so we cannot have $V_1 \pitchfork V_2$. Since there is a universal bound on the number of subsurfaces such that no two intersect transversely, this bounds the number of $V \in \mathcal{U}_W$ for which $\mathrm{T}_V \cap \mathrm{T}_W \neq \emptyset$.

It remains to bound the number of $V \in \mathcal{U}_W$ for which $\mathrm{T}_V \cap \mathrm{T}_W = \emptyset$; we will only focus on the case $\mathrm{T}_V < \mathrm{T}_W$. Suppose that $T_W = [a,b] \subset [x,y]$ and consider the set $\Omega' = \{V \in \Omega : V \subsetneq W \text{ and } d_{\mathcal{C}(V)}(x,a) \geq 2\mathsf{M}\}$. Notice that the subcollection $(\Omega')_1$ forms an antichain for $(W, x, a, 2\mathsf{M}, 4\mathsf{M})$: the only difficulty is to check that every $Y \subsetneq W$ with $d_{\mathcal{C}(Y)}(x,a) \geq 4\mathsf{M}$ is contained in an element of $(\Omega')_1$. However, this is true because the reverse triangle inequality guarantees that $d_{\mathcal{C}(Y)}(x,y) \geq d_{\mathcal{C}(Y)}(x,a) - B \geq 3\mathsf{M}$ and therefore that $Y \in \Omega'$. Since $d_{\mathcal{C}(W)}(x,a) \leq \mathsf{M}$, Lemma 27 now gives a bound on $|(\Omega')_1|$. Finally, notice that for each $V \in \mathcal{U}_W$ with $\mathrm{T}_V < \mathrm{T}_W$ the triangle inequality gives $d_{\mathcal{C}(V)}(x,a) \geq d_{\mathcal{C}(V)}(x,y) - \mathsf{M} \geq 2\mathsf{M}$ and so ensures that $V \in \Omega'$. Therefore each such V is contained in some topologically maximal $Z \in \Omega'$; that is to say, each $V \in \mathcal{U}_W$ with $\mathrm{T}_V < \mathrm{T}_W$ is contained in $\mathcal{U}_Z$ for some $Z \in (\Omega')_1$. The bound on $|\mathcal{U}_W|$ now follows by induction on the complexity of the subsurface W. $\square$

Definition 30. Define a constant $\mathsf{P} := 36 \max \left\{ \log_+(1/\epsilon_0), \log_+(3\mathsf{M}) \right\}$. Say that an annular subsurface A has an *exceptional thin interval* T_A along $[x,y]$ if $d_{\mathcal{C}(A)}(x,y) \leq 3\mathsf{M}$ but $d_A(x,y) \geq \mathsf{P}$. By Lemma 8, this is only possible if $l_x(\partial A) < \epsilon_0$ or $l_y(\partial A) < \epsilon_0$. Therefore, such an annulus must determine a nonempty thin interval along $[x,y]$ that contains either x or y. Since all annuli A with $l_x(\partial A) < \epsilon_0$ must be disjoint, we see that there is a universal bound (namely $6g-6 = h$) on the number of annuli with exceptional thin intervals along an arbitrary geodesic $[x,y]$.

We now define the *primary thin portion* $\mathcal{W}$ of a geodesic segment $[x,y]$ to be the union of thin intervals T_V for all non-annular proper subsurfaces with $d_V(x,y) \geq 3\mathsf{M}$ and all annular subsurfaces with $d_{\mathcal{C}(A)}(x,y) \geq 3\mathsf{M}$ or $d_A(x,y) \geq \mathsf{P}$. In the case that $d_S(x,y) \leq t_0$, Proposition 29 implies that $\mathcal{W}$ is the union of at most $N+h$ thin intervals, namely, those corresponding to the thin-significant subsurfaces and to annuli with exceptional thin intervals.

While $\mathcal{W}$ does contain most of the nonempty thin intervals along the geodesic, it need not cover the entire time that $[x,y]$ spends in the thin part of Teichmüller space. Nevertheless, the projections to all proper subsurfaces remain uniformly bounded on the complement of $\mathcal{W}$.

Lemma 31 (Complement of $\mathcal{W}$)**.** *There exists a constant M' with the following property. If $[a,b] \subset [x,y] \setminus \mathcal{W}$ is a connected interval in the complement of the primary thin portion of $[x,y]$, then $d_Y(a,b) \leq M'$ for all proper subsurfaces $Y \subsetneq S$.*

Proof. First suppose that Y satisfies the reverse triangle inequality (4) along $[x,y]$. If Y is non-annular and $d_Y(x,y) \geq 3\mathsf{M}$, or if Y is an annulus and $d_{\mathcal{C}(Y)}(x,y) \geq 3\mathsf{M}$ or $d_Y(x,y) \geq \mathsf{P}$, then $\mathrm{T}_Y \subset \mathcal{W}$ by definition. Therefore $d_Y(a,b) \leq \mathsf{M}$ since $[a,b] \cap \mathrm{T}_Y = \emptyset$. If this is not the case, then the reverse triangle inequality gives $d_Y(a,b) \leq d_Y(x,y) + B \leq 4\mathsf{M} + \mathsf{P}$ as claimed.

It remains to consider an annular subsurface $A \subset S$ for which the reverse triangle inequality fails. We may assume that $d_{\mathcal{C}(A)}(x,y) \leq 3\mathsf{M}$ and $d_A(x,y) \leq \mathsf{P}$, for otherwise we have $\mathrm{T}_A \subset \mathcal{W}$ and $d_A(a,b) \leq \mathsf{M}$ as above. Let B' be the constant corresponding to the threshold $5\mathsf{M}+\mathsf{P}$ in Theorem 13 (R.T.I. exception). According to that theorem, applied to the geodesic $[a,y]$, we either have $d_A(a,b) + d_A(b,y) \leq d_A(a,y) + B'$, or there exist subsurfaces W_i that satisfy the reverse triangle inequality and which have $d_{W_i}(a,b) \geq 5\mathsf{M} + \mathsf{P}$. However, as we have seen above, there are no such proper subsurfaces. Therefore the former inequality must hold. We similarly have $d(x,a) + d(a,b) \leq d(x,b) + B'$. Adding these inequalities and using the triangle inequality then gives $d_A(a,b) \leq d_A(x,y) + B' \leq \mathsf{P} + B'$. ☐

By the distance formula (3), it follows that long intervals in the complement of $\mathcal{W}$ must travel a large distance the curve complex $\mathcal{C}(S)$ of the whole surface. The following lemma says that each such subinterval contributes to the curve complex distance along the total geodesic.

Lemma 32 (Cumulative contribution of subintervals)**.** *There exist constants $0 < \rho_1 < 1$ and $D_1 > 0$ such that for all $d > D_1$, if $[x,y]$ is a Teichmüller geodesic that contains n subintervals $[x_i, y_i]$ with disjoint interiors whose endpoints satisfy $d_S(x_i, y_i) \geq d$, then*

$$d_S(x,y) \geq \rho_1 n d.$$

Proof. Applying the reverse triangle inequality (4) to the points x_i and y_i we have $d_S(x, x_i) + d_S(x_i, y_i) + d_S(y_i, y) \leq d_S(x,y) + 2B$. By recursively applying this observation to $[x, x_i]$ and $[y_i, y]$ and then throwing out the complementary intervals, we find have that

$$d_S(x, y) \geq \sum d_S(x_i, y_i) - 2nB \geq nd - 2nB.$$

Choose $D_1 > 4B$ and $\rho_1 = 1/2$. Then for $d \geq D_1$ the quantity on the right side is at least $\rho_1 nd$. $\square$

We now fix once and for all a "definite progress" constant $\mathsf{D} > 0$ sufficiently large so that $\rho_1\mathsf{D} > D_1$ (and thus $\mathsf{D} > D_1$ as well). Applying the distance formula (3) with the threshold M' given by Lemma 31, we have quasi-isometry constants K, C such that $d_{\mathcal{T}}(a, b) \leq K d_S(a, b) + C$ for any connected interval $[a, b] \subset [x, y] \setminus \mathcal{W}$. This gives rise to a fixed value L such that any interval $[a, b]$ of length at least L that lies entirely in $[x, y] \setminus \mathcal{W}$ satisfies $d_S(a, b) \geq \mathsf{D}$; for example, any value $\mathsf{L} \geq K\mathsf{D} + C$ will suffice. Thus according to Lemma 32, if I is any interval along a geodesic that contains a subinterval of length L that is disjoint from $\mathcal{W}$, then the distance in the curve complex between the endpoints of I is at least $\rho_1\mathsf{D}$.

Furthermore by Proposition 29 associated to the constant $t_0 = \rho_1\mathsf{D}$, there is a constant N so that the conclusion of Proposition 29 holds.

Proposition 33 (Definite progress)**.** *For each $0 < \delta < 1$, there exist constants $\rho, R_1 > 0$ with the following property. If $[x, y]$ is a Teichmüller geodesic of length $r \geq R_1$ that spends at least time δr in $\mathcal{T}_{\epsilon_0}$, then $d_S(x, y) \geq \rho r$.*

Proof. Let N denote the constant obtained by applying Proposition 29 with $t_0 = \rho_1\mathsf{D}$. Choose n so that $n\delta > 1$ and make the following definitions:

$$\delta' = \frac{n\delta - 1}{n - 1}, \quad T_0 \geq \frac{\mathsf{L}(\mathsf{N} + h + 1)}{\delta'}, \quad R_1 = 2T_0, \quad \rho = \frac{\rho_1^2\mathsf{D}}{2nT_0}.$$

Let $[x, y]$ be a Teichmüller geodesic of length $r \geq R_1$ that spends at least δr in the thick part. Set $m = \lfloor r/T_0 \rfloor$ and divide $[x, y]$ into m subsegments of length $r/m \geq T_0$. Let us say that a subsegment $[a, b] \subset [x, y]$ is *stalled* if $d_S(a, b) < \rho_1\mathsf{D}$ and *progressing* if $d_S(a, b) \geq \rho_1\mathsf{D}$. Suppose that m_1 of the subsegments are stalled, and thus $m_2 = m - m_1$ are progressing. Given a stalled segment $[a, b]$, we decompose it into its primary thin portion $\mathcal{W}$ and the corresponding complementary subintervals. Since the interval is stalled, Proposition 29 ensures that $\mathcal{W}$ is the union at most $\mathsf{N} + h$ thin intervals. Therefore we can conclude that $\mathcal{W}$ has at most $\mathsf{N} + h + 1$ complementary subintervals in $[a, b]$. Furthermore, each complementary subinterval has length at most L, for otherwise we would have $d_S(a, b) \geq \rho_1\mathsf{D}$ by the preceding paragraph. Since $\mathcal{W}$ is contained in the thin part, we see that the total amount of time that this interval $[a, b]$ spends in the thick part is at most

$$(\mathsf{N} + h + 1)\mathsf{L} \leq \delta' T_0 \leq \delta' r/m.$$

Therefore the total amount of time that the full interval $[x, y]$ spends in the thick part is at most

$$\left(\delta' \frac{r}{m}\right) m_1 + \left(\frac{r}{m}\right) m_2 = \frac{r}{m}(\delta' m_1 + m_2).$$

We claim that $m_2 \geq m/n$. If this were not the case, then we necessarily have $m_1 > (n - 1)m/n$. Since $\delta' < 1$, it follows that

$$\delta' \cdot m_1 + 1 \cdot m_2 < \delta' \cdot m \frac{n - 1}{n} + 1 \cdot m \frac{1}{n}$$

where the inequality is valid by the elementary fact that for any constants $a, b, c, d, \alpha, \beta$ such that $a + b = c + d$ and $0 < \alpha < \beta$ we have

$$\alpha\cdot a + \beta\cdot b < \alpha\cdot c + \beta\cdot d \iff a > c. \tag{8}$$

But then the amount of time that $[x, y]$ is thick is less than

$$\frac{r}{m}\left(\delta' m\frac{n-1}{n} + m\frac{1}{n}\right) = r\left(\frac{n\delta - 1}{n-1}\cdot\frac{n-1}{n} + \frac{1}{n}\right) = r\delta,$$

which contradicts the assumption on $[x, y]$. Therefore $m_2 \geq m/n$, as claimed.

On each of the m_2 progressing intervals, the curve complex distance between endpoints is at least $\rho_1\mathsf{D}$. Therefore, cumulative contribution of subintervals (Lemma 32) implies that

$$d_S(x, y) \geq \rho_1 m_2(\rho_1\mathsf{D}) \geq \rho_1^2\mathsf{D}\frac{m}{n} \geq \frac{\rho_1^2\mathsf{D}}{n}\left(\frac{r}{T_0} - 1\right) \geq \frac{\rho_1^2\mathsf{D}}{2nT_0}r = \rho r. \quad \square$$

4.3. **Distance between rays in the curve complex.** We next establish a technical lemma which says that if we have a pair of geodesic segments from a common basepoint, we can "back up" from the endpoints to earlier points that admit distance estimates for both $d_{\mathcal{T}}$ and d_S.

Lemma 34 (Backing up to eliminate thin parts)**.** *There exist constants k_0 and M_0 with the following property. Suppose γ_1, γ_2 are a pair of geodesic rays based at x and let $y = \gamma_1(t_1)$ and $z = \gamma_2(t_2)$. Then there are times $0 \leq t_i' \leq t_i$ and corresponding points $y' = \gamma_1(t_1')$, $z' = \gamma_2(t_2')$ such that*

(i) *$d_S(y, y') \leq k_0\cdot d_S(y, z)$ and $d_S(z, z') \leq k_0\cdot d_S(y, z)$;*
(ii) *either $d_S(y', z') \leq 6$ or $d_Y(y', z') \leq M_0$ for all proper subsurfaces Y.*

We will call these the backup points *and* backup times *for the segments $[x, y], [x, z]$. Furthermore,*

(iii) *given any d, there are N_0 and c_0 such that if $d_{\mathcal{T}}(y, z) \geq N_0$ and $d_S(y, z) \leq d$, then the backup points satisfy*

$$d_{\mathcal{T}}(y, y') \geq c_0\cdot d_{\mathcal{T}}(y, z) \qquad \text{or} \qquad d_{\mathcal{T}}(z, z') \geq c_0\cdot d_{\mathcal{T}}(y, z)$$

Conclusion (i) says that the backup points are not much farther from the endpoints in the curve complex than the endpoints are from each other. The interpretation of (ii) is that in the distance formula (3), a significant contribution is made either by the whole curve complex distance or by projection distances to proper subsurface (but not both). These properties of backup points hold in general; (iii) says that if the Teichmüller distance between endpoints is long enough relative to their curve-complex distance, then on at least one side the distance backed up was significant. (Compare (i) and (iii).)

Proof. First, we will construct backup points satisfying (i) and (ii). Then we will verify (iii).

Let

$$\Omega = \{V \subsetneq S : d_V(y, z) \geq 10\mathsf{M}\}.$$

If $\Omega = \emptyset$, let $y' = y$ and $z' = z$ and we are done, since (i) is trivially satisfied and (ii) works with $M_0 = 10\mathsf{M}$. So assume Ω is nonempty.

For each $V \in \Omega$, there is a thin interval T_V along $[y, z]$. By the triangle inequality, for each $V \in \Omega$, either $d_V(x, y)$ or $d_V(x, z)$ is at least $5\mathsf{M}$ and so there is a nonempty thin interval T_V along at least one of $[x, y]$ or $[x, z]$, as in Figure 2.

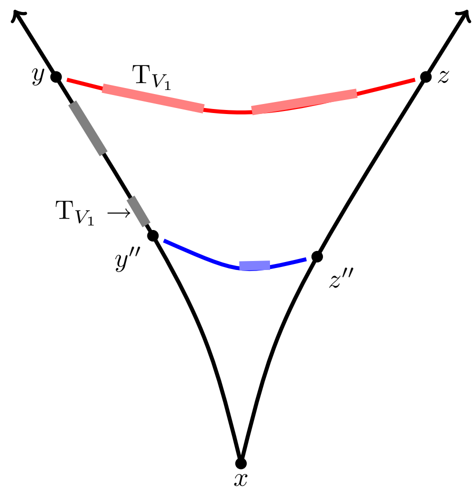

FIGURE 2. Two rays. Some thin intervals are shown, and two thin intervals for the same subsurface $V_1 \subset \Omega$ are marked.

Now let y'' and z'' be the earliest endpoints on γ_i of any T_V for $V \in \Omega$. That is, we move back in time until we have passed through each thin interval T_V for $V \in \Omega$. Let us use the notation V_1 and V_2 for the last surfaces in Ω whose thin interval one passes through in going from y to y'' and z to z'', respectively. Thus we have a short curve at y'' which is also short somewhere along $[y, z]$, which means that the image of y'' under the projection $\mathcal{T}(S) \to \mathcal{C}(S)$, which sends a surface to its Bers marking, shares a point with the image of $[y, z]$ in the curve complex. By the reverse triangle inequality, this means $d_S(y, y'')$ is bounded above relative to $d_S(y, z)$, as in (i).

Now with respect to the new points y'' and z'', let

$$\Omega'' = \{Y : d_Y(y'', z'') \geq 100\mathsf{M} + 2B'\},$$

where B' is the constant obtained by applying Theorem 13 (R.T.I. exception) with the threshold $50\mathsf{M}$. Note that $\Omega \cap \Omega'' = \emptyset$ because no $V \in \Omega$ has a thin interval along $[x, y'']$ or $[x, z'']$, so we would get a violation of the triangle inequality if V were also in Ω''. If $\Omega'' = \emptyset$, then by setting $y' = y''$, $z' = z''$, and letting $M_0 = 100\mathsf{M} + 2B'$, we are again done. So assume $\Omega'' \neq \emptyset$.

Fix some $Y \in \Omega''$. We know that either $d_Y(x, y'')$ or $d_Y(x, z'')$ is at least $50\mathsf{M}+B'$; let us suppose that the first of these is true. First assume that Y satisfies the reverse triangle inequality (Lemma 12). Then

$$d_Y(x, y) \geq 50\mathsf{M} + B' - B \geq 49\mathsf{M}.$$

Since $d_Y(y, z) < 10\mathsf{M}$, the triangle inequality then implies

$$d_Y(x, z) \geq d_Y(x, y) - 10\mathsf{M} \geq 39\mathsf{M}.$$

That is, for any $Y \in \Omega''$ that satisfies the reverse triangle inequality we have $d_Y(x, y), d_Y(x, z) \geq 39\mathsf{M}$.

If Y is an exception to the reverse triangle inequality, it must be an annulus, say with core curve α. Then by Theorem 13 we either have $d_Y(x, y) \geq d_y(x, y'') - B'$, or there exists W disjoint from α such that W satisfies the reverse triangle inequality along $[x, y]$ and $d_W(x, y'') \geq 50\mathsf{M}$. In the former case we have $d_Y(x, y) \geq 50\mathsf{M}$ so that $d_Y(x, z) \geq 40\mathsf{M}$ by the triangle inequality. In the latter case, we have $d_W(x, y) \geq d_W(x, y'') - B \geq 49\mathsf{M}$ by (4). Furthermore, since W must have a thin interval along $[x, y'']$, it cannot be in Ω and so we conclude $d_W(x, z) \geq 39$ as above.

In summary, to each $Y \in \Omega''$ we have associated a domain $W \notin \Omega$ (equal to Y except in the one case) such that both $d_W(x, y) \geq 39\mathsf{M}$ and $d_W(x, z) \geq 39\mathsf{M}$. Notice that we need only resort to the exceptional case $W \neq Y$ when $d_Y(y'', y) \geq \mathsf{M}$ (still assuming here that $d_Y(x, y'') \geq 50\mathsf{M} + B'$), for otherwise we may again conclude $d_Y(x, z) \geq 39$ by the triangle inequality. The inequalities $d_Y(x, y''), d_Y(y'', y) \geq M$ require that Y is thin along both $[x, y'']$ and $[y'', y]$; by the connectedness of T_Y, this implies that $\alpha = \partial Y$ is thin at y'' and consequently disjoint from ∂V_1. Therefore, since W is disjoint from α, we may safely assume that $d_S(\partial W, \partial V_1) \leq 2$ in the exceptional case $W \neq Y$.

We now make the following observation:

$$Z \notin \Omega, Z \pitchfork V_1 \text{ and } d_Z(x, y) \geq 39\mathsf{M} \implies \mathrm{T}_Z < \mathrm{T}_{V_1} \text{ along } \gamma_1, \tag{9}$$

and likewise for z, V_2, γ_2. We argue by contradiction. Otherwise $\mathrm{T}_{V_1} < \mathrm{T}_Z$ and so we have $d_Z(x, y'') \leq \mathsf{M}$ which implies $d_Z(y'', y) \geq 38\mathsf{M}$. Since $d_Z(y, z) < 10\mathsf{M}$ and $d_{V_1}(y, z) \geq 10\mathsf{M}$ this would violate Lemma 11 viewed from y. Namely along $[y, y'']$ we pass through T_Z and then T_{V_1} while along $[y, z]$ we pass through T_Z with a much smaller projection.

For each $Y \in \Omega''$, we now see that the associated subsurface $W \notin \Omega$ satisfies $d_S(\partial W, \partial V_i) \leq 2$ for either $i = 1$ or $i = 2$. We have already observed this when $W \neq Y$, and in the case that $W = Y$ we in fact have that Y is disjoint or nested with respect to either V_1 or V_2. Indeed, if $Y \pitchfork V_i$ for both $i = 1, 2$, then (9) would imply that $\mathrm{T}_Y \subset [x, y'']$ and $\mathrm{T}_Y \subset [x, z'']$. But then $d_Y(y'', y), d_Y(z'', z) \leq \mathsf{M}$ and since $d_Y(y, z) \leq 10\mathsf{M}$ we have contradicted $d_Y(y'', z'') \geq 50\mathsf{M}$.

We are ready to define the back up points y' and z'. Choose any $Y \in \Omega''$ (recall that we are assuming $\Omega'' \neq \emptyset$) and let W be the associated subsurface. If W is disjoint or nested with respect to V_1 then define $y' = y''$. Otherwise (9) implies that $\mathrm{T}_W < \mathrm{T}_{V_1}$ along γ_1; in this case we back up farther and define y' to be the beginning of T_W along $[x, y]$. Define the point z' similarly.

Since either V_1 or W is thin at y' and $d_S(\partial W, Z) \leq 2$ for some surface that is thin along $[y, z]$ (namely, V_1 or V_2), we conclude that $d_S(y, y')$ (and similarly $d_S(z, z')$) is bounded relative to $d_S(y, z)$, as in (i). If we have backed up farther on both sides, then W is thin at both y' and z' so that $d_S(y', z') \leq 4$ (recall that a Bers marking has diameter 2 in $\mathcal{C}(S)$). If we backed up on just one side, then we have a path of length 1 in the curve complex (either $\partial W - \partial V_2$ or $\partial V_1 - \partial W$) showing that $d_S(y', z') \leq 5$. Finally, if we didn't back up on either side, then the path $\partial V_1 - \partial W - \partial V_2$ of length 2 in the curve complex implies $d_S(y', z') \leq 6$. This verifies (ii).

We now consider (iii), so we are assuming that $d_S(y, z) \leq d$. The distance formula (3) says

$$d_{\mathcal{T}}(y, z) \leq K \left(d_S(y, z) + \sum_{\Omega} d_V(y, z) \right) + C \leq Kd + C + K \sum_{\Omega} d_V(y, z)$$

where K, C are the constants coming from threshold $M_0 = 10\mathsf{M}$. Let us write d_{12} for $d_{\mathcal{T}}(y, z)$. Take $N_0 \geq 2Kd + 2C$, so that the assumption of (iii) says that

$d_{12} \geq 2Kd + 2C$. This gives

$$\sum_{\Omega} d_V(y,z) \geq \frac{d_{12} - Kd - C}{K} \geq \frac{d_{12} - d_{12}/2}{K} = \frac{d_{12}}{2K}.$$

For each $V \in \Omega$, the definition of y' and z' ensures that $\mathrm{T}_V \cap [x,y'] = \mathrm{T}_V \cap [x,z'] = \emptyset$; therefore $d_V(y',z') \leq 2\mathsf{M}$. Therefore

$$d_V(y,z) \leq d_V(y,y') + d_V(z,z') + 2\mathsf{M}.$$

Furthermore, since $d_V(y,z) \geq 10\mathsf{M}$, it follows that

$$d_V(y,y') + d_V(z,z') \geq d_V(y,z) - 2\mathsf{M} \geq \frac{4}{5} d_V(y,z).$$

Notice that it cannot be the case that $d_V(y,y') \leq 3\mathsf{M}$ and $d_V(z,z') \leq 3\mathsf{M}$, for this would imply that $d_V(y,z) \leq 3\mathsf{M} + 3\mathsf{M} + 2\mathsf{M} < 10\mathsf{M}$, which is not the case. Therefore at least one of $d_V(y,y')$ or $d_V(z,z')$ is larger than $3\mathsf{M}$. We now have that

$$\begin{aligned}
[d_V(y,y')]_{3\mathsf{M}} + [d_V(z,z')]_{3\mathsf{M}} \geq & d_V(y,y') + d_V(z,z') - 3\mathsf{M} \\
\geq & \frac{4}{5} d_V(y,z) - 3\mathsf{M} \geq \left(\frac{4}{5} - \frac{3}{10}\right) d_V(y,z) = \frac{1}{2} d_V(y,z).
\end{aligned}$$

Let K', C' be the constants in (3) for the threshold $M_0 = 3\mathsf{M}$ and enlarge N_0 if necessary to ensure that $d_{12} \geq 16C'K'K$. The distance formula then gives

$$\begin{aligned}
d_{\mathcal{T}}(y,y') + d_{\mathcal{T}}(z,z') \geq & \frac{1}{K'}\left(d_S(y,y') + d_S(z,z') + \sum_{\Omega} [d_V(y,y')]_{3\mathsf{M}} + [d_V(z,z')]_{3\mathsf{M}}\right) - 2C' \\
\geq & \frac{1}{K'} \sum \frac{1}{2} d_V(y,z) - 2C' \geq \frac{d_{12}}{4K'K} - 2C' \\
\geq & \frac{d_{12}}{4K'K} - \frac{d_{12}}{8K'K} = \frac{d_{12}}{8K'K},
\end{aligned}$$

so we are done if we take $c_0 = \frac{1}{16K'K}$. ☐

Next we show that a pair of long geodesic segments which both stay far apart from each other in $\mathcal{T}(S)$ and spend a large fraction of their time in $\mathcal{T}_{\epsilon_0}$ must have big curve complex distance at some point. We will repeatedly use Lemma 34 to "back up" along each of the two rays in order to make distance estimates.

Theorem 35 (Big curve complex distance)**.** *For all $d > 6$, $T > 0$, and $0 < \delta < 1$, there exist two constants $D_0 \geq 0$, $R_0 \geq T$ with the following property. Let $\gamma_1 = [x,y]$ and $\gamma_2 = [x,z]$ be two Teichmüller geodesics based at x with lengths $r_1, r_2 \geq R_0$. Suppose that*

1. *the fraction of $[0, r_i]$ that γ_i spends in $\mathcal{T}_{\epsilon_0}$ is at least δ for each of $i = 1,2$; and*
2. *for all $t \geq T$, the point $\gamma_1(t)$ is not contained in the D_0–neighborhood of the geodesic $[x,z]$, and similarly for $\gamma_2(t)$ and $[x,y]$.*

Then $d_S(y,z) \geq d$.

Proof. Choose n such that $n\delta > 1$ and define $\delta' = \frac{n\delta - 1}{n-1}$; notice that $0 < \delta' < \delta$. Let ρ, R_1 and ρ', R_1' be the corresponding constants guaranteed by the Definite Progress lemma (Proposition 33). Set k_0, M_0, N_0, and c_0 to be the constants from the backing up lemma (Lemma 34) for our given d, and let $K \geq 1$ and $C \geq 0$ be large enough to be constants in the distance formula (3) for the threshold M_0. Let D_0 and R_0 be any constants which satisfy

$$D_0 > \max\left\{N_0,\ \frac{2k_0 d}{c_0\rho'},\ \frac{R_1'}{c_0},\ K(2k_0+1)d + C\right\},\ \text{and}$$
$$R_0 > \max\left\{nT,\ nR_1,\ \frac{2dn}{\rho}\right\}.$$

In order to derive a contradiction, assume furthermore that $d_S(y,z) < d$.

Set $y_0 = y$ and $z_0 = z$. By repeatedly backing up along γ_1 and γ_2, we will recursively define sequences of points $\{y_i\}, \{z_i\}$. Each step of the sequence will cover a large Teichmüller distance and a comparatively small curve complex distance. After backing up sufficiently far, we will eventually contradict the fact that each geodesic γ_i spends a large fraction of its time in $\mathcal{T}_{\epsilon_0}$.

Suppose that the points y_i, z_i have been defined and satisfy

$$(\star) \qquad d_S(y_i,z_i) < d; \qquad d_{\mathcal{T}}(y_i,z_i) \geq D_0; \qquad y_i, z_i \neq x.$$

(Notice that these conditions hold for the initial points y_0 and z_0.) Backing up along the rays as in Lemma 34, we then obtain new points $y_{i+1} \in [x, y_i]$ and $z_{i+1} \in [x, z_i]$. For these, we have that $d_S(y_{i+1}, y_i)$ and $d_S(z_{i+1}, z_i)$ are bounded above by $k_0 \cdot d_S(y_i,z_i) < k_0 d$, by property (i) of backup points. We also have that either $[y_{i+1}, y_i]$ or $[z_{i+1}, z_i]$ (or both) has length at least $c_0 \cdot d_{\mathcal{T}}(y_i,z_i) \geq c_0 D_0$, by (iii). We are free to continue defining new points in this manner as long as the conditions of $(\star)$ remain satisfied.

Suppose then that we have applied this procedure m times and arrived at points y_m and z_m. At each step of this process we traveled back a Teichmüller distance of at least $c_0 D_0$ along one of the two segments γ_1 or γ_2. Therefore on at least one of the geodesics we have traveled a total Teichmüller distance of at least $c_0 D_0 m/2$. Without loss of generality, suppose γ_1 has this property; then $d_{\mathcal{T}}(y_m, y_0) \geq c_0 D_0 m/2$. On the other hand we have $d_S(y_m, y_0) \leq k_0 dm$, since at each step we travel at most $k_0 d$ in the curve complex. Therefore, along the geodesic segment $[y_m, y_0]$ the ratio of curve-complex distance to Teichmüller distance is

$$\frac{d_S(y_m,y_0)}{d_{\mathcal{T}}(y_m,y_0)} \leq \frac{2k_0 dm}{c_0 D_0 m} < \rho'.$$

Proposition 33 now implies that the fraction of $[y_m, y_0]$ spent in $\mathcal{T}_{\epsilon_0}$ is strictly less than δ'.

Let $t_m \in [0, r_1]$ be the time for which $\gamma_1(t_m) = y_m$. We claim that if $t_m \geq r_1/n$, then the points y_m, z_m satisfy $(\star)$ so that we may reapply Lemma 34 and back up farther to points y_{m+1} and z_{m+1}. Firstly, since $t_m \geq r_1/n \geq R_0/n \geq T$, hypothesis (ii) in the theorem implies that $y_m = \gamma_1(t_m)$ is not within D_0 of any point on γ_2; whence $d_{\mathcal{T}}(y_m, z_m) \geq D_0$. Applying the triangle inequality to the series of points $y_m, y_{m-1}, z_{m-1}, z_m$ implies that $d_S(y_m, z_m) \leq (2k_0+1)d$. Now if $d_Y(y_m . z_m) \leq M_0$ for all proper subsurfaces Y, then the distance formula (3) gives

$$d_{\mathcal{T}}(y_m,z_m) \leq K \cdot d_S(y_m,z_m) + C \leq K(2k_0+1)d + C < D_0,$$

which is a contradiction. Thus by Lemma 34 we have $d_S(y_m, z_m) \leq 6 \leq d$. Finally, since the fraction of $[t_m, r_1]$ that γ_1 spends in $\mathcal{T}_{\epsilon_0}$ is at most $\delta' < \delta$, it must be that the thick fraction along $[0, t_m]$ is at least δ (since the thick fraction of the whole γ_1 is δ). By Proposition 33, since $t_m \geq R_0/n \geq R_1$, it now follows that

$d_S(y_m, x) \geq \rho t_m \geq \rho R_0/n \geq 2d$. As $d_S(y_m, z_m) \leq d$, we see that $z_m \neq x$. Thus we have verified the conditions of $(\star)$.

We have thus verified that we can repeatedly back up until we reach a point $y_m = \gamma_1(t)$ on $[x, y]$ (or similarly a point on $[x, z]$) such that $t < r_1/n$ and the fraction of $[t, r_1]$ that γ_1 spends in $\mathcal{T}_{\epsilon_0}$ is strictly less than δ'. It now follows that the amount of time that γ_1 spends in $\mathcal{T}_{\epsilon_0}$ along $[0, r_1]$ is strictly less than

$$t + \delta'(r_1 - t) < \frac{r_1}{n} + \frac{n\delta - 1}{n-1}\left(\frac{n-1}{n} r_1\right) = \delta r_1,$$

again by shifting weight as in (8). This contradicts the first hypothesis of this theorem. $\square$

4.4. **Teichmüller distance.** The previous theorem implies that, under suitable conditions, we may assume our two Teichmüller geodesics stay far apart in the curve complex beyond some radius R_0. We now show that in this situation, the distance between two points on the sphere of radius $r \gg R_0$ is on the order of $2r$.

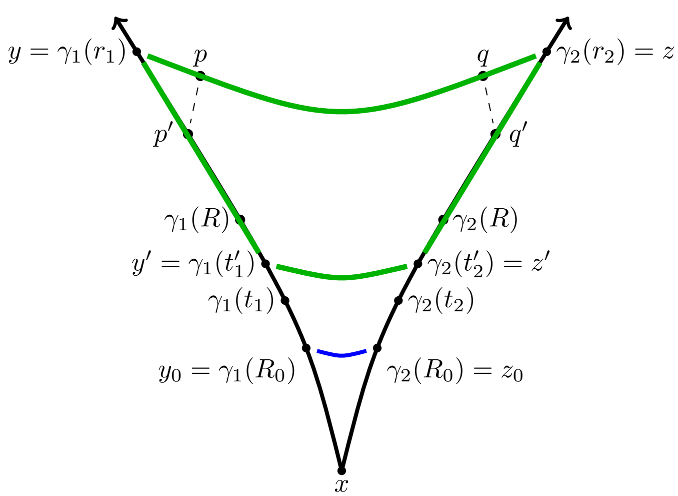

FIGURE 3. We will end up showing that $[y, z]$ "dips back" to within bounded distance of the basepoint and thus has length nearly $2r$.

Theorem 36 (Teichmüller distance estimate)**.** *Fix any $0 < \delta < 1$. Then for all sufficiently large R_0, d there exists a constant H as follows. For any geodesic rays γ_1, γ_2 based at $x \in \mathcal{T}(S)$ which satisfy*

(a) *for all $r \geq R_0$, the fraction of $[0, r]$ that γ_i spends in $\mathcal{T}_{\epsilon_0}$ is at least δ, and*
(b) *for all $t_1, t_2 \geq R_0$ we have $d_S(\gamma_1(t_1), \gamma_2(t_2)) \geq d$,*

they also must satisfy

$$d_{\mathcal{T}}(\gamma_1(r_1), \gamma_2(r_2)) \geq r_1 + r_2 - H$$

for all r_1, r_2.

Proof. We fix some constants.

- Let $K \geq 1$ and $C \geq 0$ be large enough to be quasi-isometry constants in the distance formula (3) for threshold $10\mathsf{M}$, and such that the projection of every geodesic in $\mathcal{T}(S)$ to $\mathcal{C}(S)$, concatenated with a geodesic segment of length 5, is an unparameterized (K, C)-quasigeodesic in $\mathcal{C}(S)$;
- Let $d > 10+\tau$, where τ is a constant such that (K, C)-quasigeodesic quadrilaterals in the curve complex are τ–thin;
- Let ρ, R_1 be the constants corresponding to δ from Lemma 33;
- Take $R_0 > \max(R_1, 4 + C)$.

Now write $y = \gamma_1(r_1)$, $y_0 = \gamma_1(R_0)$, $z = \gamma_2(r_2)$, $z_0 = \gamma_2(R_0)$, and let

$$\Omega = \{V \subsetneq S \mid \mathrm{T}_V \cap [y_0, z_0] \neq \emptyset\}$$

be the set of subsurfaces which are thin somewhere along the geodesic segment $[y_0, z_0]$. For each $V \in \Omega$ it is possible that it determines a nonempty thin interval T_V which intersects $\gamma_i[R_0, \infty)$. We define

$$t_i = \sup \bigcup_{V\in\Omega} \mathrm{T}_V \cap \gamma_i[R_0, \infty).$$

In other words, $t_i \in [R_0, \infty)$ is the smallest time such that for any $V \in \Omega$ and for $s > t_i$, $\gamma_i(s) \notin \mathrm{T}_V$. It follows that for each $i = 1, 2$, there exists some $V_i \in \Omega$ such that the Bers marking at $\gamma_i(t_i)$ contains ∂V_i.

By definition of $V_i \in \Omega$, there exists a point $w_i \in [y_0, z_0]$ such that V_i is thin at w_i; in particular $d_S(w_i, \gamma_i(t_i)) \leq 4$. Since w_i is at most R_0 away from an endpoint of $[y_0, z_0]$, the triangle inequality implies that $d_{\mathcal{T}}(x, w_i) \leq 2R_0$. We then have that

$$d_S(x, \gamma_i(t_i)) \leq 4 + d_S(x, w_i) \leq 4 + 2KR_0 + C < 3KR_0,$$

where the middle inequality is an application of the distance formula (3) and last inequality holds since $R_0 > 4 + C$. Furthermore, since $t_i \geq R_0$, our hypotheses imply that γ_i spends at least δt_i time during $[0, t_i]$ in $\mathcal{T}_{\epsilon_0}$. Since $R_0 \geq R_1$, by Lemma 33 we have that

$$t_i \leq \frac{1}{\rho} d_S(x, \gamma_i(t_i)) < \frac{3K}{\rho} R_0.$$

Let $R = \frac{3KR_0}{\rho\delta}$, and note that $R > t_1, t_2$. If the geodesic segment $\gamma_i[t_i, R]$ were completely contained in the thin part, then the fraction of $[0, R]$ that γ_i spends in $\mathcal{T}_{\epsilon_0}$ would be at most

$$\frac{t_i}{R} < \frac{3KR_0}{\rho R} = \delta.$$

As this is not the case, there must exist a time $t_i' \in [t_i, R]$ at which $\gamma_i(t_i') \in \mathcal{T}_{\epsilon_0}$. Let y' and z' equal $\gamma_i(t_i')$ for $i = 1, 2$, respectively. Note that in particular

$$d_{\mathcal{T}}(x, y'), d_{\mathcal{T}}(x, z') \leq \frac{3KR_0}{\rho\delta}. \tag{10}$$

By the above estimate, the theorem is trivially true if either $r_i \leq t_i'$ if we choose $H \geq \frac{6KR_0}{\rho\delta}$, so it is enough to prove the theorem for both $r_i > t_i'$. That is, $[x, y'] \subset [x, y]$ and $[x, z'] \subset [x, z]$. Now let

$$\Omega_1 = \{W \subsetneq S \mid d_W(y', y) \geq 10\mathsf{M}\}, \qquad \Omega_2 = \{W \subsetneq S \mid d_W(z', z) \geq 10\mathsf{M}\}$$

Note that a marking projects to the curve complex as a set with diameter 2. Also if $W_1 \in \Omega_1$ and $W_2 \in \Omega_2$ were disjoint or nested, then their boundaries would have distance ≤ 1 in the curve complex, so markings where those boundaries are short would have distance ≤ 5. Thus in light of assumption (b) and since $d > 10$, we have that $W_1 \pitchfork W_2$ for all $W_1 \in \Omega_1$ and $W_2 \in \Omega_2$. The same argument shows that $W_2 \in \Omega_2$ cannot determine an thin interval along $[y', y]$ and similarly for W_1 along $[z', z]$. In particular $\Omega_1 \cap \Omega_2 = \emptyset$.

Suppose V is any subsurface which is thin somewhere along $[y', y]$ (for example, if $V \in \Omega_1$). Since its thin interval T_V along $[x, y]$ is connected and y' is thick, we see that V is not thin along $[y_0, y']$. Furthermore, the definition of t_i implies that V is not thin along $[y_0, z_0]$, and assumption (b) plus the choice of d shows that V cannot be thin along $[z_0, z']$. Therefore, the triangle inequality gives $d_V(y', z') \leq 3\mathsf{M}$ for these V. Letting

$$\Omega' = \{Z \subsetneq S \mid d_Z(y', z') \geq 10\mathsf{M}\},$$

we conclude that each surface $Z \in \Omega'$ cannot be thin along $[y', y]$ (or $[z', z]$). In particular, the three collections Ω_1, Ω_2, and Ω' are pairwise disjoint.

Next we establish that for any $W \in \Omega_1 \cup \Omega_2 \cup \Omega'$,

$$d_W(y, z) \geq 6\mathsf{M}.$$

This is because we know that for each $W \in \Omega_1$, W is not thin along $[z', z]$ and that $d_W(y', z') \leq 3\mathsf{M}$ as above. Since $d_W(y', y) \geq 10\mathsf{M}$, the triangle inequality gives $d_W(y, z) \geq 10\mathsf{M} - 3\mathsf{M} - \mathsf{M} = 6\mathsf{M}$. The same argument applies to Ω_2. Similarly, for each $Z \in \Omega'$, we have both $d_Z(y', z') \geq 10\mathsf{M}$ and $d_Z(y', y), d_Z(z', z) \leq \mathsf{M}$, so that $d_Z(y, z) \geq 8\mathsf{M}$. We conclude that each $W \in \Omega_1 \cup \Omega_2 \cup \Omega'$ determines a nonempty thin interval T_W along $[y, z]$.

Let $p \in [y, z]$ denote the last point (in traveling from y to z) at which any subsurface in Ω_1 is thin and let $W_1 \in \Omega_1$ denote the corresponding subsurface. Analogously define q and $W_2 \in \Omega_2$ on $[z, y]$; p' and $Z_1 \in \Omega_1$ on $[y, y']$; and q' and $Z_2 \in \Omega_2$ on $[z, z']$. (If $\Omega_i = \emptyset$ then we take the initial endpoints of the respective intervals and leave the associated subsurfaces undefined.) Thus the Bers markings at p, q, p', q' contain $\partial W_1, \partial W_2, \partial Z_1, \partial Z_2$, respectively.

It cannot be the case that $W_1 \pitchfork Z_1$, as this would violate the time-order principle (W_1 occurs after Z_1 on $[y, z]$ but before Z_1 on $[y, y']$). Therefore ∂W_1 and ∂Z_1 either coincide or are disjoint; in either case we have $d_S(p, p') \leq 5$. The same argument shows that $d_S(q, q') \leq 5$.

Next we want to verify that p occurs before q along $[y, z]$. This is another application of time ordering: suppose for contradiction that T_{W_2} appeared before $\mathcal{T}_{W_1}$ along $[y, z]$. (Recall that $W_1 \pitchfork W_2$, since this was verified for all subsurfaces from Ω_1 and Ω_2 above.) Then since $d_{W_1}(y, z), d_{W_2}(y, z), d_{W_1}(y', y)$ are all more than $3\mathsf{M}$, Lemma 11 ensures that T_{W_2} also appears along $[y, y']$. But this contradicts the fact that no surface in Ω_2 has a thin interval along $[y, y']$.

The setup is now complete, and we are ready to prove the following:

Claim 37. *There is a uniform bound (independent of r_1, r_2) on the minimal distance between the segments $[y, z]$ and $[y', z']$.*

It is easy to see that the theorem follows from this claim. Suppose that the geodesic segment $[y, z]$ comes within distance N of the geodesic $[y', z']$. As this latter segment is contained in the ball of radius $2R = \frac{6K}{\rho\delta} R_0$ around x, this means that there must be some point w on $[y, z]$ whose distance from x satisfies $d_{\mathcal{T}}(w, x) \leq \frac{6K}{\rho\delta} R_0 + N$. Since y and z are on the spheres of radius r_1, r_2 centered at x, this proves that

$$d_{\mathcal{T}}(y,z) = d_{\mathcal{T}}(y,w) + d_{\mathcal{T}}(w,z) \geq r_1 + r_2 - 2\left(\frac{6K}{\rho\delta}R_0 + N\right),$$

establishing the theorem with $H = 2\left(\frac{6K}{\rho\delta}R_0 + N\right)$.

Proof of claim. Consider any two points $w' \in [y', z']$ and $w \in [p, q]$. Aiming to use the distance formula (3), we estimate the projections $d_V(w, w')$ for all proper subsurfaces $V \subsetneq S$.

First any subsurface $W \in \Omega_1$ is not thin along $[z', z]$ or $[w, z]$. Therefore

$$\begin{aligned} d_W(w', w) &\leq d_W(w', z') + d_W(z', z) + d_W(z, w) \\ &\leq K d_{\mathcal{T}}(y', z') + C + \mathsf{M} + \mathsf{M} \\ &= K d_{\mathcal{T}}(y', z') + 2\mathsf{M} + C, \end{aligned}$$

We have the same bound for $W \in \Omega_2$ using y' and y instead of z' and z.

Lastly, consider a proper subsurface $V \notin \Omega_1 \cup \Omega_2$ (such as $V \in \Omega'$). We have

$$\begin{aligned} d_V(w', w) &\leq d_V(w', y') + d_V(y', y) + d_V(y, w) \\ &\leq d_V(y', w') + 10\mathsf{M} + d_V(y, w) \end{aligned}$$

If V satisfies Lemma 12, then $d_V(y, w) \leq d_V(y, z) + B$ and so we have

$$\begin{aligned} d_V(w', w) &\leq d_V(y', w') + 10\mathsf{M} + d_V(y, z) + B \\ &\leq d_V(y', w') + d_V(y', z') + 30\mathsf{M} + B \\ &\leq 2K d_{\mathcal{T}}(y', z') + 2C + 30\mathsf{M} + B. \end{aligned}$$

If V does not satisfy Lemma 12 it must be an annulus. Applying Theorem 13 with constant $35\mathsf{M}$, there exists B' such that either $d_V(y, w) \leq d_V(y, z) + B'$, in which case we have $d_V(w', w) \leq 2K d_{\mathcal{T}}(y', z') + 2C + 30\mathsf{M} + B'$ as above, or there is a collection $\{U_j\}$ of subsurfaces that do satisfy Lemma 12 along $[y, z]$, are disjoint from V, and satisfy

$$d_{U_j}(y, w) \geq 35\mathsf{M} \quad \text{and} \quad d_V(y, w) \leq \sum_{U_j} d_{U_j}(y, w).$$

Together with Lemma 12, the first of these inequalities implies $d_{U_j}(y, z) \geq 34\mathsf{M}$ so that each U_j has a large projection along both $[y, w]$ and $[y, z]$. The only such subsurfaces are contained in $\Omega_1 \cup \Omega'$. Running the argument with y' and y replaced by z' and z, we obtain another such collection contained in $\Omega_2 \cup \Omega'$. Since all of these subsurfaces are disjoint from V, the fact that elements of Ω_1 and Ω_2 are far apart in $\mathcal{C}(S)$ implies that one of these collections must, in fact, be contained in Ω'. Therefore we may assume that $\{U_j\} \subset \Omega'$. We then have the bound

$$d_{U_j}(y, w) \leq d_{U_j}(z, y) + B \leq d_{U_j}(y', z') + 3\mathsf{M} \leq 5 d_{U_j}(y', z')$$

and so

$$d_V(y, w) \leq \sum_{U_j} d_{U_j}(y, w) \leq 5 \sum_{Y \subsetneq S} [d_Y(y', z')]_{10\mathsf{M}} \leq 5K d_{\mathcal{T}}(y', z') + 5C.$$

In this case we conclude

$$d_V(w', w) \leq 6K d_{\mathcal{T}}(y', z') + 6C + 10\mathsf{M}.$$

Putting these estimates together and taking a maximum we conclude that for any proper subsurface V we have a bound

$$d_V(w', w) \leq 6K d_{\mathcal{T}}(y', z') + 6C + 31\mathsf{M} + B'.$$

Set

$$M_1 := 6K d_{\mathcal{T}}(y', z') + 6C + 32\mathsf{M} + B' \leq \frac{36K^2}{\rho\delta} R_0 + 6C + 32\mathsf{M} + B'.$$

We now apply the distance formula (3) with this threshold to conclude that for some constants K' and C' that just depend on M_1 (hence on fixed constants),

$$d_{\mathcal{T}}(w', w) \leq K' d_S(w', w) + C' \tag{11}$$

for all points $w' \in [y', z']$ and $w \in [p, q]$.

Thus we have bounded the Teichmüller distance across the green quadrilateral in Figure 3 by the corresponding curve complex distance; now we will use the fact that quasigeodesic quadrilaterals in the curve complex are thin. Let π be the projection map to the curve complex. On the left, we let γ_L be the concatenation of the quasi-geodesic segment $\pi[y', p']$ with the geodesic $[\pi(p'), \pi(p)]$; likewise, on the right, γ_R is the concatenation of $\pi[z', q']$ and $[\pi(q'), \pi(q)]$. These are actually (K, C)-quasigeodesics because the second segment has bounded length: $d_S(p, p'), d_S(q, q') \leq 5$ (recall the definition of K and C). Thus the quasi-geodesic segments γ_L, $\pi[p, q]$, γ_R, and $\pi[y', z']$ form a (K, C)-quasigeodesic quadrilateral, and we conclude that each side is contained in the τ-neighborhood of the union of the other three sides.

Now, the separation hypothesis (b) of the theorem implies that no point on $\pi[y', p']$ is within d of any point on $\pi[z', q']$. Therefore, no point on γ_L is within $(d - 10)$ of any point on γ_R. Recall that $d - 10 > \tau$, which implies that no point on γ_L is contained in the τ–neighborhood of γ_R. This means that there exist points on $\pi[p, q]$ and $\pi[y', z']$ that are within $2\tau + 1$ of each other (since each point on γ_L, say, is within τ of one or the other, and the points of γ_L are separated by one). By definition of projection distance, these correspond to points $w \in [p, q], w' \in [y', z']$ with $d_S(w, w') \leq 2\tau + 1$. Combining this with (11) we find that

$$d_{\mathcal{T}}(w, w') \leq K'(2\tau + 1) + C'.$$

This completes the proof of Claim 37 and Theorem 36. □

5. Statistical hyperbolicity

We can now assemble the results from §4 to prove Theorems 2 and 1.

Theorem 2. *For any basepoint $x \in \mathcal{T}(S)$, and either of the normalized standard visual measures $\mu_x = \mathrm{Vis}_r(\nu_x)$ or $\mathrm{Vis}_r(s_x)$ on $\mathcal{S}_r(x)$, we have*

$$\lim_{r_1, r_2 \to \infty} \tfrac{1}{r_1 + r_2} \int_{\mathcal{S}_{r_1}(x) \times \mathcal{S}_{r_2}(x)} d_{\mathcal{T}}(y, z) \quad d\mu_x(y) d\mu_x(z) = 1.$$

In particular, taking $r_1 = r_2 = r$, we get

$$E(\mathcal{T}(S), \mu_x) = 2.$$

Proof. Fix any $\epsilon > 0$. We have shown (Proposition 24) that except for a subset of $\mathcal{S}_r(x)$ of μ_x–measure at most ϵ, all long enough geodesics spend a definite fraction δ of time in the thick part. We can use this to choose $R_0 = R_0(\epsilon)$ large enough so that all but measure ϵ of *pairs* of rays both stay thick for fraction δ once they are longer than R_0. We just established (Theorem 36) that if rays are thick for fraction δ and if they stay d-separated in curve complex distance, then the Teichmüller geodesic between the points at distance r_1, r_2 has length at least $r_1 + r_2 - H$ (where the constant H depends on the threshold R_0 which in turn depends on ϵ). But curve complex separation of d is ensured for our geodesics as long as they remain separated by D_0 in Teichmüller distance (Theorem 35). And that in turn is guaranteed after throwing out another set of measure ϵ (and possibly increasing R_0) because of Proposition 23.

We conclude that except for a set of measure at most 2ϵ, every pair $y \in \mathcal{S}_{r_1}(x), z \in \mathcal{S}_{r_2}(x)$ satisfies

$$d_{\mathcal{T}}(y,z) \geq r_1 + r_2 - H.$$

Therefore

$$\liminf_{r_1,r_2\to\infty} \tfrac{1}{r_1+r_2} \int d_{\mathcal{T}}(y,z)\; d\mu_x(y) d\mu_x(z) \geq \liminf_{r_1,r_2\to\infty} (1-2\epsilon) \tfrac{1}{r_1+r_2}(r_1+r_2-H) = 1-2\epsilon.$$

But the limsup is also ≤ 1, by the triangle inequality. Since ϵ was arbitrary, we have shown that the limit exists and is 1. In the particular case that $r_1 = r_2 = r$ we conclude that $E(\mathcal{T}(S)) = 2$. $\square$

Theorem 1. *Fix a basepoint x and let μ refer to either of the standard visual measures* $\mathrm{Vis}(\nu_x)$ *or* $\mathrm{Vis}(s_x)$*, to the holonomy measure* $\mathbf{m}$*, or to the Hausdorff measure η on $\mathcal{T}(S)$. Then*

$$\lim_{r\to\infty} \frac{1}{r} \frac{1}{\mu(\mathcal{B}_r(x))^2} \int_{\mathcal{B}_r(x)\times\mathcal{B}_r(x)} d_{\mathcal{T}}(y,z)\; d\mu(y)d\mu(z) = 2.$$

Proof. Fix any $0 < \rho < 1$. Since the volume of balls for any of these measures grows exponentially with the radius, almost all of the measure of $\mathcal{B}_r(x)$ is contained in the shell $\mathcal{B}_r(x) \setminus \mathcal{B}_{\rho r}(x)$ as $r \to \infty$; furthermore, ρr is eventually larger than any fixed threshold R_0. By Theorem 2, the average distance between points in $\mathcal{B}_r(x) \setminus \mathcal{B}_{\rho r}(x)$ is at least $2\rho r$, and since ρ can be taken arbitrarily close to 1 we are done. $\square$

References

[1] J.C. Álvarez Paiva, A. Thompson, *Volumes on normed spaces and Finsler spaces*, MSRI Publications 50 (2004).

[2] J. Athreya, A. Bufetov, A. Eskin, and M. Mirzakhani, *Lattice point asymptotics and volume growth on Teichmüller space*, arxiv 0610715

[3] L. A. Santaló, *An affine invariant for convex bodies of n-dimensional space*, Portugaliae Math. 8 (1949), 155–161.

[4] H. Busemann, *Intrinsic area*, Ann. of Math. 48 (1947), 234-267.

[5] M. Duchin, S. Lelièvre, and C. Mooney, *Statistical hyperbolicity in groups*, preprint.

[6] A. Eskin and M. Mirzakhani, *Counting closed geodesics in moduli space*, arxiv 08112362

[7] A. Fathi, F. Laudenbach, P. Poenaru, *Travaux de Thurston sur les surfaces*, Asterisque 66-67 (1979).

[8] F. John, *Extremum problems with inequalities as subsidiary conditions*, 187–204 in Studies and essays presented to R. Courant on his 60th birthday, edited by K. O. Friedrichs, O. Neuegebauer, and J. J. Stoker, Interscience, New York, 1948.

[9] H. Masur, *On a class of geodesics in Teichmüller space*, Ann. of Math. (2) 102 (1975), no. 2, 205–221.

[10] H. Masur, *Uniquely ergodic quadratic differentials*, Comment. Math. Helvetici 55 255–266 (1980).

[11] H. Masur, *Interval exchange transformations and measured foliations*, Annals of Math. 115 (1982) 169–200.

[12] H. Masur, *Hausdorff dimension of the set of nonergodic foliations of a quadratic differential*, Duke Math. Journal 66 387–442 (1992).

[13] H. Masur, *The Teichmüller flow is Hamiltonian*, Proceedings of the American Mathematical Society 123 (1995) **12** 3739–3747.

[14] H. Masur, and Y. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. 138 (1999), no. 1, 103–149.

[15] H. Masur and M. Wolf, *Teichmüller space is not Gromov hyperbolic*, Ann. Acad. Sci. Fenn. Ser. A I Math., 20(2) (1995), 259–267.

[16] Y.N. Minsky, *Extremal length estimates and product regions in Teichmüller space*, Duke Math. J. 83:2(1996), 249–286.

[17] K. Rafi, *A combinatorial model for the Teichmüller metric*, Geom. Funct. Anal. 17:3 (2007), 936–959.

[18] K. Rafi, *Hyperbolicity in Teichmüller space*, preprint.

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

[20] H. L. Royden, *Automorphisms and isometries of Teichmüller space*, Advances in the Theory of Riemann Surfaces (Proc. Conf., Stony Brook, N.Y., 1969), Ann. of Math. Studies, No. 66. Princeton Univ. Press, Princeton, N.J., (1971), 369–383.

Spencer Dowdall, Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
E-mail address: dowdall@math.uiuc.edu

Moon Duchin, Mathematics, Tufts University, Medford, MA 02155
E-mail address: moon.duchin@tufts.edu

Howard Masur, Mathematics, University of Chicago, Chicago, IL 60637
E-mail address: masur@math.uchicago.edu