STATISTICAL HYPERBOLICITY IN TEICHMÜLLER SPACE

SPENCER DOWDALL, MOON DUCHIN, AND HOWARD MASUR

Abstract. In this paper we explore the idea that Teichmüller space is hyperbolic “on average.” Our approach focuses on studying the geometry of geodesics which spend a definite proportion of time in the thick part of Teichmüller space, and we establish a statistical version of the thin triangle property that holds in this setting. We consider several different measures on Teichmüller space and find that this behavior for geodesics is indeed typical. With respect to each of these measures, we moreover show that 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 metric properties of Teichmüller space \( \mathcal{T}(S) \), which is the parameter space for several types of geometric structures on \( S \). Equipped with the Teichmüller metric \( d_T \), it is a complete metric space homeomorphic to \( \mathbb{R}^{6g-6} \). It is not \( \delta \)-hyperbolic [18], 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 [12], and there are large “thin parts” of the space which, up to bounded additive error, are isometric to product spaces equipped with sup metrics (and therefore rule out hyperbolicity in the space as a whole) [19]. However, these exceptions to negative curvature seem to come from rare occurrences, while a long list of properties associated with hyperbolicity do hold globally or in specialized situations. Thus one may expect such properties to hold generically or “on average.” This paper aims to show that this is indeed the case.

Our motivating goal is to understand the generic geometry of Teichmüller space, particularly with regard to negative-curvature phenomena. For example, geodesics that stay in the thick part of \( \mathcal{T}(S) \) are well understood and exhibit many properties characteristic of hyperbolicity. Geodesics lying completely in the thin part are also well understood, but these exhibit no negative curvature characteristics. However, much more typical is for a geodesic to spend time in both the thick and thin parts of \( \mathcal{T}(S) \)—indeed, a generic geodesic ray will switch between these parts infinitely often. In this paper we develop tools to study these types of geodesics, and we discover that certain negative curvature phenomena do hold in this setting. For example, we obtain the following variant of thin triangles.

**Theorem A.** For any \( \epsilon > 0 \) and \( 0 < \theta \leq 1 \), there exist constants \( C, L \) such that if \( I \subset [x, y] \subset \mathcal{T}(S) \) is a geodesic subinterval of length at least \( L \) and at least
proportion $\theta$ of $I$ is $\epsilon$–thick, then for all $z \in \mathcal{T}(S)$, we have

$$I \cap \text{Nbd}_{C}(\{[x, z] \cup [y, z]\}) \neq \emptyset.$$  

(Note that Minsky’s product regions theorem [19] shows that some condition on thickness is necessary.) This theorem is a generalization of Rafi’s result, Theorem 3.12 below, in which he obtains the same conclusion under the stronger hypothesis that the entire interval $I$ is thick.

As a consequence (Corollary 3.14 below), we can assert for instance that if the three sides of a triangle each spend half of their time in the $\epsilon$–thick part, then (almost) half of each side is within a uniformly bounded distance of the union of the other two sides.

Having established that geodesics spending a definite fraction of time in the thick part exhibit hyperbolic characteristics, we go on to verify that such geodesics are in fact generic. To this end, we investigate a number of a priori different measures on Teichmüller space which are natural from various points of view. For instance, as a metric space $\mathcal{T}(S)$ carries a $(6g - 6)$–dimensional Hausdorff measure $\eta$. Other measures come from the Finsler structure (Busemann measure $\mu_{B}$ and Holmes–Thompson measure $\mu_{HT}$), from the holonomy coordinates on the cotangent bundle (holonomy, or Masur–Veech, measure $\mathbf{m}$), and from the symplectic structure. In §4, we find that all of these measures are mutually absolutely continuous and in fact are related by explicit inequalities. Still more interesting measures are provided by the identification of the metric $r$–sphere $S_{r}(x)$ with the unit sphere $Q^{1}(x)$ in the vector space of quadratic differentials on $x$ via the Teichmüller map. The latter has various natural measures, and corresponding measures on $S_{r}(x)$ will be called visual measures; we will pay special attention to two standard visual measures, denoted $\text{Vis}_{r}(\nu_{x})$ and $\text{Vis}_{r}(s_{x})$. We will also use the term visual measures for the induced measures on $\mathcal{T}(S)$, denoted $\text{Vis}(\nu_{x})$ and $\text{Vis}(s_{x})$, obtained by integrating radially.

Using Theorem A we can compute a statistic built by combining a metric and a measure to quantify how fast a space “spreads out.” Suppose we are given a family of probability measures $\mu_{r}$ on the spheres $S_{r}(x)$ of a metric space $(X, d)$. Then let $E(X) = E(X, x, \bar{d}, \{\mu_{r}\})$ be the average normalized distance between points on large spheres:

$$E(X) := \lim_{r \to \infty} \frac{1}{\int_{S_{r}(x) \times S_{r}(x)} d(y, z) \ d\mu_{r}(y) d\mu_{r}(z)},$$

if the limit exists. This creates a numerical index varying from 0 (least spread out) to 2 (most spread out). It is shown in [6] 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 $E(\mathbb{R}^{n}) < \sqrt{2}$ for all $n$, and that $E(\mathbb{Z}^{n}, S) < 2$ for all $n$ and $S$, with nontrivial dependence on $S$. (See [6] for more examples.) Motivated by these findings, we may regard a measured metric space with $E = 2$ as being “statistically hyperbolic.”

We note that the finding that $E = 2$ for hyperbolic groups makes use of homogeneity. In contrast, it is easy to build (highly non-regular) locally finite trees, equipped with counting measure on spheres, for which $E$ obtains any value from 0 to 2; see [6, p.4]. Thus neither $\delta$–hyperbolicity nor exponential growth is sufficient to ensure $E = 2$. Indeed, since the measures are normalized, the growth rate of the space has no direct effect on $E$. As an illustration, note that the Euclidean
plane could be endowed with a visual measure, constructed just like the ones we
study below in §4.3, which would give it exponential volume growth while leaving
$E = 4/\pi$ unchanged. On the other hand, other measures on $\mathbb{R}^2$ would give different
values of $E$; the statistic is quite sensitive to the choice of measure.

The following theorem concerns the average distance between points in the ball
$B_r(x)$ of radius $r$ centered at $x$. We show that this average distance is asymptotic
to $2r$, which, in light of the triangle inequality, is the maximum possible distance.

**Theorem B.** Let $\mu$ denote the Hausdorff measure $\eta$, holonomy measure $m$, or
either standard visual measure $\text{Vis}(\nu_x)$ or $\text{Vis}(s_x)$. Then for every point $x \in \mathcal{T}(S)$,

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

Of course, by the remarks above, this also holds for all the other measures
discussed in the paper. Indeed, we will work with properties of measures on $\mathcal{T}(S)$
that suffice to guarantee this conclusion: a **thickness property** (P1) defined in §5.2
guaranteeing that typical rays spend a definite proportion of their time in the thick
part, and a **separation property** (P2) defined in §6 asserting that typical pairs of
rays will exceed any definite amount of separation. In some places we use a stronger
separation property (P3) which is a quantified version with an exponential bound.

With respect to the standard visual measures, the same methods produce the
following result.

**Theorem C.** For every point $x \in \mathcal{T}(S)$ and either family $\{\mu_r\}$ of standard visual
measures $\mu_r = \text{Vis}_r(\nu_x)$ or $\text{Vis}_r(s_x)$ on the spheres $\mathcal{S}_r(x)$, we have

$$
E(\mathcal{T}(S), x, d\mathcal{T}, \{\mu_r\}) = 2.
$$

We sketch here the main ideas in the proofs of the theorems. Theorem A
is put together with distance estimates coming from subsurface projections, us-
ing reverse triangle inequalities (following Masur–Minsky and Rafi), and antichain
bounds (Rafi and Schleimer). A crucial ingredient is to show that geodesics spend-
ing a definite proportion of time in the thick part have shadows that make definite
progress in the curve complex (Theorem 3.9). We note that an alternate proof of
this could be given using work of Hamenstädt instead (see Remark 3.11).

The idea for Theorem B is that the separation property (P2) ensures that most
pairs of geodesics will have stopped fellow-traveling in the Teichmüller metric by 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.

One obstruction, as alluded to above, is that the pair of geodesics can enter thin
parts corresponding to disjoint subsurfaces, making the length of the third side
smaller than the sum, as on the right in Figure 1. However Theorem A says that
if one side satisfies the thickness proportion for an interval close to the basepoint,
then this shortcut behavior is ruled out. We note that an alternate proof could be
given using Rafi’s Theorem 3.12 instead of Theorem A, together with the stronger
version of statistical thickness provided by Theorem 5.10; see Remark 7.2.

Some of the delicate work is in establishing the thickness and separation prop-
erties in §5.2 and §6. Here we use a variety of recently developed tools such as vol-
ume asymptotics in Teichmüller space (Athreya–Bufetov–Eskin–Mirzakhani) and
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 thin parts corresponding to disjoint subsurfaces, then Minsky’s product regions theorem shows that the connecting geodesic can take a “shortcut.” We show that this effect is rare.

We also develop several facts that can serve as future tools, including a simplified formula for Teichmüller distance (§2.5), comparisons of measures (§4), and consequences of the thickness property: a typical ray makes definite progress in the curve complex, and typical pairs of rays become arbitrarily separated in Teichmüller space.

1.1. Acknowledgments. We would like to thank Benson Farb, Curtis McMullen, and especially Alex Eskin and Kasra Rafi for numerous helpful comments and explanations. In particular, we are indebted to Kasra Rafi for suggesting the idea of the proof for Theorem 5.10. We are also grateful for the suggestions of an anonymous referee, which led to major improvements in the paper.

2. Background

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 \((\rho, f)\); namely, the markings are 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 \(f_2\).

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 typically denote a point of \(\mathcal{T}(S)\) by \(x\), regarding it either as a Riemann surface or a hyperbolic surface, and suppressing the marking \(f\).

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

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

where the minimum is taken over all quasiconformal maps \(F\) and \(K(F)\) is the maximal dilatation of \(F\). Equipped with this metric, Teichmüller space becomes a unique geodesic metric space. For \(x, y \in \mathcal{T}(S)\), the Teichmüller geodesic segment
joining $x$ to $y$ will usually be denoted $[x,y]$. We will also write $y_t$ for the time-$t$ point on the ray based at $x$ and going through $y$.

A \textit{quadratic differential} on a Riemann surface $X$ is a 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 \textit{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 \textit{mapping class group} of $S$, defined by

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

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 [23], $\text{Mod}(S)$ is the full group of (orientation-preserving) isometries of $(\mathcal{T}(S),d_{\mathcal{T}})$.

2.2. Curve complex. When we speak of a \textit{curve} on $S$, this will mean an isotopy class of essential simple closed curves. Given $x \in \mathcal{T}(S)$, the length $l_2(\alpha)$ of a curve $\alpha$ is the length of the geodesic in the isotopy class in the hyperbolic metric $x$.

We recall the definition of the \textit{curve complex} (or curve graph) $\mathcal{C}(S)$ of $S$. The vertices of $\mathcal{C}(S)$ are the 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 [17]. That is, there exists a constant $\delta > 0$ such that every geodesic triangle in $\mathcal{C}(S)$ is $\delta$–thin: each side of the triangle is contained in the union of the $\delta$–neighborhoods of the other two sides. Furthermore, in any $\delta$–hyperbolic metric space and for any quasi-isometry constants $(K,C)$, there exists a constant $\tau$, depending only on $\delta,K,C$, such that any two $(K,C)$–quasi-geodesic segments with the same endpoints remain within $\tau$ of each other. Since actual geodesics are $(1,0)$–quasi-geodesics, this implies that every $(K,C)$–quasi-geodesic triangle is $(\delta + 2\tau)$–thin.

2.3. Thick parts and subsurface projections. For any given $\epsilon$, we say a curve is $\epsilon$–short if its hyperbolic length is less than $\epsilon$. Then define the $\epsilon$–\textit{thick part} of Teichmüller space to be the subset $\mathcal{T}_\epsilon \subset \mathcal{T}(S)$ corresponding to those hyperbolic
surfaces on which no curve is $\epsilon$–short. Its complement is called the $\epsilon$–thin part or, when $\epsilon$ is understood, simply the thin part.

For each $x \in \mathcal{T}(S)$ there is associated a Bers marking $\mu_x$. To construct $\mu_x$, greedily choose a shortest pants decomposition of the surface (a collection of $3g-3$ disjoint simple geodesics). Then for each pants curve $\beta$, choose a shortest geodesic crossing $\beta$ 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. Notice that the curves comprising $\mu_x$ for a diameter–2 subset of $\mathcal{C}(S)$.

Recall that there exists a universal Margulis constant such that any two curves with hyperbolic length (on any surface $x \in \mathcal{T}(S)$) less than this value are disjoint. When discussing the $\epsilon$–thick part $\mathcal{T}_\epsilon$, we always assume $\epsilon$ is less than the Margulis constant. In particular, this ensures that for $x \in \mathcal{T}(S) \setminus \mathcal{T}_\epsilon$, the Bers marking $\mu_x$ contains every curve $\alpha$ with $l_x(\alpha) \leq \epsilon$.

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 $\partial V$ consisting of a disjoint union of curves on $S$. We say that two subsurfaces $V$ and $W$ transversely intersect, denoted $V \cap W$, if they are neither (isotopically) disjoint nor nested. In this case, $\partial V$ necessarily intersects $W$, and $\partial W$ 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 $\beta$ and $\partial V$ as geodesics (in any hyperbolic metric on $S$). If $\beta \subset V$, then $\pi_V(\beta)$ is defined to be $\beta$. If $\beta$ 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 $\partial V$, and we obtain $\pi_V(\beta)$ by choosing any arc and performing a surgery along $\partial V$ 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 $\beta$ in the Bers marking at $x$. This is a non-empty subset of the curve complex $\mathcal{C}(V)$ with uniformly bounded diameter.

**Definition 2.1** (Non-annular projection distance). For a non-annular subsurface $V \subset S$, the projection distance in $V$ of a pair of points $x, y \in \mathcal{T}(S)$ is defined to be

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

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 $\alpha$ and is analogous to the definition above, and a second which also incorporates the length of $\alpha$. Any simple closed curve $\beta$ that crosses $\alpha$ 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 \( \alpha \), with the Gromov compactification. For a pair \( \beta, \gamma \) of such curves, we may then consider the intersection number \( i(\tilde{\beta}, \tilde{\gamma}) \) in \( \tilde{A} \). The twisting distance in \( A \) of a pair of points \( x, y \in T(S) \) is then defined as

\[
d_{C(A)}(x, y) := \sup_{\beta \in \mu_x, \gamma \in \mu_y} i_{\tilde{A}}(\tilde{\beta}, \tilde{\gamma}).
\]

We additionally define a hyperbolic projection distance as follows.

**Definition 2.2** (Annular projection distance). For an annular subsurface \( A \subset S \) with core curve \( \alpha = \partial A \), we let \( H_\alpha \) denote a copy of the standard horoball \( \{ \text{Im}(z) \geq 1 \} \subset \mathbb{H}^2 \). Given \( x, y \in T(S) \), we consider the points \( (0, 1/l_x(\alpha)) \) and \( (d_{C(A)}(x, y), 1/l_y(\alpha)) \) in \( \mathbb{H}^2 \) and denote their closest point projections to the horoball \( H_\alpha \) by

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

The projection distance in \( A \) (or hyperbolic distance \( d_{H_\alpha} \)) between \( x \) and \( y \) is then defined to be

\[
d_A(x, y) := d_{H^2}(\pi_\alpha(x), \pi_\alpha(y)).
\]

### 2.4. Notation.

Following Rafi [20, §2.4], we fix a parameter \( \epsilon_0 > 0 \) for the entirety of this paper which is smaller than the Margulis constant and small enough for a few other fundamental results to hold (Minsky’s product regions theorem and Rafi’s distance estimates described in the following section). Note that the definition of \( \epsilon_0 \) depends only on the topology of the surface \( S \), and we therefore view \( \epsilon_0 \) as a global constant.

Our analysis involves many inequalities that have controlled multiplicative and additive error. To streamline the presentation, we will often avoid explicitly writing the constants involved and will instead rely on the following notation: For real-valued expressions \( A \) and \( B \), we use the notation

\[
A \precsim B
\]

to mean that there exists a universal constant \( c \geq 1 \), depending only on the topology of the surface \( S \), such that \( A \leq cB \). We will use \( A \asymp B \) to mean that \( A \precsim B \) and \( A \asymp B \) both hold. (The dot in the symbols indicates that the error is only multiplicative.) When allowing for multiplicative and additive error we will instead use symbols \( <, >, \text{ and } \asymp \). Thus \( A \asymp B \) means that there exists a universal constant \( c \geq 1 \) so that \( A \leq cB + c \) and \( B \leq cA + c \).

When the implied constant depends on additional parameters we will list these as subscripts of the binary relation. For example, \( A \precsim_{\epsilon, \theta} B \) means that there exists a constant \( c \) depending only on \( \epsilon, \theta \), and the topology of \( S \) such that \( A \leq cB \).

### 2.5. Distance formula.

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 \( \mu_x \) and \( \mu_y \). Recall the global constant \( \epsilon_0 > 0 \) introduced in §2.4 above.
Theorem 2.3 (Distance formula, Rafi [20]). Given any sufficiently large threshold $M_0$, for all $x, y \in \mathcal{T}(S)$ we have
\[
d_T(x, y) \asymp_{M_0} d_S(x, y) + \sum_v d_V(x, y)_{M_0} + \max_{\alpha \in \Gamma_x} d_{\mathbb{H}_\alpha}(x, y)
\]
\[
+ \sum_{A : \partial A \not\in \Gamma_{xy}} \log_+ \left( \frac{1}{l_x(A)} \right)_{M_0} + \max_{\alpha \in \Gamma_x} \log_+ \left( \frac{1}{l_y(\alpha)} \right),
\]
where $\Gamma_{xy}$ is the set of $\epsilon_0$–short curves in both $x$ and $y$, $\Gamma_x$ is the set of curves that are $\epsilon_0$–short in $x$ but not in $y$, and $\Gamma_y$ 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.

By instead making all annular measurements with the hyperbolic distance on $\mathbb{H}_\alpha$ we will obtain a particularly simple restatement of this formula.

Proposition 2.4 (Repackaged distance formula). Given any sufficiently large threshold $M_0$, for all $x, y \in \mathcal{T}(S)$ we have:
\[
(2.5) \quad d_T(x, y) \asymp_{M_0} d_S(x, y) + \sum_v d_V(x, y)_{M_0}
\]
Here, the sum is over all (annular and non-annular) proper subsurfaces.

Remark 2.6. The definition of $d_{\mathbb{H}_\alpha} = d_A$ given above is technically different than that used by Rafi in [20]; however, the two definitions agree up to bounded additive error.

This version of the distance formula treats annular and non-annular subsurfaces on equal footing. For simplicity and without loss of generality, below we suppose that $\epsilon_0$ has been chosen small enough that $\log_+ (1/\epsilon_0) \geq 100$, say. We begin with a straightforward reformulation.

Lemma 2.7. For any sufficiently large threshold $M_0$, for all $x, y \in \mathcal{T}(S)$ we have:
\[
d_T(x, y) \asymp_{M_0} d_S(x, y) + \sum_v d_V(x, y)_{M_0} + \sum_{A : \partial A \not\in \Gamma_{xy}} [d_A(x, y)]_{M_0} + \sum_{A : \partial A \not\in \Gamma_{xy}} \left[ \max \left\{ \log_+ (d_C(A)(x, y))_{M_0}, \log_+ \left( \frac{1}{l_x(\partial A)} \right), \log_+ \left( \frac{1}{l_y(\partial A)} \right) \right\} \right]_{M_0}
\]

Proof. Since $\Gamma_{xy}$, $\Gamma_x$ and $\Gamma_y$ 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 $\Gamma_x$ we have
\[
\sum_{\alpha \in \Gamma_x} \log_+ \left( \frac{1}{l_x(\alpha)} \right) = \sum_{\alpha \not\in \Gamma_{xy}} \log_+ \left( \frac{1}{l_x(\alpha)} \right)_{1/\epsilon_0}.
\]
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 (in fact with the implied multiplicative constant equal to 3)
\[
\log_+ [f]_{M_0} + \log_+ [g]_{M_0} + \log_+ [h]_{M_0} \leq \left[ \max \{ \log_+ f, \log_+ g, \log_+ h \} \right]_{\log M_0}.
\]
We now show that each term in the last summand is bilipschitz equivalent to the corresponding hyperbolic distance $d_A(x,y)$.

**Lemma 2.8.** 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

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

If $x, y \in \mathcal{T}(S)$ are such that $\alpha \not\in \Gamma_{xy}$ and either $d_A(x,y)$ or $H_A(x,y)$ is greater than $36 \log_+(1/\epsilon_0)$, then $6^{-1}d_A(x,y) \leq H_A(x,y) \leq 6d_A(x,y)$.

**Proof.** Choose points $x, y \in \mathcal{T}(S)$ that satisfy the hypotheses. To fix notation, set $\pi'_\alpha(x) = (0,1)$ and $\pi'_\alpha(y) = (d_{C(A)}(x,y),1)$. 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

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

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

$$\log_+ d_{C(A)}(x,y) \leq B \leq 4 \log_+ d_{C(A)}(x,y)$$

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

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

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

Then (2.9), (2.10), and the definition of $H_A$ imply that $d_A(x,y) \leq 6H_A(x,y)$ in the case that $B \geq 3$. 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 $d_A(x,y) \leq 3L \leq 3H_A(x,y)$, where $L$ denotes the larger of the other two terms. Indeed, if $B$ were the largest term and $B < 3$, then (2.10) would imply $d_A(x,y) < 9$, and (2.9) would necessitate $\log_+ d_{C(A)}(x,y) < 3$ 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 $d_A(x,y) \geq 36 \log_+(1/\epsilon_0)$ implies that $H_A(x,y) \geq 6 \log_+(1/\epsilon_0)$; therefore all cases will be covered by proving that this in turn implies $H_A(x,y) \leq 6d_A(x,y)$. Without loss of generality, we may assume that $l_x(\alpha) \leq l_y(\alpha)$; since $\alpha \not\in \Gamma_{xy}$, this guarantees $l_y(\alpha) \geq \epsilon_0$. First suppose that

$$\log_+ d_{C(A)}(x,y) \geq \log_+ d_{C(A)}(x,y) \geq 3 \log_+(1/l_x(\alpha)),$$

in which case we have $H_A(x,y) \geq 6 \log_+(1/\epsilon_0)$. In particular we certainly have $d_{C(A)}(x,y) \geq 3$; thus (2.9) and the triangle inequality give

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

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

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

$$\log_+ \left( \frac{1}{l_x(\alpha)} \right) \leq d_{\mathbb{H}^2}(\pi_\alpha(x), \pi_\alpha(y)) \leq d_A(x,y) + \log_+ \left( \frac{1}{l_y(\alpha)} \right).$$
The assumptions $3 \log_2(1/l_y(\alpha)) \geq H_A(x, y) \geq 6 \log_2(1/\epsilon_0)$ and $l_y(\alpha) \geq \epsilon_0$ now ensure that $H_A(x, y) \leq 6d_A(x, y)$.

**Corollary 2.11.** Let $H_A(x, y)$ be defined as in Lemma 2.8. Then for any threshold $M_0 \geq 36 \log_2(1/\epsilon_0)$ and any $x, y \in T(S)$ we have

$$\sum_{\partial A \not\in T_{xy}} 6^{-1} [d_A(x, y)]_{6M_0} \leq \sum_{\partial A \in T_{xy}} [H_A(x, y)]_{M_0} \leq \sum_{\partial A \not\in T_{xy}} 6 [d_A(x, y)]_{M_0/6}$$

With these estimates, we can derive the simplified distance formula.

**Proof of Repackaged Distance Formula.** Choose any sufficiently large threshold $M_0$ such that Lemma 2.7 holds for both $e^{6M_0}$ and $M_0/6$ and such that $M_0/6 \geq 36 \log_2(1/\epsilon_0)$. Notice that, in any sum of the form $\sum [f]_{M_1}$, raising the threshold can only decrease the value of the sum, and lowering the threshold can only increase its value. Therefore, combining Lemma 2.7 and Corollary 2.11 we find that for any $x, y \in T(S)$ the various distances satisfy

$$d_T \leq \sum_{V} [d_V]_{6M_0} + \sum_{\partial A \in T_{xy}} [d_A]_{e^{6M_0}} + \sum_{\partial A \not\in T_{xy}} [H_A]_{e^{6M_0}}$$

$$\leq 6 \left( \sum_{V} [d_V]_{M_0} + \sum_{\partial A \in T_{xy}} [d_A]_{M_0} + \sum_{\partial A \not\in T_{xy}} [d_A]_{M_0} \right),$$

where we have suppressed the $x$ and $y$ in the notation. The lower bound on $d_T(x, y)$ is similar.

### 2.6. Thin intervals

We will use some results from Rafi’s work combinatorializing the Teichmüller metric. Specifically, Corollary 3.4 and Proposition 3.7 of [20] show that for every Teichmüller geodesic and every proper subsurface $V$, there is a (possibly empty) interval along the geodesic where $\partial V$ 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: for each positive $\epsilon \leq \epsilon_0$ there are positive constants $M_\epsilon$ and $\epsilon' \leq \epsilon$ such that for any pair of points $x, y \in T(S)$ there is a possibly empty (and not uniquely defined) connected interval $I'_V$ along the geodesic segment $[x, y]$ such that

- for $a \in I'_V$, the length each component of $\partial V$ on $a$ is at most $\epsilon$;
- for $a \in [x, y] \setminus I'_V$, some component $\beta$ of $\partial V$ has $l_\alpha(\beta) \geq \epsilon'$;
- for $a, b$ in the same component of $[x, y] \setminus I'_V$, we have $d_V(a, b) < M_\epsilon$; and
- if $V \nparallel W$ then $I'_V \cap I'_W = \emptyset$.

This $I'_V$ is called the $\epsilon$–thin interval for $V$ (or active interval in some papers), or just the thin interval when $\epsilon$ is understood. While $[x, y]$ is suppressed in the notation, the geodesic with respect to which the interval $I'_V$ is defined should be clear from context. We also note that for us $I'_V$ is a segment in Teichmüller space, whereas Rafi works with the corresponding time interval $I \subset \mathbb{R}$.

If $I'_V \neq \emptyset$ we will say that $V$ becomes thin along $[x, y]$. In particular if $d_V(x, y) \geq M_\epsilon$, then $I'_V \neq \emptyset$ and so $V$ becomes thin along $[x, y]$. Note that the second condition above says that the complement of the union of thin intervals (for all proper subsurfaces) lies in the $\epsilon'$–thick part of $T(S)$.

We always assume that $M_\epsilon$ is chosen large enough to be a valid threshold in the distance formula (2.5). In the case $\epsilon = \epsilon_0$ we will omit the parameter and
simply write $M$ and $I_V$ for $M_{\epsilon_0}$ and $I_V^0$. Thus $M$ is a global constant that depends only on the topology of $S$.

### 2.7. 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 [17] and then for general subsurfaces by Rafi [21, Thm B].

**Lemma 2.12** (Reverse triangle inequality). There exists a global constant $B > 0$ such that for any non-annular subsurface $V$ (including $S$ itself) and for any geodesic interval $[x,y] \subset T(S)$ and any point $a \in [x,y]$ we have

\[ d_V(x,a) + d_V(a,y) \leq d_V(x,y) + B. \]

(2.13)

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

In the exceptional annulus case, we have the following theorem from Rafi [20].

**Theorem 2.14** (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 $d_A(x,a) + d_A(a,y) - d_A(x,y) \geq B'$ (i.e., the reverse triangle inequality fails), then there exists a collection of disjoint annuli $A_i$, one of which is $A$ and for each the reverse triangle inequality fails, and a collection of subsurfaces $W_j$ disjoint from the $A_i$ such that

- for each $W_j$ the reverse triangle inequality (2.13) holds along $[x,y]$,
- $d_{W_j}(a,y) \geq M_0$ for each $j$, and
- $d_{A_i}(a,y) \leq \sum_j d_{W_j}(a,y)$ for each $i$.

We remark that this is not exactly how the result in [20] 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 a pair of expanding annuli. 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 $H^2$ defined by the length and twist coordinates by computing the modulus of these annuli. He shows that the distance traveled in $H^2$ 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 $H^2$ is much smaller than Teichmüller length forces, by his distance formula, Proposition 2.4, the presence of the domains $W_j$ as in the statement of the theorem.

Going forward, for each $\epsilon \leq \epsilon_0$ we additionally assume that $M_\epsilon$ is chosen large enough to satisfy $M_\epsilon \geq B$ and to be a valid threshold in Theorem 2.14 above.

### 3. The geometry of statistically thick geodesics

In the study of Teichmüller geometry, one finds that the thick part $T_\epsilon$ behaves very much like a negatively curved space. For example in Theorems 4.4 and 7.6 of [11] it is shown that geodesic triangles contained entirely within $T_\epsilon$ are $\delta$–thin for some $\delta$, and that the projection of any geodesic $\gamma \subset T_\epsilon$ to the curve complex
\( \mathcal{C}(S) \) is an honest parametrized quasi-geodesic which, in particular, must progress at a linear rate. These facts can also be deduced from the distance formula (2.5) together with properties of thin intervals \( \Gamma' \) (§2.6 above).

All of these negative curvature properties are lost when geodesics are allowed to enter the thin part. For example, Minsky’s product region theorem [19] shows that geodesic triangles in \( T(S) \setminus T_e \) need not be \( \delta \)-thin for any \( \delta \), and it is easy to construct arbitrarily long geodesics in \( T(S) \setminus T_e \) that project to uniformly bounded diameter sets in \( \mathcal{C}(S) \).

However, both of these extremes – living entirely in \( T_e \) or entirely in \( T(S) \setminus T_e \) – are quite rare, as a typical Teichmüller geodesic will spend part of its time in \( T_e \) and part of its time in \( T(S) \setminus T_e \). In this section we develop tools to study geodesics with exactly this behavior (later on, in §5 we will show that this behavior is in fact generic in a certain quantifiable sense).

Our techniques rely on controlling the fraction of time a Teichmüller geodesic spends in a given thick part \( T_e \). We call this quantity the thick-stat; for a non-degenerate geodesic segment \( [x,y] \subset T(S) \) it is denoted by

\[
\text{Thk}^s_{T_e}[x,y] = \frac{|\{0 \leq s \leq d_T(x,y) : y_s \in T_e\}|}{d_T(x,y)},
\]

where \( y_s \) is the time–s point on the geodesic ray from \( x \) through \( y \). Thus \( [x,y] \subset T_e \) is equivalent to \( \text{Thk}^s_{T_e}[x,y] = 1 \). In the following subsections, we will show that negative curvature properties similar to those mentioned above for thick geodesics also hold when \( \text{Thk}^s_{T_e} \) is merely bounded away from zero.

3.1. Progress in the curve complex. The goal of this subsection is to prove Theorem 3.9, which says that geodesics that spend a definite fraction of their time in the thick part \( T_e \) must move at a definite linear rate in the curve complex.

The idea is that long subintervals contained in \( T_e \) contribute to progress in \( \mathcal{C}(S) \); alternately, one could consider intervals in the complement of all the \( \epsilon \)-thin intervals \( \Gamma' \) for proper subsurfaces \( V \subset S \). For this analysis, we would like to bound the number of connected components of \( \bigcup_{V} \Gamma'_V \) 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 \( \Gamma'_V \), some of these will be redundant in the sense that \( \Gamma'_V \subset \Gamma'_W \) for some other subsurface \( W \).

**Definition 3.1** (Thin-significance). Fix \( 0 < \epsilon \leq \epsilon_0 \). A proper subsurface \( V \subset S \) is said to be \( \epsilon \)-thin-significant (or simply thin-significant) for the geodesic segment \( [x,y] \) if \( d_{\mathcal{C}(V)}(x,y) \geq 3M_\epsilon \) and for every other proper subsurface \( Z \subset S \) with \( d_{\mathcal{C}(Z)}(x,y) \geq 3M_\epsilon \) we have \( \Gamma'_V \not\subset \Gamma'_Z \).

**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 first goal is to bound the number of thin-significant subsurfaces along an arbitrary geodesic. For this, we will use the work of Rafi–Schleimer [22] bounding the size of an antichain in the poset of subsurfaces of \( S \).

**Definition 3.2** (Antichain). Given a subsurface \( \Sigma \subset S \) a pair of points \( x, y \in T(S) \) and constants \( T_0 \geq T_0 > 0 \), a collection \( \Omega \) of proper subsurfaces of \( \Sigma \) is an antichain for \( (\Sigma, x, y, T_0, T_1) \) if the following hold:
• if $Y,Y' \in \Omega$, then $Y$ is not a proper subsurface of $Y'$;
• if $Y \in \Omega$, then $d_{C(S)}(x,y) \geq T_0$; and
• if $Z \subsetneq \Sigma$ and $d_{C(S)}(x,y) \geq T_1$, then $Z \subset Y$ for some $Y \in \Omega$.

**Lemma 3.3** (Antichain bound [22, 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 $\Omega$ is an antichain for $(\Sigma, x, y, T_0, T_1)$ then

$$|\Omega| \leq A \cdot d_{C(S)}(x,y).$$

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

1. $V \preceq_1 W \iff V \subset W$,
2. $V \preceq_2 W \iff I_V \subset I_W$, and
3. $V \preceq_3 W \iff V \subset W$ and $I_V \subset I_W$.

The subcollection of $\Omega$ consisting of maximal elements with respect to $\preceq_3$ will be denoted $(\Omega)_3$; 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 $\preceq_3$.

**Proposition 3.5** (Progress from thin-significant subsurfaces). For any positive $\epsilon \leq \epsilon_0$ and any $t_0$, there is a constant $N$ such that if $d_{C(S)}(x,y) \leq t_0$, then the number of $\epsilon$-thin-significant subsurfaces along $[x,y]$ is at most $N$.

**Proof.** Let $\Omega = \{ V \subsetneq S : d_{C(S)}(x,y) \geq 3M_\epsilon \}$ be the collection of proper subsurfaces which have a large projection. By definition, the set of $\epsilon$-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 $(S, x, y, 3M_\epsilon, 3M_\epsilon)$. By Lemma 3.3, 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 $I_V \cap I_W \neq \emptyset$. The definition of $\preceq_3$ implies that $I_V \not\subseteq I_W$; therefore $I_V$ must overlap with at least one endpoint of $I_W$. If $I_V \preceq_3 I_W$, both contain the initial endpoint of $I_W$, then $I_V \cap I_W \neq \emptyset$ and we cannot have $V_1 \cap V_2$. 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 $I_V \cap I_W \neq \emptyset$.

It remains to bound the number of $V \in \mathcal{U}_W$ for which $I_V \cap I_W = \emptyset$; we will only focus on the case that $I_V$ occurs before $I_W$ when traveling from $x$ to $y$. Suppose that $I_W = [a,b] \subset [x,y]$ and consider the set $\Omega' = \{ V \in \Omega : V \subsetneq W \text{ and } d_{C(S)}(x,a) \geq 2M_\epsilon \}$. Notice that the subcollection $(\Omega')_1$ forms an antichain for $(W, x, a, 2M_\epsilon, 4M_\epsilon)$: the only difficulty is to check that every $Y \subsetneq W$ with $d_{C(S)}(x,a) \geq 4M_\epsilon$ is contained in an element of $(\Omega')_1$. However, this is true because the reverse triangle inequality guarantees that $d_{C(S)}(x,y) \geq d_{C(S)}(x,a) - B \geq 3M_\epsilon$ and therefore that...
$Y \in \Omega'$. Since $d_{c(V)}(x, a) \leq M_\epsilon$, Lemma 3.3 now gives a bound on $|\{\Omega'\}|$. Finally, notice that for each $V \in \mathcal{U}_W$ with $I_V$ occurring before $I_W$ along $[x, y]$, the triangle inequality gives $d_{c(V)}(x, a) \geq d_{c(V)}(x, y) - M_\epsilon \geq 2M_\epsilon$ 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 $I_V$ occurring before $I_W$ along $[x, y]$ is contained in $\mathcal{U}_Z$ for some $Z \in \{\Omega'\}$. The bound on $|\mathcal{U}_W|$ now follows by induction on the complexity of the subsurface $W$. 

\section*{Definition 3.6.} For $\epsilon \leq \epsilon_0$, define a constant $P_\epsilon := 36 \max \{\log_4(1/\epsilon), \log_4(3M_\epsilon)\}$. We say that an annular subsurface $A$ has an exceptional $\epsilon$–thin interval $I_A^\epsilon$ along $[x, y]$, if $d_{c(A)}(x, y) \leq 3M_\epsilon$ but $d_A(x, y) \geq P_\epsilon$. By Lemma 2.8, this is only possible if $I_A^\epsilon(\partial A) < \epsilon$ or $I_A^\epsilon(\partial A) > \epsilon$. Therefore, such an annulus must determine a nonempty $\epsilon$–thin interval along $[x, y]$ that contains either $x$ or $y$. Since all annuli $A$ with $I_A^\epsilon(\partial A) \leq \epsilon$ are 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 $\epsilon$–thin portion $W_\epsilon$ of a geodesic segment $[x, y]$ to be the union of thin intervals $I_A^\epsilon$ for all non-annular proper subsurfaces with $d_{c(V)}(x, y) \geq 3M_\epsilon$ and all annular subsurfaces with $d_{c(A)}(x, y) \geq 3M_\epsilon$ or $d_A(x, y) \geq P_\epsilon$. In the case that $d_Y(x, y) \leq t_0$, Proposition 3.5 implies that $W_\epsilon$ is the union of at most $N + h$ thin intervals, namely, those corresponding to the $\epsilon$–thin-significant subsurfaces and to annuli with exceptional $\epsilon$–thin intervals.

While $W_\epsilon$ does contain most of the nonempty $\epsilon$–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 $W_\epsilon$.

\section*{Lemma 3.7 (Complement of $W_\epsilon$).} For each positive $\epsilon \leq \epsilon_0$, if $[a, b] \subset [x, y] \setminus W_\epsilon$ is a connected interval in the complement of the primary $\epsilon$–thin portion of $[x, y]$, then $d_Y(a, b) <_{\epsilon} 1$ for all proper subsurfaces $Y \subseteq S$.

\section*{Proof.} First suppose that $Y$ satisfies the reverse triangle inequality (2.13) along $[x, y]$. If $Y$ is non-annular and $d_Y(x, y) \geq 3M_\epsilon$, or if $Y$ is an annulus and $d_{c(Y)}(x, y) \geq 3M_\epsilon$ or $d_Y(x, y) \geq P_\epsilon$, then $I_Y^\epsilon \subset W_\epsilon$ by definition. Therefore $d_Y(a, b) \leq M_\epsilon$ since $[a, b] \cap I_Y^\epsilon = \emptyset$. If this is not the case, then the reverse triangle inequality gives $d_Y(a, b) \leq d_Y(x, y) + 2B \leq 5M_\epsilon + P_\epsilon$ as claimed.

It remains to consider an annular subsurface $A \subset S$ for which the reverse triangle inequality fails. We may assume that $d_{c(A)}(x, y) \leq 3M_\epsilon$ and $d_A(x, y) \leq P_\epsilon$, for otherwise we have $I_A^\epsilon \subset W_\epsilon$ and $d_A(x, y) \leq M_\epsilon$ as above. Let $B'$ be the constant corresponding to the threshold $6M_\epsilon + P_\epsilon$ in Theorem 2.14 (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 $W_j$ that satisfy the reverse triangle inequality and which have $d_{W_j}(a, b) \geq 6M_\epsilon + P_\epsilon$. However, as we have seen above, there are no such proper subsurfaces. Therefore the former inequality must hold. We similarly have $d_A(x, a) + d_A(a, b) \leq d_A(x, b) + B'$. Adding these inequalities and using the triangle inequality then gives $d_A(a, b) \leq d_A(x, y) + B' \leq P_\epsilon + B'$.

By the distance formula (2.5), it now follows that long intervals in the complement of $W_\epsilon$ 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 3.8** (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

$$d_{S}(x, y) \geq \rho_1 nd.$$  

**Proof.** Applying the reverse triangle inequality (2.13) to the points $x_i$ and $y_i$ we have $d_{S}(x, x_i) + d_{S}(x_i, y_i) + d_{S}(y_i, y) \leq d_{S}(x, y) + 2\mathcal{B}$. By recursively applying this observation to $[x, x_i]$ and $[y_i, y]$ and then throwing out the complementary intervals, we find that

$$d_{S}(x, y) \geq \sum d_{S}(x_i, y_i) - 2n\mathcal{B} \geq nd - 2n\mathcal{B}.$$  

Choose $D_1 > 4\mathcal{B}$ and $\rho_1 = 1/2$. Then for $d \geq D_1$ the quantity on the right side is at least $\rho_1 nd$. \( \Box \)

We now fix once and for all a “definite progress” constant $D > 0$ sufficiently large so that $\rho_1 D > D_1$ (and thus $D > D_1$ as well). Applying the distance formula (2.5) using the implied bound from Lemma 3.7 as the threshold, we see that $d_{\tau}(a, b) \prec_{\epsilon} d_{S}(a, b)$ for any connected interval $[a, b] \subset [x, y] \setminus \mathcal{W}_\epsilon$. This gives rise to a fixed value $L_\epsilon$ such that any interval $[a, b]$ of length at least $L_\epsilon$ that lies entirely in $[x, y] \setminus \mathcal{W}_\epsilon$ satisfies $d_{S}(a, b) \geq D$. Thus according to Lemma 3.8, if $I$ is any interval along a geodesic that contains a subinterval of length $L_\epsilon$ that is disjoint from $\mathcal{W}_\epsilon$, then the distance in the curve complex between the endpoints of $I$ is at least $\rho_1 D$.

**Theorem 3.9** (Definite progress). For every $\epsilon > 0$ and $0 < \theta < 1$, there exists a constant $R_1 > 0$ such that

$$d_{S}(x, y) \succ_{\epsilon, \theta} d_{\tau}(x, y)$$

for every Teichmüller geodesics $[x, y]$ satisfying $d_{\tau}(x, y) \geq R_1$ and $\text{Thk}_{\epsilon}^{\rho_1}[x, y] \geq \theta$.

**Proof.** Since shrinking $\epsilon$ preserves the hypothesis $\text{Thk}_{\epsilon}^{\rho_1}[x, y] \geq \theta$, we may assume $\epsilon \leq \epsilon_0$. Let $N$ denote the constant obtained by applying Proposition 3.5 with the parameters $\epsilon$ and $\tau_0 = \rho_1 D$. Choose $n$ so that $n\theta > 1$ and make the following definitions:

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

Let $[x, y]$ be a Teichmüller geodesic of length $r \geq R_1$ satisfying $\text{Thk}_{\epsilon}^{\rho_1}[x, y] \geq \theta$. 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 $d_{S}(a, b) < \rho_1 D$ and **progressing** if $d_{S}(a, b) \geq \rho_1 D$. 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 $\epsilon$–thin portion $\mathcal{W}_\epsilon$ and the corresponding complementary subintervals. Since the interval is stalled, Proposition 3.5 ensures that $\mathcal{W}_\epsilon$ is the union at most $N + h$ thin intervals. Therefore we can conclude that $\mathcal{W}_\epsilon$ has at most $N + h + 1$ complementary subintervals in $[a, b]$. Furthermore, each complementary subinterval has length at most $L_\epsilon$, for otherwise we would have $d_{S}(a, b) \geq \rho_1 D$ by the preceding
paragraph. Since $W_e$ is contained in the $\epsilon$–thin part, we see that the total amount of time that this interval $[a, b]$ spends in the thick part is at most
\[(N + h + 1)\epsilon \leq \theta'T_0 \leq \theta'r/m.\]
Therefore the total amount of time that the full interval $[x, y]$ spends in the thick part is at most
\[\left(\frac{\theta'}{m}\right) m_1 + \left(\frac{r}{m}\right) m_2 = \frac{r}{m}(\theta'm_1 + m_2).\]

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 $\theta' < 1$, it follows that
\[\theta'm_1 + 1 \cdot m_2 < \theta'm_\frac{n-1}{n} + 1 \cdot \frac{1}{n}\]
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
\[(\alpha \cdot a + \beta \cdot b < \alpha \cdot c + \beta \cdot d) \iff a > c.\]
But then the amount of time that $[x, y]$ is thick is less than
\[\frac{r}{m}\left(\theta'm_\frac{n-1}{n} + \frac{1}{n}\right) = r\left(\frac{n\theta - 1}{n-1} \cdot \frac{n-1}{n} + \frac{1}{n}\right) = r\theta,\]
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 $\rho_1D$. Therefore, cumulative contribution of subintervals (Lemma 3.8) implies that
\[d_S(x, y) \geq \rho_1m_2(\rho_1D) \geq \rho_1^2D \frac{m}{n} \geq \frac{\rho_1^2D}{n} \left(\frac{r}{T_0} - 1\right) \geq \frac{\rho_1^2D}{2nT_0} r = pr. \quad \square\]

Remark 3.11. After developing our proof of Theorem 3.9 we learned of an independent yet closely related result of Hamenstädt’s, namely Proposition 2.1 of [9], which under the same hypotheses provides a lower bound on $d_S(x, y)$ that is constant rather than linear in $d_T(x, y)$. In fact, the linear bound in Theorem 3.9 may be deduced from Hamenstädt’s result by breaking $[x, y]$ into subintervals, applying [9, Proposition 2.1] to those with large thick-stat, and adding the resulting contributions using Lemma 3.8, much as we have done above. With this approach [9, Proposition 2.1] would effectively replace the use of Proposition 3.5 and Lemma 3.7 in our argument. However, we have decided to retain our original argument using Proposition 3.5 and Lemma 3.7 as we believe these to be of independent interest.

3.2. A statistical thin triangles statement. In this subsection we prove Theorem A and obtain thinness results for geodesic triangles whose sides satisfy various thick-stat conditions. Recall that given $\epsilon > 0$ there is a $\delta > 0$ such that every geodesic triangle whose sides live entirely in $T_\epsilon$ is $\delta$–thin. This fact can be deduced from the following theorem of Rafi, which gives specific information under much more general conditions.

Theorem 3.12 (Rafi [21, Theorem 8.1]). For every $\epsilon > 0$ there exist constants $C_1, L_1$ such that if $I \subset [x, y] \subset T(S)$ is a geodesic subinterval of length at least $L_1$ lying entirely in the $\epsilon$–thick part, then for all $z \in T(S)$, we have
\[I \cap Nbd_{C_1}([x, z] \cup [y, z]) \neq \emptyset.\]

We weaken the hypothesis to only require a definite thick-stat.
Theorem A. For any $\epsilon > 0$ and $0 < \theta \leq 1$, there exist constants $C, L$ such that if $I \subset [x, y] \subset \mathcal{T}(S)$ is a geodesic subinterval of length at least $L$ and at least proportion $\theta$ of $I$ is $\epsilon$-thick, then for all $z \in \mathcal{T}(S)$, we have

$$I \cap \text{Nbd}_C([x, z] \cup [y, z]) \neq \emptyset.$$ 

Before proving this result, we discuss two consequences. Firstly we observe that there is not merely one point in the subinterval $I$ which is close to $[x, z] \cup [y, z]$, but in fact this conclusion holds for a large fraction of the interval $I$.

Proposition 3.13. For any $\epsilon > 0$ and $0 < \theta' < \theta \leq 1$, there are constants $L', C'$ so that if a side $[x, y]$ of a geodesic triangle $\triangle(x, y, z) \subset \mathcal{T}(S)$ contains a subinterval $I \subset [x, y]$ of length at least $L'$ with $\text{Thk}_{\epsilon}^C(I) \geq \theta$, then at least proportion $\theta'$ of $I$ is within distance $C'$ of $[x, z] \cup [y, z]$. That is, if length denotes Lebesgue measure along a geodesic segment,

$$\text{length} \left( \{ I \cap \text{Nbd}_{C'}([x, z] \cup [y, z]) \} \right) \geq \theta' \cdot \text{length}(I).$$

Proof. Let $\rho = \frac{\theta - \theta'}{1 - \theta'}$. Apply Theorem A to $\mathcal{T}_\epsilon$ with the fraction $\rho$ and let $L' = L$ and $C$ be the corresponding constants. Given a subinterval $I \subset [x, y]$ satisfying the hypotheses of the theorem, divide $I$ into $n = \lfloor \text{length}(I)/L \rfloor$ subintervals of equal length (the length will be between $L$ and $2L$). Let $a$ denote the fraction of these subintervals that have $\text{Thk}_{\epsilon}^C \geq \rho$ (so $a = \frac{k}{n}$ for some $k \in \{0, \ldots, n\}$). Each of these $na$ subintervals can spend at most all of their time in $\mathcal{T}_\epsilon$ and the other $n(1 - a)$ subintervals spend less than proportion $\rho$ of their time in $\mathcal{T}_\epsilon$. Therefore the maximum amount of time the whole interval $I$ can spend in $\mathcal{T}_\epsilon$ is less than

$$1 \cdot na \cdot \frac{\text{length}(I)}{n} + \rho \cdot n(1 - a) \cdot \frac{\text{length}(I)}{n} = \text{length}(I)(a + \rho - pa).$$

Since we have $\text{Thk}_{\epsilon}^C(I) \geq \theta$ by hypotheses, this implies $\theta < a + \rho - ap$. That is,

$$a > \frac{\theta - \rho}{1 - \rho} = \theta'.$$

Now, Theorem A implies that each subinterval with $\text{Thk}_{\epsilon}^C \geq \rho$ contains a point within distance $C$ of $[x, z] \cup [y, z]$. Therefore, each of the $na$ subintervals satisfying $\text{Thk}_{\epsilon}^C \geq \rho$ is contained entirely within the $C' = C + 2L$ neighborhood of $[x, z] \cup [y, z]$. As the union of these $na$ subintervals comprise proportion $a > \theta'$ of the interval $I$, the statement follows. \hfill $\square$

From this we obtain the following immediate corollary.

Corollary 3.14 (Statistically thin triangles). For all $\epsilon > 0$ and $0 < \theta' < \theta \leq 1$ there exists a constant $\delta$ with the following property. For any geodesic triangle in $\mathcal{T}(S)$ whose three sides have $\text{Thk}_{\epsilon}^C \geq \theta$, at least proportion $\theta'$ of each side of the triangle is contained within $\delta$ of the union of the other two sides.

We now give the proof of Theorem A.

Proof of Theorem A. We will find $L$ so that the conclusion of the theorem applies to any subinterval $I$ with $L \leq \text{length}(I) \leq 2L$ and $\text{Thk}_{\epsilon}^C(I) \geq \theta$. This will suffice because any long interval with $\text{Thk}_{\epsilon}^C \geq \theta$ can be partitioned into subintervals satisfying this length condition, one of which must have $\text{Thk}_{\epsilon}^C \geq \theta$. 

Recall that the (coarsely defined) projection \( \pi_S : \mathcal{T}(S) \to \mathcal{C}(S) \) sends a point \( w \in \mathcal{T}(S) \) to a choice of simple closed curve in the Bers marking \( \mu_w \). (The set of choices lies in a set of diameter at most 2 in the curve complex). The work of Masur–Minsky [17] shows that there are universal constants \( K, C \) so that Teichmüller geodesics project to (unparametrized) \((K,C)\)-quasi-geodesics under \( \pi_S \). By the hyperbolicity of \( \mathcal{C}(S) \) (see §2.2), each quasi-geodesic fellow travels any geodesic with the same endpoints, and so there is a constant \( \tau > 0 \) so that every \((K,C)\)-quasi-triangle in \( \mathcal{C}(S) \) is \( \tau \)-thin.

By Theorem 3.9 the interval \( I \) moves a definite amount in \( \mathcal{C}(S) \); that is, by making \( L \) large, we can arrange for \( I \) to project to an arbitrarily long subsegment of the (unparametrized) quasi-geodesic \( \pi_S([x,y]) \). In particular, by choosing \( L \) sufficiently large, we can ensure that there is a point \( w \in I \) so that either

\[
d_S(\mu_w, \pi_S([y,z])) \geq 2\tau + 6 \quad \text{or} \quad d_S(\mu_w, \pi_S([x,z])) \geq 2\tau + 6.
\]

By choosing such \( w \) with \( \mu_w \) near the center of \( \pi_S(I) \), we can moreover ensure that

\[
d_S(w, t) \geq 2\tau + 6
\]

for all points \( t \in [x,y] \) outside of \( I \).

Assuming without loss of generality that

\[
d_S(\mu_w, \pi_S([y,z])) \geq 2\tau + 6 > \tau,
\]

hyperbolicity implies that there is a point \( u \in [x,z] \) so that \( d_S(w, u) \leq \tau \). We will show that \( d_V(w, u) \leq 2\epsilon, \theta \leq 1 \) for all proper subsurfaces \( V \subset S \). The result will then follow from the distance formula (2.5). We first establish the following

**Claim 3.15.** There is a constant \( M_0 \) (depending only on \( \epsilon \) and \( \theta \)) so that for any proper subsurface \( V \subset S \) satisfying \( d_S(\partial V, \mu_w) \leq 2\tau + 3 \) we have

\[
d_V(x,y), d_V(y,z), d_V(x,z) \leq M_0.
\]

To see this, first observe that for any such \( V \) the triangle inequality implies

\[
d_S(\partial V, \pi_S([y,z])) \geq 3.
\]

Therefore \( V \) does not become thin along \([y,z]\) and so we may conclude \( d_V(y,z) \leq M \). If \( V \) does not become thin along \([x,y]\), then we have the same bound on \( d_V(x,y) \). However, \( V \) may become thin along \([x,y]\) in which case there is a point \( t \in [x,y] \) at which the length of \( \partial V \) is smaller than \( \epsilon_0 \). Therefore \( \mu_t \) contains \( \partial V \), which implies

\[
d_S(w, t) \leq 2\tau + 5 < 2\tau + 6
\]

and consequently that \( t \in I \subset [x,y] \). Thus the entire active interval \( I_V \) for \( V \) is contained within \( I \) and in particular has length at most \( \text{length}(I) \leq 2L \). Since the projection \( \pi_V \) to \( \mathcal{C}(V) \) is a Lipschitz map and, up to an additive error, the projection of \([x,y]\) to \( \mathcal{C}(V) \) can only change in the active interval \( I_V \) (see §2.6), we conclude that \( d_V(x,y) \) is bounded in terms of \( L \) (and \( L \) depends only on \( \epsilon \) and \( \theta \)). Finally, the triangle inequality and the above bounds on \( d_V(y,z) \) and \( d_V(x,y) \) together provide a uniform bound on \( d_V(x,z) \). This completes the proof of Claim 3.15.

We now show that \( d_V(w, u) \) is uniformly bounded for all proper subsurfaces. Consider any \( V \) with \( d_V(w, u) \geq M \). Then \( V \) becomes thin along \([w,u]\) and so \( \partial V \) lies within distance \( \tau + 2 \) of the \( \mathcal{C}(S) \)-geodesic from \( \mu_w \) to \( \mu_u \). In particular

\[
d_V(\partial V, \mu_w) \leq \tau + 2 + d_S(w, u) \leq 2\tau + 2
\]

and so Claim 3.15 implies that \( d_V(x,y) \) and \( d_V(x,z) \) are at most \( M_0 \).
If $V$ is a non-annular surface, then the reverse triangle inequality, applied to $[x, y]$ and $[x, z]$, yields bounds on $d_V(x, w)$ and $d_V(x, u)$ so that we may bound $d_V(w, u)$ by the triangle inequality.

If $V$ is an annulus we instead appeal to Theorem 2.14 (R.T.I. exception): According to that theorem applied to $2M_0 + B$, there is a constant $B'$ such that if
\begin{equation}
    d_V(x, w) + d_V(w, y) - d_V(x, y) \geq B',
\end{equation}
then there exists a collection of subsurfaces $W_j$ disjoint from $V$ that each satisfy $d_{W_j}(x, w) \geq 2M_0 + B$. But this is not possible: disjointness ensures that $d_S(\partial W_j, \mu_w) \leq d_S(\partial W_j, \partial V) + d_S(\partial V, \mu_w) \leq 2\tau + 3$, and thus Claim 3.15 implies $d_{W_j}(x, y) \leq M_0$; by the reverse triangle inequality, this contradicts $d_{W_j}(x, w) \geq 2M_0 + B$. Therefore the inequality (3.16) is false and we have $d_V(x, w) \leq B' + d_V(x, y) \leq B' + M_0$.

The same argument yields a bound on $d_V(x, u)$, and so the triangle inequality gives the desired bound on $d_V(w, u)$. □

4. Comparing measures

To address genericity and 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 4.1** (Hausdorff measure). The $n$–dimensional Hausdorff measure on a metric space will be denoted by $\eta$. It is defined by
\[
    \eta(E) := \lim_{\delta \to 0} \left[ \inf \sum \text{diam}(U_i)^n \right],
\]
where the infimum is over countable covers $\{U_i\}$ of $E$ with $\text{diam} U_i < \delta \forall i$.

For the Teichmüller metric, there is a nontrivial $h$–dimensional Hausdorff measure (recalling that $h = 6g - 6$). 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.

4.1. 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 $M = T(S)$. The treatment closely follows the survey by Álvarez and Thompson [1].

Recall that a Finsler metric on an $n$–dimensional Finsler manifold $M$ is a continuous function $F: T(M) \to \mathbb{R}$ 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
\begin{equation}
    \phi: T_x(M) \to \mathbb{R}^n \quad \text{and} \quad \psi: T^*_x(M) \to \mathbb{R}^n
\end{equation}
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

\[
\begin{align*}
  f(x) &= \frac{\varepsilon_n}{\lambda(\phi(B_x))} \\
  g(x) &= \frac{\lambda(\psi(B_x^*))}{\varepsilon_n},
\end{align*}
\]

where \( \lambda \) is Lebesgue measure and \( \varepsilon_n := \lambda(\text{Ball}^n) \) 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

\[
f(x) \ dx_1 \wedge \cdots \wedge dx_n \quad \text{and} \quad g(x) \ dx_1 \wedge \cdots \wedge dx_n
\]

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 \( \omega \) 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

\[
(x_1, y_1, \ldots, x_n, y_n) \mapsto (x_1, \ldots, x_n, \sum_{i=1}^n y_i \ dx_i).
\]

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 \( \mu_{sp} = \omega^n/n! \) on \( T^* M \). By restricting to the unit disk bundle \( T^* M \leq (M) \) and pushing forward by the projection \( \pi: T^* M \to M \), we obtain a symplectic measure \( n \) 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 \( \Omega^\circ \subset (\mathbb{R}^n)^\star = \mathbb{R}^n \) via

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

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

\[
\frac{\varepsilon_n^2}{n^{n/2}} \leq M(\Omega) \leq \varepsilon_n^2 = M(\text{Ball}^n).
\]

The first inequality was established by John [10], and the latter, which gives an equality if and only if the norm is Euclidean, is known as the Blaschke–Santaló inequality [4].

**Theorem 4.5** (Assembling facts on Finsler measures). Suppose that \( M^n \) is a continuous Finsler manifold. Then

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

Note that it is still possible for $\mu_{\text{HT}}$ and $\mu_n$ 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 [5] 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:

$$
\int_E g(x) \, dx_1 \wedge \cdots \wedge dx_n = \int_E \left( \int_{\psi(B^*_x)} \frac{1}{\varepsilon_n} \, d\lambda \right) \, dx_1 \wedge \cdots \wedge dx_n
$$

$$
= \frac{1}{\varepsilon_n} \int_{\pi^{-1}(E) \cap T^* \cdot \leq 1(M)} dy_1 \wedge \cdots \wedge dy_n \wedge dx_1 \wedge \cdots \wedge dx_n
$$

$$
= \frac{1}{\varepsilon_n} \mathbf{n}(E).
$$

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

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

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 $\phi$ defined in (4.2). 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 (4.4) now implies that $n^{-n/2}f(x) \leq g(x) \leq f(x)$ for all $x \in M$. We conclude that $\mu_{\text{HT}}(E) \leq \mu_n(E) \leq n^{n/2} \mu_{\text{HT}}(E)$ for all $E \subset M$. Finally, since Blaschke–Santaló can only give equality for a Euclidean norm, it follows that $\mu_n$ and $\mu_{\text{HT}}$ can only be equal for a Riemannian metric. □

### 4.2. Measures coming from quadratic differentials

Recall that quadratic differential space $Q(S)$ is naturally identified with the cotangent bundle $T^*(\mathcal{T}(S))$ of Teichmüller space, and that each quadratic differential $q \in Q(S)$ 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

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

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

The space $Q(S)$ also carries a natural $\text{Mod}(S)$-invariant measure $\mu_{\text{hol}}$ 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 [14] for details). This measure has been studied extensively, for instance to establish ergodicity results for the geodesic flow. The measure $\mu_{\text{hol}}$ is also related to the “Thurston measure” $\mu_{\text{TH}}$ on the space of measured foliations $\mathcal{MF}$ induced by the piecewise-linear structure of $\mathcal{MF}$ [8]. Indeed, as seen in [14], $\mu_{\text{hol}}$ is equal to the pullback of $\mu_{\text{TH}} \times \mu_{\text{TH}}$ under the $\text{Mod}(S)$-invariant map $Q(S) \to \mathcal{MF} \times \mathcal{MF}$ that sends a quadratic differential to its vertical and horizontal foliations.
Just as \( \mu_{wp} \) induces \( n \), the holonomy measure \( \mu_{hol} \) descends to a measure \( m \) on \( T(S) \). Explicitly, the \( m \)-measure of a set \( E \subset T(S) \) is given by
\[
m(E) := \mu_{hol} \left( \pi^{-1}(E) \cap Q^{\leq 1}(S) \right).
\]
This measure \( m \) has been studied previously in [3] and [7].

**Proposition 4.6.** [16, p.3746] There is a scalar \( k > 0 \) such that \( \mu_{wp} = k \cdot \mu_{hol} \).

We recall the outline of the argument here. In [16], it was shown that the Teichmüller geodesic flow on \( Q(S) \) is a Hamiltonian flow for the function
\[
H(q) = \frac{\|q\|^2}{2}.
\]
As such, the Teichmüller flow preserves the symplectic form \( \omega \) and the corresponding measure \( \mu_{wp} \). The measures \( \mu_{wp} \) and \( \mu_{hol} \) both descend to the quotient space \( Q(S)/\text{Mod}(S) \); furthermore, the latter defines an ergodic measure for the Teichmüller flow on \( Q(S)/\text{Mod}(S) \) [14]. Since \( \mu_{wp} \) is absolutely continuous with respect to \( \mu_{hol} \), the proposition follows.

We therefore also have \( n = km \), and combining Proposition 4.6 with Theorem 4.5 we get:

**Corollary 4.7.** There are scalars \( k_2 > k_1 > 0 \) such that
\[
k_1m \leq \eta \leq k_2m.
\]

### 4.3. Visual measures.

The unit sphere subbundle of \( Q(S) \) will be denoted by
\[
Q^1(S) = \{ q \in Q(S) : \|q\| = 1 \}.
\]
For each \( x \in T(S) \), the fiber \( Q^1(x) \) is identified with the “space of directions” at \( x \), and the Teichmüller geodesic flow \( \varphi_t : Q(S) \to Q(S) \) gives rise to a homeomorphism
\[
\Psi_x : Q^1(x) \times (0, \infty) \to T(S) \setminus \{x\}
\]
\[
(q, r) \mapsto \pi(\varphi_r(q)),
\]
which serves as “polar coordinates” centered at \( x \). Furthermore, this conjugates \( \varphi_t \) 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 \( T(S) \) that are compatible with these polar coordinates and with the radial flow.

**Definition 4.8** (Visual measure). Given any measure \( \kappa_x \) on the unit sphere \( Q^1(x) \cong S^{h-1} \), we define the corresponding visual measures on \( S_r(x) \) and \( T(S) \) as follows.

Firstly, the visual measure \( \text{Vis}_r(\kappa_x) \) on the sphere \( S_r(x) \) of radius \( r \) is just the push-forward of \( e^{hr}\kappa_x \) under the homeomorphism \( Q^1(x) \times \{r\} \cong S_r(x) \). Integrating these over \( (0, \infty) \) then gives a visual measure on \( T(S) \) defined by
\[
\text{Vis}(\kappa_x)(E) := \int_{(q,r) \in E \subset Q^1(S) \times (0, \infty)} e^{hr}d\kappa_x(q)d\lambda(r).
\]
Said differently, \( \text{Vis}(\kappa_x) \) is equal to the push-forward of \( \kappa_x \times \lambda_0 \) under the homeomorphism \( \Psi_x \), where \( \lambda_0 \) 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
\[
\frac{\text{Vis}_{r+t}(\kappa_x)(\hat{\varphi}_t(E))}{\text{Vis}_{r+t}(\kappa_x)(\mathcal{S}_r(x))} = \frac{\text{Vis}_r(\kappa_x)(E)}{\text{Vis}_r(\kappa_x)(\mathcal{S}_r(x))}.
\]
The same invariance holds for $\text{Vis}(\kappa_x)$ 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 $Q(x)$ carries a unique translation-invariant measure $\nu_x$ normalized so that $\nu_x(B^*_x) = 1$; recall that the unit ball $B^*_x$ is just the intersection $Q^{\leq 1}(S) \cap Q(x)$. This induces a measure (also denoted $\nu_x$) on the unit sphere $Q^1(x)$ via the usual method of coning off: $\nu_x(E) := \nu_x([0, 1] \times E)$ for $E \subset Q^1(x)$.

Secondly, since $Q(S)$ has the structure of a fiber bundle over $T(S)$, we can define a conditional measure $s_x$ on $Q(x)$ by disintegration from $\mu_{\text{hol}}$. More precisely, $s_x$ is the unique measure on $Q(x)$ such that the $\mu_{\text{hol}}$-measure of $E \subset Q(S)$ is given by
\[
\mu_{\text{hol}}(E) = \int_{T(S)} s_x(E \cap Q(x)) \, dm(x).
\]

Via the process of coning off, we again think of $s_x$ as a measure on $Q^1(x)$.

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

4.4. Summary. The measures on $T(S)$ considered above are $\mathbf{n}$ and $\mathbf{m}$ (induced by the symplectic and holonomy measures on $Q(S)$, respectively, via the covering map), Hausdorff measure $\eta$, the visual measures $\text{Vis}(\kappa_x)$ created by radially flowing measures on the sphere of directions $Q^1(x)$, and the measures $\mu_{\text{hol}}$ and $\mu_{\text{hol,y}}$ coming from the Finsler structure.

We found that $\mathbf{n}$, $\mathbf{m}$, and $\mu_{\text{hol,y}}$ are scalar multiples of each other, Hausdorff measure and Busemann measure coincide, and all five of these are mutually comparable in the sense of being bounded above and below by scalar multiples of each other. In the following section we will establish results about the structure of generic geodesic rays with respect to these measures and the visual measures.

5. Thickness statistics for geodesic rays

In §3 we studied the behavior of Teichmüller geodesics that spend a definite fraction of their time in some thick part $T_r$. In this section we will show that most Teichmüller geodesics in fact satisfy this property. Therefore the tools developed in §3 apply generically and we may use them in studying averaging questions such as Theorems B and C.

In all of what follows, if $x$ is a fixed basepoint and $y \in B_r(x)$ is a point in the ball centered at $x$, then we will write $y_t$ to denote the time–$t$ point on the geodesic ray based at $x$ and traveling through $y$. 

5.1. Volume estimates. We begin by recalling some estimates on the volume of Teichmüller balls and using these to reduce to the case of annular shells.

Athreya, Bufetov, Eskin, and Mirzakhani [3] have found the following asymptotic estimate for the $m$–volume of a ball of radius $r$.

**Theorem 5.1** (Volume asymptotics [3, Theorem 1.3]). There is a (bounded) function $f : \mathcal{T}(S) \to (0, \infty)$ such that for each $x \in \mathcal{T}(S)$

$$\lim_{r \to \infty} \frac{m(B_r(x))}{e^{hr}} = f(x).$$

**Corollary 5.2** (Definite exponential growth). Let $\mu$ denote Hausdorff measure $\eta$, holonomy measure $m$, or any visual measure $\mu_x = \text{Vis}(\kappa_x)$. For each $x \in \mathcal{T}(S)$, there exist constants $C_1 \leq C_2$ such that for all sufficiently large $r$ (depending on $x$) we have

$$C_1 e^{hr} \leq \mu(B_r(x)) \leq C_2 e^{hr}.$$

**Proof.** This is built into the definition of the visual measure $\text{Vis}(\kappa_x)$. For the holonomy measure $m$, this follows from Theorem 5.1 above. The same estimate then holds for $\eta$ by Proposition 4.6 and Corollary 4.7. □

Of course, this holds for the other measures discussed in this paper as well by the comparisons in the last section.

For any $r > k > 0$, let $A_k^r(x) = B_r(x) \setminus B_r(-k)(x)$ denote the annular shell between radii $r$ and $r - k$. The fact that the volume of a ball grows exponentially in the radius means that we can focus our attention on annuli rather than on balls.

**Lemma 5.3** (Reduction to annuli). Fix $x \in \mathcal{T}(S)$ and let $\mu$ be any measure with definite exponential growth (i.e., satisfying the conclusion of Corollary 5.2). Suppose that for all $k > 0$ we have

$$\lim_{r \to \infty} \frac{1}{r \mu(A_k^r(x))} \int_{A_k^r(x) \times A_k^r(x)} d\tau(y, z) \, d\mu(y) d\mu(z) = 2.$$

Then the same holds when $A_k^r(x)$ is replaced by $B_r(x)$.

**Proof.** Let $C_1, C_2$ be as in Corollary 5.2 above. For each $k$ sufficiently large (satisfying $C_1 e^{-hk} < 1$) and all sufficiently large $r$ we have

$$2 \geq \frac{1}{r \mu(B_r(x))^2} \int_{B_r(x) \times B_r(x)} d\tau(y, z) \, d\mu(y) d\mu(z) \geq \left( \frac{\mu(B_r(x))}{\mu(B_r(x))} \right)^2 \frac{1}{r \mu(A_k^r(x))^2} \int_{A_k^r(x) \times A_k^r(x)} d\tau(y, z) \, d\mu(y) d\mu(z) \geq \left( 1 - \frac{C_2}{C_1} e^{-hk} \right)^2 \frac{1}{r \mu(A_k^r(x))^2} \int_{A_k^r(x) \times A_k^r(x)} d\tau(y, z) \, d\mu(y) d\mu(z).$$

The claim now follows since, by assumption, the latter becomes arbitrarily close to 2 when $r$ and $k$ are sufficiently large. □

5.2. The thickness property. Recall from §3 that the thick-stat of a nondegenerate geodesic $[x, y] \subset \mathcal{T}(S)$ is defined by

$$\text{Thk}^m_
u [x, y] = \frac{\left| \{ 0 \leq s \leq d\tau(x, y) : y_s \in \mathcal{T}_\nu \} \right|}{d\tau(x, y)}.$$
holds for all geodesics into sample paths of a random walk. We begin the setup by combining considera-
ably more involved. The proof uses ideas of Eskin and Mirzakhani on discretizing
5.3. Random walks and thickness. We next verify (P1) for \( m \), which is considerably more involved. The proof uses ideas of Eskin and Mirzakhani on discretizing geodesics into sample paths of a random walk. We begin the setup by combining some results on the volume of balls from Athreya–Bufetov–Eskin–Mirzakhani [3] and Eskin–Mirzakhani [7].
Lemma 5.7 (Volume of balls [3, Theorem 1.2], [7, Lemma 3.1]). There exists a constant $c > 0$ (depending only on the topology of $S$) such that $\mathfrak{m}(B_\epsilon(y)) \asymp 1$ for all $y \in \mathcal{T}(S)$. Additionally, given $\epsilon > 0$ we have $\mathfrak{m}(B_\epsilon(y)) \preceq e^{\epsilon r}$ for all $y \in \mathcal{T}_\epsilon$.

Proof. The first claim is exactly Lemma 3.1 of [7]. For the second claim, choose a point $x \in \mathcal{T}_\epsilon$. Note that our choice of $x$ depends only on $\epsilon$. By the volume asymptotics (Theorem 5.1), there exists constant $R_0$ such that

$$\mathfrak{m}(B_r(x)) \preceq e^{\epsilon r}$$

for all $r \geq R_0$. Furthermore, by increasing $R_0$ if necessary, we may assume that the $\text{Mod}(S)$–translates of $B_{R_0}(x)$ cover $\mathcal{T}_\epsilon$. It follows that for any $y \in \mathcal{T}_\epsilon$ and any $r \geq 0$ we have

$$\mathfrak{m}(B_r(y)) \leq \mathfrak{m}(B_{r+R_0}(x')) = \mathfrak{m}(B_{r+R_0}(x)) \preceq e^{\epsilon R_0} e^{\epsilon r}$$

for some $\text{Mod}(S)$–translate $x'$ of $x$. This establishes the second claim. □

To define a random walk on $\mathcal{T}(S)$ with basepoint $x$, first choose a net $\mathcal{N}$ of points in $\mathcal{T}(S)$, choosing so that $x \in \mathcal{N}$ and such that the net points are $c$–separated and $(2c)$–dense (i.e., the distances between net points are at least $c$ but the $(2c)$–balls about net points cover Teichmüller space). Here $c$ is the constant from Lemma 5.7, depending only on the topology of $S$.

Given a parameter $\tau$, a sample path of length $s$ (starting at $x$) is a map

$$\lambda: \{0, \ldots, \lfloor s/\tau \rfloor \} \to \mathcal{N}$$

such that $\lambda(0) = x$ and for each index, $d_\tau(\lambda(k), \lambda(k+1)) \leq \tau$. Let $P_\tau(s)$ be the set of sample paths $\lambda$ of length at most $s$, and let $P_\tau$ be the set of all sample paths of any length. By (36) of [7], for any $\delta > 0$ and sufficiently large $\tau$ (depending on $\delta$) we have,

$$|P_\tau(s)| \leq e^{s(h+\delta)}$$

for all $s \geq 0$. (Note that the constant $C_2$ in (36) of [7], coming from [7, Proposition 4.5], can be taken to equal 1.)

Now given $\tau$ we define a map $F_\tau: \mathcal{T}(S) \to P_\tau$ which takes a point $y$ and “discretizes” the geodesic $[x, y]$ to a sample path. For any $[x, y]$ we mark off points along the geodesic starting at $x$ and spaced by time $\tau - 2c$. For each such point along the geodesic we choose a nearest point in $\mathcal{N}$. This is the sample path $F_\tau(y)$ associated to $y$. For any $\epsilon_1 > 0$, we can choose $\tau$ sufficiently large so that the image under $F_\tau$ of the ball $B_r(x)$ is contained in $P_\tau(r(1+\epsilon_1))$. Furthermore, after increasing $\tau$ if necessary, the above estimate on the cardinality of $P_\tau(s)$ shows that we additionally have

$$|P_\tau(r(1+\epsilon_1))| \leq e^{hr(1+2\epsilon_1)}$$

for all $r$. We are now ready to show that the measure of points determining a ray with a too-large thick-stat decays exponentially in $r$.

Theorem 5.8 (Thickness estimate for holonomy measure). For all $0 < \theta, \sigma < 1$, $x \in \mathcal{T}(S)$, and $k > 0$, there exist $\epsilon = \epsilon(\theta) > 0$ and $\alpha > 0$ such that

$$\mathfrak{m} \left( \left\{ y \in A^k_\tau(x) : \text{Thk}^k_\tau[x, y_t] < \theta \text{ for some } t \in [\sigma r, r] \right\} \right) < e^{-\alpha r} \mathfrak{m}(A^k_\tau(x))$$

for all sufficiently large $r$, where $y_t$ is the time-$t$ point on the geodesic ray from $x$ through $y$. 

Proof. We start with a geodesic \([x, y]\) for \(y \in A_k^b(x)\). As in the above discussion, for any \(\epsilon_1\), we can choose \(\tau\) large enough so that the geodesic determines a sample path \(\lambda\) in \(P_\tau(r(1+\epsilon_1))\) and so that \(|P_\tau(r(1+\epsilon_1))| \leq e^{hr(1+2\epsilon_1)}\). Let \(F_\tau : A_k^b(x) \rightarrow P_\tau\) be the map carrying \(y\) to the sample path associated to \([x, y]\). We quote the proof of Theorem 5.1 of [7] (which itself quotes [2]) to say that given \(\theta\), there exist \(\delta' > 0\) and \(\epsilon'\) such that for all large \(\tau\) we have the following: For each \(\tau \leq t \leq r\), the estimate

\[
\frac{1}{[t/\tau]} \left| \{ 1 \leq i \leq [t/\tau] : \lambda(i) \in T_\tau \} \right| \geq \theta
\]

holds for all but at most

\[
e^{-\delta t} |P_\tau(r(1+\epsilon_1))| \leq e^{-\delta t} e^{hr(1+2\epsilon_1)}
\]

of the sample paths \(\lambda \in P_\tau(r(1+\epsilon_1))\). This is the crucial ingredient that lets us control our thickness statistic. For the given \(\sigma\), we now choose \(\epsilon_1\) small enough (forcing \(\tau\) to be large) so that \(\kappa := \frac{\sigma}{hr} - 2\epsilon_1 > 0\). In particular, note that \(\tau\) and \(\delta'\) depend only on \(\theta\) and \(\sigma\). If we let \(\Omega \subset P_\tau(r(1+\epsilon_1))\) denote the union of these exceptional sample paths corresponding to any \(\sigma \tau \leq t \leq r\), it follows that

\[
|\Omega| \leq \sum_{k=0}^{\infty} e^{-\delta (kr+\delta r)} e^{hr(1+2\epsilon_1)} \lesssim \theta \sigma e^{-\delta \sigma r} e^{hr(1+2\epsilon_1)} = e^{hr(1-\kappa)}.
\]

We know that the ball of radius \(c\) centered at any point has \(m\)-measure \(O(1)\) by Lemma 5.7. Consequently, when \(r\) is sufficiently large we have

\[
m(F_\tau^{-1}(\Omega)) \lesssim |\Omega| \lesssim \theta \sigma e^{-\delta \sigma r} \lesssim m(A_k^b(x)) \cdot e^{-\sigma r},
\]

for a suitable choice of \(\alpha\). We conclude that for all \(y \in A_k^b(x)\) except for a set of at most this measure, the sample path associated to the geodesic from \(x\) to \(y\) has \(\text{Thk}_y^b \geq \theta\) for ending times \(t \geq \sigma r\). Now every point on the geodesic is within distance \(\tau\) of a point on the sample path. Let \(\epsilon = \epsilon' e^{-\tau}\). A point at distance at most \(\tau\) from a point in \(T_\tau\) lies in \(T_\tau\). This concludes the proof.

This says that, except for set of endpoints \(y\) of exponentially small measure, geodesics have the property that they eventually have spent a definite fraction of their time in the thick part. In particular, this implies the following.

**Corollary 5.9.** The measures \(m\) and \(\eta\) satisfy the thickness property (P1).

We will again use random walks to now show that a typical geodesic has a long interval where it stays in the thick part. Specifically we say that a geodesic segment \([x, y]\) contains an \(\epsilon\)-thick interval \(I\) if there is a geodesic subsegment \(I\) along \([x, y]\) such that \(I \subset T_\tau\).

**Theorem 5.10** (Thick intervals). For all \(0 < \sigma, \theta < 1, M > 0\) and sufficiently small \(\epsilon\), there exists \(\beta > 0\) such that for all sufficiently large \(r\),

\[
m \left( \{ y \in A_k^b(x) : [y_{\sigma r}, y_{2\sigma r}] \ contains \ no \ \epsilon\text{-thick interval of length} \ M \} \right) < e^{-\beta r},
\]

where \(y_t\) is the time–\(t\) point on the geodesic ray from \(x\) through \(y\).

**Proof.** Let \(\tau\) be fixed as in the last theorem. Given a small \(\epsilon > 0\), let \(\epsilon_2 = \epsilon e\), so that if a point is \(\epsilon\)-thin, then the \(\tau\)-ball centered at that point is \(\epsilon_2\)-thin. We first claim that there exists \(\rho > 0\) such that for any net point \(y \in T_\tau\), the probability that the next step in the random walk starting at \(y\) remains in \(T_\tau\) is at least \(\rho\).
By Lemma 5.7 the volume of a ball $B_r(x)$ centered at $x \in T_{\epsilon_2}$ is $\approx \epsilon_2^{hr}$. The number of net points in the ball is also $\approx \epsilon_2^{hr}$. By [3] the number of orbit points of $x$ under the action of the mapping class group that are contained in the ball is also $\approx \epsilon_2^{hr}$, so the number of net points contained in $T_{\epsilon_2}$ is $\approx \epsilon_2^{hr}$. Thus the ratio of net points contained in $T_{\epsilon_2}$ to the total number of net points is bounded away from 0, proving the claim.

Let $\kappa := 1 - \rho^{M/\tau} < 1$. Then given a point $y \in T_{\epsilon_2}$, the probability that in the next $[M/\tau]$ steps in the random walk (that is, the sample path of length $M$) starting at $y$ at least one of the points is $\epsilon_2$–thick is at most $\kappa$.

Consider a sample path of length $L$ (so having $[L/\tau]$ points) for which at least $\theta$ proportion of its points are in $T_{\epsilon_2}$. Divide the path into subpaths with $[M/\tau]$ points, so that the number of pieces is additively close to $L/M$ if $M \gg \tau$. Let $J_1$ be the collection of odd-index pieces and $J_2$ the collection of even-index pieces. Suppose without loss of generality that $J_1$ contains at least as many pieces with an $\epsilon_2$–thick point as $J_2$ does. The number of $J_1$ pieces containing a point in $T_{\epsilon_2}$ is then at least $N := \frac{Ld}{2M}$. The probability that the sample path of length $M$ starting at a thick point enters the thin part is at most $\kappa$, as we have seen. If two thick starting points are in different $J_1$ pieces, then these events are independent, because between the two points is a $J_2$ piece of length $M$. Thus the probability that the random path does not have any pieces of length $M$ that lie entirely in $T_{\epsilon_2}$ is at most $\kappa^N$.

Now as in the discussion in the previous theorem, we have the map $F_\tau : A^k_\epsilon(x) \rightarrow P_\tau(r(1+\epsilon_1))$. Let $y \in A^k_\epsilon(x)$. We consider the segment $[y_\sigma r, y_{2r}] \subset [x, y]$. Its image under $F_\tau$ is a path of length $L = \sigma r$. The probability that this random path fails to have the desired segment of length $M$ in the $\epsilon_2$–thick part is at most $\kappa^{\sigma r/2M}$, by the statement at the end of the last paragraph. Since $\kappa$ is a fixed number smaller than 1, we can upper-bound this proportion by $e^{-\beta r}$ for some $\beta > 0$. As we saw in the proof of the last theorem, since this property holds for an exponentially small proportion of sample paths, the corresponding property holds for an exponentially small proportion of endpoints $y \in A^k_\epsilon(x)$.

Note that if the geodesic $[x, y]$ enters the $\epsilon$–thin part, then the associated random path enters the $\epsilon_2$–thin part. We conclude, as in the last theorem, that the measure of the set of points $y \in A^k_\epsilon(x)$ such that the geodesic $[x, y]$ does not have an $\epsilon$–thick interval of length $M$ in $[\sigma r, 2\sigma r]$ is at most $e^{-\beta r}$. \hfill \Box

6. Separation statistics for pairs of rays

We need one final ingredient before proving Theorems B and C. Namely, in order to apply Theorem A to show that the geodesic $[y, z]$ connecting two generic points $y, z \in B_r(x)$ must “dip back” towards $x$, we must first know that $[x, y]$ and $[x, z]$ become $C$–separated, where $C$ is the constant from Theorem A. Thus we need an estimate for the probability that two geodesic rays based at $x$ fellow-travel past a given radius. The appropriate sort of control is ensured by the following property.

Definition 6.1 (Property P2). We say a measure $\mu$ on $T(S)$ satisfies the separation property (P2) if for all $M_0, k > 0$, $0 < \sigma < 1$, and $x \in T(S)$, we have

$$\lim_{r \to \infty} \frac{\mu \times \mu \left( \{ (y, z) \in A^k_\epsilon(x) \times A^k_\epsilon(x) : d_T(y_t, z_t) \geq M_0 \text{ for all } t \in [\sigma r, r] \} \right)}{\mu \times \mu (A^k_\epsilon(x) \times A^k_\epsilon(x))} = 1.$$ 

We will also consider the following time-specific version this separation property:
Definition 6.2 (Property P3). We say that a measure $\mu$ on $T(S)$ satisfies the strong separation property (P3), or has exponential decay of fellow travelers, if for all $x \in T(S)$, $M_0, k > 0$, $0 < \sigma < 1$, there exist $\alpha, R_0 > 0$ such that

$$\mu \times \mu \left( \{ (y, z) \in A^k_t(x) \times A^k_t(x) : d_T(y_t, z_t) < M_0 \} \right) \leq e^{-\alpha t},$$

whenever $R_0 \leq \sigma r \leq t \leq r$.

Theorem 6.3 (Strong separation for visual measures). All rotation-invariant visual measures $\mu_x = \text{Vis}(\kappa_x)$ on $T(S)$, and in particular $\text{Vis}(\nu_x)$ and $\text{Vis}(s_x)$, have the strong separation property (P3).

Proof. Choose $\sigma r \leq t \leq r$, fix a point $y \in A^k_t(x)$, and let $E = \{ z \in A^k_t(x) : d_T(y_t, z_t) < M_0 \}$. Looking instead in the sphere $S_t(x)$, we have the set $E_t = \{ z \in S_t(x) : d_T(y_t, z) < M_0 \}$. Notice that, by definition,

$$E = \bigcup_{s \in [r-t-k, r-t]} \hat{\phi}_s(E_t).$$

(Recall that $\hat{\phi}_s$ denotes the radial geodesic flow based at $x$.) Therefore, by the normalized invariance, we have

$$\mu(E) = \int_{r-k}^{r} \text{Vis}_s(\kappa_x)(\hat{\phi}_{s-1}(E_t)) d\lambda(s)$$

$$= \int_{r-k}^{r} \text{Vis}_s(\kappa_x)(E_t) \frac{\text{Vis}_s(\kappa_x)(S_t(x))}{\text{Vis}_s(\kappa_x)(S_t(x))} d\lambda(s)$$

$$= \frac{\kappa_x(E_t)}{\kappa_x(Q^1_t(x))} \mu_x(A^k_t(x)),$$

where, in the last line, we have identified $E_t$ with its image in $Q^1_t(x) \cong S_t(x)$.

It remains to find $R_0$ (independent of $y$) such that $\kappa_x(E_t)/\kappa_x(Q^1_t(x)) \lesssim e^{-t}$ when $t \geq R_0$. Recall that $S^1$ acts freely on $Q^1(x)$ by rotations. Choosing orbit representatives, we may realize $Q^1_t(x)$ as a setwise product $(Q^1_t(x)/S^1) \times S^1$. The measure $\kappa_x$ pushes forward to a measure on $Q^1_t(x)/S^1$. By disintegration, we then obtain a measure on each fiber $S^1$ which, by the rotation-invariance of $\kappa_x$, must agree with a measure on a similar 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 $E_t$ is $\lesssim e^{-t}$. Using the product structure and integrating over the $Q^1_t(x)/S^1$ factor, Fubini’s theorem then implies that $\kappa_x(E_t)/\kappa_x(Q^1_t(x)) \lesssim e^{-t}$ as well. \hfill $\square$

Theorem 6.4 (Strong separation for holonomy measure). The measure $\mathbf{m}$ satisfies the strong separation property (P3).

Proof. For a given $r, t$, let

$$D_{r,t} = \{ (y, z) : d_T(y_t, z_t) < M_0 \} \subset A^k_t(x) \times A^k_t(x)$$

denote the set in question. Set $\theta = 1/2$ and choose $\epsilon$ and $\alpha > 0$ as in Theorem 5.8 so that the $m$-measure of the set
$$E_r = \{y : \text{Thk}_k^{\approx}[x, y] < \theta \text{ for some } t \in [\sigma r, r] \} \subset A_k^t(x)$$
is at most $m(A_k^t(x))e^{-\alpha t}$ for all large $r$. We may now write $D_{r,t} = D'_{r,t} \cup D''_{r,t}$, where
$$D'_{r,t} = \{(y, z) \in D_{r,t} : y \in E_r\} \quad \text{and} \quad D''_{r,t} = \{(y, z) \in D_{r,t} : y \notin E_r\}.$$By the above, we have that
$$m(D'_{r,t}) \leq m(E_r)m(A_k^t(x)) \leq e^{-\alpha t}$$for all large $r$ and all $t \leq r$.

Choose any point $y \in A_k^t(x) \setminus E_r$ and fix some $t \in [\sigma r, r]$. Then any point $z \in A_k^t(x)$ satisfying $d_T(y_t, z_t) < M_0$ must be contained in the ball of radius $r - t + M_0$ about $y_t$. Since $t \geq \sigma r$ and $y \notin E_r$, we know that $\text{Thk}_k^{\approx}[x, y_t] \geq \theta$.

Choosing any $0 < \delta < \theta$, there must exist a time $t' \in [\delta t, t]$ for which $y_{t'} \in T_e$. Furthermore, each such $z$ lies within the ball of radius $(1 - \delta)t + M_0 + r - t$ about $y_{t'}$. Applying Lemma 5.7, we see that
$$m(D''_{r,t}) \leq m(E_{r-\delta t+M_0}(y_{t'})) \leq e^{hM_0\epsilon_h r e^{-\delta h t}}.$$The Theorem now follows from (6.5), (6.6) and the fact that, by Corollary 5.2, we have $m(A_k^t(x)) \sim_{x,k} e^{hr}$ for all large $r$.

Finally, we see that for any measure enjoying exponential decay of fellow travelers, most pairs of geodesic rays are in fact never near each other beyond some threshold.

**Proposition 6.7.** The strong separation property (P3) implies the separation property (P2).

**Proof.** If $(y, z) \in A_k^t(x) \times A_k^t(x)$ does not lie in the set
$$E_r = \{(y, z) \in A_k^t(x) \times A_k^t(x) : d_T(y_t, z_t) \geq M_0 \text{ for all } t \in [\sigma r, r]\},$$then there is some $n \in \mathbb{N}$, $n \leq (1 - \sigma)r$, such that $d_T(y_{r+n}, z_{r+n}) < M_0 + 2$. Thus all such points are contained in the union of exceptional sets corresponding to the radii $\sigma r, \sigma r + 1, \ldots, \sigma r + |r - \sigma r|$. Using the exponential bound provided by property (P3), we see that for large $r$ the complement of $E_r$ has measure at most
$$\mu(A_k^t(x))^2 \left(e^{-\alpha \sigma r} + e^{-\alpha \sigma r+1} + \ldots + e^{-\alpha \sigma r+|r-\sigma r|}\right) \leq \mu(A_k^t(x))^2 \left(\frac{1}{1 - e^{-\sigma}}\right) e^{-\alpha \sigma r}.$$\qed

**Corollary 6.8.** The Hausdorff measure $\eta$, holonomy measure $m$, and the visual measures $\text{Vis}(\nu_x)$ and $\text{Vis}(s_x)$ all satisfy property (P2).

Thus we can conclude that after throwing out a subset of $A_k^t(x) \times A_k^t(x)$ of an arbitrarily small proportional measure, all pairs of geodesics stay separated by an arbitrarily chosen distance in Teichmüller space after a threshold time $\sigma r$ has elapsed.
7. Statistical hyperbolicity

We can now assemble our results to prove Theorems B and C.

**Theorem 7.1** (Annulus version of statistical hyperbolicity). Let \( \mu \) be any measure on \( T(S) \) satisfying the thickness property (P1) and the separation property (P2). Fix a basepoint \( x \in T(S) \) and an arbitrary \( k > 0 \). Then

\[
\lim_{r \to \infty} \frac{1}{r} \frac{1}{\mu(A^k_r(x))} \int_{A^k_r(x) \times A^k_r(x)} d\tau(y, z) \, d\mu(y) \, d\mu(z) = 2.
\]

We have shown that these hypotheses are satisfied by the standard visual measures \( \text{Vis}(\nu_x) \) and \( \text{Vis}(s_x) \), the holonomy measure \( \mathbf{m} \), and the Hausdorff measure \( \eta \) (Corollaries 5.6, 5.9, 6.8), and thus also all the other measures considered in this paper. Therefore, combining Theorem 7.1 with Lemma 5.3 (Reduction to annuli) we immediately obtain Theorem B:

**Theorem B.** Let \( \mu \) denote the Hausdorff measure \( \eta \), holonomy measure \( \mathbf{m} \), or either standard visual measure \( \text{Vis}(\nu_x) \) or \( \text{Vis}(s_x) \). Then for every point \( x \in T(S) \),

\[
\lim_{r \to \infty} \frac{1}{r} \frac{1}{\mu(B_r(x))} \int_{B_r(x) \times B_r(x)} d\tau(y, z) \, d\mu(y) \, d\mu(z) = 2.
\]

**Proof of Theorem 7.1.** Set \( \theta = \frac{3}{4} \) and let \( \epsilon > 0 \) be the corresponding thickness parameter guaranteed by Property (P1). For this \( \epsilon \) and \( \theta' = \frac{1}{2} \), let \( C \) and \( L \) be the corresponding constants provided by Theorem A. For any \( 0 < \delta, \sigma < 1/3 \), properties (P1) and (P2) together imply that for all large \( r \), we may restrict to a subset \( E_r \subset A^k_r(x) \times A^k_r(x) \) whose complement has proportional \( \mu \)-measure at most \( \delta \) and such that all pairs \((y, z) \in E_r \) satisfy

\[
\text{Thk}^\#_\theta[x, y_t], \text{Thk}^\#_\theta[x, z_t] \geq \theta \quad \text{and} \quad d\tau(y_t, z_t) \geq 3C
\]

for all \( t \in [\sigma r, r] \), where \( y_t \) and \( z_t \) are the time-\( t \) points on the geodesic rays from \( x \) through \( y \) and \( z \), respectively. Notice that, in this case, the point \( y_t \) cannot be within \( C \) of any point on \([x, z] \) (by the triangle inequality, any point on \([x, z] \) within \( C \) of \( y_t \) must lie in \([z_t - C, z_t + C]\)).

We now let \( t = 2\sigma r \). Since \( \text{Thk}^\#_\theta[x, y_{2\sigma r}] \geq \theta = \frac{3}{4} \), it follows that the interval \( I_r = [y_{\sigma r}, y_{2\sigma r}] \) satisfies \( \text{Thk}^\#_{\theta'} I_r \geq \frac{1}{2} = \theta' \) for all large \( r \). We also have length(\( I_r \)) = \( \sigma r \geq L \) when \( r \) is large, and \( \sigma < 1/3 \) ensures \( I_r \subset [x, y] \). Since \( I_r \cap \text{Nbhd}_C((y, z)) = \emptyset \), as noted above, Theorem A now implies that \( I_r \cap \text{Nbhd}_C([y, z]) \neq \emptyset \). Therefore \([y, z] \) contains a point in the ball \( B_{2\sigma r + C}(x) \), and so we conclude that

\[
d\tau(y, z) \geq 2(r - k - (2\sigma r + C))
\]

for all \((y, z) \in E_r \). Putting the above estimates together, we find that

\[
\lim_{r \to \infty} \inf_{r \to \infty} \frac{1}{r} \frac{1}{\mu(A^k_r(x))} \int_{A^k_r(x) \times A^k_r(x)} d\tau(y, z) \, d\mu(y) \, d\mu(z)
\]

\[
\geq \lim_{r \to \infty} \frac{1}{r} \frac{1}{r}(1 - \delta)(2r - 2k - 4\sigma r - 2C) = (1 - \delta)(2 - 4\sigma).
\]

Since \( \delta \) and \( \sigma \) can be chosen arbitrarily small, the result follows.

**Remark 7.2.** One could give an alternate proof of Theorem B that does not rely on Theorem A, but rather on Rafi’s Theorem 3.12, which reaches the same conclusion under the stronger hypothesis that the interval \( I \) lies entirely in the \( \epsilon \)-thick part.
That is: as we show in §5.2, a sufficiently long geodesic has a totally thick subinterval of definite length with arbitrarily high probability. (This holds in spite of the fact that the probability of such a subinterval occurring at any specified time is small). However, Theorem A generalizes Theorem 3.12 and thus may have further applications in the future.

We similarly obtain Theorem C:

**Theorem C.** For every point \(x \in \mathcal{T}(S)\) and either family \(\{\mu_r\}\) of standard visual measures \(\mu_r = \text{Vis}_r(\nu_x)\) or \(\text{Vis}_r(s_x)\) on the spheres \(S_r(x)\), we have
\[
E(\mathcal{T}(S), x, d_{\mathcal{T}}, \{\mu_r\}) = 2.
\]

**Proof.** Visual measures on \(\mathcal{T}(S)\) are constructed by radially integrating these visual measures on spheres. In fact, Proposition 5.5 and Theorem 6.3 were proved for annuli by first verifying them for spheres, and so analogous formulations of properties (P1) and (P2) also hold for the visual measures \(\{\mu_r\}\) on spheres. The result thus follows by the same argument used to prove Theorem 7.1 above. \(\square\)

**References**

[1] J.C. Álvarez Paiva, A. Thompson, *Volumes on normed spaces and Finsler spaces*, MSRI Publications 50 (2004).
[2] J.S. Athreya, Quantitative recurrence and large deviations for Teichmüller geodesic flow. Geom. Dedicata 119 (2006), 121–140.
[3] J. Athreya, A. Bufetov, A. Eskin, and M. Mirzakhani, *Lattice point asymptotics and volume growth on Teichmüller space*, Duke Math. J. 161 (2012), no. 6, 1055–1111.
[4] L. A. Santaló, *An affine invariant for convex bodies of n-dimensional space*, Portugaliae Math. 8 (1949), 155–161.
[5] H. Busemann, *Intrinsic area*, Ann. of Math. 48 (1947), 234-267.
[6] M. Duchin, S. Lelièvre, and C. Mooney, *Statistical hyperbolicity in groups*, Alg. and Geom. Top. 12 (2012), 1–18.
[7] A. Eskin and M. Mirzakhani, *Counting closed geodesics in moduli space*, J. Modern Dynamics 5 (2011), no. 1, 71–105
[8] A. Fathi, F. Laudenbach, P. Poenaru, *Travaux de Thurston sur les surfaces*, Asterisque 66-67 (1979).
[9] U. Hamenstädt, *Distance in the curve graph*, arxiv 1104.4001, to appear in Geom. Dedicata.
[10] 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. Neugebauer, and J. J. Stoker, Interscience, New York, 1948.
[11] R. Kent, C. Leininger, *Shadows of mapping class groups: capturing convex cocompactness*, Geom. Funct. Anal. 18 (2008), 1270-1325.
[12] H. Masur, *On a class of geodesics in Teichmüller space*, Ann. of Math. (2) 102 (1975), no. 2, 205–221.
[13] H. Masur, *Uniquely ergodic quadratic differentials*, Comment. Math. Helvetici 55 255–266 (1980).
[14] H. Masur, *Interval exchange transformations and measured foliations*, Annals of Math. 115 (1982) 169–200.
[15] H. Masur, *Hausdorff dimension of the set of nongeodetic foliations of a quadratic differential*, Duke Math. Journal 66 387–442 (1992).
[16] H. Masur, *The Teichmüller flow is Hamiltonian*, Proceedings of the American Mathematical Society 123 (1995) 12 3739–3747.
[17] H. Masur, and Y. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. 138 (1999), no. 1, 103–149.
[18] H. Masur and M. Wolf, *Teichmüller space is not Gromov hyperbolic*, Ann. Acad. Sci. Fenn. Ser. A I Math., 20(2) (1995), 259–267.
[19] Y.N. Minsky, *Extremal length estimates and product regions in Teichmüller space*, Duke Math. J. 83:2(1996), 249–286.
[20] K. Rafi, *A combinatorial model for the Teichmüller metric*, Geom. Funct. Anal. 17:3 (2007), 936–959.

[21] K. Rafi, *Hyperbolicity in Teichmüller space*, arxiv 1011.6004.

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

[23] 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