

CONVEX COCOMPACTNESS AND STABILITY IN MAPPING CLASS GROUPS

MATTHEW GENTRY DURHAM AND SAMUEL J. TAYLOR

ABSTRACT. We introduce a strong notion of quasiconvexity in finitely generated groups, which we call *stability*. Stability agrees with quasiconvexity in hyperbolic groups and is preserved under quasi-isometry for finitely generated groups. We show that the stable subgroups of mapping class groups are precisely the convex cocompact subgroups. This generalizes a well-known result of Behrstock [Beh06] and is related to questions asked by Farb-Mosher [FM02] and Farb [Far06].

1. INTRODUCTION

In order to understand the structure of a finitely generated group G , one often investigates subgroups $H \leq G$ whose geometry reflects that of G . One successful application of this approach is to the study of quasiconvex subgroups of hyperbolic groups. In this setting, H is finitely generated and undistorted in G and these properties are preserved under quasi-isometries of G . Quasiconvexity, however, is not as useful for arbitrary finitely generated groups. Without hyperbolicity of G , quasiconvexity depends on a choice of generating set for G and, in particular, is not preserved under quasi-isometry. To address this situation, we introduce the stronger notion of *stability*, which agrees with quasiconvexity when G is hyperbolic. Specifically, we define the following:

Definition 1. *Let G be a finitely generated group. A subgroup $H \leq G$ is stable if H is undistorted in G and for all $L \geq 0$ there exists an $R = R(L) \geq 0$ satisfying the following: for any pair of L -quasigeodesics of G that share common endpoints in H , each is contained in the R -neighborhood of the other.*

Our primary motivation for defining stable subgroups of a finitely generated group is the mapping class group $\text{Mod}(S)$ of a connected, orientable surface S . In this note, we relate stable subgroups of $\text{Mod}(S)$ to *convex cocompact subgroups* of the mapping class group, introduced by Farb and Mosher in [FM02]. These are much studied subgroups of $\text{Mod}(S)$ that have important connections to the geometry of Teichmüller space, the curve graph, and surface group extensions. (See Section 2.3 for definitions.) Our main result can be interpreted as a generalization of a theorem of Behrstock [Beh06] (also see [DMS10]). Behrstock proves that the stable (or Morse) elements of $\text{Mod}(S)$ are exactly the pseudo-Anosov mapping classes.

Our main theorem provides a purely group theoretic characterization of convex compactness, which does not involve the geometry of either Teichmüller space or the curve graph. This distinguishes our characterization of convex cocompact subgroups of mapping class groups from those appearing in [FM02, KL08, Ham05]. We prove

Theorem 1.1. *The subgroup $G \leq \text{Mod}(S)$ is stable if and only if it is convex cocompact.*

Theorem 1.1 partially answers questions appearing in [FM02] and [Far06]. In particular, Farb and Mosher ask how their notion of convex cocompactness (which they define as having quasiconvex orbit in Teichmüller space) is related to quasiconvexity in the mapping class group [FM02]. Also, in Problem 3.8 of [Far06], Farb asks what subgroups of mapping class groups are quasiconvex with respect to fixed generating sets. Theorem 1.1 characterizes the subgroups of the mapping class group that satisfy our strong notion of quasiconvexity and implies that convex cocompact subgroups are quasiconvex in $\text{Mod}(S)$ with respect to any generating set (Proposition 5.7). It is our hope that this notion of stability will also be useful in other finitely generated groups.

2. BACKGROUND

We keep this section brief and refer the reader to [BH09] for background on coarse geometry and [FM12] for mapping class group basics. Throughout, we assume that the reader has some familiarity with subsurface projections and hierarchies for the mapping class group, as introduced in [MM00]. See [Min10, Beh06] for additional references.

2.1. Coarse geometry. Let (X, d_X) and (Y, d_Y) be metric spaces. Recall that $f : X \rightarrow Y$ is a K -quasi-isometric embedding if for all $x_1, x_2 \in X$,

$$\frac{1}{K}d_X(x_1, x_2) - K \leq d_Y(f(x_1), f(x_2)) \leq Kd_X(x_1, x_2) + K.$$

We remark that what we have defined is usually called a (K, K) -quasi-isometric embedding in the literature, but our definition will reduce the number of constants appearing throughout this note. If $f : X \rightarrow Y$ has the additional property that every point in Y is within K of the image $f(X)$, then f is a K -quasi-isometry and X and Y are *quasi-isometric*.

If I is a subinterval of either \mathbb{R} or \mathbb{Z} , then a K -quasi-isometric embedding $f : I \rightarrow X$ is called a *K -quasigeodesic*. We will often refer to f as a quasigeodesic and call K the *quasigeodesic constant* for f . When f is an isometric embedding, it is called a *geodesic*. The metric space X is called *geodesic* if for any $x_1, x_2 \in X$ there is a geodesic $f : [0, N] \rightarrow X$ with $f(0) = x_1$ and $f(N) = x_2$, i.e. there is a geodesic joining x_1 to x_2 . We will sometimes write $[x_1, x_2]$ to denote an arbitrary geodesic joining x_1 and x_2 . For any path $\gamma : I \rightarrow X$, we will continue to use the symbol γ to denote the image of γ in X , as what is meant will be clear from context.

Recall that a subset C of a geodesic metric space X is K -*quasiconvex* if for any $c_1, c_2 \in C$ and any geodesic $[c_1, c_2]$ in X , $[c_1, c_2] \subset N_K(C)$. Here, $N_K(C)$ denotes the closed K -neighborhood of C . For any $\epsilon > 0$, two subsets A and B of X have *Hausdorff distance no greater than ϵ* if $A \subset N_\epsilon(B)$ and $B \subset N_\epsilon(A)$. The infimum over all such ϵ is the *Hausdorff distance* between A and B , denoted by $d_{\text{Haus}(X)}(A, B)$. Throughout this note, we reserve the notation $d(A, B)$ to denote the diameter of the union of A and B . In symbols, $d(A, B) = \text{diam}(A \cup B)$.

We make one further remark on notation. The expression $A < B$ is defined to mean that there exists a $K \geq 1$ so that $A \leq K \cdot B + K$. In different contexts the constant K will depend on particular parameters but not on the numbers A and B directly. We define $A > B$ similarly and write $A \asymp B$ if both $A < B$ and $B < A$. When using this notation below, we will be clear about the dependence of K .

2.2. Hyperbolic geometry. A geodesic metric space X is δ -hyperbolic if for any $x, y, z \in X$ and any geodesic segments $[x, y], [y, z], [x, z]$ joining them, $[x, z] \subset N_\delta([x, y] \cup [y, z])$. That is, X has δ -thin triangles. In this note, we will need a few well-known properties about the nearest point retraction from a hyperbolic metric space X to a quasigeodesic γ in X . See [BH09] for additional details.

Let $\gamma : [0, N] \rightarrow X$ be a K -quasigeodesic. The *nearest point retraction* from X to γ is a map $\mathbf{n} = \mathbf{n}_\gamma : X \rightarrow \text{im}(\gamma)$ defined as follows: for $x \in X$, $\mathbf{n}(x)$ is any point in the image of γ such that $d_X(x, \mathbf{n}(x)) = \min_{i \in [0, N]} d_X(x, \gamma(i))$. In the case that X is δ -hyperbolic, there is a $p \geq 0$ depending only on K and δ such that if $\gamma(j)$ is a different point on γ minimizing distance to x , then $d_X(\mathbf{n}(x), \gamma(j)) \leq p$. Moreover, $\mathbf{n}(\gamma(i)) = \gamma(i)$ for any $i \in [0, N]$ and for any $x, y \in X$,

$$d_X(\mathbf{n}(x), \mathbf{n}(y)) \leq p \cdot d(x, y) + p.$$

In Section 5, the nearest point retraction will be used to define a projection from the space X to the domain interval of a quasigeodesic.

2.3. Curves, markings, and hierarchy paths. In this section, we recall the work of Masur-Minsky on the curve and marking graphs. Fix a orientable surface S with genus $g \geq 0$ and $p \geq 0$ punctures so that $\omega(S) = 3g - 3 + p \geq 1$; $\omega(S)$ is called the *complexity* of S . The *curve graph*, denoted $\mathcal{C}(S)$, is a locally infinite simplicial graph whose vertices are isotopy classes of essential simple closed curves on S and two (isotopy classes of) curves are joined by an edge if they can be realized disjointly on S . The curve graph is the 1-skeleton of a simplicial complex introduced by Harvey in [Har81].

Remark 2.1. The above definition is for S with $\omega(S) \geq 2$. When $\omega(S) = 1$, the definition is modified so that $\mathcal{C}(S)$ is the Farey graph. See Subsection 2.4 of [MM00] for when S is an annulus, i.e. $\omega(S) = -1$.

Endow $\mathcal{C}(S)$ with the graph metric. We frequently use the following foundational result of Masur-Minsky [MM99]:

Theorem 2.2 ([MM99]). *For any S , there is a $\delta > 0$ so that $\mathcal{C}(S)$ is δ -hyperbolic.*

A (complete clean) *marking*, μ , on S is a pants decomposition called the *base* of μ , $\text{base}(\mu)$, and, for each $\alpha \in \text{base}(\mu)$, a *transversal* $t_\alpha \in \mathcal{C}(S)$ which intersects α and no other base curve. The *marking graph*, $\mathcal{M}(S)$, is a simplicial graph whose vertices are markings, with two markings connected by an edge if they differ by a Dehn (half) twist around a base curve $(\alpha, t_\alpha) \mapsto (\alpha, T_\alpha \cdot t_\alpha)$ called a *twist move*, or a *flip move*, which switches a base curve and its transversal, $(\alpha, t_\alpha) \mapsto (t_\alpha, \alpha)$ (see Section 2.5 of [MM00] for more details). Masur and Minsky show:

Theorem 2.3 ([MM00]). *$\text{Mod}(S)$ is $\text{Mod}(S)$ -equivariantly quasi-isometric to $\mathcal{M}(S)$.*

Often we want to compare two curves or markings on a subsurface. For any curve $\alpha \in \mathcal{C}(S)$ and nonannular subsurface $Y \subset S$, the *subsurface projection* of α to Y is the subset $\pi_Y(\alpha) \subset \mathcal{C}(Y)$ obtained by restricting α to Y and completing the resulting arcs to curves along ∂Y in a natural way (see Section 2.3 of [MM00] for more details and the definition when Y is an annulus). In the case of a marking $\mu \in \mathcal{M}(S)$, one projects only the base, that is $\pi_Y(\mu) = \pi_Y(\text{base}(\mu))$. For $\mu_1, \mu_2 \in \mathcal{M}(S)$, we write $d_Y(\mu_1, \mu_2) = \text{diam}_{\mathcal{C}(Y)}(\pi_Y(\mu_1) \cup \pi_Y(\mu_2))$.

One of the main constructions from [MM00] is the hierarchy machinery, from which we need only a few features of the induced hierarchy paths (see Section 4 of [MM00]). Given

two markings $\mu_1, \mu_2 \in \mathcal{M}(S)$, a *hierarchy* J between μ_1 and μ_2 is a collection of geodesics in various subsurface curve graphs whose interrelations encode the combinatorial relationship between μ_1 and μ_2 .

For any $A \geq 0$ we call a subsurface $Y \subseteq S$ (possibly $Y = S$) an *A-large link* for μ_1 and μ_2 if $d_Y(\mu_1, \mu_2) \geq A$. The following theorem says that the distance between any two markings is coarsely determined by the large links between them:

Theorem 2.4 (The distance formula; Theorem 6.12 in [MM00]). *There is a constant $A_0 \geq 0$, depending only on S , so that for all $A \geq A_0$ there exists $K \geq 1$ such that for any pair of markings $\mu_1, \mu_2 \in \mathcal{M}(S)$, we have*

$$\frac{1}{K} \cdot d_{\mathcal{M}(S)}(\mu_1, \mu_2) - K \leq \sum_{Y \subseteq S} [d_Y(\mu_1, \mu_2)]_A \leq K \cdot d_{\mathcal{M}(S)}(\mu_1, \mu_2) + K$$

where $[X]_A = X$ if $X \geq A$ and 0 otherwise.

By contrast, we say that two markings $\mu_1, \mu_2 \in \mathcal{M}(S)$ are *E-cobounded* if $d_Y(\mu_1, \mu_2) \leq E$ for every proper subsurface $Y \subsetneq S$. More generally, we say that a collection of markings $M \subset \mathcal{M}(S)$ is *E-cobounded* if every pair of markings in M is *E-cobounded*. Coboundedness is a strong condition and paths between cobounded markings have hyperbolic behavior, a central idea in Section 5.

Though hierarchies are technical objects with many applications, for this note their utility lies in their ability to be built into *hierarchy paths*. We collect some properties of hierarchy paths in the following theorem:

Theorem 2.5. *There are $M, M_1, M_2 \geq 0$ depending only on S , such that for any $\mu_1, \mu_2 \in \mathcal{M}(S)$, the following hold:*

- (1) *There is a hierarchy path $H : [0, N] \rightarrow \mathcal{M}(S)$ with $H(0) = \mu_1$ and $H(N) = \mu_2$, and every hierarchy path is an M -quasigeodesic.*
- (2) *For each $Y \subseteq S$, the projection of the hierarchy path H to $\mathcal{C}(Y)$ via subsurface projection is an unparameterized quasigeodesic with uniform constants.*
- (3) *If $d_Y(\mu_1, \mu_2) \geq M_1$, then the set of markings in H whose bases contain ∂Y is a contiguous subpath denoted H_Y . Further, if α_Y and β_Y denote the initial and terminal markings of H_Y , respectively, then*

$$d_Y(\alpha_Y, \beta_Y) \geq d_Y(\mu_1, \mu_2) - 2M_2.$$

- (4) *For any $E > 0$ there is an $E' > 0$ depending only on E and S such that if μ_1, μ_2 are E -cobounded, H is a hierarchy path between them, and $\mu'_1, \mu'_2 \in H$, then μ'_1 and μ'_2 are E' -cobounded.*

Remark 2.6. The above theorem essentially follows from the work in [MM00], with (1) being [Theorem 6.10, [MM00]], (2) following from the construction, and (3) a consequence of [Lemma 5.16, [Min10]]. Part (4) follows from (2) and (3). These statements also appear in [BMM11].

Remark 2.7 (Hierarchy paths between cobounded markings). If two markings $\mu_1, \mu_2 \in \mathcal{M}(S)$ are *E-cobounded*, then Theorem 2.4 implies that $d_{\mathcal{M}(S)}(\mu_1, \mu_2) \asymp d_{\mathcal{C}(S)}(\mu_1, \mu_2)$. It follows then from Theorem 2.5 (2) that the projection to $\mathcal{C}(S)$ of any hierarchy path between μ_1 and μ_2 is a genuine quasigeodesic. See Section 5 below.

2.4. Convex cocompactness in $\text{Mod}(S)$. Convex cocompact subgroups of mapping class groups were introduced by Farb and Mosher in [FM02]. A finitely generated $G \leq \text{Mod}(S)$ is *convex cocompact* if for some $x \in \mathcal{T}(S)$ the orbit $G \cdot x$ is quasiconvex with respect to the Teichmüller metric on $\mathcal{T}(S)$. Farb-Mosher verify that convex cocompactness is independent of the chosen $x \in \mathcal{T}(S)$ and relate convex cocompact subgroups of mapping class groups to hyperbolic extensions of surfaces groups. See [FM02, Ham05] for details.

Kent-Leininger [KL08] and, independently, Hamenstädt [Ham05] gave a characterization of convex cocompactness in terms of the curve graph $\mathcal{C}(S)$:

Theorem 2.8 (Kent-Leininger, Hamenstädt). *Let $G \leq \text{Mod}(S)$ be finitely generated. Then G is convex cocompact if and only if some (any) orbit map $G \rightarrow \mathcal{C}(S)$ is a quasi-isometric embedding.*

Our main goal in this note is to provide a characterization of convex cocompactness in $\text{Mod}(S)$ that uses only the geometry of $\text{Mod}(S)$ itself, and neither Teichmüller space nor the curve graph. This geometric characterization leads us to define the notion of *stability*, which is defined for arbitrary finitely generated groups.

3. STABILITY

In this section, we define stability and provide some basic properties. Informally, a quasi-isometrically embedded subspace is stable if all quasigeodesics beginning and ending in the space are forced to follow travel. This strong notion of convexity forces hyperbolic-like behavior around the subspace.

Definition 2. *Let $f : Y \rightarrow X$ be a quasi-isometric embedding between geodesic metric spaces. We say Y is stable in X if for any $L \geq 0$ there is a $R = R(L) \geq 0$ so that if $\gamma : [a, b] \rightarrow X$ and $\gamma' : [a', b'] \rightarrow X$ are L -quasigeodesics with $\gamma(a) = \gamma'(a') \in f(Y)$ and $\gamma(b) = \gamma'(b') \in f(Y)$, then*

$$d_{\text{Haus}}(\gamma, \gamma') \leq R.$$

Note that when we say Y is stable in X we mean that Y is stable in X with respect to a particular quasi-isometric embedding $Y \rightarrow X$. Such a quasi-isometric embedding will always be clear from context, e.g., an undistorted subgroup H of a finitely generated group G .

Remark 3.1. The condition that $f : Y \rightarrow X$ is a K -quasi-isometric embedding for some $K \geq 1$ implies that if γ is an L -quasigeodesic that begins and ends on the image of Y then it remains within an R' -neighborhood of $f(Y)$ where $R' = R(\max\{K, L\})$. In particular, $f(Y)$ is quasiconvex in X . To see these statements, note that if we let σ be a geodesic in Y whose end points map under f to the end points of γ , then $f(\sigma)$ is a K -quasigeodesic and $d_{\text{Haus}}(\gamma, \sigma) \leq R'$.

It is well-known that when X is δ -hyperbolic, the preimage of a quasiconvex subspace through a quasi-isometric embedding is itself quasiconvex. This property, however, fails when the space X is not hyperbolic. An important property of stability is that it is preserved under quasi-isometric embeddings. This will be especially important when characterizing stable subgroups of mapping class groups.

Proposition 3.2. *Suppose that X, Y, Z are geodesic metric spaces and $X \rightarrow Y \rightarrow Z$ are quasi-isometric embeddings. If X is stable in Z , then X is stable in Y .*

Proof. Let γ_1 and γ_2 be L -quasigeodesics in Y which share endpoints in X . If $f : Y \rightarrow Z$ is a K -quasi-isometric embedding, then $f(\gamma_1), f(\gamma_2)$ are L' -quasigeodesics in Z that share endpoints in X , where L' depends only on L and K . By stability of X in Z , these quasigeodesics remain within an R -neighborhood of one another, for R depending on L' . We conclude that γ_1 and γ_2 have Hausdorff distance no greater than $K(R + K)$. Since these constants depend only on K and L , this completes the proof. \square

Lemma 3.3. *If Y is stable in X then Y is δ -hyperbolic for some $\delta \geq 0$.*

Proof. This follows from well-known arguments, so we only provide a sketch. See Lemma 6.2 of [MM99] for details. Fix a K -quasi-isometric embedding $f : Y \rightarrow X$. Let $x, y, z \in Y$ and consider geodesics $[x, y], [y, z], [x, z]$ in Y joining these points. It suffices to show that $[x, y]$ is contained in the δ -neighborhood of the other two geodesics, for δ depending only on K and the stability constants.

If z' denotes a point on $[x, y]$ nearest to z in Y , then both $[z, z'] \cup [z', x]$ and $[z, z'] \cup [z', y]$ are 3-quasigeodesics, where $[x, y] = [x, z'] \cup [z', y]$. The point is that the images of these (quasi-) geodesics under $f : Y \rightarrow X$ are quasigeodesics with uniform constants. Hence, there is an $R \geq 0$ depending only on these constants so that $f([z, z'] \cup [z', x]) \subset N_R(f([x, z]))$ and $f([z, z'] \cup [z', y]) \subset N_R(f([y, z]))$. Since f is a K -quasi-isometric embedding, this implies that every point on $[x, y]$ is within $K(R + K)$ of some point on either $[x, z]$ or $[y, z]$. This completes the proof. \square

Although we have defined stability in a general setting, our focus will be the case of a finitely generated group G . Fix a finite generating set S of G and let $|\cdot|_S$ be the associated word metric. Recall that any two generating sets of G give quasi-isometric metrics and that a finitely generated subgroup $H \leq G$ is called *undistorted* if the inclusion $H \rightarrow G$ is a quasi-isometric embedding for some (any) word metrics on H and G .

Definition 3. *Let G be a finitely generated group with word metric $|\cdot|_S$. Then $H \leq G$ is stable if H is undistorted in G and $H \subset (G, |\cdot|_S)$ is stable (as in Definition 2) for any choice of word metric on H .*

Note that in the definition of stability for $H \leq G$, since H is undistorted in G one can use any word metric on H when defining stability. The next lemma, whose proof follows directly from Lemma 3.2, shows that the stability of $H \leq G$ is also independent of the word metric on G .

Lemma 3.4. *Let G be a finitely generated group. If $H \leq G$ is stable with respect to some word metric $|\cdot|_S$ on G , then it is stable with respect to any word metric on G .*

Remark 3.5. For a finitely generated group G , the property of stability of a subgroup H is well studied in the case where H is cyclic. In this case, $H = \langle h \rangle$, the generating element h is usually called either *stable* or *Morse*. See [DMS10] and the references found there.

4. THE MASUR-MINSKY CRITERIA FOR STABILITY

To show that convex cocompact subgroups of mapping class groups are stable, we use the criterion for hyperbolicity developed by Masur-Minsky in [MM99] which we adapt for our purposes.

We say that a family of paths Γ in X is *transitive* for a subspace $Y \subset X$ if any two points in Y can be connected by a path in Γ .

Definition 4. Let X be a metric space with subspace $Y \subset X$. Let Γ be a transitive family of paths in X between points in Y with the following property: for each $\beta : I \rightarrow X$ there exists a map $\pi : X \rightarrow I$ and constants $a, b, c > 0$ such that

- (1) For any $t \in I$, $\text{diam}_X(\beta([t, \pi(\beta(t))])) \leq c$.
- (2) If $d(x_1, x_2) \leq 1$, then $\text{diam}(\beta([\pi(x), \pi(y)])) \leq c$.
- (3) If $d(x, \beta(\pi(x))) \geq a$ and $d(x, x') \leq b \cdot d_X(x, \beta(\pi(x)))$, then

$$\text{diam}(\beta[\pi(x), \pi(x')]) \leq c.$$

Then we call Γ a family of uniformly contracting paths for $Y \subset X$.

Remark 4.1. The projection $\beta(\pi(x)) \in \text{im}(\beta)$ need not be a bounded distance from the points on β which are closest to x .

Proposition 4.2 ([MM99]). Let Γ be a family of uniform contracting paths for $Y \subset X$. Then for any $L \geq 0$ there exists an $R = R(L) \geq 0$ with the following property: if γ is a L -quasigeodesic that begins and ends on Y , then for any $\beta \in \Gamma$ with β having the same endpoints of γ , we have $d_{\text{Haus}}(\gamma, \beta) \leq R$.

Proof. This is proven in Lemma 6.1 of [MM99], which states that any space X with a family of uniformly contracting paths for X has stability of quasigeodesics. The proof of Lemma 6.1 shows that any L -quasigeodesic α whose endpoints agree with a $\beta \in \Gamma$ is contained in an R' -neighborhood of β , where R' depending only on L and the uniform contracting constants $(a, b, c$ from Definition 4). This proves our proposition. \square

The following consequence is now immediate:

Corollary 4.3. Suppose that $f : Y \rightarrow X$ is a quasi-isometric embedding between geodesic metric spaces and that $f(Y) \subset X$ has a family of uniform contracting paths. Then Y is stable in X .

5. CONVEX COCOMPACTNESS IMPLIES STABILITY

Our starting point is the following characterization of convex cocompact subgroups of the mapping class group, which follows easily from [KL08] or by combining results of [FM02] and [Raf10]. We provide a few details using these references.

Recall that a collection of markings $M \subset \mathcal{M}(S)$ is called E -cobounded if for any $\mu, \nu \in M$ and any proper subsurface $Y \subsetneq S$, we have $d_Y(\mu, \nu) \leq E$.

Lemma 5.1. Let G be a finitely generated subgroup of $\text{Mod}(S)$. If G is convex cocompact, then, for any marking $\mu \in \mathcal{M}(S)$, there is an $E \geq 0$ so that the orbit $G \cdot \mu \subset \mathcal{M}(S)$ is E -cobounded. Conversely, if G is undistorted in $\text{Mod}(S)$ and there is a marking $\mu \in \mathcal{M}(S)$ and an $E \geq 0$ so that $G \cdot \mu$ is E -cobounded, then G is convex cocompact.

Proof. The first statement is contained in the proof of Theorem 7.4 of [KL08], where the assumption on G is that the orbit map from G into $\mathcal{C}(S)$ is a quasi-isometric embedding.

Alternatively, we can see E -coboundedness of the orbit $G \cdot \mu$ using the fact that orbits of G in $\mathcal{T}(S)$ are quasiconvex. Let $x \in \mathcal{T}(S)$ be such that every curve in μ has bounded length in x . If there exists subsurfaces $Y_i \subsetneq S$ and $g_i \in G$ with

$$d_{Y_i}(\mu, g_i \cdot \mu) \rightarrow \infty$$

then the Teichmüller geodesics τ_i joining x and $g_i \cdot x$ become ϵ_i -thin for $\epsilon_i \rightarrow 0$ [Theorem 5.5, [Raf10]]. (See also Theorem 4.1 of [RS09].) However, the orbit $G \cdot x$ is in some fixed

thick part of $\mathcal{T}(S)$ and so we must have that points along τ get arbitrary far from the orbit $G \cdot x$. This contradicts orbit quasiconvexity of G in $\mathcal{T}(S)$.

The second statement follows from the Masur-Minsky distance formula (Theorem 2.4). Since G is undistorted, we may coarsely measure distance in G by distance in the orbit $G \cdot \mu \subset \mathcal{M}(S)$. That is, for any $g_1, g_2 \in G$

$$(1) \quad d_G(g_1, g_2) \asymp d_{\mathcal{M}(S)}(g_1 \cdot \mu, g_2 \cdot \mu) \asymp d_S(g_1 \cdot \mu, g_2 \cdot \mu) + \sum_{Y \subsetneq S} [d_Y(g_1 \cdot \mu, g_2 \cdot \mu)]_A,$$

where the symbol \asymp depends only on the surface S and the quasi-isometry constant of the orbit map $G \rightarrow \mathcal{M}(S)$. Choosing the threshold A in the distance formula (Theorem 2.4) to be larger than E shows that distance in G is coarsely distance in its curve graph orbit. Hence, the orbit map $G \rightarrow \mathcal{C}(S)$, given by $g \mapsto g \cdot \mu$, is a quasi-isometric embedding and so G is convex cocompact by Theorem 2.8. This completes the proof. \square

For the remainder of this section suppose that $G \leq \text{Mod}(S)$ is convex cocompact. In order to show that G is stable in $\text{Mod}(S)$, it suffices to show that $G \cdot \mu$ is stable in $\mathcal{M}(S)$. This follows from the fact that $\text{Mod}(S)$ is quasi-isometric to $\mathcal{M}(S)$ and G is undistorted in $\text{Mod}(S)$. By Lemma 5.1, $G \cdot \mu$ is \bar{E} -cobounded for some $\bar{E} \geq 0$. We proceed by showing that hierarchy paths form a family of uniform contracting paths for $G \cdot \mu \subset \mathcal{M}(S)$.

Our use of hierarchy paths is motivated by the Slice Comparison Lemma [Lemma 6.7, [MM00]], Behrstock's work on the asymptotic cone of $\text{Mod}(S)$ [Beh06], and more recent work of Sisto [Sis11]. To provide the most direct proof of Theorem 5.6, we have chosen to use a theorem of Duchin-Rafi [DR09] (stated as Theorem 5.3 below), which is compatible with the sort of projection from Definition 4. The strong contraction property of hierarchy paths between cobounded pants decompositions was also proven in Theorem 4.4 of [BMM11] in their work on the Weil-Petersson geometry of $\mathcal{T}(S)$. We have included a proof of Proposition 5.5 here for completeness and as an application of the Masur-Minsky criteria (Theorem 4.2).

Let $p : \mathcal{M}(S) \rightarrow \mathcal{C}(S)$ be the map which associates to a marking μ the collection of curves which appear in the base of μ , i.e. $a \in p(\mu)$ if and only if $a \in \text{base}(\mu)$. This map, called the shadow map, is coarsely 4-Lipschitz [Lemma 2.5, [MM00]]. For any markings $\mu, \nu \in \mathcal{M}(S)$, let

$$H = H(\mu, \nu) : [0, N] \rightarrow \mathcal{M}(S)$$

be a hierarchy path with $H(0) = \mu$ and $H(N) = \nu$. Recall that H is a M -quasigeodesic in $\mathcal{M}(S)$, where M depends only on the topology of S (Theorem 2.5 (1)). If μ and ν are \bar{E} -cobounded, then all the markings that appear in H are E -cobounded, for some $E \geq 0$ that depends only \bar{E} and the surface S (Theorem 2.5 (4)). By the argument in Lemma 5.1, $h = p \circ H : [0, N] \rightarrow \mathcal{C}(S)$ is a quasigeodesic in $\mathcal{C}(S)$ from $p(\mu)$ to $p(\nu)$ whose quasigeodesic constant depends only on E . More precisely, if we choose the cut off A in the distance formula (Theorem 2.4) to be larger than E , then for any $i, j \in [0, N]$

$$\begin{aligned} |i - j| &\leq M \cdot d_{\mathcal{M}(S)}(H(i), H(j)) + M \\ &\leq MK \cdot d_S(h(i), h(j)) + MK^2 + M, \end{aligned}$$

where K depends only on E and S . Hence, $h = p \circ H$ is a K_E -quasigeodesic in $\mathcal{C}(S)$, where $K_E = K^2 M$.

As $h : [0, N] \rightarrow \mathcal{C}(S)$ is a K_E -quasigeodesic into the δ -hyperbolic space $\mathcal{C}(S)$, there is a nearest point retraction $\mathbf{n}_h : \mathcal{C}(S) \rightarrow h$ as discussed in Subsection 2.2. Define the projection

$\text{proj}_h : \mathcal{C}(S) \rightarrow [0, N]$ to the domain of the path h so that for $c \in \mathcal{C}^0(S)$,

$$h(\text{proj}_h(c)) = \mathbf{n}_h(c).$$

That is, $\text{proj}_h(c)$ is a parameter $i \in [0, N]$ so that the distance from c to the image of h is minimized at $h(i)$. By the properties of \mathbf{n}_h stated in Subsection 2.2, it is immediate that there is an L (depending only on E) so that this projection is both L -coarsely well-defined and coarsely L -Lipschitz. We emphasize that this uses only the facts that $\mathcal{C}(S)$ is hyperbolic and $h : [0, N] \rightarrow \mathcal{C}(S)$ is a quasigeodesic.

The projection $\text{proj}_h : [0, N] \rightarrow \mathcal{C}(S)$ induces a corresponding map from $\mathcal{M}(S)$ to $[0, N]$. Let

$$\text{Proj}_H : \mathcal{M}(S) \rightarrow [0, N]$$

be defined as follows: for any $\alpha \in \mathcal{M}(S)$, set $\text{Proj}_H(\alpha) = \text{proj}_h(a)$, for some choice of curve $a \in p(\alpha) \subset \mathcal{C}(S)$. Note that for different choices of curves $a, a' \in p(\alpha)$ we have $d_S(a, a') \leq 4$ and so $|\text{proj}_h(a) - \text{proj}_h(a')| \leq 4L$.

Remark 5.2. It may seem slightly unnatural to define the projection Proj_H to the domain of the path H , rather than to its image in $\mathcal{M}(S)$. We have done so for two reasons. First, this allows for a direct application of Theorem 4.2, which verifies that quasigeodesics follow travel in a uniform way. Indeed, projecting to the domain of a path is the approach of Masur-Minsky in [MM99]. Second, it seems that such a projection of a marking μ to a hierarchy path H need not be a uniformly bounded distance from the closet point to μ on H . Using Theorem 4.2 avoids this subtlety.

Our goal for the rest of this section, achieved in Proposition 5.5 below, is to show that the collection of hierarchy paths between markings in a fixed orbit of a convex cocompact subgroup G is a family of uniform contracting paths for the orbit in $\mathcal{M}(S)$. That convex cocompact subgroups are stable, Theorem 5 below, follows quickly from Proposition 5.5 and Corollary 4.3. The proof of the contracting property uses the following theorem of Duchin and Rafi [DR09] (see also Theorem 4.3 of [BMM11]):

Theorem 5.3 (Theorem 4.2, [DR09]). *Given E there exist B_1 and B_2 so that if H is a hierarchy path in $\mathcal{M}(S)$ between E -cobounded markings μ and ν , then for any $\alpha \in \mathcal{M}(S)$ with $d(\alpha, H) \geq B_1$ and $R = d(\alpha, H)/B_1$, we have*

$$\text{diam}_{\mathcal{C}(S)}(h(\text{Proj}_H(B_R(\alpha)))) \leq B_2,$$

where B_R denote the R -ball in $\mathcal{M}(S)$.

Remark 5.4. In the statement of Theorem 5.3 in [DR09], the authors allow any quasigeodesic in $\mathcal{M}(S)$ all of whose markings are uniformly cobounded. This is automatically satisfied by the hierarchy path H (Theorem 2.5 (4)).

Proposition 5.5. *Let $M \subset \mathcal{M}(S)$ be a collection of E -cobounded markings. The set of all hierarchy paths between markings in M is a family of uniformly contracting paths for $M \subset \mathcal{M}(S)$.*

Proof. Let $\mu, \nu \in M$ be arbitrary and let $H : [0, N] \rightarrow \mathcal{M}(S)$ be an hierarchy path with $H(0) = \mu$ and $H(N) = \nu$. We show that the conditions from Proposition 4.2 are satisfied for the projections Proj_H defined above, with constants a, b, c depending only on E . For (1), we must show that for any $i \in I$,

$$\text{diam}_{\mathcal{M}(S)} H([i, \text{Proj}_H(H(i))])$$

is bounded by a constant depending only on K_E . As H is an M -quasigeodesic, this quantity is bounded by $M|i - \text{Proj}_H(H(i))| + M$. Since $h(i) = p(H(i))$, for any curve $a \in p(H(i))$ the difference $|\text{proj}_h(a) - i|$ is bounded by L . Hence,

$$M|i - \text{Proj}_H(H(i))| + M \leq M|i - \text{Proj}_h(a)| + M \leq M \cdot L + M,$$

as required.

For (2), we show that for any $\alpha, \beta \in \mathcal{M}(S)$ with $d_{\mathcal{M}(S)}(\alpha, \beta) \leq 1$

$$\text{diam}_{\mathcal{M}(S)}(H([\text{Proj}_H(\alpha), \text{Proj}_H(\beta)]))$$

is bounded by a constant depending only on E . This is similar to (1), since again it suffices to bound $|\text{Proj}_H(\alpha) - \text{Proj}_H(\beta)|$. Let $a \in p(\alpha)$ and $b \in p(\beta)$, then $d_S(a, b) \leq 8$ (p is 4-Lipschitz) and so $|\text{proj}_h(a) - \text{proj}_h(b)| \leq 8L$. Hence,

$$|\text{Proj}_H(\alpha) - \text{Proj}_H(\beta)| \leq |\text{proj}_h(a) - \text{proj}_h(b)| \leq 8L.$$

For (3), we will apply Theorem 5.3. Before doing so, we must first show that there is a $B_3 \geq 0$ so that for any $\alpha \in \mathcal{M}(S)$,

$$d_{\mathcal{M}(S)}(\alpha, H(\text{Proj}_H(\alpha))) \leq B_3 \cdot d_{\mathcal{M}(S)}(\alpha, H),$$

whenever $d_{\mathcal{M}(S)}(\alpha, H)$ is sufficiently large. Here, as in Theorem 5.3, $d_{\mathcal{M}(S)}(\alpha, H)$ is the minimum distance from α to any marking in the image of H .

Let $n(\alpha)$ be a marking on H that is closest to α and set $\bar{\alpha} = H(\text{Proj}_H(\alpha))$. By construction, $d_S(\alpha, \bar{\alpha}) \leq d_S(\alpha, n(\alpha))$. For a proper subsurface $Y \subsetneq S$,

$$(2) \quad d_Y(\alpha, \bar{\alpha}) \leq d_Y(\alpha, n(\alpha)) + d_Y(n(\alpha), \bar{\alpha}) \leq d_Y(\alpha, n(\alpha)) + E$$

Plugging the above inequality (2) into the distance formula (Theorem 2.4) with threshold $A \geq 2E$ gives

$$d_{\mathcal{M}(S)}(\alpha, \bar{\alpha}) \leq B_3 \cdot d_{\mathcal{M}(S)}(\alpha, n(\alpha)),$$

for some $B_3 \geq 0$. Note that to eliminate the additive constant in the distance formula, we have used that $d_{\mathcal{M}(S)}(\alpha, n(\alpha)) \neq 0$.

Now set $a = B_1 B_3$, $b = \frac{1}{B_1 B_3}$, and $c = M(K_E \cdot B_2 + K_E + 1)$, where B_1 and B_2 are as in Theorem 5.3 and B_3 was determined above. Let $\alpha \in \mathcal{M}(S)$ with $d_{\mathcal{M}(S)}(\alpha, H(\text{Proj}_H(\alpha))) \geq a$ and $\beta \in \mathcal{M}(S)$ with $d_{\mathcal{M}(S)}(\alpha, \beta) \leq b \cdot d_{\mathcal{M}(S)}(\alpha, H(\text{Proj}_H(\alpha)))$. Then

$$d_{\mathcal{M}(S)}(\alpha, H) \geq 1/B_3 \cdot d_{\mathcal{M}(S)}(\alpha, H(\text{Proj}_H(\alpha))) \geq a/B_3 = B_1$$

and

$$\begin{aligned} d_{\mathcal{M}(S)}(\alpha, \beta) &\leq b \cdot d_{\mathcal{M}(S)}(\alpha, H(\text{Proj}_H(\alpha))) \\ &\leq \frac{1}{B_1 B_3} \cdot d_{\mathcal{M}(S)}(\alpha, H(\text{Proj}_H(\alpha))) \\ &\leq \frac{1}{B_1} \cdot d_{\mathcal{M}(S)}(\alpha, H). \end{aligned}$$

Hence, $\beta \in B_R(\alpha)$ for $R = \frac{1}{B_1} d_{\mathcal{M}(S)}(\alpha, H)$, and so by Theorem 5.3

$$d_{\mathcal{C}(S)}(h(\text{Proj}_H(\alpha)), h(\text{Proj}_H(\beta))) \leq B_2.$$

Since h is a K_E -quasigeodesic, $|\text{Proj}_H(\alpha) - \text{Proj}_H(\beta)| \leq K_E \cdot B_2 + K_E$, and so we conclude that

$$\text{diam}_{\mathcal{M}(S)}(H([\text{Proj}_H(\alpha), \text{Proj}_H(\beta)])) \leq M(K_E \cdot B_2 + K_E) + M = c.$$

This completes the proof of condition (3), and shows that the collection of hierarchy paths between markings in M is a family of uniformly contracting paths for M in $\mathcal{M}(S)$. \square

Theorem 5.6. *If $G \leq \text{Mod}(S)$ is convex cocompact, then G is stable.*

Proof. Since G is convex cocompact, it is undistorted in $\text{Mod}(S)$. Fix $\mu \in \mathcal{M}(S)$ and recall that the orbit map $\text{Mod}(S) \rightarrow \mathcal{M}(S)$, given by $g \mapsto g \cdot \mu$, is a quasi-isometry. Hence, by Proposition 3.2, it suffices to show that $G \rightarrow \mathcal{M}(S)$ is stable. By Theorem 5.1, the orbit $G \cdot \mu$ is E -cobounded for some $E \geq 0$. Proposition 5.5 then implies that the set of all hierarchy paths between vertices in the orbit $G \cdot \mu$ is a family of uniform contracting paths for $G \cdot \mu$ in $\mathcal{M}(S)$. By Corollary 4.3, this implies that $G \rightarrow G \cdot \mu \subset \mathcal{M}(S)$ is stable. Thus G is a stable subgroup of $\text{Mod}(S)$. \square

The following corollary is now immediate.

Corollary 5.7. *Let $G \leq \text{Mod}(S)$ be convex cocompact. Then G is quasiconvex in $\text{Mod}(S)$ with respect to any (finite) generating set.*

6. STABILITY IMPLIES CONVEX COCOMPACTNESS

The proof of the converse to Theorem 5.6 is a straightforward contradiction argument using the structure of hierarchy paths and the marking complex (see Section 2.3). Before we proceed with the proof, we recall some notions from [BM08] about the product regions of $\mathcal{M}(S)$ associated to simplices of $\mathcal{C}(S)$. By a simplex of $\mathcal{C}(S)$, we mean a collection of pairwise adjacent vertices of $\mathcal{C}(S)$.

Let $\Delta \subset \mathcal{C}(S)$ be a simplex, and let $\mathcal{Q}(\Delta) \subset \mathcal{M}(S)$ be the set of markings whose bases contain Δ . We note that $\mathcal{Q}(\Delta)$ is quasi-isometric to $\text{Stab}_{\text{Mod}(S)}(\Delta)$, the stabilizer of Δ in $\text{Mod}(S)$. Let $\sigma(\Delta) \subset S$ be the components of $S \setminus \Delta$ which are not pairs of pants, including the annuli about the curves of Δ . There is a map

$$\mathcal{Q}(\Delta) \rightarrow \prod_{Y \in \sigma(\Delta)} \mathcal{M}(Y)$$

where $\mathcal{M}(Y) = \mathcal{C}(Y)$ if Y is an annulus. The map is given by restricting (or projecting) a marking $\mu \in \mathcal{Q}(\Delta)$ to markings on the subsurfaces in $\sigma(\Delta)$ and, for each $\alpha \in \Delta$, associating the transversal to α in μ to a corresponding point in $\mathcal{C}(\alpha)$. The following lemma is essentially an application of the distance formula (Theorem 2.4):

Lemma 6.1 (Lemma 2.1, [BM08]). *The correspondence*

$$\mathcal{Q}(\Delta) \rightarrow \prod_{Y \in \sigma(\Delta)} \mathcal{M}(Y)$$

is a P -quasi-isometry, where $P \geq 0$ depends only on the surface S .

We can now give the idea of the proof of Theorem 6.3. If a group $G \leq \text{Mod}(S)$ is stable but not convex cocompact, then Lemma 5.1 implies that the G -orbit of some marking $\mu \in \mathcal{M}(S)$ does not have bounded subsurface projections. Thus, for any $E > 0$, we can find a marking $\nu \in G \cdot \mu$ such that $d_Y(\mu, \nu) > E$ for some proper subsurface $Y \subsetneq S$. Theorem 2.5 implies that there is a hierarchy path H from μ to ν and a subsegment $H_Y \subset H$ with $H_Y \subset \mathcal{Q}(\partial Y)$ so that $|H_Y| > E$. If $\alpha_Y, \beta_Y \in H_Y$ are the initial and terminal markings of H_Y , respectively, then stability of G implies that there are markings $\mu_1, \mu_2 \in G \cdot \mu$ such that μ_1 and μ_2 are within some uniform distance of α_Y and β_Y . Using the product structure in Lemma 6.1, we can use H_Y to build two quasigeodesics between μ_1 and μ_2 with constants

depending only on S , whose Hausdorff distance is coarsely at least E . Since $E > 0$ was chosen arbitrarily, this contradicts the stability assumption for $G \cdot \mu$.

In this last step, we are taking advantage of the well-known fact that quasigeodesics in product spaces need not fellow-travel, a variation of which we record in the following lemma:

Lemma 6.2. *Let \mathcal{X} and \mathcal{Y} be connected, infinite-diameter graphs and let \mathcal{Z} be the 1-skeleton of $\mathcal{X} \times \mathcal{Y}$, endowed with the graph metric. For any vertices $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ of \mathcal{Z} , there are 3-quasigeodesics γ_a, γ_b each from z_1 to z_2 so that*

$$d_{\text{Haus}}(\gamma_a, \gamma_b) \geq \max\{d_{\mathcal{X}}(x_1, x_2), d_{\mathcal{Y}}(y_1, y_2)\}.$$

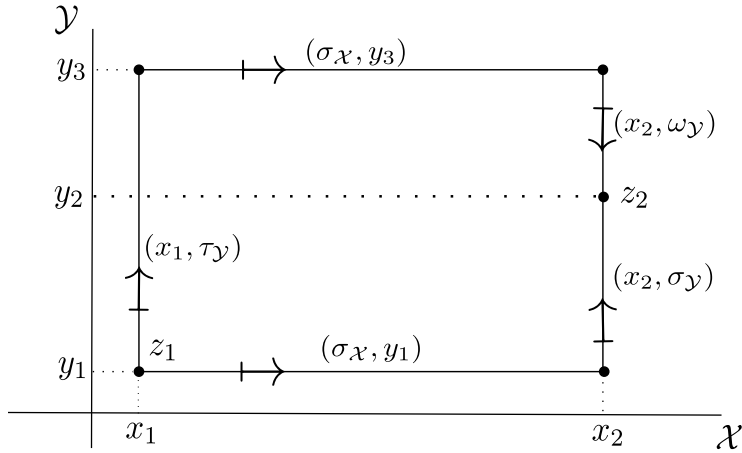


FIGURE 1. The path $\gamma_a = (\sigma_{\mathcal{X}}, y_1) \cdot (x_2, \sigma_{\mathcal{Y}})$ is far from the path $\gamma_b = (x_1, \tau_{\mathcal{Y}}) \cdot (\sigma_{\mathcal{X}}, y_3) \cdot (x_2, \omega_{\mathcal{Y}})$.

Proof. The proof is easily seen with Figure 1, but we provide the written details here. Since we are working only with graphs, all paths will be considered as sequences of adjacent vertices indexed by intervals of integers. Hence, for a path $\gamma : [0, N] \rightarrow \mathcal{Z}$, we have $\ell(\gamma|_{[i, j]}) = |j - i|$, where the length of such a path is the number of edges it traverses. In this case, to show that γ is a 3-quasigeodesic, it suffice to show that for any $i \leq j$,

$$\ell(\gamma|_{[i, j]}) \leq 3 \cdot d_{\mathcal{Z}}(\gamma(i), \gamma(j)).$$

Also, recall that for any two vertices $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ of \mathcal{Z} , the graph metric is

$$(3) \quad d_{\mathcal{Z}}(z_1, z_2) = d_{\mathcal{X}}(x_1, x_2) + d_{\mathcal{Y}}(y_1, y_2).$$

Now, suppose that $d := d_{\mathcal{X}}(x_1, x_2) \geq d_{\mathcal{Y}}(y_1, y_2)$ and let $\sigma_{\mathcal{X}}$ be a geodesic path in \mathcal{X} that joins x_1 to x_2 . Similarly, let $\sigma_{\mathcal{Y}}$ be a geodesic path in \mathcal{Y} that joins y_1 to y_2 . For any $y \in \mathcal{Y}$, we denote by $(\sigma_{\mathcal{X}}, y)$ the corresponding geodesic path in \mathcal{Z} whose first coordinate entries are the vertices of $\sigma_{\mathcal{X}}$ and whose second coordinate is y . With this notation, let γ_a be the path in \mathcal{Z} that is the concatenation (read from left to right)

$$\gamma_a = (\sigma_{\mathcal{X}}, y_1) \cdot (x_2, \sigma_{\mathcal{Y}}).$$

It is clear that γ_a is a path of adjacent vertices of \mathcal{Z} which joins z_1 to z_2 . Moreover, γ_a is a geodesic path. This follows from Equation 3 and the observation that γ_a does not backtrack in either coordinate.

We now construct a 3-quasigeodesic γ_b which also joins z_1 to z_2 but travels far from γ_a . Let y_3 be a vertex of \mathcal{Y} with the property that $d_{\mathcal{Y}}(y_1, y_3) = d_{\mathcal{X}}(x_1, x_2) = d$. Let $\tau_{\mathcal{Y}}$ be a geodesic path in \mathcal{Y} between y_1 and y_3 and let $\omega_{\mathcal{Y}}$ be a geodesic path in \mathcal{Y} between y_3 and y_2 . Now define γ_b to be the path of adjacent vertices of \mathcal{Z} given by

$$\gamma_b = (x_1, \tau_{\mathcal{Y}}) \cdot (\sigma_{\mathcal{X}}, y_3) \cdot (x_2, \omega_{\mathcal{Y}}).$$

As before, γ_b is a path from z_1 to z_2 . Also, since γ_b contains the point (x_1, y_3) , the Hausdorff distance from γ_b to γ_a is no less than the distance from (x_1, y_3) to γ_a . Since,

$$\begin{aligned} d_{\mathcal{Z}}((x_1, y_3), \gamma_a) &= \min\{d_{\mathcal{Z}}((x_1, y_3), (\sigma_{\mathcal{X}}, y_1)), d_{\mathcal{Z}}((x_1, y_3), (x_2, \sigma_{\mathcal{Y}}))\} \\ &\geq \min\{d_{\mathcal{Z}}((x_1, y_3), (x_1, y_1)), d_{\mathcal{Z}}((x_1, y_3), (x_2, y_3))\} \\ &= \min\{d_{\mathcal{Y}}(y_3, y_1), d_{\mathcal{X}}(x_1, x_2)\} = d, \end{aligned}$$

we have $d_{\text{Haus}(\mathcal{Z})}(\gamma_a, \gamma_b) \geq d = d_{\mathcal{X}}(x_1, x_2)$. It remains to show that γ_b is a 3-quasigeodesic.

As in the construction of γ_a , both $(x_1, \tau_{\mathcal{Y}}) \cdot (\sigma_{\mathcal{X}}, y_3)$ and $(\sigma_{\mathcal{X}}, y_3) \cdot (x_2, \omega_{\mathcal{Y}})$ are geodesic subpaths of γ_b . Hence, let $z_i \in (x_1, \tau_{\mathcal{Y}})$ and let $z_j \in (x_2, \omega_{\mathcal{Y}})$ and note that $d_{\mathcal{Z}}(z_i, z_j) \geq d_{\mathcal{X}}(x_1, x_2)$. Denote by $\gamma_b|_{[z_i, z_j]}$ the portion of γ_b between z_i and z_j . We compute

$$\begin{aligned} \ell(\gamma_b|_{[z_i, z_j]}) &= d_{\mathcal{Z}}(z_i, (x_1, y_3)) + d_{\mathcal{Z}}((x_1, y_3), (x_2, y_3)) + d_{\mathcal{Z}}((x_2, y_3), z_j) \\ &\leq 3 \cdot d_{\mathcal{X}}(x_1, x_2) \\ &\leq 3 \cdot d_{\mathcal{Z}}(z_i, z_j). \end{aligned}$$

By our remark in the first paragraph of this proof, we are done. \square

The following theorem completes the proof of Theorem 1.1.

Theorem 6.3. *Suppose that $G \leq \text{Mod}(S)$ is stable. Then G is convex cocompact.*

Proof. Assume towards a contradiction that $G \leq \text{Mod}(S)$ is stable but not convex cocompact. For a fixed $\mu \in \mathcal{M}(S)$, the orbit map $G \rightarrow \mathcal{M}(S)$ given by $g \mapsto g \cdot \mu$ is a K -quasi-isometric embedding for some $K \geq 1$. By Proposition 3.2, we have that $G \cdot \mu$ is stable in $\mathcal{M}(S)$.

Let $L = \max\{K, M\}$, where M is the quasigeodesic constant for a hierarchy path (Theorem 2.5 (1)), and set $A_1 = R(L)$, the stability constant for L -quasigeodesics in $\mathcal{M}(S)$ which begin and end on $G \cdot \mu$. Since any two markings in the orbit $G \cdot \mu$ are joined by both hierarchy paths and K -quasigeodesics which are contained in $G \cdot \mu$, any hierarchy path between markings in $G \cdot \mu$ is contained in the A_1 -neighborhood of $G \cdot \mu$ (see Remark 3.1). Finally, set $A_2 = R(3P^2 + 2A_1)$, where P is as in Theorem 6.1.

Since G is undistorted but not convex cocompact, Proposition 5.1 implies that for any $E \geq 0$ there is a $g \in G$ and a proper subsurface $Y \subset S$ such that

$$d_Y(\mu, g \cdot \mu) \geq E.$$

For $E \geq M_1$, Theorem 2.5 (3) implies that any hierarchy path $H : [0, N] \rightarrow \mathcal{M}(S)$ with $H(0) = \mu$ and $H(N) = g \cdot \mu$ contains a subpath H_Y such that $\partial Y \subset \text{base}(\mu)$ for each $\mu \in H_Y$, i.e. $H_Y \subset \mathcal{Q}(\partial Y)$. If we denote the initial marking and terminal markings of H_Y by α_Y and β_Y , respectively, then

$$(4) \quad d_{\mathcal{M}(Y)}(\alpha_Y, \beta_Y) \geq \frac{1}{4} d_Y(\alpha_Y, \beta_Y) = \frac{1}{4} (E - 2M_2)$$

with the last equality in (4) following from Theorem 2.5 (3). Set $E' = \frac{1}{4}(E - 2M_2)$.

By Lemma 6.1, $\mathcal{Q}(\partial Y)$ is P -quasi-isometric to $\prod_{X \subset \sigma(\partial Y)} \mathcal{M}(X)$. To apply Lemma 6.2, set $\mathcal{Y} = \mathcal{M}(Y)$ and $\mathcal{X} = \prod_{X \subset \sigma(\partial Y) \setminus \{Y\}} \mathcal{M}(X)$. Then

$$\mathcal{Z} = \mathcal{X} \times \mathcal{Y} = \prod_{X \subset \sigma(\partial Y)} \mathcal{M}(X).$$

Since $\alpha_Y, \beta_Y \in \mathcal{Q}(\partial Y)$, we may use the correspondence of Lemma 6.1 to view α_Y and β_Y in the product space \mathcal{Z} . Equation (4) implies that $d_{\mathcal{Y}}(\alpha_Y, \beta_Y) \geq E'$, where $d_{\mathcal{Y}}(\alpha_Y, \beta_Y)$ is just the distance between α_Y and β_Y in $\mathcal{M}(Y)$. Lemma 6.2 implies that there exists 3-quasigeodesics γ^1 and γ^2 in \mathcal{Z} that join the markings α_Y and β_Y in \mathcal{Z} and whose Hausdorff distance in \mathcal{Z} is greater than or equal to E' .

Using the P -quasi-isometry in Theorem 6.1, we may view γ^1 and γ^2 as $3P^2$ -quasigeodesics in $\mathcal{M}(S)$ that join the markings α_Y and β_Y . Measuring Hausdorff distance in $\mathcal{M}(S)$, we have

$$(5) \quad d_{\text{Haus}(\mathcal{M}(S))}(\gamma^1, \gamma^2) \geq \frac{E' - P}{P}.$$

Since the original hierarchy path H joins markings in $G \cdot \mu$, it is contained in an A_1 -neighborhood of the orbit $G \cdot \mu$. Hence, there are markings $\mu_1, \mu_2 \in G \cdot \mu$ so that

$$(6) \quad d_{\mathcal{M}(S)}(\alpha_Y, \mu_1) \leq A_1 \quad \text{and} \quad d_{\mathcal{M}(S)}(\beta_Y, \mu_2) \leq A_1,$$

where A_1 depends only on S , as above. By appending initial and terminal geodesic segments of length no more than A_1 to γ^1 and γ^2 , we may consider these paths as $(3P^2 + 2A_1)$ -quasigeodesics in $\mathcal{M}(S)$ that join the orbit points $\mu_1, \mu_2 \in G \cdot \mu$. By our choice of $A_2 = R(3P^2 + 2A_1)$, the stability constant for $(3P^2 + 2A_1)$, any two $(3P^2 + 2A_1)$ -quasigeodesics between markings in $G \cdot \mu$ have Hausdorff distance no greater than A_2 .

Since $E \geq 0$ was arbitrary, we may choose $E > 4(PA_2 + P) + 2M_2$ so that

$$E' = \frac{1}{4}(E - 2M_2) > PA_2 + P.$$

Then by Equation 5

$$d_{\text{Haus}(\mathcal{M}(S))}(\gamma^1, \gamma^2) \geq \frac{E' - P}{P} > A_2.$$

This contradicts the assertion that the $(3P^2 + 2A_1)$ -quasigeodesics γ^1 and γ^2 between $\mu_1, \mu_2 \in G \cdot \mu$ must be A_2 -Hausdorff close. We conclude that stable subgroups of $\text{Mod}(S)$ are convex compact, as required. \square

REFERENCES

- [Beh06] Jason Behrstock, *Asymptotic geometry of the mapping class group and Teichmüller space*, *Geom. Topol.* **10** (2006), 1523–1578.
- [BF06] Jeffrey Brock and Benson Farb, *Curvature and rank of Teichmüller space*, *American Journal of Mathematics* (2006), 1–22.
- [BH09] M.R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Springer, 2009.
- [BM08] Jason A Behrstock and Yair N Minsky, *Dimension and rank for mapping class groups*, *Annals of Mathematics* **167** (2008), 1055–1077.
- [BMM11] Jeffrey Brock, Howard Masur, and Yair Minsky, *Asymptotics of weil–petersson geodesics ii: bounded geometry and unbounded entropy*, *Geometric and Functional Analysis* **21** (2011), no. 4, 820–850.

- [DMS10] Cornelia Druţu, Shahar Mozes, and Mark Sapir, *Divergence in lattices in semisimple Lie groups and graphs of groups*, Transactions of the American Mathematical Society **362** (2010), no. 5, 2451–2505.
- [DR09] Moon Duchin and Kasra Rafi, *Divergence of geodesics in Teichmüller space and the mapping class group*, Geometric and Functional Analysis **19** (2009), no. 3, 722–742.
- [Far06] Benson Farb, *Some problems on mapping class groups and moduli space*, Proceedings of symposia in pure mathematics, vol. 74, Providence, RI; American Mathematical Society; 1998, 2006, p. 11.
- [FM02] Benson Farb and Lee Mosher, *Convex cocompact subgroups of mapping class groups*, Geom. Topol. **6** (2002), 91–152 (electronic).
- [FM12] Benson Farb and Dan Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, Princeton, NJ, 2012.
- [Ham05] Ursula Hamenstädt, *Word hyperbolic extensions of surface groups*, preprint (2005), arXiv:0807.4891v2.
- [Har81] W. J. Harvey, *Boundary structure of the modular group*, 245–251.
- [KL08] Richard P. Kent, IV and Christopher J. Leininger, *Shadows of mapping class groups: capturing convex cocompactness*, Geom. Funct. Anal. **18** (2008), no. 4, 1270–1325.
- [Min10] Yair N. Minsky, *The classification of Kleinian surface groups, I: Models and bounds*, Annals of Mathematics **171** (2010), 1–107.
- [MM99] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149.
- [MM00] ———, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974.
- [Raf10] Kasra Rafi, *Hyperbolicity in Teichmüller space*, arXiv preprint arXiv:1011.6004 (2010).
- [RS09] Kasra Rafi and Saul Schleimer, *Covers and the curve complex*, Geometry & Topology **13** (2009), 2141–2162.
- [Sis11] Alessandro Sisto, *Contracting elements and random walks*, arXiv preprint arXiv:1112.2666 (2011).