QUASICONVEXITY IN THE CURVE COMPLEX

HOWARD MASUR AND YAIR MINSKY

1. Introduction

Let S be a surface of genus g with n punctures, and assume $3g-3+n>1$. Associated to S is an object $\mathcal{C}(S)$ called the *complex of curves*, whose vertices are homotopy classes of nontrivial, nonperipheral simple closed curves. A k-simplex of $\mathcal{C}(S)$ is a collection of $k+1$ disjoint nonperipheral homotopically distinct simple closed curves. The dimension of the complex is $3g-4+n$. We are interested in the 1*-skeleton* $\mathcal{C}_1(S)$ which is endowed with a metric $d(\cdot,\cdot)$ which assigns length 1 to every edge.

In [6] we proved that $\mathcal{C}_1(S)$, with this metric, is δ-hyperbolic in the sense of Gromov and Cannon. This fact has applications to the study of mapping class groups and hyperbolic 3-manifolds [7, 9, 2]. See also Bowditch [1] for a new and more concrete proof of the hyperbolicity theorem.

In the study of Heegaard splittings it is of interest to consider the set of essential curves in the boundary of a handlebody which bound disks – see Casson-Gordon [3], Hempel [4], Johannson [5] and Schleimer [11]. In slightly more generality, for a compact, orientable 3-manifold M with boundary component S we can define the set of (homotopy classes of) boundaries of essential disks:

$$\Delta(M,S)=\{[\partial D]:(D,\partial D)\subset(M,S)\text{ is an essential disk}\}\subset\mathcal{C}(S).$$

The purpose of this note is to prove

Theorem 1.1. *$\Delta(M,S)$ is a K-quasiconvex subset of $\mathcal{C}(S)$, where K depends only on the genus of S.*

Recall that a subset Y of a geodesic metric space (X,d) is K*-quasiconvex* if for any pair of points $y_1,y_2\in Y$ any geodesic in X joining them stays in a K-neighborhood of Y. This theorem can be viewed as something of an analogy to a result on quasiconvexity of sublevel sets in $\mathcal{C}(S)$ for certain length functions arising from Kleinian representations, which plays a central role in [8, 9].

Theorem 1.1 will be a special case of more general theorems about quasi-convex subsets of $\mathcal{C}(S)$ obtained by the combinatorial processes of *curve replacement* and *train-track nesting.*

Curve replacement is the standard process of simplifying intersections with a closed curve in a surface by surgery, which is usually used to obtain upper bounds on distance in $\mathcal{C}(S)$. We will treat this with a bit of care in Section 2, because we need to consider *well-nested curve replacements*. In Section 4 we will prove:

Date: February 10, 2003.

Theorem 1.2. *There exists $K = K(S)$ such that the vertices of a well-nested curve replacement sequence form a K-quasiconvex set in $\mathcal{C}(S)$, which moreover is Hausdorff distance K from a geodesic in $\mathcal{C}(S)$.*

In order to prove this theorem we will consider *nested train tracks* in Section 3, and prove the following statement (stated more precisely later on):

Theorem 1.3. *The vertices of a nested train-track sequence with bounded steps form a K-quasiconvex set in $\mathcal{C}(S)$, which moreover is Hausdorff distance K from a geodesic in $\mathcal{C}(S)$.*

These two theorems, together with Proposition 2.1 which produces curve-replacement sequences lying in $\Delta(M, S)$, will be enough for the proof of Theorem 1.1, which will appear at the end of Section 4.

2. Curve Replacement

A closed curve in S is *essential* if it is homotopically nontrivial and nonperipheral (not homotopic to a puncture of S). Two essential curves a and b are in *minimal position* if they intersect transversely and the number of intersection points $|a \cap b|$ is minimal over all representatives of their respective homotopy classes. This number is the geometric intersection number of the homotopy classes, written $i([a], [b])$ or just $i(a, b)$.

Consider two simple curves a and b in minimal position, and let $J \subset a$ be an interval with $J \cap b \neq \emptyset$. For any $x, y \in J$ let $[x, y]$ denote the subinterval of J with endpoints x and y, and similarly (x, y), $(x, y]$ etc.

A *curve replacement* with respect to (a, b, J) is the construction of a new curve a_1 and a new interval $J_1 \subset J$ in one of the following ways:

(1) Let w be an interval of b with $int(w) \cap a = \emptyset$, and endpoints $p, q \in int(J)$.

If w is incident to opposite sides of a at its two endpoints, Let $J_1 = [p', q']$ where p' is slightly to the right of p and q' is slightly to the right of q, so that $J_1 \cap b = (p, q] \cap b$. Let w' be a slight perturbation of w in a regular neighborhood so that the endpoints of w' are p', q' and w' and w are disjoint. Define a_1 to be the composition $w' * J_1$.

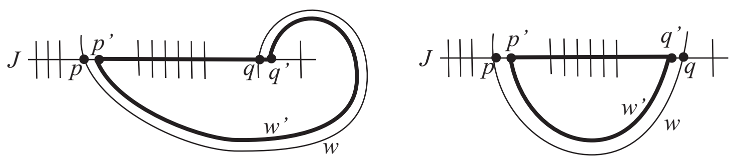

Figure 1. The two curve replacements using a single arc w. In each case a_1 is thickened.

If w is incident to the same side of a at both endpoints, we call w a *wave*. In this case, let us also assume that w is *innermost*, i.e. that there is no other wave w_2 with endpoints inside $[p, q]$ incident to w on the same side of J such that w and w_2 are homotopic with endpoints on J. In this case let $J_1 = [p', q']$ where $p', q' \in (p, q)$ and $J_1 \cap b = (p, q) \cap b$. Again move w in a regular neighborhood to obtain w' with endpoints p', q', and let $a_1 = w' * J_1$.

(2) Let w_1 and w_2 be two waves incident to J on opposite sides of a, with endpoints p_1, q_1 and p_2, q_2 respectively. Assume that $p_1 < q_1$ and $p_2 < q_2$ after fixing some orientation on J. Suppose either that $q_1 = p_2$ or $q_1 = q_2$. Move each of w_1 and w_2 slightly inward, as above; in the case $q_1 = q_2$ do this so that $q_1' = q_2'$. In either case we obtain w_i' with endpoints $p_i' < q_i'$. Now let $a_1 = w_1' * [p_1', p_2'] * w_2' * [q_2', q_1']$. Let $p' = \min(p_1', p_2')$. In the case that $q_1 = p_2$ set $q' = q_2'$, and in the case of $q_1 = q_2$ set $q' = \max(p_1', p_2')$. Finally set $J_1 = [p', q']$.

We call this a *double wave* curve replacement.

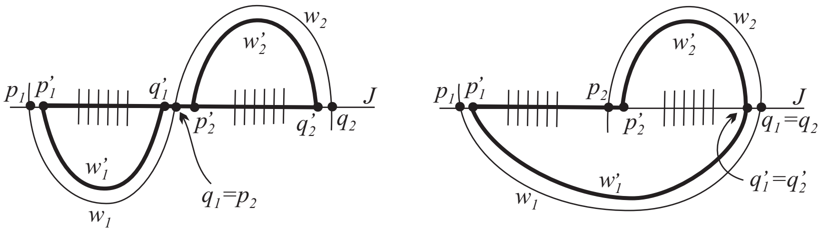

FIGURE 2. The two types of double wave replacements. Note that in the $q_1 = p_2$ case a_1 is not embedded, but can be perturbed to be.

Note that in case (1), and case (2) when $q_1 = q_2$, a_1 is simple, and in case (2) when $q_1 = p_2$ it is homotopic to a simple curve (the arcs $[p_1, p_2]$ and $[q_1, q_2]$ overlap but after moving them slightly to opposite sides of a we get a simple curve). In all cases a_1 is homotopically nontrivial, which follows from the assumption that a and b are in minimal position.

In the non-wave and double wave case it is easy to check that a_1 is also nonperipheral. If S is closed then of course a_1 is always nonperipheral.

We note the special case of a double-wave replacement with $p_1 = p_2$ and $q_1 = q_2$. In this case $i(a, b) = 2$ and a_1 is homotopic to b.

One can also check that a_1 and b are in minimal position – here in the single wave case we must use the condition that the wave is innermost.

Note that all intersections of a_1 with b lie in J_1. For the curve replacements in case (1) and the double wave replacement when $q_1 = q_2$, we have in particular

$$i(a_1, b) \le |J_1 \cap b| < |J \cap b|. \tag{1}$$

For the double wave replacement when $q_1 = p_2$ the counting is slightly different and we get:

$$i(a_1, b) \le |J_1 \cap b| + 1 < |J \cap b|. \tag{2}$$

We also define, for completeness, a curve replacement in the case that $J \cap b = \emptyset$ to be $a_1 = b$, and $J_1 = \emptyset$. If $J \cap b$ contains exactly one point then we are in the non-wave case of part 1 with $p = q$, and the replacement is again $a_1 = b$, $J_1 = \emptyset$.

Nested curve replacement. Given a and b in minimal position, a *nested curve replacement sequence* is a sequence $\{(a_i, J_i)\}$ of curves $a = a_0, a_1, a_2, \ldots, a_n$ and segments $a \supset J_0 \supset J_1 \supset \cdots \supset J_n$ such that J_0 contains all the points of $a \cap b$, and a_{i+1}, J_{i+1} are obtained by a curve replacement from (a, b, J_i).

Proposition 2.1. *Given any a, b in minimal position, and an interval $J_0 \subset a$ containing $a \cap b$, there exists a nested curve replacement sequence $\{(a_i, J_i)\}$ such that*

- *a_i is nonperipheral for all i.*
- *If S is a boundary component of a compact 3-manifold M and a and b are boundaries of compressing discs, then the a_i can be chosen to be boundaries of compressing discs.*
- *The sequence terminates with a_n homotopic to b.*

Proof. The sequence is chosen inductively. To satisfy the first statement, at each stage we choose a curve replacement configuration as in the definition, noting that if there are no non-wave curve replacements and every single wave replacement yields a peripheral curve, then there must be a double wave replacement, which will yield a non-peripheral curve. The intervals J_i are automatically nested by the definition. The process terminates, with a_n homotopic to b, because the numbers $|J_i \cap b|$ are strictly decreasing by (1) and (2).

To prove the second statement, note first that since S is closed the non-peripheral property is automatic. We will inductively construct each a_i using a wave curve replacement. We recall from the definition of curve replacements and induction, that for $i > 0$, a_i intersects b transversely and $a_i \cap b \subset J_i$ (for $i = 0$ this is true by hypothesis). Let B be a properly embedded disk in M with boundary b, and suppose by induction that there exists a properly embedded disk A_i with boundary a_i. If $i(a_i, b) = 0$ we are done, so suppose $i(a_i, b) > 0$.

We may assume that A_i and B intersect transversely. Their intersection locus can be assumed to consist of properly embedded arcs, since the closed-curve components can be removed by an exchange that does not affect a_i and b. Now let e be an innermost intersection arc on B, i.e. an arc that, together with an arc w on b bounds a disk E in B whose interior is disjoint from A_i. Thus $int(w)$ is disjoint from a_i and hence from J_i. The endpoints p, q of w lie on a_i, and hence on J_i. The arc $[p, q] \subset J_i$ together with e bounds a subdisk of F of A_i. The disk $E \cup F$ has boundary $[p, q] * w$ which, after a slight isotopy, is the curve replacement a_{i+1} obtained from (a, b, J_i) using the wave w. □

Distance in $\mathcal{C}(S)$. Using curve replacements one can bound distance in $\mathcal{C}(S)$ in terms of intersection number. One such bound is:

$$d(\alpha, \beta) \leq i(\alpha, \beta) + 1 \tag{3}$$

(see Lemma 1.1 of Bowditch [1]. For large intersection numbers a better bound is logarithmic in $i(\alpha, \beta)$ – see also Hempel [4] and Masur-Minsky [6].)

If $\{(a_i, J_i)\}$ is a curve replacement sequence, we note that $|a_{i+1} \cap J_i| \leq 2$. It follows (taking a little care when there are two double-wave replacements in a row) that

$$i(a_i, a_{i+1}) \leq 4. \tag{4}$$

From this and (3) we obtain

$$d(a_i, a_{i+1}) \leq 5. \tag{5}$$

We note that (5) is probably not sharp, even in the case of double-wave replacements, but it will suffice for our purposes. In the case of one-wave replacements (the only case relevant to Theorem 1.1), we always have $i(a_i, a_{i+1}) = 0$ and $d(a_i, a_{i+1}) \leq 1$.

3. Train tracks

A train track on a surface S is an embedded 1-complex τ satisfying the following properties (see Penner-Harer [10] for more details). Each edge (called a branch) is a smooth path with well-defined tangent vectors at the endpoints, and at any vertex, (called a switch) the incident edges are mutually tangent. The tangent vector at the switch pointing toward the interior of an edge can have two possible directions, dividing the ends of edges into "outgoing" and "incoming" directions, neither of which can be empty. For the tracks that we consider there will be a single switch. Each component of $S \setminus \tau$ has negative generalized Euler characteristic which is the usual Euler characteristic minus 1/2 for every outward pointing cusp (internal angle 0). This means that we exclude annuli, once-punctured discs with smooth boundary, or unpunctured discs with 0, 1, or 2 cusps. The last are called bigons.

A *train route* is a nondegenerate smooth path in τ. It traverses a switch only by passing from incoming to outgoing edge or vice versa. A *transverse measure* on τ is a non-negative function μ on branches satisfying the switch condition that for any switch the sums of μ over outgoing edges equals the sum over incoming. A closed train route induces counting measure on τ.

A train track is *recurrent* if every branch is contained in a closed train route. A curve β is *carried* on τ if there is a homotopy of S taking β to a set of train routes. Then β induces a measure on τ which uniquely determines it.

A train track σ is *carried* by τ, written $\sigma \prec \tau$, if there is a homotopy of the surface such that every train route of σ is taken to a train to a train route of τ. We also say that σ is *nested* in τ. We write $\sigma < \tau$ if σ is a *subtrack* of τ; that is, σ is a train track which is a subset of τ.

For a recurrent train track τ let $P(\tau)$ denote the polyhedron of transverse measures supported on τ. We note that $P(\tau)$ is preserved by scaling so it is a cone on a compact polyhedron in a projective space. We can consider all of these polyhedra as subsets of the measured lamination space $\mathcal{ML}(S)$, and then $\sigma \prec \tau$ is equivalent to $P(\sigma) \subset P(\tau)$, and $\sigma < \tau$ is equivalent to $P(\sigma)$ being a face of $P(\tau)$. By $int(P(\tau))$ we will mean the set of measures on τ which are positive on every branch (recurrence implies that this is nonempty).

A *vertex* of the projectivization of $P(\tau)$ is an extreme point, and corresponds to a line in $\partial P(\tau)$. By abuse of notation we will let "vertex" denote any non-zero point on this line. After scaling, a vertex can always be realized by the counting measure on a single simple closed curve (see [6, §4.1]), which we call a *vertex cycle.*

Because the set $vert(\tau)$ of vertex cycles of τ is finite and there are finitely many homeomorphism classes of train tracks in S, there is a constant B such that

$$diam_{\mathcal{C}(S)}\{vert(\tau)\} \leq B.$$

For tracks τ and σ, define

$$d_T(\tau, \sigma) = \min_{\substack{\beta \in vert(\tau) \\ \alpha \in vert(\sigma)}} d(\beta, \alpha)$$

Note that this is not a distance function on the set of tracks.

Our goal in this section is to prove Theorem 1.3, which we restate here more precisely:

Theorem 1.3. *Given $M > 0$ there exists K such that if $\cdots \prec \tau_n \prec \tau_{n-1} \prec \cdots \prec \tau_1 \prec \tau_0$ is a sequence of nested train tracks such that $d_T(\tau_{i+1}, \tau_i) \leq M$ then the set of vertices of the τ_i is K-quasiconvex in $\mathcal{C}(S)$.*

A track is *large* if all complementary components of the track are disks or once-punctured disks. We have:

Lemma 3.1. *If a train track τ is not large then for any α, β carried by τ, $d(\alpha, \beta) \leq 2$.*

Proof. Since τ is not large, there is an essential simple closed curve γ which misses τ, and hence both α and β; hence $d(\alpha, \beta) \leq 2$. □

This lemma will mean that we will be able to restrict our attention to large tracks. Let σ be a large track. A *diagonal extension* of σ is a track κ such that $\sigma < \kappa$ and every branch of $\kappa \setminus \sigma$ is a diagonal of σ. This means that it is an edge that terminates in corners of a complementary region of σ.

Let $E(\sigma)$ denote the (finite) set of all recurrent tracks which are diagonal extensions of σ. Let

$$PE(\sigma) = \bigcup_{\kappa \in E(\sigma)} P(\kappa)$$

Let $int(PE(\sigma))$ denote the set of measures $\mu \in PE(\sigma)$ which are positive on every branch of σ. Next for τ a large track, let

$$N(\tau) = \bigcup_{\substack{\sigma < \tau \\ \sigma \text{ large}}} E(\sigma),$$

and define

$$PN(\tau) = \cup_{\kappa \in N(\tau)} P(\kappa).$$

This can be thought of as a neighborhood of $P(\tau)$. Let

$$int(PN(\tau)) = \bigcup_{\substack{\sigma < \tau \\ \sigma \text{ large}}} int(PE(\sigma)).$$

We will need the following lemmas proved in [6].

Lemma 3.2. *There is a constant D such that if $\sigma \prec \tau$ are a pair of large recurrent tracks and $d_T(\tau, \sigma) \geq D$, then $PN(\sigma) \subset int(PN(\tau))$*

This lemma is stated in [6] for *generic tracks*, i.e. those with trivalent switches, but any track can be perturbed to a generic track without changing its vertex set or polyhedron of measures (by "combing" – see Penner-Harer [10, §1.4]), and it is easy to see that this can be done simultaneously for σ and τ.

Lemma 3.3. *If σ is a large track, $\gamma \in int(PE(\sigma))$ and $d(\gamma, \beta) \leq 1$, then $\beta \in PE(\sigma)$.*

In fact this implies

Lemma 3.4. *If σ is a large track then*

$$\mathcal{N}_1(int(PN(\sigma))) \subset PN(\sigma)$$

where $\mathcal{N}_k(int(PN(\sigma)))$ denotes the k-neighborhood in $\mathcal{C}(S)$ of the vertices of $\mathcal{C}(S)$ contained in $int(PN(\sigma))$.

We will also need the following lemma.

Lemma 3.5. *Let τ be a large recurrent track and v a vertex of τ. Then $v \notin int(PN(\tau))$.*

Proof. First we claim that $dim(P(\tau)) \geq 2$ for any large recurrent track τ. Let b be the number of branches, s the number of switches and f the number of complementary discs. Then

$$\chi(S) = f - b + s$$

so

$$dim(P(\tau)) \geq b - s = f - \chi(S),$$

since $b - s$ is a lower bound for the dimension of the solution set to the switch conditions and the recurrence of τ gives us at least one solution with all positive weights. The latter quantity is at least 2 except in the case that S is a punctured torus, where $\chi(S) = -1$ and $f = 0$. However in this case it is easily checked directly that $dim(P(\tau)) = 2$, so the claim is proven.

Suppose now that the lemma is false. By definition then, there must be some large subtrack $\sigma < \tau$, and track $\kappa \in E(\sigma)$ such that a representative α_1 of v is carried by κ and assigns positive weights to the branches of σ. We may assume that α_1 assigns positive weights to every branch of κ as well – deleting branches of κ if necessary. Let α_2 be a representative of v carried on τ. We may assume that α_2 puts positive measure on every branch of $\tau \setminus \sigma$, otherwise we may remove those branches, obtaining a smaller track containing σ for which v is a vertex. We now wish to arrive at a contradiction.

By an *admissible deformation* of an edge e of $\kappa \setminus \sigma$, we mean a homotopy, preserving endpoints and tangency to σ, to a path e' such that $\sigma \cup e'$ is still a train track. In particular e' is allowed to have segments tangent to σ.

We say an edge e of $\kappa \setminus \sigma$ intersects an edge f of $\tau \setminus \sigma$ *inessentially* if e can be deformed admissibly to an edge e' which is either disjoint from the interior of f, or traverses a train route in $\sigma \cup f$. Suppose that all intersections of $\tau \setminus \sigma$ with $\kappa \setminus \sigma$ are inessential. Then κ can be deformed to be carried in a track $\hat{\tau}$ containing σ, on which v puts a measure positive on every edge, and such that τ and $\hat{\tau}$ are subtracks of a common train track ω.

Suppose first that $\hat{\tau} < \tau$. Since σ is large so is $\hat{\tau}$, but then by the first claim $dim(P(\hat{\tau})) \geq 2$, which means the projectivized polyhedron has dimension at least 1, and any extreme point could not be in its interior. Thus $v \in int(P(\hat{\tau}))$ means that v is not extreme in $P(\hat{\tau})$, and hence not in $P(\tau)$.

Next suppose that $\hat{\tau}$ is not a subtrack of τ. In this case v is represented by two distinct measures on ω, which contradicts the injectivity of the map $P(\omega) \to \mathcal{ML}(S)$.

We are left with the possibility that an edge b_τ of $\tau \setminus \sigma$ has essential intersection with b_κ in $\kappa \setminus \sigma$. Both edges lie in some complementary disc or once punctured disc R of $S \setminus \sigma$, and we note that b_κ is a diagonal but b_τ does not need to be. The representative α_2 of v assigns positive weight to b_τ, and the representative α_1 assigns positive weight to b_κ.

We may assume that b_τ and b_κ have a unique intersection point x if R is a disk, or possibly two intersection points if R is a punctured disk. Let x be one of the points in the latter case. We may also assume, possibly after admissible deformations of κ, that there are no inessential intersections of κ and τ, and that every edge of κ that is admissibly deformable into τ already lies in τ.

Let $\widetilde{\tau}$ and $\widetilde{\kappa}$ be the lifts of the tracks τ and κ to the universal cover. Each contains the lift $\widetilde{\sigma}$ of σ. Let $\widetilde{x} \in \widetilde{R}$ be lifts of x and R. Let $\widetilde{\alpha}_1, \widetilde{\alpha}_2$ be lifts of α_1, α_2 that pass through $\widetilde{x}$, and let $\widetilde{b}_\tau, \widetilde{b}_\kappa$ be the lifts of b_τ and b_κ that intersect at $\widetilde{x}$. Since $\alpha_1 \sim \alpha_2$, $\widetilde{\alpha}_1$ and $\widetilde{\alpha}_2$ must intersect at some other point $\widetilde{y}$ so that the segments of $\widetilde{\alpha}_1$ and $\widetilde{\alpha}_2$ from $\widetilde{x}$ to $\widetilde{y}$ bound a disc D. The point $\widetilde{y}$ may or may not be in the interior of a complementary domain of $\widetilde{\sigma}$.

We will now build a track $\widetilde{\sigma}_1$ that contains $\widetilde{\sigma}$. To form $\widetilde{\sigma}_1$, add to $\widetilde{\sigma}$ any branch of $\widetilde{\alpha}_1$ and $\widetilde{\alpha}_2$ that is entirely contained in ∂D. Since ∂D is embedded none of these branches cross and we obtain a train track. Notice that we do not introduce any bigons, since neither τ nor κ have any, and we have already deformed κ so that no edge of $\kappa \setminus \tau$ is admissibly deformable into τ. Thus any complementary disc of $\widetilde{\sigma}_1$ still has negative generalized Euler characteristic.

Next consider $\widetilde{R} \cap D$. It is bounded by one or more edges of $\widetilde{\sigma}$, a subsegment of $\widetilde{b}_\tau$ that goes from a switch $\widetilde{P}_1$ to $\widetilde{x}$ and a subsegment of $\widetilde{b}_\kappa$ from $\widetilde{x}$ to a switch $\widetilde{P}_2$. Replace this latter pair of segments with a smooth path joining $\widetilde{P}_1$ to $\widetilde{P}_2$ in the same homotopy class rel endpoints (and maintaining tangency at the switches). The assumption that the intersection is essential implies that the subdisc $\tilde{R}'$ of D bounded by this smooth path and the edges of $\widetilde{\sigma}$ is not a monogon. If $\tilde{R}'$ is a not a bigon, we add this edge to $\widetilde{\sigma}_1$ and replace D with $D' = (D \setminus \widetilde{R}) \cup \widetilde{R}'$. Note that $\widetilde{\sigma}_1$ still has no bigons. If $\tilde{R}'$ is a bigon, the path can be admissibly deformed into $\widetilde{\sigma}$ and we do not add it to $\widetilde{\sigma}_1$. Now let $D' = D \setminus \widetilde{R}$

We perform the same construction if $\widetilde{y}$ is in the interior of a complementary domain as well, noting that $\widetilde{y}$ is also an essential intersection point. The new disc $\widetilde{D}''$ has smooth boundary except possibly for one cusp at the point $\widetilde{y}$ (if $\widetilde{y}$ was not an interior point). Thus its generalized Euler characteristic is positive. However it is a union of complementary discs of $\widetilde{\sigma}_1$, each of which has negative generalized Euler characteristic. Since generalized Euler characteristic adds, we have a contradiction. $\square$

We are now ready to give the proof of the quasiconvexity theorem for nested sequences of tracks.

Proof of Theorem 1.3. Let $\{\tau_n\}$ be our nested sequence, so that $\tau_{i+1} \prec \tau_i$ and $d_T(\tau_i, \tau_{i+1}) \leq M$. We may assume that all τ_i are large. For if some τ_i is not large so are all τ_j for $j \geq i$, and since all are carried in τ_i, Lemma 3.1 implies that $d_T(\tau_j, \tau_k) < 3$ for all $j, k \geq i$, and hence we may ignore these tracks. Using the number D given by Lemma 3.2, the condition $d_T(\tau_i, \tau_{i+1}) \leq M$ implies that we may inductively find a subsequence τ_{i_j} (with $i_0 = 0$) such that

$$D \leq d_T(\tau_{i_j}, \tau_{i_{j+1}}) < D + M \tag{6}$$

and for any τ_n we have

$$d_T(\tau_{i_j}, \tau_n) < D \tag{7}$$

for some i_j.

Lemma 3.2 then says that

$$PN(\tau_{i_{j+1}}) \subset int(PN(\tau_{i_j})),$$

and Lemma 3.4 says that

$$\mathcal{N}_1(int(PN(\tau_{i_j}))) \subset PN(\tau_{i_j}).$$

Combining these and applying induction we find that

$$\mathcal{N}_{k-1}(PN(\tau_{i_{j+k}})) \subset int(PN(\tau_{i_j})).$$

Applying this to the vertices, and using Lemma 3.5 we find

$$d_T(\tau_{i_j}, \tau_{i_k}) \geq |k-j| \tag{8}$$

for any j,k in the subsequence.

Inequality (6), together with the bound $diam\{vert(\tau)\} \leq B$, implies that any sequence of vertices of the τ_{i_j} forms a K-quasigeodesic with $K = K(D,M,B)$. By δ-hyperbolicity of $\mathcal{C}(S)$, any geodesic joining the vertices of this sequence stays within some fixed distance $M'(M,B,D,\delta)$ of it. Thus $\bigcup vert(\tau_{i_j})$ is M'-quasi-convex.

Now by (7), $vert(\tau_i)$ for any i is within $D+B$ of some $vert(\tau_{i_j})$. Again by δ-hyperbolicity, there exists $M'' = M''(B,D,\delta)$ such that any geodesic joining vertices of the τ_i stays within distance M'' of a geodesic joining some pair of vertices of the τ_{i_j}.

Combining these two facts we see that $\bigcup vert(\tau_i)$ is a $M' + M''$ quasiconvex set, and moreover that it is Hausdorff distance $M' + M''$ from a geodesic connecting vertices of τ_{i_0} and the last τ_{i_j}. □

4. Quasiconvexity of curve replacements

In this section we will restate Theorem 1.2 more precisely, and apply Theorem 1.3 to prove it. In order to do this, given a nested curve replacement sequence (a_i, J_i) we will construct a sequence of nested train tracks τ_i carrying b such the vertices of τ_i are a bounded $\mathcal{C}(S)$-distance from $[a_i]$.

One-switch train tracks. Let a and b be two essential curves intersecting minimally and let J be an interval in a such that $J \cap b$ is non-empty. We wish to construct a train track τ_J closely associated to the "first return map" of b to J. The quotient of $J \cup b$ obtained by squeezing J to a point P is a 1-complex, and we may impose a switch structure on P by declaring that tangent vectors at P pointing to interiors are all outgoing for edges leaving a on one side and incoming for edges leaving a on the other. This makes the quotient, which we call τ_J^0, a "bigon track": it satisfies all the conditions for being a train track except for the possibility of bigon complementary regions. The other disallowed types of complementary regions fail to occur because a and b intersect minimimally.

If we identify the opposite sides of any bigon in τ_J^0 we obtain a new 1-complex τ_J, still with a switch structure. This will be a train track provided that the identification can be performed as a homotopy of the surface S, and in this case we say that J is a *good interval.* The following lemmas will allow us to use this construction effectively:

Lemma 4.1. *Let a and b be essential curves in S intersecting minimally.*

1. *There exists a good interval $J \subset a$ containing all points of $a \cap b$.*
2. *Any subinterval of a good interval is good.*

Proof. A component Q of $S \setminus a \cup b$ is a *rectangle* if it is a disk and its boundary, traversed from the inside, is a path meeting exactly four intersection points of $a \cap b$. By an Euler characteristic argument, there must be at least one region Q which is not a rectangle. Let $I \subset a$ be an open interval between two successive points x_1, x_2 of $a \cap b$, which lies on the boundary of a non-rectangle Q. We claim that $J = a \setminus I$ is a good interval.

To see this, squeeze J to a point and let Ω be the complementary region of the resulting τ_J^0 which contains Q. For each $i = 1, 2$ the arc of b that passes through x_i is on the boundary of Ω and traverses the switch P smoothly. If Ω were a bigon, its boundary would consist of two arcs of b passing smoothly through x_1 and x_2 and meeting in cusps at both ends. The preimage of these cusps would consist of two arcs of $b \setminus a$, and this would force Q to be a rectangle, a contradiction. Furthermore, because τ_J^0 has a single switch P, the arcs of b through x_i are the only smooth arcs through P on the boundary of complementary regions. As they are on the boundary of Ω which is not a bigon, every bigon of τ_J^0 meets the switch only at its cusps, and hence has exactly two branches of τ_J^0 in its boundary. Consider a connected component of the union of bigons in the complement of the switch. It must be a disk (with punctured boundary) with two branches in its boundary – the only other way of gluing such bigons together gives a sphere minus two points, which is impossible. Hence the collapse of bigons can be performed on each connected component, yielding the desired track τ_J.

For part (2): Suppose $J \subset a$ is good, let $J' \subset J$ be an interval which intersects b, and suppose B is a bigon complementary region of $\tau_{J'}^0$. Every arc of $b \setminus J'$ is a train route for τ_J^0, so clearly $\tau_{J'}^0 \prec \tau_J^0$. Thus after a homotopy of S we can assume the branches of $\tau_{J'}^0$ travel along branches of τ_J^0, and B is taken by this homotopy to a collection of complementary regions of τ_J^0, attached smoothly along arcs of τ_J^0. The generalized Euler characteristic of any complementary region is nonpositive and zero if and only if it is a bigon or an annulus. Since B is simply connected and the characteristic of a union of complementary regions is additive, these regions must necessarily be bigons. Hence the bigon collapse $\tau_J^0 \to \tau_J$ also collapses B. Since this collapse can be performed by a homotopy of S, this same homotopy collapses all the bigons of $\tau_{J'}^0$, so J' is good. ☐

With this in mind, let us say that $\{(a_i, J_i)\}$ is a *well-nested curve replacement sequence* if it is a nested curve replacement sequence and $J_0 \subset a$ is a good interval. We can now restate Theorem 1.2.

Theorem 1.2. *There exists $K = K(S)$ such that, if a and b are in minimal position and $\{(a_i, J_i)\}$ is a well-nested curve replacement sequence, then the vertices $[a_i]$ form a K-quasiconvex set in $\mathcal{C}(S)$ which moreover are Hausdorff distance K from a geodesic in $C(S)$.*

Proof. Now let $J_0 \subset a$ be a good interval which contains $a \cap b$, and let $\{(a_i, J_i)\}$ be a curve replacement sequence, terminating with $[a_n] = [b]$. Using Lemma 4.1, each J_i is good, and so we obtain a sequence of tracks $\tau_i = \tau_{J_i}$. As in the proof of Lemma 4.1, we have:

Lemma 4.2. *For each i, $\tau_{i+1} \prec \tau_i$, and furthermore each τ_i carries b.*

Next we observe that

$$d(a_i, vert(\tau_i)) \leq 5 \tag{9}$$

To prove (9), we observe that a_i is composed of one or two arcs of b and one or two arcs along J_i. Hence after after isotopy it intersects τ_i at most twice at the switch P. Each vertex cycle of τ_i passes through the switch at most twice. Hence $i(a_i, v) \leq 4$ for any vertex v of τ_i, and the distance bound follows from (3).

Since, by (5), $d(a_i, a_{i+1}) \leq 5$, we may conclude by the triangle inequality that

$$d_T(\tau_i, \tau_{i+1}) \leq 15.$$

(As in Section 2, we note that this is probably not sharp. In the case of one-wave replacements the bound obtained is 7.)

Thus the hypotheses of Theorem 1.3 are satisfied, so that $\{vert(\tau_i)\}$ form a quasiconvex set. Inequality (9) then gives a quasiconvexity bound for $\{[a_i]\}$, completing the proof of Theorem 1.2. $\square$

Proof of the main theorem. The proof of Theorem 1.1 can now be easily assembled:

Given two vertices of $\Delta(M, S)$, represent them in minimal position as a, b and let $J_0 \subset a$ be a good interval containing $a \cap b$ (by Lemma 4.1). Proposition 2.1 gives us a nested curve replacement starting with (a_0, J_0), which is well-nested by choice of J_0, and whose vertices are all in $\Delta(M, S)$. By Theorem 1.2, this set is K-quasiconvex. It follows that $\Delta(M, S)$ is K-quasiconvex. $\square$

References

[1] B. Bowditch, *Intersection numbers and the hyperbolicity of the curve complex*, Preprint, Southampton.

[2] J. Brock, R. Canary, and Y. Minsky, *Classification of Kleinian surface groups II: the ending lamination conjecture*, in preparation.

[3] A. J. Casson and C. McA. Gordon, *Reducing Heegaard splittings*, Topology Appl. **27** (1987), no. 3, 275–283.

[4] J. Hempel, *3-manifolds as viewed from the curve complex*, Topology **40** (2001), no. 3, 631–657.

[5] K. Johannson, *Topology and combinatorics of 3-manifolds*, Lecture Notes in Mathematics, vol. 1599, Springer-Verlag, Berlin, 1995.

[6] H. A. Masur and Y. Minsky, *Geometry of the complex of curves I: Hyperbolicity*, Invent. Math. **138** (1999), 103–149.

[7] ———, *Geometry of the complex of curves II: Hierarchical structure*, Geom. Funct. Anal. **10** (2000), 902–974.

[8] Y. Minsky, *Bounded geometry in Kleinian groups*, Invent. Math. **146** (2001), 143–192.

[9] ———, *Classification of Kleinian surface groups I: models and bounds*, preprint, 2002.

[10] R. Penner and J. Harer, *Combinatorics of train tracks*, Annals of Math. Studies no. 125, Princeton University Press, 1992.

[11] S. Schleimer, *The disjoint curve property*, preprint, 2002.