



RESEARCH ARTICLE

From curves to currents

Dídac Martínez-Granado ¹ and Dylan P. Thurston ²

¹Department of Mathematics, University of California, Davis, One Shields Ave, Davis, 95616, USA;
E-mail: dmartinezgranado@math.ucdavis.edu.

²Department of Mathematics, Indiana University, 831 East 3rd St., Bloomington, 47405, USA;
E-mail: dpthurst@indiana.edu.

Received: 1 November 2020; **Revised:** 22 August 2021; **Accepted:** 18 September 2021

Keywords: geodesic currents, extremal length, curve counting

2020 Mathematics Subject Classification: *Primary* – 57M50; *Secondary* – 37E30

Abstract

Many natural real-valued functions of closed curves are known to extend continuously to the larger space of geodesic currents. For instance, the extension of length with respect to a fixed hyperbolic metric was a motivating example for the development of geodesic currents. We give a simple criterion on a curve function that guarantees a continuous extension to geodesic currents. The main condition of our criterion is the smoothing property, which has played a role in the study of systoles of translation lengths for Anosov representations. It is easy to see that our criterion is satisfied for almost all known examples of continuous functions on geodesic currents, such as nonpositively curved lengths or stable lengths for surface groups, while also applying to new examples like extremal length. We use this extension to obtain a new curve counting result for extremal length.

Contents

1	Introduction	1
2	Background on curves and currents	5
3	Convexity and continuity	12
4	Examples	14
5	Counting problems	24
6	Proof outline	26
7	Defining the extension	27
8	Constructing global cross sections	34
9	Join lemma	36
10	Continuity of the extension	41
11	The extension extends	44
12	Hyperbolic geometry estimates	46
13	Stable functions	48

1. Introduction

Geodesic currents on surfaces are measures that realise a suitable closure of the space of weighted (multi-)curves on a surface. They were first introduced by Bonahon in his seminal paper [10]. Many metric structures can be embedded in the space of currents, such as hyperbolic metrics [8, Theorem 12]

or half-translation structures [20, Theorem 4]. Thus, geodesic currents allow one to treat curves and metric structures on surfaces as the same type of object. Via this unifying framework, counting curves of a given topological type and counting lattice points in the space of deformations of geometric structures become the same problem [51, Main Theorem]. Geodesic currents also play a key step in the proof of rigidity of the marked length spectrum for metrics, via an argument by Otal [47, Théorème 2]. Finally, they provide a boundary of the Teichmüller space, in both the compact [8, Proposition 17] and noncompact [11, Theorem 2] cases.

In this article we consider the problem of extending continuously a function defined on the space of weighted multi-curves to its closure, the space of geodesic currents.

Previous work of Bonahon extended the notion of geometric intersection number as a continuous function of two geodesic currents [10, Proposition 4.5]. This allowed him to extend hyperbolic length to geodesic currents by following the principle of realising it as an intersection number with a distinguished geodesic current [8, Proposition 14].

The same principle using intersection numbers has been used by many authors to extend length for many other metrics: Otal for negatively curved Riemannian metrics [47, Proposition 3], Croke–Fathi–Feldman for nonpositively curved Riemannian metrics [16, Theorem A], Hersonsky–Paulin for negatively curved metrics with conical singularities [32, Theorem A], Bankovic–Leininger for nonpositively curved Euclidean cone metrics [2], Duchin–Leininger–Rafi more explicitly for singular Euclidean structures associated to quadratic differentials [20, Lemma 9] and Erlandsson for word length with respect to simple generating sets of the fundamental group [22, Theorem 1.2].

Another line of results on extending functions to geodesic currents was also started by Bonahon, who showed how to extend stable lengths to geodesic currents, not just for surface groups but for general hyperbolic groups [9, Proposition 10]. This result was recently improved by Erlandsson–Parlier–Souto [23, Theorem 1.5], who used the return map of the geodesic flow to remove technical assumptions. These constructions apply, for instance, to arbitrary Riemannian metrics and the stable version of word lengths for arbitrary generating sets.

The problem of extending functions to geodesic currents is interesting in itself, since, by a result of Rafi and Souto reviewed in Section 5, it provides a way to compute asymptotics of the number of curves of a fixed type with a bounded ‘length’, for a notion of ‘length’ that extends to currents [51]. Their result builds on work by Mirzakhani [44, Theorem 7.1] and Erlandsson–Souto [24, Proposition 4.1]. Recently, Erlandsson and Souto gave a new argument to compute these asymptotics [26, Theorem 8.1].

Our main theorem gives a simple criterion on functions defined on multi-curves that guarantees they extend to geodesic currents. Our result subsumes most of the previous extension results mentioned above and provides new extensions for other notions of ‘length’, such as extremal length, thus yielding counting asymptotics for them.

Our proof does not use Bonahon’s principle on intersection numbers. Although we drew some inspiration from the dynamics of Erlandsson–Parlier–Souto [23], our techniques are distinct.

1.1. Main results

We start by summarising our main results. Complete definitions of the terms are deferred to Section 2.

Definition 1.1. For S a compact topological surface without boundary, let $f: \mathcal{C}^+(S) \rightarrow \mathbb{R}$ be a function defined on the space of oriented multi-curves, not-necessarily-simple oriented curves; see Definition 2.1, and see Table 1 for a summary of notation. We will also refer to f as a *curve functional* for short. (*Functional* means that it takes values in scalars; it is not assumed to be linear.) We will also refer to *unoriented* or *weighted* curve functionals for real-valued functions defined on the appropriate type of multi-curves. We define several properties that f might satisfy.

Table 1. Notation for the objects related to surfaces, curves and geodesic currents.

Notation	Meaning
S	topological surface
Σ	Riemannian surface
$UT\Sigma$	unit tangent bundle over Σ
ϕ_t	geodesic flow on $UT\Sigma$
τ	cross section to the geodesic flow
ψ	bump function on a cross section
$\mathcal{S}(S)$	unoriented simple multi-curves
$\mathcal{C}(S)$	unoriented multi-curves
$\mathcal{S}^+(S)$	oriented simple multi-curves
$\mathcal{C}^+(S)$	oriented multi-curves
$\tilde{G}(S)$	unoriented geodesics on \tilde{S}
$G^+(S)$	oriented geodesics on \tilde{S}
$\mathbb{R}\mathcal{S}(S)$	weighted unoriented simple multi-curves
$\mathbb{R}\mathcal{C}(S)$	weighted unoriented multi-curves
$\mathbb{R}\mathcal{S}^+(S)$	weighted oriented simple multi-curves
$\mathbb{R}\mathcal{C}^+(S)$	weighted oriented multi-curves
$\mathcal{GC}(S)$	unoriented geodesic currents
$\mathcal{GC}^+(S)$	oriented geodesic currents
γ	concrete multi-curve on S
C	multi-curve on S

- **Quasi-smoothing:** There is a constant $R \geq 0$ with the following property. Let C be an oriented curve on S , and let x be an essential crossing of C . Let C' be the oriented smoothing of C at x . Then $f(C) \geq f(C') - R$. Schematically, we have

$$f\left(\text{diagram with crossing}\right) \geq f\left(\text{diagram with smoothing}\right) - R \tag{1.2}$$

See Definition 2.6 for ‘essential crossing’. Loosely, it is a crossing that cannot be removed by homotopy. See Definition 2.7 for ‘oriented smoothing’.

- **Smoothing:** We take $R = 0$ in the above definition of quasi-smoothing:

$$f\left(\text{diagram with crossing}\right) \geq f\left(\text{diagram with smoothing}\right). \tag{1.3}$$

- **Convex union:** Let C_1 and C_2 be two oriented curves on S . Then

$$f(C_1 \cup C_2) \leq f(C_1) + f(C_2). \tag{1.4}$$

- **Additive union:** The inequality in convex union becomes an equality:

$$f(C_1 \cup C_2) = f(C_1) + f(C_2). \tag{1.5}$$

Many natural curve functionals satisfy the additive union property; for instance, length with respect to an arbitrary length metric on S satisfies it by definition. The square root of extremal length is an example of a curve functional satisfying convex union but not additive union (Subsection 4.8). The name ‘convex union’ comes from the fact that if we extend to weighted curves and additionally assume homogeneity (Definition 1.7), then for a fixed oriented multi-curve with varying weights, f is a convex as a function of the weights (Proposition 3.4).

There are many curve functionals satisfying the smoothing property, such as hyperbolic length, extremal length, intersection number with a fixed curve or length from a length metric on S . For an example of a natural curve functional that satisfies quasi-smoothing but not smoothing, we have the

word length with respect to an arbitrary generating set of $\pi_1(S)$ (Example 4.10). The (quasi-)smoothing property is usually easy to check.

The smoothing property plays an important role in the study of translation lengths associated to Anosov representations, as we discuss in Subsection 4.7, following Martone and Zhang [38] and Burger et al. [13]. These works use the smoothing property to reduce the study of length systoles to the case of simple closed curves. Although these papers point out the parallelism between the smoothing property of Anosov translation lengths and that of hyperbolic length ([13, Section 4] or negatively curved lengths [38, Corollary 1.3], the results in our article reveal the much more universal nature of the smoothing property. Indeed, we show many other natural notions of lengths which are not associated to negatively curved structures satisfy the smoothing condition, such as lengths on Riemannian metrics with no curvature assumption, lengths coming from more general length space structures, extremal lengths or word lengths with respect to certain generating sets; see Section 4.

Definition 1.6. For C an oriented multi-curve, nC is the oriented multi-curve that consists of n parallel copies of C (so with n times as many components or, in the context of weighted oriented multi-curves, with weights multiplied by n), and C^n is the oriented multi-curve with as many components as C , in which each component of C is covered by an n -fold cover. That is, if $g \in \pi_1(S, x)$ represents C , g^n represents C^n .

Definition 1.7. Let f be a curve functional and let $n > 0$ be an integer. We define some properties f might satisfy:

- **Homogeneity:** For an arbitrary oriented multi-curve C ,

$$f(nC) = nf(C). \quad (1.8)$$

- **(Weak) stability:** For an arbitrary oriented multi-curve C ,

$$f(C^n) = f(nC). \quad (1.9)$$

- **Strong stability:** For arbitrary oriented multi-curves C, D ,

$$f(D \cup C^n) = f(D \cup nC). \quad (1.10)$$

Additive union implies homogeneity, and if f satisfies convex union, $f(nC) \leq nf(C)$. If f satisfies quasi-smoothing, then $f(nC) - nR \leq f(C^n)$, since the self-crossings in C^n are essential crossings by definition (see Definition 2.6).

We furthermore note that curve functionals are not necessarily positive.

With this background, we can state our main theorems on extensions of curve functionals to the space $\mathcal{GC}^+(S)$ of oriented geodesic currents.

Theorem A. *Let f be a curve functional satisfying the quasi-smoothing, convex union, stability and homogeneity properties. Then there is a unique continuous homogeneous function $\bar{f}: \mathcal{GC}^+(S) \rightarrow \mathbb{R}_{\geq 0}$ that extends f .*

In the case of unoriented curves, there are two possible smoothings of an essential crossing, not distinguished from each other. Then we have the following version of the theorem, deduced from Theorem A in Subsection 2.5.

Corollary 1.11. *Let f be an unoriented curve functional satisfying quasi-smoothing for both possible smoothings of a crossing, in addition to the convex union, stability and homogeneity properties. Then there is a unique continuous homogeneous function $\bar{f}: \mathcal{GC}(S) \rightarrow \mathbb{R}$ that extends f .*

Theorem A should be thought of as an analogue of the classical theorem that a convex function defined on the rational points in a finite-dimensional vector space automatically extends continuously to a convex function defined on the whole vector space (Proposition 3.1(iv)). As in the classical case, the functions on geodesic currents arising from this construction are restricted, as the next example shows.

(This example is almost the only function we are aware of where our techniques do not suffice to prove continuity of the extension. See Subsection 4.1.)

Example 1.12. Consider the curve curves given by the square root of self-intersection number; that is, $f(C) := \sqrt{i(C, C)}$. Since intersection number is a continuous two-variable function [10, Proposition 4.5], it follows that f extends continuously to geodesic currents. However, f does not satisfy convex union. For instance, take C_1 and C_2 to be two simple closed curves intersecting once. For any multi-curve, $i(C, C)$ is twice the self-intersection number of C . Thus, $f(C_1 \cup C_2) = \sqrt{2}$, but $f(C_1) + f(C_2) = 0$, contradicting convex union. On the other hand, f clearly satisfies smoothing.

On the other hand, the stability and homogeneous properties are necessary conditions for an extension to exist for elementary reasons, as the multi-curves nC and C^n should represent the same currents (Example 4.10). However, a curve functional that satisfies quasi-smoothing and convex union can be modified to get a curve functional satisfying all the hypotheses of Theorem A.

Theorem B. *Let f be a curve functional satisfying quasi-smoothing and convex union. Then the stabilised curve functional*

$$\|f\|(C) := \lim_{n \rightarrow \infty} \frac{f(C^n)}{n}$$

satisfies quasi-smoothing, convex union, strong stability and homogeneity and thus extends to a continuous function on $\mathcal{GC}^+(S)$.

Theorem B is proved in Subsection 13, although the implication that weak stability implies strong stability is used in the proof of Theorem A.

If the convex union property of f is strengthened to additive union and the quasi-smoothing property is strengthened to smoothing, then in fact this extension to geodesic currents comes from intersection with a fixed current (as in the proofs of extension that used Bonahon's principle). This will appear in a forthcoming paper. For this stronger result, the strict smoothing property is necessary, since intersection number cannot increase after smoothing an essential crossing.

2. Background on curves and currents

Throughout this article, S is a fixed oriented compact 2-manifold without boundary. (For a discussion of the more general surface case, see Remark 2.26.) If we fix an arbitrary (hyperbolic) metric on S , we will denote it by Σ . The various types of curves and associated objects we consider are summarised in Table 1.

2.1. Curves

Definition 2.1 (multi-curve). A *concrete multi-curve* γ on a surface S is a smooth 1-manifold without boundary $X(\gamma)$ together with a map (also called γ) from $X(\gamma)$ into S . $X(\gamma)$ is not necessarily connected. We say that γ is *trivial* if it is homotopic to a point. Two concrete multi-curves γ and γ' are *equivalent* if they are related by a sequence of the following moves:

- *homotopy* within the space of all maps from $X(\gamma)$ to S ;
- *reparametrisation* of the 1-manifold; and
- *dropping* trivial components.

The equivalence class of γ is denoted by $[\gamma]$, and we will call it a *multi-curve*. If $X(\gamma)$ is connected, we will call $[\gamma]$ a *curve*; a curve is equivalent to a conjugacy class in $\pi_1(S)$. When we just want to refer to the equivalence class of a (multi-)curve, without distinguishing a representative, we will use capital letters such as C . A concrete multi-curve γ is *simple* if γ is injective, and a multi-curve is simple if it

has a concrete representative that is simple. We write $\mathcal{S}(S)$ for the space of simple multi-curves on S and $\mathcal{C}(S)$ for the space of all multi-curves.

We also consider *oriented multi-curves*, which we will still denote by γ , in which $X(\gamma)$ is oriented. We add the further condition in the equivalence relation that the reparametrisations must be orientation-preserving. In this article, unless stated otherwise, we will be working with oriented multi-curves. The spaces of oriented simple and general multi-curves are denoted $\mathcal{S}^+(S)$ and $\mathcal{C}^+(S)$, respectively.

Definition 2.2 (weighted multi-curve). A *weighted multi-curve* $C = \bigcup_i a_i C_i$ is a multi-curve in which each connected component is given a nonnegative real coefficient a_i . If coefficients are not specified, they are 1. We add further moves to the equivalence relation:

- *merging* two parallel components and adding their weights and
- *nullifying*, deleting a component with weight 0.

For instance, $C \cup C$ is equivalent to $2C$. The space of weighted multi-curves up to equivalence is denoted by appending an \mathbb{R} in front of their nonweighted names, so $\mathbb{R}\mathcal{S}(S)$ is the space of weighted simple multi-curves and $\mathbb{R}\mathcal{C}(S)$ is the space of weighted general multi-curves. This is a slight abuse of notation since the weights are required to be nonnegative.

Remark 2.3. Since the weighted curve functionals we are considering are not necessarily positive, they may increase after dropping a component (see Definition 2.2).

2.2. Crossings

Loosely speaking, an essential crossing is a crossing of a multi-curve that cannot be homotoped away. We make this definition precise as follows. We cover cases where γ does not have transverse crossings for convenience of some of the examples.

Definition 2.4 (linked points on a circle). We say that two sets of two points $\{a, b\}$ and $\{c, d\}$ in S^1 are *linked* if the four points are distinct and both connected components of $S^1 - \{a, b\}$ have an element of $\{c, d\}$.

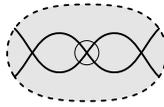
Definition 2.5 (lift of a concrete curve). Given a concrete multi-curve γ on S and a choice $p \in X(\gamma)$, set $x = \gamma(p) \in S$. Pick a lift $\tilde{x} \in \tilde{S}$ of x . The unique lifting property gives a unique lift $\tilde{\gamma}_p : \tilde{X}(\gamma; p) \rightarrow \tilde{S}$ of γ with $\tilde{\gamma}_p(\tilde{p}) = \tilde{x}$, where $X(\gamma; p)$ is the component of $X(\gamma)$ containing p and $\tilde{X}(\gamma; p)$ is its universal cover with basepoint \tilde{p} .

Definition 2.6 (essential crossing). Let γ be a concrete multi-curve on S , and suppose we have points $p, q \in X(\gamma)$ so that $x := \gamma(p) = \gamma(q) \in S$. Pick a lift $\tilde{x} \in \tilde{S}$ of x , and let $\tilde{\gamma}_p$ and $\tilde{\gamma}_q$ be the corresponding lifts of components of γ following Definition 2.5. Then the pair (p, q) form an *essential crossing* if the following two conditions hold:

- (1) both components of $X(\gamma)$ containing p and q are not null-homotopic, so that $\tilde{\gamma}_p$ and $\tilde{\gamma}_q$ are quasi-geodesic components of $\tilde{\gamma}$ and
- (2) either
 - (a) the endpoints $\{a, b\}$ of $\tilde{\gamma}_p$ and the endpoints $\{c, d\}$ of $\tilde{\gamma}_q$ are linked in S^1_∞ or
 - (b) p and q lie on the same component of $X(\gamma)$, $[\gamma] = [\delta^n]$ for some $n > 1$ and some primitive $\delta \in \pi_1(S, x)$, the loop from p to q in $X(\gamma)$ maps to $[\delta^m]$ for some $0 < m < n$ and so the loop from q to p maps to $[\delta^{n-m}]$.

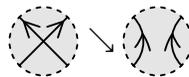
In case (2)b, the endpoints of $\tilde{\gamma}_p$ and $\tilde{\gamma}_q$ are the same.

This definition might be somewhat looser than expected. For instance, in the chain of three crossings



the circled middle crossing is essential iff the other two are, since ‘linking at infinity’ does not see the direction of crossing. This does not matter for our purposes.

Definition 2.7 (smoothings). Let $(p, q) \in X(\gamma)$ be an essential crossing of γ on S . To make a *smoothing* γ' of (p, q) , cut $X(\gamma)$ at p and q and reglue the resulting four endpoints in one of the two other possible ways, getting a new 1-manifold $X(\gamma')$. The map γ' agrees with γ ; this is well-defined since $\gamma(p) = \gamma(q)$. In pictures we will homotop γ' slightly to round out the resulting corners. If γ is oriented, then the *oriented smoothing* is the smoothing that respects the orientation on $X(\gamma)$:



If we obtain a concrete curve γ' from γ by a sequence of k smoothings of essential crossings, we will write $\gamma \searrow_k \gamma'$. (We check whether the crossings are essential at each stage of this process; this is more restrictive than checking at the beginning.)

Lemma 2.8. *Essential crossings are unavoidable in a homotopy class, in the sense that if γ and γ' are homotopic concrete multi-curves and $(p, q) \in X(\gamma)$ is an essential crossing of γ , then there is an essential crossing $(p', q') \in X(\gamma')$ so that the smoothings of (p, q) and (p', q') are homotopic.*

Proof. For both types of essential crossings, there is a representative γ' with minimal crossing number for which the result is clear:

- (a) For crossings of the first type, take the geodesic representative on S , perturbed slightly to make it transverse.
- (b) For crossings of the second type $[\gamma] = [\delta^n]$, take the geodesic representative and perturb it slightly in a standard way in a neighbourhood of δ .

In general, take the given representative γ and perturb it slightly to make it transverse (without introducing new types of crossings). If γ is connected, then by a result of Hass and Scott [30, Theorems 1.8 and 2.1] (see also de Graaf and Schrijver [18]), γ can be turned into any desired minimal form γ' using only Reidemeister I, II and III moves, with the Reidemeister I and II moves being used only in the forward (simplifying) direction. Since we know that γ' has a crossing of the desired type, we can trace the crossings backwards through these moves: a Reidemeister III move does not change the homotopy types of curves achievable by a single smoothing, and we can ignore the additional crossings created by backwards Reidemeister I and II moves.

For multi-curves, the papers above also prove that any diagram can be connected to a minimal diagram by a series of forward Reidemeister moves and that the only obstruction to connecting two minimal diagrams for a multi-curve is swapping the location of two components γ_1, γ_2 that are homotopically powers of the same primitive curve δ [30, pp. 31–32]. But any of these minimal representatives in a neighbourhood of δ contain all essential crossings between γ_1 and γ_2 (necessarily related to an essential self-crossing of δ). □

Remark 2.9. We can also see directly that essential crossings of the first type exist by considering the lift to the universal cover. Lemma 2.8 is false on nonorientable surfaces (consider the double cover of the core curve of a Möbius strip).

We have similar notions for weighted curves.

Definition 2.10. Let $C \in \mathbb{R}C^+(S)$ be a weighted oriented curve, and let γ be a concrete representative of the underlying unweighted curve. Let $(p, q) \in X(\gamma)$ be an essential crossing of γ on S , and let γ' be the smoothing as defined above. If the corresponding components of C have equal weight $w > 0$, then we can make a weighted curve C' by giving every component in $[\gamma']$ not involved in the smoothing the same weight it had in C and giving the one or two new components weight w . In this case we say that C' is obtained from C by a *smoothing of weight w* and write $C \searrow_w C'$.

Using this, we define conditions on a weighted curve functional, extending Definition 1.1.

- **Weighted quasi-smoothing:** There is a constant $R > 0$ so that $C \searrow_w C'$ are weighted curves related by a smoothing of weight w ,

$$f(C) \geq f(C') - wR.$$

- **Weighted smoothing:** Take $R = 0$ in the definition of weighted quasi-smoothing.

See Proposition 3.6 for justification for these definitions.

2.3. Space of geodesics

Definition 2.11 (Boundary at infinity). Endow S with a complete hyperbolic metric g ; we denote the pair (S, g) by Σ . Then we can consider the metric universal covering $p: \tilde{\Sigma} \rightarrow \Sigma$, with $\tilde{\Sigma}$ isometric to the hyperbolic plane. Two quasi-geodesic rays $c, c': [0, \infty) \rightarrow \tilde{\Sigma}$ are said to be *asymptotic* if there exists a constant K for which $d(c(t), c'(t)) \leq K$ for all $t \geq 0$. We define $\partial_\infty \Sigma$, the *boundary at infinity* of S , to be the set of equivalence classes of asymptotic quasi-geodesic rays. This boundary at infinity is independent of the hyperbolic structure on S up to canonical homeomorphism.

Definition 2.12 (Space of oriented geodesics). Let $G^+(\Sigma)$ denote the space of *oriented geodesics* in $\tilde{\Sigma}$; that is,

$$G^+(\Sigma) := \partial_\infty \Sigma \times \partial_\infty \Sigma - \Delta.$$

Since this is independent of the hyperbolic structure, we will also write $G^+(S)$.

2.4. Geodesic currents

Definition 2.13 (Geodesic current definition 1). We define $\mathcal{G}C^+(S)$, the space of oriented geodesic currents on S , to be the space of $\pi_1(S)$ -invariant (positive) Radon measures on $G^+(S)$.

Since the action of $\pi_1(S)$ on $G^+(S)$ is not discrete, this definition is hard to visualise. We give alternate definitions that play a key role in our proofs. For a hyperbolic surface Σ , let $UT\Sigma$ be the unit tangent bundle and let ϕ_t be the geodesic flow on it.

Definition 2.14 (Geodesic current definition 2). We can also define oriented geodesic currents to be the space of (positive) finite Radon measures μ on $UT\Sigma$ which are invariant under the geodesic flow, in the sense that $(\phi_t)_*(\mu) = \mu$ for all $t \in \mathbb{R}$.

We can also look at induced measures on cross sections.

Definition 2.15 (Geodesic current definition 3). A geodesic current is a *transverse invariant measure*: a family of measures $\{\mu_\tau\}_\tau$, where $\tau \subset UT\Sigma$ is a submanifold-with-boundary of the unit tangent bundle of real codimension 1 transverse to the geodesic foliation \mathcal{F} , with the following invariance property: if $x_1 \in \tau_1, x_2 \in \tau_2$ are two points on transversal submanifolds on the same leaf of \mathcal{F} and $\phi: U_1 \rightarrow U_2$ a holonomy diffeomorphism between neighbourhoods of x_1 and x_2 respectively, then $\phi_*\mu_{\tau_1} = \mu_{\tau_2}$.

The equivalence of the three definitions was known to Bonahon [10, Chapter 4]. Details can be found in [1, Section 3.4]. Briefly, given a measure μ on $UT\Sigma$ as in Definition 2.14 and a cross section τ ,

there is an induced flux μ_τ on τ , as explained in Definition 7.15; this gives a geodesic current in the sense of Definition 2.15. Lifting to the universal cover then gives a geodesic current in the sense of Definition 2.13. We can also relate Definitions 2.13 and 2.14 directly by connecting both to measures on $UT\mathbb{H}^2$ that are invariant under both $\pi_1(\Sigma)$ and the geodesic flow, as described by Benoist and Oh [7, Proposition 8.1].

Because Definitions 2.13 and 2.15 are invariant under the mapping class group, we will also write $G^+(S)$ and $\mathcal{GC}^+(S)$ in the sequel. We will also write $\pi_1(S)$. On the other hand, we will emphasise the dependence of $UT\Sigma$ on the hyperbolic structure.

Remark 2.16. For Σ, Σ' hyperbolic surfaces, any homeomorphism $\psi: \Sigma \rightarrow \Sigma'$, there is a homeomorphism $\hat{\psi}: UT\Sigma \rightarrow UT\Sigma'$ that is an orbit equivalence, and it is tempting to use this to define an induced map between geodesic currents in the sense of Definition 2.14. But this does not quite work: $\hat{\psi}_*$ does not take geodesic currents to geodesic currents. Orbit equivalence means that $\psi(\phi_t(x)) = \phi_{f(t)}(\psi(x))$ for some monotonic function $f: \mathbb{R} \rightarrow \mathbb{R}$, but this is not enough to guarantee that $(\phi_t)_*(\hat{\psi}_*\mu) = \hat{\psi}_*\mu$, and, indeed, this is usually false. See Wilkinson [61, Theorem 3.6].

2.5. Oriented vs unoriented currents

We will be mostly working in the setting of oriented geodesic currents, but most of our natural examples (like measured laminations) use unoriented currents.

Definition 2.17 (Unoriented geodesic currents). To define the subspace $\mathcal{GC}(S) \subset \mathcal{GC}^+(S)$ of *unoriented geodesics currents*, let $\sigma: G^+(S) \rightarrow G^+(S)$ be the *flip map* that switches the two factors in the definition of $G^+(S)$, reversing the orientation of the geodesic. This induces a map $\sigma_*: \mathcal{GC}^+(S) \rightarrow \mathcal{GC}^+(S)$. Set

$$\mathcal{GC}(\Sigma) := \{\mu \in \mathcal{GC}^+(\Sigma) \mid \sigma_*(\mu) = \mu\}.$$

There is a map $\Pi: \mathcal{GC}^+(S) \rightarrow \mathcal{GC}^+(S)$ given by $\Pi(\mu) := \frac{1}{2}(\mu + \sigma_*(\mu))$ with image the subset of unoriented currents.

In the proof of the main result, we shall work with oriented currents $\mathcal{GC}^+(S)$; oriented currents are more general and just as easy to work with for our proof.

The maps σ and Π have obvious analogues for curves.

Proof of Corollary 1.11, assuming Theorem A. For a curve functional as in the statement, let $g: \mathbb{RC}^+(S) \rightarrow \mathbb{R}$ be $f \circ \Pi$. Then g satisfies quasi-smoothing, with the same constant as f , and thus by Theorem A extends uniquely to a continuous function $\bar{g}: \mathcal{GC}^+(S) \rightarrow \mathbb{R}$. The desired extension \bar{f} is the restriction of \bar{g} to the subspace of unoriented currents. \square

2.6. Curves as currents

For an oriented multi-curve C on a hyperbolic surface Σ , we can construct a geodesic current as follows.

For Definition 2.13, consider all lifts of all nontrivial components of C to $\tilde{\Sigma}$. Each lift gives a quasi-geodesic in $\tilde{\Sigma}$ and thus a unique fellow-travelling geodesic in $G^+(S)$; we thus get an infinite countable subset of $G^+(S)$, which is easily seen to be discrete and $\pi_1(S)$ -invariant. Define the geodesic current to be the δ -function of this subset.

For Definition 2.14, take the geodesic representative γ of C and consider the canonical lift $\tilde{\gamma}$ of γ to $UT\Sigma$; this is an orbit of ϕ_t . Let μ_C be the length-normalised δ -function on this orbit. That is, for an open set U we set $\mu_C(U)$ to be the total length of $\tilde{\gamma} \cap U$ with respect to the natural Riemannian metric on $UT\Sigma$.

For Definition 2.15 on a cross section τ , again take $\tilde{\gamma} \subset UT\Sigma$, and let μ_τ be the δ -function on the discrete set of points $\tilde{\gamma} \cap \tau$. (This is compatible with the length normalisation in the previous paragraph.)

From any of these points of view the inclusion naturally extends to weighted multi-curves.

Weighted closed (multi-)curves are dense in the space of geodesic currents [10, Proposition 4.4].

Geometric intersection number extends continuously to geodesic currents, as shown by Bonahon [10, Proposition 4.5]. The space of measured laminations (defined by Harer and Penner [49, Section 1.7]) can be characterised, as in Bonahon [8, Proposition 17], as a subset of (unoriented) geodesic currents:

$$\mathcal{ML}(S) := \{\alpha \in \mathcal{GC}(S) \mid \sigma(\alpha) = \alpha, i(\alpha, \alpha) = 0\}.$$

The following square of inclusions is useful to keep in mind:

$$\begin{array}{ccc} \mathcal{RS}(S) & \hookrightarrow & \mathcal{ML}(S) \\ \downarrow & & \downarrow \\ \mathcal{RC}(S) & \hookrightarrow & \mathcal{GC}(S). \end{array}$$

Here the horizontal inclusions have dense image: Douady and Hubbard showed that weighted simple multi-curves are dense in \mathcal{ML} [19, Theorem]. Soon after, Masur showed that weighted simple curves are also dense [39, Theorem 1].

2.7. Topology on currents and measures

Let $\mathcal{M}(X)$ denote the space of positive Borel measures on a topological space X . $\mathcal{M}_1(X)$ will denote the space of Borel probability measures on X . The topology on $\mathcal{M}(X)$ is the *weak* topology*; that is, the smallest topology so that, for all continuous, compactly supported functions f on $G^+(S)$, the functional

$$\mu \mapsto \int_{G^+(S)} f d\mu$$

is continuous.

The topology on $\mathcal{GC}^+(S)$ (in Definition 2.13) is the weak* topology as a subspace of measures on $G^+(S)$, We could also look at the weak* topology on currents as a subspace of measures on $UT\Sigma$ (Definition 2.14); these two points of view give the same topology [7, Proposition 8.1]. On the other hand, if we take τ to be a closed cross section (including the boundary), the map $\mu \mapsto \mu_\tau$ relating Definitions 2.14 and 2.15 is *not* usually continuous with respect to the weak* topologies, so it is delicate to use the weak* topology on $\mathcal{M}(\tau)$; see Lemma 10.4 and Remark 10.6.

There are in fact two topologies on spaces of measures that are sometimes called the weak* topology; the one above is also called the *wide topology* [41, *interalia*]. There is also the *narrow topology* on measures $\mathcal{M}(X)$ on a space X , defined as the smallest topology so that, for all continuous bounded f on X , the functional $\mu \mapsto \int_X f d\mu$ is continuous. (That is, replace *compactly supported* with *bounded* in the functions considered.)

Remark 2.18. Some authors call the weak* topology the vague topology and use the term weak topology for the narrow one (for example, Bauer’s textbook [6]). However, this conflicts with the notion of weak topology used for Banach spaces, and we prefer the wide/narrow usage.

In general, the weak* or wide topology is weaker than the narrow topology, but in some particular cases they are equivalent.

A topological space X is called *Polish* if its topology has a countable base and can be defined by a complete metric.

Theorem 2.19 ([6, Theorem 31.5]). *Let X be a locally compact topological space. Then X is Polish if and only if $\mathcal{M}(X)$ is Polish with respect to the weak*-topology.*

Thus, $\mathcal{GC}^+(S)$ is second countable completely metrisable and second countable and so in particular sequential continuity is the same as continuity. Although we will be dealing with Radon measures, for Polish spaces it is equivalent to consider the a priori more general class of Borel measures.

Theorem 2.20 ([6, Theorem 26.3]). *On a Polish space, a locally finite Borel measure is a σ -finite Radon measure.*

The narrow and wide topology agree in certain sequences on locally compact spaces.

Theorem 2.21 ([6, Theorem 30.8]). *Let X be a locally compact topological space and μ_n a sequence of Radon measures of uniformly bounded mass converging to a Radon measure μ in the wide topology. Then μ_n converges to μ in the narrow topology if and only if $\lim_n \mu_n(X) \rightarrow \mu(X)$.*

Proposition 2.22 ([6, Corollary 30.9]). *Let X be a locally compact topological space and μ, μ_n Borel probability measures. Then $\mu_n \rightarrow \mu$ in the wide topology if and only if $\mu_n \rightarrow \mu$ in the narrow topology.*

In particular, when X is Polish, the two topologies agree for the space $\mathcal{M}_1(X)$ of Borel probability measures.

Proposition 2.23. *If X is a locally compact Polish space, the weak* and narrow topologies agree on $\mathcal{M}_1(X)$.*

Convention 2.24. *For any topological space X , we will always use the weak* topology on $\mathcal{M}(X)$. We will also work with the dense subspace $\mathbb{R}X \subset \mathcal{M}(X)$ of finitely supported measures on X (also called weighted linear combinations of X), with its inherited subspace topology. (The weights are positive, but we usually omit that from the notation.)*

Remark 2.25. If we limit to sums with at most k terms in the linear combination (or points in the support of the measure), we get a further subspace temporarily denoted $\mathbb{R}_{(k)}X \subset \mathbb{R}X \subset \mathcal{M}(X)$. We can view $\mathbb{R}_{(k)}X$ as a quotient of $(\mathbb{R}_{\geq 0} \times X)^k$, quotienting by the action of the symmetric group and other evident equivalences; as such, it inherits an obvious topology, which agrees with the subspace topology.

Remark 2.26. Geodesic currents can also be defined more generally for finite type hyperbolic surfaces. Depending on if we consider ends as cusps or funnels, we get two different spaces, which we will call $\mathcal{G}_{\text{cusp}}(S)$ and $\mathcal{G}_{\text{open}}(S)$, respectively. In the first case, we define the space of geodesic currents analogously to Definition 2.13 for closed surfaces; that is, as invariant measures supported on the space of geodesics of the universal cover, noting that now the space of geodesics contains arcs going from cusp to cusp. In the second case, we consider geodesic currents supported on geodesics projecting to the convex core of the surface. Extending continuously curve functionals to these spaces is more delicate.

In the case of $\mathcal{G}_{\text{cusp}}(S)$, let S be a surface with two open ends and let Σ be a complete hyperbolic metric of finite area, with respect to which the ends of S are cusps. Let a be an arc going from cusp to cusp. Let C_n be the closed curve going along for some time a , winding n times around one cusp, going along a again and winding n times around the other cusp. Observe that although $C_n \rightarrow a$ in the weak* topology and $i(a, a) = 0$, we have $i(a, C_n) = 2n$, so intersection number is not a continuous function on geodesic currents.

The case of $\mathcal{G}_{\text{open}}(S)$ is different, since intersection number is continuous. Indeed, let $\bar{\Sigma}$ denote the convex core of the complete hyperbolic surface of infinite area, which is a compact surface with geodesic boundary. We can consider the intersection number on the double $D(\bar{\Sigma})$ of $\bar{\Sigma}$, which is a closed surface $\bar{\Sigma}$ embeds into. This intersection number on $D(\bar{\Sigma})$ is continuous [10, Proposition 4.5]. Restricting this intersection number to $\bar{\Sigma}$, we obtain continuity of intersection number on $\mathcal{G}_{\text{open}}(S)$. However, the conditions of Theorem A alone are not enough to guarantee a continuous extension $f: \mathcal{G}_{\text{open}}(S) \rightarrow \mathbb{R}$. Indeed, let ℓ be the restriction to $\bar{\Sigma}$ of the hyperbolic length on $D(\bar{\Sigma})$ and consider the modified curve functional ℓ' obtained by setting

$$\ell'(C) := \begin{cases} 0 & C \text{ is a boundary curve} \\ \ell(C) & \text{otherwise.} \end{cases}$$

We note that ℓ' satisfies additivity, stability and homogeneity properties because ℓ does. Also, it satisfies smoothing because ℓ does and nonboundary curves do not intersect boundary curves. However, ℓ' does

not extend to a continuous function on $\mathcal{GC}_{\text{open}}(S)$: let $\gamma, \beta \in \pi_1(S, p)$ be elements based at a point $p \in S$ and denote $C = [\beta]$, $D = [\gamma]$. Assume that C is a boundary curve and D is not. For each n , define a nonsimple, nonboundary parallel curve by $C_n := [\gamma\beta^n]$. Observe that the sequence $\frac{1}{n}C_n$ converges to C in the weak* topology but

$$\ell'(C_n) \asymp n,$$

whereas

$$\ell'(C) = 0,$$

so ℓ' cannot be a continuous function on $\mathcal{GC}_{\text{open}}(S)$. So additional conditions on f are needed to guarantee a continuous extension to $\mathcal{GC}_{\text{open}}(S)$.

3. Convexity and continuity

3.1. Convexity on the reals

The curve functionals f we study have some convexity property as a function of the weights, because of the convex union and homogeneity properties. We first review some background on convex functions and their continuity properties.

A function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is called \mathbb{R} -convex (respectively \mathbb{Q} -convex) if

$$f(ax + (1 - a)y) \leq af(x) + (1 - a)f(y)$$

for all $x, y \in \mathbb{R}^n$ and $a \in [0, 1]$ (respectively text $a \in [0, 1] \cap \mathbb{Q}$). A function $f: \mathbb{Q}^n \rightarrow \mathbb{R}$ might also be \mathbb{Q} -convex, with the same definition. We furthermore say that f is *midpoint-convex* if

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y).$$

Proposition 3.1. *The following are true:*

- (i) *A midpoint-convex function $f: \mathbb{Q}^n \rightarrow \mathbb{R}$ is \mathbb{Q} -convex.*
- (ii) *An \mathbb{R} -convex function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous.*
- (iii) *A \mathbb{Q} -convex function $f: \mathbb{Q}^n \rightarrow \mathbb{R}$ is continuous.*
- (iv) *Every \mathbb{Q} -convex function $f: \mathbb{Q}^n \rightarrow \mathbb{R}$ has a unique continuous extension to an \mathbb{R} -convex function $\tilde{f}: \mathbb{R}^n \rightarrow \mathbb{R}$.*

Proof.

- (i) This proof is due to Ivan Meir [40], following Hardy, Littlewood, and Pólya [29, P. 17]. We first prove that midpoint inequality extends to arbitrary means:

$$g\left(\frac{x_1 + \dots + x_m}{m}\right) \leq \frac{g(x_1) + \dots + g(x_m)}{m}$$

for any $m \in \mathbb{Z}_{\geq 1}$. We can prove this first for $m = 2^k$ by using midpoint convexity repeatedly. For general $m \leq 2^i$, we take x_1, \dots, x_m plus $2^i - m$ copies of $x' = (x_1 + \dots + x_m)/m$, yielding

$$g(x') = g\left(\frac{(2^i - m)x' + x_1 + \dots + x_m}{2^i}\right) \leq \frac{(2^i - m)g(x') + g(x_1) + \dots + g(x_m)}{2^i},$$

which implies $g(x') = g((x_1 + \dots + x_m)/m) \leq (g(x_1) + \dots + g(x_m))/m$.

To prove \mathbb{Q} -convexity, taking a copies of x and b copies of y we obtain

$$g\left(\frac{ax + by}{a + b}\right) \leq \frac{ag(x) + bg(y)}{a + b} = \left(\frac{a}{a + b}\right)g(x) + \left(\frac{b}{a + b}\right)g(y)$$

for $a, b \in \mathbb{Z}_{\geq 0}$ not both zero.

- (ii) See Kuczma [37, Theorem 7.1.1].
- (iii) The proof of [37, Theorem 7.1.1] can be adapted for functions on \mathbb{Q}^n . The proof relies on Bernstein–Doetsch theorem, which works in high generality for topological vector spaces (see Kominek and Kuczma [36, Theorem B]) and the fact that any point $x \in \mathbb{Q}^n$ is the interior of some full-dimensional \mathbb{Q} -simplex on which f is bounded.
- (iv) Define the extension by

$$\bar{f}(x) := \liminf_{\substack{y \rightarrow x \\ y \in \mathbb{Q}^n}} f(y).$$

By continuity of f on \mathbb{Q}^n , \bar{f} is an extension of f . To study $\bar{f}(ax + (1 - a)y)$, let x_i, y_i be sequences in \mathbb{Q}^n with $\lim x_i = x, \lim y_i = y, \liminf f(x_i) = f(x)$ and $\liminf f(y_i) = f(y)$. Let $a_i \in [0, 1] \cap \mathbb{Q}$ be a sequence with $\lim a_i = a$. Then

$$\begin{aligned} \bar{f}(ax + (1 - a)y) &\leq \liminf f(a_i x_i + (1 - a_i)y_i) \\ &\leq \liminf (a_i f(x_i) + (1 - a_i)f(y_i)) \\ &= a\bar{f}(x) + (1 - a)\bar{f}(y). \end{aligned}$$

Thus, \bar{f} is convex and therefore continuous here

□

It is not true that all \mathbb{Q} -convex functions $f: \mathbb{R}^n \rightarrow \mathbb{R}$ must be continuous, but all counterexamples are highly pathological. In particular, any measurable \mathbb{Q} -convex function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is necessarily continuous (see [37, Theorem 9.4.2]).

Remark 3.2. Proposition 3.1(ii) does not hold for infinite-dimensional topological vector spaces: an unbounded linear functional is convex but not continuous.

3.2. Convexity for curve functionals

We now apply the results above to our setting of real-valued functions on curves.

As an immediate consequence of Proposition 3.1 (iv), for a curve functional f satisfying convex union and homogeneity, we can extend f to a weighted curve functional that is convex and therefore continuous for a fixed set of components. This will play a role in the proof of Theorem A, specifically in Proposition 10.3.

First, for any curve functional satisfying homogeneity, we adopt the convention that we extend f to rationally weighted curves $\mathbb{Q}\mathcal{C}^+(S)$ in the usual way by clearing denominators: set

$$f\left(\sum a_i C_i\right) := \frac{1}{d} f\left(\sum da_i C_i\right) \tag{3.3}$$

for some integer d sufficiently large so all the da_i are integers. By homogeneity of f , the extension does not depend on d .

Proposition 3.4. Let $C = (C_i)_{i=1, \dots, n}$ be a finite sequence of multi-curves and consider combinations $\sum_{i=1}^n a_i C_i$. Let f be a curve functional that satisfies homogeneity and convex union. Define a function $f_C: \mathbb{Q}^n \rightarrow \mathbb{R}$ by

$$f_C(a_1, \dots, a_n) := f\left(\sum_{i=1}^n a_i C_i\right).$$

Then f_C is \mathbb{Q} -convex and thus continuous.

Proof. It is immediate from the definitions that f_C is midpoint convex. The result follows from Proposition 3.1. \square

Corollary 3.5. *If a curve functional f satisfies homogeneity and convex union, then there is a unique continuous homogeneous extension of f to weighted curve functional.*

Proposition 3.6. *If a curve functional satisfies convex union, homogeneity and quasi-smoothing, the extension from Corollary 3.5 satisfies weighted quasi-smoothing with the same constant.*

Proof. We first observe that f , as a function on integrally weighted multi-curves, satisfies weighted quasi-smoothing. If $C = [\gamma]$, $\gamma \searrow \gamma'$ and k is an integer, then $k\gamma \searrow_k k\gamma'$, since $k\gamma$ are k disjoint parallel copies of γ . Thus,

$$f(kC) \geq f(kC') - kR.$$

By the method of clearing denominators and homogeneity, as in equation (3.3), we obtain rational weighted quasi-smoothing.

Finally, by continuity of f as a function of the weights of a fixed multi-curve, we get real weighted quasi-smoothing. \square

Theorems A and B as stated start from a curve functional of various types. Many curve functionals naturally come as functions on weighted curves (see Section 4). On the other hand, we have seen in Proposition 3.6 that a curve functional satisfying convex union and homogeneity and stability properties yields weighted curve functional satisfying the same properties.

4. Examples

We give several examples of curve functionals that extend to functions on currents, mostly as a consequence of our main theorems. This includes known results, such as hyperbolic lengths and intersection numbers, or, more generally, lengths for any length metric structure, as well as new results, such as extremal lengths with respect to a conformal structure or with respect to a graph. In the following applications we consider unoriented curves unless otherwise stated.

4.1. Intersection number

Fix a multi-curve D and consider the curve functional $\mathcal{C}(S)$ defined by $f(C) = i(C, D)$, where $i(C, D)$ is the minimal number of intersection points between representatives of C and D in general position. Then f is homogeneous, additive and stable.

There is a simple geometric argument, which we will use repeatedly, to see that f satisfies smoothing. Fix a minimal representative δ for D . Take a curve C with an essential self-intersection x and a representative γ with minimal intersection with δ . Then γ has a self-intersection point x' of the homotopy type of x . If we consider the curve representative $\gamma' \in C'$ obtained by smoothing at x' , then, since $i(C', D)$ is an infimum, we have

$$i(C', D) \leq i(\gamma', \delta) = i(\gamma, \delta) = i(C, D),$$

as desired.

By Theorem A, intersection number with D extends to a continuous function on geodesic currents

$$i(\cdot, D): \mathcal{GC}(S) \rightarrow \mathbb{R}.$$

We can then fix C and vary D to show that, for μ a geodesic current, $i(\cdot, \mu)$ is a continuous function on $\mathcal{GC}(S)$. In [10, Proposition 4.5], Bonahon shows that the geometric intersection number $i: \mathbb{RC}(S) \times$

$\mathbb{R}\mathcal{C}(S) \rightarrow \mathbb{R}_{\geq 0}$ between two weighted multi-curves extends to a continuous two-variable function

$$i: \mathcal{GC}(S) \times \mathcal{GC}(S) \rightarrow \mathbb{R}_{\geq 0}.$$

Question 4.1. Can the arguments in this article be extended to give an alternate proof that geometric intersection number is a continuous two-variable function?

Following Example 1.12, proving that there is a continuous extension of $\sqrt{i(C, C)}$ to a function on $\mathcal{GC}(S)$ is equivalent to proving continuity of i as a two-variable function, by a simple polarisation argument:

$$i(C, D) = \frac{i(C \cup D, C \cup D) - i(C, C) - i(D, D)}{2}.$$

4.2. Hyperbolic length

We continue with the original motivating example for geodesic currents in Bonahon’s paper [8, Proposition 14]. Fix a hyperbolic metric g on S and denote the hyperbolic structure by Σ . Then, for any closed curve C on S (not necessarily simple), we can consider its hyperbolic length with respect to the Riemannian metric. In terms of the holonomy representation $\rho_g : \pi_1(S) \rightarrow PSL_2(\mathbb{R})$, this is given by

$$\ell_g(C) = 2 \cosh^{-1} \left(\frac{1}{2} \operatorname{tr}(\rho(c)) \right)$$

where $c \in \pi_1(S)$ is a representative of C . We extend ℓ_g to a weighted curve functional by additivity and homogeneity:

$$\ell_g(t_1 C_1 \cup \dots \cup t_n C_n) = \sum_{i=1}^n t_i \ell_g(C_i).$$

By definition, ℓ_g is additive and homogeneous. Stability follows from properties of the trace of 2×2 matrices or geometrically from the length. Smoothing follows by the argument for intersection number. Thus, by Theorem A, ℓ_g extends to a continuous function on geodesic currents.

We recall that Bonahon shows that

$$\ell_g(C) = i(\mathcal{L}_\Sigma, C)$$

where $\ell_g(C)$ denotes the hyperbolic length of C – that is, the length of the g -geodesic representative – and \mathcal{L}_Σ denotes the *Liouville current*, a geodesic current induced by the volume form on $UT\Sigma$. (For an equivalent formulation in terms of Definition 2.13, see Bonahon [8, Section 2].)

4.3. Length with respect to arbitrary metrics

The argument from Subsection 4.2 applies equally well to show that for any Riemannian or, more generally, length metric g on S , length ℓ_g with respect to g satisfies smoothing. For completeness and later use, we prove that these curve functionals are stable. (Freedman–Hass–Scott give a proof in the Riemannian case [27, Lemma 1.3].)

Lemma 4.2. *For any orientable surface S and length metric g on S , the curve functional ℓ_g is stable: $\ell_g(C^n) = n\ell_g(C)$.*

Proof. One inequality is true in any length space: by taking the obvious n -fold representative of C^n , we see that $\ell_g(C^n) \leq n\ell_g(C)$. The other inequality follows from the smoothing property: since $C^n \searrow nC$, we have $\ell_g(C^n) \geq n\ell_g(C)$. □

Remark 4.3. Lemma 4.2 is false if S is not orientable. For instance, if S is the projective plane, g is any metric and C is the nontrivial curve on Σ , then $\ell_g(C) > 0$ but C^2 is null-homotopic so $\ell_g(C^2) = 0$. We can get a similar inequality without torsion on a Möbius strip by removing a small disk from this projective plane.

4.4. Length with respect to embedded graphs

We can generalise further beyond length metrics. Let $\iota: \Gamma \hookrightarrow S$ be an embedding of a finite graph in S that is filling, in the sense that the complementary regions are disks or, equivalently, ι_* is surjective on π_1 . Endow Γ with a length metric g . Then any closed multi-curve C on S can be homotoped so that it factors through Γ , in many different ways. Let $\ell_\Gamma(C)$ be the length of the smallest multi-curve D on Γ so that $\iota(D)$ is homotopic to C . It is easy to see that this length is realised and is positive. (In fact, we can see ℓ_Γ as a limit of lengths with respect to Riemannian metrics, by fixing an embedding of Γ and making the metric on the complement of a regular neighbourhood of Γ be very large, following Shepard [53].)

As before, ℓ_Γ is clearly additive and homogeneous and is stable by the argument of Lemma 4.2. To see that ℓ_Γ satisfies smoothing at an essential crossing, take a minimal-length concrete representative δ of D on Γ . Since the image $\iota \circ \delta$ has a corresponding crossing by Lemma 2.8 and ι is an embedding, there is a corresponding crossing of δ that can be smoothed and then tightened to get the desired inequality.

As a special case, we can consider the case when Γ is a *rose graph* with only one vertex $*$ and edges of length 1. Since ι is filling, the image of the edges of Γ give generators for $\pi_1(S, \iota(*))$. Then the length $\ell_\Gamma(C)$ of a curve C is the length of C as a conjugacy class in $\pi_1(S)$ with respect to these generators. This is a simple generating set in the sense of Erlandsson [22], who proved this continuity and constructed an explicit multi-curve K so that $\ell_\Gamma(C) = i(C, K)$.

4.5. Stable lengths

Generalising the previous example, let $\iota: \Gamma \rightarrow S$ be an *immersion* from a finite graph to S so that $\iota_*: \pi_1(\Gamma) \rightarrow \pi_1(S)$ is surjective, and again give a length metric on Γ . For instance, if Γ has a single vertex and all edges have length 1, this is equivalent to giving an arbitrary generating set for $\pi_1(S)$. We can define $\ell_\Gamma(C)$ as before, as the minimum length of any multi-curve D on Γ so that $\iota_*(D) = C$.

The curve functional ℓ_Γ is still additive, but unlike the previous examples it is not stable (see Example 4.10). Thus, we cannot hope to extend ℓ_Γ to currents but rather extend the stable curve functional $\|\ell_\Gamma\|$ (defined in Section 13). We do have quasi-smoothing.

Lemma 4.4. *For any connected, π_1 -surjective immersion of a length graph $\iota: \Gamma \rightarrow S$, the curve functional ℓ_Γ satisfies quasi-smoothing.*

This is a special case of a more general result. Let V be a connected length space, with a continuous, π_1 -surjective map $\iota: V \rightarrow S$. Then for C a multi-curve on S , a *lift* of C is a multi-curve \tilde{C} in V so that $\iota_*\tilde{C} = C$. (Both \tilde{C} and C are defined up to homotopy.) Define $\ell_{\iota,V}(C)$ to be the infimum, over all lifts \tilde{C} of C , of the length of \tilde{C} in V .

Proposition 4.5. *Let V be a connected, compact length space, with a continuous, π_1 -surjective map $\iota: V \rightarrow S$. Then $\ell_{\iota,V}: \mathcal{C}(S) \rightarrow \mathbb{R}_+$ satisfies quasi-smoothing.*

As a corollary of Proposition 4.5 and Theorems B and A, $\|\ell_V\|$ extends continuously to a function on $\mathcal{GC}(S)$. This continuous extension was first proved by Bonahon [9, Proposition 10] in the context of a hyperbolic group acting discretely and cocompactly on a length space (replacing $\pi_1(S)$ acting on \tilde{V}), with an additional technical assumption that the space is uniquely geodesic at infinity. Later, Erlandsson, Parlier, and Souto [23, Theorem 1.5] lifted this assumption.

We remark that given a properly discontinuous action of $\pi_1(S)$ on X a CW-complex, we can construct a $\pi_1(S)$ -surjective map $\iota: X/\pi_1(S) \rightarrow S$. In fact, we will construct a $\pi_1(S)$ -equivariant map $\iota: X \rightarrow \tilde{S}$. First, we define $\iota: X_0 \rightarrow \tilde{S}$ by picking a value on each $\pi_1(S)$ orbit of the 0-skeleton arbitrarily and

extending equivariantly. Similarly, on the 1-skeleton X_1 , for each $\pi_1(S)$ orbit on X_1 , pick a path in \tilde{S} between the images of the endpoints. Continue the construction inductively. This construction works because $\pi_1(S)$ acts freely – since $\pi_1(S)$ is torsion-free and acts properly discontinuously – and \tilde{S} is contractible.

In this construction, proper discontinuity of the action is crucial. For example, it was shown by Bonahon in [9, Proposition 11] that if W is a finite graph which is a deformation retract of S , the action of $\pi_1(S)$ on the universal cover \tilde{W} of W is cocompact but not properly discontinuous and translation length of conjugacy classes of $\pi_1(S)$ acting on X does not extend continuously to $\mathcal{GC}(S)$.

Proof of Proposition 4.5. Let $\tilde{\iota}: \tilde{V} \rightarrow \tilde{S}$ be the pull-back of ι along the universal cover π_S of S , part of the pullback square

$$\begin{array}{ccc}
 \tilde{V} & \overset{\pi_V}{\dashrightarrow} & V \\
 \tilde{\iota} \downarrow & & \downarrow \iota \\
 \tilde{S} & \xrightarrow{\pi_S} & S
 \end{array} \tag{4.6}$$

where $\tilde{V} = V \times_S \tilde{S}$. Since ι is π_1 -surjective, \tilde{V} is a connected covering space of V . For $\tilde{x} \in \tilde{S}$, let $\text{diam}_\iota(\tilde{x})$ be the diameter of $\tilde{\iota}^{-1}(\tilde{x}) \subset \tilde{V}$. (Set $\text{diam}_\iota(\tilde{x}) = 0$ if \tilde{x} is not in the image of $\tilde{\iota}$.) Set

$$\text{diam}_\iota(V) := \sup_{\tilde{x} \in \tilde{S}} \text{diam}_\iota(\tilde{x}).$$

We wish to see that $\text{diam}_\iota(V)$ is finite. First, since ι is π_1 -surjective, for every $\gamma \in \pi_1(S, x)$ there exists $\delta_\gamma \in \pi_1(V, \tilde{x})$ so that $\iota_* \delta_\gamma = \gamma$. This implies

$$\delta_\gamma \cdot \tilde{\iota}^{-1}(\tilde{x}) = \tilde{\iota}^{-1}(\gamma \cdot \tilde{x}).$$

But δ_γ is a deck transformation and thus acts as an isometry on \tilde{V} , so $\text{diam}(\tilde{\iota}^{-1}(\gamma \cdot \tilde{x})) = \text{diam}(\tilde{\iota}^{-1}(\tilde{x}))$ and so we can define

$$\text{diam}_\iota(x) = \text{diam}_\iota(\tilde{x}) \tag{4.7}$$

for $x \in S$ and any lift \tilde{x} of x . (Note that $\text{diam}_\iota(x)$ is not in general the diameter of $\iota^{-1}(x)$; rather than looking at the length of a shortest path connecting two points in $\iota^{-1}(x)$, we restrict to paths that map to null-homotopic loops.)

Lemma 4.8. *In the above setting, the diameter $\text{diam}_\iota(x)$ is upper semi-continuous as a function of x .*

Proof. For each $x_0 \in S$, consider an evenly covered neighbourhood U of x_0 and fix a lift $\tilde{x}_0 \in \tilde{S}$. We want to show that for all sequences $\{x_i\} \subset U$ with $x_i \rightarrow x$ and for all $\varepsilon > 0$, there exists i_0 so that for all $i \geq i_0$

$$\text{diam}(\tilde{\iota}^{-1}(\tilde{x}_i)) < \text{diam}(\tilde{\iota}^{-1}(\tilde{x}_0)) + \varepsilon.$$

Now, $\tilde{\iota}$ is a pullback of a proper map, so it is a closed map.

Lemma 4.9 ([54, Theorem 005R]). *Let X be a metric space and $f: X \rightarrow Y$ a proper map. For any continuous map $g: Z \rightarrow Y$, the pullback map $X \times_Y Z \rightarrow Z$ is closed.*

By definition of diameter, and since the fibres $\tilde{\iota}^{-1}(\tilde{x})$ are compact for any \tilde{x}_i , we can find points $p_i, q_i \in \tilde{\iota}^{-1}(\tilde{x}_i)$ so that $\text{diam}(\tilde{\iota}^{-1}(\tilde{x}_i)) = d(p_i, q_i)$. Also, by closedness of $\tilde{\iota}$, a subsequence of the p_i and

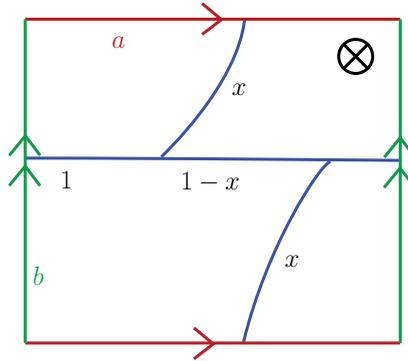


Figure 4.1. A punctured torus with, in blue, a train track carrying a slice of measured laminations on the punctured torus depending on a parameter $x \in [0, 1]$. In red, the parallel loop a . In green, the meridian loop b .

q_i converges to points $p, q \in \tilde{\tau}^{-1}(\tilde{x})$. Furthermore, by continuity of distance, for any $\varepsilon > 0$, we have, for i large enough,

$$\text{diam}(\tilde{\tau}^{-1}(\tilde{x}_i)) = d(p_i, q_i) < d(p, q) + \varepsilon \leq \text{diam}(\tilde{\tau}^{-1}(\tilde{x})) + \varepsilon,$$

finishing the proof of Lemma 4.8. □

As a result of Lemma 4.8, the function $\text{diam}_\iota(x)$ is bounded on S ; let $R(V)$ be this global bound.

We now finish the proof of Proposition 4.5. We are given a curve C with an essential crossing p and corresponding smoothing C' . Pick a concrete curve $\tilde{\gamma}$ on V that comes within ε of realising $\ell_{\iota, V}(C)$; in particular, $\iota \circ \tilde{\gamma}$ represents C . By Lemma 2.8, there are points $x, y \in X(\tilde{\gamma})$ so that $\iota(\tilde{\gamma}(x)) = \iota(\tilde{\gamma}(y))$ is a crossing corresponding to p . We wish to find another curve $\tilde{\gamma}'$ on V , with length not too much longer, so that $\iota \circ \tilde{\gamma}'$ represents C' . We can do this by cutting $\tilde{\gamma}$ at x and y , yielding endpoints x_1, x_2 and y_1, y_2 and reconnecting x_1 to y_2 and y_1 to x_2 by paths in V that project to the identity in $\pi_1(S)$.

But the maximal length of a path connecting any two points $x, y \in V$ with $\iota(x) = \iota(y)$ that projects to a null-homotopic path is exactly $\text{diam}_\iota(\iota(x))$. We can therefore construct a desired representative $\tilde{\gamma}'$ with

$$\ell_V(\tilde{\gamma}') \leq \ell_V(\tilde{\gamma}) + 2R(V) \leq \ell_{\iota, V}(C) + \varepsilon + 2R(V).$$

Since ε was arbitrary, we have proved the result with quasi-smoothing constant $2R(V)$. □

We show now an example of a curve functional that satisfies quasi-smoothing but not strict smoothing.

Example 4.10. Consider the torus with one puncture with fundamental group generated by the usual horizontal loop a and vertical loop b , as in Figure 4.1. (This does not, strictly speaking, fit in the context of closed surfaces considered in this article, but we can embed this punctured torus in a larger surface without essential change.) Its fundamental group is the free group $F_2 = \langle a, b \rangle$. We will consider word length f with respect to the generating set (a, a^2, b) . Word length satisfies quasi-smoothing and additive union but not stability. (For instance, $f(a^2) = 1 \neq f(2a) = 2$.) The stable word length $\|f\|$ satisfies stability and still satisfies quasi-smoothing, but it does not satisfy strict smoothing. We will show it behaves more erratically than word length with respect to embedded generating sets.

Consider, for example, the collection of weighted curves $C(x)$ carried by the train track in Figure 4.1, with weights depending on a rational parameter $x \in [0, 1]$. For instance, for $x = 2/5$, the curve is $1/5[aabab]$, with stable length $(1/5) \cdot 4$. If we plot the stable word length of $C(x)$ multiplied by the weight, we obtain the saw-tooth graph in Figure 4.2. We note the erratic behavior as a function of x . In particular, it is far from convex. If $\|f\|$ satisfied the smoothing property, then it would be a convex function of the train track weights, since if w_1 and w_2 are two rational weights on a train track T , the

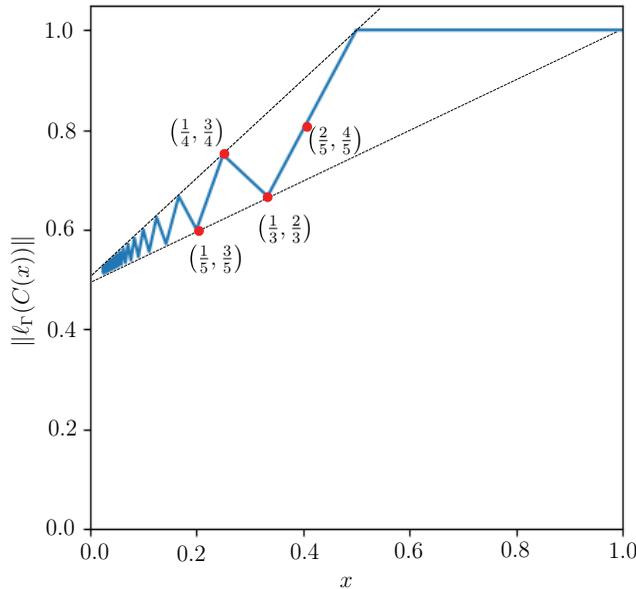


Figure 4.2. We consider the curve $C(x)$ carried by a train track depending on a parameter x and plot the stable word length of $C(x)$ as a function of x . The graph has vertices at $(\frac{1}{2n+1}, \frac{n+1}{2n+1})$ and $(\frac{1}{2n}, \frac{n+1}{2n})$.

weighted multi-curve $T(w_1) \cup T(w_2)$ can be smoothed to $T(w_1 + w_2)$ (observed by Mirzakhani in [42, Appendix A] and the second author in [55, Subsection 3.2]).

4.6. Asymmetric lengths

The arguments in Subsection 4.5 apply equally well to cases where distances may be zero or not symmetric. For instance, we can take a directed graph Γ with a nonnegative length on each edge, together with a map $\iota: \Gamma \rightarrow S$ so that the corresponding cover $\tilde{\Gamma}$ is strongly connected (every vertex can be reached from any other vertex). The same arguments apply to show that $l_\Gamma(\vec{C})$ satisfies the oriented quasi-smoothing property and so its stabilisation $\|l_\Gamma\|$ extends to a continuous function $\mathcal{GC}^+(S) \rightarrow \mathbb{R}_{\geq 0}$.

One example would be to take a generating set for $\pi_1(S)$ as a monoid. This corresponds to taking Γ to be a graph with a single vertex and one edge for each monoid generator.

4.7. Generalised translation lengths from higher representations

Let G be a real, connected, noncompact, semi-simple, linear Lie group. Let K denote a maximal compact subgroup of G , so that $X = G/K$ is the Riemannian symmetric space of G . Let $[P]$ be the conjugacy class of a parabolic subgroup $P \subset G$. Then there is a natural notion of $[P]$ -Anosov representation $\rho: \pi_1(S) \rightarrow G$; see, for example, Kassel’s notes [34, Section 4]. When $\text{rank}_{\mathbb{R}}(G) = 1$ there is essentially one class $[P]$, so we can simply refer to them as Anosov representations, and they can be defined as those injective representations $\rho: \pi_1(S) \rightarrow G$ where $\Gamma := \rho(\pi_1(S))$ preserves and acts cocompactly on some nonempty convex subset V of X .

Rank 1 Anosov representations include two familiar examples:

- (1) Fuchsian representations into $G = PSL(2, \mathbb{R})$. Here, $K = SO(2)$ and $X = \mathbb{H}^2$. The convex set V in this case is the lift of the convex core of the hyperbolic surface.
- (2) Quasi-Fuchsian representations of surface groups into $G = PSL(2, \mathbb{C})$. In this case, $K = SU(2)$, $X = \mathbb{H}^3$, and V is the lift of the convex core of the hyperbolic quasi-Fuchsian manifold.

In general, the conjugacy classes of parabolic subgroups of G correspond to subsets θ of the set of restricted simple roots Δ of G . For a given $[P]$ -Anosov representation and each $\alpha \in \theta$, Martone and Zhang define [38, Definition 2.21] a curve functional

$$l_\alpha^\rho: \mathcal{C}(S) \rightarrow \mathbb{R}_{\geq 0}$$

and show that for a certain subset of Anosov representations, these can be extended to geodesic currents as intersection numbers with some fixed geodesic current.

For the two rank 1 examples above, this length $l_\alpha^\rho(C)$ corresponds to hyperbolic length of the closed geodesic in the homotopy class C in the quotient hyperbolic manifold $\mathbb{H}^2/\rho(\pi_1(S))$ or $\mathbb{H}^3/\rho(\pi_1(S))$.

Bonahon showed the length in the Fuchsian case extends to geodesic currents [8, Proposition 14]. Our techniques give an extension to geodesic currents of the quasi-fuchsian length; that is, the hyperbolic length of the geodesic representatives in the quasi-Fuchsian manifold.

Proposition 4.11. *Translation length l_ρ , for $\rho: \pi_1(S) \rightarrow PSL(2, \mathbb{C})$ a quasi-Fuchsian representation, extends to geodesic currents.*

Proof. Let V to be the convex core of $\mathbb{H}^3/\rho(\pi_1(S))$. We have an obvious retract $r: V \rightarrow S$ (defined up to homotopy). Now use Proposition 4.5, taking $\iota = r$. \square

Question 4.12. Bridgeman and Taylor show that a *complex* length function of a quasi-Fuchsian representation coming from a Patterson–Sullivan measure also extends continuously to a function on geodesic currents, as [12, Section 6]. Is there a version of our main theorem that would prove that a complex-valued curve functional like this extends to currents?

Remark 4.13. Extending the definition of convex cocompact representation for higher rank groups turns out to yield products of representations of rank 1 (see Kleiner–Leeb [35, Theorem 1.3] and Quint [50, Théorème]) so it is not clear that the approach using Proposition 4.5 will allow one to extend curve functionals in higher rank to geodesic currents.

For $G = PSL(3, \mathbb{R})$, there is another cocompact action, not on a convex subset of the symmetric space but on a convex subset of $G/P = \mathbb{R}P^2$. In this case, there is a natural metric on this convex subset, the Hilbert metric. One can easily show using this metric that smoothing is satisfied. In general, in higher rank one can construct similar cocompact actions on convex domains of G/P (see Guichard–Wienhard [28]), but there is not a known canonical choice of metric. Martone and Zhang show [38, Theorem 2.1] that some types of representations known as *positively ratioed* can be realised as intersection numbers with a distinguished geodesic current. This immediately implies this subclass of representations satisfy the smoothing property. It would be interesting to subsume their extension result under our scope. More specifically, we have the following.

Question 4.14. Can we prove quasi-smoothing for the translation length for a subclass of $[P]$ -Anosov representations, as in Martone–Zhang [38, Definition 2.25], directly from the definition of translation length?

4.8. Extremal length

We now turn to curve functionals that satisfy only convex union and not additive union, starting with the original motivation for this work, extremal length.

Definition 4.15. Fix Σ a Riemann surface with a metric g . Let $C = \bigcup t_i C_i$ be a weighted multi-curve on Σ . For $\rho: \Sigma \rightarrow \mathbb{R}_{\geq 0}$ a measurable rescaling function, the *area* of ρ is

$$\text{Area}(\rho g) := \int_{x \in \Sigma} \rho(x)^2 \mu_g(x),$$

where μ_g is the Lebesgue measure of g . The *length* of C is

$$\ell_{\rho g}(C) := \inf_{\gamma \in C} \sum_i t_i \int_{x \in \gamma_i} \rho(x) dx,$$

where dx is measured with respect to g arc-length and the infimum runs overall all representatives $\gamma = \bigcup_i \gamma_i$ of C , where γ_i is a representative of the component C_i of C . When ρ is continuous, $\ell_{\rho}(C)$ is the length with respect to the metric g rescaled by ρ . The square root of the *extremal length* of C is

$$\sqrt{\text{EL}}(C) := \sup_{\rho} \frac{\ell_{\rho g}(C)}{\sqrt{\text{Area}(\rho g)}}.$$

Observe that the supremum is unchanged under multiplying ρ by a positive constant. It is a standard result that the supremum is realised by some generalised metric (not necessarily Riemannian) [52, Theorem 12] and that, when C is a simple multi-curve, the optimum metric ρg is the cone Euclidean metric associated to a quadratic differential, as shown by Jenkins [33]. Very little is known about the optimum metric when C is not simple, except in special cases [62, 14, 31, 46].

Lemma 4.16. *As a function of C with fixed Σ , $\sqrt{\text{EL}}$ satisfies homogeneity, stability and smoothing.*

Proof. This follows since $\ell_{\rho g}$ satisfies these properties for each ρ ; here is the argument for smoothing.

Let C be a multi-curve with an essential crossing and let C' be the curve obtained by smoothing at the crossing. Then, for any scaling function ρ ,

$$\frac{\ell_{\rho g}(C')}{\sqrt{\text{Area}(\rho g)}} \leq \frac{\ell_{\rho g}(C)}{\sqrt{\text{Area}(\rho g)}}.$$

Since $\sqrt{\text{EL}}(C')$ and $\sqrt{\text{EL}}(C)$ are the suprema of such terms, the result follows. □

Convex union is different, since on one side of the inequality we have a sum of values of $\sqrt{\text{EL}}$. (Extremal length does not satisfy additivity.)

Lemma 4.17. *$\sqrt{\text{EL}}$ satisfies convex union.*

Proof. Fix a curve split as a union $C = C_1 \cup C_2$, and let $\rho: \Sigma \rightarrow \mathbb{R}$ be the function realising the supremum in the definition of extremal length for C . Then

$$\begin{aligned} \sqrt{\text{EL}}(C_1 \cup C_2) &= \frac{\ell_{\rho g}(C_1)}{\sqrt{\text{Area}(\rho g)}} + \frac{\ell_{\rho g}(C_2)}{\sqrt{\text{Area}(\rho g)}} \\ &\leq \sqrt{\text{EL}}(C_1) + \sqrt{\text{EL}}(C_2), \end{aligned}$$

where the last inequality holds by the supremum in the definition of EL. □

Thus, by Theorem A, $\sqrt{\text{EL}}$ (and EL) extend uniquely to continuous functions on geodesic currents. With this extension, we propose the following conjecture.

Conjecture 4.18. *For some universal constant C ,*

$$\text{EL}_{\Sigma}(\mathcal{L}_{\Sigma}) = C \text{Area}(\Sigma),$$

where \mathcal{L}_{Σ} is the Liouville current (compare Subsection 4.2).

Remark 4.19. For the extremal length *without* the square root, we instead have inequalities

$$\text{EL}(C_1) + \text{EL}(C_2) \leq \text{EL}(C_1 \cup C_2) \leq 2(\text{EL}(C_1) + \text{EL}(C_2)).$$

The second inequality is a simple consequence of Lemma 4.17. To see the first inequality, take optimal rescaling functions ρ_i for $\text{EL}(C_i)$, normalised so that $\ell_{\rho_i g}(C_i) = \text{Area}(\rho_i g) = \text{EL}(C_i)$. Then using $\rho_1 + \rho_2$ as the test function for $\text{EL}(C_1 \cup C_2)$ gives the desired inequality after elementary manipulations.

4.9. Extremal length with respect to elastic graphs

There is a parallel notion of extremal length with respect to elastic graphs, as introduced by the second author [56], just as there is for ordinary lengths (Subsection 4.4).

An *elastic graph* (Γ, α) is a 1-dimensional CW complex Γ (i.e., allowing multiple edges and loops) together with an assignment of positive real numbers $\alpha(e)$ for each $e \in \text{Edge}(\Gamma)$, where the *edges* are the 1-dimensional cells of Γ .

By a concrete *multi-curve* γ on Γ we mean a 1-manifold $X(\gamma)$ and a PL map $\gamma: X(\gamma) \rightarrow \Gamma$. Given a scaling function $\rho: \text{Edge}(\Gamma) \rightarrow \mathbb{R}_{\geq 0}$, the length metric $\rho\alpha$ on Γ gives edge e the length $\rho(e)\alpha(e)$. We define the *length* of γ as

$$\ell_{\rho\alpha}(\gamma) := \sum_{e \in \text{Edge}(\Gamma)} n_\gamma(e)\rho(e)\alpha(e),$$

where $n_\gamma(e)$ is the weighted number of times that γ runs over e . We can likewise define the length of a multi-curve D on Γ as the infimum over of concrete multi-curves in D .

The *area* of Γ with respect to $\rho\alpha$ is defined to be

$$\text{Area}_\rho(\Gamma, \alpha) := \sum_{e \in \text{Edge}(\Gamma)} \rho(e)^2\alpha(e).$$

Intuitively, each edge is turned into a rectangle of width $\rho(e)$, aspect ratio $\alpha(e)$ and thus area $\rho(e)^2\alpha(e)$.

As for extremal length for surfaces, we define the square root of extremal length of a multi-curve on Γ by

$$\sqrt{\text{EL}}(D; \Gamma, \alpha) := \sup_{\rho: \text{Edge}(\Gamma) \rightarrow \mathbb{R}_{\geq 0}} \frac{\ell_{\rho\alpha}(D)}{\sqrt{\text{Area}_\rho(\Gamma, \alpha)}}. \tag{4.20}$$

It is easy to do this optimisation. We get a more interesting quantity by incorporating a filling embedding $\iota: \Gamma \hookrightarrow S$ of Γ in a surface S ; that is, an embedding ι that is π_1 -surjective. Then, for a multi-curve C on S and scaling ρ , the length is defined as in Subsection 4.4:

$$\ell_{\rho\alpha; \iota}(C) := \inf_{\substack{D \text{ on } \Gamma \\ \iota_*(D) = C}} \ell_{\rho\alpha}(D).$$

(The ‘filling’ condition guarantees that there are such multi-curves D with $\iota_*(D) = C$.) We can then define a version of extremal length, following equation (4.20), with respect to ι :

$$\sqrt{\text{EL}}(C; \Gamma, \alpha, \iota) := \sup_{\rho: \text{Edge}(\Gamma) \rightarrow \mathbb{R}_{\geq 0}} \frac{\ell_{\rho\alpha; \iota}(C)}{\sqrt{\text{Area}_\rho(\Gamma, \alpha)}}. \tag{4.21}$$

Proposition 4.22. *For $\iota: \Gamma \rightarrow \Sigma$ a filling embedding, $\sqrt{\text{EL}}(C; \Gamma, \alpha, \iota)$ satisfies the convex union, stability, homogeneity and smoothing properties.*

Proof. As in Lemma 4.16, smoothing, stability and homogeneity follow because $\ell_{\rho\alpha}(C)$ satisfies them for any ρ . Convex union follows as in Lemma 4.17. □

We can also consider extremal length with respect to an immersion ι (rather than an embedding), defined in the same way.

Proposition 4.23. For $\iota: \Gamma \rightarrow \Sigma$ a π_1 -surjective immersion, $\sqrt{\text{EL}}(C; \Gamma, \alpha, \iota)$ satisfies the convex union, homogeneity and quasi-smoothing properties.

Proof. Convex union and homogeneity still hold by the same argument. We need an extra argument for quasi-smoothing. Instead of taking the supremum over all ρ , rewrite equation (4.21) as

$$\sqrt{\text{EL}}(C; \Gamma, \alpha, \iota) = \sup_{\substack{\rho: \text{Edge}(\Gamma) \rightarrow \mathbb{R}_{\geq 0} \\ \text{Area}_\rho(\Gamma, \alpha) = 1}} \ell_{\rho\alpha}(C).$$

The immersed graph $\iota(\Gamma)$ has finitely many self-intersections. For each self-intersection x , take the supremum over the compact set of metrics $\{\rho \mid \text{Area}_\rho(\Gamma, \alpha) = 1\}$ of the diameter $\text{diam}_\iota(x)$ defined in equation (4.7). By the replacement argument in Proposition 4.5, all $\ell_{\rho\alpha}$ for ρ in this set satisfy quasi-smoothing with a uniform quasi-smoothing constant. It follows that $\text{EL}(C; \Gamma, \alpha, \iota)$ also satisfies quasi-smoothing with the same constant. \square

Remark 4.24. We can relate extremal length for elastic graphs and surfaces by making a choice of a ribbon structure to Γ . Given any $\varepsilon > 0$, a ribbon elastic graph G can be thickened into a conformal surface with boundary $N_\varepsilon(G)$ by replacing each edge e of G by a rectangle of size $\alpha(e) \times \varepsilon$ and gluing the rectangles at the vertices by using the given ribbon structure. There are then inequalities relating $\text{EL}(C; \Gamma, \alpha)$ and $\varepsilon \text{EL}(C; N_\varepsilon(G))$, to within a multiplicative factor as shown by the second author [57, Props. 4.8 and 4.9]. For graphs immersed or embedded in a surface, the situation is less clear. By suitably choosing the elastic weights on an embedded graph, it appears that one can approximate extremal length well; Palmer gives one approach [48]. We are not aware of precise theorems.

4.10. p -Extremal-length with respect to immersed graphs

Extremal length fits into a family of energies for graphs, as explored by the second author in [56, Appendix A]. For Γ a metric graph with metric g , a constant p with $1 \leq p \leq \infty$ and C a curve on Γ , define

$$E_p(C; \Gamma, g) := \sup_{\sigma: \text{Edge}(\Gamma) \rightarrow \mathbb{R}_{\geq 0}} \frac{\ell(C; \sigma g)}{\|\sigma\|_p} \tag{4.25}$$

where the L^p norm $\|\sigma\|_p$ is taken with respect to the metric g . As in the previous section, we can also consider a π_1 -surjective immersion $\iota: \Gamma \rightarrow S$ and consider C to be a curve on S rather than on Γ .

For $p = \infty$, $E_\infty(C)$ is in fact just the length with respect to g (as in Subsection 4.5). Indeed, let σ be any scaling factor, and let γ be the shortest representative of C on Γ with respect to g (not with respect to σg). Then

$$\ell(C; \sigma g) \leq \ell(\gamma; \sigma g) \leq \|\sigma\|_\infty \ell(\gamma; g),$$

from which the result easily follows.

Proposition 4.26. For any π_1 -surjective immersion $\iota: \Gamma \rightarrow S$, the curve functional $E_p(\cdot; \Gamma, g, \iota)$ satisfies convex union, homogeneity and quasi-smoothing, and thus its stabilisation extends continuously to a function on geodesic currents. If ι is a filling embedding, then E_p in addition satisfies stability and smoothing.

Proof. This follows as in Propositions 4.22 and 4.23. To prove quasi-smoothing in the immersed case, we restrict to those functions σ on $\text{Edge}(\Gamma)$ where $\|\sigma\|_p = 1$; as in the proof of Proposition 24, this set is compact. \square

5. Counting problems

One direct application of Theorem A is to obtain new counting results for curves on surfaces of a given topological type.

A *filling current* is a geodesic current $\alpha \in \mathcal{GC}(\Sigma)$ so that $i(\alpha, \mu) > 0$ for all $\mu \in \mathcal{GC}(\Sigma) \setminus \{0\}$. One example is a filling multi-curve, one whose complement in S consists of disks.

Rafi and Souto [51] proved the following.

Definition 5.1. A function f on currents is *positive* if $f(\mu) > 0$ for all $\mu \neq 0$.

For a fixed continuous, homogeneous and positive function $f : \mathcal{GC} \rightarrow \mathbb{R}_+$, α a current and L a positive real number, let

$$N(f, \alpha, L) := \#\{\phi \in MCG \mid f(\phi(\alpha)) \leq L\}.$$

Theorem 5.2 (Rafi-Souto [51, Main Theorem]). *For a fixed continuous, homogeneous and positive function $f : \mathcal{GC} \rightarrow \mathbb{R}_+$ and for a fixed filling current $\alpha \in \mathcal{GC}(\Sigma)$, the limit*

$$\lim_{L \rightarrow \infty} \frac{N(f, \alpha, L)}{L^{6g-6}}$$

exists and is equal to

$$\frac{m(f)m(\alpha)}{\mathfrak{m}_g}$$

where $m(f)$, $m(\alpha)$ and \mathfrak{m}_g are constants depending only on f , α and the genus g , respectively:

$$m(f) = \mu_{\text{Thu}}(\{\lambda \in \mathcal{ML} \mid f(\lambda) \leq 1\}) \tag{5.3}$$

$$m(\alpha) = \mu_{\text{Thu}}(\{\lambda \in \mathcal{ML} \mid i(\alpha, \lambda) \leq 1\}) \tag{5.4}$$

$$\mathfrak{m}_g = \int_{\mathcal{M}_g} m(Y) d\omega_{\text{WP}}(Y). \tag{5.5}$$

Here μ_{Thu} is the Thurston measure on \mathcal{ML} . One way to obtain it is as a scaling limit of Dirac measures on measured laminations supported on integral simple multi-curves (see Erlandsson–Souto’s book [26, Proposition 8.8]). Another way is as a certain normalisation of the volume form induced by the symplectic structure on measured laminations [26, Theorem 4.16].

Dumas [21] communicates a proof of the following theorem (attributed to Mirzakhani).

Theorem 5.6 (Dumas–Mirzakhani [21, Theorem 5.10]). *The function $\Lambda : \mathcal{M}_g \rightarrow \mathbb{R}_{\geq 0}$ given by $\Sigma \mapsto m(\text{EL}_\Sigma)$ is constant, where $m(\text{EL}_\Sigma)$ is defined by equation (5.3).*

Let $\sqrt{\text{EL}_\Sigma} : \mathcal{GC} \rightarrow \mathbb{R}_{\geq 0}$ be the continuous extension of square root of extremal length to currents provided by Theorem A. In order to be able to apply Theorem 5.2, it remains to check that the square root of extremal length is nonzero.

Lemma 5.7. *For any $\Sigma \in \text{Teich}(S)$, the functional $\sqrt{\text{EL}_\Sigma}$ is positive on $\mathcal{GC}(\Sigma)$.*

Proof. Let $A = \sqrt{-2\pi\chi(S)}$. By Definition 4.15 applied to the hyperbolic metric, $\ell_\Sigma(C)/A \leq \sqrt{\text{EL}_\Sigma}(C)$ for all curves C and all $\Sigma \in \text{Teich}(S)$. For any $\mu \in \mathcal{GC}(\Sigma)$, there exists a sequence $(\lambda_i C_i)_{i \in \mathbb{N}}$ of weighted curves so that $\lambda_i C_i \rightarrow \mu$ in the weak* sense. It thus follows that $\ell_\Sigma(\mu)/A \leq \sqrt{\text{EL}_\Sigma}(\mu)$. Since ℓ_Σ is a positive function on currents, $\sqrt{\text{EL}_\Sigma}$ is as well. □

We thus get solutions to counting problems for extremal length.

Corollary 5.8. *For any filling current α and $\Sigma \in \text{Teich}(S)$, the limit*

$$\lim_{L \rightarrow \infty} \frac{N(\sqrt{EL_\Sigma}, \alpha, L)}{L^{6g-6}}$$

exists, is independent of Σ and is equal to

$$\frac{\Lambda \cdot m(\alpha)}{\mathfrak{m}_g}. \tag{5.9}$$

Proof. Using Theorem A and the results in Subsection 4.8, we can extend $\sqrt{EL_X}$ as a continuous, real-homogeneous functional on geodesic currents. Furthermore, $\sqrt{EL_X}$ is positive on currents, by Proposition 5.7. Thus, by Theorem 5.2, the result follows. Independence of Σ follows from Theorem 5.6. \square

A similar counting result is true for α a simple multi-curve by Mirzakhani’s work, but one has to count slightly differently. For f a curve functional, α a simple curve and $L > 0$, set

$$n(f, \alpha, L) := \#\{\phi(\alpha) \in MCG \mid f(\phi(\alpha)) \leq L\}.$$

In general, note that $n(f, \alpha, L) \neq N(f, \alpha, L)$; in fact, N will be infinite if α is not filling. Even for α filling, N will be bigger than n if α has nontrivial stabiliser in the mapping class group.

We state the corresponding result for simple multi-curves.

Proposition 5.10. *For any simple multi-curve α , there is a constant $c(\alpha)$ so that, for any $f : \mathcal{GC}^+(S) \rightarrow \mathbb{R}$ continuous, positive and real-homogeneous function, the limit*

$$\lim_{L \rightarrow \infty} \frac{n(f, \alpha, L)}{L^{6g-6}} \tag{5.11}$$

exists and is equal to

$$\frac{m(f)c(\alpha)}{\mathfrak{m}_g}. \tag{5.12}$$

Proof. First note that the set $\{\lambda \in \mathcal{ML}(\Sigma) \mid f(\lambda) \leq 1\}$ is compact because f is positive on nonzero measured laminations. Let $A := \{\lambda \in \mathcal{ML} \mid f(\lambda) = 1\}$; then $\mu_{\text{Thu}}(A) = 0$, as proved by Rafi and Souto [51, p. 879]. Finally, we apply Mirzakhani’s counting result [43, Theorem 1.3] and the Portmanteau theorem (see [6, Theorem 30.12]) to conclude the limit exists. \square

Corollary 5.13. *For any simple multi-curve α and $\Sigma \in \text{Teich}(S)$, the limit*

$$\lim_{L \rightarrow \infty} \frac{n(\sqrt{EL_\Sigma}, \alpha, L)}{L^{6g-6}}$$

exists, is independent of Σ and is equal to

$$\frac{\Lambda \cdot c(\alpha)}{\mathfrak{m}_g}. \tag{5.14}$$

Remark 5.15. The constant $c(\alpha)$ in Proposition 5.13 is not the same as the constant $m(\alpha)$ in Theorem 5.2. For details on how $c(\alpha)$ is defined, see Mirzakhani’s paper [43, Equation (1.2)]. This is related to the fact that in Proposition 5.13 we count multi-curves instead of mapping classes because the stabiliser of a simple multi-curve under the mapping class group is infinite.

The counting problem $n(\ell_\Sigma, \alpha, L)$ with α an arbitrary essential multi-curve (possibly neither simple nor filling) and $f = \ell_\Sigma$ a hyperbolic length is established in more recent work of Mirzakhani [44,

Theorem 1.1]. Relying on her work, Erlandsson–Parlier–Souto [23, Theorem 1.6] give the corresponding result where f is allowed to be intersection number with other filling currents (not just a hyperbolic Liouville current). From Mirzakhani’s work and work of Erlandsson–Souto [24, Corollary 4.4], one also can get the corresponding counting problems where α is allowed to be a current (not just a multi-curve). In fact, from Erlandsson–Souto’s work one can also allow f to be any continuous, positive and real-homogeneous function on currents, although they do not explicitly state this in their paper. This is done in Rafi–Souto’s work (Theorem 5.2 above) which also gives the expression (5.9) for the limit of the counting problem. Rafi–Souto also relies on Mirzakhani’s work [44].

Finally, recently Erlandsson–Souto [26, Theorem 8.1] gave an independent proof of the counting argument in [44] illuminating the connection between counting problems for simple and nonsimple multi-curves. Corollary 5.8 also follows from Proposition 5.13 and forthcoming work of Erlandsson and Souto [25, 26].

We remark that Erlandsson–Souto’s work shows that if one knows a counting result for simple closed curves, then one can obtain a counting result for nonsimple closed curves (for curve functionals extending continuously to currents). The connection between these two types of counting problems is perhaps that, in some sense, the simple closed curves are the extremal points of the space of currents, in the sense of convex sets [58]. For instance, the systole of positive curve functional satisfying smoothing and convex union is always a simple curve. (Here by *systole* we mean a weight 1, nontrivial multi-curve C with a minimal value of $f(C)$.)

6. Proof outline

In this section, we prove the core theorem of the article, Theorem 6.1, giving a continuous extension to geodesic currents of a functional f on weighted multi-curves satisfying convex union, stability, homogeneity and weighted quasi-smoothing.

Theorem 6.1. *Let f be a weighted curve functional defined on weighted oriented multi-curves satisfying the weighted quasi-smoothing, convex union, stability and homogeneity properties. Then there is a unique continuous homogeneous function $\tilde{f}: \mathcal{GC}^+(S) \rightarrow \mathbb{R}_{\geq 0}$ that extends f .*

Proof of Theorem 6.1. The proof proceeds by studying the geodesic flow on the unit tangent bundle to S (with respect to an arbitrary hyperbolic metric), picking a suitable global cross section with boundary τ and looking at a ‘smeared first return map’ to τ .

The proof breaks up into the following steps.

- o **Step 1:** In Section 7, we define the cross sections we consider, though we delay proving existence. We introduce *bump functions* and the associated *smeared first return map* (Definition 7.4); indeed, there are several varieties of return maps (Table 2). The main advantage of smeared return maps is that they are continuous (Proposition 7.7). We use these smeared returns to define our purported extension f_τ to geodesic currents as a limit (Definition 7.25), assuming a suitable global cross section τ exists.

Table 2. Various types of return maps. $M = UT\Sigma$ is the domain of the flow ϕ_t .

Notation	Type	Meaning
p	$\tau \rightarrow \tau$	Ordinary first return
P	$\tau \rightarrow \mathbb{R}_1 \tau$	Smeared first return
q	$\tau \rightarrow \tau \times \pi_1(M)$	Complete homotopy return
Q	$\tau \rightarrow \mathbb{R}_1(\tau \times \pi_1(M))$	Smeared homotopy return
m	$\tau \rightarrow \pi_1(M)$	Return curve (projection of q)
$[m]$	$\tau \rightarrow \mathcal{C}^+(S)$	Conjugacy class of projection of m to S
M	$\tau \rightarrow \mathbb{R}_1 \pi_1(M)$	Smeared return curve (projection of Q)
$[M]$	$\tau \rightarrow \mathbb{R}_1 \mathcal{C}^+(S)$	
R	$\mathcal{GC}^+(S) \rightarrow \mathbb{R} \mathcal{C}^+(S)$	Integral of $[M]$

- **Step 2:** In Section 8, we find a suitable ‘good’ τ , defined in Definition 8.4, by considering certain *wedge subsets* (Definition 8.1) of $UT\Sigma$, based at a collection of geodesic sub-segments of a closed geodesic δ (see Proposition 8.3).
- **Step 3:** In Section 9, Proposition 9.6, we show the limit defining f_τ exists, by using convex union, quasi-smoothing, stability and homogeneity assumptions on f and applying a version of Fekete’s Lemma (Lemma 9.5).
- **Step 4:** In Section 10, Proposition 10.8, we show that f_τ is a continuous function on the space of oriented geodesic currents. We do this by showing that the iterates f_τ^k used in the definition of f_τ are continuous for each k , using the continuity of the smeared homotopy return map from Step 1, and f is a continuous function of the weights of a fixed multi-curve (Proposition 3.5). Propositions 9.4 and 10.1 give enough control to ensure uniform convergence of the iterates f_τ^k to f_τ (see Lemma 10.2), and thus continuity of f_τ follows.
- **Step 5:** In Section 11, Proposition 11.5, we show that f_τ extends f for oriented curves, by analysing the image of the smeared return map in that case and mixing and matching the components of the multi-curve.

The definition of f_τ depends on many choices: the hyperbolic metric on S , a choice of global cross section τ and, in fact, nested cross sections $\tau_0 \subset \tau \subset \tau'$ and a choice of bump function ψ on τ (see Definition 7.1). Different choices yield, a priori, different extensions f_τ . But we have proved that f_τ is a continuous function on the space of geodesic currents and, moreover, it restricts to f on multi-curves. Since weighted multi-curves are a dense subset of the space of geodesic currents (see Subsection 2.7), the extension does not depend on these choices. This proves Theorem 6.1. \square

Now, we prove Theorem A.

Proof of Theorem A. By Proposition 3.6, f extends uniquely to a weighted curve functional satisfying convex union, homogeneity, stability and weighted quasi-smoothing with the same constant. Then, by Theorem 6.1, the theorem follows. \square

7. Defining the extension

We now turn to the proof of Theorem A. As mentioned above, we will fix a hyperbolic structure Σ on S (with no relation to the curve functional f) and use the geodesic flow ϕ_t on the unit tangent bundle to define the extension to currents.

In this section we will deal with *return maps* for this flow. After some generalities about return maps for cross sections with boundary, we introduce a *smeared return map* that is continuous. We also define a *homotopy return map* (and a smeared version of it) that keeps track of homotopy classes of closures of trajectories. Then we use the smeared homotopy return map to define the extension f_τ of f . The several variants of the return map are summarised in Table 2.

7.1. Smeared first return map

Let Y be a smooth closed manifold with a smooth flow ϕ_t . For us, a *cross section* is a compact smooth codimension 1 submanifold-with-boundary τ that is smoothly transverse to the foliation of Y given by ϕ_t . A *global cross section* is a cross section τ so that its interior τ° intersects all forward and backward orbits: for all $x \in Y$, there exists $s < 0$ and $t > 0$ with $\phi_s(x), \phi_t(x) \in \tau^\circ$. Any flow on a compact manifold has global cross section consisting of a union of finitely many disks, although not all flows on compact manifolds admit a global cross section consisting of a single connected component. (For instance, the Reeb foliation on \mathbb{T}^2 has no connected global cross section.) By the implicit function theorem, for any cross section τ there is a larger cross section τ' with $\partial\tau \subset \tau'^\circ$, which we also write $\tau \Subset \tau'$.

Let $t_\tau: Y \rightarrow \mathbb{R}$ be the first return time defined by $t_\tau(x) := \min\{t > 0 \mid \phi_t(x) \in \tau\}$, and let $p_\tau(x) := \phi_{t_\tau(x)}(x)$. Then p_τ (restricted to τ) is the *first return map* associated to the cross section τ .

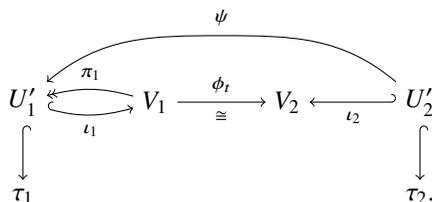
We will omit the subscript on p_τ if it is clear from context. We also have the first return time to the interior, denoted t_τ° . (Recall we assume $t_\tau^\circ(x)$ is finite.)

If τ has no boundary, then p is a homeomorphism. On the other hand, if the cross section has a noninvariant boundary (i.e., $p(\partial\tau) \neq \partial\tau$), then t_τ and p will have discontinuities. This necessarily happens for the geodesic flow on the unit tangent bundle of a hyperbolic surface. See Cossarini–Dehornoy [15, Sec. 1] for a justification and examples of global cross sections with boundary for this flow; we construct our own cross section in Section 8. However, by the continuity of ϕ_t with respect to initial parameters, we have the following ‘local continuity’ claim.

Lemma 7.1. *Let τ_1, τ_2 be cross sections (not necessarily global) of ϕ . Let $x_1 \in \tau_1^\circ$, and suppose we are given $t > 0$ so $x_2 = \phi_t(x_1) \in \tau_2^\circ$. Then there exists a neighbourhood U_1 of x_1 in τ_1 , a neighbourhood U_2 of x_2 in τ_2 and a continuous function $t_2^1: U_1 \rightarrow \mathbb{R}_{>0}$ so that for $x \in U_1$, $\phi_{t_2^1}(x) \in U_2$. Furthermore, $\phi_{t_2^1}: U_1 \rightarrow U_2$ is a diffeomorphism. If t is the first return to τ_2 , then we can choose U_1 and U_2 so that $t_2^1(x) = t_{\tau_2}(x)$ is also the first return time.*

This is presumably standard (Basener gives this as ‘a useful technical lemma, the proof of which is trivial’ [3, Lem. 1]), but we give a proof for completeness.

Proof. Pick an initial neighbourhood U'_1 of x_1 in τ , and let ε be small enough so that $V_1 := \phi_{(-\varepsilon, \varepsilon)}(U'_1)$ is a 3-dimensional flow-box neighbourhood of x_1 in Y . Then the restriction of ϕ_t to V_1 is a homeomorphism to a neighbourhood V_2 of x_2 in Y . Set $U'_2 := V_2 \cap \tau_2$. Now consider the composition $\psi := \pi_1 \circ \phi_{-t} \circ \iota_2: U'_2 \rightarrow U'_1$, where $\iota_i: U'_i \hookrightarrow V_i$ is the inclusion and $\pi_1: V_1 \rightarrow U'_1$ is the flow projection:



Then ψ is a map from U'_2 to U'_1 , taking x_2 to x_1 . By transversality of τ_1 and τ_2 , the differential of ψ at x_2 is invertible. Thus, by the inverse function theorem, there is a neighbourhood U_2 of x_2 and U_1 of x_1 so that the restriction of ψ is a diffeomorphism from U_2 to U_1 . For $x \in U_1$, set $t_2^1(x) := t + \pi_t(\phi_{-t}(\psi^{-1}(x)))$, where $\pi_t: V_1 \rightarrow (-\varepsilon, \varepsilon)$ is the projection onto the time coordinate of the flow box. We have $\phi_{t_2^1}(x) = \psi^{-1}(x) \in U_1$, as desired for the first claim.

For the second claim, by hypothesis, the compact sets $\phi_{[0, t]}(x_1)$ and τ_2 do not intersect, so $\phi_{[0, t]}(x_1)$ has an open neighbourhood that does not intersect τ_2 . It follows that we can shrink U_1 and U_2 so that ψ^{-1} restricted to U_1 is the first return map to τ_2 . □

Lemma 7.2. *Let $x \in \tau$. If $p(x) \in \tau^\circ$, then p_τ and t_τ are continuous in a neighbourhood of x in τ .*

Proof. Let $\tau' \ni \tau$ be a slightly enlarged global cross section (to cover cases when $x \in \partial\tau$). By Lemma 7.1, there exists a neighbourhood U of x in τ' such that $p_\tau(U) \subset \tau^\circ$. By taking $V := U \cap \tau$, we get the desired neighbourhood in τ . □

Lemma 7.3. *Let τ be a global cross section. Then, on τ , t_τ is lower semi-continuous and t_τ° is upper semi-continuous. There are thus positive global upper and lower bounds on t_τ .*

Proof. Fix $x \in \tau$. If $p_\tau(x) \in \tau^\circ$, then t_τ is continuous at x by Lemma 7.2. Otherwise, find a cross section τ' with $\tau \Subset \tau'$. Then $t_{\tau'}$ is continuous at x , and since $t_\tau(y) \geq t_{\tau'}(y)$ we have proved that t_τ is lower semi-continuous.

On the other hand, for any $x_1 \in \tau$, we can set $x_2 = p_\tau^\circ(x_1)$ and find a neighbourhood U_1 of x_1 in τ with a function t_2^1 as in Lemma 7.1. But then $t_\tau^\circ(x) \leq t_2^1(x)$ for $x \in U_1$. (Note that x_2 is the first return point to τ° , not to τ , so we cannot conclude that t_τ° is continuous.) □

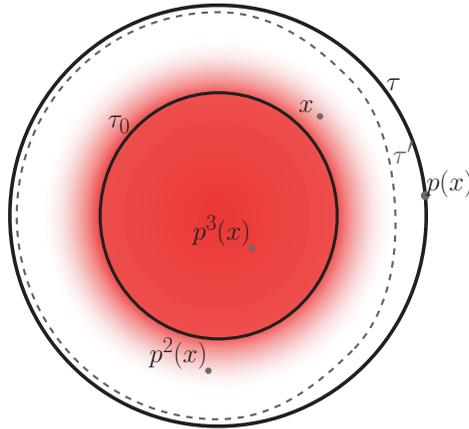


Figure 7.1. Smeared first return map, illustrating the proof of continuity in the case $n = 3$ (before shrinking τ). The bump function ψ is indicated by the density of red.

In the C^1 setting, Basener showed that a global cross section can be perturbed slightly so that the first return map is piecewise continuous with a cellular structure [4]. However, we want a *continuous* version of the first return map, so we proceed in a different direction.

Definition 7.4. Fix a nested pair of global cross sections $\tau_0 \Subset \tau$. A bump function ψ for this pair is a continuous function $\psi : \tau \rightarrow [0, 1]$ so that ψ is 1 on τ_0 and 0 on an open neighbourhood of $\partial\tau$. Set $\bar{\psi}(x) = 1 - \psi(x)$. Let $p : \tau \rightarrow \tau$ be the first return map with respect to τ . Then the *smeared first return map* of ψ is a function $P_\psi : \tau \rightarrow \mathbb{R}_1\tau$ defined by

$$P_\psi(x) := \begin{cases} p(x) & p(x) \in \tau_0 \\ \psi(p(x)) \cdot p(x) + \bar{\psi}(p(x)) \cdot P_\psi(p(x)) & p(x) \in \tau - \tau_0. \end{cases}$$

Here, $\mathbb{R}_1\tau \subset \mathcal{M}_1(\tau)$ is the subspace of measures with finite support and total mass 1; see Convention 2.24.

The convention here is that a *smeared map* takes values in finite linear combinations of the target space (or maybe in measures). We use capital letters for smeared maps.

Intuitively, we iterate x forward, stopping at each iterate with probability given by ψ . More visually, imagine the original cross section as a disk. As we look along the flow lines, we see an overlapping set of disks, with hard edges between them. To find the smeared first return map, we ‘feather’ the edges by giving the disks partially transparent boundaries made out of cellophane. If we continuously increase the transparency towards the boundary, the resulting image will have soft edges. See Figure 7.1.

We will usually omit ψ from the notation and denote the smeared first return map by P .

Since τ_0 is a global cross section, in the definition of P we eventually take the first choice, and so $P(x)$ is a finite sum of elements of τ_1 as claimed. A little more is true.

Lemma 7.5. *If $\tau_0 \Subset \tau_1$ is a nested pair of global cross sections, there is an $N > 0$ so that for any $x \in \tau_1$, there is an integer $k < N$ so that $p_1^k(x) \in \tau_0$, where p_1 denotes the first return map for τ_1 .*

Proof. By Lemma 7.3, there is an upper bound on the return time from τ_1 to τ_0 and thus an upper bound on the number of intersections of the return path to τ_0 with the compact set τ_1 . \square

As a consequence of Lemma 7.5, we can rewrite P directly. Let N be the bound from Lemma 7.5. Then

$$\begin{aligned}
 P(x) &= \psi(p(x)) \cdot p(x) + \bar{\psi}(p(x))\psi(p^2(x)) \cdot p^2(x) + \dots \\
 &= \sum_{k=1}^N \bar{\psi}(p(x)) \cdot \dots \cdot \bar{\psi}(p^{k-1}(x))\psi(p^k(x)) \cdot p^k(x).
 \end{aligned}
 \tag{7.6}$$

If we extend the upper limit of the sum beyond N , the additional terms will be 0.

Proposition 7.7. *For any nested global cross sections $\tau_0 \Subset \tau$ and bump function ψ , the smeared first return map P is continuous.*

Proof. We wish to show that P is continuous at $x \in \tau$. There is some first $n > 0$ such that $p^n(x) \in \tau_0^\circ$. If there is any i between 1 and n so that $p^i(x) \in \partial\tau$, we first find a smaller cross section τ' with $\tau_0 \Subset \tau' \Subset \tau$ without this problem, as follows. Recall that we assumed that ψ vanishes in a neighbourhood of $\partial\tau$. Since there are finitely many points $p^i(x)$ in the open set $\tau \setminus \text{supp}(\psi)$, we can pick τ' containing $\text{supp}(\psi)$ so that its boundary avoids those finitely many $p^i(x)$. Since ψ vanishes on $\tau \setminus \tau'$, the smeared first return map defined with respect to τ' agrees with that defined with respect to τ . By replacing τ by τ' , we may thus assume that $p^i(x) \notin \partial\tau$. Similarly, shrink τ_0 so that $p^i(x) \notin \partial\tau_0$ for $0 < i < n$.

We proceed by induction on n . If $n = 1$ – that is, if $p(x) \in \tau_0^\circ$ – then p is continuous at x by Lemma 7.2, and P is continuous since the map taking a point x to the delta function δ_x is continuous. Otherwise, note that p_1 (the return map for τ) is continuous at x (again by Lemma 7.2, since $p_1(x) \in \tau^\circ$). By induction, P is continuous at $p_1(x)$, and therefore

$$P(x) = \psi(p_1(x)) \cdot p_1(x) + \bar{\psi}(p_1(x)) \cdot P(p_1(x))$$

is continuous at x . □

To define iterates of P , we first extend P and other functions to act on measures.

Definition 7.8. When X, Y are measure spaces and $f: X \rightarrow Y$ is a measurable function, by convention we extend f to a function $\mathcal{M}_1(X) \rightarrow \mathcal{M}_1(Y)$ acting on measures, denoted f_* (or simply f), by setting

$$f_*(\mu)(S) := \mu(f^{-1}(S))$$

for $\mu \in \mathcal{M}_1(X)$ and $S \subset Y$ a measurable set. If f is continuous, then this extension is continuous with respect to the weak* topology on $\mathcal{M}_1(X)$ and $\mathcal{M}_1(Y)$. (This uses Proposition 2.23 and the fact that if $g: Y \rightarrow \mathbb{R}$ is bounded, then $g \circ f: X \rightarrow \mathbb{R}$ is also bounded.) If f is invertible and $\psi: X \rightarrow \mathbb{R}_{\geq 0}$ is a scaling factor, then

$$f(\psi \cdot \mu) = (\psi \circ f^{-1}) \cdot f(\mu). \tag{7.9}$$

In practice, we will often be interested in the subspace of finitely supported measures, in which case the extension $f: \mathbb{R}_1 X \rightarrow \mathbb{R}_1 Y$ is given by

$$f\left(\sum a_i x_i\right) := \sum a_i f(x_i).$$

For $F: X \rightarrow \mathcal{M}_1(Y)$ a smeared function, we extend F to a function $\mathcal{M}_1(X) \rightarrow \mathcal{M}_1(Y)$ from measures to measures, denoted \tilde{F} (or simply F), by setting, for any measurable function $\varphi: Y \rightarrow \mathbb{R}_{\geq 0}$,

$$\begin{aligned}
 F_\varphi(x) &:= \int_{y \in Y} \varphi(y)(F(x))(y) \\
 \int_{y \in Y} \varphi(y)\tilde{F}(\mu)(y) &:= \int_{x \in X} F_\varphi(x)\mu(x),
 \end{aligned}$$

where $F_\varphi: X \rightarrow \mathbb{R}_{\geq 0}$ is an auxiliary function. See also equations (7.11) and (7.12).

Proposition 7.10. *If $F : X \rightarrow \mathcal{M}_1(Y)$ is continuous, then the extension $\tilde{F} : \mathcal{M}_1(X) \rightarrow \mathcal{M}_1(Y)$ is continuous.*

Proof. Let $(x_i)_{i=0}^\infty$ be a sequence approaching $x \in X$. By assumption, $F(x_i)$ approaches $F(x)$ in the weak* topology. By Proposition 2.23, this is equivalent to saying that for all continuous bounded functions $\varphi : Y \rightarrow \mathbb{R}_{\geq 0}$ the function F_φ above is continuous. Furthermore, F_φ is bounded since F takes values in probability measures. We now show that \tilde{F} is continuous. Let $\mu_i \rightarrow \mu \in \mathcal{M}(X)$. We want to show that $\tilde{F}(\mu_i) \rightarrow \tilde{F}(\mu) \in \mathcal{M}_1(Y)$; that is, for any continuous bounded function $\varphi : Y \rightarrow \mathbb{R}_{\geq 0}$,

$$\int_{x \in X} F_\varphi(x) \mu_i(x) \rightarrow \int_{x \in X} F_\varphi(x) \mu(x).$$

This is true by definition of the weak* topology in $\mathcal{M}_1(X)$ and Proposition 2.23, since F_φ is continuous and bounded. □

In our applications, F takes values in finitely supported measures, with a bound on the size of the support. Concretely, if $F : X \rightarrow \mathbb{R}Y$ can be written as a finite sum

$$F(x) = \sum_{i=1}^N \psi_i(x) f_i(x)$$

for real-valued functions ψ_i and invertible Y -valued functions f_i , then, by equation (7.9), the extension is defined by

$$F(\mu) = \sum_{i=1}^N f_i(\psi_i \cdot \mu) = \sum_{i=1}^N (\psi_i \circ f_i^{-1}) \cdot f_i(\mu) \tag{7.11}$$

where the middle expression is a sum of pushforwards of scaled measures, and in the last expression we have pulled the scaling factors out. If μ is also finitely supported, we have

$$F\left(\sum_i a_i x_i\right) = \sum_{i,j} a_i \psi_j(x_i) f_j(x_i). \tag{7.12}$$

Definition 7.13. With the above extension of notation, the iterates of the smeared return map P are defined by

$$\begin{aligned} P^0(x) &:= x \\ P^n(x) &:= P(P^{n-1}(x)). \end{aligned} \tag{7.14}$$

Definition 7.15. A measure ν on Y that is invariant under the flow ϕ_t induces a flux $\mu = \nu_\tau$ on a global cross section τ that is invariant under the first return map p (see Viana and Oliveira [59, Section 3.4.2]). Concretely, pick $\varepsilon > 0$ small enough so that the map $b : [0, \varepsilon] \times \tau \rightarrow Y$ defined by $b(t, x) = \phi_t(x)$ is an embedded flow box. Then for $S \subset \tau$, define $\mu(S) := \nu(\phi_{[0, \varepsilon]}(S))/\varepsilon$. (Since our cross sections are compact manifolds-with-boundary, we can always find such an ε .)

We have the following invariance property.

Proposition 7.16. *If ν is a measure on $UT\Sigma$ that is invariant under ϕ_t , then ν_τ is invariant under p and $\psi \nu_\tau$ is invariant under P .*

For motivation for the factor of ψ in the proposition statement, think about extending the definitions to allow ψ to be the (noncontinuous) characteristic function of $\tau_0 \Subset \tau$; then P is the ordinary first return map to τ_0 , and $\psi \nu_\tau = \nu_{\tau_0}$ is the flux of τ_0 . See also Example 9.3.

Proof. The first part is standard. For the second part, let $\mu = \nu_\tau$. Then we have

$$\begin{aligned} P(\psi \cdot \mu) &= p(\psi \cdot (\psi \circ p) \cdot \mu) + p^2(\psi \cdot (\bar{\psi} \circ p) \cdot (\psi \circ p^2) \cdot \mu) + \dots \\ &= \sum_{k=1}^N p^k(\psi \cdot (\bar{\psi} \circ p) \cdots (\bar{\psi} \circ p^{k-1}) \cdot (\psi \circ p^k) \cdot \mu) \\ &= \sum_{k=1}^N (\psi \circ p^{-k}) \cdot (\bar{\psi} \circ p^{-k+1}) \cdots (\bar{\psi} \circ p^{-1}) \cdot \psi \cdot p^k(\mu) \\ &= \sum_{k=1}^N (\psi \circ p^{-k}) \cdot (\bar{\psi} \circ p^{-k+1}) \cdots (\bar{\psi} \circ p^{-1}) \cdot \psi \cdot \mu \end{aligned}$$

using the definition of P (in the form of equation (7.6)), rewriting as a sum, equation (7.11) and invariance of μ under p . Since $\psi + \bar{\psi} = 1$, this sum telescopes, and the result is $\psi\mu$. □

We will sometimes blur the distinction between a geodesic current and its flux and write, for instance, $\nu(\tau)$ for the total mass of ν_τ on τ , or $\nu(\psi\tau)$ for $\int_{x \in \tau} \psi(x)\nu_\tau(x)$.

7.2. Homotopy type of return

We will additionally need to track *how* a point returns to the cross section. For this, we suppose that we have a global cross section τ contained in a simply connected cross section τ' . (For a C^1 flow on a manifold of dimension at least 3, there is always a simply connected cross section [4].) For such a cross section, from the first return for $x \in \tau$, we can extract another piece of information: the homotopy class of the return trajectory.

Definition 7.17. Let ϕ_t be a flow on a manifold Y and τ be a global cross section, contained in a larger compact simply connected cross section τ' . Fix a basepoint $* \in \tau'$. For $x \in \tau$, define the *return trajectory* $m(x) \in \pi_1(Y, *)$ by taking the homotopy class of a path that runs in τ' from $*$ to x , along the flow trajectory from x to $p_\tau(x)$, and then in τ' from $p_\tau(x)$ back to $*$. Since τ' is simply connected, $m(x)$ is independent of the choice of path.

Lemma 7.18. *Let Y be a compact manifold with flow ϕ_t and τ be a global cross section. As x varies in τ , the return trajectory $m(x)$ takes on only finitely many values.*

Proof. Since there are upper bounds on the return time (Lemma 7.3), on the speed of ϕ_t with respect to a Riemannian metric on Y and on the diameter of τ' , the length of the path representing $m(x)$ is bounded. On a compact manifold, there are only finitely many elements of $\pi_1(Y, *)$ that have representatives of bounded length. □

To get the return map for iterates, we also incorporate the point of first return.

Definition 7.19. The *homotopy return map* is the map $q: \tau \rightarrow \tau \times \pi_1(Y, *)$ defined by

$$q(x) := (p(x), m(x)).$$

We can iterate q by inductively defining q^{n+1} to be the composition

$$\tau \xrightarrow{q^n} \tau \times \pi_1(Y) \xrightarrow{q \times \text{id}} \tau \times \pi_1(Y) \times \pi_1(Y) \xrightarrow{(x,g,h) \mapsto (x,hg)} \tau \times \pi_1(Y).$$

Define $m^n(x) \in \pi_1(Y, *)$ to be the second component of $q^n(x)$.

Remark 7.20. An alternative approach to defining homotopy types of return trajectories is to pick a cross section in the universal cover, as done by Erlandsson–Parlier–Souto in [23, Section 3.2].

Definition 7.21. For τ a global cross section with basepoint $*$, $\tau_0 \subseteq \tau$ a smaller global cross section, ψ a bump function for this pair and $\tau' \supset \tau$ a simply connected cross section, the *smearred homotopy return map* $Q: \tau \rightarrow \mathbb{R}_1(\tau \times \pi_1(Y, *))$ is defined by

$$Q(x) := \begin{cases} q(x) & p(x) \in \tau_0 \\ \psi(p(x)) \cdot q(x) + \bar{\psi}(p(x)) \cdot L_{m(x)}Q(p(x)) & p(x) \in \tau - \tau_0 \end{cases}$$

where L_g is left translation by $g \in \pi_1(Y, *)$:

$$L_g \left(\sum_i a_i(x_i, h_i) \right) := \sum_i a_i(x_i, gh_i).$$

There is once again a natural notion of iteration, defined by inductively setting Q^{n+1} to be the composition

$$\tau \xrightarrow{Q^n} \mathbb{R}_1(\tau \times \pi_1(Y)) \xrightarrow{\mathbb{R}_1(Q \times \text{id})} \mathbb{R}_1(\mathbb{R}_1(\tau \times \pi_1) \times \pi_1(Y)) \xrightarrow{\text{join}} \mathbb{R}_1(\tau \times \pi_1(Y)),$$

where join is the somewhat more involved operation

$$\text{join} \left(\sum_i a_i \left(\left(\sum_j b_{ij}(x_{ij}, g_{ij}) \right), h_i \right) \right) := \sum_{i,j} a_i b_{ij}(x_{ij}, h_i g_{ij}).$$

(The terminology comes from the theory of monads [45, 60]. See equation (7.12).)

Definition 7.22. We define the *smearred nth return trajectory* $M^n: \tau \rightarrow \mathbb{R}_1\pi_1(Y)$ to be the composition

$$\tau \xrightarrow{Q^n} \mathbb{R}_1(\tau \times \pi_1(Y)) \longrightarrow \mathbb{R}_1\pi_1(Y)$$

where at the second step we lift the projection on the second component to act on weighted objects as in Definition 7.8.

Let $\Lambda(n, \tau)$ be the set of curves that appear with nonzero coefficient in $M^n(x)$ for some $x \in \tau$.

Lemma 7.23. $\Lambda(n, \tau)$ is finite.

Proof. Immediate from Lemmas 7.5 and 7.18. □

Lemma 7.24. The maps Q^k and M^k are continuous.

Proof. The proof of Proposition 7.7 also proves that Q is continuous. It then follows that Q^k and M^k are continuous. □

7.3. Return maps for the geodesic flow

We now turn to the specifics of our situation. Let Σ be the surface S endowed with an arbitrary hyperbolic Riemannian metric g . Points in $UT\Sigma$ will be denoted \vec{x} , meaning a pair of a point $x \in \Sigma$ and a unit tangent vector at x . Let $\phi_t: UT\Sigma \rightarrow UT\Sigma$ be the geodesic flow associated to g .

We pick nested global cross sections $\tau_0 \subseteq \tau \subset \tau'$, with τ' simply connected, and a bump function ψ for the pair (τ_0, τ) . We thus get a smearred nth return trajectory $M^n: \tau \rightarrow \mathbb{R}_1\pi_1(UT\Sigma, *)$. We want to work with curves in Σ rather than its unit tangent bundle, so compose with the projection $\pi_\Sigma: UT\Sigma \rightarrow \Sigma$ to get a linear combination of elements of $\pi_1(S, \pi_\Sigma(*))$. Then take conjugacy classes (to pass to unbased curves) to get an element of $\mathbb{R}_1C^+(S)$. We call the resulting map $[M^n]$, which has type

$$[M^n]: \tau \rightarrow \mathbb{R}_1C^+(S).$$

From Lemmas 7.23 and 7.24, $[M^n]$ is a continuous function with values in the finite-dimensional subspace $\mathbb{R}^{[\Lambda(n,\tau)]} \subset \mathbb{R}\mathcal{C}^+(S)$, where $[\Lambda(n,\tau)]$ is the projection of $\Lambda(n,\tau)$.

7.4. Definition of the extension

Now, we will use the above return map $[M^n]$ to define the extension of f to geodesic currents in Theorem A. With $\tau_0 \Subset \tau \subset \tau'$ as above, we define

$$R^n : \mathcal{GC}^+(S) \rightarrow \mathbb{R}\mathcal{C}^+(S)$$

$$R^n(\mu) := \int_{\tau} [M^n(\vec{x})]\psi(\vec{x})d\mu_{\tau}(\vec{x}).$$

Observe that, for fixed n , $R^n(\mu)$ is a weighted multi-curve with a fixed set of possible connected components but with weights depending on μ . As we will explain in Section 10, because $[M^n(\vec{x})]$ is continuous on τ , R^n is continuous with respect to the weak* topology on $\mathcal{GC}^+(S)$.

We can now finally define our extension of f .

Definition 7.25. Let τ be a good cross section and f a weighted curve functional satisfying stability, homogeneity, weighted quasi-smoothing and convex union. We define

$$f_{\tau}^n(\mu) := f(R^n(\mu)) \tag{7.26}$$

$$f_{\tau}(\mu) := \lim_{n \rightarrow \infty} \frac{f_{\tau}^n(\mu)}{n}. \tag{7.27}$$

We will prove that the limit exists (at least for our cross section) in Proposition 9.6.

Warning 7.28. We work with weighted linear combinations of objects (or, more generally, measures) at many places in the article. Some functions (like R^n) are by definition additive under linear combinations, and in Definition 7.8 we also silently extend other functions (like p) to apply additively to linear combinations of points or measures. But the main curve functional f we are interested is *not* necessarily additive. (We only assume that f satisfies convex union in the main theorems.)

8. Constructing global cross sections

Next we define the specific global cross section we use. We make choices that are convenient for guaranteeing that certain crossings are essential.

Definition 8.1. For Σ a hyperbolic surface, c an oriented geodesic segment on Σ and $0 < \theta < \pi/2$ an angle, the *wedge set* $W(c, \theta) \subset UT\Sigma$ is the set of vectors that cross c nearly perpendicularly:

$$W(c, \theta) := \{ \vec{x} = (x, v) \mid x \in c, |\text{ang}(T_x c, v) - \pi/2| \leq \theta \}.$$

(Angles $\text{ang}(v, w)$ are measured by the counterclockwise rotation from v to w .) We can likewise define the wedge set $W(\{c_i\}, \theta)$ for a collection of geodesic segments $\{c_i\}_{i=1}^k$.

We wish to find a wedge set $W(\{c_i\}, \theta)$ that is an embedded global cross section for the geodesic flow ϕ_t .

Fix $\theta = \pi/6$. For any geodesic arc c , the wedge set $W(c, \theta)$ intersects any geodesic that passes through a nonempty open set. Thus, by compactness of $UT\Sigma$, there exist a finite collection of immersed arcs $(c_i)_{i=1}^n$ so that $\bigcup_{i=1}^n W(\{c_i\}, \theta)_{i=1}^n$ is a disconnected, not necessarily embedded, global cross section of the geodesic flow. We will produce an embedded global cross section from it. Immersed points come from intersection points between the geodesic segments c_i , but not all of them produce immersed points of the global cross section.

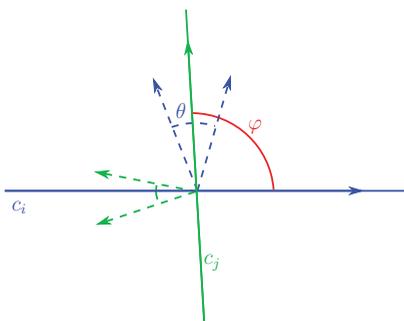


Figure 8.1. Wedge sets not intersecting, in the case $2\theta < \varphi$.

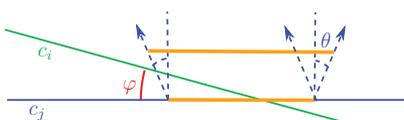


Figure 8.2. Cut and flow operation when the angle of intersection is small, $\theta < \pi/2 - \varphi$.

Indeed, suppose $\text{ang}(c_i, c_j) = \varphi$. There are two good cases:

- (1) If $2\theta < |\varphi|$, the wedge sets do not intersect, as shown in Figure 8.1.
- (2) If $\theta < \frac{\pi}{2} - |\varphi|$, then the corresponding wedge sets do intersect, but we can perturb the c_j slightly to avoid the intersection. Given a small interval $[a, b]$ of the geodesic segment c_j , containing one intersection point between c_i and c_j , we consider the wedge $W([a, b], \theta)$. By pushing the endpoints a, b forward along the extremal angles and removing $[a, b]$ from c_j , as shown in Figure 8.2, we obtain new interval with endpoints a', b' and a new wedge set $W([a', b'], \theta')$, for some $\theta' > \theta$, so that $W([a', b'], \theta')$ is disjoint from c_i and $W([a', b'], \theta)$ intersects every geodesic that $W([a, b], \theta)$ does (so we still have a global cross section).

Remark 8.2. Note that the inequality must be strict in case (2). Indeed, if $\theta = \pi/2 - |\varphi|$, the wedge sets do intersect, but this is not a good case: we cannot guarantee to make the wedge sets locally disjoint by flowing a segment of c_j forward slightly.

Our choice $\theta = \pi/6$ guarantees that one of these two cases happens.

Next, we will construct an immersed connected global cross section τ' containing this disconnected wedge set.

Proposition 8.3. For any wedge set $W(\{c_i\}, \theta)$ and any $\varepsilon > 0$, there exists a closed geodesic δ so that for each i , there is a subsegment $\delta_i \subset \delta$ so that $c_i \subset B_\varepsilon(\delta_i)$ and every geodesic that intersects $W(c_i, \theta)$ also intersects $W(\delta_i, \theta + \varepsilon)$.

Proof. Use [5, Theorem 2.4] to construct the closed geodesic δ with $c_i \subset B_\varepsilon(\delta)$. If ε is small enough, by following the geodesic flow from $W(c_i, \theta)$ we hit δ in a geodesic segment δ_i , so that every geodesic intersecting $W(c_i, \theta)$ also intersects $W(\delta_i, \theta + \varepsilon')$ for some ε' . Since ε' goes to 0 as ε goes to 0, the result follows. □

Observe that if ε is small enough and the c_i in Proposition 8.3 are disjoint, then the δ_i will be disjoint as well. Thus, combining the above propositions (and redefining θ to be $\theta + \varepsilon$), we have found a closed geodesic δ and disjoint geodesic segments $\delta'_i \subset \delta$ so that we have the following global cross sections:

- A wedge set $\tau_0 := W(\{\delta'_i\}, \theta)$ giving a disconnected embedded global cross section.
- A global cross section $W(\delta, \theta)$ containing the previous one which is connected but not embedded (as δ will self-intersect).

We would like to use the second global cross section $W(\delta, \theta)$ to close up the homotopy return trajectories of the smeared return map. However, $W(\delta, \theta)$ is not simply connected, so the homotopy return map will depend on which path along the cross section we choose. This can be easily fixed by setting, for some small open interval $I \subset \delta \setminus \bigcup_i \{\delta_i\}$,

$$\tau' := W(\delta - I, \theta)$$

so that $\tau_0 \subset \tau'$. Strictly speaking, τ' is not simply connected as a subset of $UT\Sigma$; rather, it is the image of an immersed disk. Since $\tau_0 \subset \tau'$ lies in a portion where the immersion is injective, there is no ambiguity about how to connect up the return paths to τ_0 within τ' .

Definition 8.4. A good cross section is the data of cross sections τ_0, τ, τ' and bump function ψ , where τ is a slight enlargement of the embedded cross section τ_0 so that $\tau_0 \Subset \tau \subset \tau'$, for τ' a cross section as above. For simplicity, we will refer to a good cross section just as τ .

A good cross section gives the complete setup of Subsection 7.4.

9. Join lemma

We now turn to the heart of the proof, proving *join lemmas* to show that we can smooth essential crossings to relate the return maps of order k , order ℓ and order $k + \ell$. We chose the global cross sections $\tau_0 \Subset \tau \subset \tau'$ in Section 8 to be wedge sets in order to connect to hyperbolic geometry and prove the necessary crossings are essential. Recall that we refer to the data of the nested cross sections from wedge sets (including the bump function ψ , when relevant) as a good cross section (Definition 8.4), which we refer to as τ .

Lemma 9.1 (Classical join lemma). *Let τ be a good cross section. There is a curve K_τ and integer w_τ so that for large enough $k, \ell \geq 0$, we have, for all $\vec{x} \in \tau$,*

- (a) $[m^k(\vec{x})] \cup [m^\ell(p^k(\vec{x}))] \cup K_\tau \searrow_{w_\tau} [m^{k+\ell}(\vec{x})]$
- (b) $[m^{k+\ell}(\vec{x})] \cup K_\tau \searrow_{w_\tau} [m^k(\vec{x})] \cup [m^\ell(p^k(\vec{x}))]$.

As a corollary, we will prove a corresponding join lemma for the smeared return map.

Lemma 9.2 (Smeared join lemma). *Let τ be a good cross section. There is a curve K_τ and weight w_τ so that for large enough $k, \ell \geq 0$, we have, for all $\vec{x} \in \tau$,*

- (a) $[M^k(\vec{x})] \cup [M^\ell(P^k(\vec{x}))] \cup K_\tau \searrow_{w_\tau} [M^{k+\ell}(x)]$
- (b) $[M^{k+\ell}(\vec{x})] \cup K_\tau \searrow_{w_\tau} [M^k(\vec{x})] \cup [M^\ell(P^k(\vec{x}))]$.

Example 9.3. As an example of smeared first return map and to illustrate how the join lemma is applied, consider the case when the geodesic current μ is δ_γ for γ a closed curve whose lift to the unit tangent bundle intersects the global cross section τ at two points \vec{x}_0 and \vec{x}_1 . We assume further that $\vec{x}_0 \notin \tau_0$, $\psi(\vec{x}_0) = t \in (0, 1)$, and $\vec{x}_1 \in \tau_0$. Then, as illustrated in Figure 9.1, $[M^1(\vec{x}_0)]$ is a curve $C_{0,1}$ with weight 1, since $p_\tau(\vec{x}_0) = \vec{x}_1 \in \tau_0$. On the other hand, $[M^1(\vec{x}_1)]$ consists of a weighted multi-curve with two components $C_{1,2}$ and $C_{1,3}$ starting from x_1 and landing at $\vec{x}_2 = \vec{x}_0$ and $\vec{x}_3 = \vec{x}_1$, with weights t and $1 - t$, respectively. Then (with $\mu = \delta_\gamma$) we have

$$\begin{aligned} \psi \mu_\tau &= t \delta_{\vec{x}_0} + \delta_{\vec{x}_1} \\ R^1(\mu) &= t C_{0,1} + t C_{1,2} + (1 - t) C_{1,3}. \end{aligned}$$

Now, the join lemma asserts that we can to join the curves $C_{0,1}$ and $C_{1,2}$, together with an extra curve K , to get $C_{0,2}$. Assuming all of the relevant intersections are essential, we can do it with K being two copies of δ , one oriented in each direction, in the steps shown in Figure 9.2.

In the full proof, to guarantee the analogous intersections are essential we will add more copies of δ .

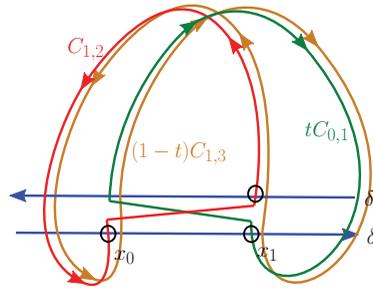


Figure 9.1. Example of a first iteration of the smeared first return map; that is, $[M^1]$ on a geodesic current corresponding to a closed curve intersecting the cross section τ twice at points \vec{x}_0, \vec{x}_1 and the cross section τ_0 once at point \vec{x}_1 . The weight of the bump function ψ at \vec{x}_0 is t . We obtain three weighted curves. $C_{0,1}$ consisting of the geodesic trajectory that goes from \vec{x}_0 to x_1 and closes off by following the cross section in some coherent way. $C_{0,1}$ has weight 1 by definition of smeared return map, since $\vec{x}_1 \in \tau_0$. $C_{1,2}$ has weight t , whereas $C_{1,3}$ has weight $1 - t$.

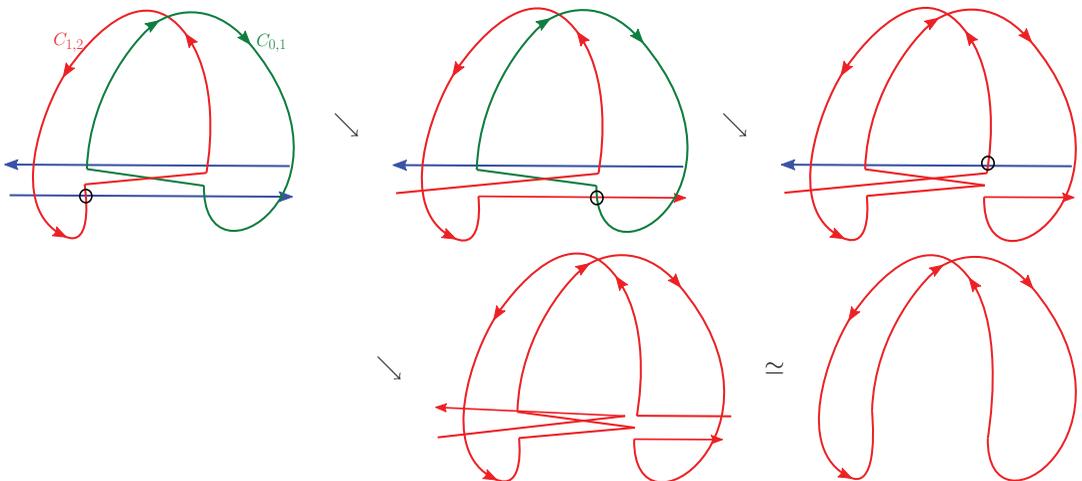


Figure 9.2. Applying the join lemma in Example 9.3. At each step, we smooth at the circled crossing.

Proof of Lemma 9.1. First we look at the geometry of a return trajectory $m^k(\vec{x})$ when k is large, as a concrete curve either on Σ or lifted to the universal cover. Since there is a lower bound on the first return time (Lemma 7.3), the n th return time grows at least linearly in k . We may therefore assume that the portion of $m^k(\vec{x})$ that follows $\phi_t(\vec{x})$ is very long. The lift of $m^k(\vec{x})$ to the universal cover is thus a *broken path*: for some large L , it alternates between

- (a) long segments of length at least L following $\phi_t(\vec{x})$ and
- (b) short segments of some length following δ ,

with turns between them that are within ε of a right angle, alternating left and right. (See Definition 12.1 for a precise definition.)

We study the geometry of broken paths in Section 12. In particular, we prove several lemmas there guaranteeing that broken paths intersect essentially in certain circumstances. If L is large enough, we have the following results:

- A broken path and a lift of δ intersect essentially (Lemma 12.2).
- Broken paths with short segments that are different enough in length intersect essentially (Lemma 12.3).

We first prove part (a) in the lemma statement. We will use the following steps. By convention, δ is oriented to the right and δ^{-1} is the same curve oriented to the left. Let α , β and γ be $m^k(\vec{x})$ and $m^\ell(p^k(\vec{x}))$, respectively.

- (1) We start by smoothing $[\alpha]$ with a large number N of copies of $[\delta]$. Each one of these intersections is essential by Lemma 12.2. This yields a new curve $[\alpha_1]$ with lift a broken path with an lengthened short segment.
- (2) We then smooth $[\alpha_1]$ against $[\beta]$. The corresponding lifts are broken paths with short segments of different enough lengths, so Lemma 12.3 guarantees that the crossing is essential, yielding a new curve $[\gamma_1]$.
- (3) Finally, we smooth $[\gamma_1]$ against $N[\delta^{-1}]$. This returns to the correct homotopy class, again using Lemma 12.2 to guarantee that the crossings are essential. The result is $[\gamma]$, as desired.

We need to use a large enough number N of copies of δ that guarantees that the crossing in the second step above is essential. Let ε be the angle of the wedge set, let ℓ be the length of δ and let $\kappa(\varepsilon)$ be the constant from Lemma 12.3. Then we claim that it suffices to take $N = 2M$ with $M = 1 + \lceil \kappa(\varepsilon)/\ell \rceil$, so that overall constants in the statement are

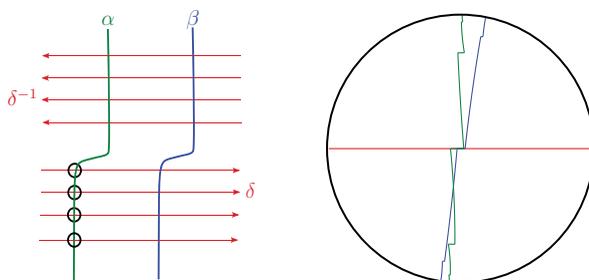
$$K_\tau = N([\delta] + [\delta^{-1}])$$

$$w_\tau = 2N + 1 = 3 + 4\lceil \kappa(\varepsilon)/\ell \rceil.$$

In order to be explicit about how to apply Lemmas 12.2 and 12.3, we will work with concrete lifts of our curves to broken paths in the universal cover; to pick out a lift, we work with elements of π_1 and so pick a basepoint. For concreteness, choose the basepoint $*$ to be at the far left end of the segment on δ defining τ' . (Recall that we removed a short interval to make τ' simply connected.) We are particularly interested in the short segments on the lift of δ ; for that purpose, parametrise the lift of δ by length in \mathbb{R} , with 0 at the lift of the basepoint $*$ and δ oriented in the positive direction so that δ itself lifts to a curve ending at ℓ .

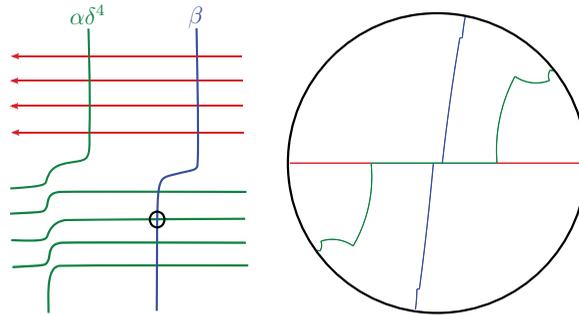
Now we state precisely the sequence of smoothings that we will perform, illustrating them with slightly schematic figures of both the curves on the surface and of the corresponding broken paths realising the lifts in the universal cover. At each step we circle the crossings that we smooth at the next step.

- (i) Let δ , α and β be elements of $\pi_1(S)$ representing the transversal curve, $m^k(\vec{x})$ and $m^\ell(p^k(\vec{x}))$, respectively:



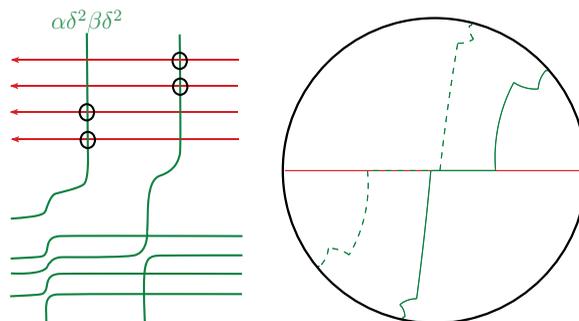
The endpoints of the central short segments of lifts of α and β are both in $(0, \ell)$, in the parametrisation above.

- (ii) Smooth $[\alpha]$ with $N[\delta]$ a total of N times to get $[\alpha_1]$ with $\alpha_1 = \delta^M \alpha \delta^M$. The crossings are essential by Lemma 12.2.



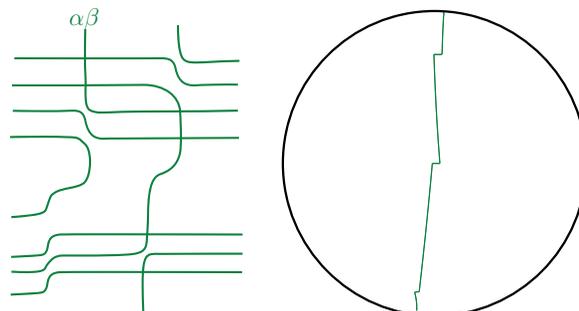
The endpoints of the central short segment of the lift of α_1 are in $(-M\ell, -(M - 1)\ell)$ and in $(M\ell, (M + 1)\ell)$.

- (iii) Smooth $[\alpha \delta^N] \cup [\beta]$ at a middle crossing to make $[\gamma_1] = [\alpha \delta^M \beta \delta^M]$. Since $(M - 1)\ell \geq \kappa(\varepsilon)$, the crossing is essential by Lemma 12.3.

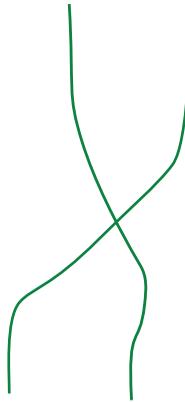


Here, in the picture in the universal cover, two different lifts of γ_1 are shown (one dashed), to make it clearer what happened in the smoothing; these are the lifts of $\delta^M \alpha \delta^M \beta$ (solid) and $\beta \delta^M \alpha \delta^M$ (dashed).

- (iv) Smooth $[\gamma_1]$ with $N[\delta^{-1}]$ a total of N times at appropriate crossings to make $[\alpha \beta] = m^{k+\ell}(\bar{x})$. The crossings are essential by Lemma 12.2.



(v) The result is $[\alpha\beta]$ as desired.



This completes the proof of part (a) of the statement. Part (b) is very similar. Precisely, we do the following steps.

- (i') Let $\delta = \tau$, $\alpha = m^k(\vec{x})$ and $\beta = m^\ell(p^k(\vec{x}))$ be as before, so that we start with $[\alpha\beta] = [m^{k+\ell}(\vec{x})]$.
- (ii') Use Lemma 12.2 to smooth with $N[\delta]$ a total of N times to get $[\gamma_1]$ with $\gamma_1 = \delta^M \alpha \beta \delta^M$. If we set $\gamma_2 = \beta \delta^N \alpha$, then $[\gamma_1] = [\gamma_2]$, but these two curves have different canonical lifts to the universal cover: the endpoints of the primary short segment γ_1 are in $(-M\ell, -(M-1)\ell)$ and $(M\ell, (M+1)\ell)$ on the lift of δ , while the endpoints of the (zero-length) primary ‘short segment’ of γ_2 are at the same point in $(0, \ell)$.
- (iii') Smooth $[\gamma_1]$ with itself to make $[\alpha_1] \cup [\beta_1]$, with $\alpha_1 = \delta^M \alpha$ and $\beta_1 = \beta \delta^M$. The crossing corresponds to the lifts given by γ_1 and γ_2 and is essential by Lemma 12.3.
- (iv') Smooth $[\alpha_1]$ and $[\beta_1]$ each M times with $M[\delta^{-1}]$ to make $[\alpha]$ and $[\beta]$, respectively, using Lemma 12.2.
- (v') The result is $[\alpha] \cup [\beta]$ as desired. □

Proof of Lemma 9.2. By definition of M^k and P^k , we have nonnegative constants K, a_i, L and $b_{i,j}$ so that

$$\begin{aligned}
 M^k(\vec{x}) &= \sum_{i=k}^K a_i m^i(\vec{x}) & P^k(\vec{x}) &= \sum_{i=k}^K a_i \vec{y}_i \\
 M^\ell(\vec{y}_i) &= \sum_{j=\ell}^L b_{i,j} m^j(\vec{y}_i) & M^{k+\ell}(\vec{x}) &= \sum_{i=k}^K \sum_{j=\ell}^L a_i b_{i,j} m^i(\vec{x}) m^j(\vec{y}_i).
 \end{aligned}$$

Furthermore, $\sum_i a_i = 1$ and, for fixed i , $\sum_j b_{i,j} = 1$. The result follows by distributing and applying Lemma 9.1 repeatedly. □

As an immediate consequence, we have the following.

Proposition 9.4. *For a fixed good cross section τ as constructed above and for every curve functional f satisfying quasi-smoothing and convex union, there is a constant $\kappa(\tau)$ so that, for sufficiently large k, ℓ and every geodesic current μ , we have*

$$f_\tau^{k+\ell}(\mu) \leq f_\tau^k(\mu) + f_\tau^\ell(\mu) + \kappa(\tau)\mu(\psi\tau).$$

Proof. We will prove this with $\kappa(\tau) = f(K_\tau) + R w_\tau$, where K_τ and w_τ are from Lemma 9.2 and R is the quasi-smoothing constant from equation (1.2).

We have

$$\begin{aligned}
 f_\tau^{k+\ell}(\mu) &= f\left(\int_\tau M^{k+\ell}(\vec{x})\psi(\vec{x})\mu(\vec{x})\right) \\
 &\leq f\left(\int_\tau (M^k(\vec{x}) + M^\ell(P^k(\vec{x})) + K_\tau)\psi(x)\mu(x)\right) + \int_\tau R_{w_\tau}\psi(\vec{x})\mu(\vec{x}) \\
 &\leq f\left(\int_\tau M^k(x)\psi(\vec{x})\mu(\vec{x})\right) + f\left(\int_\tau M^\ell(P^k(\vec{x}))\psi(\vec{x})\mu(\vec{x})\right) + \kappa(\tau) \int_\tau \psi(\vec{x})\mu(\vec{x}) \\
 &= f^k(\mu) + f\left(\int_\tau M^\ell(\vec{x})P_*^k(\psi(\vec{x})\mu(\vec{x}))\right) + \kappa(\tau)\mu(\psi\tau) \\
 &= f^k(\mu) + f^\ell(\mu) + \kappa(\tau)\mu(\psi\tau),
 \end{aligned}$$

where we use, successively:

- o the definition of $f^{k+\ell}$;
- o Lemma 9.2 and the quasi-smoothing property of f ;
- o the convex union property of f and the definition of $\kappa(\tau)$;
- o change of variables and the definitions of f^k and $\mu(\psi\tau)$; and
- o Proposition 7.16 and the definition of f^ℓ . □

We recall a slight variation of Fekete’s lemma, which follows from standard versions, such as de Bruijn and Erdős [17, Theorem 22].

Lemma 9.5 (Fekete’s lemma). *Let $(a_n)_{n=1}^\infty$ be a sequence of real numbers and suppose there exists N such that for all $m, n \geq N$, $a_{n+m} \leq a_n + a_m$. Then*

$$\lim_{n \rightarrow \infty} \frac{a_n}{n} \text{ exists and is equal to } \inf_{n \geq N} \frac{a_n}{n}.$$

Finally, we can show that the limit defining the extension of f exists.

Proposition 9.6. *For any curve functional f satisfying quasi-smoothing and convex union, the limit defining f_τ in equation (7.27) exists.*

Proof. Use Proposition 9.4 and apply Lemma 9.5 to the sequence $f_\tau^k(\mu) + \kappa(\tau)\mu(\psi\tau)$. □

10. Continuity of the extension

In order to prove continuity of the extension, we will prove continuity of f_τ^k and then get upper and lower bounds on the limit $f_\tau(\mu)$ in terms of $f_\tau^k(\mu)$. Proposition 9.4 lets us use Fekete’s lemma to get upper bounds. To get lower bounds, we have the following.

Proposition 10.1. *For a fixed good cross section τ and any weighted curve functional f satisfying homogeneity, weighted quasi-smoothing and convex union, there is a constant $K(\mu) = \kappa(\tau)\mu(\psi\tau)$ so that for all sufficiently large k and every geodesic current μ we have*

$$2f_\tau^k(\mu) \leq f_\tau^{2k}(\mu) + K(\mu).$$

Proof. By Lemma 9.2(b) in the case $k = \ell$,

$$M^{2k}(\vec{x}) \cup_{K_\tau} \searrow_{w_\tau} M^k(\vec{x}) \cup M^k(P^k(\vec{x})).$$

Integrating this statement with respect to the measure $\psi\mu$ (which is invariant under P^k), we find that

$$M^{2k}(\psi\mu) \cup_{K_\tau \cdot \mu(\psi\tau)} \searrow_{w_\tau \cdot \mu(\psi\tau)} 2M^k(\psi\mu).$$

Applying f to both sides and using homogeneity of f gives the desired result. □

Since f is not in general additive, by comparison to Proposition 9.4, Proposition 10.1 is more restrictive, requiring $k = \ell$. This still suffices to show that the f_τ^k approximate f_τ well.

Lemma 10.2. *Let τ be fixed good cross section and f be a weighted curve functional satisfying homogeneity, weighted quasi-smoothing and convex union. For any sufficiently large k ,*

$$|f_\tau(\mu) - \frac{f_\tau^k(\mu)}{k}| \leq \frac{K(\mu)}{k}$$

where $K(\mu) = \kappa(\tau)\mu(\psi\tau)$ is the constant from Propositions 9.4 and 10.1.

Proof. From Propositions 9.4 and 10.1, for large enough k we have

$$\left| \frac{f_\tau^{2k}(\mu)}{2k} - \frac{f_\tau^k(\mu)}{k} \right| \leq \frac{K(\mu)}{2k}.$$

We also have

$$\begin{aligned} \frac{f_\tau^k(\mu)}{k} + \left(\frac{f_\tau^{2k}(\mu)}{2k} - \frac{f_\tau^k(\mu)}{k} \right) + \left(\frac{f_\tau^{4k}(\mu)}{4k} - \frac{f_\tau^{2k}(\mu)}{2k} \right) + \left(\frac{f_\tau^{8k}(\mu)}{8k} - \frac{f_\tau^{4k}(\mu)}{4k} \right) + \dots \\ = \lim_{n \rightarrow \infty} \frac{f_\tau^{2^n k}(\mu)}{2^n k} = \lim_{n \rightarrow \infty} \frac{f_\tau^{n k}(\mu)}{n k} = f_\tau(\mu), \end{aligned}$$

where the first equality follows by telescoping and the second one because we have already proved that the limit exists. We can then give bounds:

$$\begin{aligned} |f_\tau(\mu) - \frac{f_\tau^k(\mu)}{k}| &\leq \left| \frac{f_\tau^{2k}(\mu)}{2k} - \frac{f_\tau^k(\mu)}{k} \right| + \left| \frac{f_\tau^{4k}(\mu)}{4k} - \frac{f_\tau^{2k}(\mu)}{2k} \right| + \left| \frac{f_\tau^{8k}(\mu)}{8k} - \frac{f_\tau^{4k}(\mu)}{4k} \right| + \dots \\ &\leq \frac{K(\mu)}{2k} + \frac{K(\mu)}{4k} + \frac{K(\mu)}{8k} + \dots \\ &= \frac{K(\mu)}{k}. \end{aligned}$$

□

We next prove that the f_τ^k are continuous.

Proposition 10.3. *Let τ be fixed good cross section and f be a weighted curve functional satisfying homogeneity, weighted quasi-smoothing and convex union. Then the functions $f_\tau^k : \mathcal{GC}^+(S) \rightarrow \mathbb{R}$ are continuous for every k .*

We will break the proof into lemmas.

Lemma 10.4. *For τ a (closed) global cross section with interior τ° , the map $\mu \mapsto \mu_{\tau^\circ}$ from $\mathcal{GC}^+(S)$ to $\mathcal{M}(\tau^\circ)$ is continuous.*

Proof. We first adjust the definition of the flux μ_{τ° . Let ε be small enough so that the corresponding flow box is embedded. Pick a nonzero continuous function $\omega : [0, \varepsilon] \rightarrow \mathbb{R}_{\geq 0}$ so that $\omega(0) = \omega(\varepsilon) = 0$ and $\int_0^\varepsilon \omega(t) dt = 1$. Then, for any measurable function r on τ° , the flux μ_{τ° satisfies

$$\int_{\vec{x} \in \tau^\circ} r(\vec{x}) \mu_{\tau^\circ}(\vec{x}) = \int_{t=0}^\varepsilon \int_{\vec{x} \in \tau^\circ} \omega(t) r(\vec{x}) \mu(\phi_t(\vec{x})). \tag{10.5}$$

Now suppose that we have a sequence of measures μ_i approaching μ in the weak* topology, and let r be a continuous function on τ° with compact support. By Theorem 2.19, it suffices to show that $\int_{\vec{x} \in \tau^\circ} r(\vec{x}) \mu_i(\tau^\circ)(\vec{x})$ converges to $\int_{\vec{x} \in \tau^\circ} r(\vec{x}) \mu_{\tau^\circ}(\vec{x})$.

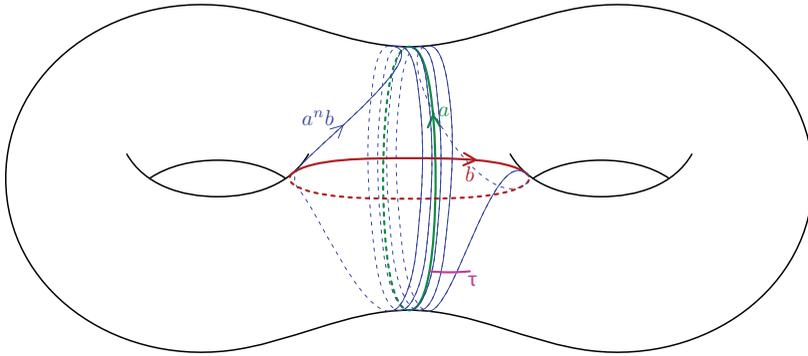


Figure 10.1. An example showing that the flux map $\mu \mapsto \mu_\tau$ is not continuous. The sequence of curves $[a^n b]/n$ approaches $[a]$, but they have very different intersections with τ .

Consider the function s on $UT\Sigma$ defined by

$$s(\vec{x}) = \begin{cases} r(\vec{x})\omega(t) & \text{if } y = \phi_t(\vec{x}) \text{ for } \vec{x} \in \tau, t \in [0, \varepsilon] \\ 0 & \text{otherwise.} \end{cases}$$

Then s is continuous, since r and ω vanish on the boundaries of their domains of definition, so $\int_{\vec{y} \in UT\Sigma} s(\vec{x})\mu_i(\vec{y})$ converges to $\int_{\vec{y} \in UT\Sigma} s(\vec{y})\mu(\vec{y})$. The result follows from equation (10.5). \square

Remark 10.6. The map $\mu \mapsto \mu_\tau$ from $\mathcal{GC}^+(S)$ to positive measures $\mathcal{M}(\tau)$ on the closed cross section is not continuous with respect to the weak* topology at points where $\mu_\tau(\partial\tau) \neq 0$. Indeed, let μ be the geodesic current corresponding to a closed curve $[a]$, let $[b]$ be another closed curve intersecting $[a]$ and let μ_n be the geodesic current corresponding to $\frac{1}{n}[a^n b]$ so that $\lim_{n \rightarrow \infty} \mu_n = \mu$, as shown in Figure 10.1. Take a (noncomplete) transversal τ that intersects $\text{supp}(\mu)$ only once on $\partial\tau$. Then (for appropriate choices, as shown) the total mass of μ_n – that is, $(\mu_n)_\tau(\tau)$ – is approximately 1/2, while $\mu_\tau(\tau) = 1$.

Similarly, for the open transversals τ° , we have $\mu_{\tau^\circ}(\tau^\circ) = 0$ while $(\mu_n)_{\tau^\circ}(\tau^\circ)$ is approximately 1/2 for large n . This does not contradict Lemma 10.4; it just says that total mass is not a continuous function in the weak* topology on a noncompact space.

Lemma 10.7. The extension of $[M^k]$ to measures, as a map from $\mathcal{M}(\tau)$ to $\mathbb{R}\mathcal{C}(S)$, is continuous.

Proof. Recall from Lemma 7.23 that $[M^k(\mu)]$ takes values in the finite-dimensional subspace $\mathbb{R}[\Lambda(k, \tau, \tau_0)] \subset \mathbb{R}\mathcal{C}(S)$. By continuity of $[M^k(\vec{x})]$ (Lemma 7.24), we can write

$$[M^k(\vec{x})] = \sum_{C \in [\Lambda(k, \tau, \tau_0)]} a_C(\vec{x}) \cdot C$$

where a_C is a continuous function on τ . (Recall that a function to a finite-dimensional vector space is continuous iff each of the coordinate functions is continuous; see Remark 2.25.) But then

$$[M^k(\mu)] = \sum_{C \in [\Lambda(k, \tau, \tau_0)]} \left(\int_{\vec{x} \in \tau} a_C(\vec{x}) \mu(\vec{x}) \right) \cdot C.$$

The integrals are continuous functions of μ by definition of the weak* topology on $\mathcal{M}(\tau)$. \square

Proof of Proposition 10.3. f_τ^k is the composition of maps

$$\mathcal{GC}^+(S) \xrightarrow{\mu \mapsto \mu_{\tau^\circ}} \mathcal{M}(\tau^\circ) \xrightarrow{\cdot\psi} \mathcal{M}(\tau) \xrightarrow{[M^k]} \mathbb{R}[\Lambda(k, \tau, \tau_0)] \xrightarrow{f} \mathbb{R}.$$

The component maps are continuous by, respectively, Lemma 10.4, the fact that ψ vanishes on a neighbourhood of $\partial\tau$, Lemma 10.7 and Proposition 3.4. □

Proposition 10.8. *Let τ be fixed good cross section and f be a curve functional satisfying homogeneity, weighted quasi-smoothing and convex union. Then $f_\tau : \mathcal{GC}^+(S) \rightarrow \mathbb{R}_{\geq 0}$ is a continuous function.*

Proof. By Proposition 10.3, it suffices to show that f_τ is a uniform limit of f_τ^k . The constant $K(\mu)$ in Lemma 10.2 does depend on μ ; however, if we bound μ within a ball so that $\int_\tau \psi \mu_\tau$ is bounded, the constant in the approximation becomes uniform and tends to 0 as $k \rightarrow \infty$. □

11. The extension extends

In this section we prove that when restricted to weighted curves, the purported extension f_τ coincides with the original curve functional f . More precisely, let $\gamma \in C$ be the geodesic representative of an oriented closed curve with corresponding geodesic current μ_C , and let τ be a good cross section of the geodesic flow (Definition 8.4). We wish to show that $f_\tau(\mu_C) = f(C)$.

Let $\tilde{\gamma}$ be the canonical lift of γ to the unit tangent bundle, and let n be the number of times that $\tilde{\gamma}$ intersects τ , with intersections at $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_{n-1}$ in order (so $p(\vec{x}_i) = \vec{x}_{i+1}$). Then $(\mu_C)_\tau = \sum_{i=0}^{n-1} \vec{x}_i$. Let $a_i = \psi(\vec{x}_i)$, so that

$$C \cap \psi\tau := \psi \cdot (\mu_C)_\tau = \sum_{i=0}^{n-1} a_i \vec{x}_i.$$

By Proposition 7.16, this sum (which we call $C \cap \psi\tau$, in an abuse of notation) is invariant under the smeared return map P :

$$P^k \left(\sum_{i=0}^{n-1} a_i \vec{x}_i \right) = \sum_{i=0}^{n-1} a_i \vec{x}_i. \tag{11.1}$$

We need a slightly stronger fact. Recall that each term in $[M^k(C \cap \psi\tau)]$ is a curve that follows the geodesic trajectory $\vec{x}_i \rightarrow \vec{x}_{i+1} \rightarrow \dots$ for some time and then travels along τ to close up. We say that a segment of the return map $\vec{x}_i \rightarrow \vec{x}_{i+1}$ is *covered* with degree r in $[M^k]$ if the weighted number of times that segment appears in $[M^k(C \cap \psi\tau)]$ is r .

Lemma 11.2. *For any closed curve C and a good cross section τ with bump function ψ as above, in $[M^k(C \cap \psi\tau)]$, every segment $\vec{x}_i \rightarrow \vec{x}_{i+1}$ is covered with degree k .*

(See Example 9.3 for one concrete case.)

Proof. Fix $a_i = \psi(\vec{x}_i)$ as above and consider the case $k = 1$. By the assumption that τ_0 is a complete cross section, we have $a_i = 1$ for some i . By rotating the indices, assume $a_0 = 1$. We prove the statement for each segment $\vec{x}_i \rightarrow \vec{x}_{i+1}$ by induction on i . For $i = 0$, it is clear, since $a_0 = 1$ and no earlier trajectories continue through \vec{x}_0 . For $i > 0$, we have a_i trajectories starting at \vec{x}_i and going to \vec{x}_{i+1} . By the induction hypothesis, we also have weight 1 of trajectories arriving at \vec{x}_i from \vec{x}_{i-1} and so a weight of $1 - a_i$ for those continuing on to \vec{x}_{i+1} . These two types of trajectories have a total weight of 1, as desired.

The statement for $k > 1$ follows from equation (11.1) and induction. □

Now, for $i < j$, let C_{ij} be the curve that starts at \vec{x}_i , passes through $j - i - 1$ intermediate points to \vec{x}_j and closes up along τ , with indices interpreted modulo n . Then, for some coefficients w_{ij} , we can write

$$\left[M^k \left(\sum a_i \vec{x}_i \right) \right] = \sum w_{ij} C_{ij}.$$

The nonzero coefficients w_{ij} that appear will have $k \leq j - i \leq kn$, so as k gets large the C_{ij} that appear in the weighted sum also get long.

The invariance from equation (11.1) tells us that for all i_0 ,

$$\sum_{i \equiv i_0} w_{ij} = \sum_{j \equiv i_0} w_{ij} = a_{i_0} \tag{11.3}$$

while the fact that all n steps $x_i \rightarrow x_{i+1}$ are covered with degree k implies that

$$\sum_{i,j} w_{ij} = kn. \tag{11.4}$$

Proposition 11.5. *If μ_C is the geodesic current associated to a weighted closed multi-curve C and τ is a good cross section of the geodesic flow, then $f_\tau(\mu_C) = f(C)$.*

Proof. We first suppose C is a single curve with weight 1.

As above, let $\sum_{i=0}^{n-1} a_i \vec{x}_i = C \cap \psi\tau$. For the k th iterate set

$$C_0^k = \left[M^k \left(\sum a_i \vec{x}_i \right) \right] = \sum_{i < j} w_{ij} C_{ij}.$$

Note that $C_{i,i+rn} = C^r$; other C_{ij} have a more complicated relation to C . For k sufficiently large, we will use Lemma 9.2 to simplify the sum so that only curves of the form $C_{i,i+rn}$ appear.

For each $i = 0, \dots, n-1$ (in any order), consider all of the curves that either start or end at \vec{x}_i , starting with $i = i_0$. By equation (11.3),

$$\sum_{\substack{i \equiv i_0 \\ j \neq i_0}} w_{ij} = \sum_{\substack{i \neq i_0 \\ j \equiv i_0}} w_{ij} \leq a_{i_0}.$$

We can therefore pair the corresponding components of C_0^k against each other using Lemma 9.2 pairwise in any order, getting a reduction

$$C_0^k \cup a_{i_0} K \searrow_{a_{i_0} w} C_1^k$$

where K and w are the curve and weight from Lemma 9.2, and C_1^k is another weighted combination of the C_{ij} in which each component that starts at i_0 also ends at i_0 .

This join operation does not change the degree by which segments of the curves are covered, so equation (11.4) still holds, and equation (11.3) still holds at the other indices. So we can repeat this at each index. In the end we get a reduction

$$C_0^k \cup aK \searrow_{aw} \sum_j b_j C^j =: C_n^k$$

where $a = \sum_i a_i$ and C_n^k is another weighted curve. By considering the degrees, we see that $\sum_j b_j = k$.

Similar considerations (similar to part (b) of Lemma 9.2) show that

$$C_n^k \cup aK \searrow_{aw} C_0^k.$$

By Corollary 13.5, f satisfies strong stability. The homogeneity and strong stability properties then yield

$$kf(C) = f(kC) = f\left(\sum_j b_j C^j\right) = f(C_n^k).$$

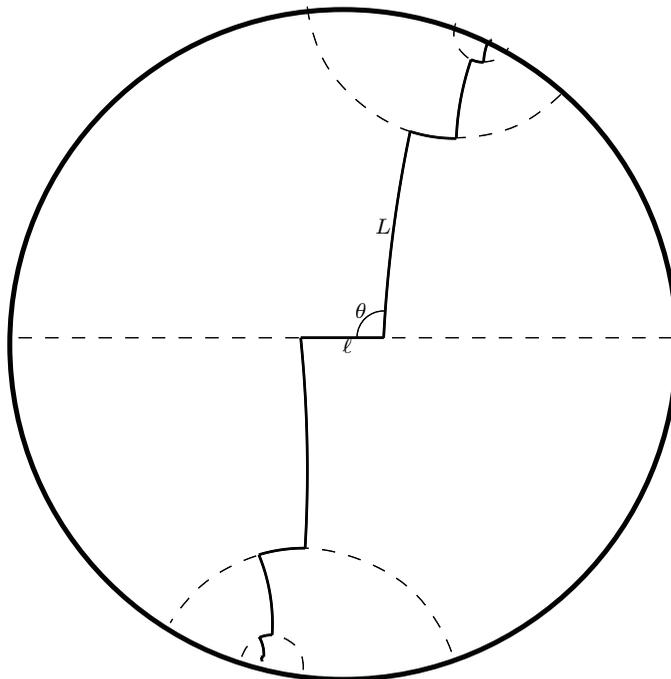


Figure 12.1. A broken path in the disk model. Here $|\pi/2 - \theta| < \varepsilon$.

Therefore, since aK and aw are independent of k ,

$$f_\tau(\mu) = \lim_{k \rightarrow \infty} \frac{f(C_0^k)}{k} = \lim_{k \rightarrow \infty} \frac{f(C_n^k)}{k} = \lim_{k \rightarrow \infty} \frac{f(kC)}{k} = f(C).$$

We have thus proved that f_τ extends f on unweighted curves.

For the case of a general weighted curve $C = \sum w_\ell C_\ell$, the proof proceeds as above, except that we start with the weighted intersection of C with the smeared cross section. More precisely, let $\vec{x}_{\ell,i}$ be the intersections of C_ℓ with τ ; then we work with $\sum_{\ell,i} w_\ell \psi(x_{\ell,i}) \cdot \vec{x}_{\ell,i}$, in the same way as above. \square

12. Hyperbolic geometry estimates

We complete the proof of Theorem A by proving facts about the geometry of broken paths, as used in Section 9.

Definition 12.1. Fix a real length L and angle $\varepsilon < \pi/2$. A broken path $b(L, \varepsilon)$ is a concatenation of geodesic segments in \mathbb{H}^2 that alternate between ‘long’ segments of length at least ℓ and ‘short’ segments of unconstrained length, so that the angle between the long and short segments is within ε of $\pi/2$, alternately turning left and right. See Figure 12.1 for an example. We will denote by a_i the hyperbolic line containing the i th short segment.

We prove some basic facts about when broken paths cross.

Lemma 12.2. For any $0 < \varepsilon < \pi/2$, there is a constant $L_0(\varepsilon)$ so that, for any $L > L_0(\varepsilon)$, any broken path $b(L, \varepsilon)$ converges to unique points at infinity that are on opposite sides of the hyperbolic line containing any short segment. As ε approaches 0, the constant $L_0(\varepsilon)$ approaches 0 as well.

That is, in Figure 12.1, the broken path crosses the dashed paths.

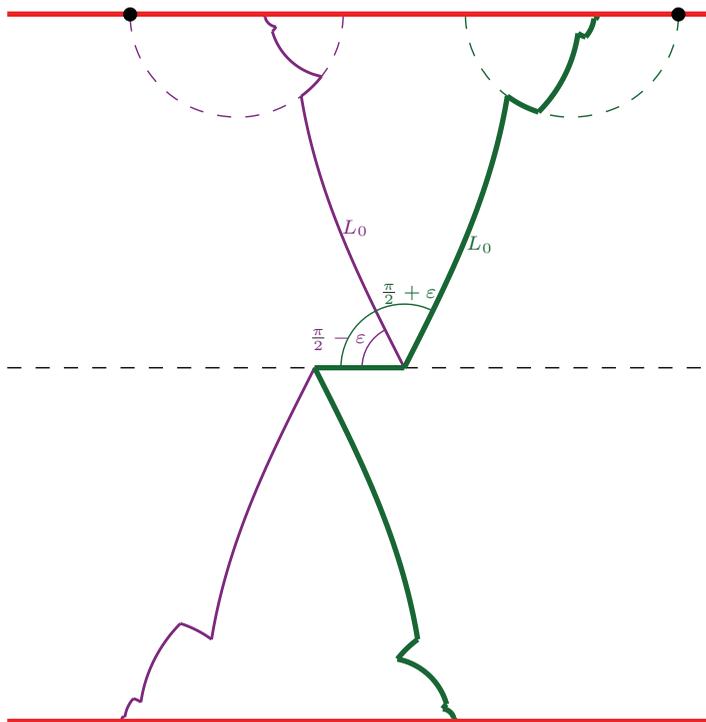


Figure 12.3. The window of possible endpoints of a broken path. The marked points are bounds on ends of broken paths $b(L, \varepsilon)$ with $L \geq L_0$.

Similar arguments apply to the other endpoints of γ and γ' , implying that for large enough $\kappa(\varepsilon)$ the paths cross essentially. □

13. Stable functions

Some curve functionals satisfy quasi-smoothing and convex union but are not stable or homogeneous on the nose. For example, the length of a curve with respect to an arbitrary generating set is of this form (Example 4.10). We fix this by passing to a stable length as in Theorem B. Recall that the *stable curve functional* $\|f\|$ is defined by

$$\|f\|(C) := \lim_{n \rightarrow \infty} \frac{f(C^n)}{n}.$$

As in the proof of Theorem 6.1, we will consider weighted curve functionals. In this section we will prove the following theorem.

Theorem 13.1. *Let f be a weighted curve functional satisfying weighted quasi-smoothing and convex union. Then the stabilised curve functional*

$$\|f\|(C) := \lim_{n \rightarrow \infty} \frac{f(C^n)}{n}$$

satisfies weighted quasi-smoothing, convex union, strong stability and homogeneity and thus extends to a continuous function on $\mathcal{QC}^+(S)$.

We first prove some lemmas.

Lemma 13.2. For any connected curve C and any sufficiently large $n, m \geq 0$, there exists some curve K and weight w so that $C^n \cup C^m \cup K \searrow_w C^{n+m}$.

Proof. Let p be the number of intersections of our cross section τ with the canonical lift of C to $UT\Sigma$ and apply Lemma 9.1(b), taking $k = np$ and $l = mp$. (We reuse the same cross section τ for convenience; nothing here depends on the definition of the extension.) Then for sufficiently large n, m ,

$$C^{n+m} \cup K_\tau \searrow_{w_\tau} C^m \cup C^n,$$

proving the lemma with $K = K_\tau$ and $w = w_\tau$. □

Lemma 13.3. For any connected curve C , we have $C^n \searrow_{n-1} nC$.

Proof. Self-crossings of an n -fold cover of a curve are essential by definition. □

Lemma 13.4. There is a curve K and a constant w so that, for any curve C on S and any $n \geq 2$, we have

$$nC \cup (n - 1)K \searrow_{(n-1)w} C^n.$$

Proof. Pick $0 < \varepsilon < \pi/2$ so that $L_0(\varepsilon)$ from Lemma 12.2 is less than the systole of Σ , the length of the shortest closed geodesic on Σ . As in Section 8, find a curve K and a complete global cross section $\tau \subset W(K, \varepsilon)$. Then, by the arguments of Lemma 9.1(a), there is some integer w so that

$$2C \cup K \searrow_w C^2.$$

(We are not directly applying Lemma 9.1, since we do not let the iteration in the return map go to infinity, but all of the long segments of the broken paths are long enough to make the arguments there work.) Iterating in this way, we deduce the desired result. □

Proof of Theorem 13.1. We must show that $\|f\|$ is well-defined and satisfies convex union, weighted quasi-smoothing, strong stability and homogeneity. Let $R \geq 0$ be the quasi-smoothing constant of f , and let K and w be the curves and constants from Lemmas 13.2 and 13.4 (which we can take to be the same), and let $f(K)^+$ be $\max(f(K), 0)$.

- **Well-defined:** Lemma 13.2 shows that the sequence $(f(C^n) + wR + f(K)^+)_{n \in \mathbb{N}}$ is subadditive for large enough n , and thus by Lemma 9.5 the limit defining $\|f\|$ exists.
- **Convex union:** This follows immediately from the fact that $(C_1 \cup C_2)^k = C_1^k \cup C_2^k$, the definition of $\|f\|$ and convex union property of f .
- **Strong stability:** For any multi-curve D and any curve C , by Lemmas 13.4 and 13.3 (applied to C^k) and the quasi-smoothing and convex union properties of f , we have

$$\begin{aligned} f(D^k \cup C^{nk}) - (n - 1)wR &\leq f(D^k \cup nC^k) + (n - 1)f(K)^+ \\ f(D^k \cup nC^k) - (n - 1)R &\leq f(D^k \cup C^{nk}). \end{aligned}$$

Combining the inequalities, dividing by k and letting k go to infinity, we obtain

$$\|f\|(D \cup C^n) = \|f\|(D \cup nC).$$

We can iterate this to prove the result when C is a multi-curve.

- **Homogeneity:** It is clear from the definition of $\|f\|$ that $\|f\|(C^n) = n\|f\|(C)$. Homogeneity then follows from stability.
- **Weighted Quasi-smoothing:** Let $C = C_1 \cup C_2$ be a multi-curve, where the smoothing involves the component(s) in C_1 , so that C_1 and its smoothing C'_1 each have at most two components. Thus,

$C_1^k \searrow_{2(k-1)} kC_1$ and $kC_1 \searrow_k kC'_1$. Then

$$\begin{aligned} \|f\|(C) &= \lim_{k \rightarrow \infty} \frac{f(C_1^k \cup C_2^k)}{k} \\ &\geq \lim_{k \rightarrow \infty} \frac{f(kC_1 \cup C_2^k) - 2kR}{k} && \text{(Lemma 13.3)} \\ &\geq \lim_{k \rightarrow \infty} \frac{f(kC'_1 \cup C_2^k) - 3kR}{k} && \text{(quasi-smoothing for } f\text{)} \\ &\geq \lim_{k \rightarrow \infty} \frac{f((C'_1)^k \cup C_2^k) - 3kR - 2kwR - 2kf(K)^+}{k} && \text{(Lemma 13.4)} \\ &= \|f\|(C') - (3 + 2w)R - 2f(K)^+, \end{aligned}$$

so $\|f\|$ satisfies quasi-smoothing, with constant $(3 + 2w)R + 2f(K)^+$. By Proposition 3.6, $\|f\|$ also satisfies weighted quasi-smoothing. □

Finally, we show that with other hypotheses, (weak) stability implies strong stability, so that we do not need to assume strong stability in the statement of Theorem 6.1.

Corollary 13.5. *Let f be a weighted curve functional satisfying weighted quasi-smoothing, convexity, stability and homogeneity. Then f also satisfies strong stability.*

Proof. By the definition of $\|f\|$ and stability and homogeneity of f , we have, for all oriented multi-curves C ,

$$\|f\|(C) = f(C).$$

By Theorem B, $\|f\|$ satisfies strong stability. □

The proof of this part of Theorem 13.1 does not use Theorem 6.1, so we can use Corollary 13.5 in the proof of Theorem 6.1.

We finish by proving Theorem B.

Proof of Theorem B. By Proposition 3.6, f extends uniquely to a weighted curve functional satisfying convex union, homogeneity, stability and weighted quasi-smoothing with the same constant. Theorem 13.1 applied to this extension gives the result. □

Acknowledgements. We thank Francisco Arana, Martin Bridgeman, Maxime Fortier Bourque, Maria Beatrice Pozzetti, Kasra Rafi, and Tengren Zhang for helpful conversations and the anonymous referee for careful reading and suggestions. The first author was supported by the Mathematics Department Indiana University Bloomington, via the Hazel King Thompson fellowship and the Indiana University Graduate School under the Dissertation Research Fellowship. The second author was supported by the National Science Foundation under Grant Numbers DMS-1507244 and DMS-2110143.

Conflict of Interest: None.

References

[1] J. Aramayona and C. J. Leininger, ‘Hyperbolic structures on surfaces and geodesic currents’, *Algorithmic and Geometric Topics around Free Groups and Automorphisms*, Adv. Courses Math. CRM Barcelona (Birkhäuser/Springer, Cham, Switzerland, 2017), 111–149.
 [2] A. Bankovic and C. J. Leininger, ‘Marked-length-spectral rigidity for flat metrics’, *Trans. Amer. Math. Soc.* **370**(3) (2018), 1867–1884. MR 3739194.
 [3] W. Basener, ‘Global cross sections and minimal flows’, *Topol. Appl.* **121** (2002), 415–442.
 [4] W. Basener, ‘Every nonsingular C^1 flow on a closed manifold of dimension greater than two has a global transverse disk’, *Topol. Appl.* **135**(1–3) (2004), 131–148.
 [5] A. Basmajian, H. Parlier and J. Souto, ‘Geometric filling curves on surfaces’, *Bull. Lond. Math. Soc.* **49**(4) (2017), 660–669. MR 3725486.

- [6] H. Bauer, *Measure and Integration Theory, De Gruyter Studies in Mathematics*, Vol. **26** (Walter de Gruyter & Co., Berlin, 2001). Translated from the German by Robert B. Burckel. MR 1897176.
- [7] Y. Benoist and H. Oh, ‘Equidistribution of rational matrices in their conjugacy classes’, *Geom. Funct. Anal.* **17**(1) (2007), 1–32.
- [8] F. Bonahon, ‘The geometry of Teichmüller space via geodesic currents’, *Invent. Math.* **92**(1) (1988), 139–162.
- [9] F. Bonahon, ‘Geodesic currents on negatively curved groups’, in *Arboreal Group Theory (Berkeley, CA, 1988)*, *Math. Sci. Res. Inst. Publ.*, Vol. **19** (Springer, New York, 1991), 143–168. MR 1105332.
- [10] F. Bonahon, ‘Bouts des variétés hyperboliques de dimension 3’, *Ann. Math. (2)* **124**(1) (1986), 71–158.
- [11] F. Bonahon and D. Šarić, ‘A Thurston boundary for infinite-dimensional Teichmüller spaces’, *Math. Ann.* **380**(3-4) (2021), 1119–1167.
- [12] M. J. Bridgeman and E. C. Taylor, ‘An extension of the Weil–Petersson metric to quasi-Fuchsian space’, *Math. Ann.* **341**(4) (2008), 927–943.
- [13] M. Burger, A. Iozzi, A. Parreau and M. B. Pozzetti, ‘Currents, systoles, and compactifications of character varieties’, Preprint 2019, [arXiv:1902.07680](https://arxiv.org/abs/1902.07680).
- [14] E. Calabi, ‘Extremal isosystolic metrics for compact surfaces’, *Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992)*, *Sémin. Congr.*, Vol. 1 (Soc. Math. France, Paris, 1996), 167–204.
- [15] M. Cossarini and P. Dehornoy, ‘Intersection norms on surfaces and Birkhoff cross sections’, Preprint, 2016, [arXiv:1604.06688](https://arxiv.org/abs/1604.06688).
- [16] C. Croke, A. Fathi and J. Feldman, ‘The marked length-spectrum of a surface of nonpositive curvature’, *Topology* **31**(4) (1992), 847–855. MR 1191384.
- [17] N. G. de Bruijn and P. Erdős, ‘Some linear and some quadratic recursion formulas. II’, *Nederl. Akad. Wetensch. Proc. Ser. A. 55 = Indagationes Math.* **14** (1952), 152–163. MR 0047162.
- [18] M. de Graaf and A. Schrijver, ‘Making curves minimally crossing by Reidemeister moves’, *J. Combin. Theory Ser. B* **70**(1), (1997), 134–156.
- [19] A. Douady and J. Hubbard, ‘On the density of Strebel differentials’, *Invent. Math.* **30**(2) (1975), 175–179. MR 396936.
- [20] M. Duchin, C. J. Leininger and K. Rafi, ‘Length spectra and degeneration of flat metrics’, *Invent. Math.* **182**(2) (2010), 231–277. MR 2729268.
- [21] D. Dumas, ‘Skinning maps are finite-to-one’, *Acta Math.* **215**(1) (2015), 55–126. MR 3413977.
- [22] V. Erlandsson, ‘A remark on the word length in surface groups’, *Trans. Amer. Math. Soc.* **372**(1) (2019), 441–455. MR 3968775.
- [23] V. Erlandsson, H. Parlier and J. Souto, ‘Counting curves, and the stable length of currents’, *J. Eur. Math. Soc.* **22**(6) (2020), 1675–1702.
- [24] V. Erlandsson and J. Souto, ‘Counting curves in hyperbolic surfaces’, *Geom. Funct. Anal.* **26**(3) (2016), 729–777.
- [25] V. Erlandsson and J. Souto, ‘Mirzakhani’s curve counting’, Preprint, 2019, [arXiv:1904.05091](https://arxiv.org/abs/1904.05091).
- [26] V. Erlandsson and J. Souto, ‘Geodesic currents and Mirzakhani’s curve counting’, (2020).
- [27] M. Freedman, J. Hass and P. Scott, ‘Closed geodesics on surfaces’, *Bull. London Math. Soc.* **14**(5) (1982), 385–391.
- [28] O. Guichard and A. Wienhard, ‘Anosov representations: domains of discontinuity and applications’, *Invent. Math.* **190**(2) (2012), 357–438.
- [29] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities, Cambridge Mathematical Library* (Cambridge University Press, Cambridge, 1988). Reprint of the 1952 edition. MR 944909.
- [30] J. Hass and P. Scott, ‘Shortening curves on surfaces’, *Topology* **33**(1) (1994), 25–43. MR 1259513.
- [31] M. Headrick and B. Zwiebach, ‘Minimal-area metrics on the Swiss cross and punctured torus’, *Comm. Math. Phys.* **377**(3) (2020), 2287–2343.
- [32] S. Hersonsky and F. Paulin, ‘On the rigidity of discrete isometry groups of negatively curved spaces’, *Comment. Math. Helv.* **72**(3) (1997), 349–388. MR 1476054.
- [33] J. A. Jenkins, ‘On the existence of certain general extremal metrics’, *Ann. Math. (2)* **66**(3) (1957), 440–453.
- [34] F. Kassel, ‘Geometric structures and representations of discrete groups’, in *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018*. Vol. **II. Invited lectures** (World Scientific, Hackensack, NJ, 2018), 1115–1151. MR 3966802.
- [35] B. Kleiner and B. Leeb, ‘Rigidity of invariant convex sets in symmetric spaces’, *Invent. Math.* **163**(3) (2005), 657–676.
- [36] Z. Kominek and M. Kuczma, ‘Theorems of Bernstein-Doetsch, Piccard and Mehdi and semilinear topology’, *Arch. Math. (Basel)* **52**(6) (1989), 595–602. MR 1007635.
- [37] M. Kuczma, *An Introduction to the Theory of Functional Equations and Inequalities*, 2nd ed. (Birkhäuser, Basel, 2009).
- [38] G. Martone and T. Zhang, ‘Positively ratioed representations’, *Comment. Math. Helv.* **94**(2) (2019), 273–345. MR 3942786.
- [39] Howard Masur, ‘The Jenkins–Strebel differentials with one cylinder are dense’, *Comm. Math. Helv.* **54**(1) (1979), 179–184.
- [40] I. Meir, Does midpoint-convex imply rationally convex? (2019). URL: <https://mathoverflow.net/q/347499>.
- [41] R. A. Minlos, Convergence of measures, *Encyclopedia of Mathematics* (EMS Press, 2010). URL: <http://encyclopediaofmath.org/index.php?title=Convergenceofmeasures&oldid=3008>.
- [42] M. Mirzakhani, *Simple Geodesics on Hyperbolic Surfaces and the Volume of the Moduli Space of Curves*, Ph.D. Thesis (Harvard University, 2004).

- [43] M. Mirzakhani, 'Growth of the number of simple closed geodesics on hyperbolic surfaces', *Ann. Math. (2)* **168**(1) (2008), 97–125.
- [44] M. Mirzakhani, 'Counting mapping class group orbits on hyperbolic surfaces', Preprint, 2016, [arXiv:1601.03342](https://arxiv.org/abs/1601.03342).
- [45] E. Moggi, 'Notions of computation and monads', *Inform. and Comput.* **93**(1) (1991), 55–92.
- [46] U. Naseer and B. Zwiebach, 'Extremal isosystolic metrics with multiple bands of crossing geodesics', Preprint, 2019, [arXiv:1903.11755](https://arxiv.org/abs/1903.11755).
- [47] J.-P. Otal, 'Le spectre marqué des longueurs des surfaces à courbure négative', *Ann. Math. (2)* **131**(1) (1990), 151–162.
- [48] D. R. Palmer, *Toward Computing Extremal Quasiconformal Maps via Discrete Harmonic Measured Foliations*, A.B. Thesis (Harvard University, Cambridge, MA, 2015).
- [49] R. C. Penner and J. L. Harer, 'Combinatorics of train tracks', *Annals of Mathematics Studies*, Vol. **125** (Princeton University Press, Princeton, NJ, 1992). MR 1144770.
- [50] J.-F. Quint, 'Groupes convexes cocompacts en rang supérieur', *Geom. Dedicata* **113** (2005), 1–19.
- [51] K. Rafi and J. Souto, 'Geodesics currents and counting problems', *Geom. Funct. Anal.* **29**(3) (2019), 871–889.
- [52] B. Rodin, 'The method of extremal length', *Bull. Amer. Math. Soc.* **80**(4) (1974), 587–607.
- [53] M. Shepard, *The Topology of Shortest Curves in Surfaces*, Ph.D. Thesis (U.C. Berkeley, 1991).
- [54] The Stacks Project Authors, *Stacks Project* (2018). URL: <https://stacks.math.columbia.edu>.
- [55] D. P. Thurston, 'From rubber bands to rational maps: A research report', *Res. Math. Sci.* **3** (2016), Art. 15.
- [56] D. P. Thurston, 'Elastic graphs', *Forum Math. Sigma* **7** (2019), e24, 84.
- [57] D. P. Thurston, 'A positive characterization of rational maps', *Ann. Math. (2)* **192**(1) (2020), 1–46.
- [58] R. Tyrell Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).
- [59] M. Viana and K. Oliveira, *Foundations of Ergodic Theory*, *Cambridge Studies in Advanced Mathematics*, Vol. **151** (Cambridge University Press, Cambridge, 2016). MR 3558990.
- [60] P. Wadler, 'Monads for functional programming', in *Marktoberdorf Summer School on Program Design Calculi (M. Broy, ed.)*, *NATO ASI Series 118: Computer and systems sciences*, Vol. **118** (Springer, 1993), 233–264.
- [61] A. Wilkinson, 'Lectures on marked length spectrum rigidity', in *Geometric Group Theory*, *IAS/Park City Math. Ser.*, Vol. **21** (American Mathematical Society, Providence, RI, 2014), 283–324. MR 3329731.
- [62] M. Wolf and B. Zwiebach, 'The plumbing of minimal area surfaces', *J. Geom. Phys.* **15**(1) (1994), 23–56.