Karen Butt

Approximate rigidity of the marked length spectrum


https://doi.org/10.5802/mrr.18

© The authors, 2023.

This article is licensed under the Creative Commons Attribution 4.0 International License.

http://creativecommons.org/licenses/by/4.0/
Approximate rigidity of the marked length spectrum

Karen Butt

(Recommended by Boris Hasselblatt)

Abstract. We report on recent work investigating the extent to which finitely many closed geodesics approximately determine a negatively curved metric on a closed manifold. It is known in certain cases—and conjectured to be true in general—that the lengths of all closed geodesics (as a function of their free homotopy classes) determine the underlying negatively curved metric up to isometry. This length function is known as the marked length spectrum. Here, we consider certain pairs of Riemannian manifolds whose marked length spectra agree—only approximately—on a finite set of closed geodesics. We report on our recent results which show the two metrics are “almost isometric”. More precisely, we show the metrics are bi-Lipschitz equivalent with constant close to 1, and we obtain estimates for these constants depending only on concrete Riemannian data.

1. Introduction

1.1. Overview. We report on recent work proving quantitative versions of known marked length spectrum rigidity results. Throughout this article, we consider the setting where \((M, g)\) is a closed Riemannian manifold of negative sectional curvature. In broad terms, our work is about closed geodesics of \((M, g)\) (or, equivalently, periodic orbits of the geodesic flow) and the extent to which they determine the underlying metric \(g\). In our setting, it is known in certain cases—and conjectured to be true in general—that in order to determine the isometry type of the metric \(g\), it suffices to measure the lengths of all closed geodesics (as a function of their free homotopy classes). This phenomenon is known as marked length spectrum rigidity.

Here, we consider certain pairs of Riemannian manifolds whose marked length spectra agree—only approximately—on a finite set of closed geodesics. Our results show the two metrics are bi-Lipschitz equivalent with constant close to 1. More precisely, we obtain explicit estimates for this constant in terms of the measurement error and the length of the longest geodesic in the finite set. Our estimates depend only on concrete geometric information about the given metrics, such as the dimension, sectional curvature bounds, and injectivity radii.

1.2. The marked length spectrum. Closed geodesics are objects of fundamental interest—akin to prime numbers in many ways. They have been thoroughly investigated from multiple viewpoints—differential-geometric, dynamical, symplectic, number-theoretic, and spectral—which has in turn elucidated further connections between various areas of mathematics. For instance, Huber proved an asymptotic formula for

Received April 2, 2023; revised August 18, 2023.
2020 Mathematics Subject Classification. 37D40, 37D20, 37C27, 53C24, 53C22.
Keywords. marked length spectrum, closed geodesics, geodesic flow, negative curvature, locally symmetric spaces, rigidity.
the closed geodesic counting function for closed surfaces of constant negative curvature using ideas from analytic number theory, namely, the Selberg trace formula [47]. In addition, he showed that the lengths of closed geodesics and the spectrum of the Laplacian completely determine one another in this setting, spurring developments in both geometry and spectral theory.

In the more general setting of variable curvature, algebraic and number-theoretic tools are not as readily available, whereas dynamical and analytic approaches to studying closed geodesics have proved to be effective. In his seminal thesis, Margulis generalized Huber’s counting formula to variable negative curvature (and higher dimensions) in terms of the topological entropy of the geodesic flow, a central notion in dynamics measuring the complexity of a system’s orbit structure. (See [55] and [56, Chapter 6]; this was also independently proved by Bowen [9].) The relationship between lengths of closed geodesics and the spectrum of the Laplacian generalizes partially as well. For generic metrics, the latter determines the former, as shown by the trace formulae in the works of Colin de Verdière [69, 70] and Duistermaat–Guillemin [22].

A fundamental question in Riemannian geometry is determining a set of parameters which describe a metric up to isometry. In negative curvature, a natural candidate is the set of lengths of closed geodesics, also known as the length spectrum due to its close connection with the Laplace spectrum. In fact, to what extent the Laplace spectrum determines the metric is a question which falls into a broad class of inverse spectral problems, famously known by the tagline “Can one hear the shape of a drum?” [48]. It turns out that one cannot hear the shape of a negatively curved drum: the first examples of isospectral non-isometric hyperbolic surfaces were constructed by Vignéras [71], and Sunada later provided a method to generate more general counterexamples [67]. (There are also a multitude of counterexamples to this question outside of the negatively curved setting, such as [58]. See the survey [33] for more on this topic.)

Since the Laplace spectrum, and hence the length spectrum, turns out to not always describe a Riemannian manifold up to isometry, one is led to consider a few natural follow-up questions. One of these is to ask which geometric properties can in fact be recovered from spectral data; this is one of the main themes in the rich and vast field of spectral geometry, which is discussed, for instance, in the survey [73]. Another fruitful avenue, which leads to the subject of the present article, is to ask if there is some more refined spectral information which actually does determine a negatively curved metric up to isometry. From this perspective, it is natural to consider lengths of closed geodesics together with the additional information of their associated free homotopy classes, which leads to the following definition.

**Definition 1.1.** Given a closed, negatively curved Riemannian manifold \((M, g)\), the marked length spectrum \(L_g\) is the function on free homotopy classes of closed curves in \(M\) which associates to each class the length of its unique geodesic representative.

**Remark 1.2.** Whenever \((M, g)\) is a closed Riemannian manifold, every free homotopy class has a shortest-length representative, which is a closed geodesic. In other words, the existence of the geodesic representative follows simply from the compactness of \(M\). The negative curvature assumption is used to guarantee uniqueness. See, for instance, [16, p. 8-9].

**1.3. Marked length spectrum rigidity.** While the length spectrum alone is not sufficient to determine a negatively curved metric up to isometry, the additional information provided by the marking, that is, which lengths correspond to which curves, is conjectured to suffice.
Conjecture 1.1 ([13, Conjecture 3.1]). Let \((M, g)\) and \((M, g_0)\) be closed negatively curved Riemannian manifolds with \(\mathcal{L}_g = \mathcal{L}_{g_0}\). Then \(g\) and \(g_0\) are isometric by an isometry which preserves the marking, i.e., which is homotopic to the identity map on \(M\).

Remark 1.3. While it is customary to compare \(\mathcal{L}_g\) and \(\mathcal{L}_{g_0}\) for metrics \(g\) and \(g_0\) on the same manifold \(M\), this also makes sense more generally for manifolds \((M, g)\) and \((N, g_0)\) with isomorphic fundamental groups. Indeed, the set of free homotopy classes of \(M\) can be identified with conjugacy classes in the fundamental group \(\Gamma\) of \(M\), and as such we can view \(\mathcal{L}_g\) as a function on \(\Gamma\).

Conjecture 1.1 (marked length spectrum rigidity) is still open in general, but substantial progress has been made. We begin by discussing this conjecture in the special case of hyperbolic surfaces, that is, \(M\) and \(N\) have dimension 2, and \(g\) and \(g_0\) are both metrics of constant negative curvature (hyperbolic metrics). Here, marked length spectrum rigidity is a classical result due to Fricke–Klein [30] (see also [57]). In fact, a much stronger statement holds in this setting: it suffices to check the equality \(\mathcal{L}_g = \mathcal{L}_{g_0}\) on a certain finite set of free homotopy classes in order to guarantee \(g\) and \(g_0\) are isometric (see [23, Theorem 10.7] and [64]). This finiteness is closely related to the fact that the Teichmüller space of all possible hyperbolic metrics on a given topological surface has finite dimension (equal to \(6\text{genus}(M) - 6\)).

One way to see that Teichmüller space is \((6\text{genus}(M) - 6)\)-dimensional is by using Fenchel–Nielsen coordinates (see, for instance, [23, Chapter 10]). To summarize briefly, one can cut a closed surface up along certain closed geodesics so that each component is a pair of pants (topologically, a sphere with three punctures). A hyperbolic metric on each pair of pants is determined by the three “cuff lengths”. There are \(3\text{genus}(M) - 3\) total cuff lengths in any pants decomposition of \(M\), but these lengths alone do not suffice to determine the isometry type of \(M\). (They account for exactly half of the \(6\text{genus}(M) - 6\) Fenchel–Nielsen coordinates.) In addition to these cuff lengths, one needs to keep track of “twist parameters”, which dictate how the pants are glued back together to reconstruct the surface, since twisting a cuff before gluing it to another of the same length will change the isometry type of \(M\). It turns out that these twist parameters can be recovered from the lengths of finitely many additional closed geodesics (and as few as \(6\text{genus}(M) - 5\) total lengths are needed to determine the isometry type of the surface [64]).

It is natural to ask if marked length spectrum rigidity generalizes beyond this classical setting of hyperbolic surfaces. For hyperbolic metrics on higher dimensional closed manifolds, Mostow’s strong rigidity theorem says the isometry type of \(g\) (up to rescaling) is determined by the fundamental group of \(M\) alone [59], so it is not even necessary to consider lengths of closed geodesics in this case.

In variable negative curvature, the situation is much more complicated, even for surfaces. The pants decomposition method above is not at all applicable in this setting. Indeed, one can simply perturb the metric in the interior of a pair of pants without changing any of the cuff lengths, so the metric cannot be described by finitely many coordinates.

Nevertheless, Conjecture 1.1 was resolved in the case of (variably curved) surfaces independently by both Otal [60] and Croke [18]. Prior to these results, partial progress was made by Guillemin–Kazhdan in the case of one-parameter families of metrics (deformation rigidity) [39] and by Katok in the case of metrics in the same conformal class [49]. Later, Croke–Fathi–Feldman generalized Otal’s argument to surfaces of non-positive curvature [20], and recently Guillarmou–Lefeuvre–Paternain devised a new argument that proves marked length spectrum rigidity for all surfaces whose geodesic
flow is Anosov [38]. We also mention that Gogolev–Rodriguez Hertz proved a weighted version of marked length spectrum rigidity for negatively curved surfaces [32, Corollary 1.3].

In higher dimensions, Conjecture 1.1 was solved by Hamenstädt [43] in the case where one of the two metrics is locally symmetric, using the entropy rigidity theorem of Besson–Courtois–Gallot [6, 7]. A more direct proof also follows from work of Bourdon [8]. More recently, the conjecture was solved locally, that is, for two metrics which are sufficiently close in some suitable $C^k$ topology, by Guillarmou–Lefeuvre, using techniques from microlocal analysis [37].

In these above works, the methods are of course different from the ones used for hyperbolic metrics. Broadly speaking, the proofs all make use of dynamical properties of the geodesic flow.

1.4. Main results on approximate rigidity. As mentioned above, in these variably curved settings, it is not possible to describe a metric up to isometry by finitely many closed geodesics. Indeed, one can simply perturb the metric in a neighborhood outside any given finite set of closed curves. However, it is natural to ask if finitely many closed geodesics can still provide approximate information about the metric.

**Question 1.4.** Does the marked length spectrum on a sufficiently large finite set approximately determine the metric?

Question 1.4 has not been previously considered anywhere in the literature as far as we know (aside from the case of surfaces of constant negative curvature). All known proofs of marked length spectrum rigidity in the variably curved setting rely on dynamics of the geodesic flow. In particular, these proofs use limiting procedures involving longer and longer closed geodesics. For example, the topological entropy $h(g)$ of the geodesic flow is the exponential growth rate of periodic orbits, i.e., it can be obtained from the marked length spectrum $\mathcal{L}_g$ as follows:

$$h(g) = \lim_{T \to \infty} \frac{\log(\#\{\gamma \in \Gamma | \mathcal{L}_g(\gamma)\})}{T}.$$  

(See [9, 55, 56].) It is clear that changing $\mathcal{L}_g$ on a finite set does not affect the value of the entropy; however, it is not at all clear what information can be obtained about $h(g)$ from only knowing $\mathcal{L}_g$ on a finite set. On a related note, Sawyer proved that for negatively curved surfaces, it is enough for $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ to agree on a set of conjugacy classes whose complement has subexponential growth in order to conclude $g$ and $g_0$ are isometric [63]. In the local setting, Guillarmou–Knieper–Lefeuvre proved that $g$ and $g_0$ are isometric so long as $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ agree asymptotically [36]. In particular, in both of the above cases, rigidity holds if $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ coincide outside of a finite set, and Question 1.4 is a natural counterpart to this.

We approach Question 1.4 in two steps. In [14], we show that a sufficiently large finite set of closed geodesics approximately determines the full marked length spectrum, and the approximation improves as the size of the set of known closed geodesics increases. In fact, we only require that the length functions $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ coincide approximately on this finite set:

**Hypothesis 1.5.** Let $\Gamma$ denote the fundamental group of $M$. Let $f : M \to N$ be a homotopy equivalence, and let $f_*$ denote the induced map on fundamental groups. For $L > 0$, let $\Gamma_L := \{\gamma \in \Gamma | \mathcal{L}_g(\gamma) \leq L\}$. Now let $\epsilon > 0$ small and suppose

$$1 - \epsilon \leq \frac{\mathcal{L}_{g_0}(f_* \gamma)}{\mathcal{L}_g(\gamma)} \leq 1 + \epsilon.$$
for all $\gamma \in \Gamma$.

If $L$ is sufficiently large, we obtain estimates for the ratio $\mathcal{L}_g/M \mathcal{L}_g$ on all of $\Gamma$ in terms of $\varepsilon$ and $L$. Note that the estimates do not depend on the particular pair of metrics under consideration; they are uniform for all $(M, g)$ and $(N, g_0)$ with pinched sectional curvatures and injectivity radii bounded away from zero.

**Theorem 1.6** ([14, Theorem 1.2]). Let $(M, g)$ and $(N, g_0)$ be closed Riemannian manifolds of dimension $n$ with sectional curvatures contained in the interval $[-\lambda^2, -\lambda^2]$. Let $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ denote their marked length spectra. Let $\Gamma$ denote the fundamental group of $M$ and let $i_N$ denote the injectivity radius of $(N, g_0)$. Suppose there is a homotopy equivalence $f : M \to N$ and let $f_\ast$ denote the induced map on fundamental groups.

Then there is $L_0 = L_0(n, \Gamma, \lambda, \Lambda, i_N)$ so that the following holds: Suppose the marked length spectra $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ satisfy Hypothesis 1.5 for some $\varepsilon > 0$ and $L \geq L_0$. Then there exist constants $C > 0$ and $0 < \alpha < 1$, depending only on $n, \Gamma, \lambda, \Lambda, i_N$, so that

$$1 - (\varepsilon + C L^{-\alpha}) \leq \frac{\mathcal{L}_{g_0}(f_\ast \gamma)}{\mathcal{L}_g(\gamma)} \leq 1 + (\varepsilon + C L^{-\alpha})$$

for all $\gamma \in \Gamma$.

**Remark 1.7.** A similar result can be deduced from the finite Livsic theorem of Gouëzel–Lefeuvre [34, Theorem 1.2]. However, we opt for a more direct approach in order to show how the constants involved depend explicitly on concrete Riemannian data. We discuss this in greater detail in [14, Remark 2.10].

**Remark 1.8.** Recent work of Cantrell–Reyes improves the estimates in Theorem 1.6 above to $\alpha = 1$ using coarse-geometric methods [17, Theorem 8.1].

In light of this theorem, we now consider metrics $g$ and $g_0$ whose marked length spectra are multiplicatively close on the set of all free homotopy classes of closed curves, as in the hypothesis below:

**Hypothesis 1.9.** There is some small $\tilde{\varepsilon} > 0$ so that

$$1 - \tilde{\varepsilon} \leq \frac{\mathcal{L}_{g_0}(f_\ast \gamma)}{\mathcal{L}_g(\gamma)} \leq 1 + \tilde{\varepsilon}$$

for all $\gamma \in \Gamma$, where $\Gamma$ and $f_\ast$ are as in Hypothesis 1.5.

This reduces Question 1.4 to the following question for closed negatively curved manifolds in general.

**Question 1.10.** If two metrics have marked length spectra which are not equal, but are multiplicatively close (as in Hypothesis 1.9), is there a sense in which the metrics are close?

In [15], we answer Question 1.10 in dimension 2, and in higher dimensions when one of the metrics is locally symmetric. These are two of the main cases where marked length spectrum rigidity is known—due to Otal and Croke for surfaces [60, 18], and Hamenstädt and Besson–Courtois–Gallot for higher dimensions [43, 6].

Question 1.10 was previously known for hyperbolic surfaces and in general for pairs of metrics $g$ and $g_0$ on the same manifold $M$ which are sufficiently close in some suitable $C^k$ topology. The first case is due to Thurston [68]. He showed that if $(M, g)$ and $(N, g_0)$ are both surfaces of constant negative curvature, and $f : M \to N$ is a fixed homeomorphism, then the best possible Lipschitz constant for a map $F : M \to N$ in the same homotopy class as $f$ is precisely $\sup_{\gamma \in \Gamma} \frac{\mathcal{L}_{g_0}(f_\ast \gamma)}{\mathcal{L}_g(\gamma)}$. The second case is part of the previously mentioned work of Guillarmou–Knieper–Lefeuvre [36]. Their techniques provide
explicit estimates (in a suitable Sobolev norm) for how close the metrics are in terms of the ratio \( \frac{D_{g_0}}{D_g} \), or more precisely the geodesic stretch; in fact, their results hold more generally for non-positively curved metrics with Anosov geodesic flow. However, this work requires \( g \) and \( g_0 \) to be sufficiently close metrics (in some \( C^k \) topology) on the same manifold.

Our results in [14, 15] do not require the metrics to be close, nor do they require the two metrics to be on the same manifold, but only on pairs of manifolds \( M \) and \( N \) with isomorphic fundamental groups. In our setting (negative curvature), a standard result in algebraic topology states any isomorphism of fundamental groups is induced by a homotopy equivalence \( M \to N \), and a deep theorem of Farrell–Jones states that this homotopy equivalence can be upgraded to a homeomorphism (in dimensions not equal to 3 or 4) [25]. However, \( M \) and \( N \) need not be diffeomorphic \textit{a priori} [24, 26, 1, 2]. Our results cover these cases as well.

In [15, Theorem 1.1], we prove an approximate version of Otal’s theorem [60] for certain pairs of surfaces \((M, g)\) and \((M, h)\) with bounded geometry. We consider the set \( \mathcal{C}(2, \lambda, \Lambda, v_0, D_0) \) of all closed \( C^\infty \) Riemannian manifolds of dimension 2 with sectional curvatures contained in the interval \([-\Lambda^2, -\lambda^2]\), volume bounded below by \( v_0 \), and diameter bounded above by \( D_0 \). We show pairs of such spaces become more isometric as their marked length spectra get closer to one another. More precisely, we show that for any \( L > 1 \), there exists \( \tilde{\varepsilon} = \tilde{\varepsilon}(L, \lambda, \Lambda, v_0, D_0) > 0 \) small enough so that the following holds. For any pair \((M_1, g_1), (M_2, h_2)\) \( \in \mathcal{C}(2, \lambda, \Lambda, v_0, D_0) \) satisfying

\[
1 - \tilde{\varepsilon} \leq \frac{D_{g_1}}{D_{h_2}} \leq 1 + \tilde{\varepsilon},
\]

there exists an \( L \)-Lipschitz diffeomorphism \( f : (M_1, g_1) \to (M_2, h_2) \).

Our main result of [15] is in the case where \((N, g_0)\) is a negatively curved locally symmetric space of dimension at least 3. We quantify how close \( g \) and \( g_0 \) are to being isometric by estimating the derivative of a map \( F : M \to N \) in terms of \( \tilde{\varepsilon} \). This is considerably stronger than our result [15, Theorem 1.1] for surfaces, since we are able to determine how the Lipschitz constant depends on \( \tilde{\varepsilon} \). This refines the rigidity result in [43, Corollary to Theorem A], which corresponds to the case \( \tilde{\varepsilon} = 0 \) in the theorem below. As in [15, Theorem 1.1], we only assume the marked length spectra of the two metrics are close; we do not assume the metrics themselves are close in any \( C^k \) topology.

**Theorem 1.11** ([15, Theorem 1.2]). \textit{Let \((M, g)\) be a closed Riemannian manifold of dimension \( n \geq 3 \) with fundamental group \( \Gamma \) and sectional curvatures contained in the interval \([-\Lambda^2, 0]\). Let \((N, g_0)\) be a locally symmetric space. Assume there is a homotopy equivalence \( f : M \to N \) and let \( f_* \) denote the induced map on fundamental groups. Then there exists small enough \( \varepsilon_0 \) (depending on \( \Gamma \)) so that whenever \( \tilde{\varepsilon} \leq \varepsilon_0 \) and

\[
1 - \tilde{\varepsilon} \leq \frac{D_{g_0}(f_* \gamma)}{D_g(\gamma)} \leq 1 + \tilde{\varepsilon}
\]

for all \( \gamma \in \Gamma \), there is a diffeomorphism \( F : M \to N \) homotopic to \( f \) and constants \( c_1(\tilde{\varepsilon}, n, \Gamma, \Lambda) < 1 \) and \( C_2(\tilde{\varepsilon}, n, \Gamma, \Lambda) > 1 \) such that for all \( v \in TM \) we have

\[
c_1 \| v \|_g \leq \| dF(v) \|_{g_0} \leq C_2 \| v \|_g.
\]

More precisely, there exists a constant \( C = C(n, \Gamma, \Lambda) \) so that \( c_1 = 1 - C\tilde{\varepsilon}^{1/8(n+1)} \) and \( C_2 = 1 + C\tilde{\varepsilon}^{1/8(n+1)} \).

**Remark 1.12.** The conclusion of Theorem 1.11 can be restated as \( \| g - F^* g_0 \|_{C^0} \leq C\tilde{\varepsilon}^{1/8(n+1)} \).
Remark 1.13. If \( \tilde{N} \) is a real, complex or quaternionic hyperbolic space, we can take \( c_1 = 1 - C \varepsilon^{1/4(n+1)} \) and \( c_2 = 1 + C \varepsilon^{1/4(n+1)} \).

2. Dynamics of the geodesic flow

Some alternate intuition in support of Conjecture 1.1—which has led to significant partial results in the setting of variable negative curvature—comes from dynamics. The dynamical system under consideration is the geodesic flow, which we denote by \( \phi^t \). This is a flow on the unit tangent bundle \( T^1 M \) of \((M, g)\). The flow \( \phi^t \) is defined as follows: Given a unit tangent vector \( v \), first consider the unique unit speed geodesic \( c(t) \) with initial condition \( c'(0) = v \). Now for any \( t \in \mathbb{R} \), define \( \phi^t v \) to be the unit tangent vector \( c'(t) \). A simple but important observation is that periodic orbits of \( \phi^t \) correspond precisely to closed geodesics in \( M \).

In our setting, that is, when \((M, g)\) is closed and negatively curved, the geodesic flow is uniformly hyperbolic, more commonly known as Anosov (see [29, Definition 5.1.1]). This hyperbolicity turns out to reveal significant information about the overall orbit structure of the flow (despite the fact that individual trajectories are highly sensitive to small changes in initial conditions). For instance, periodic orbits of Anosov flows are dense. In our geometric setting, this means vectors tangent to closed geodesics are dense in \( T^1 M \).

There are also stronger results about approximating certain trajectories of Anosov flows with periodic ones, such as the Anosov closing lemma, which we discuss below. That is to say, from the perspective of hyperbolic dynamics, it is natural to expect periodic orbits of the geodesic flow to provide significant information about the flow.

Definition 2.1. (See [29, Definition 5.1.1].) A flow \( \phi^t \) on \( T^1 M \) is said to be Anosov if there is a splitting

\[
T(T^1 M) \cong X \oplus E^s \oplus E^u
\]

into \( d\phi^t \)-invariant subbundles, where \( X \) denotes the vector field tangent to the flow direction, and \( E^s \) and \( E^u \) denote the stable and unstable distributions, respectively. Vectors in the stable bundle are uniformly exponentially contracted under \( d\phi^t \), whereas vectors in the unstable bundle are uniformly expanded.

To illustrate why the geodesic flow is Anosov when \( M \) is negatively curved, we will describe the (strong) stable and (strong) unstable foliations \( W^ss, W^su \subset T^1 M \). The leaves of these foliations are tangent to the stable and unstable distributions \( E^s \) and \( E^u \), respectively. (See, for instance, [3, p. 72].) Let \( v \in T^1 \tilde{M} \). Let \( p \in \tilde{M} \) be the forward projection of \( v \) and let \( \xi \in \partial \tilde{M} \) be the forward projection of \( v \in T^1 \tilde{M} \) to the visual boundary at infinity. Let \( B_{\xi, p} \) denote the Busemann function on \( \tilde{M} \) and let \( H_{\xi, p} \) denote its zero set. Then the lift of \( W^ss(v) \) to \( T^1 \tilde{M} \) is given by

\[
\{ -\text{grad}B_{\xi, p}(q) \mid q \in H_{\xi, p} \}.
\]

In other words, these are vectors normal to the horosphere \( H_{\xi, p} \) which are pointing towards \( \xi \). If \( \eta \) denotes the projection of \( -v \) to the boundary \( \partial \tilde{M} \), then the lift of \( W^su(v) \) to \( T^1 \tilde{M} \) is analogously given by

\[
\{ \text{grad}B_{\eta, p}(q) \mid q \in H_{\eta, p} \}.
\]

These are vectors orthogonal to \( H_{\eta, p} \) and whose negatives point towards \( \eta \).

Such a family of vectors gives rise to a geodesic variation, and the verification that the geodesic flow on \( T^1 M \) is Anosov boils down to estimates of the norms of the Jacobi fields (together with their covariant derivatives) associated to these variations. The idea is to use the Rauch comparison theorem to compare with the constant curvature setting; here, the Jacobi equation can be solved explicitly, thereby concretely illustrating
the desired exponential divergence/convergence of nearby geodesics. For further details see [3, Proposition IV.1.13 and Proposition IV.2.15]. A slightly different approach is given in [29, Theorem 5.2.4].

Anosov flows are often described as chaotic, since a slight change of initial condition (in the unstable direction) causes exponential divergence of trajectories. Nevertheless, we have the following strong result about approximating certain trajectories with periodic ones. See, for instance, [29, Theorem 5.3.10]. In the statement of Lemma 2.2 below, \( d \) denotes the Riemannian distance on the unit tangent bundle induced by \( g \).

**Lemma 2.2** (Anosov closing lemma). There is \( \delta_0 > 0 \) sufficiently small, \( T \) sufficiently large, and a constant \( C > 0 \) so that the following holds for all \( \delta \leq \delta_0 \) and all \( t \geq T \). Suppose \( v, \phi^t v \in T^1 M \) are so that \( d(v, \phi^t v) < \delta \). Then either \( v \) and \( \phi^t v \) are on the same local flow line or there is \( w \) with \( d(v, w) < C\delta \) so that \( w \) is tangent to a closed geodesic of length \( t' \in [t - C\delta, t + C\delta] \).

Put briefly, the Anosov closing lemma says that “almost periodic” trajectories are shadowed by periodic ones. Any flow which preserves some finite measure \( \mu \) has an abundance of such almost periodic trajectories. Indeed, the Poincaré Recurrence Theorem [50, Theorem 4.1.19] implies that \( \mu \)-almost every \( v \) will return arbitrarily close to itself after flowing for a sufficiently long time. This applies to the geodesic flow of any closed Riemannian manifold \( M \) because the flow preserves the Liouville measure, a natural measure on \( T^1 M \) induced by the Riemannian volume of \( M \). If \( M \) is, in addition, negatively curved, the Anosov closing lemma implies that periodic orbits are also abundant; in particular, they are dense in \( T^1 M \).

As mentioned above, the Anosov closing lemma suggests that knowledge of the marked length spectrum should provide significant information about the underlying geodesic flow. This intuition leads to the following fact. (See [40]. Below, we outline an alternate standard approach, which uses the Livsic theorem [53].)

**Proposition 2.1.** Let \( (M, g) \) and \( (N, g_0) \) be a pair of homotopy-equivalent closed negatively curved manifolds such that their marked length spectra \( \mathcal{L}_g \) and \( \mathcal{L}_{g_0} \) coincide. Let \( \phi^t \) and \( \psi^t \) denote the associated geodesic flows on the unit tangent bundles \( T^1 M \) and \( T^1 N \), respectively. Then the flows \( \phi^t \) and \( \psi^t \) are conjugate, that is, there is a homeomorphism \( \mathcal{F} : T^1 M \to T^1 N \) so that

\[
\mathcal{F}(\phi^t v) = \psi^t \mathcal{F}(v)
\]

for all \( v \in T^1 M \).

The conjugacy \( \mathcal{F} \) is a strong form of equivalence between the flows \( \phi^t \) and \( \psi^t \) as it preserves dynamically defined invariants, such as periodic orbits (together with their lengths), the stable and unstable manifolds, and the topological entropies.

Without any assumptions on the length functions \( \mathcal{L}_g \) and \( \mathcal{L}_{g_0} \), that is, whenever \( (M, g) \) and \( (N, g_0) \) are homotopy-equivalent closed negatively curved manifolds, Gromov proved the associated geodesic flows are orbit-equivalent [35]. This means there is a homeomorphism \( \mathcal{F}^t : T^1 M \to T^1 N \) such that

\[
\mathcal{F}^t(\phi^t v) = \psi^{a(t, v)} \mathcal{F}^t(v)
\]

for some function \( a(t, v) \) on \( \mathbb{R} \times T^1 M \). This is a weaker form of equivalence than a conjugacy; for instance, the map \( \mathcal{F}^t \) takes stable/unstable manifolds to weak stable/unstable manifolds, that is, manifolds tangent to the distributions \( E^s \oplus X \) and \( E^u \oplus X \), respectively. Periodic orbits are preserved by \( \mathcal{F}^t \), but not their lengths, since \( \mathcal{L}_g \) and \( \mathcal{L}_{g_0} \) do not in general coincide.

The additional condition \( \mathcal{L}_g = \mathcal{L}_{g_0} \) is sufficient to upgrade the above orbit equivalence to a conjugacy. See [50, Section 2.2]. The key tool used is the Livsic Theorem [53]
(see also [50, Theorem 19.2.1]), whose proof in turn relies on the Anosov closing lemma, together with the fact that the geodesic flow has a dense orbit.

While the conjugacy \( \mathcal{F} : T^1 M \to T^1 N \) provides significant dynamical information about the geodesic flows \( \phi^t \) and \( \psi^t \), it does not give immediate information about the underlying metrics \( g \) and \( g_0 \) on \( M \) and \( N \). Showing \( g \) and \( g_0 \) are isometric entails finding a map from \( M \) to \( N \), and it is not clear whether the conjugacy \( \mathcal{F} : T^1 M \to T^1 N \) between unit tangent bundles is fiber-preserving, i.e., descends to the base manifolds. (Another difficulty to note is that \( \mathcal{F} \) may only be of \( C^0 \) regularity, not necessarily differentiable.) Proving marked length spectrum rigidity (Conjecture 1.1) thus requires more sophisticated considerations of the geometry and dynamics of the geodesic flow.

In Croke’s proof of marked length spectrum rigidity for surfaces [18], he obtains the desired isometry \( f : M \to N \) directly from the conjugacy \( \mathcal{F} : T^1 M \to T^1 N \). He considers the image under \( \mathcal{F} \) of the fiber \( T^1_p M \) of all unit tangent vectors based at a single point \( p \), together with the associated geodesics in the universal cover \( \tilde{N} \). Croke proves that these geodesics must all intersect at a single point \( q \) (by showing the Jacobi field arising from the geodesic variation of \( \mathcal{F}(T^1_q M) \) vanishes at some point). Thus, the conjugacy \( \mathcal{F} : T^1 M \to T^1 N \) descends to a map \( f : M \to N \). The fact that the conjugacy is time-preserving immediately implies that \( f \) is distance-preserving and hence an isometry. Otal’s methods in [60] are different from Croke’s, but the overall approach is similar in spirit. While Otal does not work directly with the conjugacy between unit tangent bundles, he shows the associated correspondence of geodesics takes the geodesics with initial vectors in \( T^1_p M \) to geodesics intersecting in a single point \( q \in \tilde{N} \), thereby obtaining a map \( f : M \to N \). Both authors’ constructions rely heavily on the dimension 2 hypothesis; simply put, this is the setting in which it is easiest for pairs of geodesics to intersect.

Hamenstädt partially resolved Conjecture 1.1 in dimensions 3 and greater [43], a setting where Otal and Croke’s constructions of the desired isometry do not readily generalize. Instead, she leverages additional rigidity properties enjoyed by negatively curved metrics in higher dimensions, namely, the celebrated entropy rigidity theorem of Besson–Courtois–Gallot [6] (see [62] for the Cayley case). Below, we state the relevant special case of the theorem, which is the subject of the survey article [7].

**Theorem 2.3** (Besson–Courtois–Gallot). Let \( (N, g_0) \) be a compact negatively curved locally symmetric space of dimension \( n \geq 3 \) and let \( (M, g) \) be a negatively curved Riemannian manifold which is homotopy-equivalent to \( N \). Suppose that the topological entropies \( h(g) \) and \( h(g_0) \) of their geodesic flows coincide and that the total volumes \( \text{Vol}_g(M) \) and \( \text{Vol}_{g_0}(N) \) coincide. Then there is an isometry \( F : (M, g) \to (N, g_0) \).

**Remark 2.4.** If \( (M, g) \) is also negatively curved and locally symmetric, this reduces to Mostow rigidity (see [28, Corollary 1.3]). Note also that this statement fails in dimension 2 for the same reasons as Mostow rigidity. Indeed, take any two non-isometric surfaces \( (M, g) \) and \( (M, g_0) \) of constant curvature \( -1 \) (there is a \( (6 \text{genus}(M) - 6) \)-dimensional Teichmüller space of such metrics). By Gauss–Bonnet, the total areas of both surfaces agree. By [54], the topological entropy of the geodesic flow is equal to the volume growth entropy, that is, the exponential growth rate of balls in the universal cover. These must also coincide because both surfaces are covered by the hyperbolic plane.

In the setting of the statement of Theorem 2.3, Hamenstädt obtains marked length spectrum rigidity by showing that \( \mathcal{L}_g = \mathcal{L}_{g_0} \) implies \( \text{Vol}_g(M) = \text{Vol}_{g_0}(N) \) [43, Theorem A]. The equality of entropies is automatic in this case, for instance using (1.1), or even Proposition 2.1. Hence, \( M \) and \( N \) are isometric by Theorem 2.3.
3. Methods

We now give an overview of the proofs of our main results. First, we discuss Theorem 1.11, which is our approximate version of Hamenstädt’s marked length spectrum rigidity result. Recall this is in the higher dimensional setting, when one of the two metrics is locally symmetric. Theorem 1.11 states that if the length functions $L_g$ and $L_{g_0}$ are multiplicatively close, the two metrics are bi-Lipschitz equivalent with constant close to 1. Afterwards, we discuss Theorem 1.6, our result that finitely many closed geodesics determine the full marked length spectrum approximately.

Before we explain our proof of Theorem 1.11, we recall Hamenstädt’s approach for the $\tilde{\epsilon} = 0$ case. In [43, Theorem A], Hamenstädt proves that two negatively curved manifolds with the same marked length spectrum have the same volume, provided one of the manifolds has geodesic flow with $C^1$ Anosov splitting, a condition which holds in particular for locally symmetric spaces. (The Anosov splitting of the geodesic flow on the unit tangent bundle $T^1N$ refers to the flow-invariant decomposition of $TT^1N$ into the stable, unstable and flow directions as in Definition 2.1 above; see, for instance, [29, Definition 5.1.1].)

Thus, if $M$ and $N$ satisfy the assumptions of Theorem 1.11 for $\tilde{\epsilon} = 0$, they must have the same volume. Then, since the marked length spectrum determines the topological entropy of the geodesic flow, the fact that the two manifolds are isometric follows from the celebrated entropy rigidity theorem of Besson–Courtois–Gallot [7, 6]. (See also Theorem 2.3 above for the statement.)

Our proof of Theorem 1.11 consists of the same two key steps as in the $\tilde{\epsilon} = 0$ case: the volume step and the entropy rigidity step. First, we first show an approximate version of Hamenstädt’s volume theorem (Theorem 3.1 below), that is, if the ratio of $L_g$ and $L_{g_0}$ is close to 1, so is the ratio of the volumes $\text{Vol}_g(M)$ and $\text{Vol}_{g_0}(N)$. It is clear from (1.1) that the ratio of the entropies must also be close to 1 in this case. The second main step in our proof of Theorem 1.11 is then an approximate version of the Besson–Courtois–Gallot entropy rigidity theorem. We estimate the derivative of the natural map $F : M \to N$ constructed in [7] under the assumption that the volumes and entropies of $M$ and $N$ are almost equal instead of equal. This assumption is satisfied when the length functions $L_g$ and $L_{g_0}$ are multiplicatively close, but otherwise our proof does not use this hypothesis on the lengths.

3.1. The volume step. To prove Theorem 1.11, we start by proving an analogue of [43, Theorem A] under the assumption the marked length spectra satisfy Hypothesis 1.9, i.e., we estimate the ratio $\text{Vol}(M)/\text{Vol}(N)$ in terms of $\tilde{\epsilon}$. In order to obtain an explicit estimate, we assume the Anosov splitting is $C^{1+\alpha}$ instead of $C^1$. (For geodesic flows on manifolds with strictly $\frac{1}{4}$-pinched negative curvature, the Anosov splitting is $C^{1+\alpha}$ for some $\alpha > 0$. The splitting is $C^1$ by work of Hirsch–Pugh [46] and $C^{1+\alpha}$ by work of Hasselblatt [44, Theorem 5, Remark after Theorem 6].) Unlike in Theorem 1.11, the constants here do not depend on $(M, g)$ in any way.

**Theorem 3.1** ([15, Theorem 1.4]). Let $(M, g)$ be a closed negatively curved Riemannian manifold with fundamental group $\Gamma$. Let $(N, g_0)$ be another closed negatively curved manifold with fundamental group $\Gamma$ and assume the geodesic flow on $T^1N$ has $C^{1+\alpha}$ Anosov splitting. Suppose the marked length spectra of $M$ and $N$ satisfy

$$1 - \tilde{\epsilon} \leq \frac{L_{g_0}(\gamma)}{L_g(\gamma)} \leq 1 + \tilde{\epsilon}$$

for all $\gamma \in \Gamma$. Then there is a constant $C$ depending only on $\tilde{\epsilon}$ such that

$$(1 - C\tilde{\epsilon}^3)(1 - \tilde{\epsilon})^n \text{Vol}(M) \leq \text{Vol}(N) \leq (1 + C\tilde{\epsilon}^3)(1 + \tilde{\epsilon})^n \text{Vol}(M).$$
If, in addition, \((N, g_0)\) is locally symmetric and \(\varepsilon\) is sufficiently small (depending on \(n = \dim N\)), then \(\alpha\) can be replaced with 2 in the above estimates and the constant \(C\) depends only on \(n\).

**Remark 3.2.** If the Anosov splitting of \(T^1 N\) is only \(C^1\), then our proof shows the quantities \((1 \pm C\varepsilon^n)\) can be replaced with constants that converge to 1 as \(\varepsilon \to 0\), but we are not able to determine the explicit dependence of these constants on \(\varepsilon\).

**Remark 3.3.** If \(N\) is locally symmetric, then \(\Vol(N) \leq (1 + \varepsilon)^n \Vol(M)\) follows from (1.1) and the proof of the main theorem in [7]. However, the lower bound for \(\Vol(N)/\Vol(M)\) in Theorem 3.1 is also crucial for the proof of Theorem 1.11.

**Remark 3.4.** If \(\dim M = \dim N = 2\), then our proof of Theorem 3.1 shows

\[
(1 - \varepsilon)^2 \Vol(M) \leq \Vol(N) \leq (1 + \varepsilon)^2 \Vol(M),
\]

which is the optimal estimate. This result also follows from [19, Theorem 1.1].

The key concept used in Hamenstädt’s proof of the above theorem in the \(\varepsilon = 0\) case is the **Liouville measure**. This is a measure on the unit tangent bundle \(T^1 M\), which we denote by \(\mu\). On the one hand, this measure is compatible with the volume on \(M\) arising from the Riemannian metric \(g\). More precisely, in a coordinate chart \(U \times S^{n-1} \subset T^1 M\), the measure \(\mu\) is locally the product of Riemannian volume on \(U \subset M\) and Lebesgue measure on the fiber \(S^{n-1}\). Thus, the total volume of \(M\) is determined by the total measure \(\mu(T^1 M)\).

On the other hand, the Liouville measure is geodesic flow–invariant, and this dynamical point of view is better suited to investigating the relationship between \(\mu\) and the marked length spectrum. This alternative description of \(\mu\) comes from a natural contact structure on the unit tangent bundle. Let \(\omega\) be the 1-form on \(T^1 M\) obtained by pulling back the canonical 1-form on \(T^* M\) to \(TM\) via the identification induced by the Riemannian metric and then restricting to \(T^1 M\). If \(X\) denotes the vector field on \(T^1 M\) generating the geodesic flow, a straightforward calculation shows that \(\omega(X) \equiv 1\). This, in turn, shows that \(\omega\) is a flow-invariant contact form, meaning \(\omega \wedge (d\omega)^{n-1}\) is a flow-invariant volume form on \(T^1 M\). The measure arising from this volume form coincides (up to a constant multiple) with the local product description of the Liouville measure given above; see, for instance, [12, 1.E].

Incidentally, the volume form \(\omega \wedge (d\omega)^{n-1}\) leads directly to an alternative local product structure for \(\mu\). Roughly speaking, \(\omega\) can be viewed as a one-dimensional measure in the flow direction, whereas \((d\omega)^{n-1}\) can be interpreted as a measure on the space of geodesics transverse to the flow. The space of geodesics \(\mathcal{M}\) in the universal cover \(\hat{M}\) is defined as the quotient of \(T^1 \hat{M}\) by the geodesic flow, that is, we identify any two unit tangent vectors on the same geodesic. (In negative curvature, the space \(\mathcal{M}\) is also identified with pairs of distinct points in the visual boundary at infinity \(\partial \hat{M}\).) Since the 2-form \(d\omega\) is flow-invariant, it descends to a 2-form on \(\mathcal{M}\), where it is a symplectic form; in other words, \((d\omega)^{n-1}\) is a volume form on the space of geodesics. We call the associated measure the **Liouville current**, which we denote by \(\lambda\). As such, we can locally write the Liouville measure as \(d\mu = dt \times d\lambda\), where \(dt\) is 1-dimensional Lebesgue measure on orbits and \(d\lambda\) is the Liouville current on the space of geodesics \(\mathcal{M}\).

To show the marked length spectrum determines, or approximately determines, the total Riemannian volume, it thus suffices to consider the measures \(dt\) (the time component) and \(d\lambda\) (the Liouville current) separately.

**3.1.1. The Liouville current.** The key tool that connects the marked length spectrum to the Liouville current is the **cross-ratio**, which goes back to Otal’s proof of marked length
spectrum rigidity for surfaces. While many of Otal’s methods contrast with Hamenstät’s, the fact that the marked length spectrum determines the Liouville current is a key step in both proofs.

The cross-ratio associated to a closed negatively curved manifold \((M, g)\) is the following function on quadruples of distinct points \(a, b, c, d\) in the visual boundary at infinity \(\partial M\). See [61, Lemma 2.1].

**Definition 3.5.** Let \(a_i, b_i, c_i, d_i \in \partial M\) be sequences converging to \(a, b, c, d \in \partial M\), respectively. Define

\[
[a, b, c, d] = \lim_{i \to \infty} d(a_i, c_i) + d(b_i, d_i) - d(a_i, d_i) - d(b_i, c_i),
\]

where \(d\) is the Riemannian distance function. By [61, Lemma 2.1], this limit exists and is independent of the chosen sequences \(a_i, b_i, c_i, d_i\). We call \([\cdot, \cdot, \cdot, \cdot]\) the cross-ratio.

The cross-ratio is completely determined by the marked length spectrum for manifolds \(M\) of any dimension [61, Theorem 2.2]. The basic idea is to approximate each of the four distances in (3.1) by lengths of closed geodesics using the Anosov closing lemma. In our setting, it is straightforward to verify that if the ratio \(\mathcal{L}_g / \mathcal{L}_{g_0}\) is between \(1 ± \varepsilon\), so is the ratio of the cross-ratios [15, Proposition 2.3].

In the proof of [60, Theorem 2], Otal explicitly relates the Liouville current and the cross-ratio in dimension 2 (though the word “cross-ratio” never appears in this paper). Let \(a, b, c, d \in \partial M\) be four distinct points. When \(\dim(M) = 2\), the boundary \(\partial M\) is a circle; hence, the pair of points \((a, b)\) determines an interval in the boundary (after fixing an orientation). Let \((a, b) \times (c, d) \in \partial M\) denote the geodesics starting in the interval \((a, b)\) and ending in the interval \((c, d)\). In the proof of [60, Theorem 2], Otal shows

\[
\lambda((a, b) \times (c, d)) = \frac{1}{2} [a, b, c, d].
\]

(See also [45, Theorem 4.4].

In [43], Hamenstät relates the cross-ratio and the Liouville current in higher dimensions. The proof uses several technical constructions, but the key idea is an insightful dynamical interpretation of the cross-ratio, which explicitly relates it to the symplectic form \(d\omega\). (This interpretation of the cross-ratio is also explained in [72, Section 2.3].)

The tool which relates the cross-ratio and \(d\omega\) is the temporal function (see also [42]). This is a ubiquitous object in the study of (contact) Anosov flows. Consider a point \(v \in T^1 M\) and let \(W^{ss}(v)\) be the leaf of the strong stable foliation through \(v\). Let \(W^s(v) = W^{ss}(v) \cap B(v, \delta)\), where \(B(v, \delta)\) is a ball of radius \(\delta\) with respect to some fixed Riemannian distance on the unit tangent bundle. Define \(W^u(v)\) analogously. Let \(w \in W^u(v)\) and let \(z \in W^s(v)\). Then there is a unique time \(\sigma = \sigma(w, z)\) so that the intersection \(W^u(w) \cap W^s(\phi^{\sigma} z)\) consists of a single point, denoted \([w, z]\) [29, Proposition 6.2.2]. We call \(\sigma(w, z)\) the temporal function. See Figure 1.

When the Anosov flow \(\phi^t\) preserves a contact form \(\omega\), the temporal function is nonzero. Indeed, the contact condition implies that the distribution \(\ker(\omega)\) is maximally non-integrable, and the Anosov condition (or more generally, hyperbolicity) implies the stable and unstable distributions \(E^s\) and \(E^u\) are contained in \(\ker(\omega)\). It then follows that \(E^s\) and \(E^u\) are not jointly integrable, and hence \(\sigma(w, z) \neq 0\).

In addition, \(\sigma(w, z)\) can be computed from the symplectic form \(d\omega\) as follows (see also [41, Lemma 2.1]). Consider a \(C^1\) surface \(S\) bounded by the five arcs in Figure 1, i.e., the ones connecting the points \(v, w, [w, z], \phi^{-\sigma}[w, z], z\). Note that with the exception of the flow line from \([w, z]\) to \(\phi^{\sigma}[w, z]\), these arcs are all tangent to vectors in \(\ker(\omega)\). By Stokes’ theorem, we obtain

\[
\int_s d\omega = \int_{\partial S} \omega = \sigma(w, z).
\]
Approximate rigidity of the marked length spectrum

It turns out that the temporal function can also be expressed in terms of the cross-ratio. More precisely, for any \( v \in T^1 M \), let \( \pi(v) \) denote the projection to \( \partial \tilde{M} \) along the geodesic ray with initial velocity \( v \). Then for \( w \) and \( z \) as above we have

\[
\sigma(w, z) = [\pi(v), \pi(w), \pi(-z), \pi(-v)].
\]

To see this, we start by specializing the above dynamical definition of the temporal function to our geometric setting (see Figure 1). Recall \( w \in W^{su}_\delta(v) \) and \( z \in W^{ss}_\delta(v) \). Now consider the stable horosphere through the basepoint of \( w \) and the boundary point \( \pi(w) \). Similarly consider the horosphere through \( \pi(-z) \) and the basepoint of \( z \). Then the vector \( [w, z] \) is the normal vector of the horosphere associated to \( W^{su}_\delta(w) \) which also lies on the geodesic starting at \( \pi(-z) \) and ending at \( \pi(w) \). Thus, \( \sigma(w, z) \) is simply the distance between the horospheres associated to \( W^{su}_\delta(w) \) and \( W^{ss}_\delta(z) \), as illustrated in the righthand panel of Figure 1.

The cross-ratio as defined in Definition 3.5 can also be directly related to distances between horospheres. This idea is used in the proof of [61, Lemma 2.1]. Let \( a, b, c, d \in \partial \tilde{M} \) and take four disjoint horospheres \( H_a, H_b, H_c, H_d \) based at each of these points. Now consider the four bi-infinite geodesics \( [a, d], [b, c], [a, c], [b, d] \). The intersection of these bi-infinite geodesics with \( \tilde{M} \setminus (H_a \cup H_b \cup H_c \cup H_d) \) consists of four finite geodesic segments. Then the limit in (3.1) is equal to the analogous sum of signed lengths of these four geodesic segments. Furthermore, this quantity does not depend on the choice of horospheres. (See the proof of [61, Lemma 2.1].) One can choose larger horospheres so that all but one of the four terms is zero, and the nonzero term is equal to \( \sigma(w, z) \). This is illustrated in [72, Figure 14].

In summary, the cross-ratio can be viewed as a “coarse version of the symplectic structure on the space of geodesics” [43, p. 115]. When \( M \) has geodesic flow with \( C^1 \) Anosov splitting, the temporal function \( \sigma \) is \( C^1 \), and in this case the value of \( d\omega(X,Y) \) can be obtained directly from the cross-ratio by differentiating \( \sigma \) in the direction \( X,Y \). This is discussed in Lemma 2.1 of Hamenstädt’s earlier paper [41]. In fact, the main result of this paper shows the marked length spectrum determines the volume when both \( M \) and \( N \) have \( C^1 \) Anosov splitting [41, Corollary 1]. In [43], Hamenstädt extends this to the case where only one of the two manifolds has \( C^1 \) Anosov splitting, which introduces significant technical difficulties. The proof involves complicated constructions of auxiliary measures \( \mathcal{S} \) and \( \mathcal{P} \) determined by the cross-ratio (and hence the marked length spectrum). These measures are in turn defined as limits of measures \( \mathcal{S}_\delta \) and \( \mathcal{P}_\delta \).
as $\delta \to 0$, where the definitions of $\mathcal{R}_g$ and $\mathcal{R}_0$ involve covering subsets of the boundary $\partial \hat{M}$ by certain sets (quasi-symplectic balls) of diameter less than $\delta$.

As mentioned before, in our setting of multiplicatively close length functions $\mathcal{L}_g$ and $\mathcal{L}_{g_0}$ (Hypothesis 1.9), it is straightforward to verify the cross-ratios arising of $(M, g)$ and $(N, g_0)$ are also multiplicatively close. One might then expect $\mathcal{F}^M$ and $\mathcal{F}^N$ approximately coincide as well (and similarly for $\mathcal{P}$); however, the approximate equality of cross-ratios only implies the measures $\mathcal{F}^M$ and $\mathcal{F}^N$ are close (where $\delta$ is the number in Hypothesis 1.9). As such, we undertake a detailed analysis of Hamenstädt’s construction to quantify the rate at which $\mathcal{F}_\delta$ converges to $\mathcal{F}$. We also investigate this convergence more precisely when $N$ is locally symmetric. Here, we use the analyticity of the Anosov splitting to compute part of the power series expansion of the temporal function in order to obtain a better volume estimate in this setting (replacing $\alpha$ with $2$ in Theorem 3.1). This estimate is the starting point for the subsequent entropy rigidity step in the proof of our main result (Theorem 1.11).

3.1.2. The time component. In Hamenstädt’s setting of $\mathcal{L}_g = \mathcal{L}_{g_0}$, Proposition 2.1 gives a time-preserving conjugacy $\mathcal{F} : T^1M \to T^1N$ of geodesic flows; in other words, $\mathcal{F}$ preserves the $dt$ component of $d\mu$. In contrast, if the length functions do not coincide exactly, the flows cannot be conjugate, though they are still orbit-equivalent by [35] (as mentioned after the statement of Proposition 2.1). Recall that an orbit equivalence of flows $\phi^t$ and $\psi^t$ on $T^1M$ and $T^1N$, respectively, is a homeomorphism $\mathcal{F} : T^1M \to T^1N$ such that

$$\mathcal{F}(\phi^t v) = \psi^a(t, v) \mathcal{F}(v)$$

for all $t \in \mathbb{R}$ and $v \in T^1M$. In [15], we prove our volume estimate in Theorem 3.1 by delicately implementing Gromov’s construction of $\mathcal{F}$ in [35]. More precisely, we use the assumption of approximately equal lengths (Hypothesis 1.9) to show the time change $a(t, v)$ is close to $t$ on sets of large measure. This, together with an estimate of the ratios of Liouville currents $\lambda_g$ and $\lambda_{g_0}$ discussed above, allows us to show the total Liouville measures of $T^1M$ and $T^1N$ are close.

In order to further explain our approach, we now explain Gromov’s construction of the orbit equivalence $\mathcal{F}$ from [35]. Recall $(M, g)$ and $(N, g_0)$ are closed negatively curved manifolds with isomorphic fundamental groups. Since the universal covers $\hat{M}$ and $\hat{N}$ are contractible, all higher homotopy groups of $M$ and $N$ are trivial; hence, the isomorphism of fundamental groups is induced by a homotopy equivalence $f : M \to N$, and this homotopy equivalence can be taken to be $C^1$, and hence Lipschitz. The compactness of $M$ and $N$ guarantees that the lift $\hat{f} : \hat{M} \to \hat{N}$ of this homotopy equivalence to the universal covers is a quasi-isometry [4, Proposition C.1.2]. This means that if $c(t)$ is a geodesic in $\hat{M}$, then $\hat{f}(c(t))$ is a quasi-geodesic in $\hat{N}$. The Morse lemma in turn implies that there is a unique geodesic $\eta$ in $\hat{N}$ which is within bounded Hausdorff distance of the quasi-geodesic $\hat{f} \circ c$ [10, Theorem III.H.1.7].

By definition, an orbit equivalence $\mathcal{F}$ of geodesic flows should take vectors tangent to the geodesic $c(t)$ in $\hat{M}$ to tangent vectors along some geodesic in $\hat{N}$, and the geodesic $\eta$ described above is a natural candidate for this. As such, let $P_\eta : \hat{N} \to \eta$ denote the orthogonal projection map. Define $\mathcal{F}_0(c^t(t))$ to be the tangent vector to $P_\eta(\hat{f}(c(t)))$ at time $t$. Let $v = c'(0)$. Then there is some function $b(t, v)$ so that $\mathcal{F}_0(\phi^t v) = \psi^{b(t, v)} \mathcal{F}_0(v)$ for all $t \in \mathbb{R}$. However, $\mathcal{F}_0$ is not an orbit equivalence, because it is possible for a fiber of the normal projection to intersect the quasi-geodesic $\hat{f}(c(t))$ in more than one point; thus, $\mathcal{F}_0$ is not necessarily injective.
In order to fix this problem, Gromov replaces the function \( b(t, v) \) in the definition of \( \mathcal{F}_0 \) with an average of itself over a sufficiently long interval. In other words, let

\[
a_c(t, v) = \frac{1}{c} \int_{t}^{t+c} b(s, v) \, ds.
\]

Then there is a large enough \( c \), depending only on the quasi-isometry constants of \( \tilde{f} \), so that \( t \mapsto a_c(t, v) \) is injective for all \( v \). Then \( \mathcal{F}(\phi^t v) := \psi^{a_c(t, v)} \mathcal{F}_0(v) \) is an orbit-equivalence of geodesic flows. (For details, see [15, Proposition 2.25] or [52].)

**Remark 3.6.** Another well-known way to obtain an orbit equivalence of geodesic flows is using **Anosov structural stability**, which states that a sufficiently nice perturbation of an Anosov flow is orbit-equivalent to the original one [29, Theorem 5.3.6]. In the Riemannian setting, the hypotheses of structural stability apply to pairs of geodesic flows arising from pairs of sufficiently nearby metrics on the same manifold. Gromov’s orbit equivalence in [35] can thus be thought of as a global version of structural stability for geodesic flows. The proof of structural stability uses the shadowing property of Anosov flows to approximate orbits of the perturbed flow by genuine orbits; in Gromov’s construction, the Morse lemma can be thought of as a global analogue of this shadowing property. (In dimension 2, Anosov structural stability can be made global directly, since the space of all negatively curved metrics on a surface is path-connected (the proof is sketched in [72, Proposition 3.1]); however, this path-connectedness fails in high enough dimensions [27].)

In our proof of Theorem 3.1, we use the assumption of approximately equal length functions (Hypothesis 1.9) to show that for very large \( c \), the time-change cocycle \( a_c(t, v) \) is close to \( t \) on sets of large measure. We consider the **speed** of the orbit equivalence, i.e., the derivative

\[
\frac{d}{dt} a_c(t, v) \bigg|_{t=0} = \frac{b(c, v)}{c},
\]

and we seek to show this expression can be made close to 1 for large enough \( c \). We do this by interpreting the righthand side above as an ergodic average, and then use the ergodic theorem to show that for almost every \( v \), the quantity \( b(c, v)/c \) converges to a limit which can in turn be well-approximated by the ratio \( L_g(\gamma)/L_{g_0}(\gamma) \) for a suitable \( \gamma \) using [65, Theorem 1]. By Hypothesis 1.9, this ratio is close to 1 as desired.

**Remark 3.7.** The above approach is different in strategy from the proof of Proposition 2.1 outlined above. Our method gives an orbit equivalence which has speed close to 1, but only on a subset of \( T^1 M \) of large measure. Since we are only trying to estimate the total volumes of \( M \) and \( N \), and not the Liouville measure of small sets in \( T^1 M \) and \( T^1 N \), this is sufficient for our purposes. Note that our proof does not make use of any Livsic theorems. There are approximate versions of the Livsic theorem which apply to our setting such as [34, Theorem 1.2] and an earlier special case [51], but the estimates in the conclusions of these theorems involve constants which a priori depend on the given flow, so we opted for a more direct approach.

**Remark 3.8.** An alternate approach to obtaining a controlled orbit equivalence under Hypothesis 1.9 follows from a construction of Bourdon [8, Section 1.4]. Bourdon constructs a conjugacy of geodesic flows between CAT\((-1)\) spaces, which are a generalization of negatively curved Riemannian manifolds, under the assumption that the boundary map \( \tilde{f} : \partial \tilde{M} \to \partial \tilde{N} \) preserves the cross-ratio. The construction starts with an identification of \( T^1 M \) with triples of distinct points in \( \partial \tilde{M} \). (Note that for dimensional reasons, such an identification is not one-to-one when \( \dim M = n \geq 3 \). The boundary map \( \tilde{f} \)
induces a map on triples of points in the boundary, and this map descends to a well-defined map of unit tangent bundles so long as \( \overline{f} \) preserves quadruples of points with zero cross-ratio. By [15, Proposition 2.3], this holds so long as the length functions satisfy 
\[
A \leq \frac{f_*}{f_0} \leq B \text{ for any } A, B > 0.
\]
By Hypothesis 1.9, this holds for \( A = 1 - \varepsilon, B = 1 + \varepsilon \). In this setting, the map \( \overline{f} \) approximately preserves the cross-ratio, and in this case Bourdon's construction of the conjugacy can be easily seen to be an orbit equivalence with speed between 1 ± \( \varepsilon \).

### 3.2. The entropy rigidity step.
In [43], Hamenstädt obtains a marked length spectrum rigidity result by using the entropy rigidity theorem of Besson–Courtois–Gallot [7] (Theorem 2.3 above). Recall this theorem says \((M, g)\) is isometric to its homotopy-equivalent locally symmetric counterpart \((N, g_0)\) whenever equality of volumes and equality of entropies are satisfied. To prove our approximate marked length spectrum rigidity result Theorem 1.11, we consider instead the entropy stability problem: if the equalities of volumes and entropies are replaced with almost equalities, is \( g \) close to the locally symmetric metric \( g_0 \)? Related questions have been considered by Bessières–Besson–Courtois–Gallot [5], and very recently by Song [66].

We begin by recalling the construction of the map \( F : M \to N \) in [7]. We then summarize the proof that \( F \) is an isometry in the case of equal entropies and volumes, before explaining how to modify it for approximately equal entropies and volumes.

Given \( p \in M \), let \( \mu_p \) be the Patterson-Sullivan measure on \( \partial M \). Let \( \overline{f} : \partial M \to \partial N \) be the boundary homeomorphism induced by the given homotopy equivalence \( f : M \to N \). (As discussed above, the map \( f : M \to N \) can be lifted to \( \tilde{f} : \tilde{M} \to \tilde{N} \), and compactness of \( M \) and \( N \) implies \( \tilde{f} \) is a quasi-isometry. Thus, \( \tilde{f} \) can be extended to a homeomorphism \( \overline{f} \) between the boundaries \( \partial M \) and \( \partial N \).

Now define \( F(p) = \text{bar}(f_* \mu_p) \), where \( \text{bar} \) denotes the barycenter map (see [7] for more details). (As discussed in [7], the technique of using barycenters of measures on the boundary originates in [31, 21, 11].)

By the definition of the barycenter, the map \( F \) has the implicit description
\[
\int_{\partial N} dB_{F(p), \xi} (\cdot) d(\overline{f}_* \mu_p)(\xi) = 0,
\]
where \( \xi \in \partial N \) and \( B_{F(p), \xi} \) is the Busemann function on \((\tilde{N}, g_0)\). This is a \( \Gamma \)-equivariant map \( \tilde{M} \to \tilde{N} \), and thus descends to a map \( M \to N \). Abusing notation, we call \( F : M \to N \) the BCG map. By the implicit function theorem, the BCG map \( F \) is \( C^1 \) (actually, \( C^2 \) since Busemann functions on \( \tilde{M} \) are \( C^2 \) [3, Proposition IV.3.2]), and its derivative \( dF_p \) satisfies
\[
\int_{\partial N} \text{Hess} B_{F(p), \xi} (dF_p(v), u) d(\overline{f}_* \mu_p)(\xi) = h(g) \int_{\partial N} dB_{F(p), \xi} (u) dB_{F(p), \xi} (v) d(\overline{f}_* \mu_p)(\xi)
\]
for all \( v \in T_p M \) and \( u \in T_{F(p)} N \) [7, (5.2)].

Without any assumptions about the volumes or entropies, the following inequality holds; see [62] for the Cayley case.

**Lemma 3.9** ([7, Proposition 5.2(ii)]).
\[
|\text{Jac}(F)| \leq \left( \frac{h(g)}{h(g_0)} \right)^n.
\]

As in the proof of [7, Theorem 5.1], the above lemma relates the volumes of \( M \) and \( N \) as follows:
\[
\text{Vol}(N, g_0) \leq \int_M |F^* dVol| = \int_M |(\text{Jac} F) dVol| \leq \left( \frac{h(g)}{h(g_0)} \right)^n \text{Vol}(M, g).
\]
Remark 3.10. This, together with (1.1), improves one of the inequalities in Theorem 3.1 in the special case where \( N \) is a locally symmetric space.

With this setup in mind, the argument in [7] showing that \( F \) is an isometry consists of the following components:

1. If the volumes and entropies are equal, then the inequalities in (3.3) are all equalities, which gives equality in Lemma 3.9.
2. This forces several other inequalities in the proof of Lemma 3.9 to be equalities, which forces certain matrices to be scalar matrices. (For example, if equality holds in the arithmetic-geometric mean inequality
   \[
   \det A \leq \left( \frac{\text{trace}(A)}{n} \right)^n,
   \]
   then \( A \) is a scalar matrix.) The end of the proof of [7, Proposition 5.2(ii)] then shows that \( dF_p \) has all its eigenvalues equal to the ratio 
   \[
   \left( \frac{h(g)}{h(g_0)} \right)^n,
   \]
   which means \( F \) is an isometry in the case where the entropies are equal. This concludes the proof of [7, Theorem 1].

Assuming instead that 
   \[
   1 - \tilde{\epsilon} \leq \frac{\text{Vol}(M)}{\text{Vol}(M)} \leq 1 + \tilde{\epsilon},
   \]
   the equalities of volumes and entropies are replaced with the conclusions of Theorem 3.1 and Lemma 1.1 respectively. Proceeding as in the above outline, we can instead obtain estimates for \( \| dF_p \| \) in terms of \( \tilde{\epsilon} \):

1. We show equality almost holds in (3.3); that is, we find a lower bound for \( \text{Jac} F(p) \) of the form \( \beta \left( \frac{h(g)}{h(g_0)} \right)^n \) for suitable \( \beta \).
2. This implies the above-mentioned matrices are almost scalar matrices, in that their eigenvalues are all approximately equal. We then mimic the proof of [7, Proposition 5.2(ii)] to obtain bounds for the eigenvalues of \( dF_p \), which completes the proof of Theorem 1.11.

The main difficulty is step (1), where we cannot simply mimic the arguments in [7]. Indeed, with the above assumptions about the entropies (1.1) and the volumes (Theorem 3.1), the inequalities in (3.3) become

\[
(1 - C\tilde{\epsilon}^2)(1 - \tilde{\epsilon})^n \frac{1}{(1 + \tilde{\epsilon})^n} \left( \frac{h(g)}{h(g_0)} \right)^n \text{Vol}(M) \leq \int_M |\text{Jac} F| \leq \left( \frac{h(g)}{h(g_0)} \right)^n \text{Vol}(M),
\]

which does not give a lower bound for the integrand. In order to obtain a lower bound for \( |\text{Jac} F| \), we use the above lower bound for its integral together with a Lipschitz bound for the function \( p \mapsto |\text{Jac} F(p)| \). The fact that this function is Lipschitz is immediate from the fact that \( F \) is \( C^2 \); however, it is not clear a priori how the Lipschitz bound depends on \( (M, g) \). Assuming 
   \[
   1 - \tilde{\epsilon} \leq \frac{\text{Vol}(M)}{\text{Vol}(M)} \leq 1 + \tilde{\epsilon}
   \]
   holds (Hypothesis 1.9) for \( \epsilon \) sufficiently small (depending on \( n \) and \( \Gamma \)), we show there is a Lipschitz bound for \( \text{Jac} F(p) \) depending only on the dimension \( n \), the fundamental group \( \Gamma \) and the lower bound \(-\Lambda^2 \) for the sectional curvatures of \( M \).

3.3. Finiteness. We now give a sketch of the proof of Theorem 1.6, which states that finitely many closed geodesics determine the full marked length spectrum approximately. The basic idea is to start by covering the unit tangent bundle \( T^1 M \) with finitely many sufficiently small "flow boxes", that is, sets obtained by flowing local transversals for some fixed time interval \((0, \delta)\). On the one hand, any periodic orbit of the flow that visits each of these boxes at most once is short, i.e., has period at most \( \delta \) times the total number of boxes. On the other hand, any periodic orbit that is long, i.e., of length more than \( \delta \) times the number of boxes, must return to at least one of the boxes more than once before it closes up. In other words, long periodic orbits contain shorter almost-periodic segments. By the Anosov closing lemma, these are in turn shadowed
by periodic orbits. This allows us to approximate the lengths of long closed geodesics with sums of lengths of short ones. We then use a Hölder continuous orbit equivalence \( F : T^1 M \to T^1 N \) to argue that similar approximations hold for the corresponding closed geodesics in \( N \). From this, we are able to estimate the ratio of \( L_g(\gamma) / L_{g_0}(\gamma) \) for all long geodesics \( \gamma \) given our assumed estimate holds for short ones (Hypothesis 1.5).

However, considerable technical difficulties arise in making some of the standard results from the theory of Anosov flows effective in geometric terms. Since these results are stated so generally, they contain a multitude of constants which depend on the given flow in arguably mysterious ways. We illustrate this by briefly mentioning one of several instances where such a difficulty occurs, namely, the well-known fact that a continuous orbit equivalence of Anosov flows is Hölder continuous [50, Theorem 19.1.5]. While the proof in [50] shows the Hölder exponent depends only on the exponential expansion/contraction rates of the flow, which in our geometric setting are determined entirely by the sectional curvature bounds \(-\lambda^2\) and \(-\Lambda^2\), the dependence of the multiplicative constant on the given flow is far less transparent. Indeed, the proof in [50] uses the fact that the starting orbit equivalence is uniformly continuous, and the final estimate ends up depending on this initial modulus of continuity. As such, to ensure the constants \( C \) and \( \alpha \) in the conclusion of Theorem 1.6 do not depend on the particular metrics \( g \) and \( g_0 \), we take great care to refine the statements of several standard dynamical facts to this geometric setting.

References

Approximate rigidity of the marked length spectrum

Karen Butt


Karen Butt: Department of Mathematics, University of Chicago, 5734 S University Ave, Chicago, IL 60637
E-mail: karenbutt@uchicago.edu