Interrogating surface length spectra and quantifying isospectrality

Hugo Parlier[*]

Abstract. This article is about inverse spectral problems for hyperbolic surfaces and in particular how length spectra relate to the geometry of the underlying surface. A quantitative answer is given to the following: how many questions do you need to ask a length spectrum to determine it completely? In answering this, a quantitative upper bound is given on the number of isospectral but non-isometric surfaces of a given genus.

1. Introduction

This article is about inverse spectral problems for hyperbolic surfaces and in particular for how length spectra relate to the geometry of the underlying surface. The idea is to understand how much information about a surface X is contained in the set of lengths (with multiplicities) of all of its closed geodesics $\Lambda(X)$.

The main goal here is to provide quantitive results which only depend on the topology of the underlying surface to several problems. Here is the first one:

Problem 1: *How many questions must one ask a length spectrum to determine it completely?*

There are variations on this problem depending on what type of question one allows. Here we'll only allow one type of question. You're allowed to constitute a (finite) list of values you're not interested in and then ask for the smallest value not on the list.

The approach taken to tackle this problem leads to another one:

Problem 2: *How many non-isometric isospectral surfaces of genus g can their be?*

This last problem is a well-studied problem in inverse spectral theory and for closed surfaces, via the Selberg trace formula, length isospectrality is equivalent to Laplace isospectrality.

There are different known techniques to produce examples of non-isometric isospectral hyperbolic surfaces. That such surfaces exist might seem surprising, and in many ways they are an extremely rare phenomenon. McKean [11] showed that at most a finite number

of other surfaces can be isospectral but non-isometric to a given hyperbolic surface and Wolpert [22] showed that outside of a certain proper real analytic subvariety of the moduli space $\mathcal{M}_g$ of genus $g \geq 2$ surfaces , all surfaces are uniquely determined by their length spectrum. The first examples were due to Vignéras [20, 21]. A multitude examples, not only in the context of surfaces, were found using a technique introduced by Sunada [17], namely those of Gordon, Webb and Wolpert [6], who also produced isospectral but non-isometric planar domains, answering a famous question of Kac [8].

As one might expect, the size of sets of isospectral non-isometric surfaces (isospectral sets) can grow with the topology of the underlying surface. In particular Brooks, Gornet and Gustafson [4] showed the existence of isospectral sets of cardinality that grows like $g^{c \log(g)}$ where g is the genus and c is an explicit constant. This improved previous results of Tse [19]. It would be difficult to review all of the literature, but the point is that a lot of effort has gone into finding examples of isospectral sets, and in particular into producing lower bounds to Problem 2 above.

On the other end, in addition to theorems of McKean and Wolpert, the only quantitative upper bounds seem to be due to Buser [5, Chapter 13]. Buser shows that the cardinality of isospectral sets is bounded above by e^{720g^2}. The proof is based on a number of ingredients, one of them being a bound on lengths of geodesics in pants decompositions (so-called Bers constants, see [2]). Since Buser's theorem, there have been improvements as to what is known about short pants decompositions which in turn lead to direct improvements of the e^{720g^2} upper bound. However, a direct application of Buser's techniques will still give bounds on the order of e^{Cg^2} for some constant C. To improve this bound significantly requires using something else and this is one of the goals of this paper.

Concerning Problem 1, the main result is the following theorem.

Theorem 1.1. *There exists an explicit universal constant A such that the following holds. The length spectrum Λ of a surface of genus g can be determined by at most g^{Ag} questions.*

The constant A can be taken to be 154, which is not in any way sharp. Note that in particular a finite number of lengths determine the full spectrum, a well known fact which was somehow at the origin of this project and the above result is one way of quantifying it.

We note there is another version of this finiteness, again due to by Buser [5, Theorem 10.1.4], used to provide a simpler (or at least a different) proof of Wolpert's theorem mentioned previously. The result is that there is a constant $B_{g,\varepsilon}$ that only depends on the genus g of the surface and a lower bound ε on the systole such that the lengths of length less than $B_{g,\varepsilon}$ determine the full length spectrum. (The systole is the length of a shortest closed

geodesic of the surface.) The proof is based on analyticity but there aren't any known quantifications of $B_{g,\varepsilon}$. Note that it must depend on g and ε, unlike the quantification of Theorem 1.1. It is also interesting to compare these results to other rigidity phenomena for length spectra, such as result of Bhagwat and Rajan [3], which states that, for even dimensional compact hyperbolic manifolds, two length spectra are either equal or they differ by an infinite number of values.

The proof of Theorem 1.1 involved determining possible isometry classes of surfaces with given length spectra. In particular, it is necessary to count the size of isospectral sets.

Theorem 1.2. *There exists a explicit universal constant B such that the following holds. Given* $X \in M_g$ *there are at most* g^{Bg} *surfaces in* $\mathcal{M}_g$ *isospectral to X.*

As before, the constant B can be taken to be 154. Although this considerably lowers the upper bound, there is still a significant discrepancy between the lower and upper bounds ($g^{c \log(g)}$ vs. g^{Cg}). Finding the true order of growth seems like a challenging problem.

One of the main ingredients for obtaining these quantifications is to find a different parameter set for the moduli space of surfaces. Buser's approach uses Fenchel-Nielsen coordinates and pants decompositions. The problem is that even the shortest pants decompositions curves can get long (as least on the order of $\sqrt{g}$). In fact, even the shortest closed geodesic can get long too as there are families of surfaces whose systoles grow on the order of $\log(g)$. The parameter set proposed here is also based on a set of curves whose geometric data determine the surface. The set of curves has two different components: Γ (the curves) and $\Gamma_{\mathcal{A}}$ (the chains). The set Γ is a set of disjoint simple curves and $\Gamma_{\mathcal{A}}$ is a set of curves somehow attached to the elements of Γ via a set of arcs $\mathcal{A}$ (see Section 3 for a proper definition). The first main feature of a curve and chain system $\Gamma, \Gamma_{\mathcal{A}}$ is that lengths and twists of elements of Γ and the lengths of elements of $\Gamma_{\mathcal{A}}$ determine a surface in moduli space (and in fact in Teichmüller space although that's not the point here).

Their second main feature - and this is what truly distinguishes them from pants decompositions - is that we can bound their lengths by constants on the order of $\log(g)$.

Theorem 1.3. *Any* $X \in \mathcal{M}_g$ *admits a curve and chain system* $\Gamma, \Gamma_{\mathcal{A}}$ *satisfying*

$$\ell(\gamma) < 2 \log(4g)$$

for all $\gamma \in \Gamma$ *and*

$$\ell(\gamma_a) < 8 \log(4g)$$

for all $a \in \mathcal{A}$.

3

This theorem is a type of Bers' constant theorem for curve and chain systems and the examples of surfaces with large systoles show that the order of growth is not improvable. The $\log(g)$ order of growth is an essential ingredient for obtaining the quantification in Theorem 1.2. We note that Theorems 1.2, 1.1 and 1.3 all remain true for complete finite area surfaces when replacing the dependence on g with a dependence on the area but, in tune with most of the literature on the subject, we focus on closed surfaces.

Any approach to bounding the cardinality of isospectral sets requires more than just short curves. This is because when surfaces have very thin parts, lengths of curves that transverse these parts are long and necessary to identify the moduli of the thin parts. Identifying possible geometries of the thin part of surfaces requires another type of approach. Buser's approach to this is to treat "very" short curves differently from just "somewhat" short curves and in both cases provides involved, although mostly elementary, arguments. The approach taken here is quite different.

First of all, we only differentiate between long and short curves where short means those of length less than $2\operatorname{arcsinh}(1)$. This constant is a natural quantity when dealing with simple closed geodesics because is the exact value beyond which, on any complete hyperbolic surface, one can guarantee that any closed geodesic that intersects it is of greater length. To deal with a short curve, we first examine the topological types of the shortest curve transversal to it. Depending on their type, we employ two strategies. The first uses a recent theorem of Przytycki [16] which bounds the number of arcs that pairwise intersect at most once. The second strategy, perhaps somewhat unexpectedly, uses McShane type identities [12, 13, 18] as a measure on embedded pairs of pants. Estimates related to the lengths of curve and chain systems are then plugged in resulting on bounds on the size of isospectral sets.

Organization. A preliminary section mostly contains standard results, sometimes adapted slightly for our purposes. This is followed by a full section on curve and chain systems culminating in the proof of Theorem 1.3. Section 4 is then dedicated to the different approaches used to deal with thin structures (although certain aspects concerning McShane identities are delayed to the appendix because they are of somewhat different nature to the rest). The results of the previous two sections are put together in Section 5 to prove Theorem 1.2. Finally, Section 6 is mainly the proof of Theorem 1.1.

2. Preliminaries

The moduli space $\mathcal{M}_g$ is the space of complete hyperbolic structures on a closed orientable topological surface Σ_g of genus $g \geq 2$ up to isometry. In order to describe $\mathcal{M}_g$, it is often useful to use lengths of closed geodesics. Given an element of $\xi \in \pi_1(\Sigma_g)$ and any $X \in \mathcal{M}_g$, there is a unique closed geodesic $\gamma_\xi \subset X$ in the free homotopy class of ξ. In that way, for every ξ, one gets a function $\ell_. : \mathcal{M}_g \to \mathbb{R}^+$ where to any X one attributes the value

$$\ell_X(\xi) := \ell(\gamma_\xi)$$

This one-to-one correspondence between elements of the fundamental group and closed geodesics is very useful and in general we won't distinguish between a homotopy class and its geodesic realization. In fact, unless specifically stated, when the surface X is clear from the context, $\ell(\gamma)$ will generally mean the length of the closed geodesic in the homotopy class of γ.

A closed geodesic is primitive if is not the iterate of a another closed geodesic. A primitive closed geodesic is called simple if doesn't have any self-intersection points. We denote by

$$\Lambda(X) := \{\ell_1 \leq \ell_2 \leq \ldots\}$$

the ordered (but unmarked) set of lengths of primitive closed geodesics of X, counted with multiplicity. This means that if X has n closed geodesics of length l, l will appear n times. As X is of finite type, the set $\Lambda(X)$ is always a discrete set. As we'll be trying to say things about a surface that has a given spectrum, we'll sometimes use Λ by itself to denote a length spectrum without knowledge of the underlying surface. We note that everything we do could be done for the full length spectrum instead (thus including the lengths of non-primitive elements) as one is determined by the other, but we use primitive here for convenience. We say that X and Y are isospectral if $\Lambda(X) = \Lambda(Y)$.

Cutting and pasting techniques will be used throughout the paper. For this, the following convention will be used. Given a simple closed geodesic α of a surface X, we can remove α from X. The result is an open surface (possibly disconnected) whose geometric closure has two boundary curves. We denote this closed manifold with boundary $X \setminus \alpha$. Similarly, we denote $X \setminus \mu$ for the corresponding geometric closure when μ is a geodesic multicurve.

The following result [9], see [5, Thm. 4.1] for this version, will be useful for our purposes.

Lemma 2.1 (Collar lemma). *Given such two disjoint simple closed geodesics $\gamma_1, \gamma_2 \subset X$, their collars*

$$\mathcal{C}(\gamma_i) := \{x \in X : d_X(x, \gamma_i) \leq w_i\}$$

of widths respectively

$$w_i := \operatorname{arcsinh}\left(\frac{1}{\sinh(\ell(\gamma_i)/2)}\right)$$

for $i = 1, 2$ are embedded cylinders and disjoint.

We can compute the lengths of the boundary curves of the collar, which are *not* geodesic, as follows.

Lemma 2.2. *The boundary lengths of the collar $\mathcal{C}(\gamma)$ for a simple closed geodesic γ are both equal to*

$$\ell(\gamma)\coth\left(\frac{\ell(\gamma)}{2}\right)$$

Proof. The length of the boundary of the set of points at distance r from a curve of length ℓ in the hyperbolic plane is $\ell\cosh(r)$ so here, as everything is embedded in X, the length of both boundary curves are

$$\ell(\gamma)\cosh\left(\operatorname{arcsinh}\left(\frac{1}{\sinh(\ell(\gamma)/2)}\right)\right) = \ell(\gamma)\coth\left(\frac{\ell(\gamma)}{2}\right)$$

as claimed. $\qquad\square$

For a given surface X, the set of closed geodesics of length less than $2\operatorname{arcsinh}(1)$ will be denoted $\Gamma_0(X)$. Non-simple closed geodesics are all of length at least $2\operatorname{arcsinh}(1)$ (see [5, Theorem 4.2.2]) so $\Gamma_0(X)$ consists of simple closed geodesics. As a corollary of the collar lemma and of the previous lemma, we have the following about elements of $\Gamma_0(X)$.

Corollary 2.3. *Given distinct $\gamma_1, \gamma_2 \in \Gamma_0(X)$ their collars satisfy*

$$\mathcal{C}(\gamma_1) \cap \mathcal{C}(\gamma_2) = \varnothing$$

Furthermore, if $\gamma \in \Gamma_0$, the boundary curves of its collar $\mathcal{C}(\gamma)$ have length b where

$$2 \le b \le 2\sqrt{2}\log(1+\sqrt{2})$$

Proof. The first statement follows directly from Lemma 2.1. The second statement follows from the monotonicity in ℓ of the function $\ell\coth\left(\frac{\ell}{2}\right)$ between 0 and $2\operatorname{arcsinh}(1)$. $\qquad\square$

In particular, geodesics in $\Gamma_0(X)$ are disjoint. Thus there are at most $3g - 3$ closed geodesics of length at most $2\operatorname{arcsinh}(1)$. The other ones are more difficult to count, and the following lemma, based on a lemma of Buser [5, Lemma 6.6.4], is a useful tool for this. Note the statement is not exactly the same but the proof contains all of the ingredients we'll need.

6

Lemma 2.4 (Counting closed geodesics). *Let $L > 0$ and $X \in \mathcal{M}_g$. Then there are at most*

$$(g-1)\, e^{L+6}$$

primitive closed geodesics of length most L.

Proof. The statement is not exactly the statement from [5, Lemma 6.6.4] but can be obtained by the proof as follows.

There are at most $3g - 3$ curves of length $2\operatorname{arcsinh}(1)$ and now we count the longer curves.

In the proof of [5, Lemma 6.6.4], Buser shows that the number of closed geodesics that are *not* iterates of the curves of length less than $2\operatorname{arcsinh}(1)$ are bounded above by

$$\frac{\cosh(L + 3r) - 1}{\cosh(r) - 1} \frac{2(g-1)}{\cosh(r/2) - 1}$$

where $r = \operatorname{arcsinh}(1)$. This bound is obtained by covering the thick part of the surface by balls of radius $\operatorname{arcsinh}(1)$ and by counting geodesic loops based in the centers of the balls above and using an area comparaison argument.

Now the total number of primitive curves is thus bounded by

$$3g - 3 + \frac{\cosh(L + 3r) - 1}{\cosh(r) - 1} \frac{2(g-1)}{\cosh(r/2) - 1} = (g - 1)\left(3 + \frac{2\cosh(L + 3r) - 2}{(\cosh(r) - 1)(\cosh(r/2) - 1)}\right)$$

and a straightforward calculus computation shows that this quantity is bounded above by

$$(g - 1)\, e^{L+6}$$

for all $g > 2$ and $L > 0$ as desired. $\qquad\square$

The above statement is in fact slightly weaker than the statement in [5] but it will be exactly the statement we need as we're dealing with the primitive length spectrum.

Remark 2.5. The number of closed geodesics of length at most L grows asymptotically like e^L/L, a result due to Huber [7], generalized to many different contexts by Margulis [10]. Although the bound in Lemma 2.4 does not exhibit the correct order of growth, it has the advantage of being effective and working for *any L*. We also note that although the asymptotic growth of the number of simple closed geodesics is considerably slower (it is polynomial in L, a result of Mirzakhani [14]), the above lemma, for sufficiently small values of L, provides an effective upper bound on the growth which is interesting even for the simple geodesics.

Another ingredient will be the following result, due to Bavard [1], which can be used to find short curves on surfaces. It will be a useful tool and so we state it as a lemma. The only thing we really need is an upper bound on the length of non-trivial curve that grows like $2\log(g)$ which can be directly obtained by an area comparison argument. This result, remarkably, is sharp and so we state it precisely.

Lemma 2.6. *] For any $X \in \mathcal{M}_g$ and any $x \in X$, there exists a geodesic loop δ_x based in x such that*

$$\ell(\delta_x) \leq 2\operatorname{arccosh}\left(\frac{1}{2\sin\frac{\pi}{12g-6}}\right)$$

For future reference we denote

$$R_g := \operatorname{arccosh}\left(\frac{1}{2\sin\frac{\pi}{12g-6}}\right)$$

That such a bound exists is relatively straightforward: the area of a ball in the hyperbolic plane grows exponentially in its radius whereas the area of a surface of genus g is $4\pi(g-1)$ thus the radius of an embedded ball cannot exceed $\log(g)$ by some large amount.

3. Curve and chain systems

We'll begin by describing a decomposition of a surface obtained by cutting along simple closed geodesics and simple orthogeodesics between them. This will lead to what will be called an *curve and arc* decomposition. We'll then show how the arcs relate to a family of curves to obtain *curve and chain systems*. The lengths of these curves will help determine isometry types of surfaces.

3.1. Definition and topological types

Definition 3.1. A curve and arc decomposition of $X \in \mathcal{M}_g$ is a non-empty collection of disjoint simple closed geodesics $\Gamma = \gamma_1, \ldots, \gamma_k$ and a collection of arcs $\mathcal{A} = a_1, \ldots, a_{6g-6}$ on $X \setminus \Gamma$ such that $X \setminus \{\Gamma, \mathcal{A}\}$ is a collection of geodesic right angled hexagons.

We shall refer to the $\gamma_1, \ldots, \gamma_k$ as the curves and the $a_1, \ldots, a_{6g-6}$ as the arcs (or as orthogeodesics, as they are orthogonal to the curves in their endpoints).

Example. A first example of curve and arc decomposition can be obtained by taking Γ to be a pants decomposition of X (a collection of disjoint simple closed geodesics that decompose X into three holed spheres or pairs of pants). The set of arcs $\mathcal{A}$ is a collection of three disjoint simple orthogeodesics between boundary curves on each pair of pants.

8

Observe that, although the definition given is geometric, it could have been made purely topological as follows. The set Γ is a set of disjoint non-isotopic essential simple closed curves and $\mathcal{A}$ is a maximal set of disjoint simple arcs on $X \setminus \Gamma$ with endpoints on Γ. Further note that any non-empty set Γ of disjoint simple curves can be completed into a curve-arc decomposition by taking $\mathcal{A}$ to be any maximal set of arcs on $X \setminus \Gamma$ with endpoints on Γ.

We'll need to count the number of different topological types of curve and arc decompositions that one can have in genus g. For future use, it will be useful to have a marking on the curves Γ and arcs $\mathcal{A}$. This is simply a labelling of the curves and arcs and we'll refer to these as marked curve and arc decompositions. Two marked curve and arc decompositions are topologically equivalent if there is a homeomorphism between them which respects the marking on the curves and arcs.

Lemma 3.2. *The number of different topological types of marked curve and arc decompositions is bounded above by*

$$\frac{1}{e^6} \left(\frac{12^6}{e^5} \right)^{g-1} (g-1)^{6g-6}$$

Proof. A curve and arc decomposition is a decomposition of X into $4g - 4$ hexagons. We see X as the result of a two step construction. First we paste together the hexagons to obtain $X \setminus \Gamma$ and then we paste together the boundary curves of $X \setminus \Gamma$ to obtain X.

We begin with a set of $4g - 4$ hexagons, each with a set of three marked non-adjacent oriented side arcs. For instance, we put the hexagons all in the plane and give the orientation to arcs induced by a fixed orientation of the plane. We'll get rid of the orientation on the arcs later, but it will be useful to retain it for the time being. The hexagons have a "frontside" and a "backside". We're going to paste the hexagons together to obtain an orientable two-sided surface where the front sides of each side hexagon are all on the same side.

To do so take a first side arc and paste it to another. Note there is a unique way of pasting so that both front sides of the hexagon are on the same side. We retain, as orientation for the resulting arc, the orientation of the first arc.

In all there are $12g - 12$ side arcs. That means that there are

$$(12g - 12)!! = (12g - 11) \cdot (12g - 9) \cdots 6 \cdot 3 = \frac{(12g - 12)!}{2^{12g - 12/2}(12g - 12/2)!}$$

ways of doing this. Using standard inequalities on factorials gives at most

$$\frac{(12g - 12)!}{2^{6g-6}(6g - 6)!} \leq \frac{(12g - 12)^{(12g-12+1/2)}}{2^{6g-6}e^{12g-11}} \frac{e^{6g}}{\sqrt{2\pi}(6g - g)^{6g-6+1/2}}$$

which after further simplifications becomes

$$\frac{1}{e^6} \left(\frac{12^6}{e^5} \right)^{g-1} (g-1)^{6g-6}$$

Now the result is a surface with boundary curves and marked oriented arcs. The marked oriented arcs mark the boundary curves, so in particular, we also have all necessary information to perform the second step of our process (the pasting of the curves Γ) without any further counting. As we don't care about orientations we can forget them. Certainly we've counted every topological type of marked curve and arc decomposition and the lemma is proved. $\qquad\square$

The main purpose of this counting lemma will be to count curve and chain systems which will be introduced in the next section.

3.2. Coordinates for moduli space

The goal is to use geometric quantities to determine the isometry class of surfaces. These coordinates are somewhat similar to Fenchel-Nielsen coordinates (but, unlike the latter, some of ours will be redundant).

Like for Fenchel-Nielsen coordinates it will be necessary to consider twist parameters, but only along curves of Γ. To do so we consider a way of marking points on each side of a curve of Γ. An example of how to do this is to choose, for every $\gamma \in \Gamma$ and each side of γ, an endpoint of an arc $a \in \mathcal{A}$ which has is attached to γ and which "leaves" on the corresponding side. The specific marked point associated to a metric structure $X \in \mathcal{M}_g$ is obtained by taking the geodesic realizations of γ and a (where a is now an orthogeodesic) and taking the appropriate intersection point between a and γ (there might be two of them). In this way, for every $\gamma \in \Gamma$ and every $X \in \mathcal{M}_g$, one has two marked points, say p_γ^+ and p_γ^-, one for each side of γ.

Definition 3.3. The twist parameter along γ is the signed distance between p_γ^+ and p_γ^-.

We denote by $\tau(\gamma)$ the twist parameter of γ. Similarly, $\tau(\Gamma)$ is the set of (marked) twist parameters of Γ. Similarly, $\ell(a)$ is the length of the unique orthogeodesic in the free homotopy class of a and we'll denote by $\ell(\mathcal{A})$ the set of (marked) lengths of $\mathcal{A}$.

Lemma 3.4. *The parameters $\tau(\Gamma)$ and $\ell(\mathcal{A})$ uniquely determine $X \in \mathcal{M}_g$.*

Proof. The length of the arcs of $\mathcal{A}$ determine the geometry of each of the hexagons. They in turn determine the lengths of Γ. The only thing remaining is how to determine how the elements of Γ are pasted but this is determined by $\tau(\Gamma)$. $\qquad\square$

Recall our goal is to relate the length spectrum to isometry classes of surfaces so although the curve-arc length parameters are convenient, we'd like a set of parameters which only use curves. To do so we replace the lengths of arcs by the lengths of curves as follows.

First given $a \in \mathcal{A}$, we define a free homotopy class (or equivalently a closed geodesic) as follows. Give a and $X \setminus \Gamma$ an orientation (this defines an orientation on the end geodesics of a, say γ_1 and γ_2). Now let γ_a be the closed geodesic in the free homotopy class of

$$\gamma_1 * a * \gamma_2 * a^{-1}$$

Note that if a is arc between distinct curves γ_1 and γ_2, then γ_a is simple. Otherwise it is a closed geodesic with two self-intersection points as in the right side of Figure 1. (Here distinct curves means distinct curves on $X \setminus \Gamma$: they could very well correspond to the same curve on X.)
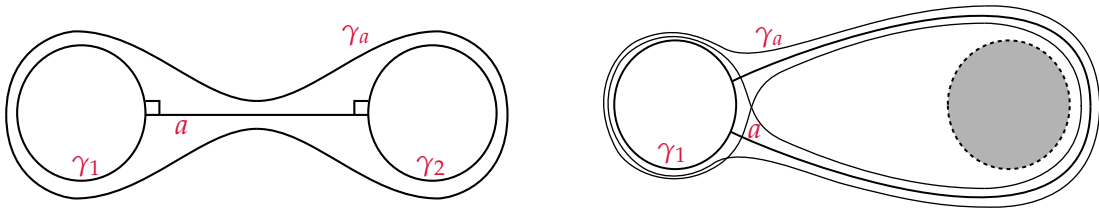

Figure 1: Different types of chains and the associated embedded and immersed pants

Lemma 3.5. *Let $a \in \mathcal{A}$ be an arc between $\gamma_1, \gamma_2 \in \Gamma$. Then $\ell(a)$ is determined by $\ell(\gamma_1), \ell(\gamma_2)$ and $\ell(\gamma_a)$.*

Proof. Note that if γ_1 and γ_2 are distinct, then they together with γ_a are the three boundary curves of an embedded pair of pants. If they are not distinct then they are still the boundary curves of a pair of pants, but this time it's immersed and not embedded. The two cases are illustrated in Figure 1.

In both cases, we can argue inside the pair of pants and use a standard fact from hyperbolic trigonometry that tells you that three lengths determine a right angled hyperbolic hexagon. $\square$

Putting these two previous lemmas together, we have the following proposition.

Proposition 3.6. *The quantities $\ell(\gamma), \tau(\gamma)$ for $\gamma \in \Gamma$ and $\ell(\gamma_a)$ for $a \in \mathcal{A}$ determine a surface $X \in \mathcal{M}_g$.*

For future reference, we'll refer to the curve γ_a for $a \in \mathcal{A}$ as the *chain* associated to a. The set of curves $\Gamma, \Gamma_{\mathcal{A}} := \{\gamma_a\}_{a \in \mathcal{A}}$ will be referred to as a *curve and chain system*.

We remark that the number of marked topological curve and chain systems, which we'll denote by $N_{cc}(g)$, is equal to the number of topological curve and arc decompositions with marked arcs. Thus, by Lemma 3.2 we have the following.

Lemma 3.7. *The number $N_{cc}(g)$ of topological types of marked curve and chain systems is bounded above by*

$$\frac{1}{e^6} \left(\frac{12^6}{e^5} \right)^{g-1} (g-1)^{6g-6}$$

3.3. Curve and chain systems of bounded length

In this section we prove the existence of curve and chain systems of lengths bounded above by a function of topology (Theorem 1.3).

We'll need to bound the distance between a geodesic loop based in a point and the collar neighborhood of the corresponding simple closed geodesic (the core curve). This comes up in the following situation. If the core curve is not too short, say greater than 1 for instance, then there is a bound on the (Hausdorff) distance between the loop and the closed geodesic than only depends on the length of the loop (see [15, Lemma 2.3] for a precise statement). However, if the core curve is arbitrarily short, the loop can be arbitrarily far away. The following lemma gives a bound, that only depends on the length of the loop, on the distance between the collar of the core curve and the loop.

Lemma 3.8. *Let $c \subset X$ be a geodesic simple loop and γ be the unique simple closed geodesic freely homotopic to c. Then*

$$\sup_{p \in c}\{d_X(p, \mathcal{C}(\gamma))\} < \log \left(\sinh \left(\frac{\ell(c)}{2} \right) \right)$$

Proof. The proof is a straightforward hyperbolic trigonometry computation. The loop c (based in a point p) and γ form the boundary curves of an embedded cylinder in X. The cylinder can further be decomposed into two isometric quadrilaterals with three right angles (sometimes called trirectangles or Lambert quadrilaterals) with opposite sides of lengths $\ell(c)/2$ and $\ell(\gamma)/2$ as in Figure 2.

The collar around γ corresponds to the shaded region in Figure 2. The segment marked d is an upper bound on the distance between c and $\mathcal{C}(\gamma)$ and this is what we want to bound. It's a subarc of the arc of length $d + w$ joining p to γ where w is the width of $\mathcal{C}(\gamma)$. Appealing
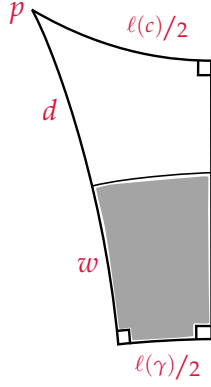
Figure 2: The quadrilateral

to hyperbolic trigonometry in the quadrilateral one obtains

$$\sinh\left(\frac{\ell(c)}{2}\right) = \sinh\left(\frac{\ell(\gamma)}{2}\right)\cosh\left(d + \operatorname{arcsinh}\left(\frac{1}{\sinh\left(\ell(\gamma)/2\right)}\right)\right)$$

From this we obtain

$$d = \operatorname{arccosh}\left(\frac{\sinh\left(\frac{\ell(c)}{2}\right)}{\sinh\left(\frac{\ell(\gamma)}{2}\right)}\right) - \operatorname{arcsinh}\left(\frac{1}{\sinh\left(\ell(\gamma)/2\right)}\right)$$

This becomes

$$
\begin{aligned}
d &= \log\left(\frac{\sinh\left(\frac{\ell(c)}{2}\right)}{\sinh\left(\frac{\ell(\gamma)}{2}\right)} + \sqrt{\frac{\sinh^2\left(\frac{\ell(c)}{2}\right)}{\sinh^2\left(\frac{\ell(\gamma)}{2}\right)} - 1}\right) - \log\left(\frac{1}{\sinh\left(\frac{\ell(\gamma)}{2}\right)} + \sqrt{\frac{1}{\sinh^2\left(\frac{\ell(\gamma)}{2}\right)} + 1}\right) \\
&< \log\left(\frac{2\sinh\left(\frac{\ell(c)}{2}\right)}{\sinh\left(\frac{\ell(\gamma)}{2}\right)}\right) - \log\left(\frac{2}{\sinh\left(\frac{\ell(\gamma)}{2}\right)}\right) \\
&= \log\left(\sinh\left(\frac{\ell(c)}{2}\right)\right)
\end{aligned}
$$

as desired. □

We now prove the existence of short chain and curve systems, the main result of this section.

Proof of Theorem 1.3. Consider $X \in \mathcal{M}_g$. We begin by considering the set of curves Γ_0 of X of length at most $2\operatorname{arcsinh}(1)$. By Corollary 2.3, their collars are disjoint. We now consider

$$X_0 := X \setminus \{\mathcal{C}(\gamma) \mid \gamma \in \Gamma_0\}$$

This set may not be connected.

13

We'll now iterate the following step starting with $k = 0$: we choose a point $x \in X_0$ such that

$$d_{X_k}(x, \partial X_k\} \leq \log(4g)$$

We consider the shortest non-trivial loop δ_x based in x. By Lemma 2.6 we have a bound on its length:

$$\ell(\delta_x) \leq 2R_g$$

It's straightforward to check that $R_g < \log(4g)$.

We consider the unique simple closed geodesic δ (on X) freely homotopic to δ_x. By Lemma 3.8, the distance between x and $\mathcal{C}(\delta)$ satisfies

$$d_{X_k}(x, \mathcal{C}(\delta)) < \log(\sinh(R_g)) < \log(2g)$$

where the last inequality is the result of comparing the two functions by standard manipulations. In particular, this implies that δ is contained in X_k and is not a boundary curve of X_k. We then set

$$X_{k+1} := X_k$$

and repeat the procedure until all $x \in X_k$ satisfy $d_{X_k}(x, \partial X_k) < \log(4g)$. The disjoint set of curves we've cut X along (which include those of Γ_0) are denoted Γ. For all $\gamma \in \Gamma$ we have

$$\ell(\gamma) < 2\log(4g)$$

as desired.

We now consider

$$X' := X_0 \setminus \Gamma$$

and consider a Voronoi cell decomposition of X around the curves of Γ. This is simply the attribution of (at least) one element of Γ to every $x \in X'$ by taking the curve (or curves) of Γ closest to x. The cells of the decomposition are

$$C_\gamma := \{x \in X' \mid d_{X'}(x, \gamma) \leq d_{X'}(x, \delta) \text{ for all } \delta \in \Gamma\}$$

Note that by construction

$$d_{X'}(x, \gamma) \leq \log(4g)$$

for all $x \in C_\gamma$.

The points that lie in several cells we refer to as the cut locus of the decomposition. It's an embedded graph, and generically the graph is trivalent. We want to find a decomposition of X' into hexagons dual the cell decomposition as follows. This process is completely

14

analogous to the construction of a Delaunay triangulation for a choice of points in the plane for instance.

Dual to each edge of the cut locus we construct an arc between the corresponding curves of Γ as in Figure 3.
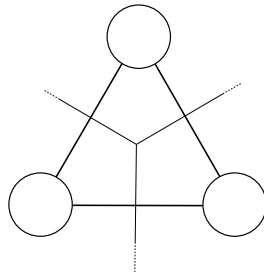

Figure 3: Constructing arcs dual to the cut locus

If the cut locus is trivalent (which it is generically) then the resulting arc decomposition is maximal in that all remaining (non trivial and non homotopic) arcs essentially cross these. As such, this provides a decomposition into hexagons.

If the cut locus is not trivalent, there are choices to be made (just like when a set of points in the plane admits several Delaunay triangulations). To do so, in any vertex v of degree $k \geq 4$ of the cut locus, consider the set of simple arcs $\{c_i\}_{i=1}^{k}$ to each of the boundary curves of X' whose Voronoi cells touch v. Each arc is contained in the corresponding cell and we suppose that they are cyclically oriented around v. Now fix one of these arcs, say c_1, and orient it towards v. Orient all of the others away from v and consider the arcs obtained by concatenating c_1 with c_i for $i = 3, \ldots, k-1$ (see Figure 4).
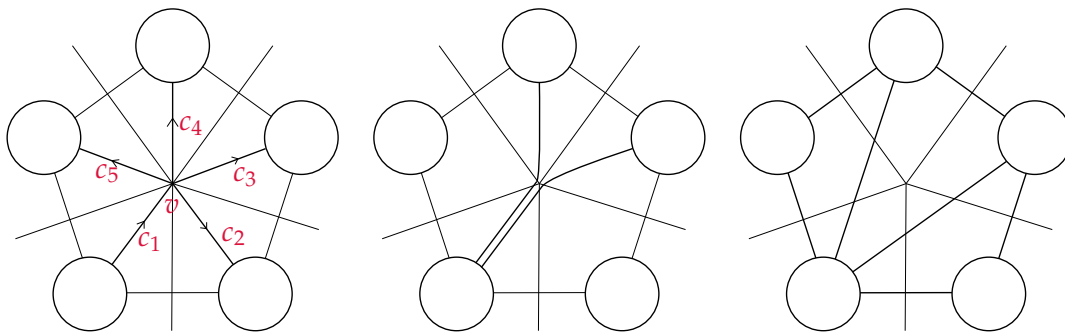

Figure 4: Constructing arcs

Note that the arcs $c_1 * c_2$ and $c_1 * c_k$ are isotopic to arcs dual to the cut locus. We add this set of $k-3$ arcs and repeat the process in all vertices of degree higher than 3.

15

Once we've chosen these arcs, we take the minimizers that join the corresponding boundary curves in X' and these are simple orthogeodesics the set of which we denote by $\mathcal{A}$. By construction, any $a \in \mathcal{A}$ satisfies

$$\ell(a) \leq 2\log(4g)$$

Now consider the curves γ_a for $a \in \mathcal{A}$ defined in Section 3.2. All that remains to show is that they have length bounded by $8\log(4g)$. In particular, for $a \in \mathcal{A}$, let the end geodesics of a be γ_1 and γ_2, and by construction they are both of length at most $2R_g$. Recall γ_a is in the free homotopy class of $\gamma_1 * a * \gamma_2 * a^{-1}$ and because $\ell(a) \leq \log(2g)$, we have

$$\ell(\gamma_a) < 4R_g + 4\log(4g) < 8\log(4g)$$

proving the result. $\qquad\square$

Before passing to the next section, we prove a lemma that we will need to control twist parameters. Here the sets Γ and $\Gamma_{\mathcal{A}}$ are the short curve and chain system from the theorem above. The goal is to do the following: for each $\gamma \in \Gamma_1 := \Gamma \setminus \Gamma_0$ we want to choose a transversal curve δ_γ that is not too long. The specific result we prove is the following.

Lemma 3.9. *For each $\gamma \in \Gamma \setminus \Gamma_0$ there exists a simple closed geodesic δ_γ such that $i(\gamma, \delta_\gamma) \leq 2$ and*

$$\ell(\delta_\gamma) < 14\log(4g)$$

Furthermore there is a choice of such a δ_γ that is only determined by the topological type of a marked curve and chain system.

Proof. Consider $\gamma \in \Gamma \setminus \Gamma_0$ and consider its two copies on $X_0 \setminus \Gamma$, say γ_1 and γ_2. Note γ can be given an orientation given by an orientation of the surface.

If there is an arc $a \in \mathcal{A}$ joining γ_1 to γ_2, then consider the closed curve δ_γ on X obtained by concatenating a with the oriented sub-arc of γ between the two endpoints of a on γ (see Figure 5).

By construction $\ell(\delta_\gamma) \leq \ell(a) + \ell(\gamma) < 2\log(4g) + 2\log(4g) = 4\log(4g)$ and $i(\gamma, \delta_\gamma) = 1$.

We could possibly do better, by taking the shortest one, but we want to make a choice that only depends on the topology of a marked curve and chain system.

If no such arc exists, then for γ_i, for $i = 1, 2$ we construct an arc that has both endpoints on γ_i. Either there is an arc $a_i \in \mathcal{A}$ which does this, in which case, we use it, or no such arc exists in which case we consider any arc $a_i \in \mathcal{A}$ with an endpoint on γ_i. The other endpoint
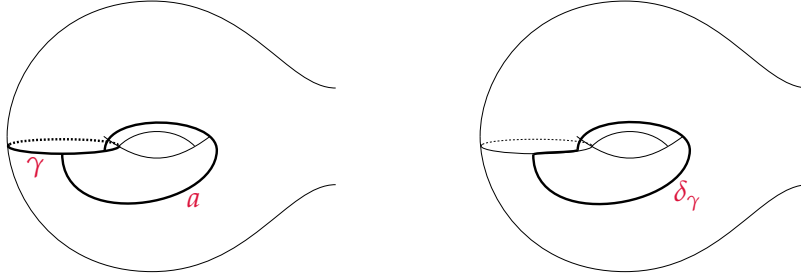
Figure 5: The one holed torus and the construction of δ_γ

of a_i must lie on another $\gamma_i' \in \Gamma$. Giving the arc and γ_i' the appropriate orientation, we construct a homotopy class of arc b_i by considering the concatenation

$$a_i * \gamma_i' * a_i^{-1}$$

Note that b_1, b_2 and γ fill a four holed sphere. We also have

$$\ell(b_i) < 6\log(4g)$$

using the bound on lengths of the concatenated paths. Now the shortest curve that essentially intersects γ exactly twice on this four holed sphere has length at most

$$\ell(b_1) + \ell(b_2) + \ell(\gamma)$$

by the same type of cut-and-concatenate argument as before (see Figure 6).
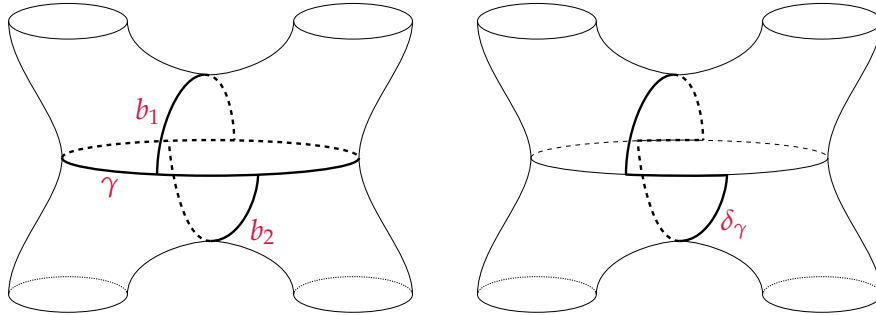

Figure 6: The four holed sphere and the construction of δ_γ

The resulting curve we denote δ_γ and we have

$$\ell(\delta_\gamma) < 14\log(4g)$$

as desired. $\qquad\qquad\square$

17

We'll denote by Δ_{Γ_1} the set of curves obtained as in the above proof. Note they are determined if we know the topological type of the curve chain decomposition $\Gamma, \Gamma_{\mathcal{A}}$.

4. Dealing with thin structures using the topology of arcs and length identities

Here we show how to identify, up to quantifiable finiteness, the isometry class of a surface using the isometry type of its thick part. The reason this requires a different analysis is because we have to determine a twist parameter using curves that could be arbitrarily long. This means we can't use estimates based on their length.

Recall that given $X \in \mathcal{M}_g$, we've denoted by $\Gamma_0(X)$ the set of closed geodesics of length less than $2 \operatorname{arcsinh}(1)$. We denote by

$$X_0 := X \setminus \Gamma_0$$

This is a slight abuse of notation as we used X_0 before for surface obtained by removing the collars of the short curves, but as they determine each other, for simplicity we'll use it here too. The goal is to determine X using X_0 and $\Lambda(X)$, and to do so we'll proceed one curve at a time.

4.1. The topology of the next shortest curve

Let $Y \subset X$ be a subsurface obtained by cutting along some subset of Γ_0 and let $\alpha \subset Y$ that belongs to Γ_0. Set $Y_\alpha := Y \setminus \alpha$. Denote by α_1, α_2 the two copies of α on Y_α. We want to determine Y knowing $\Lambda(Y)$ and Y_α.

Given Y_α and $\Lambda(Y)$, we know the set $\Lambda(Y) \setminus \Lambda(Y_\alpha)$. The first element of this set corresponds to the shortest closed geodesic of Y which is not entirely contained in Y_α. Let us denote the corresponding closed geodesic α' and observe that necessary $i(\alpha, \beta) \neq 0$. First we discuss the topology of α'.

Lemma 4.1. *The curve α' is a simple closed geodesic which is one of two types:*

- Type (I): $i(\alpha, \alpha') = 1$ and the two curves fill a one holed torus,

- Type (II): $i(\alpha, \alpha') = 2$ and the two curves fill an embedded four holed sphere.

Proof. This will all follow from standard cut and paste arguments and the fact that $\ell(\alpha) \leq 2 \operatorname{arcsinh}(1)$.

Consider the image of α' on Y_α is a collection of arcs $a_1, \ldots, a_m$ with endpoints on $\alpha_1 \cup \alpha_2$. In the course of the proof, we'll show that the arcs are all simple. Note that if there is only one and it is simple, then it lies between α_1 and α_2, thus $i(\alpha, \alpha') = 1$ and α' is of type (I).

18

Note that because $\ell(\alpha) \leq 2\operatorname{arcsinh}(1)$, by Lemma 2.1 there is an embedded collar of width $\operatorname{arcsinh}(1)$ around α. In particular, any of these arcs is of length at least $2\operatorname{arcsinh}(1)$.

In particular that if there is an arc a_i between α_1 and α_2, then it is the only arc. Indeed, if there was another one, then $\ell(\alpha') > 2\operatorname{arcsinh}(1) + \ell(a_i)$. The distance between α_1 and α_2 is at most $\ell(a_i)$. Denote by a a shortest distance path between α_1 and α_2. Note that a is simple and $\ell(a) \leq \ell(a_i)$.

Now by concatenating a and a subarc of α, one can construct an essential curve α'' that intersects α exactly once. And as $\ell(\alpha) \leq 2\operatorname{arcsinh}(1)$, this implies that

$$\ell(\alpha'') \leq 2\operatorname{arcsinh}(1) + \ell(a) \leq 2\operatorname{arcsinh}(1) + \ell(a_i) < \ell(\alpha')$$

a contradiction. This proves that if $a_1, \ldots, a_m$ contains an arc between α_1 and α_2, then $m = 1$, it is simple and α' is of type (I).

If not, then all arcs of $a_1, \ldots, a_m$ have both endpoints on either α_1 or α_2. Denote by $a_1, \ldots, a_{m_1}$ those with endpoints on α_1. Note that $m = 2m_1$ as the number of endpoints of the arcs must be the same on both α_1 and α_2.

Claim: Any arc $a \in \{a_1, \ldots, a_m\}$ must be simple.

Proof of claim: The basic idea is to do surgery in a point of self-intersection, but we want to be careful to ensure that the surgery doesn't produce a trivial arc. First observe that by minimality of α', a is a shortest non-trivial arc between its endpoints (both on α_i for $i \in \{1, 2\}$). Now if a is not simple, it contains as a subarc a simple embedded loop $\hat{a}$ based in a point p (this is true of any non-trivial non-simple arc). So the two subarcs of a between α_i and p are distance realizing and hence simple and disjoint. Denote them by d_1 and d_2. We now give orientations to a, d_1, d_2 and $\hat{a}$, so that $a = d_1 * \hat{a} * d_2$. The arc $a' = d_1 * \hat{a}^{-1} * d_2$ is non-trivial and the unique geodesic in its homotopy class (with endpoints fixed) is shorter than a, a contradiction and this proves the claim.

For $i = 1, 2$, we consider the shortest non-trivial arc b_i with both of its endpoints on α_i. As above b_1 and b_2 are simple and by minimality they satisfy the following: $\ell(b_1) \leq \ell(a_j)$ for $j \leq m_1$ and $\ell(b_2) \leq \ell(a_j)$ for $j > m_1$.

We now observe that $i(b_1, b_2) = 0$. If not, by concatenating the appropriate subarcs, we can construct a path of length at most $\max_{i \in \{1, \ldots, m\}} \{\ell(a_i)\}$ between α_1 and α_2. The same argument used when there was an arc a_i between α_1 and α_2 can be repeated, and we reach a contradiction. Similarly, if $m_1 > 1$, then we can also reproduce this argument.

To resume, this shows that unless α' is of type (I), the image of α' on Y_α consists of two simple disjoint arcs a_1, a_2 with base points on α_1 and α_2. Thus α, α' fill a four holed sphere

on Y as claimed and α' is of type (II). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We now deal with counting possible homotopy classes for α' of Type (I).

4.1.1. Type (I) next shortest curves

If α' is of type (I), then the projection (or restriction) of α' to Y_α is a simple arc between α_1 and α_2. Denote by a the unique orthogeodesic in its free homotopy class. Given a, we can basically reconstruct α' using the fact that α' is the shortest closed geodesic intersecting α. Indeed, α' is homotopic to the concatenation of the (or a) shortest subarc of α between the two endpoints of a. This is because the length of α' is a function of the length of the subarc of α and the length of a, and is monotonic increasing in both quantities. As there are at most two shortest subarcs of α between the two endpoints of a, knowing a will determine one of two possible homotopy classes for α'.
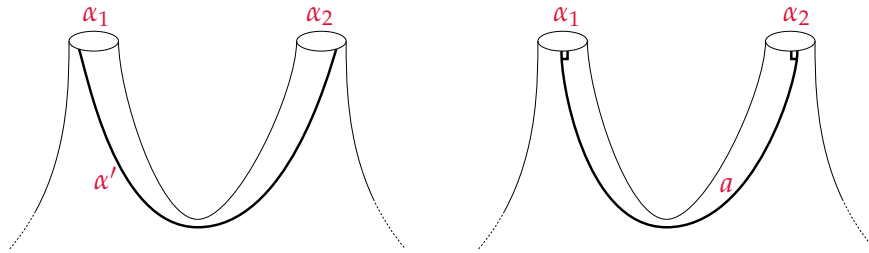

Figure 7: When α' is of type (I)

To resume the above discussion, if we determine all possible homotopy classes of orthogeodesics a, this will allow us to determine all possible homotopy classes of α'.

This observation will allow us to use a result of Przytycki [16].

Proposition 4.2. *There are at most $(8g-8)(2g-1) < 16(g-1)^2$ type (I) next shortest curves hence at most $2(8g-8)(2g-1) < 32(g-1)^2$ isometry classes of Y when α' is of type (I).*

Proof. To show this, we consider all possible surfaces Y_α^t obtained from Y_α by fixing the twist parameter of α to be t ($t \in [0.\ell(\alpha)]$). Among all of these, we consider the subset of them (possibly all or none of them) where α' is of type (I). For given t, we denote by α'_t the curve α'. For each of these we look at the restriction of α'_t to $Y_\alpha \subset Y_\alpha^t$. It is a geodesic arc a_t with endpoints on α_1 and α_2.

This arc a_t has the following property: on Y_α it is the shortest paths between its endpoints for the following reason. If there was a shorter path between its endpoints, this would give

20

rise to a shorter curve on Y_α^t which intersects α exactly once. And, by the bigon property, such a curve is both essential and essentially intersects α, contradicting the minimality of α_t'.

Consider two different such arcs, a_t and a_s coming from curves α_t' and α_s'. Because both are minimal arcs between their endpoints, they cannot cross more than once. Thus $i(a_t, a_s) \le 1$ for all $t, s \in [0, \ell(\alpha)]$.

Now we pass to free homotopy classes, relative to boundary, of arcs, or equivalently to orthogeodesics in the homotopy classes of our collection of arcs. It's a result of Przytycki [16] that their are at most $2|\chi(Y_\alpha)|(|\chi(Y_\alpha)| + 1)$ such homotopy classes. By the discussion above, this gives rise to at most

$$4|\chi(Y_\alpha)|(|\chi(Y_\alpha)| + 1)$$

possible homotopy classes for α'. It's important to note here that we can do more than just bound their number. Our knowledge of Y_α allows us to determine the homotopy classes exactly.

Now to determine Y, for each possible homotopy class of α', we insert the smallest value of $\Lambda(Y) \setminus \Lambda(Y_\alpha)$. By the convexity of geodesic length functions, there are at most 2 possible isometry classes of Y with this length of α'.

Now $\chi(Y_\alpha)| \le \chi(X) = 2g - 2$ and the lemma follows. $\qquad \square$

We will need to argue differently if α' is of type (II).

4.1.2. Type (II) next shortest curves

The projection of α' to Y_α consists of two arcs, freely homotopic to two simple orthogeodesics a_1, a_2 with the endpoints of a_1 on α_1 and the endpoints of a_2 on α_2. As previously, given a_1 and a_2 there are two possibilities for the homotopy class of α'. (Here we have to be more careful: there are more cases to consider depending on the relative positions of the endpoints of a_1 and a_2.)

Note Y_α, the arc a_1 lies in an embedded (geodesic) pair of pants P_1 with α_1 as one of its boundary curves. Similarly, a_2 and α_2 lies in an embedded (geodesic) pair of pants P_2 with α_2 as one of its boundary curves. Using curve and chain systems, we can prove the following statement about the geometry of the pants P_1 and P_2.

Lemma 4.3. *Let δ be a boundary curve of P_1 or P_2. Then*

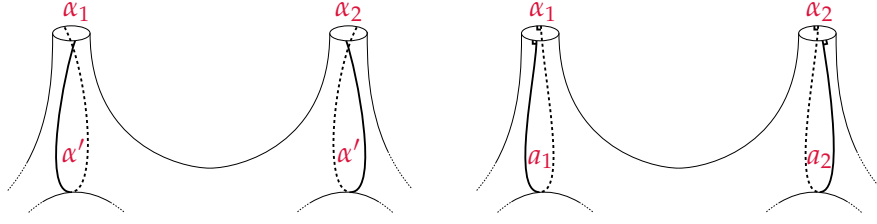$$\ell(\delta) < 6\log(4g) + \operatorname{arcsinh}(1) + 2\sqrt{2}\log(1 + \sqrt{2}) < 6\log(8g)$$

21

Figure 8: When α' is of type (II)

Proof. The lengths of the boundary curves of P_1 and P_2 depend on the lengths of α and the arcs a_1 and a_2.

The arcs a_i may not be the shortest orthogeodesics with endpoints on α_i, but they can't be off by much. In fact, if they are off by more than $\frac{1}{2}\ell(\alpha)$, we can construct a curve that essential intersects α of length less than α'.

Let a'_1 and a'_2 the restrictions of a_1 and a_2 to Y'_α. Note they are still orthogeodesics, even though they lie between non geodesic boundary. Further note that the restriction to Y'_α of any orthogeodesic of Y_α with both endpoints on α_1 or α_2 has its length reduced by exactly $2w(\alpha)$ where $w(\alpha)$ is the width of $\mathcal{C}(\alpha)$. This means that a'_1, resp. a'_2, is not more than $\frac{1}{2}\ell(\alpha)$ longer than the shortest orthogeodesics with endpoints both on α_1, resp. α_2.

In the proof of Theorem 1.3, we used short orthogeodesics to find short curve and chain decompositions. Putting α_1 and α_2 in that context, we found orthogeodesics of length at most $2\log(4g)$ attached to them. If these orthogeodesics don't have both endpoints on α_1, resp. α_2, their second endpoint is on a curve of length at most $2\log(4g)$. Using a concatenation of paths allows one to find an orthogeodesic of length at most $6\log(4g)$ with both endpoints on α_1, resp. α_2.

Thus for both $i = 1, 2$ we have $\ell(a_i) < 6\log(4g) + \frac{1}{2}\ell(\alpha)$.

Now using these orthogeodesics, we can bound the lengths of the boundary curves of P_1 and P_2. For $i = 1, 2$, the boundary curves of P_i are of length at most

$$\ell(\alpha_i) + \ell(a_i)$$

and as $\ell(\alpha) \leq 2\operatorname{arcsinh}(1)$ and $\ell(\alpha_i) \leq 2\sqrt{2}\log(1 + \sqrt{2})$ by Corollary 2.3, this proves the lemma. $\qquad\square$

In order to use this bound on the length of boundary curves of the pants, we have the following lemma. The proof is based on a McShane type identity. It is of a somewhat

different nature from the rest of the article, and possibly of independent interest, so we delay its proof to Appendix A.

Lemma 4.4. *Let Y be a surface with non-empty geodesic boundary and let β be one of the boundary curves. Let $\mathcal{P}$ be the set of all embedded geodesic pairs of pants that have β as a boundary curve and $\mathcal{P}_L$ the subset of $\mathcal{P}$ with the other two boundary curves of length at most L. Then the cardinality of $\mathcal{P}_L$ is at most e^L.*

Note that the bound on the cardinality does not depend on either the topology or the length of β.

We can now use the previous lemmas to prove the following.

Proposition 4.5. *There are at most $2(8g)^{12}$ type (II) next shortest curves, hence at most $4(8g)^{12}$ isometry classes of Y when α' is of type (II).*

Proof. By Lemma 4.3, the arcs a_1 and a_2 are found in embedded pants with boundary curves of length at most $L = 6 \log(8g)$. Now by Lemma 4.4, there are at most e^L such embedded pants. Thus there are at most e^{2L} choices for a_1 and a_2 and hence at most $2e^{2L}$ possible isotopy classes for α'. By the convexity of length functions, this means there are at most $4e^{2L}$ isometry types for Y. $\qquad\square$

4.2. From thick to thin

We now put this all together prove the following result, the main goal of this section.

Theorem 4.6. *Let $\Lambda(Y)$ be a length spectrum for some $Y \in \mathcal{M}_g$. Let X_0 be fixed isometry class of a thick part of a surface in $\mathcal{M}_g$ with k_0 curves in Γ_0. Then there are at most*

$$8^{k_0+1} g^{12 k_0}$$

possible isometry types of X with $X_0 \subset X$ and $\Lambda(X) = \Lambda(Y)$.

Proof. We proceed iteratively, curve by curve in Γ_0, adding one curve at a time and estimating the number of possible isometry types using Propositions 4.2 and 4.5. We begin by knowing the isometry type of X_0, and then choose a curve α in Γ_0, and consider the surface $X_1 := X \setminus \{\Gamma_0 \setminus \alpha\}$. Depending on the type of α', we have bounds on the number of isometry classes for X_1. We then proceed to another curve in Γ_0 and so forth. Note that $X_{k_0} = X$.

It is easy to check that our estimate for the number of isometry types for α' is of type (II) (Proposition 4.5) exceeds the estimate when the curve is of type (I) (Proposition 4.2). As

23

at each step, α' is one or the other, a rough bound is thus at most twice the estimate of $8 \cdot (8g)^{12}$ on the isometry types proved in Proposition 4.5.

Thus at the end of the process, we have at most $8^{k_0+1}g^{12k_0}$ possible isometry types for X_{k_0} which proves the theorem. □

5. Counting isometry types

We can now proceed to proving the full upper bound on the number of isospectral but non isometric surfaces in a given moduli space.

Proof of Theorem 1.2. Let X be isospectral to Y.

Let $\Gamma, \Gamma_{\mathcal{A}}$ be a short curve and chain system of X, the existence of which is guaranteed by Theorem 1.3. As before, $\Gamma_1 := \Gamma \setminus \Gamma_0$ and Δ_{Γ_1} the set of curves transversal to those of γ of length bounded in Lemma 3.9. Note $\Gamma, \Gamma_{\mathcal{A}}$ is one of $N_{cc}(g)$ different topological types. We'll take this into account at the end and for now concentrate on counting all possible isometry classes for $\Gamma, \Gamma_{\mathcal{A}}$ lying in a particular topological type. Recall the curves are marked.

By Lemma 2.4, there are at most $(g-1)e^{L+6}$ different primitive closed geodesics of length less than L.

Let k be the number of curves of Γ. Note that $k \leq 3g-3$. In particular, by Theorem 1.3, the k lengths of the curves of Γ are among a set of at most

$$(g-1)e^{2\log(4g)+6} = (g-1)e^{\log(16e^6g^2)} = 16e^6(g-1)g^2$$

lengths. A (crude) upper is that there are at most $\left(16e^6(g-1)g^2\right)^k$ choices for these curves. Similarly, the lengths of the $6g-6$ curves of $\Gamma_{\mathcal{A}}$ are among a collection of

$$(g-1)e^{8\log(4g)+6} = 16e^6(g-1)g^8$$

lengths and thus there are at most $\left(16e^6(g-1)g^8\right)^{6g-6}$ choices for these.

These choices of lengths will determine the isometry type of $X \setminus \Gamma$ but we need more information to determine X_0. For this we'll use the curves Δ_{Γ_1}. Using Lemma 3.9, we know there are at most

$$(g-1)e^{14\log(4g)+6} = 16e^6(g-1)g^{14}$$

choices for these curves. So if there are $k_1 \leq k$ curves in Γ_1, this gives at most

$$\left(16e^6(g-1)g^{14}\right)^{k_1}$$

24

However, unlike before, the choice of length might not uniquely determine the isometry type. Fortunately, length is convex along twists. Thus for $\gamma \in \Gamma_1$ and δ_γ its transversal curve, the quantity $\ell(\delta_\gamma)$ will determine τ_γ up to two possibilities. When counting possible isometry type, we must thus multiply by an additional factor of 2^{k_1}.

All in all, this shows that there are at most

$$N_{cc}(g) \cdot \left(16e^6(g-1)g^2\right)^k \cdot \left(16e^6(g-1)g^8\right)^{6g-6} \cdot \left(16e^6(g-1)g^{14}\right)^{k_1} \cdot 2^{k_1}$$

possible isometry types for X_0.

We now use the results on thin surfaces to conclude. By Theorem 4.6, given X_0, there are at most

$$8^{k_0+1} g^{12k_0}$$

possible isometry types for X where k_0 is the cardinality of Γ_0. That gives us a total of

$$N_{cc}(g) \cdot \left(16e^6(g-1)g^2\right)^k \cdot \left(16e^6(g-1)g^8\right)^{6g-6} \cdot \left(16e^6(g-1)g^{14}\right)^{k_1} \cdot 2^{k_1} \cdot 8^{k_0+1} g^{12k_0} \quad (1)$$

of possible isometry types for X, with the condition that $k_0 + k_1 = k \leq 3g - 3$.

Unfortunately, we have to get into some messy estimations using our previous estimates for $N_{cc}(g)$. As much as possible, choices will be made in terms of their simplicity and not in terms of optimality.

Using Lemma 3.2 and some obvious simplifications, the quantity 1 becomes

$$\frac{8}{e^6} \left(\frac{12^6}{e^5}\right)^{g-1} (g-1)^{6g-6} \left(256e^6(g-1)g^2\right)^k \cdot \left(16e^6(g-1)g^8\right)^{6g-6} \cdot \left(32e^6(g-1)g^{14}\right)^{k_1} \cdot g^{12k_0}$$

which further simplifies to

$$\frac{8}{e^6} \left(\frac{(12 \cdot 16 \cdot e^6)^6}{e^5}\right)^{g-1} \left(256e^6\right)^k \left(32e^6\right)^{k_1} \cdot (g-1)^{12g-12+k+k_1} g^{48g-48+2k+14k_1+12k_0}$$

Now using $k_1 + k_0 \leq k \leq 3g - 3$ and further simplications the above expression is bounded above by an expression of the form

$$\frac{8}{e^6} A^{g-1} g^{B(g-1)}$$

where $A = 3^6 2^{75} e^{67}$ and $B = 114$.

The main point is that the above bound is of the type g^{Cg} for some C. Asymptotically it is certainly dominated by g^{115g} for large enough g, but for a cleaner expression, a small manipulation shows that for all $g \geq 2$ it is bounded above by g^{154g} which proves the result. $\qquad \square$

6. Length spectra interrogations

In this final section we'll show that it suffices to ask an unknown spectrum a quantifiable finite number of questions to determine it uniquely. In some sense, most of the hard work has already been done.

We recall the setup: we have an unknown length spectrum Λ of a genus g surface which we want to determine by asking questions.

An *admissible question* is the following: what is the first value of $\Lambda \setminus \mathcal{L}$ where $\mathcal{L}$ is a finite list of values?

We want to determine the minimum number of questions we need to ask to determine Λ.

The basic strategy is as follows. By the results on lengths of curve and chain systems, for some polynomial $p(g)$, it will suffice to know the first $p(g)$ lengths to determine all possible isometry types of the thick part of the underlying surface (the number of these being quantified). For each isometry type of thick surface, by asking an additional question for each short curve, we can determine all possible isometry types for the whole surface, again the number of these being quantified. Thus with a polynomial number of questions we can determine all possible isometry types, the number of these being at most g^{Cg} for some C. We regroup isospectral isometry types and use additional questions to distinguish between them.

We can now proceed to the proof of the main result.

Proof of Theorem 1.1. To begin, we interrogate the beginning of the length spectrum Γ to determine enough lengths to reconstruct the thick part of the surface up to computable finiteness. By the methods and proof of Theorem 1.2, the lengths of curves of length at most $14 \log(4g)$ determine the thick part of a surface. Again by Lemma 3.9, these lengths are among the first

$$(g-1)\, e^{14 \log(4g)+6} = 16 e^6 (g-1) g^{14}$$

lengths of Γ. Thus the first set of admissible questions is to ask for the first $16 e^6 (g-1) g^{14}$ lengths. This gives us knowledge of the thick part of the surface up to computable finiteness, corresponding to all possible curve configurations.

Now to deal with the full surface, we proceed as in the proof of Theorem 4.6. Given a possible isometry type X_0 of the thick part of the surface, we look at $\Lambda \setminus \Lambda(X_0)$ and consider the first value in this set, say ξ_1. This can be determined via a single admissible question. Indeed, with the knowledge of $\Lambda(X_0)$, we can determine an upper bound on ξ_1, say B_1. Note that the position of ξ_1 in Λ can be arbitrarily large if ℓ_1 is arbitrarily small, but knowing

26

$\ell_1 \in \Lambda(X_0)$ tells us where to look for ξ_1. The admissible question is: what is the first value of $\Lambda \setminus \{x \in \Lambda(X_0) \mid x \leq B_1\}$?

With this is hand, we iterate this process as in Theorem 4.6: to determine the length of the k curves traversal to those of Γ_0 requires k admissible questions. So all in all, because $k \leq 3g - 3$, we've asked at most $16e^6(g-1)g^{14} + 3g - 3$ admissible questions.

This finite set of questions has allowed us to determine all possible isometry classes of surfaces that might have Λ as their length spectrum. Specifically, we have now reduced the problem to a collection of at most I_g isometry classes where I_g satisfies the bounds of Theorem 1.2. Note this might seem somewhat counter-intuitive: the bounds from Theorem 1.2 are bounds on the number of isospectral but non-isometric surfaces whereas here we can't distinguish between isospectral surfaces. The key thing is that once we know an isometry class, we also know the length spectrum.

We can regroup these isometry classes by spectrum, as we won't be able to distinguish between non-isometric but isospectral surfaces by interrogating a length spectrum. Denote by M this collection of possible spectra, each represented by an isometry class of surface with the appropriate length spectrum.

Now we need to ask additional admissible questions to figure out which one of these possible spectra Λ really is.

To do so, we take any two X, Y with $\Lambda(X) \neq \Lambda(Y)$. There is a smallest integer $m_{X,Y}$ such that

$$\ell_{m_{X,Y}}(X) \neq \ell_{m_{X,Y}}(Y)$$

We ask the following admissible question: what is the first value of

$$\Lambda \setminus \{x \in \Lambda(X) \mid x < \ell_{m_{X,Y}}(X)\}$$

If the answer is not $\ell_{m_{X,Y}}(X)$, then $\Lambda(X) \neq \Lambda$. If the answer is not $\ell_{m_{X,Y}}(Y)$ then $\Lambda(Y) \neq \Lambda$. So a single question rules out either X or Y (or possibly both). Once a spectrum is ruled out, we discard it. After at most $M - 1$ questions, a single spectrum remains and we have determined Λ with absolute certainty.

All in all we've asked

$$M - 1 + 16e^6(g-1)g^{14} + 3g - 3$$

possible questions so using the estimates from the proof of Theorem 1.2 the number of questions is bounded by

$$\frac{8}{e^6}A^{g-1}g^{B(g-1)} + 16e^6(g-1)g^{14} + 3g - 3$$

where $A = 3^6 2^{75} e^{67}$ and $B = 114$. As in the proof of Theorem 1.2, this estimate is asymptotically bounded above by g^{115g} and via a small manipulation can be shown to be bounded, for all $g \geq 2$, by

$$g^{154g}$$

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

A.   Appendix: McShane identities

The main goal of the appendix is to prove Lemma 4.4 using a version of the McShane identity. The identity we shall use is the following, due to Mirzakhani [13] and independently discovered by Tan, Wong and Zhang who also proved a version for surfaces with cone points [18].

Theorem A.1. *Let Y be a surface with non-empty geodesic boundary and let β be one of the boundary curves. Let $\mathcal{P}$ be the set of all embedded geodesic pairs of pants that have β as boundary curve. Let $\mathcal{P}'$ be the subset of $\mathcal{P}$ with two boundary curves of Y as its boundary curves.*

Then there exist explicit positive functions μ, η that depend only the geometry of P such that

$$\sum_{P \in \mathcal{P} \setminus \mathcal{P}'} \mu(P) + \sum_{P \in \mathcal{P}'} \eta(P) = 1$$

The functions μ and η, often called gap functions, are functions of the boundary lengths of P. They are real functions in three variables. By convention we set the first variable of μ to be the length of β. For η, the first variable is the length of β and the second variable is the length of the second boundary curve of the corresponding pair of pants.

The following proposition contains all the features about these functions μ and η that we will need.

Proposition A.2. *The gaps functions μ and η enjoy the following properties:*

  (i) $\mu(x, y, z) < \eta(x, y, z)$

  (ii) $\mu(x, y, z) > \dfrac{1}{e^{\frac{y+z}{2}}}$

Proof. By [13] and [18], the explicit formulae for the functions μ and η are the following:

$$\mu(x, y, z) = \frac{4}{x} \operatorname{arctanh} \left(\frac{\sinh\left(\frac{x}{2}\right)}{\cosh\left(\frac{x}{2}\right) + e^{\frac{y+z}{2}}} \right)$$

28

and

$$\eta(x,y,z) = 1 - \frac{2}{x} \operatorname{arctanh}\left(\frac{\sinh\left(\frac{x}{2}\right)\sinh\left(\frac{y}{2}\right)}{\cosh\left(\frac{z}{2}\right) + \cosh\left(\frac{x}{2}\right)\cosh\left(\frac{y}{2}\right)}\right)$$

Both statements of the proposition can be shown by function manipulation, but we'll give a geometric reason for why (i) is true. The proof of (ii) however will be a straightforward function manipulation.

We begin with (i) which, as we shall see, comes from analyzing where the functions in these identities come from.

These identities come from breaking up β into intervals as follows. For each point of β, consider the unique geodesic segment that leaves β from this point at a right angle and exponentiate until the corresponding geodesic either hits itself or hits a boundary geodesic. (The points that don't do either are measure 0.) If it hits itself, return to β by following the geodesic back to obtain a simple arc. Then we regroup the basepoints into segments according to the homotopy type of the associated simple arc. The functions (called gap functions) come from the measures of the segments divided by $\ell(\beta)$ for normalization. Each homotopy class of arc determines an embedded pair of pants.

The functions μ and η are computed in the corresponding pair of pants. The intervals of β (before normalization) are different depending on whether the pair of pants contains one or two boundary curves (these are the cases $\mathcal{P} \setminus \mathcal{P}'$ and $\mathcal{P}'$ above). The boundary points of the intervals correspond to simple geodesics that spiral indefinitely around the boundary curves of the pair of pants. For $\mathcal{P} \setminus \mathcal{P}'$, these spiraling geodesics are illustrated in [18, Figure 1, p. 91]. We reproduce a similar figure for convenience (Figure 9).
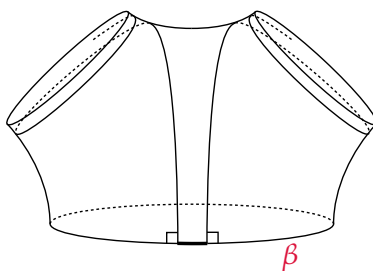

Figure 9: The "front" interval on β corresponding to this pair of pants is in bold

Note there are two of them, one for each orientation of the corresponding simple arc. A pair of pants in $\mathcal{P}'$ has a second boundary cuff β'. There is a single interval this time, but it encompasses both intervals that would have appeared had the corresponding pair of pants been in the set $\mathcal{P} \setminus \mathcal{P}'$. This is illustrated in Figure 10.
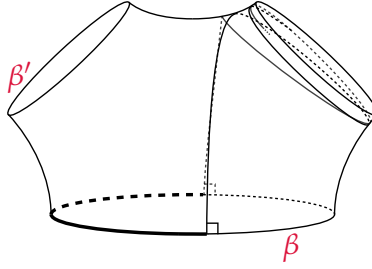
Figure 10: The whole interval on β corresponding to this pair of pants is in bold

This, without doing a single computation, shows that $x \cdot \mu(x,y,z) < x \cdot \eta(x,y,z)$, and hence that $\mu(x,y,z) < \eta(x,y,z)$.

We now pass to (ii).

Using the definition of arctanh and setting $A := e^{\frac{y+z}{2}} \geq 1$ we obtain

$$\mu(x,y,z) = \frac{2}{x} \log \left(\frac{e^{\frac{x}{2}} + A}{e^{-\frac{x}{2}} + A} \right)$$

We want to show $\mu(x,y,z) > \frac{1}{A}$ and to do this we'll show that

$$\log \left(\frac{e^{\frac{x}{2}} + A}{e^{-\frac{x}{2}} + A} \right) > \frac{x}{2A}$$

or equivalently that

$$\frac{e^{\frac{x}{2}} + A}{e^{-\frac{x}{2}} + A} > e^{\frac{x}{2A}}$$

In turn this is equivalent to showing that $F(x) > 0$ for all $x > 0$ (and $A \geq 1$) where

$$F(x) := e^{\frac{x}{2}} + A - Ae^{\frac{x}{2A}} - e^{\frac{x}{2}\left(\frac{1}{A}-1\right)}$$

Note that $F(0) = 0$. We compute the derivative

$$F'(x) = \frac{1}{2} \left(e^{\frac{x}{2}} - e^{\frac{x}{2A}} \right) + \frac{1}{2} \left(1 - \frac{1}{A} \right) e^{\frac{x}{2}\left(\frac{1}{A}-1\right)}$$

which is positive for $A \geq 1$ and the claim follows. $\qquad\square$

As a corollary, we can deduce Lemma 4.4. We recall that the statement is that cardinality of the set $\mathcal{P}_L$ is at most e^L.

Proof of Lemma 4.4. From Theorem A.1 and Proposition A.2 (i), we have

$$\sum_{P \in \mathcal{P} \setminus \mathcal{P}'} \mu(P) + \sum_{P \in \mathcal{P}'} \mu(P) = 1$$

and thus
$$\sum_{P \in \mathcal{P}_L} \mu(P) < 1$$

Now by Proposition A.2 (ii), for P with boundary lengths $\ell(\beta), y, z$ we have that $\mu(P) > \frac{1}{e^{\frac{y+z}{2}}}$. By the monotonicity of this lower bound we have

$$\mu(P) > \frac{1}{e^L}$$

for all $P \in \mathcal{P}_L$. The cardinality of $\mathcal{P}_L$ can thus be at most e^L. $\qquad \square$

References

[1] Christophe Bavard. Disques extrémaux et surfaces modulaires. *Ann. Fac. Sci. Toulouse Math. (6)*, 5(2):191–202, 1996.

[2] Lipman Bers. An inequality for Riemann surfaces. In *Differential geometry and complex analysis*, pages 87–93. Springer, Berlin, 1985.

[3] Chandrasheel Bhagwat and C. S. Rajan. On a multiplicity one property for the length spectra of even dimensional compact hyperbolic spaces. *J. Number Theory*, 131(11):2239–2244, 2011.

[4] Robert Brooks, Ruth Gornet, and William H. Gustafson. Mutually isospectral Riemann surfaces. *Adv. Math.*, 138(2):306–322, 1998.

[5] Peter Buser. *Geometry and spectra of compact Riemann surfaces*, volume 106 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1992.

[6] C. Gordon, D. Webb, and S. Wolpert. Isospectral plane domains and surfaces via Riemannian orbifolds. *Invent. Math.*, 110(1):1–22, 1992.

[7] Heinz Huber. Zur analytischen Theorie hyperbolischen Raumformen und Bewegungsgruppen. *Math. Ann.*, 138:1–26, 1959.

[8] Mark Kac. Can one hear the shape of a drum? *Amer. Math. Monthly*, 73(4, part II):1–23, 1966.

[9] Linda Keen. Collars on Riemann surfaces. In *Discontinuous groups and Riemann surfaces (Proc. Conf., Univ. Maryland, College Park, Md., 1973)*, pages 263–268. Ann. of Math. Studies, No. 79. Princeton Univ. Press, Princeton, N.J., 1974.

[10] G. A. Margulis. Certain applications of ergodic theory to the investigation of manifolds of negative curvature. *Funkcional. Anal. i Priložen.*, 3(4):89–90, 1969.

[11] H. P. McKean. Selberg's trace formula as applied to a compact Riemann surface. *Comm. Pure Appl. Math.*, 25:225–246, 1972.

[12] Greg McShane. Simple geodesics and a series constant over Teichmuller space. *Invent. Math.*, 132(3):607–632, 1998.

[13] Maryam Mirzakhani. Simple geodesics and Weil-Petersson volumes of moduli spaces of bordered Riemann surfaces. *Invent. Math.*, 167(1):179–222, 2007.

[14] Maryam Mirzakhani. Growth of the number of simple closed geodesics on hyperbolic surfaces. *Ann. of Math. (2)*, 168(1):97–125, 2008.

[15] Hugo Parlier. A short note on short pants. *Canad. Math. Bull.*, 57(4):870–876, 2014.

[16] Piotr Przytycki. Arcs intersecting at most once. *Geom. Funct. Anal.*, 25(2):658–670, 2015.

[17] Toshikazu Sunada. Riemannian coverings and isospectral manifolds. *Ann. of Math. (2)*, 121(1):169–186, 1985.

[18] Ser Peow Tan, Yan Loi Wong, and Ying Zhang. Generalizations of McShane's identity to hyperbolic cone-surfaces. *J. Differential Geom.*, 72(1):73–112, 2006.

[19] Richard M. Tse. A lower bound for the number of isospectral surfaces. In *Recent developments in geometry (Los Angeles, CA, 1987)*, volume 101 of *Contemp. Math.*, pages 161–164. Amer. Math. Soc., Providence, RI, 1989.

[20] Marie-France Vignéras. Exemples de sous-groupes discrets non conjugués de PSL(2, **R**) qui ont même fonction zéta de Selberg. *C. R. Acad. Sci. Paris Sér. A-B*, 287(2):A47–A49, 1978.

[21] Marie-France Vignéras. Variétés riemanniennes isospectrales et non isométriques. *Ann. of Math. (2)*, 112(1):21–32, 1980.

[22] Scott Wolpert. The length spectra as moduli for compact Riemann surfaces. *Ann. of Math. (2)*, 109(2):323–351, 1979.