

KISSING NUMBERS FOR SURFACES

Hugo Parlier[†]

Abstract. The so-called *kissing number* for hyperbolic surfaces is the maximum number of homotopically distinct systoles a surface of given genus g can have. These numbers, first studied (and named) by Schmutz Schaller by analogy with lattice sphere packings, are known to grow, as a function of genus, at least like $g^{4/3-\varepsilon}$ for any $\varepsilon > 0$. The first goal of this article is to give upper bounds on these numbers; in particular the growth is shown to be sub-quadratic. In the second part, a construction of (non hyperbolic) surfaces with roughly $g^{3/2}$ systoles is given.

1. INTRODUCTION

The classical kissing number problem for sphere packings is the search for an optimal upper bound on the number of n -dimensional (Euclidean) unit spheres, pairwise disjoint in their interior, that can be tangent to a fixed unit sphere. Exact values for these numbers, commonly called kissing numbers for simplicity, are only known in a finite number of cases ($n = 1 - 4, 8, 24$, see [13] and references therein). A seemingly easier problem is to ask the same question as above, but with the restriction that the centers of the spheres lie on some Euclidean lattice. This gives rise to the so-called kissing number problem for lattice sphere packings. As above, values are only known in a finite number of cases ($n = 1 - 9, 24$, see [5] and references therein). In some cases the solutions to the two problems coincide but not in general: for instance they are known to be different for $n = 9$. When $n = 2$ the answer is 6 for both problems and the optimal solution is given by points lying on the hexagonal lattice. In this case, the question is equivalent to asking how many shortest (nontrivial) vectors a lattice can contain. In terms of *systoles* (meaning shortest homotopically non-trivial curves of a manifold), the question is to determine at most how many distinct isotopy classes of systoles a flat torus can have. The kissing number problem for lattices is the natural generalization: how many isotopy classes of systoles can the underlying n -torus contain?

Another possible generalization - proposed by Schmutz Schaller - is to hyperbolic closed

[†]Research supported by Swiss National Science Foundation grant number PP00P2.128557

Address: Department of Mathematics, University of Fribourg, Switzerland

Email: hugo.parlier@unifr.ch

2010 Mathematics Subject Classification: Primary: 30F10. Secondary: 32G15, 53C22.

Key words and phrases: hyperbolic surfaces, metrics on surfaces, kissing numbers, systoles

surfaces. The question remains the same but here genus plays the part of dimension. For hyperbolic surfaces there is a unique geodesic in a prescribed nontrivial isotopy class so the question is as follows.

How many systoles can a hyperbolic surface of genus g have?

By analogy with the lattice case, Schmutz Schaller defined the maximum number for each genus to be the kissing number for hyperbolic surfaces. Usually we consider simple closed geodesics, and systoles in particular, to be unoriented objects, so morally we should be dividing lattice sphere packing kissing numbers by 2 to make the two problems truly analogous. As in the case of lattices, very little is known about exact values: the only known case is for genus 2 where the answer is 12.

A related problem for systoles of hyperbolic surfaces is the search for surfaces with systole of maximal length in a prescribed genus, again initiated by Schmutz Schaller [15]. As before, there is a possible analogy with lattices: this is the hyperbolic surface version of the search for the optimal Hermite constants. (The Hermite constant can be defined as the square of the maximal systole length among unit volume flat tori.) For each genus, via a compactness argument, there exists a hyperbolic surface with longest possible systole. Again, the only known exact value is for genus 2, a result of Jenni [9]. It is realized by the same surface as for the kissing number problem - the so-called Bolza surface. The Bolza surface is also maximal among genus 2 hyperbolic surfaces for the number of self-isometries. In the non-compact cases, arithmetic surfaces coming from principal congruence subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ are known to be maximal among finite area hyperbolic surfaces in their respective moduli spaces, a result of Schmutz Schaller [16], see also [1] for another proof. For both problems, although finding exact solutions seems difficult, one can ask for upper and lower bounds, and in particular it is natural to study the asymptotic growth of how these constants vary in function of genus. Via a simple area argument, it is easy to see that the systole length of a hyperbolic surface cannot exceed $2 \log g$ ($+C$ for some constant $C > 0$). Buser and Sarnak [4] were the first to construct families of surfaces with $\sim \log g$ systole growth. More precisely, they showed

$$\limsup_{g \rightarrow \infty} \frac{\max_{S \in \mathcal{M}_g} \mathrm{sys}(S)}{\log g} \geq \frac{4}{3}$$

where \mathcal{M}_g is the moduli space of all hyperbolic surfaces up to isometry and $\mathrm{sys}(S)$ is the length of the systole of S . Since then there have been other constructions, see for example [10], but all are based on arithmetic methods of some sort. The existence of such surfaces is somewhat surprising: the radius of a maximally embedded disk in a hyperbolic surface is $\sim \log g$, so a surface with $\sim \log g$ systole looks essentially everywhere like a fat disk that is

pasted together in some clever way in order to avoid creating a short loop.

Schmutz Schaller wrote a series of papers for the hyperbolic surface kissing number problem, in both the compact and non-compact cases, where he proved some interesting lower bounds. In first instance, one might think that it might not be possible to have a surface where the number of systoles is considerably bigger than the size of a maximal isometry group. Via Hurwitz's bound, this would imply an upper bound to the number of systoles which grows linearly in genus. Schmutz Schaller's first result on this was the existence of families of surfaces with a number of systoles that grew more than linearly in genus [18, 19]. The best results for closed surfaces appeared later [20] where he showed that

$$\limsup_{g \rightarrow \infty} \frac{\log \left(\max_{S \in \mathcal{M}_g} \text{Kiss}(S) \right)}{\log g} \geq \frac{4}{3} - \varepsilon$$

for any $\varepsilon > 0$ where $\text{Kiss}(S)$ denotes the number of systoles of S . Stated otherwise, for any $\varepsilon > 0$, there exists a family of surfaces, one in each genus, with more than $g^{4/3-\varepsilon}$ systoles (for large enough g). Again, the construction is based on arithmetic methods. Based on the intuition that one cannot do better for arithmetic surfaces and that arithmetic surfaces should be optimal for these problems, Schmutz Schaller [21] made two conjectures, namely that the inequalities above are in fact equalities (with the right hand side equal to $4/3$). For the sake of avoiding repetitive repetition, we rephrase them as follows.

Conjecture 1.1. *There exists a constant $A > 0$ such that*

$$\max_{S \in \mathcal{M}_g} \text{sys}(S) \leq \frac{4}{3} \log g + A.$$

Conjecture 1.2. *There exists a constant $B > 0$ such that*

$$\max_{S \in \mathcal{M}_g} \text{Kiss}(S) \leq Bg^{4/3}.$$

The two problems seem to run in parallel because of the arithmetic nature of the interesting known examples, but Schmutz Schaller did not establish a direct link between the two. In contrast to Conjecture 1.1, there was no "easy" known upper bound for Conjecture 1.2 that behaves roughly as conjectured. One might expect an easy quadratic upper bound for the number of systoles which would seem to be the counterpart to the straightforward $2 \log g$ bound. (In fact Schmutz Schaller claims a quadratic upper bound in [17] but the argument, supposedly only based on the topological condition that two systoles pairwise intersect at most once, is faulty and does not seem to be easily repairable.)

The first result of this paper is to find an upper bound for $\text{Kiss}(S)$, which depends on the length of the systole of S .

Theorem 1.3. *There exists a constant $U > 0$ such that for any hyperbolic surface S of genus g with systole ℓ the following holds:*

$$\text{Kiss}(S) \leq U \frac{e^{\ell/2}}{\ell} g.$$

Using the $2 \log g$ bound on ℓ mentioned above, this has the following consequence.

Corollary 1.4. *There exists a constant $U > 0$ such that any hyperbolic surface of genus g has at most $U \frac{g^2}{\log g}$ systoles.*

More generally, if there exists an upper bound on length of systoles of type $C \log g$ for some $C > 0$, then Theorem 1.3 implies a bound of order $\sim g^{C/2+1}$. In particular a positive answer to Conjecture 1.1 would imply that there are at most $\sim g^{5/3}$ systoles on a genus g surface. Another consequence is that Conjecture 1.2 holds for all surfaces with systole bounded above by $2/3 \log g$.

The bound also shows that if a family of surfaces has sub-logarithmic systole growth, then the number of systoles is “almost” at most linear. More precisely:

Corollary 1.5. *Let $f(g)$ be any positive function with $\lim_{g \rightarrow \infty} f(g)/\log g = 0$. Then*

$$\max\{\text{Kiss}(S) \mid S \in \mathcal{M}_g \text{ with } \text{sys}(S) \leq f(g)\} \leq g^{1+\varepsilon}.$$

for any $\varepsilon > 0$ and large enough g .

As stated this corollary can be also be proved via an estimate in [3, Lemma 6.6.4]. Theorem 1.3 implies a more accurate estimate than the above corollary but the formulation above is given for clarity. In particular, this means that any family of surfaces with “many” systoles (by which we mean at least g^{1+a} for some $a > 0$), then the family has $\sim \log g$ systole growth as well. Of course it is not a priori easier to construct surfaces with many systoles, but if one does, the large size systoles come for free. One can also observe that having many systoles is unusual behavior in the sense that already having any two simple closed geodesics of equal length is “rare” (see for instance [12]).

In the more general context of Riemannian metrics on surfaces, similar questions can be asked. One no longer has uniqueness of a geodesic in an isotopy class, so the appropriate question on the number of systoles is an upper bound on the number of distinct isotopy classes of simple closed curves that can simultaneously be realized as systoles for some metric.

How many systoles up to isotopy can a closed Riemannian surface of genus g have?

We'll relate to optimal upper bounds to this problem as the kissing number problem for general surfaces.

It is not known what sets of topological curves on either a hyperbolic or Riemannian surface can be realized by systoles. By a cutting and pasting argument, it is not too difficult to see that on any closed surface (not necessarily hyperbolic) a systole is necessarily a simple closed curve. Likewise, any two systoles can pairwise intersect at most once. This leads us to a related purely topological problem attributed to B. Farb and C. Leininger (see [11]).

Up to isotopy, how many distinct curves can be realized on a surface of genus g such that they pairwise intersect at most once?

It seems to be a surprisingly hard question to answer. The best known upper bound is exponential and the best known lower bound quadratic [11]. The upper bound provides also provides an upper bound for kissing numbers, and to the best of the author's knowledge, these are the best known upper bounds for kissing numbers of general surfaces. Observe that Corollary 1.4 shows that optimal kissing numbers for hyperbolic surfaces cannot be the same as the numbers coming from the purely topological problem. It was already known that the topological condition was quite different from the systolic condition: in [2], there are constructions of configurations of isotopy classes curves that fail to be systoles for *any* Riemannian metric on the surface.

For completeness, we mention that the problem of finding an upper bound on systole length in the case of surfaces of variable curvature has also been studied. For a genus g surface with area normalized to the area of its hyperbolic counterparts, Gromov [7, 8] gave a $\sim \log g$ upper bound, which in light of the hyperbolic examples is roughly optimal, but the precise asymptotic growth remains completely open. Whether the asymptotic growth for both problems should be different for both the upper bound on length or the kissing numbers is also unknown.

The second main result of this paper is about lower bounds for kissing number for Riemannian surfaces.

Theorem 1.6. *There exist surfaces of genus $g > 0$ with a number of systoles of order of growth at least $g^{3/2}$.*

This shows that if Conjecture 1.2 is correct, the asymptotic growth for kissing numbers is considerably different in the case of variable curvature surfaces. Recall the best known - and conjectured optimal - lower bounds for hyperbolic surfaces are roughly $g^{4/3}$.

The proof is by construction, and the geometry of the surfaces come from embeddings of

complete graphs. One might wonder if one can use these techniques to find hyperbolic surfaces with the same behavior, but the surfaces are far from being hyperbolic. In fact, in striking contrast to what is possible in the hyperbolic case (Corollary 1.5), the systole length is quite small proportionally to area (constant systole length for area $\sim g$).

Acknowledgement.

The author is grateful to the referee for her or his helpful comments.

2. BOUNDS ON NUMBERS OF SYSTOLES OF HYPERBOLIC SURFACES

We denote $\text{Kiss}(S)$ the number of systoles of a surface S . Before proceeding to proofs of upper bounds on $\text{Kiss}(S)$ for hyperbolic surfaces, we begin with some observations on the geometry of systoles.

2.1. Geometric properties of systoles

The estimates needed rely essentially on trigonometric arguments in the hyperbolic plane. Recall the classical collar lemma.

Lemma 2.1 (Collar lemma). *Let γ be a simple closed geodesic of length ℓ . Then there is an embedded collar of width $w(\ell)$ around γ where*

$$w(\ell) = \operatorname{arcsinh} \left(\frac{1}{\sinh(\ell/2)} \right).$$

Furthermore, any simple closed geodesic δ that enters this collar essentially intersects γ .

For systoles, one can do even better.

Lemma 2.2 (Systolic collar lemma). *Let α and β be systoles of length ℓ that don't intersect. Then there are at a distance at least $2r(\ell)$ where*

$$r(\ell) = \operatorname{arcsinh} \left(\frac{1}{2 \sinh(\ell/4)} \right).$$

Proof. Take a shortest path c between α and β and consider the geodesic γ in the free homotopy class of $\alpha * c * \beta * c^{-1}$ where α, β and c are oriented so that γ is simple (see figure 1). Note that α, β, γ form a pair of pants.

Now because $\ell(\gamma) \geq \ell$, we can compute the minimal distance between α and β . The result follows from standard trigonometry in a pair of pants and the double sinh formula. \square

From this the following corollary is immediate.

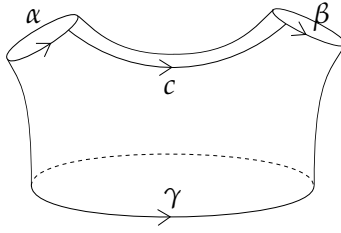


Figure 1: Orientations of α , β and c

Corollary 2.3. *If γ and δ are two systoles of length at most ℓ which pass through a same disk of radius $r(\ell)$, then they essentially intersect.*

Using only this observation, it is possible to obtain a universal polynomial bound on the number of systoles of hyperbolic surfaces, but we need to work harder to obtain our estimates.

We begin by noticing that if two systoles intersect, then their angle of intersection can be bounded below by their length.

Lemma 2.4. *Let γ and δ be systoles of length ℓ that intersect. Then $\sin \angle(\gamma, \delta) > \frac{1}{2} (\cosh(\ell/4))^{-1}$.*

Proof. Let p be the intersection point of the two curves and consider the point q of δ that is at distance exactly $\ell/2$ away from p . Fix one of the two paths of δ , say δ_1 , between q and γ . Among all paths freely homotopic to γ_1 with one fixed endpoint at q and the other on any point of γ , there is one of minimal length which we shall denote h . Denote q' the point of intersection of h and γ . We now have a right triangle p, q, q' with hypotenuse H is of length $\ell/2$, the side h , and the basis b . Observe that using the other arc of δ , say δ_2 , one obtains a symmetric situation with an isometric right angle triangle where the two triangles are linked via a rotation of angle π around q (see figure 2).

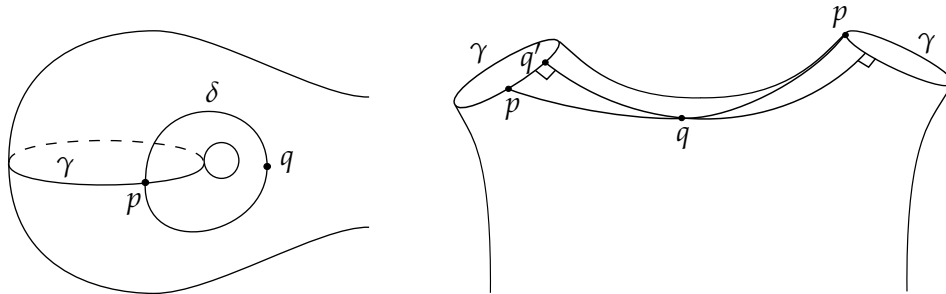


Figure 2: Two intersecting systoles

As the two triangles are isometric, we can concentrate on the first one. It is not difficult to see that b is of length at most $\ell/4$, otherwise one can construct a shorter nontrivial path than δ .

Consider the angle $\theta = \angle(\gamma, \delta)$. By the sine formula for hyperbolic triangles, it satisfies

$$\frac{\sin(\theta)}{\sinh(h)} = \frac{\sin(\pi/2)}{\sinh(\ell/2)}.$$

Now because the basis is at most $\ell/4$, h is strictly greater than $\ell/4$ (otherwise the same occurs in the other triangle, and we obtain $\ell(\delta) < \ell$). The inequality

$$\sin \angle(\gamma, \delta) > \frac{\sinh(\ell/4)}{\sinh(\ell/2)}$$

and one concludes by using the sinh formula for a double angle. \square

Notation 2.5. We denote $\theta_\ell := \arcsin \frac{1}{2 \cosh(\ell/4)}$.

For large ℓ , θ_ℓ behaves roughly like $e^{\ell/4}$. In particular, using the $2 \log g$ upper bound on length, this implies a $\frac{1}{\sqrt{g}}$ lower bound on the angle between systoles. Thus a collection of systoles that intersect in a single point cannot have cardinality greater than roughly \sqrt{g} . This is very different from the case of variable curvature where one can construct surfaces with $\sim g$ systoles that intersect in a single point, see Remark 3.3.

Corollary 2.3 ensures that two systoles γ, δ with points $p_\gamma \in \gamma$, $p_\delta \in \delta$ that satisfy $d(p_\gamma, p_\delta) < d(\ell)$ must intersect somewhere on the surface. This next lemma gives a bound on how far the intersection point can be from p_γ and p_δ .

Lemma 2.6. *Let γ, δ be systoles of length ℓ which cross a disk of radius $r(\ell)$ with center p . Then the intersection point q between γ and δ satisfies*

$$d(p, q) < \operatorname{arcsinh} \left(2 \coth \frac{\ell}{4} \right).$$

Proof. Consider points $p_\gamma \in \gamma$, $p_\delta \in \delta$ which lie in the disk centered in p . Consider the two angles $\angle(q, p, p_\gamma)$ and $\angle(q, p, p_\delta)$. By Lemma 2.4, one of the two angles must be greater than $\frac{\theta_\ell}{2}$. Without loss of generality, let us suppose that this angle is $\theta = \angle(q, p, p_\gamma)$. Note that Lemma 2.4 also provides

$$\sin(\theta) \geq \sin(\theta_\ell/2) = \frac{1}{2 \cos(\theta_\ell/2)} \sin(\theta_\ell) > \frac{1}{4 \cosh(\ell/4)}.$$

We now focus our attention to the triangle p, q, p_γ . Denote by θ' the angle opposite side \overline{pq} (see figure 3). We have the following identity:

$$\frac{\sin(\theta')}{\sinh d(p, q)} = \frac{\sin(\theta)}{\sinh d(p, p_\gamma)}.$$

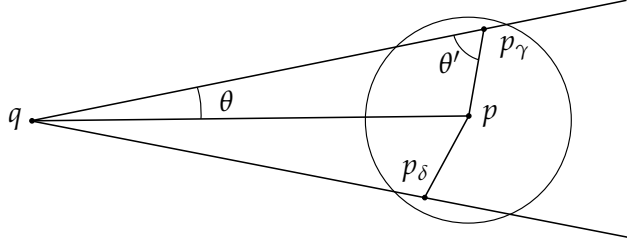


Figure 3: Computing the distance to an intersection point

Via Lemma 2.2 and what precedes we obtain

$$\sinh d(p, q) < \frac{\sin(\theta')}{\sin(\theta)} \frac{1}{2 \sinh(\ell/4)} < 2 \coth \frac{\ell}{4}$$

which concludes the proof. \square

Notation 2.7. For future reference, we fix the following notation:

$$R(\ell) := \operatorname{arcsinh} \left(2 \coth \frac{\ell}{4} \right).$$

Remark 2.8. Observe that via the same argument, for two systoles that intersect in a point q and both simultaneously pass through a disk of radius ρ and center p , the following holds:

$$\sinh d(p, q) \leq \frac{\sinh \rho}{\sin(\theta_\ell/2)} < \frac{\sinh \rho}{4 \cosh(\ell/4)}.$$

2.2. Proof of the upper bound

We can now proceed to the proof of our main upper bound. We begin by giving a more precise version of Theorem 1.3.

Theorem 2.9. *Let S be a hyperbolic surface of genus g with systole bounded by ℓ . Then*

$$\operatorname{Kiss}(S) \leq C_\ell (g - 1)$$

where C_ℓ is a constant that depends on ℓ which can be taken to be

$$C_\ell = 100 \frac{e^{\ell/2}}{\ell}.$$

Remark 2.10. The constant 100 in front can easily be improved but it is the order of growth we are interested in.

Proof. If $\ell \leq 2 \operatorname{arcsinh} 1$ then by the collar lemma all systoles are disjoint and thus, as there are at most $3g - 3$ disjoint simple closed geodesics on a genus g surface, we obtain

$$\operatorname{Kiss}(S) \leq 3g - 3.$$

We can now concentrate on the case when $\ell \geq 2 \operatorname{arcsinh} 1$.

The basic strategy will be the following. We begin by covering the surface S by balls of radius $r(\ell)$ (where $r(\ell)$ is given by Lemma 2.2). The first step will be to estimate $F(S)$, an upper bound on the minimum number of these balls required to cover S . We then find an upper bound $G(S)$ on the number of systoles that can cross such a ball. Finally, if we denote $H(S)$ the minimum number of covering balls that a systole of S must cross, we have

$$\operatorname{Kiss}(S) \leq \frac{F(S) G(S)}{H(S)}.$$

Let us now concentrate on finding bounds for these quantities in function of ℓ and g .

The number of balls required to cover S

As usual in these type of estimates, we use

$$F(S) = \{\text{Number of balls of radius } r(\ell) \text{ needed to cover } S\} \\ \leq$$

$$\{\text{Max number of balls of radius } r(\ell)/2 \text{ that embed and are pairwise disjoint}\}.$$

Now as

$$\operatorname{Area}(D_{r(\ell)/2}) = 2\pi(\cosh(r(\ell)/2) - 1) = 2\pi \left(\cosh \left(\frac{\operatorname{arcsinh} \left(\frac{1}{2 \sinh(\ell/4)} \right)}{2} \right) - 1 \right)$$

we deduce a bound for $F(S)$ that depends only on g (coming from the area of S) and ℓ :

$$F(S) \leq \frac{\operatorname{Area}(S)}{\operatorname{Area}(D_{r(\ell)/2})}.$$

To get an idea of the order of growth of this bound, observe that

$$\operatorname{Area}(D_{r(\ell)/2}) > \frac{\pi}{4} e^{-\ell/2}$$

and it follows that

$$F(S) < 16(g - 1) e^{\ell/2}.$$

The number of systoles intersecting each ball

We now proceed to find an upper bound on the number of systoles that can intersect a ball of radius $r(\ell)$. Consider a disk D_0 on S of radius $r(\ell)$. Any two given systoles that cross D_0 must intersect and via Lemma 2.6 we know that their intersection point q lies within $R(\ell)$ of the center p of D_0 . We shall now reason in the universal cover, considering D_0 and the disk D_1 of center p and radius $R(\ell)$. The geometric problem we are interested in is as follows.

How many (hyperbolic) lines, any two of which pairwise intersect in an angle of at least θ_ℓ , can intersect D_0 ?

For this consider the disk D_2 , also of center p , but of radius $R(\ell) + R'$ for a given $R' > 0$. Consider geodesics γ_1 and γ_2 which pass through D_0 and intersect in $q \in D_1$. From q , we consider one of the four angular sectors of angle $\theta > 0$ and the two rays of γ_1 and γ_2 which bound the sector. We set q_1, q_2 to be the intersection points of the rays of γ_1 and γ_2 with the boundary of D_2 as in figure 4. Note that $d(q, q_1)$ and $d(q, q_2) > R'$ and that $\theta > \theta_0$.

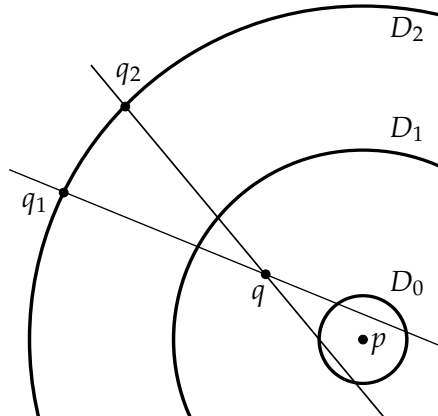


Figure 4: The disks D_0, D_1 and D_2

Considering the triangle q, q_1, q_2 , observe the following: the distance $d(q_1, q_2)$ is strictly greater than the distance one would compute if

$$d(q, q_1) = d(q, q_2) = R' \text{ and } \theta = \theta_0.$$

Via hyperbolic trigonometry in the (q, q_1, q_2) triangle:

$$d(q_1, q_2) > 2 \operatorname{arcsinh}(\sinh(R') \sin(\theta_\ell)).$$

We now consider a maximal set of lines $\gamma_1, \dots, \gamma_k$ that cross D . The lines divide ∂D_2 into $2k$ arcs. As such, the cardinality of the set of lines is bounded by the length of ∂D_2 divided by twice the minimal distance between the intersection points of the lines with D_2 . By the above estimates, we have

$$G(S) < \frac{\pi \sinh(R(\ell) + R')}{2 \operatorname{arcsinh}(\sinh(R') \sin(\theta_\ell))}.$$

Setting $R' := \operatorname{arcsinh}(1)$ this becomes

$$G(S) < \frac{\pi \sinh(R(\ell) + \operatorname{arcsinh}(1))}{2 \operatorname{arcsin} \frac{1}{2 \cosh(\ell/4)}}.$$

Note that $R(\ell)$ remains bounded as ℓ increases so we get an upper bound on $G(S)$ which has order of growth $\sim e^{\ell/4}$.

The number of balls each systole crosses

The last estimate we need is a lower bound on the number of balls $H(S)$ in our covering of S that each systole necessarily crosses. The estimate we'll use is straightforward: to cover a geodesic segment of length ℓ with balls of radius $r(\ell)$ one requires at least $2\ell/r(\ell)$ balls. Thus

$$H(S) \geq \frac{2\ell}{r(\ell)} = \frac{2\ell}{\operatorname{arcsinh}\left(\frac{1}{2 \sinh(\ell/4)}\right)}.$$

Here the order of growth is roughly $\ell e^{\ell/4}$.

We can now conclude by using our estimates for $F(S)$, $G(S)$ and $H(S)$. By our above estimates, the order of growth of the bound is

$$\sim \operatorname{Area}(S) \frac{e^{\ell/2}}{\ell}.$$

By a few elementary considerations, for $\ell \geq 2 \operatorname{arcsinh} 1$, we have the following effective inequality

$$\operatorname{Kiss}(S) < 100 (g-1) \frac{e^{\ell/2}}{\ell}$$

which concludes the proof. □

The proofs of Corollaries 1.4 and 1.5 stated in the introduction follow from straightforward calculations.

3. NON-HYPERBOLIC SURFACES WITH MANY SYSTOLES

In this section, we construct non-hyperbolic surfaces with many systoles. Recall the statement from the introduction.

Theorem 3.1. *There exist surfaces of genus $g > 0$ with a number of systoles of order of growth at least $g^{3/2}$.*

Remark 3.2. The statement is stated in terms of order of growth for simplicity. The result we shall in fact prove is stronger, namely that there exist surfaces with at least

$$g\sqrt{48g-47} + 15g + \frac{1}{3}\sqrt{48g-47} - \frac{41}{3} > 6g^{3/2}$$

number of systoles, but it is the order of growth we are really interested in.

Proof. The proof is by construction. The general idea of the proof is as follows: we'll begin with a complete graph Γ_n (with $n > 3$ for this construction to work) and imitate its geometry on a surface into which it is embedded. To give a rough idea of where the numbers are coming from, recall that one can embed a complete n -graph into a surface of genus $g \sim n^2$. On such a graph, the any nontrivial loop has length at least 3, so the short loops are those of length exactly 3 and there are $\binom{n}{3} \sim n^3 \sim g^{3/2}$ of those. The idea is then to construct an (almost) isometric embedding so that "many" of the nontrivial short cycles of the graph remain nontrivial on the surface, and so that there aren't any others of smaller length. Then the surface will require some tweaking so that the lengths of the short nontrivial curves are all exactly the same.

Our first observation is topological: consider a minimal genus surface S_{g_n} of genus g_n into which Γ_n is embedded. By Ringel's and Youngs' Theorem [14] we have

$$g_n < \frac{(n-3)(n-4)}{12} + 1.$$

This embedding will serve as our blueprint for the construction of the surface. In fact, we will begin building the geometry of our goal surface in a neighborhood of the embedded graph before describing the metric on the full surface.

Geometrically we will be thinking of our graph as a metric graph with all edges of length 1, and obtained by pasting n -pods together (each n -pod consists of a vertex with n half-edges of length $1/2$). We begin by embedding each n -pod in the Euclidean plane as follows. The vertex is on the origin and the half-edges are Euclidean segments of length $1/2$ equally distributed around the origin.

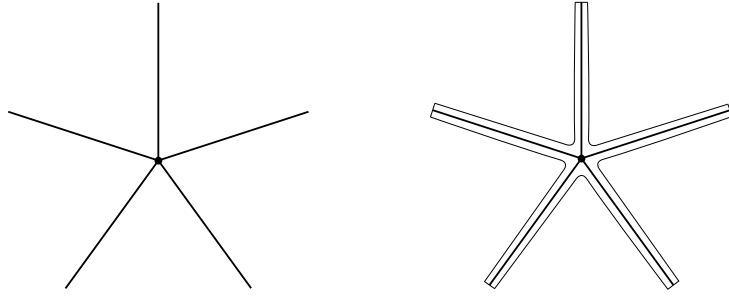


Figure 5: Constructing a ribbon n -pod

We now fix a small $\varepsilon > 0$ and consider a *ribbon type n -pod* of width $\varepsilon > 0$ around our embedded n -pod. The ends of the ribbons are flat as is illustrated in figure 5. Formally: the ribbon type n -pod is the intersection of the closed $\varepsilon > 0$ neighborhood of the ribbon and a collection of closed half-planes defined as follows. For each Euclidean ray supporting a half-edge, we consider the lower half space delimited by the unique line perpendicular to the ray and distance $\frac{1}{2}$ from the origin. The flat ends of each individual ribbon are Euclidean segments of length 2ε . Two half ribbons can be glued geometrically in the natural way to (locally) obtain a picture of a surface with boundary (see figure 6).



Figure 6: The construction of $S_{g_n}^\varepsilon$ and its embedding in S_{g_n}

Going back to the pairings induced by the original embedding, we perform this individual gluing on each of the pairings, and the result is a surface with boundary. Observe that the surface is homeomorphic to a closed neighborhood of the graph embedded in the surface. We denote this surface $S_{g_n}^\varepsilon$. We shall observe a few things about this surface.

First of all, thinking of $S_{g_n}^\varepsilon$ as a topological subsurface of S_{g_n} , observe that all the boundary curves of the embedded subsurface $S_{g_n}^\varepsilon$ are trivial (otherwise the embedding would not have minimal genus).

Via this construction, there is a natural projection (at the level of homotopy classes) from cycles in Γ_n (and curves in $S_{g_n}^\varepsilon$) to curves in S_{g_n} . We are interested in minimal length cycles

of Γ_n (those of length 3) that project to nontrivial and homotopically distinct curves on S_{g_n} . We now need to count them. Consider a cycle on Γ_n and consider it on $S_{g_n}^\varepsilon$ via our original embedding. As it passes through a vertex, it enters through one half edge, exits through another and separates the ribbon type n -pod into 2 pieces separating the remaining half-edges into two sets (one of which can be empty). This is illustrated in figure 7.

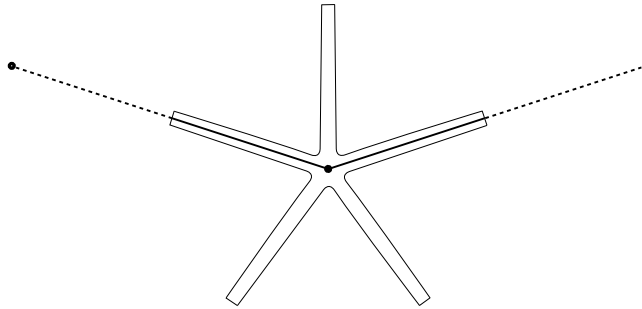


Figure 7: A local picture of a short cycle of Γ_n in S_{g_n}

Consider the subset of all length 3 cycles which contain a vertex on which the cycle separates the other half-edges of the associated n -pod into two *non-empty* sets. Now we claim that any such cycle projects to a *nontrivial* curve on S_{g_n} and that any two such cycles are non-isotopic. To see this observe that any such cycle intersects another such cycle in exactly one point. As they intersect exactly once, both of these cycles are essential and in fact homologically nontrivial. This can be seen via the so-called bigon criterion: any two curves intersect minimally is their isotopy classes if and only if they do not form any bigons in their complementary regions. In particular they are non-trivial homologically. Furthermore, observe that any two distinct such cycles must intersect at least one other such cycle differently, and as such are homologically distinct and in particular non-isotopic. We also remark that any two such cycles intersect at most once.

Remark. Although this doesn't play a role in what follows (because we are concerned with lower bounds), the cycles described above constitute the full set of short cycles of Γ_n that project to pairwise distinct nontrivial curves on S_{g_n} . To see this, observe that if a short cycle fails to be in this set, then via our embedding into S_{g_n} , it is freely homotopic to a curve that is disjoint from the embedding of Γ_n . If it were to be nontrivial on S_{g_n} then by cutting along this curve one obtains an embedding of Γ_n into a surface of smaller genus than S_{g_n} , a contradiction.

We can now count these cycles. Note that any two edges of a short cycle determine the cycle completely as they determine the 3 vertices. Around each vertex there are $\frac{n(n-3)}{2}$ choices of

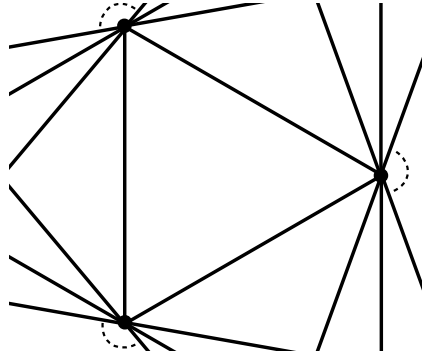


Figure 8: A nontrivial cycle in Γ_n which projects to a trivial curve in S_{g_n}

pairs of distinct edges joined in the vertex that separate the remaining edges of the vertex into two non-empty sets. We can make this number of choices for each vertex and each cycle will have been counted (at most) 3 times. The total number of cycles N is thus at least

$$N \geq \frac{n^2(n-3)}{6}.$$

These cycles determine a fixed set of (free) homotopy classes of curves on S_{g_n} and it is the set we are to going to find a geometric structure for which they are all represented by systoles. We denote this set of free homotopy classes by \mathcal{H} .

We now return to our surface with boundary $S_{g_n}^\varepsilon$. Each of the cycles on this surface admits a minimal length geodesic in its free homotopy/isotopy class. Often one refers to the minimal length of an isotopy class as the length of an isotopic geodesic of minimal length.

Claim. Each of the minimal length geodesics has length $\in [3 - 6\varepsilon, 3]$.

Proof of claim. The upper bound is obvious as each free homotopy class is represented by a cycle of length 3, passing through 3 vertices v_1, v_2, v_3 . For the lower bound, observe that the distance between two of the vertices on $S_{g_n}^\varepsilon$ is exactly 1. Now consider points p_1, p_2, p_3 along a minimal geodesic γ that are of minimal distance to v_1, v_2, v_3 . By construction, these distances are all less than ε (any curve freely homotopic to the cycle must traverse each of the three “corridors”). Consider the three subarcs of γ between the points p_1, p_2, p_3 , and denote their lengths ℓ_1, ℓ_2, ℓ_3 . Now by concatenating the paths, and using the fact that the vertices are at distance 1 from each other, we see that

$$1 \leq \ell_k + 2\varepsilon$$

for $k = 1, 2, 3$. It follows that

$$\ell(\gamma) = \ell_1 + \ell_2 + \ell_3 \geq 3 - 6\varepsilon.$$

This proves the claim.

Our next objective is to modify the metric so that the curves in \mathcal{H} all have length exactly 3. We begin by choosing, one for each of our isotopy classes, a minimal length geodesic. If necessary, we paste (arbitrarily) thin Euclidean cylinders on the boundary of the surface to ensure that all minimal geodesics we have chosen do not touch the boundary. Observe that any *other* nontrivial isotopy class on the surface has minimal length at least 4 minus a term that tends to 0 as ε tends to 0. Via [6], any two minimal geodesics intersect minimally among all representatives in their respective free homotopy classes. In particular, any two distinct and freely homotopic minimal length geodesics are disjoint. It follows that there exists a $\varepsilon' > 0$ (with $\varepsilon' < \varepsilon$) such that the closed $\varepsilon' > 0$ neighborhood of one our chosen minimal length geodesics only (completely) contains other minimal length geodesics that are in the same free homotopy class. As such, we can consider an $\varepsilon' > 0$ neighborhood of the full set of chosen minimal length geodesics for which we are sure that any two minimal length geodesics have the same “combinatorics” if and only if they are isotopic. By *combinatorics* we mean that they pass through the same ε' corridors (strips of total length ~ 1 and of width ε'). We consider this new surface with boundary $S_{g_n}^{\varepsilon'}$. Observe that the number of boundary components may have increased, but the genus remains the same (topologically we’ve just cut disjoint disks out of $S_{g_n}^{\varepsilon}$ to obtain $S_{g_n}^{\varepsilon'}$).

We can now proceed to the modification of the metric on $S_{g_n}^{\varepsilon'}$. Consider the set of *maximal length* geodesics among our set of minimal length geodesics representing \mathcal{H} .

To illustrate the construction, let us suppose that there is only one curve of maximal length, say γ , of length $L \leq 3$. We can insert a Euclidean cylinder of width ω and with both boundary lengths $\ell(\gamma)$ along γ while respecting how the two copies of γ were pasted together. (We will call this *grafting* along γ as this is the name of this type of construction for hyperbolic surfaces.) Observe that this process *increases* the minimal length of any isotopy class of a curve that crosses γ . To see this, suppose this was not the case, i.e., there is a curve δ on the modified surface of shorter length than a given minimal length geodesic on the original surface. Consider the subarc c of δ that passes through the cylinder. Consider the projection c' of c to the basis of the cylinder as in figure 9. By construction $\ell(c) > \ell(c')$. Now we “project” δ to the original surface, by concatenating the image of c' and the arc of δ that was disjoint from the cylinder to obtain a curve $\tilde{\delta}$ on the original surface of strictly shorter length than $\ell(\delta)$, a contradiction.

Also observe that via grafting along γ the minimal length of the isotopy class of γ does not change. In fact, we obtain a cylinder of minimal length geodesics parallel to the two copies



Figure 9: Grafting along γ

of γ . Also observe, that in light of the fact that minimal geodesics intersect minimally, the minimal length of *any* isotopy class that does not intersect γ remains unchanged during this process. As $\ell(\gamma) \leq 3$, and any curve going through the inserted cylinder has length strictly less than the width of the cylinder, there exists a width $\omega_0 < 3$ for which there is at least one other isotopy class among our selected classes with minimal length exactly $\ell(\gamma)$.

We now imitate this construction, and repeat it iteratively, but for more than one curve. Specifically, consider a set $\gamma_1, \dots, \gamma_k$ of curves with maximal minimal length among our chosen classes on a surface S^i with boundary, obtained after i metric modifications of $S_{g_n}^{\varepsilon'}$. We consider an ε neighborhood around them, again as in the original construction of $S_{g_n}^{\varepsilon'}$. This gives us a sub-surface of S^i with boundary curves $\delta_1, \dots, \delta_b$. We insert a Euclidean cylinder of width ω along each of the δ_k . As above, this process will make any curve that intersects at least one of the γ_k increase. And again, if there is at least one free homotopy class in our set of chosen classes that intersects one of these curves, because the length of all the γ_k is less or equal to 3, there exists an ω_0 for which a new homotopy class now has minimal length equal to L .

We now observe the following: between any two free homotopy classes h and \tilde{h} in our chosen set, there exist a sequence of $h^k \in \mathcal{H}$ with $h^0 = h$, $h^p = \tilde{h}$ and $i(h^i, h^{i+1}) = 1$. In particular, this guarantees that the above process leaves no curve in \mathcal{H} isolated and finishes in at most N steps where N is the cardinality of \mathcal{H} .

We can now summarize the result of the construction: via a finite number of insertions of cylinders into a surface with closed boundary, we now have a surface \tilde{S} of genus g_n , also with boundary, which has N distinct free homotopy classes of curves of minimal length L and where all other (non-peripheral) curves have length at least ~ 4 . We now complete the description of the metric by using a well-known trick: for each of the boundary curves $\delta \subset \partial\tilde{S}$, we glue a round hemisphere of equator length $\ell(\delta)$. On the resulting surface

S , it is easy to see that *any* minimal geodesic will be completely disjoint from any of the hemispheres and will lie completely in the subsurface \tilde{S} .

In particular, the systoles we constructed in \tilde{S} are systoles of S . As $n \sim \sqrt{g_n}$, we have $N \sim g_n^{3/2}$ and the resulting surface has the desired properties. \square

Remark 3.3. There is another construction of a surface, much simpler, that also underlies the difference between hyperbolic and non-hyperbolic surfaces. Consider a single ribbon n -pod (for some small $\varepsilon > 0$) with $n = 4m$. Now pair opposite edges in the obvious way. The result is a surface of genus $g = m = n/4$ with one boundary component. As above, the curves from the original n -pod are of length exactly 1, and because n is even, they are smooth curves on the surfaces. These curves are in fact systoles: any curve with more complicated combinatorics has length at least ~ 2 and any curve that stays in a single ribbon has length at least 1. (Each of these curves has a family of parallel isotopic systoles as well.) As above, we glue a Euclidean hemisphere to the single boundary curve of the surface. The interest in this construction is that in the center vertex of the embedded n -pod, there are exactly $n^2/2 = 2g$ non-isotopic systoles that intersect in a single point. This is again in striking difference to what is possible for hyperbolic surfaces in view of Lemma 2.4.

REFERENCES

- [1] ADAMS, C. Maximal cusps, collars, and systoles in hyperbolic surfaces. *Indiana Univ. Math. J.* 47, 2 (1998), 419–437.
- [2] ANDERSON, J. W., PARLIER, H., AND PETTET, A. Small filling sets of curves on a surface. *Topology Appl.* 158, 1 (2011), 84–92.
- [3] BUSER, P. *Geometry and spectra of compact Riemann surfaces*, vol. 106 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1992.
- [4] BUSER, P., AND SARNAK, P. On the period matrix of a Riemann surface of large genus. *Invent. Math.* 117, 1 (1994), 27–56. With an appendix by J. H. Conway and N. J. A. Sloane.
- [5] CONWAY, J. H., AND SLOANE, N. J. A. *Sphere packings, lattices and groups*, vol. 290 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1988. With contributions by E. Bannai, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.
- [6] FREEDMAN, M., HASS, J., AND SCOTT, P. Closed geodesics on surfaces. *Bull. London Math. Soc.* 14, 5 (1982), 385–391.

- [7] GROMOV, M. Filling Riemannian manifolds. *J. Differential Geom.* 18, 1 (1983), 1–147.
- [8] GROMOV, M. Systoles and intersystolic inequalities. In *Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992)*, vol. 1 of *Sémin. Congr. Soc. Math. France*, Paris, 1996, pp. 291–362.
- [9] JENNI, F. Über den ersten Eigenwert des Laplace-Operators auf ausgewählten Beispielen kompakter Riemannscher Flächen. *Comment. Math. Helv.* 59, 2 (1984), 193–203.
- [10] KATZ, M. G., SCHAPS, M., AND VISHNE, U. Logarithmic growth of systole of arithmetic Riemann surfaces along congruence subgroups. *J. Differential Geom.* 76, 3 (2007), 399–422.
- [11] MALESTEIN, J., RIVIN, I., AND THERAN, L. Topological Designs. *ArXiv e-prints* (Aug. 2010), 1008.3710.
- [12] MCSHANE, G., AND PARLIER, H. Multiplicities of simple closed geodesics and hypersurfaces in Teichmüller space. *Geom. Topol.* 12, 4 (2008), 1883–1919.
- [13] PFENDER, F., AND ZIEGLER, G. M. Kissing numbers, sphere packings, and some unexpected proofs. *Notices Amer. Math. Soc.* 51, 8 (2004), 873–883.
- [14] RINGEL, G., AND YOUNGS, J. W. T. Solution of the Heawood map-coloring problem. *Proc. Nat. Acad. Sci. U.S.A.* 60 (1968), 438–445.
- [15] SCHMUTZ, P. Riemann surfaces with shortest geodesic of maximal length. *Geom. Funct. Anal.* 3, 6 (1993), 564–631.
- [16] SCHMUTZ, P. Congruence subgroups and maximal Riemann surfaces. *J. Geom. Anal.* 4, 2 (1994), 207–218.
- [17] SCHMUTZ, P. Systoles on Riemann surfaces. *Manuscripta Math.* 85, 3-4 (1994), 429–447.
- [18] SCHMUTZ, P. Arithmetic Fuchsian groups and the number of systoles. *Math. Z.* 223, 1 (1996), 13–25.
- [19] SCHMUTZ, P. Compact Riemann surfaces with many systoles. *Duke Math. J.* 84, 1 (1996), 191–198.
- [20] SCHMUTZ SCHALLER, P. Extremal Riemann surfaces with a large number of systoles. In *Extremal Riemann surfaces (San Francisco, CA, 1995)*, vol. 201 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 1997, pp. 9–19.

- [21] SCHMUTZ SCHALLER, P. Geometry of Riemann surfaces based on closed geodesics.
Bull. Amer. Math. Soc. (N.S.) 35, 3 (1998), 193–214.