

Short homotopically independent loops on surfaces

STEVE KARAM

In this paper, we are interested in short homologically and homotopically independent loops based at the same point on Riemannian surfaces and metric graphs.

First, we show that for every closed Riemannian surface of genus $g \geq 2$ and area normalized to g , there are at least $\lceil \log(2g) + 1 \rceil$ homotopically independent loops based at the same point of length at most $C \log(g)$, where C is a universal constant. On the one hand, this result substantially improves Theorem 5.4.A of M Gromov in [7]. On the other hand, it recaptures the result of S Sabourau on the separating systole in [12] and refines his proof.

Second, we show that for any two integers $b \geq 2$ with $1 \leq n \leq b$, every connected metric graph Γ of first Betti number b and of length b contains at least n homologically independent loops based at the same point and of length at most $24(\log(b) + n)$. In particular, this result extends Bollobàs, Szemerédi and Thomason's $\log(b)$ bound on the homological systole to at least $\log(b)$ homologically independent loops based at the same point. Moreover, we give examples of graphs where this result is optimal.

30F10

1 Introduction

Short homotopically and homologically independent loops on surfaces have been of great interest. Gromov proved in [7; 8] that both the systole $\text{sys}(M)$, ie the length of the shortest noncontractible loop, and the homological systole $\text{sys}_H(M)$, ie the length of the shortest homologically nontrivial loop, of a closed Riemannian surface M of genus $g \geq 2$ with area normalized to $4\pi(g - 1)$ are at most $\sim \log(g)$. In [1], F Balacheff, S Sabourau and H Parlier found the maximal number of homologically independent loops of length at most $\sim \log(g)$. Their theorem goes as follows.

Theorem 1.1 [1] *Let $\eta: \mathbb{N} \rightarrow \mathbb{N}$ be a function such that $\lambda := \sup_g \frac{\eta(g)}{g} < 1$. Then there exists a constant C_λ such that for every closed Riemannian surface M of genus g and area normalized to g there are at least $\eta(g)$ homologically independent loops $\alpha_1, \dots, \alpha_{\eta(g)}$ that satisfy*

$$\text{length}(\alpha_i) \leq C_\lambda \log(g + 1),$$

for every $i \in \{1, \dots, \eta(g)\}$.

Moreover, they constructed some hyperbolic surfaces where their bound is optimal.

For the applications we have in mind (see Section 2), it would be nice if the loops in Theorem 1.1 were based at the same point. Unfortunately, the following example shows that in general, we cannot even find two homologically independent loops based at the same point satisfying a $\log(g)$ bound. Indeed, let M be a closed hyperbolic surface of genus g . Consider a family of $g + 1$ loops in M dividing the surface into two spheres with $g + 1$ boundary components. Pinching these loops enough, we force (by the collar theorem) every loop of M homologically independent from this family to be arbitrary long. Still, we obtain some results in this direction when the systole is bounded from below; see Theorem 4.4.

This leads us to replace the notion of homologically independent loops with the notion of homotopically independent loops defined below.

Definition 1.2 Let M be a closed Riemannian surface of genus at least one. A family of loops $(\alpha_1, \dots, \alpha_k)$ based at the same point v in M are said to be *homotopically independent* if the subgroup of $\pi_1(M, v)$ generated by $\alpha_1, \dots, \alpha_k$ is free of rank k .

Observe that k homologically independent loops based at the same point on a closed surface M of genus g are homotopically independent for $k < 2g$; see Theorem 4.3.

Now, for how many homotopically independent loops based at the same point does the $\log(g)$ bound hold?

One might wonder or even doubt the benefit of finding short homotopically independent loops based at the same point. We show the benefits of such a choice in Section 2. To the author's best knowledge, the only answer to the previous question is due to Gromov.

Theorem 1.3 [8, 5.4.B] *Let (M, h) be a closed Riemannian surface of genus $g \geq 2$ and of area normalized to g . For every $\alpha < 1$, there exist two homotopically independent loops γ_1 and γ_2 based at the same point in M such that*

$$\sup(\text{length}(\gamma_1), \text{length}(\gamma_2)) \leq C_\alpha g^{1-\alpha},$$

where C_α is a positive constant that depends only on α .

Note that Theorem 1.3 does not hold for $\alpha = 1$. Indeed, P Buser and P Sarnak constructed in [4] hyperbolic surfaces with injectivity radius $\sim \log(g)$ at every point. We improve Theorem 1.3 by showing the following result.

Throughout this paper for a positive real number R , we denote by $\lceil R \rceil$ the smallest integer greater or equal to R .

Theorem A *Let M be a closed Riemannian surface of genus $g \geq 2$ and area normalized to g . Then there are at least $\lceil \log(2g) + 1 \rceil$ homotopically independent loops $\alpha_1, \dots, \alpha_{\lceil \log(2g) + 1 \rceil}$ based at the same point in M , such that for every $i \in \{1, \dots, \lceil \log(2g) + 1 \rceil\}$,*

$$\text{length}(\alpha_i) \leq C \log(g),$$

where C is a universal constant independent from the genus.

Theorem A substantially improves Theorem 1.3. Under the same hypothesis as Theorem 1.3, Theorem A guarantees the existence of $\lceil \log(2g) + 1 \rceil$ homotopically independent loops based at the same point (instead of two) of length roughly bounded by $\log(g)$ (instead of g^α). Note that, if the homotopical systole of the surface M in Theorem A is bounded away from zero, then the $\lceil \log(2g) + 1 \rceil$ loops can be chosen to be homologically independent (see Theorem 4.4). Also Theorem A recaptures the following result by S Sabourau.

Theorem 1.4 (Sabourau [12]) *There exists a positive constant C such that every closed Riemannian surface M of genus $g \geq 2$ and area normalized to g satisfies*

$$\text{sys}_0(M) \leq C \log(g),$$

where $\text{sys}_0(M)$ is defined as the length of the shortest noncontractible loop in M which is trivial in $H_1(M, \mathbb{Z})$.

Note that Sabourau splits his proof into two cases. In the first case, he supposes that $\text{sys}_0(M) \leq 4 \text{sys}(M)$ and then he deduces the result from Gromov's $\log(g)$ bound on the systole. Meanwhile, Theorem A provides a unified proof of this theorem without referring to Gromov's asymptotic systolic inequality.

Gromov's $\log(g)$ bound on the systole has an analog for metric graphs. Note that for a metric graph Γ , the homotopical systole coincides with the homological systole. We will denote it by $\text{sys}(\Gamma)$. The best bound on the systole of a metric graph is due to Bollobás and Szemerédi [2] and Bollobás and Thomason [3]. Specifically, they proved that the systole of every connected metric graph of first Betti number $b \geq 2$ and length normalized to b satisfies

$$(1-1) \quad \text{sys}(\Gamma) \leq 4 \log(b + 1).$$

Exactly as for surfaces, given a metric graph of first Betti number $b \geq 2$ and of length normalized to b , one might wonder about the number of homologically independent loops based at the same point satisfying the Bollobás, Szemerédi and Thomason $\log(b)$ bound. We answer this question here.

Theorem B *Let Γ be a connected metric graph of first Betti number $b \geq 2$ and of length normalized to b . Let $n \in \{1, \dots, b\}$. There exist at least n homologically independent loops in Γ based at the same point and of length at most $24(\log(b) + n)$.*

An interesting value of n is $n = \lfloor \log(b) \rfloor$, ie, the integral part of $\log(b)$. In this case, Theorem B asserts that for every connected metric graph Γ of first Betti number $b \geq 2$ and of length b , there exist at least $\lfloor \log(b) \rfloor$ homologically independent loops based at the same point of length at most $48 \log(b)$. This extends Bollobàs, Szemerédi and Thomason's $\log(b)$ bound on the homological systole of Γ to $\lfloor \log(b) \rfloor$ homologically independent loops of Γ based at the same point.

One might wonder how far from being optimal Theorem B is. We show that it cannot be substantially improved. Indeed, let b and n be two integers such that $b \geq 2$ and $1 \leq n \leq b$. There exists a connected metric graph of first Betti number b and length normalized to b , such that there are at most $\lfloor 24(\log(b) + n) \rfloor + 1$ homologically independent loops in Γ based at the same point of length at most $24(\log(b) + n)$ (cf Theorem 3.2). In particular, this result shows that for $n \geq \lceil \log(b) \rceil$, there exists a connected metric graph Γ of first Betti number and length normalized to b , such that there are at most $49n$ homologically independent loops in Γ based at the same point of length at most $24(\log(b) + n)$.

This paper is organised as follows. In Section 2, we show the benefits of short homotopically independent loops based at the same point. In Section 3, we give the proof of Theorem B. In Section 4, we show how to extend Theorem B to closed surfaces with systole bounded away from zero. In Section 5, we show that on a given closed surface the cut locus of a simple closed geodesic captures its topology. In Section 6, we prove Theorem A.

Acknowledgments The author would like to thank his advisor, Stéphane Sabourau, for many useful discussions and valuable comments. He also would like to thank Florent Balacheff for reading and commenting on this paper.

2 Benefits of short homotopically independent loops based at the same point

In this section, we show two applications of homotopically independent loops based at the same point of bounded length.

Let M be a closed Riemannian surface of genus $g \geq 2$. If α and β are two homotopically independent loops based at the same point in M , then

$$\text{sys}_0(M) \leq \text{length}(\alpha\beta\alpha^{-1}\beta^{-1}).$$

In particular, if $\sup(\text{length}(\alpha), \text{length}(\beta)) \leq C \log(g)$, then

$$\text{sys}_0(M) \leq 4C \log(g).$$

Notice that the above observation allows us to recapture the result of Theorem 1.4 on the separating systole by means of Theorem A. Also we would like to point out that Gromov’s upper bound $C_\alpha g^{1-\alpha}$ on the length of two homotopically independent loops based at the same point in Theorem 1.3 is not sufficient to prove that the length of the separating systole of a closed Riemannian surface of genus $g \geq 2$ and area g is bounded above by $\sim \log(g)$.

Another use of homotopically independent loops based at the same point v of a closed Riemannian surface M is to contribute to the area of balls centered at a lift \tilde{v} of v in the universal cover \tilde{M} of M . Let us clarify this idea here. Consider a system $S = \{\alpha_1, \dots, \alpha_k\}$ of pairwise nonhomotopic loops based at v . Let

$$L = \sup_{1 \leq i \leq k} \text{length}(\alpha_i).$$

Denote by s half the systole of M at the point v , ie half the length of the shortest noncontractible loop based at v . Let H'_r (resp. N_r) be the set of elements of $H = \langle S \rangle$ (resp. $\pi_1(M, v)$) of length less than r , where the length of $\alpha \in \pi_1(M, v)$ is defined as $\text{length}(\alpha) = \text{dist}(\tilde{v}, \alpha.\tilde{v})$. It is the minimal length of a loop based at v representing α . Let $R > s + L$. Consider the ball $B = B_{\tilde{M}}(\tilde{v}, r_0)$, where $r_0 = R - s$. Every element γ_i of N_{r_0} yields a point $\tilde{v}_i = \gamma_i.\tilde{v}$ in B . The balls $B_{\tilde{M}}(\tilde{v}_i, s)$ are disjoint and of the same area. We have

$$(2-1) \quad \text{Area } B_{\tilde{M}}(\tilde{v}, R) \geq \text{card}(N_{r_0}) \text{Area } B_M(v, s),$$

where $\text{card}(N_{r_0})$ is the cardinal of N_{r_0} .

Also notice that

$$(2-2) \quad \text{card}(N_{r_0}) \geq \text{card}(H'_{r_0}).$$

Thus, a lower bound on the cardinal $\text{card}(H'_{r_0})$ of H'_{r_0} yields also a lower bound on $\text{card}(N_{r_0})$. One way to bound $\text{card}(H'_{r_0})$ from below is the following. We define a norm $\|\cdot\|$ on H as follows. For β in H , we define the word length $\|\beta\|$ of β as the smallest integer n such that $\beta = \gamma_1 \cdots \gamma_n$, where $\gamma_i \in S \cup S^{-1}$. Denote by H_r^w the set of elements of H of word length less than r . We have

$$(2-3) \quad \text{card}(N'_r) \geq \text{card}(H_{r/L}^w).$$

Combining (2-1), (2-2) and (2-3) we get

$$(2-4) \quad \text{Area } B_{\tilde{M}}(\tilde{v}, R) \geq \text{card}(H_{r/L}^w) \text{Area } B_M(v, s).$$

Now let $r' > 1$. Notice that $H_{r'}^w$ is maximal if H is free of rank k . That is guaranteed if the loops $\alpha_1, \dots, \alpha_k$ are homotopically independent in M . It is now clear how homotopically independent loops based at the same point v contribute to the area of the balls centered at points in the fiber over v in \tilde{M} whenever the radii R of these balls is longer than $s + L$. Moreover, since R must be at least $s + L$, it is straightforward to see that the shorter the L , the better the result. This means that the upper bound of the lengths of the α_i is also important.

3 Short homologically independent loops on graphs

In this section we prove Theorem B. Recall that this theorem extends the Bollobás–Szemerédi–Thomason $\log(b)$ bound on the homological systole of graphs to $\lceil \log(g) \rceil$ homologically independent loops based at the same point.

First let us recall some definitions. By definition, a graph Γ is a finite one-dimensional CW-complex (multiple edges and loops are allowed). The first Betti number of a graph Γ can be computed as

$$b(\Gamma) = e - v + n,$$

where e, v and n are respectively the number of edges, vertices and connected components of Γ . A metric graph (Γ, h) is a graph endowed with a length space metric h . The length of a subgraph of Γ is its one-dimensional Hausdorff measure. For more details on graphs we refer the reader to Diestel [6].

Definition 3.1 Let Γ be a connected graph of first Betti number $b \geq 1$. A family of loops $(\alpha_1, \dots, \alpha_k)$ in Γ is said to be *homologically independent* if their homology classes in $H_1(\Gamma, \mathbb{R})$ are so.

Note that this definition extends also to closed Riemannian manifolds.

Now we prove Theorem B.

Theorem B *Let Γ be a connected metric graph of first Betti number $b \geq 2$ and of length normalized to b . Let $n \in \{1, \dots, b\}$. There exist at least n homologically independent loops in Γ based at the same point and of length at most $24(\log(b) + n)$.*

Proof By definition of the first Betti number b , there exist b homologically independent loops $\alpha_1, \dots, \alpha_b$ in Γ . Fix a point x of α_1 . For $i = 1, \dots, b$, let C_i be a minimizing curve from x to α_i . We have $\text{length}(C_i \alpha_i C_i^{-1}) \leq \text{length}(C_i) + \text{length}(\alpha_i) + \text{length}(C_i)$. Notice that $\text{length}(C_i) + \text{length}(\alpha_i) \leq b$. Thus, there exists b

homologically independent loops in Γ based at the same point of length at most $2b$ ($\leq 24(\log(b) + b/2)$). This yields the desired result for $n \in \{b/2, b\}$. Now we consider the case $n < b/2$. In particular, we suppose $b \geq 3$. By a short cycle of Γ we mean a simple loop of length at most $12 \log(b)$. Let X be a maximal set of homologically independent short cycles of Γ and denote by N its cardinal. We claim that

$$N \geq \frac{b}{2}.$$

Indeed, we construct $k = \lceil b/2 \rceil$ graphs $\Gamma_k \subset \dots \subset \Gamma_1 = \Gamma$ and k simple loops as follows. Remove an edge from a systolic loop γ_1 of Γ_1 and denote by Γ_2 the resulting graph. The graph Γ_2 is connected and of first Betti number $b_2 = b - 1$. Now remove an edge from a systolic loop γ_2 of Γ_2 and denote by Γ_3 the resulting graph. By induction, we keep doing this until we get Γ_k . From the inequality (1-1) and since $k = \lceil b/2 \rceil$ we have, for every $i = 1, \dots, k$,

$$\text{length}(\gamma_i) \leq 4 \frac{\log(1 + b - i + 1)}{b - i + 1} \text{length}(\Gamma_i) \leq 12 \log(b).$$

By construction, the k loops $\{\gamma_i\}_{i=1}^k$ are homologically independent in Γ . So the claim is proved.

We divide the set X as follows. Take any element α_1 of X and denote by Y_1 the set $\{\beta \in X \mid \text{dist}(\beta, \alpha_1) \leq 4n\}$. Let α_2 be an element of $X \setminus Y_1$ and denote by Y_2 the set $\{\beta \in X \mid \text{dist}(\beta, \alpha_2) \leq 4n\}$. By induction we continue this process which eventually ends since X is finite. Let $\alpha_j \in X$ be the last short cycle obtained from this process, ie, let α_j be an element of $X \setminus Y_1 \cup \dots \cup Y_{j-1}$ such that $Y_1 \cup \dots \cup Y_{j-1} \cup Y_j = X$. For $i = 1, \dots, j$, we denote by T_i the cardinal of Y_i . We claim that there exists an i_0 such that

$$T_{i_0} \geq n.$$

Indeed, suppose the opposite. We have

$$\frac{b}{2} \leq N = \text{card}(X) \leq \sum_{i=1}^j T_i < jn.$$

So $j > \frac{b}{2n} > 1$. For $i \neq i'$, we have $\text{dist}(\alpha_i, \alpha_{i'}) > 4n$. This means that the open neighborhoods of radius $2n$ around the α_i are pairwise disjoint. Since Γ is connected, the length of the neighborhood of radius $2n$ around each short cycle α_i is at least $\text{length}(\alpha_i) + 2n$. This implies that

$$\text{length}(\Gamma) > 2nj > b.$$

Hence we have a contradiction. So there is an i_0 such that $T_{i_0} \geq n$.

Now fix a vertex a of α_{i_0} and let β be any element of $Y_{i_0} \setminus \{\alpha_{i_0}\}$. Let b and c be two vertices of α_{i_0} and β respectively such that $\text{dist}(\alpha_{i_0}, \beta) = \text{dist}(b, c)$. Also, let C_{ab} be a minimizing curve from a to b and C_{bc} be a minimizing curve from b to c . The following hold:

- $\text{length}(C_{ab}) \leq \text{length}(\alpha_{i_0})/2$
- $\text{length}(C_{bc}) \leq 4n$

The loop $\beta' = C_{ab}C_{bc}\beta C_{cb}C_{ba}$ is homologous to β and satisfies

$$\text{length}(\beta') \leq 24 \log(b) + 8n.$$

So the T_{i_0} short cycles of Y_{i_0} give rise to T_{i_0} homologically independent loops of Γ based at the same point a and of length at most $24(\log(b) + n)$. \square

Before stating our next theorem, we construct a connected metric graph Γ_\star that will be useful to the rest of this section. Let m and p be two positive integers with $m \geq p$. Denote by q and r the quotient and the remainder in the division of m by p , that is, $m = pq + r$ with $r \in \{0, \dots, p-1\}$. Also let L and l be two positive constants.

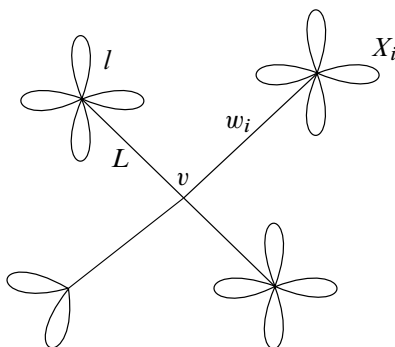


Figure 1: The graph Γ_\star for $m = 12$, $p = 4$, $q = 3$ and $r = 2$

Fix a vertex v . We construct q bouquets X_1, \dots, X_q of p circles and a bouquet X_{q+1} of r circles. We define Γ_\star by joining the vertex of each bouquet X_i to the vertex v by an edge w_i ; see Figure 1. We define a metric h on Γ_\star such that (Γ_\star, h) is a length metric space as follows. For $i = 1, \dots, q$, set $\text{length}(w_i) = L$, and $\text{length}(X_i) = l$. Also set $\text{length}(X_{q+1}) + \text{length}(w_{q+1}) = r$. It is straightforward to see that the graph Γ_\star is connected, of first Betti number m and of length $q(L + l) + r$. We claim that there are at most $p + r$ ($\leq 2p - 1$) homologically independent loops based at the same point of

length at most $2L$. Indeed, notice that there exist at most r homologically independent loops based at v of length less than $2L$. So let m be any point of Γ_\star other than the point v . There exists a unique i such that $m \in X_i \cup w_i$. Now notice that if we want to find more than $p + r$ homologically independent loops based at m , one of them must cross at least two times one of the edges w_j , with $j \in \{1, \dots, q\} \setminus \{i\}$. Thus, the length of this loop exceeds $2L$.

Our next theorem shows that one cannot substantially improve Theorem B, thus it is roughly optimal.

Theorem 3.2 *Let b and n be two integers such that $b \geq 2$ and $1 \leq n \leq b$. Let $\lambda > 0$. There exists a connected metric graph of first Betti number b , of length normalized to b , such that there are at most $\lfloor \lambda(\log(b) + n) \rfloor + 1$ homologically independent loops in Γ based at the same point of length at most $\lambda(\log(b) + n)$*

Proof We only need to consider the case when $b \geq \lfloor \lambda(\log(b) + n) \rfloor + 1$ since the other case is trivial. Denote by q and r respectively the quotient and the remainder in the division of b by $\lfloor \frac{\lambda}{2}(\log(b) + n) \rfloor + 1$. Let $\varepsilon > 0$ be such that

$$\lfloor \frac{\lambda}{2}(\log(b) + n) \rfloor + 1 = \frac{\lambda}{2}(\log(b) + n) + \varepsilon.$$

Consider the graph Γ_\star given by the previous construction with:

- $m = b$
- $p = \lfloor \frac{\lambda}{2}(\log(b) + n) \rfloor + 1$
- $L = \frac{\lambda}{2}(\log(b) + n)$
- $l = \varepsilon$

The graph Γ_\star is connected, of first Betti number b , of length b and has at most $\lfloor \lambda(\log(b) + n) \rfloor + 1$ homologically independent loops based at the same point of length at most $\lambda(\log(b) + n)$. \square

4 Short homologically independent loops on surfaces with homotopical systole bounded from below

In this section we combine ideas from [1] and the author [10] to extend Theorem B to closed surfaces with systole bounded below.

Definition 4.1 Let (M, h) be a closed Riemannian surface of genus g . The image in M of an abstract graph by an embedding will be referred to as a graph in M . The metric h on M naturally induces a metric on a graph Γ in M . Despite the risk of confusion, we will also denote by h such a metric on Γ .

Proposition 4.2 Let (M, h) be a closed Riemannian surface of genus $g \geq 1$. Suppose that the homotopical systole of M is at least ℓ . Then, there exists a graph Γ in M such that

- (1) the inclusion map $i: \Gamma \rightarrow M$ is distance nonincreasing,
- (2) the homomorphism $i_*: H_1(\Gamma, \mathbb{R}) \rightarrow H_1(M, \mathbb{R})$ induced by the inclusion is an isomorphism,
- (3)
$$\text{length}(\Gamma) \leq \frac{2^9 \text{Area}(M, h) + g}{\min\{1, \ell\}}.$$

Proof Without loss of generality, we suppose that $\ell \leq 1$. This proposition is the same as [10, Proposition 6.1], where ℓ was taken to be $\frac{1}{2}$ and the area is equal to $\frac{1}{2^{11}}(2g - 1)$ instead of g . The proof of Proposition 6.1 in [10] starts by fixing $r_0 = \frac{1}{2^5}$. In our case we fix $r_0 = \frac{\ell}{2^4}$ and reproduce the argument. \square

Before stating out next theorem, let us recall the following theorem.

Theorem 4.3 (Jaco [9]) Let M be a closed Riemann surface of Euler characteristic $\chi(M) \leq 0$. Any subgroup of $\pi_1(M)$ generated by k elements, where $k < 2 - \chi(M)$, is a free group.

Now we can prove the following result.

Theorem 4.4 Let M be a closed orientable Riemannian surface of genus $g \geq 1$ with homotopical systole at least ℓ and area normalized to g . Let $n \in \{1, \dots, 2g\}$ be an integer. There exist at least n homologically independent loops $\gamma_1, \dots, \gamma_n$ based at the same point in M such that for every $i = 1, \dots, n$, we have

$$\text{length}(\gamma_i) \leq 24C_\ell(\log(2g) + n),$$

where $C_\ell = \frac{2^9}{\min\{1, \ell\}}$. Moreover, if $n < 2g$ then $\langle \gamma_1, \dots, \gamma_n \rangle$ is free of rank n .

Proof Let Γ be a graph in M that satisfies (1), (2) and (3) of Proposition 4.2. The first Betti number of Γ is $2g$. By Theorem B, there are at least n homologically independent loops in Γ based at the same vertex of length at most $24C_\ell(\log(2g) + n)$. The images of these loops by the inclusion map i yield the desired loops. The second assumption follows from Theorem 4.3. \square

Remark 4.5 A nonorientable version of Theorem 4.4 holds. Let M be a closed nonorientable surface of genus $g \geq 1$ with homotopical systole at least ℓ and area normalized to g . Let $n \in \{1, \dots, g\}$. There are at least n loops $\gamma_1, \dots, \gamma_n$ based at the same point v in M whose homology classes in $H_1(M, \mathbb{Z}_2)$ are independent such that for every $i = 1, \dots, n$, we have

$$\text{length}(\gamma_i) \leq 24C'_\ell(\log(g) + n),$$

where $C'_\ell = C/\min\{1, \ell\}$ for some positive constant C . Moreover, if $n < g$ then $\langle \gamma_1, \dots, \gamma_n \rangle$ is free of rank n .

5 Cut loci and capturing the topology

In this section we extend the notion of cut locus, defined originally for points in a Riemannian manifold, to simple closed geodesics (this might be already defined but the author did not find a reference in the literature) and we give some basic results for the new notion.

Let M be a closed surface and p be a point in M . The cut point of p along a geodesic C_p starting at p is the first point $q \in C_p$ such that the arc of C_p between p and any point r on C_p after q is no longer minimizing. The set $\text{Cut}(p)$ of all cut points along all the geodesics issued from p is called the cut locus of p . We extend this notion to simple closed geodesics as follows.

Let $\alpha: [0, l] \rightarrow M$ be a simple closed geodesic in M and β be another geodesic that starts orthogonally from α at some point p . The cut point of α along β is the first point $q \in \beta$ such that, for any point r on β beyond q the length of the arc of β between p and r no longer agrees with the distance from r to α . The set $\text{Cut}(\alpha)$ of all the cut points of all the geodesics issued orthogonally from α is called the cut locus of α . An alternative useful way to view $\text{Cut}(\alpha)$ is the following. Denote by $N\alpha$ the normal bundle to α . Each vector $v_t \in N\alpha$ gives rise to a geodesic C_t starting at $\alpha(t)$ such that $C_t'(0) = v_t$. Denote by q_t the cut point of α along the geodesic C_t . The point q_t is the image by the exponential map of some vector v'_t parallel to v_t . Let N_1 be the set of the vectors v'_t and N_2 be the set of the vectors $\lambda v'_t$, where $\lambda \in [0, 1)$. Then $\text{Cut}(\alpha) = \exp(N_1)$.

Lemma 5.1 We have $M = \exp(N_1) \cup \exp(N_2)$, where the union is disjoint.

Proof Let x be a point in M . There exists a minimizing geodesic σ_x^{-1} from x to α parametrized by arc length such that $\text{length}(\sigma_x^{-1}) = \text{dist}(x, \alpha)$. The geodesic σ_x^{-1} hits α orthogonally in a point $\alpha(t)$ (cf do Carmo [5]). Since σ_x is

minimizing, the point x is not after the cut point of α along σ_x . That means that the vector $\text{dist}(\alpha(t), x)\sigma'_x(0) \in N_1 \cup N_2$. Notice that $x = \exp(\text{dist}(\alpha(t), x)\sigma'_x(0))$. Thus, $M = \exp(N_1) \cup \exp(N_2)$.

Now let us prove that the union is disjoint. Let $y \in \exp(N_1) \cap \exp(N_2)$. Since $y \in \exp(N_2)$, there exists a minimizing geodesic $\sigma_y: [0, \ell] \rightarrow M$ from α to y , parametrized by arc length such that σ_y is still minimizing for some time after y ie there exists an $\varepsilon > 0$ such that $\sigma_y: [0, \ell + \varepsilon]$ is a minimizing geodesic from α to $\sigma_y(\ell + \varepsilon)$. On the other hand, since $y \in \exp(N_1)$, there exists a minimizing geodesic δ_y from α to y parametrized by arc length such that δ_y is no longer minimizing after y . Let ϕ be the curve defined by $\phi(t) = \delta_y(t)$ if $t \in [0, \ell]$, and $\phi(t) = \sigma_y(t)$ for $t \in [\ell, \ell + \varepsilon]$. Let $0 < \varepsilon' < \varepsilon$. There exists a minimizing geodesic from $\phi(\ell - \varepsilon')$ to $\phi(\ell + \varepsilon')$ which is of length strictly less than the arc of ϕ between these two points since ϕ is not smooth at $\phi(\ell)$. We conclude that $\text{dist}(\sigma_y(\ell + \varepsilon'), \alpha)$ is strictly less than the length of σ_y between $\sigma_y(0)$ and $\sigma_y(\ell + \varepsilon')$. Hence we have a contradiction. So the proof is finished. \square

Lemma 5.2 *The set $\text{Cut}(\alpha)$ is a deformation retract of $M \setminus \{\alpha\}$.*

Proof Let x be a point of M not in α or $\text{Cut}(\alpha)$. Denote by σ_x the unique minimizing geodesic from x to α . Let x' be the cut point of α along the geodesic σ_x . Clearly, $x' \in \text{Cut}(\alpha)$. Now we can shrink $M \setminus \{\alpha\}$ to $\text{Cut}(\alpha)$ by sliding each point x of M not in α or $\text{Cut}(\alpha)$ to $\text{Cut}(\alpha)$ along the arc of the geodesic σ_x between x and x' . \square

In view of this lemma, we will say that $\text{Cut}(\alpha)$ captures the topology of $M \setminus \{\alpha\}$.

Proposition 5.3 *Let (M, g) be a closed real analytic Riemannian surface and α be a simple closed geodesic in M . Then $\text{Cut}(\alpha)$ is a finite graph.*

We omit the proof of Proposition 5.3 since it is essentially the same proof as in Myers [11, page 97].

6 Short homotopically independent loops on Riemannian surfaces

In this section we prove Theorem A. Before doing that, let us give some definitions and some independent propositions that will be useful to the rest of this section.

Lemma 6.1 Let $F = \langle a, b \rangle$ be a free subgroup of rank 2 of the fundamental group of a closed Riemannian manifold. The subgroup $H = \langle b, a^1ba^{-1}, \dots, a^{n-1}ba^{-(n-1)} \rangle$ of F is free of rank n for every integer $n \geq 1$. Moreover, if $\text{length}(a) = l_a$ and $\text{length}(b) = l_b$, then

$$\sup_{0 \leq i \leq n-1} \text{length}(a^i b a^{-i}) \leq 2(n-1)l_a + l_b.$$

Proof Since the subgroup of a free group is free then H is free. Next, we claim that the generator $a^p b a^{-p}$ is not an element of the free subgroup G generated by the elements $a^q b a^{-q}$ for $q \in \{0, \dots, n-1\} \setminus \{p\}$. Indeed, a reduced word in G starts with a^q with $q \neq p$. So H is of rank n . The length inequality is immediate. \square

Proposition 6.2 Let (M, g) be a compact Riemannian cylinder. Denote by α and β the two boundary components of M . Suppose that

$$\text{length}(\alpha) < 1 < \text{length}(\beta).$$

Then there exists a noncontractible simple loop γ in M of length 1 such that the systole of the cylinder R_γ bounded by β and γ is equal to 1.

In particular, the loop γ is a systolic loop of R_γ .

Proof Let $X = \{\sigma \text{ simple noncontractible loop in } M \text{ such that } \text{sys}(R_\sigma) = 1\}$, where by R_σ we mean the cylinder of boundary components β and σ . Clearly the set X is nonempty. Let $\ell = \inf_{\sigma \in X} \text{length}(\sigma)$ and ε be a small positive constant. By the definition of the infimum, there exists a simple noncontractible loop σ_0 such that $\text{sys}(R_{\sigma_0}) = 1$ with $\ell \leq \text{length}(\sigma_0) \leq \ell + \varepsilon$. The systolic loop γ of R_{σ_0} is a simple noncontractible loop in M . Moreover, we have $R_\gamma \subset R_{\sigma_0}$. Thus

$$1 = \text{sys}(R_{\sigma_0}) \leq \text{sys}(R_\gamma) \leq \text{length}(\gamma) = 1.$$

So $\text{sys}(R_\gamma) = 1$. This finishes the proof. \square

In the proof of Theorem A below, we will need the following definition.

Definition 6.3 Let M be a closed Riemann surface of genus g (with possibly one disk removed). It is well known that such a surface can be obtained from a polygon P (with possibly one disk removed) by pairwise identifications of its sides where all the vertices of P get identified to a single point on x of M . Such a polygon will be called a normal representation of M . After identification, the edges of P give rise to $2g$ simple loops (if M is orientable) or to g simple loops (if M is nonorientable) based at x and intersecting each other only at x . Such a set of loops is called a *canonical system of loops*.

Now we prove Theorem A.

Theorem A *Let M be a closed orientable Riemannian surface of genus $g \geq 2$ and area normalized to g . There are at least $n = \lceil \log(2g) + 1 \rceil$ homotopically independent loops $\alpha_1, \dots, \alpha_n$ based at the same point such that for all $i = 1, \dots, n$,*

$$\text{length}(\alpha_i) \leq 2^{20} \log(g).$$

Proof of Theorem A Since every smooth metric can be approximated by a real analytic one, we can assume that M is a real analytic Riemannian surface. We only need to consider the case where the homotopical systole of M is less than 1, since the other case is settled down by Theorem 4.4. Consider a maximal set X of simple closed geodesics $\alpha_1, \dots, \alpha_p$ of length at most 1 which are pairwise disjoint in M and not freely homotopic. Let k be the number of elements of X that are separating. Note that $k \leq p$. The main idea of the proof is to go back to the case where the homotopical systole is at least 1.

Remark 6.4 At first, we were tempted to cut the surface M open along the loops α_i of X and to attach an hemisphere along each of the $2p$ boundary components. This yields at least $k + 1$ new closed surfaces M_1, \dots, M_{k+1} , where k is the number of geodesics in X that are separating. We hoped to find the desired loops or two short homotopically independent loops based at the same point in one of the closed surfaces M_i . Recall that the homotopical systole of each M_i is at least 1 so we can use Theorem 4.4. Afterwards we wanted to show that these loops do not cross the hemispheres and so lie in the original surface M . It does not take much time to realize that this idea is naive. One can run into many problems. Let us imagine the case where $p = g$ and all of the geodesics α_i are nonseparating like the surface in Figure 2. In this case, the surface obtained by cutting M along the loops α_i and attaching hemispheres is of genus 0 and so the proof collapses. Instead we will cut M along each α_i , chop off some “maximal” cylinders and then glue the boundary components back together to obtain a new surface with systole bounded away from zero.

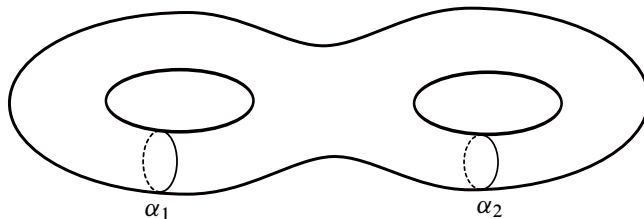


Figure 2

Let $\varepsilon \in \{-, +\}$. We divide the proof into 5 steps.

Step 1 In this step we chop off cylinders corresponding to short separating loops. If $k = 0$, we skip this step and start directly at the second step. By renumbering the α_i if needed, we can suppose that for $i = 1, \dots, k$, the simple closed geodesic α_i is separating. Cut the surface M open along α_1 . We obtain two compact surfaces M^- and M^+ with signatures $(g - m, 1)$ and $(m, 1)$, where m is some positive integer less than g . Denote by α_1^ε the boundary of the surface M^ε and let S^ε be one of its canonical systems of loops. Notice that since the genus of M^ε is at least 1, we have $\text{card}(S^\varepsilon) \geq 2$. We can suppose that for every pair of loops a and b in S^ε , we have $\sup(\text{length}(a), \text{length}(b)) > 1$. Otherwise the proof is finished by Lemma 6.1 since a and b do not commute and so generate a free group of rank 2. Cut the surface M^ε open along the loops in S^ε . This gives rise to a cylinder T^ε with two boundary components α_1^ε and β_1^ε such that $\text{length}(\beta_1^\varepsilon) > 1$. So the cylinder T^ε satisfies the hypothesis of Proposition 6.2. Thus, there exists a noncontractible simple loop γ_1^ε of length 1 which is a systolic loop of the cylinder R_1^ε bounded by β_1^ε and γ_1^ε is 1. Cut T^ε along γ_1^ε and throw away the cylinder C_1^ε bounded by α_1^ε and γ_1^ε . Now reglue R_1^ε by pairwise identifications of the edges of β_1^ε . This gives rise to a compact surface M_1^ε with one boundary component γ_1^ε of length 1. Glue the surfaces M_1^- and M_1^+ along their boundaries γ_1^- and γ_1^+ . The resulting surface M_1 , satisfies the following:

- The surface M_1 has the same genus as the surface M .
- $\text{Area}(M_1) \leq \text{Area}(M)$.
- A minimal representative in M_1 of the free homotopy class of α_1 is given by the simple loop γ_1 of length 1 obtained by gluing γ_1^- and γ_1^+ together.

Repeat the above process with the $k - 1$ remaining elements of X that are separating. This gives rise to a closed surface M_k of the same genus as the surface M such that $\text{Area}(M_k) \leq \text{Area}(M)$. Moreover, any simple closed geodesic of M_k of length less than 1 is nonseparating. Perturbing the metric again, we can suppose again that it is a real analytic one.

Step 2 In this step, we chop off cylinders corresponding to short nonseparating loops. Cut the surface M_k open along α_{k+1} . This leads to a surface N_k with genus $g - 1$ and with two boundary components α_{k+1}^- and α_{k+1}^+ . By Lemma 5.2, we know that the cut locus $\text{Cut}(\alpha_{k+1})$ of α_{k+1} is a deformation retract of $M \setminus \{\alpha_{k+1}\}$. So the fundamental group of $\text{Cut}(\alpha_{k+1})$ is isomorphic to the fundamental group of N_k . Now cut the surface N_k open along $\text{Cut}(\alpha_{k+1})$. This gives rise to two cylinders. The cylinder T_{k+1}^- with boundary components $(\alpha_{k+1}^-, \beta_{k+1}^-)$ and the cylinder T_{k+1}^+

with boundary components $(\alpha_{k+1}^+, \beta_{k+1}^+)$. Arguing as in Step 1, we can suppose that $\text{length}(\beta_{k+1}^\varepsilon) > 1$. So the cylinder T_{k+1}^ε satisfies the hypothesis of Proposition 6.2. Thus there exists a noncontractible simple loop γ_{k+1}^ε of length 1 which is a systolic loop of the cylinder R_{k+1}^ε of boundary components $(\beta_{k+1}^\varepsilon, \gamma_{k+1}^\varepsilon)$ is 1. Cut T_{k+1}^ε open along γ_{k+1}^ε and throw away the cylinder C_{k+1}^ε bounded by α_{k+1}^ε and γ_{k+1}^ε . Now reglue the cylinder R_{k+1}^ε by reidentifying the sides of β_{k+1}^ε . This gives rise to two compact surfaces M_{k+1}^- and M_{k+1}^+ with boundary components that can be pairwise identified. Gluing these two surfaces together we get a closed surface M_{k+1} that satisfies the following:

- The surface M_{k+1} has the same genus as the surface M_k .
- $\text{Area}(M_{k+1}) \leq \text{Area}(M_k)$.
- A minimal representative of the free homotopy class of α_{k+1} in M_{k+1} is given by the simple loop γ_{k+1} of length 1, obtained by gluing γ_{k+1}^- and γ_{k+1}^+ together.

Repeat the above process with the $p-k-1$ remaining elements of X . This gives rise to a closed surface M_p of the same genus as the surface M such that $\text{Area}(M_p) \leq \text{Area}(M)$.

Before proceeding to the next step, recall that the simple closed geodesics $\alpha_1, \dots, \alpha_p$ in the original surface M correspond to the simple closed geodesics $\gamma_1, \dots, \gamma_p$ in the surface M_p . Also recall that the cylinders C_i^- and C_i^+ in M share the same boundary component α_i . We denote by C_i the cylinder with boundary components (γ_i^-, γ_i^+) , that is, $C_i = C_i^+ \cup C_i^-$.

Step 3 In this step, we show we can suppose two different cylinders C_j and $C_{j'}$ in M are distant from each other. Specifically, we have $\text{dist}_M(C_j, C_{j'}) > 2^{18} \log(g)$. In other words, we have

$$(6-1) \quad \text{dist}_{M_p}(\gamma_j, \gamma_{j'}) > 2^{18} \log(g).$$

Indeed, suppose the opposite. Without loss of generality, suppose that the distance between C_j and $C_{j'}$ is equal to $\text{dist}(\gamma_j^-, \gamma_{j'}^-)$. Let z_1 be a point on C_j and z_2 be a point on $C_{j'}$ such that $\text{dist}(z_1, z_2) = \text{dist}(\gamma_j^-, \gamma_{j'}^-)$. Consider the loop μ that starts at z_1 , travels along a minimizing geodesic between z_1 and z_2 , makes a complete tour along $\gamma_{j'}^-$ and then comes back to z_1 . We have that $\text{length}(\mu) \leq 2^{19} \log(g) + 1$. Notice also that μ and γ_j^- do not commute. In particular, they are homotopically independent. So by Lemma 6.1 (take $a = \gamma_j^-$ and $b = \mu$), the proof of the Theorem is finished.

Step 4 In this step, we show that we can suppose that

$$\text{sys}(M_p) \geq 1.$$

Indeed, by contradiction, suppose that there is a systolic loop μ of M_p of length less than 1. We claim that the geodesic μ transversally intersects at least one of the γ_i . Indeed, suppose the opposite, and denote by μ' the simple closed geodesic in the original surface M that corresponds to μ . Since μ does not transversally intersects any of the γ_i , the loop μ' is disjoint from all the cylinders C_i . In particular, μ' does not intersect any of the loops α_i . This contradicts the maximality of X , since $\text{length}(\mu') < 1$.

Let $j \in \{1, \dots, n\}$ be such that μ transversally intersects γ_j . That means that in the surface M , the loop μ' goes across the cylinder C_j . Now we claim that μ intersects only one γ_j . Indeed, the length of μ' is less than 1 and the distance between any pair of cylinders C_j and $C_{j'}$ is greater than 1. Therefore, μ intersects only one γ_j . Moreover, the two minimizing simple loops μ and γ_j do not commute as $g \geq 2$ by assumption.

Lemma 6.5 *Let β be a loop in M_p of length less than L that transversally intersects only one geodesic γ_j and does not commute with it. Then there exist two noncommutative loops a, b in the original surface M based at the same point such that $\text{length}_M(a) = 1$ and $\text{length}_M(b) \leq 2L + 1$. In particular, the loops a and b are homotopically independent.*

Proof We give β and γ_j some orientation. Let x_1, \dots, x_q be the transversal intersection points of β and γ_j counted with multiplicity and ordered in the sense that if we start walking on β , then x_i is the i^{th} time β intersects γ_j . Suppose that $q \geq 2$ (the case $q = 1$ will be treated in the end of the proof). Let $\beta_{i,i+1}$ be the simple loop based at x_i defined as the concatenation of the oriented arc of β between x_i and x_{i+1} and the oriented arc $c_{i+1,i}$ of γ_j between x_{i+1} and x_i . The loop β is homotopic to the loop $\beta_{1,2}c_{1,2} \cdots \beta_{q,q+1}c_{q,q+1}$, where by convention $c_{i,i+1}$ is the inverse of $c_{i+1,i}$, and $x_{q+1} = x_1$.

We claim that one of the loops $\beta_{k,k+1}$ does not commute with γ_j . Indeed, suppose the opposite. Since γ_j is a minimizing simple closed geodesic, every loop commuting with γ_j is homotopic to a power of it. The loop $\beta_{1,2}c_{1,2} \cdots \beta_{q,q+1}c_{q,q+1}$ is homotopic to a power of γ_j since for all i the loop $\beta_{i,i+1}$ commutes with γ_j . Thus $\beta_{1,2}c_{1,2} \cdots \beta_{q,q+1}c_{q,q+1}$ commutes with γ_j . So the loop β commutes with γ_j since it is homotopic to $\beta_{1,2}c_{1,2} \cdots \beta_{q,q+1}c_{q,q+1}$. That is a contradiction.

Recall that the surface M can be obtained from the surface M_p by cutting along the γ_i and reinserting the cylinders C_i . Thus, the loop in M that corresponds to β decomposes into a union of curves whose endpoints lie on one of the two boundary components γ_j^- and γ_j^+ of the cylinder C_j . Denote by x'_k and x'_{k+1} the points in M corresponding to the points x_k and x_{k+1} of $\beta_{k,k+1}$ in M_p . We have two cases.

Case 1 The points x'_k and x'_{k+1} lie both the same boundary component, say γ_j^+ . In this case, let β' be the simple loop in M that corresponds to $\beta_{k,k+1}$; see Figure 3.

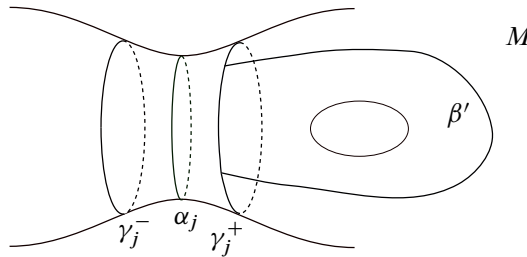


Figure 3

Take $a = \gamma_j^+$ and $b = \beta'$. These two loops are based at the same point and do not commute. Moreover, we have $\text{length}(a) = 1$ and $\text{length}(b) \leq L + 1$.

Case 2 The points x'_k and x'_{k+1} do not lie both on γ_j^- or γ_j^+ . In this case, let β' be the arc in M that corresponds to the arc of β between x_k and x_{k+1} .

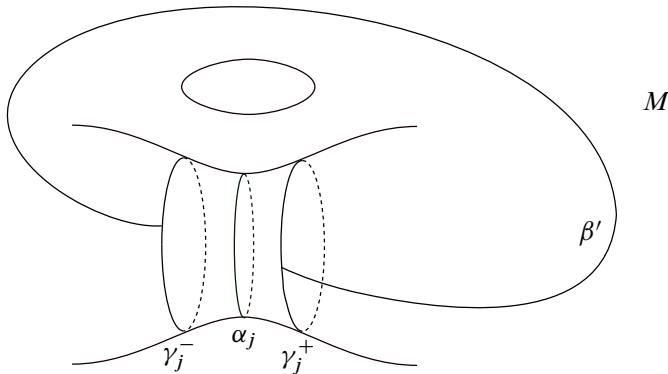


Figure 4

Take $a = \gamma_j^+$ and $b = \beta' \gamma_j^+ \beta'^{-1}$. These two loops are based at the same point and do not commute. Moreover we have $\text{length}(a) = 1$ and $\text{length}(b) \leq 2L + 1$.

Finally, if the number of intersections $q = 1$, we argue exactly like in case 2 above, supposing that $x_{k+1} = x_k$. That finishes the proof of the Lemma. \square

Now, apply Lemma 6.5 with $\beta = \mu$ and make use of Lemma 6.1 to finish the proof.

Step 5 By Theorem 4.4, there are at least $n = \lceil \log(2g) + 1 \rceil$ homotopically independent geodesic loops $\mu_1 \dots, \mu_n$ based at the same point in M_p with

$$\text{length}(\mu_i) \leq 2^{18} \log(g).$$

If these loops are in the original surface M , ie, they do not transversally intersect any of the loops γ_i in M_p , then the proof is finished. So suppose the opposite. Let μ be one the loops $\mu_1 \dots, \mu_n$ that transversally intersects at least one of the γ_i in M_p . From (6-1), the loop μ (transversally) intersects exactly one loop γ_j in M_p . By Lemma 6.5, we show that there exist two loops a, b in the original surface M based at the same point with $\text{length}(a) = 1$ and $\text{length}(b) \leq 2^{19} \log(g) + 1$. The result follows from Lemma 6.1. \square

Remark 6.6 Theorem A extends to nonorientable surfaces with multiplicative constant 2^{22} instead of 2^{20} by passing to the double oriented cover.

Corollary 6.7 *There exists a positive constant C such that the separating systole of every closed Riemannian surface M of genus $g \geq 2$ and area g satisfies*

$$\text{sys}_0(M) \leq C \log(g).$$

Proof From Theorem A, there exist two noncommuting loops a and b based at the same point of length at most $c \log(g)$ for some positive constant c . The commutator $[a, b]$ of a and b , of length at most $4c \log(g)$, yields a bound on the separating systole of M . \square

References

- [1] **F Balacheff, H Parlier, S Sabourau**, *Short loop decompositions of surfaces and the geometry of Jacobians*, *Geom. Funct. Anal.* 22 (2012) 37–73 MR2899682
- [2] **B Bollobás, E Szemerédi**, *Girth of sparse graphs*, *J. Graph Theory* 39 (2002) 194–200 MR1883596
- [3] **B Bollobás, A Thomason**, *On the girth of Hamiltonian weakly pancyclic graphs*, *J. Graph Theory* 26 (1997) 165–173 MR1475897
- [4] **P Buser, P Sarnak**, *On the period matrix of a Riemann surface of large genus*, *Invent. Math.* 117 (1994) 27–56 MR1269424
- [5] **M P do Carmo**, *Riemannian geometry*, Birkhäuser, Boston, MA (1992) MR1138207
- [6] **R Diestel**, *Graph theory*, 4th edition, Graduate Texts in Mathematics 173, Springer, Heidelberg (2010) MR2744811
- [7] **M Gromov**, *Filling Riemannian manifolds*, *J. Differential Geom.* 18 (1983) 1–147 MR697984
- [8] **M Gromov**, *Systoles and intersystolic inequalities*, from: “Actes de la Table Ronde de Géométrie Différentielle”, (A L Besse, editor), Sémin. Congr. 1, Soc. Math. France, Paris (1996) 291–362 MR1427763

- [9] **W Jaco**, *On certain subgroups of the fundamental group of a closed surface*, Proc. Cambridge Philos. Soc. 67 (1970) 17–18 MR0258030
- [10] **S Karam**, *Uniform growth of balls in the universal cover of graphs and surfaces*, to be published in Trans. Amer. Math. Soc.
- [11] **S B Myers**, *Connections between differential geometry and topology II: Closed surfaces*, Duke Math. J. 2 (1936) 95–102 MR1545908
- [12] **S Sabourau**, *Asymptotic bounds for separating systoles on surfaces*, Comment. Math. Helv. 83 (2008) 35–54 MR2365407

*Laboratoire de Mathématiques et de Physique Théorique, Université de Tours
UFR Sciences et Technologie, Parc de Grandmont, 37200 Tours, France*

`steve.karam@lmpt.univ-tours.fr`

Received: 21 October 2013