ENTROPY AND SYSTOLES ON SURFACES

STÉPHANE SABOURAU

Abstract. We show that the volume entropy of surfaces with unit systole is bounded from above by a constant which does not depend on the metric, which answers a question raised by A. Katok. This upper bound follows from a comparison between the geometric and algebraic lengths on the fundamental groups of surfaces.

1. Introduction

Let $(\tilde{M}, \tilde{g})$ be the universal Riemannian covering of a closed n-dimensional Riemannian manifold (M, g). Fix $x_0 \in M$ and a lift $\tilde{x}_0 \in \tilde{M}$ of x_0.

The volume entropy (or asymptotic volume) of (M, g) is defined as

$$h_{vol}(M,g) = \lim_{R \to +\infty} \frac{\log(\mathrm{Vol}_{\tilde{g}} B(\tilde{x}_0, R))}{R} \tag{1.1}$$

where $\mathrm{Vol}_{\tilde{g}} B(\tilde{x}_0, R)$ is the volume of the ball centered at $\tilde{x}_0$ with radius R in $\tilde{M}$. Since M is compact, the limit in (1.1) exists and does not depend on the point $\tilde{x}_0 \in \tilde{M}$ (see [13]). This asymptotic invariant describes the exponential growth rate of the volume on the universal covering. It is related to the geometry, the dynamic and the topology of manifolds.

The volume entropy can be normalized by the volume. This amounts to defining the normalized volume entropy $h_{vol}(M,g) \cdot \mathrm{Vol}(M,g)^{\frac{1}{n}}$, which is invariant under scaling. The infimum of the normalized volume entropy over the space of all metrics has been studied by A. Katok in [10], M. Gromov in [6], I. Babenko in [1] and G. Besson, G. Courtois and S. Gallot in [2]. These authors specifically found universal lower bounds (i.e., bounds which do not depend on the metric) on the normalized volume entropy. In particular, the following result holds.

Theorem 1.1 ([10] and [2]). *Let (M, g_0) be a closed negatively curved locally symmetric n-manifold. Every metric g on M satisfies*

$$Vol(M,g) \cdot h_{vol}(M,g)^n \geq Vol(M,g_0) \cdot h_{vol}(M,g_0)^n \tag{1.2}$$

with equality if and only if g is locally symmetric.

2000 *Mathematics Subject Classification.* 53C20.

Key words and phrases. systole, entropy, surface, geometric length, algebraic length, exponential growth.

Current address: Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6395, USA.

E-mail: sabourau@lmpt.univ-tours.fr.

The volume entropy can also be normalized by the diameter. In this case, every closed Riemannian manifold (M, g) satisfies

$$\mathrm{Diam}(M, g) \cdot h_{vol}(M, g) \geq \frac{1}{2} h_{alg}(\Gamma), \tag{1.3}$$

where $h_{alg}(\Gamma)$ is the minimal algebraic entropy of the fundamental group Γ of M (see [18], [16] and also [8, 5.16]). We refer to Section 2 for a definition of $h_{alg}(\Gamma)$. Note that the minimal algebraic entropy of the closed surfaces of genus h is bounded from below by $\log(4h - 3)$ (see [9, p. 195]).

In this article, we are interested in another normalization of the volume entropy. Namely, we are interested in the scale invariant product $\mathrm{sys}(M, g) \cdot h_{vol}(M, g)$ where $\mathrm{sys}(M, g)$ is the systole of a nonsimply connected closed Riemannian manifold (M, g). Recall that the systole of (M, g) is defined as the length of the shortest noncontractible loop of M. Such a shortest noncontractible loop is called a systolic loop of M.

While the previous normalizations by the volume and the diameter yield lower bounds on the volume entropy, the normalization by the systole leads to upper bounds on the volume entropy. Namely, we show that, for surfaces, the volume entropy normalized by the systole is bounded from above.

Theorem 1.2. *Let M be a nonsimply connected closed surface. Then, there exists a positive constant C such that every metric g on M satisfies*

$$sys(g) \cdot h_{vol}(g) \leq C \tag{1.4}$$

The constant C we obtain in inequality (1.4) is explicit but nonsharp. Note that the optimal constant C necessarily depends on the topology of the surface M as it must tend to infinity with the genus. Indeed, there exist hyperbolic surfaces of large genus with large systole (see [3, Sect. 4]) while the volume entropy of every hyperbolic surface equals 1. These surfaces are obtained as congruence coverings of an arithmetic Riemann surface.

Metrics which maximize $\mathrm{sys}(M, g) \cdot h_{vol}(M, g)$, if they exist, seem rather difficult to describe.

In general, the universal inequality (1.4) does not hold in higher dimension, see Section 3 for counterexamples.

The topological entropy of the geodesic flow, noted $h_{top}(M, g)$, is related to the volume entropy. More precisely, this dynamical invariant agrees with the volume entropy when the metric has no conjugate points (see [4] and [13]). In general, we have only $h_{top}(M, g) \geq h_{vol}(M, g)$. Thus, the statement of Theorem 1.1 and the inequality (1.3) still hold if we replace the volume entropy by the topological entropy. However, such a substitution in the statement of Theorem 1.2 is not valid. Indeed, arbitrary perturbations of the metric in a sufficiently small neighborhood of a point through which no systolic loop passes do not change the value of the systole. Thus, it is possible to make the topological entropy arbitrarily large (see [14]) while keeping the systole fixed.

Theorem 1.2 answers a question raised by A. Katok in [10, p. 358] where the following result with an extra assumption on the Gaussian curvature has been established.

Theorem 1.3 ([10]). *Let M be a closed surface. Then, for every real number K_0, there exists $C > 0$ such that every metric g on M with Gaussian curvature bounded from below by K_0 satisfies*

$$inj(M, g) \cdot h_{vol}(M, g) \leq C \tag{1.5}$$

where $inj(M, g)$ is the injectivity radius of (M, g).

Here the constant C depends on the lower bound K_0 on the curvature.

In order to state our next result, we need to introduce the algebraic and geometric lengths on the fundamental groups of closed Riemannian manifolds.

Let Σ be a finite generating set of a group Γ. The algebraic (or word) length of an element $\alpha \in \Gamma$ with respect to Σ is noted $|\alpha|_\Sigma$. It is defined as the smallest integer $k \geq 0$ such that $\alpha = \alpha_1 \dots \alpha_k$ where $\alpha_i \in \Sigma \cup \Sigma^{-1}$. By definition, the neutral element e is the only element of Γ with null algebraic length. Every element $\alpha \in \Gamma$ can be written in a reduced form, i.e., as a product $\alpha = \alpha_1 \dots \alpha_k$ such that $\alpha_i \in \Sigma \cup \Sigma^{-1}$ with $k = |\alpha|_\Sigma$.

Fix a point x_0 on a closed Riemannian manifold (M, g). The geometric length on $\Gamma := \pi_1(M, x_0)$ of an element $\alpha \in \Gamma$, denoted $|\alpha|_g$, is defined as the minimal length of a loop based at x_0 representing α.

The proof of Theorem 1.2 rests on the following comparison between the algebraic and geometric lengths on the fundamental groups of surfaces.

Theorem 1.4. *Let M be a nonsimply connected closed surface of Euler characteristic $\chi(M)$ and $x_0 \in M$. Then, there exists $\lambda > 0$ such that every metric g on M with $sys(g) \geq 1$ satisfies*

$$|.|_g \geq \lambda |.|_\Sigma \tag{1.6}$$

for some generating set Σ of $\pi_1(M, x_0)$ with at most $2|\chi(M)| + 3$ elements ($|\chi(M)| + 2$ in the orientable case).

As in the case of Theorem 1.2, this result does not hold in general for higher dimensional manifolds, see Section 3.

The multiplicative constants we obtain in our inequalities, even though explicit, are far from being optimal. Though many of them can easily be improved to some extent. But since the nature of our arguments leaves little hope of obtaining good constants, we will not examine that here. Instead, we will decompose our arguments into simple independent results.

Other similar bounds on the entropy have been established in [11] for systolically extremal surfaces (see [17] for a generalization in higher dimension).

This article is organized as follows. In Section 2, we recall some properties of the algebraic and geometric lengths on the fundamental groups of closed Riemannian manifolds. In particular, we observe that Theorem 1.4 implies Theorem 1.2. In Section 3, we present counterexamples to Theorem 1.2 and Theorem 1.4 in higher dimension. In Section 4, we show that cutting off the long spikes and the long tubes of surfaces does not change the value of the systole and possibly increases the volume entropy (without loss of generality, we only consider metrics with unit systole). This permits us to consider only metrics of "small" diameter without long spikes and long tubes. We show then that such metrics admit "standard" systems of loops with bounded lengths. In Section 5, we prove Theorem 1.4 first for punctured surfaces, then for arbitrary closed surfaces.

Roughly speaking, the proof proceeds as follows. The restriction of a standard system of loops in a surface M (found in Section 4) to the surface N obtained by removing a small disk centered at their based point is formed of a collection Γ of disjoint minimizing arcs. Further, this collection of arcs gives rise to generating sets Σ and S in the fundamental groups of M and N. Given a minimizing loop γ of M, the minimal writing of $[\gamma]$ in $\pi_1(N)$ with respect to S is related to the intersection of γ with the arcs of Γ. One can be read off through the other. If the minimal writing of γ is long enough, there is a subarc of γ corresponding to a subword of the minimal writing of γ with bounded length which intersects an arc of Γ a large number of times. We show that the length of this subarc is bounded from below and remove it from γ using a cut-and-paste argument. An induction argument permits us to obtain a lower bound on the length of γ in terms of the S-algebraic length of $[\gamma]$. A comparison of the S- and Σ-algebraic lengths yields the result. The nonorientable case is treated at the very end.

The author would like to thank the referee for suggestions that helped improve the presentation of this article.

2. Preliminaries

In this section, we introduce the algebraic and geometric entropies of the algebraic and geometric lengths of the fundamental groups of closed Riemannian manifolds. We also recall how these entropies are related. For the sake of completeness, we present the proofs of these classical results. Finally, we observe that Theorem 1.2 is an immediate consequence of Theorem 1.4.

Let Σ be a finite generating set of a group Γ and $|.|_\Sigma$ be the algebraic length on Γ with respect to Σ.

The algebraic entropy of Γ with respect to Σ is defined by

$$h_{alg}(\Gamma, \Sigma) = \limsup_{R \to +\infty} \frac{\log(N_\Sigma(R))}{R}, \tag{2.1}$$

where $N_\Sigma(R) = \mathrm{card}\{\alpha \in \Gamma \mid |\alpha|_\Sigma \leq R\}$ is the cardinality of the R-ball of $(\Gamma, |.|_\Sigma)$ centered at its origin. If Γ is the fundamental group of a closed manifold, then the sup-limit is a limit (see [13]).

The minimal algebraic entropy of Γ is defined as

$$h_{alg}(\Gamma) = \inf_\Sigma h_{alg}(\Gamma, \Sigma), \tag{2.2}$$

where Σ runs over all the finite generating sets of Γ.

The following well-known result provides an upper bound on the algebraic entropy.

Lemma 2.1. *Let Γ be a group generated by a finite set Σ of cardinality $|\Sigma|$. Then,*

$$h_{alg}(\Gamma, \Sigma) \leq \log(2|\Sigma| - 1). \tag{2.3}$$

Proof. We can assume that $e \notin \Sigma$. Let $k = |\Sigma|$. To construct an element of word length n with respect to Σ, we must choose the first letter of its reduced form among the $2k$ letters of $\Sigma \cup \Sigma^{-1}$. Then, for each remaining letter, we have at most $2k-1$ possible choices since that letter cannot be the inverse of the previous letter. Thus, $N_\Sigma(n) \leq 1 + \sum_{p=0}^{n-1} 2k(2k-1)^p = 1 + 2k\frac{(2k-1)^n - 1}{2k-2}$. Therefore, $h_{alg}(\Gamma, \Sigma) \leq \log(2k-1)$. □

Let (M, g) be a closed Riemannian manifold. Fix x_0 in M and a lift $\tilde{x}_0$ of x_0 in the universal covering $\tilde{M}$. The group $\Gamma := \pi_1(M, x_0)$ acts on $\tilde{M}$ by isometries. By definition of the geometric length, for every $\alpha \in \Gamma$, we have

$$|\alpha|_g = d_{\tilde{g}}(\alpha.\tilde{x}_0, \tilde{x}_0). \tag{2.4}$$

Recall that the volume entropy of (M, g) may be obtained from the geometric length $|.|_g$ on Γ as follows

Lemma 2.2. *Let (M, g) be a closed Riemannian manifold. Then,*

$$h_{vol}(g) = \lim_{R \to +\infty} \frac{\log(N_g(R))}{R} \tag{2.5}$$

where $N_g(R) = card\{\alpha \in \Gamma \mid |\alpha|_g \leq R\}$ is the cardinality of the R-ball of $(\Gamma, |.|_g)$ centered at the origin.

Proof. The orbit of $\tilde{x}_0$ under the group Γ is noted $\Gamma.\tilde{x}_0$. Consider a fundamental domain Δ for the action of Γ containing $\tilde{x}_0$. Denote by D the diameter of Δ. The cardinality of $\Gamma.\tilde{x}_0 \cap B(\tilde{x}_0, R)$ agrees with $N_g(R)$. Further, it is bounded from above by the number of translated fundamental domains $\alpha.\Delta$ contained in $B(\tilde{x}_0, R+D)$ and bounded from below by the number of translated fundamental domains $\alpha.\Delta$ contained in $B(\tilde{x}_0, R)$. Hence,

$$\frac{\mathrm{vol}(B(\tilde{x}_0, R))}{\mathrm{vol}(M, g)} \leq N_g(R) \leq \frac{\mathrm{vol}(B(\tilde{x}_0, R + D))}{\mathrm{vol}(M, g)}. \tag{2.6}$$

Take the log of these terms, multiply them by $\frac{1}{R}$ and let R go to infinity. The left-hand term and the right-hand term both yield $h(g)$. The result follows. $\square$

The volume entropy of a Riemannian manifold is related to the algebraic entropy of its fundamental group through the following classical result.

Lemma 2.3. *Let (M, g) be a closed Riemannian manifold and Σ be a finite generating set of $\Gamma = \pi_1(M, x_0)$. If $\lambda|.|_\Sigma \leq |.|_g \leq \mu|.|_\Sigma$ for some constants $\lambda, \mu > 0$, then*

$$\frac{1}{\mu} h_{alg}(\Gamma, \Sigma) \leq h_{vol}(g) \leq \frac{1}{\lambda} h_{alg}(\Gamma, \Sigma). \tag{2.7}$$

Proof. Since $\lambda|.|_\Sigma \leq |.|_g \leq \mu|.|_\Sigma$, we have $N_\Sigma(R/\mu) \subset N_g(R) \subset N_\Sigma(R/\lambda)$. In particular, $\frac{1}{\mu}\frac{1}{R/\mu}\log(N_\Sigma(R/\mu)) \leq \frac{1}{R}\log(N_g(R)) \leq \frac{1}{\lambda}\frac{1}{R/\lambda}\log(N_\Sigma(R/\lambda))$. Thus, as R goes to infinity, we get $\frac{1}{\mu} h_{alg}(\Gamma, \Sigma) \leq h_{vol}(g) \leq \frac{1}{\lambda} h_{alg}(\Gamma, \Sigma)$. $\square$

Note that if Σ is a finite generating set of Σ, then the ball $B_\Sigma(k)$ of radius k centered at the origin of $(\Gamma, |.|_\Sigma)$ is also a finite generating set.

Lemma 2.4. *Let Γ be a group generated by a finite set Σ. Then, for every integer k,*

$$h_{alg}(\Gamma, B_\Sigma(k)) \geq \frac{k}{2} h_{alg}(\Gamma, \Sigma).$$

Proof. Let $\alpha \in \Gamma$. The element α can be decomposed as a product $\alpha = \alpha_1 \cdots \alpha_n$, where $\alpha_i \in \Sigma$ and $n = |\alpha|_\Sigma$. This product can also be written as $\alpha = (\alpha_1 \cdots \alpha_k) \cdot (\alpha_{k+1} \cdots \alpha_{2k}) \cdots (\cdots \alpha_n)$, where each factor $(\alpha_{ik+1} \cdots \alpha_{(i+1)k})$ is the product of k elements except possibly for the last one $(\cdots \alpha_n)$, which is the product of k elements or less. Thus, $|\alpha|_{B_\Sigma(k)} \leq [\frac{n}{k}] + 1$. This leads to $|.|_{B_\Sigma(k)} \leq \frac{2}{k}|.|_\Sigma$. Using the same arguments as in the proof of Lemma 2.3, we conclude that $\frac{k}{2} h_{alg}(\Gamma, \Sigma) \leq h_{alg}(\Gamma, B_\Sigma(k))$. $\square$

Remark 2.5. Using better asymptotic estimates, we can show that the inequality in Lemma 2.4 can be improved by a factor 2.

Remark 2.6. Since the product $\mathrm{sys}(g) \cdot h_{vol}(g)$ is scale invariant, we can assume that $\mathrm{sys}(g) = 1$. Thus, Theorem 1.2 may immediately be deduced from Theorem 1.4, Lemma 2.1 and Lemma 2.3.

3. Counterexamples in higher dimension

In this section, we show that Theorem 1.2 (and therefore Theorem 1.4) does not hold in higher dimensions in general.

First notice that the fundamental group of a closed manifold M does not have an exponential growth (that is, its algebraic entropy with respect to a finite generating set is zero) if and only if its volume entropy $h(g)$ is zero for every metric g on M (see [16] and [18]). In this case, the inequality (1.4) trivially holds.

Otherwise, we have the following result.

Proposition 3.1. *Let M be a closed manifold of dimension $n \geq 3$ whose fundamental group Γ has exponential growth. Then, the product $sys(g) \cdot h_{vol}(g)$ is unbounded when g runs over all the metrics on M.*

Proof. Fix $x_0 \in M$. Consider a finite generating set Σ of $\pi_1(M, x_0)$. Since $n \geq 3$, the ball $B_\Sigma(k)$ of radius k in $(\Gamma, |.|_\Sigma)$ is induced by a collection of simple loops γ_i that intersect one another only at their basepoint x_0. Take a metric g on M and modify it in the tubular neighborhood of these simple loops so that, after normalization, the systole of the new metric g_k equals 1 and the length of the loops γ_i is between 1 and 2. Clearly, we have $|.|_{g_k} \leq 2|.|_{B_\Sigma(k)}$. By the lemmas 2.3 and 2.4, we deduce that

$$h_{vol}(g_k) \geq \frac{1}{2} h_{alg}(\Gamma, B_\Sigma(k)) \geq \frac{k}{4} h_{alg}(\Gamma, \Sigma).$$

Since $h_{alg}(\Gamma, \Sigma)$ is positive by assumption, we obtain the desired result by letting k go to infinity. $\square$

4. Change of the metric

The goal of this section is to prove Propositions 4.1 and 4.6 below. These propositions will permit us to consider in the remainder of this article only metrics of "small" diameter with "short" systems of loops. These features will turn out to be useful to prove Theorem 1.4 in Section 5.

Let us state the first result of this section.

Proposition 4.1. *Let M be a closed surface of genus h and $x_0 \in M$. There exists a homotopy equivalence $p : M \longrightarrow \overline{M}$, where $\overline{M}$ is a surface homeomorphic to M, and a metric $\overline{g}$ on $\overline{M}$ such that*

i) $sys(\overline{g}) \geq 1$;
ii) $Diam(\overline{g}) \leq 32h$;
iii) $|p_*(a)|_{\overline{g}} \leq 3|a|_g$ *for every* $a \in \pi_1(M, x_0)$.

Remark 4.2. Since the homeomorphism p_* induced by p between the fundamental groups is an isomorphism, the preimage by p_* of every generating set $\overline{\Sigma}$ of $\pi_1(\overline{M})$ is a generating set of $\pi_1(M)$. Furthermore, $|p_*(a)|_\Sigma = |a|_{p_*^{-1}(\overline{\Sigma})}$ for every $a \in \pi_1(M)$. Thus, if we can prove Theorem 1.4 for every metric $\overline{g}$ with $\mathrm{Diam}(\overline{g}) \leq 32h$, namely $|.|_{\overline{g}} \geq C|.|_{\overline{\Sigma}}$ for some generating set $\overline{\Sigma}$ with $2|\chi(M)| + 3$ elements, we would immediately obtain $|.|_g \geq \frac{C}{3}|.|_\Sigma$ with $\Sigma = p_*^{-1}(\overline{\Sigma})$. Therefore, in order to prove Theorem 1.4, we can consider only metrics with diameter less than $32h$.

Before proving Proposition 4.1, let us introduce some notations and definitions, and state some facts.

Let g be a Riemannian metric on M such that $\mathrm{sys}(M,g) = 1$. Fix $x_0 \in M$. We can approximate (in the sup-norm topology) the distance function $d_g(x_0, .)$ by a Morse function f with only one critical point on each of its critical levels and with x_0 as its only local minimum. Such an approximation has also been used in [12].

Definition 4.3. *A cylinder of M is said to be admissible if it is maximal (for the inclusion) among all open cylinders C such that*

- *the boundary components c_1 and c_2 of C lie in some critical levels of f, denoted $f^{-1}(\lambda_1)$ and $f^{-1}(\lambda_2)$, with $\lambda_1 < \lambda_2$;*
- *some neighborhood of c_1 in C lies in $f^{-1}(]\lambda_1, +\infty[)$;*
- *some neighborhood of c_2 in C lies in $f^{-1}(]-\infty, \lambda_2[)$;*
- *the loops c_1 and c_2 are noncontractible.*

Remark that an admissible cylinder with boundaries c_1 and c_2 lying in $f^{-1}(\lambda_1)$ and $f^{-1}(\lambda_2)$ is not necessarily included in $f^{-1}(]\lambda_1, \lambda_2[)$ (for instance, a "long finger" of the cylinder can go above the λ_2-level).

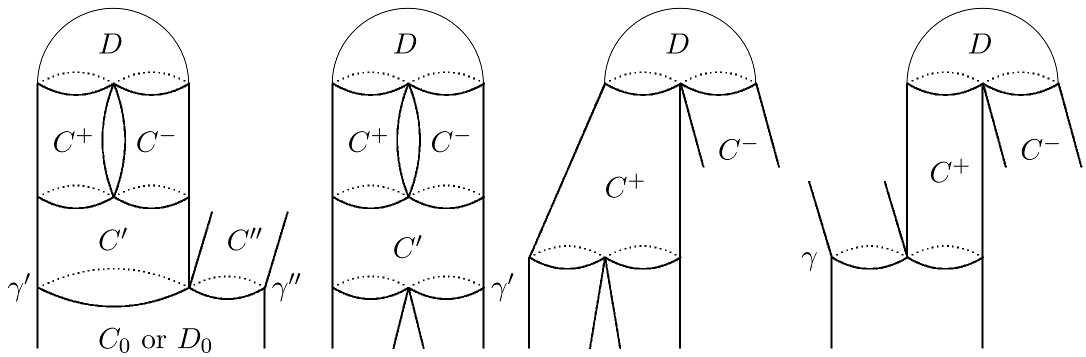
Denote by $\mathcal{C}$ the collection of (disjoint) admissible cylinders. Exactly one critical point of f lies in each boundary component of the cylinders of $\mathcal{C}$. Therefore, the boundary components of the cylinders of $\mathcal{C}$ are either simple loops or figure-eight loops (recall that the cylinders of $\mathcal{C}$ are open). Since the surface M is orientable, the connected components of $M \setminus \cup_{C \in \mathcal{C}} \overline{C}$ are open topological disks. They are called *admissible disks* and form a collection $\mathcal{D}$. The boundary of a disk of $\mathcal{D}$ is a figure-eight curve which lies in some critical level of f and decomposes into two simple loops along which two cylinders of $\mathcal{C}$ are glued.

Lemma 4.4. *The collection $\mathcal{C}$ is composed of at most $4h-2$ cylinders. The collection $\mathcal{D}$ is composed of at most $h+1$ disks.*

Proof. Let us argue by induction on h.

If M is a torus, then M is composed of two admissible cylinders and two admissible disks. Recall that x_0 is the only local minimum of f.

Assume now that M is a surface of genus $h \geq 2$. Take $D \in \mathcal{D}$ which does not contain x_0. Two cylinders $C^+, C^- \in \mathcal{C}$ with boundary components $c_1^{\pm}, c_2^{\pm}$ are glued "below" along the boundary ∂D of D, i.e., $\partial D = c_2^+ \cup c_2^-$.



Figures 1-4

If c_1^+ and c_1^- form a figure-eight curve (see Figures 1 and 2), then there exists a cylinder $C' \in \mathcal{C}$ glued "below" along $c_1^+ \cup c_1^-$. Denote by γ' the other boundary component of C'. Two cases may occur: either γ' is simple or γ' is a figure-eight loop. In the former case (Figure 1), there exists a loop γ'' lying in some critical level of f and forming with γ' a figure-eight curve such that a cylinder C'' of $\mathcal{C}$ is glued "above" along γ'' and a cylinder C_0 of $\mathcal{C}$ or a disk D_0 of $\mathcal{D}$ is glued "below" along $\gamma' \cup \gamma''$. In the latter case (Figure 2), there exist two cylinders of $\mathcal{C}$ glued "below" along each simple loop of the figure-eight loop γ' (recall that x_0 is the only local minimum of f). In both cases, replacing $C' \cup C^+ \cup C^- \cup D$ by a hemisphere cap H glued "above" along γ' (which is not a torsion element of $H_1(M, \mathbb{Z})$ since M is orientable) gives rise to a surface N of genus $h-1$. By construction, if $\mathcal{C}_N$ is the collection of admissible cylinders of N, we have $\mathrm{card}(\mathcal{C}_N) \geq \mathrm{card}(\mathcal{C}) - 4$ (in the first case, $C_0 \cup H \cup C''$ is an admissible cylinder of N or $D_0 \cup H \cup C''$ is an admissible disk of N). Similarly, N has fewer admissible disks than M. Therefore, $\mathrm{card}(\mathcal{C}) \leq 2 + 4(h-1)$ and $\mathrm{card}(\mathcal{D}) \leq h-1$ by induction.

Otherwise, c_1^+ and c_1^- are disjoint (see Figures 3 and 4). As previously, two cases may occur: either c_1^+ is a figure-eight loop or c_1^+ is simple. In the former case (Figure 3), two cylinders of $\mathcal{C}$ distinct from $C^\pm$ are glued "below" along c_1^+. In the latter case (Figure 4), c_1^+ forms with a loop γ a figure-eight curve such that two cylinders of $\mathcal{C}$ are glued along γ and $c_1^+ \cup \gamma$ (recall that x_0 is the only local minimum of f). The same goes for c_1^-. Replacing $C^+ \cup C^- \cup D$ by two hemisphere caps glued "above" along c_1^+ and c_1^-, which are not torsion elements, gives rise to a closed orientable surface N. Since x_0 is the only local minimum of f, the surface N is connected of genus $h-1$. By construction, $\mathrm{card}(\mathcal{C}_N) \geq \mathrm{card}(\mathcal{C}) - 4$ and N has fewer admissible disks than M. The result follows as previously by induction. $\square$

The proof of Proposition 4.1 consists in four steps. In the first two steps, we "chop off the spikes and remove the long tubes" of the surface. Then, in the last two steps, we apply combinatorial and counting arguments on the new surface so as to evaluate the lengths of curves.

Proof of Proposition 4.1. Without loss of generality, we can assume that $\mathrm{sys}(g) = 1$. We will also assume that f agrees with $d_g(., x_0)$. The general case requires only minor modifications of the proof. Given an admissible cylinder C, denote by $\mathcal{D}_C$ the collection of maximal (for the inclusion) disks in C whose boundary lies in some (critical) level of f.

Step 1: Lower the spikes of M.

Take a disk D in $\mathcal{D}$ or in $\mathcal{D}_C$ for some $C \in \mathcal{C}$. Consider the metric $g_t = e^{-t\varphi(x)} g$ where $\varphi(x) = d_g(U_{1/2}(M \setminus D), x)$ and $U_{1/2}(M \setminus D) = \{x \in M \mid d_g(M \setminus D, x) < \frac{1}{2}\}$. We clearly have $g_t \leq g$ and $\mathrm{sys}(g_t) = 1$ for all $t \geq 0$. Furthermore, for t large enough, we have $d_{g_t}(\partial D, x) < 1$ for every $x \in D$. Note also that the collections $\mathcal{C}, \mathcal{D}$ and $\mathcal{D}_C$ defined with respect to g agree with the same collections defined with respect to g_t. Therefore, changing g on each disk of $\mathcal{D}$ or of $\mathcal{D}_C$ for every $C \in \mathcal{C}$ if necessary, we can assume that $d_g(\partial D, x) < 1$ for every $C \in \mathcal{C}$, $D \in \mathcal{D}_C$ and $x \in D$.

Step 2: Shorten the admissible cylinders.

Let C be an admissible cylinder whose boundary components c_1 and c_2 lie in $f^{-1}(\lambda_1)$ and $f^{-1}(\lambda_2)$ with $\lambda_2 > \lambda_1 + 4 + 2\varepsilon$ for some positive $\varepsilon < \frac{1}{2}$.

Let $U_i = \{x \in C \mid \frac{1}{2} < d_g(c_i, x) < \frac{1}{2} + \varepsilon\}$ for $i = 1, 2$. We can modify the metric g on U_1 and U_2 into a metric g' conformal to g with $g' \leq g$ and $\mathrm{sys}(g') = 1$ such that some g'-systolic loops γ_1 and γ_2 pass through U_1 and U_2. The g'-systolic loop γ_i lie in $V_i = \{x \in C \mid d_g(c_i, x) < 1 + \varepsilon\}$ for $i = 1, 2$. Therefore, since $d_g(c_1, c_2) > 4 + 2\varepsilon$, we have $d_{g'}(\gamma_1, \gamma_2) > 2$. Furthermore, the g'-systolic loops γ_1 and γ_2 bound a cylinder C' in C. Let x_i be the endpoint on γ_i of a g'-minimizing segment of C' joining γ_1 to γ_2 for $i = 1, 2$. Shrinking the cylinder C' of C along its height without a twist so that the boundaries γ_1 and γ_2 of $M \setminus C'$ and the points x_1 and x_2 agree gives rise to a homotopy equivalence from M to a surface homeomorphic to $(M \setminus C')/\gamma_1 \sim \gamma_2$. Further, the metric g' on M induces a metric on the target space.

Performing this operation on all admissible cylinders of (M, g) with $\lambda_2 > \lambda_1 + 4$ yields a homotopy equivalence $p : M \longrightarrow \overline{M}$ (by shrinking all the cylinders C' for all the admissible cylinder C) and a metric $\overline{g}$ on $\overline{M}$. By construction, we have $\mathrm{sys}(\overline{g}) = 1$.

Step 3: Bound from above the $\overline{g}$-diameter.

Given $\overline{x} \in \overline{M}$, let c be a g-minimizing segment of M joining x_0 to x, where $x \in p^{-1}(\overline{x})$. The g-geodesic c starts at x_0, exits the disk D of $\mathcal{D}$ containing x_0, then goes through some cylinders of $\mathcal{C}$ and finally ends at x by possibly going into some disk D of $\mathcal{D}$. Furthermore, the g-minimizing geodesic c goes through each cylinder of $\mathcal{C}$ at most once. In the general case, that is when f is a small perturbation of $d_g(x_0, .)$, the geodesic c might go through the boundary of each cylinder C more than once, but this is unessential, since it would occur along a short segment of c.

Consider an admissible cylinder C of M not containing x whose boundary components c_i lie in $f^{-1}(\lambda_i)$ with $i = 1, 2$. The arc $\gamma = c \cap C$ joins c_1 to c_2 and the g-length $L_g(\gamma)$ of γ satisfies $L_g(\gamma) = \lambda_2 - \lambda_1$. In the general case, these two quantities might not be equal, but they are close. If $L_g(\gamma) \leq 4$, then $L_{\overline{g}}(p(\gamma)) = L_g(\gamma) \leq 4$. Otherwise, $\lambda_2 > \lambda_1 + 4$. In this case, the arc α_i of γ lying in $C \setminus C'$ and joining c_i to γ_i (see Step 2 for the notations) satisfies $L_{\overline{g}}(p(\alpha_i)) \leq L_{g'}(\alpha_i) \leq 1 + \varepsilon$. Furthermore, the $\overline{g}$-distance between the endpoints of $p(\alpha_i)$ on $p(\gamma_1) = p(\gamma_2)$ for $i = 1, 2$ is at most $\frac{1}{2}$. Thus, the $\overline{g}$-distance between the endpoints of $p(\gamma)$ is at most $2(1 + \varepsilon) + \frac{1}{2} < 4$.

Consider now the admissible cylinder C containing x (it might not exist) and the arc $\gamma = c \cap C$. If $\lambda_2 \leq \lambda_1 + 4$, then $L_{\overline{g}}(p(\gamma)) = L_g(\gamma) \leq 4 + \frac{1}{2} + \varepsilon_0 = 5$ since γ goes into at most one disk of $\mathcal{D}_C$. Otherwise, we show as previously that the $\overline{g}$-distance between the endpoints of $p(\gamma)$ is at most $2(1 + \varepsilon) + \frac{1}{2} + \frac{1}{2} + \varepsilon_0 < 5$.

Note also that the images by p of the arcs of c lying in some disk of $\mathcal{D}$ have $\overline{g}$-lengths less than $\frac{1}{2} + \varepsilon_0 = 1$.

Since $\mathrm{card}(\mathcal{C}) \leq 4h - 2$, we have $L_{\overline{g}}(p(c)) \leq 4(4h - 2) + 2$. In particular, $d_{\overline{g}}(p(x_0), \overline{x}) \leq 16h - 6$. Thus, $\mathrm{Diam}(\overline{g}) \leq 32h - 12$.

Step 4: Compare the lengths of curves.

Let C be an admissible cylinder as in Step 2. Consider an arc α of (M, g') lying in C' whose endpoints lie in $\partial C' = \gamma_1 \cup \gamma_2$ (see Step 2 for the notations). We are going to construct an arc $\overline{\alpha}$ of $(\overline{M}, \overline{g})$ homotopic to $p(\alpha)$ with the same endpoints such that $L_{\overline{g}}(\overline{\alpha}) \leq 3L_{g'}(\alpha)$.

Suppose that the endpoints of α lie in the same g'-systolic loop γ_i. Since γ_i is a g'-systolic loop, the arc $\overline{\alpha}$ of γ_i homotopic to $p(\alpha)$ with the same endpoints is not longer than α. Therefore, $L_{\overline{g}}(\overline{\alpha}) \leq L_{g'}(\alpha)$.

Suppose that the endpoints, z_1 and z_2, of α lie in γ_1 and γ_2. Recall that x_i is the endpoint on γ_i of a g'-minimizing segment γ' of C' joining γ_1 to γ_2. Denote by σ_i a shortest arc of γ_i joining x_i to z_i. We have $L_{g'}(\sigma_i) = L_{\overline{g}}(p(\sigma_i)) \leq \frac{1}{2}$. Consider the loop $c = \sigma_1 \cup \alpha \cup \sigma_2^{-1} \cup \gamma'^{-1}$ of C'. Since C' is a two-cylinder and $\mathrm{sys}(g') = L_{g'}(\gamma_1) = 1$, we have $L_{g'}(c) \geq n$ where $[c] = n[\gamma_1] \in H_1(M, \mathbb{Z})$. Thus, $L_{g'}(\alpha) \geq n - L_{g'}(\gamma') - 1$. We also have $L_{g'}(\alpha) \geq d_{g'}(\gamma_1, \gamma_2) = L_{g'}(\gamma')$. As previously, this latter equality is only an approximation in the general case. Therefore, $L_{g'}(\alpha) \geq \frac{1}{3}(n + 1)$.

This can be checked for n less than, equal to or greater than $3L_{g'}(\gamma') - 1$ (recall that $L_{g'}(\gamma') = d_{(C',g')}(\gamma_1, \gamma_2) \geq 2$). Define the arc $\overline{\alpha}$ of $(\overline{M}, \overline{g})$ by $\overline{\alpha} = p(\sigma_1^{-1} \cup \gamma_1^n \cup \sigma_2)$. By construction, $\overline{\alpha}$ is homotopic to $p(\alpha)$ with the same endpoints and $L_{\overline{g}}(\overline{\alpha}) \leq n + 1 \leq 3L_{g'}(\alpha)$.

We can now prove *iii)*. Let γ be a g-shortest representative of $a \in \pi_1(M, x_0)$. Replacing the arcs $p(\alpha)$ of $p(\gamma)$ by the arcs $\overline{\alpha}$ defined above (for each arc α of γ lying in C') yields a loop $\overline{\gamma}$ in $\overline{M}$ homotopic to $p(\gamma)$ with fixed based point $p(x_0)$. By construction, we have $L_{\overline{g}}(\overline{\gamma}) \leq 3L_{g'}(\gamma) \leq 3L_g(\gamma)$. Hence *iii)*. □

We need another definition.

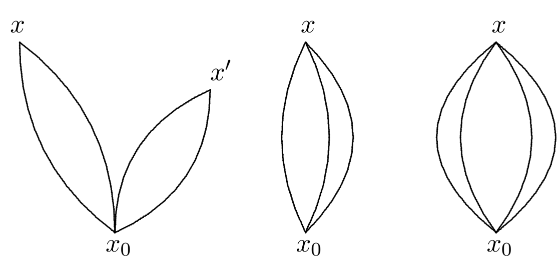
Definition 4.5. *Let $\mathcal{S} = (\alpha_1, \ldots, \alpha_h, \beta_1, \ldots, \beta_h)$ be a system of loops based at some point x_0 of a surface M of genus h. The system $\mathcal{S}$ is said to be standard if*

- *it induces a basis in $\pi_1(M, x_0)$;*
- *the loops of $\mathcal{S}$ are simple;*
- *the loops of $\mathcal{S}$ intersect each other at a single point, namely x_0;*
- *the intersection matrix of the loops of $\mathcal{S}$ is given by* $\begin{bmatrix} 0 & I_h \\ -I_h & 0 \end{bmatrix}$, *where I_h is the $h \times h$ identity matrix*

Let us now state the second result of this section.

Proposition 4.6. *Let M be a Riemannian surface of genus h and diameter D. Then, there exists a standard system of loops based at $x_0 \in M$ whose lengths are bounded from above by $L_{max} := 5^h D$.*

Proof. Let us argue by induction on the genus of M. Consider the digons formed by two minimizing segments (of length at most D) from x_0 to a point x, where the endpoint x varies.



A minimizing segment arising from x_0 cannot intersect another minimizing segment arising from x_0 in its interior. Therefore, the digons are simple loops of lengths at most $2D$ and the intersection between two of them reduces to the singleton $\{x_0\}$, a pair $\{x_0, x\}$ or a minimizing segment from x_0 to x. Furthermore, these digons induce in homology a generating set of

$H_1(M,\mathbb{Z})$. Actually, they induce in homotopy a generating set of $\pi_1(M,x_0)$ (see [8, p. 90]). The index between two of them, γ_1 and γ_2, is equal to ± 1; otherwise the index on $H_1(M,\mathbb{Z})$ would be even. If γ_1 and γ_2 intersect transversally at x, we can construct with the minimizing segments forming these two digons two other digons intersecting transversally only at x_0. Thus, we can always assume that γ_1 and γ_2 intersect transversally only at x_0 or have a minimizing segment from x_0 to x in common. In this latter case, it is possible to slightly perturb γ_1 or γ_2 into a shorter simple loop intersecting (transversally) the other digon only at x_0. This yields two simple loops of lengths at most $2D$ intersecting transversally only at x_0. Hence the result in the genus 1 case.

The commutator of these two simple loops can be represented by a simple loop γ based at x_0 of length less than $8D$ (recall that the surface is orientable). If $h \geq 2$, this simple loop is noncontractible and separates M into two parts as it vanishes in $H_1(M,\mathbb{Z})$. Gluing a hemisphere cap along the boundary of each of these two parts gives rise to two closed surfaces, M_1 and M_2, of genus less than h. By construction, we have $\mathrm{Diam}(M_i) \leq D + \frac{1}{2}L(\gamma) < 5D$, where $L(\gamma)$ is the length of γ. There exists, by induction, a standard system $\mathcal{S}_i$ of loops of M_i based at x_0 whose lengths are bounded from above by $5^{h-1}(5D)$. The loops of $\mathcal{S}_i$ can be pushed off the hemisphere cap of M_i through length-nonincreasing homotopies. Therefore, they can be considered as lying in M. Putting together the standard systems $\mathcal{S}_1$ and $\mathcal{S}_2$ yields a standard system of loops of M based at x_0 whose lengths are bounded from above by $5^h D$. $\square$

5. Comparison of the geometric and algebraic lengths

In this part, we first show an inequality similar to (1.6) for punctured surfaces. Then, we deduce Theorem 1.4 for closed orientable surfaces. The case of nonorientable surfaces is treated in the last section.

5.1. **Punctured surfaces of genus h.**

Let (N,g) be a closed orientable Riemannian surface of genus h with one topological disk removed such that its boundary ∂N is smooth. Let α_i and β_i be arcs with their endpoints in ∂N which induce a standard system of loops in the closed quotient surface $N/\partial N$ of genus h (see Definition 4.5). Without loss of generality, we can assume that the arcs α_i and β_i are length-minimizing in their homotopy classes in $\pi_1(N,\partial N)$ (see [5]). Thus, the arcs α_i and β_i are orthogonal to ∂N. Denote by L an upper bound on the lengths of α_i and β_i. Cutting N along the α_i and β_i gives rise to a Riemannian $8h$-gon Δ with right angles. Denote by

$$a_1, c_1, b_1, d_1, a_{-1}, e_1, b_{-1}, f_1, \ldots, a_h, c_h, b_h, d_h, a_{-h}, e_h, b_{-h}, f_h \tag{5.1}$$

the sides of the $8h$-gon Δ in clockwise order, and orient a_i and b_i clockwise if $i > 0$ and anticlockwise if $i < 0$. Renaming the arcs $\alpha_i, \beta_i, a_i, b_i, c_i, d_i, e_i$

and f_i if necessary, this can be done so that the surface obtained by gluing a_{-i} to a_i and b_{-i} to b_i agrees with N, and that, after gluing, the sides a_{-i} and a_i (resp. b_{-i} and b_i) identify with α_i (resp. β_i). See Figure 6 below.

Let $\pi : (\tilde{N}, \tilde{g}) \longrightarrow (N, g)$ be the Riemannian universal cover. The deck transformation group G agrees with the fundamental group of N. Hence, G is a nonabelian free group of rank $2h$. The $8h$-gon Δ identifies with a fundamental domain of N in $\tilde{N}$. Thus, the universal cover $\tilde{N}$ is tiled by the orbit $G.\Delta$ of Δ under G. Let A_i (resp. B_i) be the orientation preserving isometry of $\tilde{N}$ taking a_{-i} to a_i (resp. b_{-i} to b_i). The set $S = \{A_i, B_i\}$ is a free generating system of the deck transformation group G. In particular, every element of G has a unique minimal writing with respect to S.

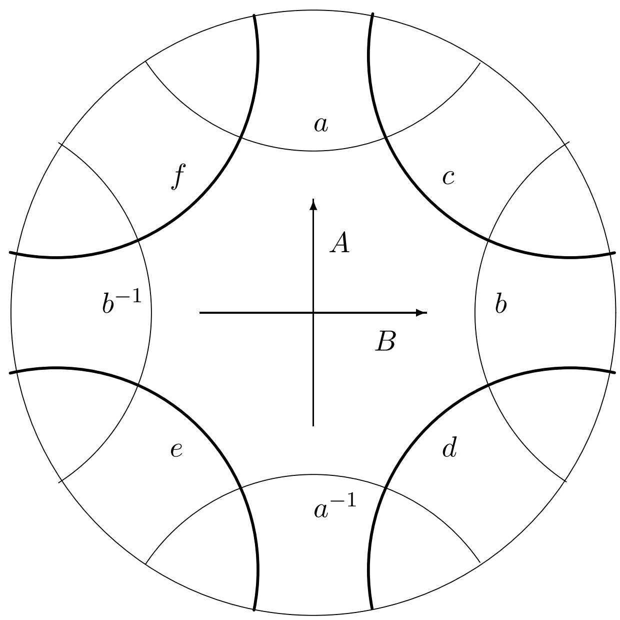


Figure 6

The following discussion relates the minimal writing of an element σ of $G \simeq \pi_1(N)$ with respect to S to the trajectory of some minimal geodesic loop γ representing this element. Let Γ be the union of the α_i and the β_i in N and $\tilde{\Gamma}$ be the preimage of Γ in $\tilde{N}$. Since the arcs α_i and β_i are length-minimizing in their homotopy classes in $\pi_1(N, \partial N)$, every minimizing segment of $\tilde{N}$ intersects the minimizing segments $\sigma.a_i$ and $\sigma.b_i$ at most once, where σ runs over G. Therefore, a minimizing segment γ joining a point x in the interior of $\Delta \subset \tilde{N}$ to $\sigma.x$, where σ is an element of G of S-length n, intersects

$$\tilde{\Gamma} = \bigcup_{\sigma \in G} \sigma.a_i \cup \sigma.b_i$$

in exactly n points. Furthermore, if $\sigma = \sigma_2 A_i^{\varepsilon} \sigma_1$ (resp. $\sigma_2 B_i^{\varepsilon} \sigma_1$) where $\varepsilon = \pm 1$, and $|\sigma_i|_S = n_i$ with $n = n_1 + n_2 + 1$, the (n_1+1)-th intersection point of γ with $\tilde{\Gamma}$ agrees with $\gamma \cap \sigma_1.a_{\varepsilon i}$ (resp. $\gamma \cap \sigma_1.b_{\varepsilon i}$).

Conversely, let γ be a minimizing segment joining a point x in the interior of $\Delta \subset \tilde{N}$ to another point $x' = \sigma.x$ in the G-orbit of x. The successive intersections of γ with the orbit of the fundamental domain Δ yield the minimal writing of σ with respect to S. For instance, if γ leaves Δ through the side $a_{\varepsilon i}$ (resp. $b_{\varepsilon i}$), then the minimal writing of σ ends with A_i^{ε} (resp. B_i^{ε}). More generally, if γ goes through $\Delta, \sigma_1.\Delta, \sigma_2\sigma_1.\Delta, \ldots, \sigma_n \ldots \sigma_2\sigma_1.\Delta$ in this order, then $\sigma = \sigma_n \ldots \sigma_2 \sigma_1$.

Assume that the systole of the surface N_0 of genus h obtained by gluing isometrically a round hemisphere along the boundary ∂N of N is greater or equal to 1, i.e., $\mathrm{sys}(N_0) \geq 1$. In particular, $L \geq 1$. Note that $N \subset N_0$.

Proposition 5.1. *With the previous notations and assumptions, let γ be a loop of N based at some point x_0 such that γ is length-minimizing in its homotopy class in $\pi_1(N_0, x_0)$. Then, we have*

$$L(\gamma) \geq C_h |\gamma|_S, \tag{5.2}$$

where $L(\gamma)$ is the length of γ and $C_h = \frac{1}{48Lh}$.

Remark 5.2. We emphasize that in this proposition, the loop γ is length-minimizing not only in its homotopy class in N, but also in its homotopy class in N_0. Such a condition is needed since the length of ∂N can be arbitrarily small, providing no lower bound on $|\partial N|_g$ solely in terms of $|\partial N|_S$.

Remark 5.3. Note also that we have identified γ with its homotopy class $[\gamma]$ in $G \simeq \pi_1(N, x_0)$. Despite the risk of confusion, we will continue to do so from now on.

Proof of Proposition 5.1. Without loss of generality, we can assume that x_0 does not lie in Γ. Set $n := |\gamma|_S$. Let us show by induction on n that $L(\gamma) \geq C_h |\gamma|_S$ where $C_h = \frac{1}{48Lh}$.

If $|\gamma|_S \leq 24Lh$, the result is obvious. Indeed, $24Lh \leq \frac{1}{C_h}$ and $L(\gamma) \geq \mathrm{sys}(N_0) \geq 1$ by assuming that the class of γ is nontrivial in $\pi_1(N_0, x_0)$. Thus, we can assume that $|\gamma|_S > 24Lh$.

In this case, from the pigeon hole principle, the element γ can be written as $[\gamma] = \zeta_1 \eta \zeta_2$ in G, where $|\gamma|_S = |\zeta_1|_S + |\eta|_S + |\zeta_2|_S$ and η is the first subword of $[\gamma]$ from the right, starting and ending with the same letter $C \in S \cup S^{-1}$, and containing $m := [4L] + 2$ times the letter C in its minimal writing. Recall that $S \cup S^{-1}$ contains $4h$ elements, that $L \geq 1$ and that $[x]$ represents the integer part of the real number x. Note that $|\eta|_S + |\zeta_2|_S \leq 24Lh$, in particular, $|\eta|_S \leq 24Lh$.

From the discussion preceding Proposition 5.1, the geodesic trajectory γ starts from x_0 and intersects Γ in exactly n points $y_1, \dots, y_n$, in this order along γ. Strictly speaking, we should say that there exist $0 < t_1 < t_2 < \dots < t_n < 1$ such that the $\gamma(t_i)$ are the only points of $\gamma : [0,1] \longrightarrow N$ lying in Γ. The points y_i agree with the $\gamma(t_i)$. The points y_p and y_q, where $p = 1 + |\zeta_2|_S$ and $q = |\eta|_S + |\zeta_2|_S$, lie in α_i if $C = A_i^{\pm 1}$, or in β_i if $C = B_i^{\pm 1}$. We will denote by δ this segment in which y_p and y_q lie, i.e., $\delta = \alpha_i$ if $C = A_i^{\pm 1}$ and $\delta = \beta_i$ if $C = B_i^{\pm 1}$. Actually, since the subword η contains m times the letter C in its minimal writing, there are m points $z_1, \dots, z_m$ in $\{y_p, \dots, y_q\}$ with $z_1 = y_p$ and $z_m = y_q$ such that γ intersects δ at $z_1, \dots, z_m$ (through the same direction). Since $m = [4L] + 2$ and $L(\delta) \leq L$, the length of some arc of δ joining two points of $\{z_1, \dots, z_m\}$ and not containing x_0 is at most $\frac{1}{4}$. Therefore, there exists an arc τ of δ with $L(\tau) \leq \frac{1}{4}$ that intersects γ only at its endpoints u and v (different from x_0). The points x_0, u and v decompose γ into three arcs γ', γ'' and γ'''.

We will need the following result.

Lemma 5.4. *We have* $L(\gamma'') \geq \frac{3}{4}$.

Proof. We can assume that $\gamma'' \cup \tau$ is simple. Otherwise, there exists a subarc c of γ'' forming a loop (recall that γ'' intersects the nonself-intersecting arc τ only at its endpoints). Since γ is length-minimizing in $\pi_1(N_0, x_0)$, the loop c is noncontractible in N_0. Therefore,

$$L(\gamma'') \geq L(c) \geq \mathrm{sys}(N_0) \geq 1,$$

which implies the desired inequality.

We want to show by contradiction that $\gamma'' \cup \tau$ is noncontractible in N_0 to conclude that

$$L(\gamma'') \geq \mathrm{sys}(N_0) - L(\tau) \geq \frac{3}{4}.$$

If the loop $\gamma'' \cup \tau$ is contractible in N_0, then the arc γ'' is no longer than τ since γ is length-minimizing in $\pi_1(N_0, x_0)$. Furthermore, $\gamma'' \cup \tau$ and ∂N bound a cylinder C in N. Thus, since γ'' and τ are length-minimizing in their homotopy classes, the nonself-intersecting trajectory δ starts from ∂N, leaves C to run across $N \setminus C$ and comes back into C only to end at ∂N. In particular, the arc $(\delta \setminus \tau) \cup \gamma''$ lies in the same homotopy class as δ in $\pi_1(N, \partial N)$. Since γ'' is no longer than τ, we can shorten δ within its homotopy class. This yields a contradiction. $\square$

We can now conclude the proof of Proposition 5.1. From the discussion preceding the statement of this proposition, the loop $\gamma' \cup \tau \cup \gamma'''$, based at x_0, satisfies $|\gamma' \cup \tau \cup \gamma'''|_S \geq |\zeta_1|_S + |\zeta_2|_S + 1$. Hence $|\gamma' \cup \tau \cup \gamma'''|_S \geq n - 24Lh$ since $|\eta|_S \leq 24Lh$. Furthermore, Lemma 5.4 and the inequality $L(\tau) \leq \frac{1}{4}$

imply that $L(\gamma'') \geq L(\tau) + \frac{1}{2}$. Therefore, by induction on n,

$$L(\gamma) = L(\gamma' \cup \gamma'' \cup \gamma''') \geq L(\gamma' \cup \tau \cup \gamma''') + \frac{1}{2} \tag{5.3}$$

$$\geq C_h|\gamma' \cup \tau \cup \gamma'''|_S + \frac{1}{2} \tag{5.4}$$

$$\geq C_h(n - 24Lh) + \frac{1}{2} \tag{5.5}$$

$$\geq C_h n \tag{5.6}$$

since $C_h \leq \frac{1}{48Lh}$. Hence, $L(\gamma) \geq C_h|\gamma|_S$. □

5.2. **Surfaces of genus h.** In this section, we prove Theorem 1.4 for orientable closed surfaces.

Let (M, g) be an orientable closed Riemannian surface of genus h with $\text{sys}(g) \geq 1$. Fix $x_0 \in M$. As previously, we will identify the loops with their homotopy classes. From Remark 4.2, we can assume that the diameter of g is at most $32h$. Let us choose a shortest representative for each class of $\pi_1(M, x_0)$ and a point x which does not lie in the union of these representatives. From Proposition 4.6, there exists a standard system of loops $\mathcal{S}_M$ based at x whose lengths are bounded from above by $L := 32h5^h$.

Let γ be one of the previous shortest representatives inducing a nontrivial class in $\pi_1(M, x_0)$. Consider the surface N obtained from M by removing a small disk D centered at x such that γ does not intersect D and every loop of $\mathcal{S}_M$, based at x, intersects D along an arc. The system of loops $\mathcal{S}_M$ restricted to N gives rise to a collection of arcs α_i and β_i with their endpoints in ∂N and their lengths bounded from above by L. This collection of arcs induces a free generating set S of $\pi_1(N)$ with $2h$ elements (see Section 5.1).

From Proposition 5.1, we have $L(\gamma) \geq C_h|\gamma|_S$. Furthermore, the inclusion $i : N \hookrightarrow M \simeq N/\partial N$ induces an epimorphism between the fundamental groups. Therefore, the system $\Sigma := i_*(S)$ is a generating set of $\pi_1(M)$ with $2h$ elements, which, by construction, does not depend on γ. In particular, $|\gamma|_S \geq |\gamma|_\Sigma$. Hence the inequality $|.|_g \geq C_h|.|_\Sigma$.

In the general case (i.e., with no assumption on the diameter of g), we obtain $|.|_g \geq \lambda|.|_\Sigma$ with $\lambda = \frac{1}{3}C_h$.

5.3. **Nonorientable surfaces.** In the previous section, we proved Theorem 1.4 for orientable surfaces. Let us show now how to extend this result to nonorientable closed surfaces.

For the projective plane, the result is obvious (take $\lambda = 1$).

Let (M, g) be a nonorientable closed Riemannian surface with $\text{sys}(g) \geq 1$, nonhomeomorphic to the projective plane. Denote by $p : (\overline{M}, \overline{g}) \longrightarrow (M, g)$ its orientable double cover. We have $\text{sys}(\overline{g}) \geq 1$. Fix $x_0 \in M$ and a lift $\overline{x}_0 \in \overline{M}$ of x_0. The map p induces an injective homomorphism p_* in homotopy such that $p_*\pi_1(\overline{M}, \overline{x}_0)$ is a subgroup of $\pi_1(M, x_0)$ of index 2. Let $\overline{\Sigma}$ be a generating set of $\pi_1(\overline{M}, \overline{x}_0)$ with $2\overline{h}$ elements such that $|.|_{\overline{g}} \geq \lambda_{\overline{h}} |.|_{\overline{\Sigma}}$, where $\overline{h}$ is the genus of $\overline{M}$. Let $\Sigma = p_*(\overline{\Sigma}) \cup \{[c]\}$, where c is a shortest loop based at x_0 with $[c] \in \pi_1(M, x_0) \setminus \mathrm{Im} p_*$. The set Σ forms a generating system of $\pi_1(M, x_0)$ with $|\Sigma| \leq 2\overline{h} + 1$.

Consider a loop γ based at x_0. If a lift $\overline{\gamma}$ of γ to $\overline{M}$ closes up, then $|\overline{\gamma}|_{\overline{g}} \geq \lambda_{\overline{h}} |\overline{\gamma}|_{\overline{\Sigma}}$. Since $|\overline{\gamma}|_{\overline{g}} = |\gamma|_g$ and $|\overline{\gamma}|_{\overline{\Sigma}} \geq |\gamma|_{\Sigma}$, we obtain the desired result. Otherwise, consider a lift γ' of the loop $\gamma \cup c$. The lift γ' closes up and $|\gamma'|_{\overline{g}} \leq |\gamma|_g + |c|_g \leq 2|\gamma|_g$. Further, $|\gamma'|_{\overline{g}} \geq \lambda_{\overline{h}} |\gamma'|_{\overline{\Sigma}} \geq \lambda_{\overline{h}} |p(\gamma')|_{\Sigma}$. Since γ is homotopic to $p(\gamma').c^{-1}$, we have $|\gamma|_{\Sigma} \leq |p(\gamma')|_{\Sigma} + 1$. Note that $|\gamma|_{\Sigma} - 1 \geq \frac{1}{2} |\gamma|_{\Sigma}$ for $|\gamma|_{\Sigma} \geq 2$. Therefore, $|\gamma|_g \geq \frac{\lambda_{\overline{h}}}{4} |\gamma|_{\Sigma}$. This latter inequality still holds for $|\gamma|_{\Sigma} < 2$ since $\mathrm{sys}(g) \geq 1$ and $\lambda_{\overline{h}} \leq 1$.

Remark 5.5. Theorem 1.2 reduces to the case when M is orientable. Indeed, let $(\overline{M}, \overline{g})$ be the orientable double cover of (M, g). Then, $h_{vol}(\overline{g}) = h_{vol}(g)$ since the volume entropy depends only on the universal cover. Furthermore, $\mathrm{sys}(\overline{g}) \geq \mathrm{sys}(g)$ since the quotient map from $\overline{M}$ to M is distance nonincreasing and induces an injective homomorphism in homotopy. Therefore, $\mathrm{sys}(\overline{g}) \cdot h_{vol}(\overline{g}) \geq \mathrm{sys}(g) \cdot h_{vol}(g)$.

References

[1] I. Babenko, *Asymptotic invariants of smooth manifolds*, Russian Acad. Sci. Izv. Math. **41** (1993) 1-38.

[2] G. Besson, G. Courtois, S. Gallot, *Volume et entropie minimale des espaces localement symétriques*, Invent. Math. **103** (1991) 417–445.

[3] P. Buser, P.. Sarnak, *On the period matrix of a Riemann surface of large genus (with an Appendix by J.H. Conway and N.J.A. Sloane)*, Invent. Math. **114** (1994) 27–56.

[4] A. Freiré, R. Mañé, *On the entropy of the geodesic flow in manifolds without conjugate points*, Invent. Math. **69** (1982) 375–392.

[5] M. Freedman, J. Hass, P. Scott, *Closed geodesics on surfaces*, Bull. London Math. Soc. **14** (1982) 385–391.

[6] M. Gromov, *Volume and bounded cohomology*, Publ. Math. IHES. **56** (1982) 5–100.

[7] M. Gromov, *Filling Riemannian manifolds*, J. Diff. Geom. **18** (1983) 1–147.

[8] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progr. in Mathematics, vol. 152, Birkhäuser, Boston, 1999.

[9] P. de la Harpe, *Topics in geometric group theory*, Chicago Lectures in Mathematics, Univ. Chicago Press, Chicago, 2000.

[10] A. Katok, *Entropy and closed geodesics*, Ergod. Th. Dynam. Sys. **2** (1983) 339–365.

[11] M. Katz, S. Sabourau, *Entropy of systolically extremal surfaces and asymptotic bounds*, Ergod. Th. Dynam. Sys. **25** (2005) 1209–1220.

[12] S. Kodani, *On two dimensional isosystolic inequalities*, Kodai Math. J. **10** (1987) 314–327.

[13] A. Manning, *Topological entropy for geodesic flows*, Ann. of Math. (2) **110** (1979) 567–573.

[14] A. Manning, *More topological entropy for geodesic flows*, Dynamical systems and turbulence, Warwick 1980 (Coventry, 1979/1980), pp. 243–249, Lecture Notes in Math., 898, Springer, Berlin-New York, 1981.

[15] W. Massey, *A basic course in algebraic topology*, GTM 127, Springer-Verlag, New York, 1991.

[16] J. Milnor, *A note on curvature and the fundamental group*, J. Diff. Geom. **2** (1968) 1–7.
[17] S. Sabourau, *Systolic volume and minimal entropy of aspherical manifolds*, preprint.
[18] A. Švarc, *A volume invariant of covering*, Dokl. Akad. Nauk. SSSR **105** (1955) 32–34.

Laboratoire de Mathématiques et Physique Théorique, Université de Tours, Parc de Grandmont, 37400 Tours, France

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6395, USA

E-mail address: sabourau@lmpt.univ-tours.fr