ISSN 0012-9593

quatrième série - tome 43 fascicule 4 juillet-août 2010

ANNALES SCIENTIFIQUES de L'ÉCOLE NORMALE SUPÉRIEURE

Florent BALACHEFF & Stéphane SABOURAU

Diastolic and isoperimetric inequalities on surfaces

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Ann. Scient. Éc. Norm. Sup.
4^{e} série, t. 43, 2010, p. 579 à 605

DIASTOLIC AND ISOPERIMETRIC INEQUALITIES ON SURFACES

BY FLORENT BALACHEFF AND STÉPHANE SABOURAU

ABSTRACT. – We prove a universal inequality between the diastole, defined using a minimax process on the one-cycle space, and the area of closed Riemannian surfaces. Roughly speaking, we show that any closed Riemannian surface can be swept out by a family of multi-loops whose lengths are bounded in terms of the area of the surface. This diastolic inequality, which relies on an upper bound on Cheeger's constant, yields an effective process to find short closed geodesics on the two-sphere, for instance. We deduce that every Riemannian surface can be decomposed into two domains with the same area such that the length of their boundary is bounded from above in terms of the area of the surface. We also compare various Riemannian invariants on the two-sphere to underline the special role played by the diastole.

RÉSUMÉ. – Nous démontrons une inégalité universelle entre la diastole, définie par un procédé de minimax sur l'espace des 1-cycles, et l'aire d'une surface riemannienne fermée. De manière informelle, nous prouvons que toute surface riemannienne fermée peut être balayée par une famille de multi-lacets dont les longueurs sont contrôlées par l'aire de la surface. Cette inégalité diastolique, qui repose sur une majoration de la constante de Cheeger, fournit en particulier un procédé effectif pour trouver de courtes géodésiques fermées sur une 2-sphère. Nous déduisons que toute surface riemannienne peut être décomposée en deux domaines de même aire dont la longueur du bord commun est majorée à l'aide de l'aire de la surface. Nous comparons également divers invariants riemanniens sur la 2-sphère afin de souligner le rôle spécial joué par la diastole.

1. Introduction

The topology of a manifold M and the topology of its loop space ΛM are closely related, through their homotopy groups, for instance. The critical points of the length or energy functionals on the loop space of a Riemannian manifold can be studied using this connection. These critical points have a special geometric meaning since they agree with the closed geodesics on the manifold.

The first author was supported by the Swiss National Science Foundation (grant 20-118014/1) during the redaction of this article. The second author has been partially supported by the Swiss National Science Foundation.

From the isomorphism between $\pi_1(\Lambda M, \Lambda^0 M)$ and $\pi_2(M)$ for the two-sphere (here, $\Lambda^0 M$ denotes the space of constant curves), G. D. Birkhoff proved the existence of a nontrivial closed geodesic on every Riemannian two-sphere using a minimax argument on the loop space (this method was extended in higher dimension by A. Fet and L. Lyusternik, *cf.* [18]).

Relationships on the lengths of these closed geodesics have then been investigated. For instance, C. Croke [9] showed that every Riemannian two-sphere M has a nontrivial closed geodesic of length

$$\mathrm{scg}(M) \leq 31\sqrt{\mathrm{area}(M)}. \tag{1.1}$$

The notation 'scg' stands for the (length of a) shortest closed geodesic. The inequality (1.1) has been improved in [20], [28], and [25].

The closed geodesic on the two-sphere obtained in C. Croke's theorem does not always arise from a minimax argument on the loop space, even though such an argument is used in the proof. Indeed, the closed geodesics obtained from a minimax argument on the loop space have a positive index when they are nondegenerate, which is the case for generic (bumpy) metrics. Now, consider two-spheres with constant area and three spikes arbitrarily long (*cf.* [28, Remark 4.10] for further detail). On these spheres, the closed geodesics satisfying the inequality (1.1) have a null index while the closed geodesics with positive index are as long as the spikes. Therefore, a minimax argument on the loop space does not provide an effective way to bound the area of a Riemannian two-sphere from below. More precisely, given a Riemannian two-sphere M, we define the diastole, $\mathrm{dias}_\Lambda(M)$, over the loop space ΛM as

$$\mathrm{dias}_\Lambda(M) := \inf_{(\gamma_t)} \sup_{0\leq t\leq 1} \mathrm{length}(\gamma_t)$$

where (γ_t) runs over the families of loops inducing a generator of $\pi_1(\Lambda M, \Lambda^0 M) \simeq \pi_2(M)$. Then,

$$\text{the ratio } \frac{\mathrm{dias}_\Lambda(M)}{\sqrt{\mathrm{area}(M)}} \text{ is unbounded} \tag{1.2}$$

on the space of Riemannian metrics on M.

Note that a diastolic inequality between the diastole over the loop space and the volume of the convex hypersurfaces in the Euclidean spaces does hold true, *cf.* [30], [9].

The existence of a closed geodesic on the two-sphere can also be proved by using a minimax argument on a different space, namely the one-cycle space $\mathcal{Z}_1(M;\mathbb{Z})$. (Recall that one-cycles are unions of loops, *cf.* Section 2 for a precise definition.) This minimax argument relies on F. Almgren's isomorphism [1] between $\pi_1(\mathcal{Z}_1(M;\mathbb{Z}), \{0\})$ and $H_2(M;\mathbb{Z})$, which holds true for every closed manifold M. Given a closed Riemannian surface M, we define the diastole over the one-cycle space as

$$\mathrm{dias}_\mathcal{Z}(M) := \inf_{(z_t)} \sup_{0\leq t\leq 1} \mathbf{M}(z_t) \tag{1.3}$$

where (z_t) runs over the families of one-cycles inducing a generator of $\pi_1(\mathcal{Z}_1(M;\Bbbk), \{0\})$ and $\mathbf{M}(z_t)$ represents the mass (or length) of z_t. Here $\Bbbk = \mathbb{Z}$ if M is orientable and $\Bbbk = \mathbb{Z}/2\mathbb{Z}$ otherwise. For short, we will write $\mathrm{dias}(M)$ for $\mathrm{dias}_\mathcal{Z}(M)$. From a result of J. Pitts [22, p. 468], [23, Theorem 4.10] (see also [6]), this minimax principle gives rise to a union of closed geodesics (counted with multiplicity) of total length $\mathrm{dias}(M)$.

Hence, $\mathrm{scg}(M) \leq \mathrm{dias}(M)$. This principle has been used in [6], [20], [28], [25], [26], [4] and [29] in the study of closed geodesics on Riemannian two-spheres.

On nonsimply connected surfaces, no minimax principle is required to show the existence of a closed geodesic. In this setting, we directly define the systole as the minimum of the length functional over the connected components not containing the trivial loops of the loop space. Every closed Riemannian surface M of genus $g \geq 1$ satisfies the following asymptotically optimal systolic inequality

$$\mathrm{sys}(M) \leq C \, \frac{\log(g+1)}{\sqrt{g}} \, \sqrt{\mathrm{area}(M)} \tag{1.4}$$

where $\mathrm{sys}(M)$ is the systole of M and C is a universal constant, *cf.* [12], [3] and [17] for three different proofs.

The goal of this article is to establish curvature-free inequalities similar to (1.1) and (1.4). The use of the one-cycle space, rather than the loop space, introduces some flexibility. It allows us to cut and paste closed curves, and to deal with both simply and nonsimply connected surfaces. We show a difference of nature between the diastole over the loop space and the diastole over the one-cycle space on the two-sphere, and between the systole and the diastole for surfaces of large genus. More precisely, we obtain the following diastolic inequality.

THEOREM 1.1. – *There exists a positive constant $C \leq 10^8$ such that every closed Riemannian surface M of genus $g \geq 0$ satisfies*

$$\mathrm{dias}(M) \leq C \, \sqrt{g+1} \sqrt{\mathrm{area}(M)}. \tag{1.5}$$

Since the minimax principle (1.3) gives rise to a union of closed geodesics of length $\mathrm{dias}(M)$, Theorem 1.1 yields a construction of *short* closed geodesics on surfaces through Morse theory over the one-cycle space. In particular, it sheds some light on the nature of the closed geodesics whose lengths provide a lower bound on the area of the two-spheres. Compare with (1.1) and (1.2).

The dependence on the genus in the inequality (1.5) is optimal, *cf.* Remark 7.3, and should be compared with the one in (1.4).

A version of the diastolic inequality (1.5) holds true for compact surfaces with boundary, in particular for disks, *cf.* Remark 2.4.

The inequality (1.5) is derived from a stronger estimate, where the diastole is replaced with the technical diastole introduced in Definition 2.2. This estimate also yields the following result.

COROLLARY 1.2. – *There exists a positive constant C such that every closed Riemannian surface M of genus $g \geq 0$ decomposes into two domains with the same area whose length of their common boundary γ satisfies*

$$\mathrm{length}(\gamma) \leq C \, \sqrt{g+1} \sqrt{\mathrm{area}(M)}. \tag{1.6}$$

The two domains in the previous result are not necessarily connected, even on two-spheres. A counterexample is given by two-spheres with three long fingers, whose area is equally concentrated in the tips of these fingers (*cf.* [28, Remark 4.10]).

The second author [28] showed that the area of a bumpy Riemannian two-sphere can be bounded from below in terms of the length of its shortest one-cycle of index one. Our inequality (1.5) on the two-sphere extends this result since the minimax process used in the definition of the diastole gives rise to a one-cycle of index one when the metric is bumpy. However, the relation between the filling radius of a bumpy Riemannian two-sphere and the length of its shortest one-cycle of index one established in [28] cannot be extended to the diastole. Indeed, from [28, Theorem 1.6], there exists a sequence M_n of Riemannian two-spheres such that

$$\lim_{n\to+\infty} \frac{\mathrm{FillRad}(M_n)}{\mathrm{dias}(M_n)} = 0.$$

This result illustrates the difference of nature between the length of the shortest closed geodesic or of the shortest one-cycle of index one, which can be bounded by the filling radius on the two-sphere, and the diastole. It also shows that the proof of Theorem 1.1 requires different techniques.

The arguments used throughout this article are rather "elementary" and robust. It is our hope that they can be adapted to investigate further problems.

At the end of this article, we consider a minimax principle on the one-cycle space for another functional and establish further geometric inequalities on the two-sphere.

Higher dimensional analogs of the diastole, sometimes called k-widths, have been investigated in [2], [23], [12] and [15]. In particular, L. Guth [15] obtained upper bounds on the k-width of Euclidean domains in terms of their n-dimensional volumes. He also showed that no such bound holds in the Riemannian setting for the $(n-1)$-width with $n \geq 3$. Our main result provides a positive result on Riemannian surfaces, answering a question raised in [15, p. 1148] in a particular case.

Let us present the structure of the article and an outline of the proof of the diastolic inequality (1.5). The one-cycle space and the definition of the diastole are presented in Section 2. In Section 3, we show how to replace a Riemannian surface with a simplicial piecewise flat surface with comparable area and diastole. This will enable us to prove the main diastolic inequality by induction on the number of simplices in an approaching piecewise flat surface. In Section 4, we prove an upper bound on Cheeger's constant, and on some related invariant, in terms of the area of the surface alone. This upper bound yields an isoperimetric inequality which permits us to split the surface into two domains D_1 and D_2 whose boundary lengths are small in comparison to their areas. This isoperimetric inequality can be thought of as a discrete version of the diastolic inequality. At this stage, we could cap off the two domains D_1 and D_2 and apply the discrete version of the diastolic inequality to the two resulting surfaces with the hope to derive a continuous version of it by iterating this principle sufficiently many times. Unfortunately, passing from a discrete parameter family of one-cycles to a continuous parameter family on a smooth surface is technically challenging. Arguing by induction on the number of simplices of a piecewise flat surface approaching the initial surface turns out to be more manageable in this case. More specifically, we argue as follows. In Section 5, we prove a simplicial version of the isoperimetric inequality obtained