WHAT IS...

a Systole?

Marcel Berger

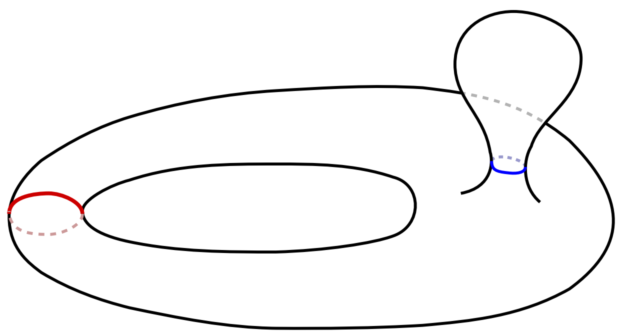

Figure 1. The closed geodesic on the right is not a systole since it is contractible.

Let us take a look at Figure 1, which represents a surface S in ordinary three-dimensional Euclidean space (throughout we will assume surfaces and other objects are compact). A closed curve on S is a curve that looks topologically like a circle. Because our surface S has the topology of a torus, there are closed curves on S that are not contractible to a point. We define the *systole* of S, $Sys(S)$, to be the smallest length of such curves. By a compactness argument, this lower bound is positive and is realized by at least one curve, which is a *closed geodesic*. On surfaces or on general Riemannian manifolds geodesics are those curves that are locally length-minimizing.

The medical term *systole* comes from the Greek word for "contraction". (If you have extra systolic beats in the medical and not geometrical sense, you had better consult your cardiologist.) The mathematical term *systole* was coined in 1980.

Marcel Berger is a former director of the Institut des Hautes Études Scientifiques. His email address is `mberger@ihes.fr`.

As the figure intuitively shows, given $Sys(S)$, the total area of S, $Area(S)$, cannot be too small. A natural question is, What is the relationship between $Sys(S)$ and $Area(S)$? We are looking for a kind of isoperimetric inequality, but in this case it's a game without boundary. The first person to tackle this problem was Loewner, who in 1949 proved that for any surface of the topological type of a torus, $Area(S) \geq \frac{\sqrt{3}}{2} Sys^2(S)$. This is an *isosystolic inequality*. Loewner proved that equality is attained only and exactly for the flat equilateral torus. The proof is not too hard if one knows the basic conformal representation theorem.

Now, any mathematical mind will ask for generalizations of Loewner's theorem, and they could go in at least three directions. First: Consider surfaces more general than the torus, the simplest being the projective plane. Pu, a student of Loewner, proved in 1952, with the same method that Loewner had used, that $Area(S) \geq \frac{2}{\pi} Sys^2(S)$, with equality holding only for the standard metric on S. Second: Consider generalizations in higher dimension. Third: Consider generalizations to submanifolds (that is, generalizing the initial case of curves) of any dimension. In any of these problems, the question splits in two: 1) What is the optimal ratio if it is nonzero? and 2) For which metric is it attained?

Today geometers are fascinated by those problems for several reasons. For one thing, despite the efforts of many, not a single one of those generalizations was obtained before 1983, when Gromov was the first to crack the nut. Also, Gromov's proofs are technically extremely hard, and, even more importantly, they introduced some completely new concepts in geometry. And finally, many "elementary" questions remain open today. But what is even more fascinating for Riemannian geometers is the fact that we finally get inequalities on a Riemannian manifold that are true without any curvature restriction; i.e., they are valid for any Riemannian structure on the manifold.

Let us now see what the state of affairs is today and look at the main ideas, concepts, and techniques that enter into the proof. The first case to look at is that of surfaces with any number of holes (the torus is a surface with one hole). In 1960 Accola and Blatter got an inequality, but with a constant that was getting smaller and smaller as the number of holes became larger and larger. Their papers launched the search in this subject. It is interesting to remark that their proofs, which were quite similar, were extensions of Loewner's method in that they used the conformal representation. For us it is the only case where complex analysis on surfaces gives a result that is dead wrong. One had to wait for Gromov in order to have a constant that grows with the genus. Exact constants are not known, and in fact are not that interesting, but one has optimal asymptotic results for them when the number of holes goes to infinity.

Metrics for which the equality is attained are not known, except for the Klein bottle and, as seen above, for the torus and the projective plane. In fact, they are forced to be singular (not smooth) even for this case.

Now let's leave surfaces and consider Riemannian structures on manifolds M of higher dimension d, first for closed curves and for the same definition for the systole $Sys(M)$. We ask if $Vol(M)/Sys^d(M)$ has, for the set of all Riemannian structures on M, a positive lower bound. We of course need M to be nonsimply connected (the algebraic topology wording for asking that not all closed curves in M be contractible). But this condition is not enough; look for example at the product of circle by a sphere. Then the systole is the length of the circle, but the volume of the sphere can be as small as desired. We should have enough families of noncontractible curves; in other words, those curves should fill M in every direction. Such manifolds are called *essential* by Gromov. A typical example is the projective space (of any dimension) $\mathbf{R}P^d$. Gromov proved that there is in fact an isosystolic inequality for any essential manifold. Still, as opposed to the case of $\mathbf{R}P^2$, we do not know for $d \geq 3$ if this constant is that of the canonical metric on $\mathbf{R}P^d$, with equality holding only for that canonical metric.

The proof of Gromov is one of the most baffling we know in geometry and exemplifies his mathematical style. He attacks mathematical problems in three characteristic ways: with radically new techniques, with new and extremely simple invariants that are incredibly hard to study, and with very involved and hard calculations.

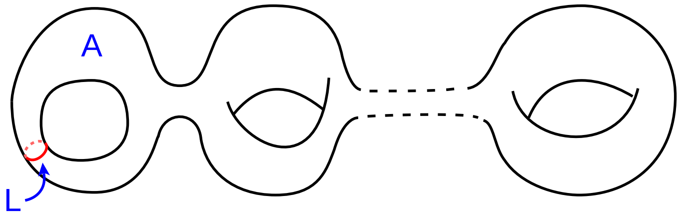

Figure 2. Intuitively, it seems clear that the ratio A/L^2, where L is the length of the shortest non-contractible curve, grows as the genus does. But it is very difficult to prove.

One trick and two new invariants. One first embeds M isometrically in the infinite-dimensional space $C_0(M)$, which consists of the continuous functions on M. The embedding simply sends a point to the function that is the distance to that point. And now we fill up the image of M in $C_0(M)$ by submanifolds of dimension $d+1$ whose boundary is this image. In this situation one can prove (hard) an infinite-dimensional isoperimetric inequality between the volume of M and the volume of the filler. This is an inequality like that for minimal surfaces. One needs thereafter to study the filling volume and the filling radius of M. Those two invariants are so deep, even though they are elementary and natural, that the filling volume of the circle is only conjectured to be 2π (think of a hemisphere). One finishes the proof with an inequality between the filling volume and the filling radius, and the main point is that the filling radius is directly linked to the systole.

Systolic freedom almost everywhere. The natural generalization is to look at what replaces, for higher-dimensional submanifolds in Riemannian manifolds of any dimension, the notion of noncontractible. From algebraic topology the most important notion is that of the homology class of such a submanifold, which is an integral homology class. For a Riemannian manifold M of dimension d we define its k-dimensional systole $Sys_k(M)$ as the infimum of the (k-dimensional) volume of all its submanifolds of dimension k whose homology class is nonzero. Then we look at the quotient $Vol^{k/d}(M)/Sys_k(M)$ and ask whether this quotient, considering all the Riemannian structures on M, is always bounded away from zero. When it is zero, one can talk of *systolic freedom* (or systole softness). The surprising discovery, completely nonintuitive for us and initiated by Gromov, is that this infimum is zero for most manifolds and all pairs ($d \geq 3, k \geq 2$). The first shock comes with the complex projective plane, for which the pair is

(4,2), because the projective complex lines fill up our space completely in every direction.

To get more definitive results one has to introduce a notion of algebraic topology called stable homology, and this time the invariant attached to a submanifold is its stable homology class, which is a real homology class. Here also it is Gromov who gave deep impetus to the subject. But here one classical tool is available, namely the calculus of differential forms, as well as the basic relation between topology and differential calculus and the basic interplay given by the theorem of de Rham.

For the stable systole problem, the freedom question was completely solved in 2007 by M. Brunnbauer (his papers are available on the `arXiv`).

Despite their quite recent introduction, systoles are already used in various domains, namely in algebraic geometry (to characterize Jacobians among flat tori; this is the so-called *Schottky problem*, which has many algebraic solutions but here is given a geometrical one) and in deep algebraic topology. One-dimensional systoles, which arise when one studies displacements by isometries in essential manifolds, are linked to the notions of entropy and the spherical volume.

Further Reading

Pages 325–353 of the author's book contain a more explicit and detailed exposition of the state of systolic affairs up to 2003 and all the references and credits. The book by Katz covers almost all the results and references for more recent developments and gives fascinating historical data.

[1] M. Berger, *A Panoramic View of Riemannian Geometry*, Springer, 2003.

[2] M. Katz, *Systolic Geometry and Topology*, Amer. Math. Soc., 2007.