A characterization of round spheres in terms of blocking light

Benjamin Schmidt and Juan Souto

Abstract. A closed Riemannian manifold M is said to have cross (compact rank one symmetric space) blocking if whenever $p \neq q$ are less than the diameter apart, all light rays from p can be shaded away from q with at most two point shades. Similarly, a closed Riemannian manifold is said to have sphere blocking if for each $p \in M$ all the light rays from p are shaded away from p by a single point shade. We prove that Riemannian manifolds with cross and sphere blocking are isometric to round spheres.

In this note we characterize constant curvature spheres in terms of light blocking properties.

Definition (Light). Let X, Y be two nonempty subsets of a Riemannian manifold M, and let $G_M(X, Y)$ denote the set of unit speed parametrized geodesics $\gamma : [0, L_\gamma] \to M$ with initial point $\gamma(0) \in X$ and terminal point $\gamma(L_\gamma) \in Y$. The *light from X to Y* is the set

$$L_M(X, Y) = \{\gamma \in G_M(X, Y) \mid \text{interior}(\gamma) \cap (X \cup Y) = \emptyset\}.$$

A subset $Z \subset M$ *blocks* the light from X to Y if the interior of every $\gamma \in L_M(X, Y)$ meets Z.

Intuitively, we are postulating that X emits light traveling along geodesics, that Y consists of receptors, and that X and Y are opaque while the remaining medium $M \setminus \{X \cup Y\}$ is transparent. From this point of view, $L_M(X, Y)$ is the set of light rays from X to Y and a set Z blocks the light from X to Y if it completely shades X away from Y. This simple model ignores diffraction, the dual nature of light, and all aspects of quantum mechanics.

A well-known result of Serre [Se51] asserts that for closed M and points $x, y \in M$, the set $G_M(x, y)$ of geodesic segments joining x and y is always infinite. In contrast,

$L_M(x, y)$ is sometimes infinite and sometimes not. For instance, if x and y are different points on the standard round sphere $\mathbb{S}^n$ with distance less than π, then $L_{\mathbb{S}^n}(x, y)$ consists of exactly two elements. In particular, we see that, under the same assumptions, it suffices to declare two additional points in $\mathbb{S}^n$ to be opaque in order to block all the light rays from x to y.

Definition (Blocking number). Let $x, y \in M$ be two (not necessarily distinct) points in M. The *blocking number* $b_M(x, y)$ for $L_M(x, y)$ is defined by

$$b_M(x, y) = \inf\{n \in \mathbb{N} \cup \{\infty\} \mid L_M(x, y) \text{ is blocked by } n \text{ points}\}.$$

The study of blocking light (also known as security) seems to have originated in the study of polygonal billiard systems and translational surfaces (see e.g. [Fo90], [Gu05], [Gu06], [GuA], [HS98], [Mo04], [Mo05], [MoA], [MoB], and [Ta]). More recently, blocking light has been studied in Riemannian spaces (see e.g. [BaGu], [BG], [GB], [GS06], [He], and [LS07]). Here we give a characterization of the round sphere in terms of its blocking properties.

If x, y are two distinct points in the standard round sphere $\mathbb{S}^n$ closer than π then, as remarked above, $b_{\mathbb{S}^n}(x, y) \leq 2$. This property does not characterize the round sphere amongst all closed Riemannian manifolds. In fact, every compact rank one symmetric space, or CROSS for short, has the following property:

Cross blocking: For every pair of distinct points $x, y \in M$ with $d_M(x, y) < \operatorname{diam}(M)$, we have $b_M(x, y) \leq 2$.

Apart from cross blocking, the round sphere also has the following property:

Sphere blocking: For every point $x \in M$, we have $b_M(x, x) = 1$.

The CROSSes are classified and consist of the round spheres $\mathbb{S}^n$, the projective spaces $K\mathbb{P}^n$ where K denotes one of $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and the Cayley projective plane, each one endowed with its symmetric metric. It is not difficult to check that the round sphere is the only CROSS with sphere blocking.

In [LS07] it was conjectured that a closed Riemannian manifold with cross and sphere blocking is isometric to a round sphere. We prove that this is the case:

Theorem 1. *A closed Riemannian manifold M has cross and sphere blocking if and only if M is isometric to a round sphere.*

In order to prove Theorem 1 we show that manifolds as in the statement are Blaschke manifolds. Recall that a compact Riemannian manifold M is said to be *Blaschke* if its injectivity radius and diameter coincide. Berger [Be78] proved that a Blaschke manifold diffeomorphic to the sphere is in fact isometric to a round sphere. This was used in [LS07] to prove Theorem 1 for Blaschke manifolds.

In [LS07] it was also conjectured that a closed Riemannian manifold with cross blocking is isometric to a compact rank one symmetric space. We prove that this is the case in dimension two:

Theorem 2. *A closed Riemannian surface M has cross blocking if and only if M is isometric to a constant curvature sphere or projective plane.*

Section 1 contains some preliminary material concerning Morse theory for path spaces and properties of totally convex subsets in Riemannian manifolds. In Section 2 we prove Theorems 1 and 2.

1. Preliminaries

In order to fix notation we start reviewing some well-known definitions and results in differential geometry. We then review the basic aspects about Morse theory on path spaces and about totally convex subsets in Riemannian manifolds needed in Section 2. Good references for this material include Milnor's *Morse Theory* [Mi63] and Cheeger and Ebin's *Comparison Theorems in Riemannian Geometry* [CE75].

1.1. Basic definitions and notation. Let M be a closed manifold with a Riemannian metric $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. The length and energy of a piecewise smooth curve $\gamma \colon [0, T] \to M$ are given by

$$
\begin{aligned}
L_M(\gamma) &= \int \|\dot{\gamma}(t)\| \, dt, \\
E_M(\gamma) &= \int \|\dot{\gamma}(t)\|^2 \, dt.
\end{aligned}
\tag{1.1}
$$

The Cauchy–Schwarz inequality implies that $L(\gamma)^2 \leq T \cdot E(\gamma)$ with equality if and only if γ has constant speed $\|\dot{\gamma}\|$. A curve with constant speed 1 is said to be parametrized by arc-length. The distance $d_M(x, z)$ between two points in M is the infimum of the lengths of curves joining them and the diameter $\mathrm{diam}(M)$ is the maximal distance between points in M. A parametrized curve $\gamma \colon (0, T) \to M$ is a geodesic if it is locally distance minimizing. Equivalently, γ fulfills the geodesic differential equation; hence, geodesics are smooth. We will often say that the image

of a geodesic is a geodesic as well. Geodesics will usually be denoted by Greek letters $\gamma, \eta, \tau, \ldots$. A variation of geodesics is a smooth map $(s, t) \rightarrow \gamma_s(t)$ where γ_s is a geodesic for all s. The vector field $\frac{\partial}{\partial s}\gamma_s(t)$ along the curve γ_0 is said to be a Jacobi field. A vector field along a geodesic is a Jacobi field if and only if it satisfies the so called Jacobi equation, a second order ordinary differential equation. In particular, the space of Jacobi fields along a geodesic is a finite dimensional vector space and every Jacobi field J is determined by its initial value and derivative. Two points x and y in M are conjugate along a geodesic arc γ joining them if there is a nonzero Jacobi field along γ vanishing at x and y.

By the Hopf–Rinow theorem, any two points in M are joined by a geodesic segment whose length realizes the distance between them. Moreover, for every point $p \in M$ and for every direction $v \in T_p M$ there is a geodesic $t \mapsto \exp_p(tv)$ starting at p with direction v. Thus we obtain the so called exponential map

$$\exp_p \colon T_p M \rightarrow M.$$

The exponential map is a local diffeomorphism in some small neighborhood of $0 \in T_p M$. The injectivity radius $\mathrm{inj}_p(M)$ is the maximum of those $r > 0$ such that exponential map is injective on the ball $B(0, r) = \{v \in T_p M \mid \|v\| < r\}$. The map $p \mapsto \mathrm{inj}_p(M)$ is continuous and hence attains a minimum, the injectivity radius $\mathrm{inj}(M)$ of the manifold.

For the sake of concreteness we will always assume that the manifolds in question have injectivity radius $\mathrm{inj}(M) \geq 2$ and will simply denote the length and energy functions by L and E instead of L_M and E_M.

1.2. The space of broken geodesics. Given $k \in \mathbb{N}$ let $\mathcal{L}_k$ be the set of piecewise geodesic curves consisting of at most k edges of at most length 1. To be more precise, elements $\gamma \in \mathcal{L}_k$ are continuous curves

$$\gamma \colon [0, k] \rightarrow M$$

such that for all $i = 0, 1, \ldots, k - 1$ the curve $\gamma|_{[i,i+1]}$ is a geodesic segment with length at most 1. When we endow $\mathcal{L}_k$ with the compact open topology, the valuation map

$$\mathcal{L}_k \rightarrow M^{k+1}, \quad \gamma \mapsto (\gamma(0), \ldots, \gamma(k+1))$$

is continuous. Moreover, the assumption that $\mathrm{inj}(M) \geq 2$ implies that this map is injective and hence a homeomorphism onto its image. The interior $\mathcal{L}_k^\circ$ of $\mathcal{L}_k$, as a subset of M^{k+1}, is the set of those elements γ consisting of geodesic arcs of length strictly less than 1. The tangent space $T_\gamma \mathcal{L}_k^\circ$ at $\gamma \in \mathcal{L}_k^\circ$ is naturally identified with the space of the continuous vector fields J along γ such that $J|_{[i,i+1]}$ is Jacobi for all $i = 0, 1, \ldots, k - 1$. Observe that this identification of $T_\gamma \mathcal{L}_k^\circ$ is consistent with

the identification of $\mathcal{L}_k^0$ with an open subset of M^{k+1}. In particular, the later point of view induces a Riemannian metric $\langle\langle\cdot,\cdot\rangle\rangle$ on $\mathcal{L}_k^\circ$.

Given two points $p, q \in M$ set

$$\mathcal{L}_k(p, q) = \{\gamma \in \mathcal{L}_k(p, q) \mid \gamma(0) = p, \gamma(k) = q\}.$$

Obviously $\mathcal{L}_k(p, q)$ is a closed subset of $\mathcal{L}_k$ homeomorphic to a closed subset of M^{k-1}. Moreover, from the description above we obtain that the interior $\mathcal{L}_k(p, q)^\circ$ of $\mathcal{L}_k(p, q)$ as a subset of M^{k-1} coincides with the intersection of $\mathcal{L}_k(p, q) \cap \mathcal{L}_k^\circ$. In particular, the tangent space $T_\gamma \mathcal{L}_k(p, q)^\circ$ of $\mathcal{L}_k(p, q)^\circ$ at some curve γ is given by the space of continuous vector fields J along γ which vanish at 0 and k and such that $J|_{[i,i+1]}$ is Jacobi for all $i = 0, 1, \ldots, k-1$. The energy function $E(\cdot)$ is smooth in $\mathcal{L}_k^\circ$ and the first variation formula asserts that the derivative of $E|_{\mathcal{L}_k(p,q)^\circ}$ at some point γ is given by

$$d(E|_{\mathcal{L}_k(p,q)^\circ})_\gamma(\cdot) = 2\sum_{i=1}^{k-1}\langle-\Delta\gamma(i),\cdot\rangle \tag{1.2}$$

where $\Delta\gamma(t) = \partial^+\gamma(t) - \partial^-\gamma(t)$ and $\partial^+\gamma(t)$ and $\partial^-\gamma(t)$ are the right and left derivatives at t. Let $\mathcal{X}$ be the negative gradient of $E|_{\mathcal{L}_k(p,q)^\circ}$, i.e.,

$$d(E|_{\mathcal{L}_k(\gamma(0),\gamma(k))^\circ})_\gamma(\cdot) = -\langle\langle\mathcal{X}_\gamma,\cdot\rangle\rangle$$

and let ϕ be the associated negative gradient flow

$$\phi'(t) = \mathcal{X}_{\phi(t)}, \quad \phi(0) = \gamma \tag{1.3}$$

Observe that since the vector-field $\Delta\gamma(t)$ is smooth not only on $\mathcal{L}_k(p, q)^\circ$ but on the whole space $\mathcal{L}_k^\circ$, the vector field $\mathcal{X}$ and the flow (ϕ_t) are also smooth when considered on the whole of $\mathcal{L}_k^\circ$.

In general, gradient lines are not defined for all $t \in \mathbb{R}$ but just for some open sub-interval. However we claim that the flow ϕ is defined for all non-negative t. In fact, consider the function

$$\lambda\colon \mathcal{L}_k(p, q) \to [0, 1], \quad \lambda(\gamma) = \text{length of the longest segment in } \gamma.$$

It is easy to check that

$$\lim_{t\to 0,\, t>0} \frac{\lambda(\phi_\gamma(t)) - \lambda(\gamma)}{t} \le 0.$$

This implies that λ is non-increasing and hence that flow lines never come close to the boundary in positive times since $\mathcal{L}_k(p, q)^\circ = \{\lambda < 1\}$. Thus, we have:

Lemma 1 ($\mathscr{L}_k^\circ$ is a cage). *There is a semi-flow*

$$\phi \colon \mathscr{L}_k^\circ \times [0, \infty) \to \mathscr{L}_k^\circ, \quad (\gamma, t) \mapsto \phi_\gamma(t),$$

such that $\phi_\gamma(0) = \gamma$ and $\frac{d}{dt}\phi_\gamma(t) = \mathcal{X}_{\phi_\gamma(t)}$ for all γ and t. Moreover, the semi-flow preserves $\mathscr{L}_k(p, q)^\circ$ for all $p, q \in M$. ☐

We now consider the restriction of the energy function E to $\mathscr{L}_k(p, q)^\circ$ for some pair of points $p, q \in M$. In order to relax notation we write E instead of $E|_{\mathscr{L}_k(p,q)^\circ}$. It follows directly from the first variation formula (1.2) that the critical points of E are precisely the geodesics of length less than k joining p and q.

Lemma 2 (Third geodesic). *Assume that $\gamma_0, \gamma_1 \in \mathscr{L}_k(p, q)$ are minimizing geodesics joined by a continuous curve $\gamma \colon [0, 1] \to \mathscr{L}_k(p, q)$, $s \mapsto \gamma_s$. Then there is a third geodesic $\alpha \in \mathscr{L}_k(p, q)$ joining p to q with $E(\alpha) \le \max_{s \in [0,1]} E(\gamma_s)$.*

Proof. Let $c = E(\gamma_0) = E(\gamma_1)$, $C = \max_{s \in [0,1]} E(\gamma_s)$, and assume that $\gamma_0, \gamma_1 \in \mathscr{L}_k(p, q)$ are the only geodesic segments joining p and q with energy not more than C. Then for each $s \in (0, 1)$, $c < E(\gamma_s) \le C$ and by Lemma 1 and the paragraph following that lemma, $\phi_{\gamma_s}(t)$ converges to either γ_0 or γ_1 as $t \to \infty$. Choose $t_0 > 0$ so that for all $s \in [0, 1]$, $d(\phi_{\gamma_s}(t_0), \{\gamma_0, \gamma_1\}) < d(\gamma_0, \gamma_1)/3 := d_0$. The assumption that both γ_0 and γ_1 are minimizing implies that for $s > 0$ sufficiently close to zero (resp. close to 1), $d(\phi_{\gamma_s}(t_0), \gamma_0) < d_0$ (resp. $d(\phi_{\gamma_s}(t_0), \gamma_1) < d_0$.) Finally, for $i = 0, 1$, define $S_i \subset [0, 1]$ by $S_i = \{s \in (0, 1) \mid d(\phi_{\gamma_s}(t_0), \gamma_i) < d_0\}$. Then $(0, 1)$ is the disjoint union of the two nonempty open sets S_0 and S_1, a contradiction. ☐

1.3. Totally convex subsets

Definition. A set C in a complete Riemannian manifold M is called totally convex if whenever $p, q \in C$ and η is a geodesic segment from p to q, then $\eta \subset C$.

A closed totally convex set $C \subset M$ has the structure of an embedded topological submanifold with smooth interior and possibly nonempty and nonsmooth boundary (see e.g. [CE75, Chapter 8]). The next result is Theorem 8.14 in [CE75].

Theorem 3. *Let C be a compact boundaryless totally convex set C in M. Then the inclusion $C \subset M$ is a homotopy equivalence.*

The idea behind the proof is to apply the negative gradient flow of the energy functional on the space $\mathscr{L}_C$ consisting of curves in M with endpoints in C. As C is totally convex, the only critical points are the constant curves into C. It follows that $C \subset \mathscr{L}_C$ is a deformation retract, proving that the relative homotopy groups $\pi_i(M, C)$ vanish. The next corollary is an easy consequence of Theorem 3.

Corollary 1. *Assume that $\gamma \subset M$ is a closed geodesic in a closed Riemannian manifold M of dimension at least two. Then there exists a geodesic segment $\eta \colon [0, 1] \to M$ with endpoints in γ but not completely contained in γ.*

Proof. If not, then $\gamma \subset M$ is a totally convex subset and hence by Theorem 3, M is homotopy equivalent to γ. This is a contradiction since the fundamental class $[M] \in H_{\dim(M)}(M, \mathbb{Z}/2\mathbb{Z})$ is a nonzero element. $\qquad\square$

2. Main theorems

In this section we prove Theorems 1 and 2. The bulk of the work lies in proving the following technical result.

Proposition 1. *Suppose that M is a closed Riemannian manifold with cross blocking. If M is not a Blaschke manifold, then there is simple closed geodesic $\gamma \subset M$ of length $2\operatorname{inj}(M)$.*

Proof. We assume that M has been scaled so that $\operatorname{inj}(M) = 2$. Choose $p \in M$ with $\operatorname{inj}_p(M) = \operatorname{inj}(M)$ and let $\operatorname{cut}(p) \subset T_p M$ be its cut-locus. Choose $\theta \in \operatorname{cut}(p)$ with $\|\theta\| = 2$ realizing the injectivity radius. For $r > 0$ and $v \in T_p M$ denote by $B(v, r) \subset T_p(M)$, the open ball with radius r and center v. We first argue that there is an open neighborhood $U \subset T_p M$ of θ for which the restriction of $\exp_p \colon T_p M \to M$ to $U \cap \operatorname{cut}(p)$ is one-to-one.

Indeed, if this were not the case, then the restriction of $\exp_p$ to $B(\theta, r) \cap \operatorname{cut}(p)$ is not one-to-one for each $r > 0$. Fix a positive ε' smaller than $\frac{1}{2}$. By continuity of the exponential map and the distance function in M, there is a sufficiently small $r_0 > 0$ so that for all $\theta_0, \theta_1 \in B(\theta, r_0)$ we have that

$$ d_M\left(\exp_p\left(\frac{\theta_0}{2}\right), \exp_p\left(\frac{\theta_1}{2}\right) \right) < \frac{\varepsilon'}{2}. $$

Let $\varepsilon < \min\{\varepsilon', r_0, \operatorname{diam}(M) - 2\}$ and choose $\theta_0, \theta_1 \in B(\theta, \varepsilon) \cap \operatorname{cut}(p)$ with $\exp_p(\theta_0) = \exp_p(\theta_1) := q$. Define $\gamma_i \colon [0, 4] \to M$ by $\gamma_i(t) := \exp_p\left(t\frac{\theta_i}{4}\right)$ for $i = 0, 1$. Note that both γ_0 and γ_1 are minimizing geodesics between p and q with $L(\gamma_i) \le 2 + \varepsilon$ for $i = 0, 1$. We consider the curve

$$ \sigma_p \colon [0, 1] \to T_p M, \quad \sigma_p(s) = (1 - s)\frac{\theta_0}{2} + s\frac{\theta_1}{2} $$

in the tangent space to M at p and its image under the exponential map

$$ \sigma \colon [0, 1] \to M, \quad \sigma(s) = \exp_p(\sigma_p(s)). $$

For each $s \in [0, 1]$, we have that

$$d_M(q, \sigma(s)) \le d_M(q, \sigma(0)) + d_M(\sigma(0), \sigma(s)) \le \frac{2+\varepsilon}{2} + \frac{\varepsilon'}{2} < 1 + \varepsilon' < 2.$$

Therefore, there is a unique curve $\sigma_q \colon [0, 1] \to B(0, 2) \subset T_q M$ with $\exp_q(\sigma_q(s)) = \sigma(s)$. For $s \in [0, 1]$, define the one parameter family of curves $s \mapsto \gamma_s$ by

$$\gamma_s(t) = \begin{cases} \exp_p\left(t \frac{\sigma_p(s)}{2}\right) & \text{for } t \in [0, 2], \\ \exp_q\left((4-t)\frac{\sigma_q(s)}{2}\right) & \text{for } t \in [2, 4]. \end{cases}$$
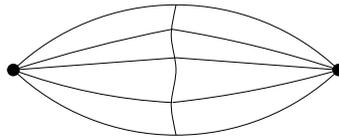
Figure 1. The variation γ_s interpolating by not much longer curves between the geodesics γ_0 and γ_1.

It is easy to check that $L(\gamma_s|_{[i,i+1]}) < 1$ for all $s \in [0, 1]$ and $i = 0, \dots, 3$ so that this family defines a continuous curve $\gamma \colon [0, 1] \to \mathcal{L}_4(p, q)^\circ$, $s \mapsto \gamma_s$, connecting γ_0 and γ_1. One also checks easily that for each $s \in [0, 1]$, the curve γ_s has at most energy $(1 + \varepsilon')^2$ so that by Lemma 2, there is a third geodesic $\alpha \in \mathcal{L}_4(p, q)^\circ$ joining p to q with $E(\alpha) \le (1 + \varepsilon')^2$. It follows that $L(\alpha) \le 2 + 2\varepsilon' < 3$. Note that since each of α, γ_0, and γ_1 have length strictly less than 4, no two can intersect in their interiors without contradicting $\mathrm{inj}(M) = 2$. Hence, $b_M(p, q) \ge 3$, a contradiction to cross blocking since $d_M(p, q) \le 2 + \varepsilon < \mathrm{diam}(M)$.

We have proved that there is some open neighborhood $U \subset T_p M$ of θ such that the restriction of $\exp_p$ to $U \cap \mathrm{cut}(p)$ is one-to-one. From now on, let U be such a neighborhood.

We argue next that there are at least two distinct unit speed minimizing geodesics $\gamma_0, \gamma_1 \colon [0, 2] \to M$ joining p and $q := \exp_p(\theta)$ (and hence exactly two by the cross blocking condition). Define

$$r_p \colon \partial \bar{B}(0, 1) \to (0, \mathrm{diam}(M)],$$
$$r_p(v) = \sup\{t \in (0, \mathrm{diam}(M)] \mid d_M(p, \exp_p(tv)) = t\}.$$

It is well known that the function r_p is continuous. Hence, the function

$$i_p \colon \bar{B}(0, 1) \setminus \{0\} \to T_p M, \quad i_p(x) = r_p\left(\frac{x}{\|x\|}\right) x$$

is continuous as well. Therefore, $i_p^{-1}(U)$ is an open subset of $\frac{\theta}{2}$ in $\bar{B}(0, 1)$. Choose $\delta > 0$ sufficiently small so that the set $V_\delta := \bar{B}\left(\frac{\theta}{2}, \delta\right) \cap \bar{B}(0, 1)$ is contained in

$i_p^{-1}(U_\theta)$. Note that V_δ is homeomorphic to a basic closed set of 0 in the upper half space $\mathbb{R}^n_{x_n \geq 0}$ and that the map $\exp_p \circ i_p$ is continuous and one-to-one on V_δ. Hence, $\exp_p(i_p(V_\delta))$ does not cover an entire neighborhood of q so that we find a sequence of points $q_i \in M - \exp_p(i_p(V_\delta))$ converging to q. For each, i, let

$$\eta_i : [0, 2] \to M$$

be a minimizing geodesic joining p to q_i and define $\gamma_0 : [0, 2] \to M$ by $\gamma_0(t) = \exp_p(t\frac{\theta}{2})$. Up to passing to a subsequence the minimizing geodesics η_i converge to a second unit speed geodesic $\gamma_1 : [0, 2] \to M$ joining p to q.

Next we argue that γ_0 and γ_1 together form a closed geodesic. If not, then either $\dot{\gamma}_0(0) \neq -\dot{\gamma}_1(0)$ or $\dot{\gamma}_0(2) \neq -\dot{\gamma}_1(2)$. We assume the latter, the former case being handled symmetrically. Fix a positive $\varepsilon < 1$ and choose $v \in T_q^1 M$ making obtuse angle with both $\dot{\gamma}_0(2)$ and $\dot{\gamma}_1(2)$. Note that for all sufficiently small s, the distance between the points $\gamma_i(2 - \varepsilon)$ and $\sigma(s) = \exp_q(sv)$ is less than one and in particular they are connected by a unique minimizing geodesic segment $\sigma_s^i : [0, 1] \to M$. By the first variation formula the energy $E(\sigma_s^i)$ is strictly decreasing for sufficiently small s. Fix $s_0 < \varepsilon$ positive and small enough such that $E(\sigma_{s_0}^i) < E(\sigma_0^i) = \varepsilon^2$.

For $i = 0, 1$ define broken geodesics $\alpha_i : [0, 3] \to M$ by

$$\alpha_i(t) = \begin{cases} \gamma_i\left(\frac{2-\varepsilon}{2}t\right) & \text{for } t \in [0, 2], \\ \sigma_{s_0}^i(t - 2) & \text{for } t \in [2, 3]. \end{cases}$$

The curves α_0, α_1 belong to $\mathcal{L}_3(p, \exp_q(s_0v))^\circ$ and have at most energy

$$E(\alpha_i) \leq \frac{(2 - \varepsilon)^2 + 2\varepsilon^2}{2} < 2.$$

Since $d_M(p, \exp_q(s_0v)) < 2$ is less than the injectivity radius, the points p and $\exp_p(s_0v)$ are connected by a unique geodesic segment α shorter than 2. The uniqueness of α implies that the flow lines $\tau \mapsto \phi_{\alpha_i}(\tau)$ of the flow provided by Lemma 1 and starting in α_0 and α_1 respectively converge to α with $\tau \to \infty$. We conclude that α_0 and α_1 are homotopic through piecewise geodesics with three segments having energy not more than $\frac{(2-\varepsilon)^2+2\varepsilon^2}{2}$. See Figure 2.
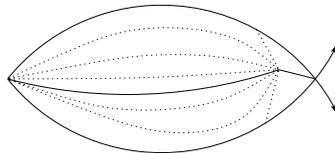

Figure 2. Flowing α_0 and α_1 to the geodesic α.

Similarly, the once broken geodesics joining $\gamma_i(2 - \varepsilon)$ to q defined by concatenating $\sigma_{s_0}^i$ with $\sigma|_{[0,s_0]}$ traversed in the opposite direction are homotopic to $\gamma_i|_{[2-\varepsilon,2]}$ through once broken geodesics of total energy not more than $2\varepsilon^2$. Combining these homotopies with those between α_0 and α_1 yields a continuous curve $\gamma \colon [0,1] \to \mathcal{L}_4(p,q)^\circ$, $s \mapsto \gamma_s$, joining γ_0 and γ_1 with $\max_{s \in [0,1]} E(\gamma_s) \leq \frac{(2-\varepsilon)^2 + 4\varepsilon^2}{2}$. By Lemma 2, there is a third geodesic $\beta \colon [0,4] \to M$ joining p to q with $E(\beta) < \frac{(2-\varepsilon)^2 + 4\varepsilon^2}{2}$. One easily checks that $L(\beta) < 4$. Therefore, β cannot intersect γ_0 or γ_1 in their interiors without contradicting $\operatorname{inj}(M) = 2$. Hence $b_M(p,q) \geq 3$, contradicting cross blocking since $d(p,q) = 2 < \operatorname{diam}(M)$.

We obtain that $\dot{\gamma}_0(0) = -\dot{\gamma}_1(0)$ and $\dot{\gamma}_0(2) = -\dot{\gamma}_1(2)$, completing the proof of Proposition 1. $\square$

Next, we prove Theorems 1 and 2.

Theorem 1. *A closed Riemannian manifold M has cross and sphere blocking if and only if M is isometric to a round sphere.*

Proof. We first scale the metric on M so that $\operatorname{inj}(M) = 2$. To begin with we claim that M is a Blaschke manifold. Otherwise there is simple closed geodesic $\gamma \subset M$ with

$$L(\gamma) = 2\operatorname{inj}(M) < 2\operatorname{diam}(M) \tag{2.1}$$

by Proposition 1. By Corollary 1, there is a geodesic segment $\eta \colon [0,1] \to M$ with end-points in γ not entirely contained in γ. Up to replacing η by a subsegment whose end-points are again in γ we can assume that the interior of η is disjoint from γ. Let x and y be the end-points of η. If $x = y$ then γ and η are two light rays from x to itself with disjoint interior. Hence one needs at least two points to block x from itself contradicting the assumption that M has sphere blocking. Assume now that $x \neq y$. Then η and the two subsegments of γ connecting x and y are three light rays with disjoint interior. This implies that x and y have blocking number $b_M(x,y) \geq 3$. Since M is assumed to have cross blocking we obtain that x and y are at distance $\operatorname{diam}(M)$ and hence γ has at least length $2\operatorname{diam}(M)$ contradicting (2.1).

We have proved that M is Blaschke. As mentioned in the introduction, Theorem 1 follows now from [LS07, Corollary 3.7] where it was shown that Blaschke manifolds with sphere blocking are isometric to round spheres. $\square$

Theorem 2. *A closed Riemannian surface M has cross blocking if and only if M is isometric to a constant curvature sphere or projective plane.*

Proof. Assume that M is a closed Riemannian surface with cross blocking that does not have constant positive curvature. By [Be78], M is not Blaschke. By Proposition 1, there is a simple closed geodesic $\gamma \subset M$ of length $2\operatorname{inj}(M) < 2\operatorname{diam}(M)$.

We first claim that γ must generate $\pi_1(M)$. To see this, fix $p \in \gamma$ and an essential map $f : ([0, 1], \{0, 1\}) \rightarrow (M, p)$ representing an element in $\pi_1(M, p)$ not in the subgroup generated by γ. Let $\varepsilon > 0$ be small, choose a point $q \in \gamma \cap B(p, \varepsilon)$ different from p, and let γ' denote the subsegment of γ joining p to q of length less than ε. Concatenating f with γ' yields a map $f' : [0, 1] \rightarrow M$ with $f(0) = p$, $f(1) = q$, and with the property that any other curve joining p to q homotopic to f' relative to the endpoints must have energy strictly greater than ε. A curve τ minimizing energy in this homotopy class is a geodesic segment joining p to q with image not entirely contained in γ. Up to passing to a subsegment of τ with distinct endpoints p' and q', we may assume that the interior of τ never intersects γ. But then τ, and the two subsegments of γ joining p' to q' are three light rays with distinct interiors. As M is assumed to be cross blocked, $d(p', q') = \text{diam}(M)$, a contradiction, completing the proof that γ generates $\pi_1 M$. In particular M is diffeomorphic to either S^2 or $\mathbb{RP}^2$.

Next, assume that M is diffeomorphic to S^2. Then γ bounds a Riemannian 2-disc $\mathbb{D} \subset S^2$. By [HS94], there is a geodesic segment $\tau : [0, 1] \rightarrow \mathbb{D}$ making right angles at both ends with γ. In particular, the endpoints of τ are distinct. Hence, τ and the two subsegments of γ joining $\tau(0)$ to $\tau(1)$ are three light rays between these points with distinct interiors. Again, as M is cross blocked, the distance between these endpoints is $\text{diam}(M)$, a contradiction.

Thus, M is diffeomorphic to $\mathbb{RP}^2$. Lift γ to a closed geodesic $\tilde{\gamma} \subset S^2$. Let $A : S^2 \rightarrow S^2$ be the order two covering transformation corresponding to γ and for $p \in \tilde{\gamma}$, let $p' = A(p)$. We shall say that such a pair of points $p, p' \in \gamma$ are an antipodal pair. Note that the same reasoning as in the above two paragraphs shows that any geodesic segment $\tau : [0, 1] \rightarrow S^2$ with endpoints in $\tilde{\gamma}$ and interior disjoint from $\tilde{\gamma}$ satisfies $\tau(0) = \tau(1)$ or $\tau(1) = \tau(0)'$. It follows easily that $\tilde{\gamma}$ has no transversal self-intersections and that each pair of subsegments of $\tilde{\gamma}$ joining antipodal pairs are minimizing. Fix a hemisphere Σ bounded by $\tilde{\gamma}$. We will next prove that Σ is isometric to a constant curvature hemisphere, contradicting the assumption that M does not have constant curvature, and completing the proof of the theorem.

For $p \in \tilde{\gamma}$, let $T_p^+(\tilde{\gamma}) \subset T_p S^2$ denote the set of unit tangent vectors based at p either tangent to $\tilde{\gamma}$ or pointing into the hemisphere Σ.

By [HS94], there is a constant speed paramaterized geodesic segment $\tau : [0, 1] \rightarrow \Sigma$ making right angles with $\tilde{\gamma}$ at both endpoints. Let $p = \tau(0)$ and $L = \text{length } \tau$. Then by the above remarks, $\tau(1) = p'$. Note that p' is conjugate to p along τ for otherwise there are geodesic segments arbitrarily close to τ joining p to a point $p'' \in \tilde{\gamma}$ distinct from but arbitrarily close to p'. Let $\text{Conj}(p) \subset T_p S^2$ denote the tangential conjugate locus to p and C the component containing $\dot{\tau}(0)$. By work of Warner in [Wa65], C is a smooth 1-submanifold of $T_p S^2$ transverse to the radial directions.

Let $v := \frac{\dot{\tau}(0)}{\|\dot{\tau}(0)\|} \in T_p^+(\tilde{\gamma})$ and note that for all vectors v' sufficiently close to v in $T_p^+(\tilde{\gamma})$, the geodesic ray $\tau_{v'}(t) = \exp_p(tv')$ crosses $\tilde{\gamma}$ in a small neighborhood of p'

at time close to L. By the above remarks, the point of intersection of each such ray must be p' and the antipodal pair p and p' are conjugate along each such ray. Let $U \subset T_p^+(\tilde{\gamma})$ be the largest open interval around v with the property that each ray in a direction through U first leaves Σ through the point p'. Note that the times the rays in directions from U leave Σ through p' vary smoothly with $v' \in U$. This follows since p and p' are conjugate along each such ray and since C is a smooth curve. By the first variation formula, they actually all leave at time exactly L. It follows that U is closed and hence that $U = T_p^+(\tilde{\gamma})$ and $L = 2\operatorname{inj}(M) = \operatorname{length} \tilde{\gamma}/2$.

It now follows that for each $q \in \tilde{\gamma}$ sufficiently close to p, there is a geodesic ray entering Σ from q and leaving Σ at a point in $\tilde{\gamma}$ close to p'. By repeating the argument in the last paragraph, it follows that for q sufficiently close to p, every geodesic entering Σ at q first exits Σ at its antipodal point q' at time exactly L. Let U' denote the largest open interval around p in $\tilde{\gamma}$ with the property that every ray entering Σ from a point in U exits Σ at its antipode at time L. Then U' is clearly closed, whence $U' = \tilde{\gamma}$. By [Ba83], Σ is a round hemisphere, completing the proof.

$\square$

References

[Ba83] V. Bangert, Manifolds with geodesic chords of constant length. *Math. Ann.* **265** (1983), 273–281. Zbl 0508.53055 MR 0721397 270

[BaGu] V. Bangert and E. Gutkin, Insecurity for compact surfaces of positive genus. *Geom. Dedicata*, to appear; doi:10.1007/s10711-009-9432-8. 260

[Be78] M. Berger, Blaschke's conjecture for spheres. Appendix D in *Manifolds all of whose geodesics are closed* by A. L. Besse, Ergeb. Math. Grenzgeb. 93, Springer-Verlag, Berlin 1978. Zbl 0387.53010 MR 0496885 260, 268

[BG] K. Burns and E. Gutkin, Growth of the number of geodesics between points and insecurity for Riemannian manifolds. *Discrete Contin. Dyn. Syst.* **21** (2008), no. 2, 403–413. Zbl 05304704 MR 2385699 260

[CE75] J. Cheeger and D. Ebin, *Comparison theorems in Riemannian geometry*. North-Holland Math. Lib. 9, Amsterdam 1975. Zbl 0309.53035 MR 0458335 261, 264

[Fo90] D. Fomin, *Zadaqi Leningradskih Matematitcheskih Olimpiad*. Leningrad, 1990. 260

[Gu05] E. Gutkin, Blocking of billiard orbits and security for polygons and flat surfaces. *Geom. Funct. Anal.* **15** (2005), 83–105. Zbl 1134.37334 MR 2140629 260

[Gu06] E. Gutkin, Insecure configurations in lattice translation surfaces, with applications to polygonal billiards. *Discrete Contin. Dyn. Syst.* **16** (2006), no. 2, 367–382. Zbl 1115.37046 MR 2226486 260

[GuA] E. Gutkin, Blocking of orbits and the phenomenon of (in)security for the billiard in polygons and flat surfaces. Preprint, IHES/M/03/06. 260

[GB]   E. Gutkin, Topological entropy and blocking cost for geodesics in Riemannian mani-folds. *Geom. Dedicata* **138** (2009), 13–23. Zbl 05549607 MR 2469985 260

[GS06] E. Gutkin and V. Schroeder, Connecting geodesics and security of configura-tions in compact locally symmetric spaces. *Geom. Dedicata* **118** (2006), 185–208. Zbl 1103.53027 MR 2239456 260

[HS94] J. Hass and P. Scott, Shortening curves on surfaces. *Topology* **33** (1994), no. 1, 25–43. Zbl 0798.58019 MR 1259513 269

[HS98] P. Hiemer and V. Snurnikov, Polygonal billiards with small obstacles. *J. Statist. Phys.* **90** (1998), 453–466. Zbl 0995.37019 MR 1611100 260

[He]   P. Herreros, Blocking: new examples and properties of products. *Ergodic Theory Dynam. Systems* **29** (2009), no. 2, 569–578. Zbl 1160.37339 MR 2486784 260

[LS07] J.-F. Lafont and B. Schmidt, Blocking light in compact Riemannian manifolds. *Geom. Topol.* **11** (2007), 867–887. Zbl 1135.53028 MR 2326937 260, 261, 268

[Mi63] J. Milnor, *Morse theory*. Princeton University Press, Princeton, NJ, 1963. Zbl 0108.10401 MR 0163331 261

[Mo04] T. Monteil, A counter-example to the theorem of Hiemer and Snurnikov. *J. Statist. Phys.* **114** (2004), 1619–1623. Zbl 1058.37040 MR 2039490 260

[Mo05] T. Monteil, On the finite blocking property. *Ann. Inst. Fourier* **55** (2005), 1195–1217. Zbl 1076.37029 MR 2157167 260

[MoA]   T. Monteil, Finite blocking property versus pure periodicity. *Ergodic Theory Dynam. Systems* **29** (2009), no. 3, 983–996. Zbl 05580205 MR 2505325 260

[MoB]   T. Monteil, A homological condition for a dynamical and illuminatory classification of torus branched coverings. Preprint, arXiv:math/0603352v1 [math.DS]. 260

[Se51] J.P. Serre, Homologie singuliere des éspaces fibrés. *Ann. of Math.* **54** (1951), 425–505. Zbl 0045.26003 MR 0045386 259

[Ta]   S. Tabachnikov, Birkhoff billiards are insecure. *Discrete Contin. Dyn. Syst.* **23** (2009), no. 3, 1035–1040. Zbl 1160.37025 MR 2461838 260

[Wa65] F. Warner. The conjugate locus of a Riemannian manifold. *Amer. J. Math.* **87** (1965), no. 3, 575–604. Zbl 0129.36002 MR 0208534 269

Benjamin Schmidt, Department of Mathematics, Wells Hall, Michigan State University, East Lansing, MI 48824, U.S.A.

E-mail: schmidt@math.msu.edu

Juan Souto, Department of Mathematics, University of Michigan, 2074 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, U.S.A.

E-mail: jsouto@umich.edu