GEODESICS IN RANDERS SPACES
OF CONSTANT CURVATURE

COLLEEN ROBLES

ABSTRACT. Geodesics in Randers spaces of constant curvature are classified.

1. INTRODUCTION

Randers metrics have received much attention lately as solutions to Zermelo's problem of navigation; largely because this navigation structure provides the framework for a complete classification of constant flag curvature Randers spaces. (*Flag curvature* is the Finslerian analog of Riemannian sectional curvature. See [BR04].) Briefly, a Randers metric is of constant flag curvature if and only if it solves Zermelo's problem of navigation on a Riemannian manifold of constant sectional curvature under the influence of an infinitesimal homothety W. See Subsection 1.1 for a sketch of the navigation problem, and Theorem 3 for an explicit statement of the classification result.

The aim of this paper is to develop a geometric description of the geodesics in these spaces of constant curvature. Intuitively, these paths minimize travel time across a Riemannian landscape under windy conditions. Presently we will show that these curves are given by composing geodesics of the Riemannian metric with the flow generated by W. This claim is formalized by Theorem 2.

Geodesics on surfaces of constant, nonpositive curvature are illustrated in Section 3. We then turn, in Section 4, to the constant flag curvature $K = 1$ Randers metrics on S^n. The case of the sphere is especially interesting; it is possible to endow this closed manifold with a metric whose geodesics display distinctly non-Riemannian behaviors. For example:

(1) A metric is *projectively flat* if every point admits coordinates in which the geodesics are straight lines. Beltrami's theorem assures us that *a Riemannian metric is of constant sectional curvature if and only if it is projectively flat.* In contrast few Randers spaces of constant flag curvature are projectively flat. There are infinitely many nonisometric Randers metrics of constant, *positive* curvature. Only the trivial (Riemannian) metrics are projectively flat. See Section 7 of [BRS04] for a thorough discussion.

(2) On the Riemannian n-sphere all geodesics close with length 2π. In contradistinction, it is possible to equip the sphere with a non-Riemannian Randers metric of constant curvature $K = 1$ for which either (i) all the geodesics close,

or (ii) only finitely many of the geodesics close! In either case, the geodesics will not all have the same length. This is a consequence of the fact that non-Riemannian Randers metrics are not *reversible* (a.k.a. *symmetric*): in general, the Randers length $F(V)$ of a vector V is not equal to the length $F(-V)$ of $-V$.

Studies of conjugate and cut points occupy the last two sections of the paper. Section 5 establishes a correspondence between points conjugate to p with respect to the Riemannian metric, and points conjugate to p with respect to the Randers metric solving the navigation problem on the Riemannian manifold. The details are spelled out in Proposition 6.

The final section of the paper investigates minimizing geodesics. We will see that a Randers geodesic is a global minimizer if and only if its associated Riemannian geodesic is as well. See Proposition 8 for a precise statement.

The remainder of the Introduction is given to a brief discussion of Randers metrics as solutions to Zermelo's problem of navigation. This material has been carefully presented elsewhere (references are provided), so we will present only a cursory introduction with the principle goal of establishing notation. To begin,

- Points on the n-dimensional manifold M are denoted by p or x.
- Tangent vectors are given by $y \in T_p M$, with components $y^i \partial_{x^i}$ relative to local coordinates $x = (x^i)$ on M.
- Partial derivatives are denoted by the subscripts x^i and y^i.

1.1. Randers metrics and Zermelo navigation.

In 1941 G. Randers [R41] introduced a Finsler metric by modifying a Riemannian metric $a := a_{ij}(x)\, dx^i \otimes dx^j$ by a linear term $b := b_i(x)\, dx^i$. The resulting Minkowski norm on $T_x M$ is given by

$$F(x, y) := \alpha(x, y) + \beta(x, y) = \sqrt{a_{ij}(x)y^i y^j} + b_i(x)y^i, \quad y = y^i \partial_{x^i} \in T_x M.$$

By requiring $a(b, b) < 1$, we ensure that F is positive. This simple condition also guarantees that the metric is strongly convex. That is, the Hessian $g_{ij}(x, y) := (\frac{1}{2}F^2)_{y^i y^j}$ is positive definite for all nonzero y. See [BCS00, BR04] for this result and a through treatment of Randers metrics.

Z. Shen [S02] has identified Randers metrics with solutions to a navigation problem. In 1931, Zermelo posed and answered the following question [Z31, C99]: Suppose a ship sails the sea on calm waters. Imagine a mild breeze comes up. How must the captain guide the ship to reach a given destination in the shortest time?

Zermelo assumes that the sea is $\mathbb{R}^2$, with the flat/Euclidean metric. Shen extended the seascape to an arbitrary Riemannian manifold $(\mathcal{M}, h)$. Assuming a time-independent wind, he found that the paths minimizing travel-time are exactly the geodesics of a Randers metric

$$(1) \qquad F(x, y) = \alpha(x, y) + \beta(x, y) = \frac{\sqrt{\lambda |y|^2 + W_0{}^2}}{\lambda} - \frac{W_0}{\lambda}.$$

Here $W = W^i \partial_{x^i}$ is the velocity vector field of the wind,

$$|y|^2 = h(y, y), \quad \lambda = 1 - |W|^2 \quad \text{and} \quad W_0 = h(W, y).$$

We say F *solves Zermelo's problem of navigation*. The defining Riemannian metric a and 1-form b of the Randers metric are

$$a_{ij} = \frac{\lambda h_{ij} + W_i W_j}{\lambda^2} \quad \text{and} \quad b_i = -\frac{W_i}{\lambda}, \quad \text{where } W_i := h_{ij}W^j.$$

Requiring $|W| < 1$ ensures that this a_{ij} is indeed a positive definite Riemannian metric. Additionally, $h(W, W) = a(b, b)$. Therefore, the condition $|W| < 1$ also ensures the positivity of F. Consequently F is defined on the open sub-manifold $\{|W| < 1\} \subset \mathcal{M}$. A straightforward computation establishes

Lemma 1. $F(y) = 1 \iff |y - W| = 1$. *Geometrically, this means the unit sphere of F in $T_x M$ differs from the unit sphere of h by a translation along $W(x)$.*

At this point it is natural to ask if *every* Randers space (M, F) arises as the solution to Zermelo's problem of navigation for a canonical choice of (h, W). The answer is yes [BR04]. As a result, *Randers metrics are naturally identified with solutions to the navigation problem.*

1.2. **The result.** It is now possible to state the main result of the paper. The notation $\mathcal{L}$ below denotes Lie differentiation.

Theorem 2. *Assume $(\mathcal{M}, h)$ is a Riemannian manifold equipped with an infinitesimal homothety W; $\mathcal{L}_W h = \sigma h$, σ a constant. Let F denote the Randers metric solving Zermelo's problem of navigation on $M = \{|W| < 1\} \subset \mathcal{M}$. Then the unit speed geodesics $\mathcal{P} : (-\varepsilon, \varepsilon) \to M$ of F are given by $\mathcal{P}(t) = \varphi(t, \rho(t))$, where*

- ∘ *$\rho : (-\varepsilon, \varepsilon) \to \mathcal{M}$ is a geodesic of h parameterized so that $|\dot{\rho}(t)|^2 = e^{-\sigma t}$,*
- ∘ *shrinking ε if necessary, $\varphi : (-\varepsilon, \varepsilon) \times U \to M$ is the flow of W defined on a neighborhood U of $\rho(0)$ so that $\rho(t) \in U$, for all $t \in (-\varepsilon, \varepsilon)$.*

In the abstract we promised to classify the geodesics in Randers spaces of constant flag curvature. The theorem does just that. This is because a Randers metric has constant flag curvature if and only if it solves the navigation problem on a Riemannian space of constant curvature under the influence of an infinitesimal homothety, and the geodesic structure of Riemannian space forms is standard knowledge.

Theorem 3 (Constant flag curvature classification [BRS04])**.** *Let F be a Randers metric on a manifold M with navigation data (h, W), $|W| < 1$. Then F is of constant flag curvature K if and only if*

(1) *The Riemannian space (M, h) is of constant sectional curvature $K + \frac{1}{16}\sigma^2$.*
(2) *The vector field W is an infinitesimal homothety of h, $\mathcal{L}_W h = \sigma h$.*

The constant σ vanishes when h is not flat, and W is an infinitesimal isometry.

Note that Theorem 2 allows an arbitrary Riemannian metric h. As a result it describes the geodesics of a larger class of Randers metrics than simply those of constant flag curvature. In this context the theorem may only be considered a 'description,' not a 'classification,' since the geodesic structure of these more general Riemannian metrics is typically not known.

1.3. **Geodesic equations.** A curve $\rho : (-\varepsilon, \varepsilon) \to \mathcal{M}$ is a geodesic of h if it satisfies the *geodesic equation*

$$(2) \qquad \ddot{\rho}^i + 2\,\mathcal{G}^i(\rho, \dot{\rho}) = \frac{d}{dt}\left(\ln |\dot{\rho}|\right) \dot{\rho}^i \,.$$

Note that the right-hand side is zero when ρ is parameterized with constant speed. The geodesic spray coefficients $\mathcal{G}^i$ are given by

$$\mathcal{G}^i(x, y) = \tfrac{1}{2}\gamma^i{}_{jk}(x)y^j y^k \,,$$

where $\gamma^i{}_{jk} = \frac{1}{2}h^{is}(h_{sj,x^k} - h_{jk,x^s} + h_{ks,x^j})$ denotes the Christoffel symbols of h.

Similarly, a curve $\mathcal{P} : (-\varepsilon, \varepsilon) \to M$ will be a geodesic of the Randers metric F if it satisfies the *geodesic equation*

$$\ddot{\mathcal{P}}^i + 2G^i(\mathcal{P}, \dot{\mathcal{P}}) = \frac{d}{dt}\left(\ln F(\dot{\mathcal{P}})\right)\dot{\mathcal{P}}^i\,.$$

(See [BCS00] for a discussion of Finslerian geodesics.) The geodesic coefficients of F are related to those of the Riemannian metric a by (11.3.12) of [BCS00]. (*Beware!* Their G is *twice* ours.) Similarly, the relationship between the geodesic spray coefficients of a and h is derived on page 235 of [BR04]. As a result, the geodesic coefficients of F are related to the those of h by

$$G^i = \mathcal{G}^i + \zeta^i\,,$$

where

(3) $\quad \zeta^i = \frac{1}{4}\left(\frac{1}{F}y^i - W^i\right)\left(2F\mathcal{S}_0 - \mathcal{L}_{00} - F^2\mathcal{L}_{WW}\right) - \frac{1}{4}F^2\left(\mathcal{S}^i + \mathcal{T}^i\right) - \frac{1}{2}F\mathcal{C}^i{}_0\,.$

Above we have utilized the notation

$$\begin{array}{rclcrcl}
\mathcal{L}_{ij} &=& W_{i:j} + W_{j:i}, & \quad & \mathcal{C}_{ij} &=& W_{i:j} - W_{j:i}, \\
\mathcal{S}_i &=& W^s\mathcal{L}_{si}, & & \mathcal{T}_i &=& W^s\mathcal{C}_{si}\,.
\end{array}$$

The colon ':' denotes covariant differentiation so that $W_{i:j} = W_{i,x^j} - W_s\gamma^s{}_{ij}$. Indices on these tensors are raised with the inverse of h. For example, $\mathcal{S}^i = h^{ij}\mathcal{S}_j$. As before the subscript 0 denotes contraction with y, $\mathcal{C}^i{}_0 = h^{ij}\mathcal{C}_{jk}y^k$. Finally, $\mathcal{L}_{WW} = W^iW^j\mathcal{L}_{ij}$.

2. Proof of Theorem 2

2.1. **Preliminaries.** We are given three items. First, a Randers space (M, F) induced by Zermelo navigation on the Riemannian manifold $(\mathcal{M}, h)$ under the influence of an infinitesimal homothety W;

(4) $$\mathcal{L}_W h = \sigma h\,, \quad \sigma \text{ a constant.}$$

Second, a geodesic of h, $\rho : (-\varepsilon, \varepsilon) \to U \subset \mathcal{M}$, parameterized so that $|\dot{\rho}(t)|^2 = e^{-\sigma t}$. And third, the flow $\varphi : (-\varepsilon, \varepsilon) \times U \to M$ of W, defined on a neighborhood U of $\rho(0)$. There is no loss of generality in assuming that U admits coordinates (x^i).

With these data in hand define a curve $\mathcal{P} : (-\varepsilon, \varepsilon) \to M$ by

$$\mathcal{P}(t) = \varphi(t, \rho(t))\,.$$

The proof is complete once we establish

(i) that the curve $\mathcal{P}$ is parameterized with unit F-speed, $F(\dot{\mathcal{P}}) = 1$; and

(ii) $\mathcal{P}$ is a geodesic of F, $\ddot{\mathcal{P}} + 2G(\mathcal{P}, \dot{\mathcal{P}}) = \ddot{\mathcal{P}} + 2\mathcal{G}(\mathcal{P}, \dot{\mathcal{P}}) + 2\zeta(\mathcal{P}, \dot{\mathcal{P}}) = 0$.

Begin by computing

$$\dot{\mathcal{P}} = W + d\varphi(\dot{\rho})\,,$$

where $d\varphi$ denotes the differential of the map $\varphi(t, \cdot) : U \to M$. With respect to local coordinates (x^i) this reads $\dot{\mathcal{P}}^i = W^i + \varphi^i{}_{,x^j}\dot{\rho}^j$. Similarly,

(5) $\quad \ddot{\mathcal{P}}^i = W^i{}_{,x^j}W^j + 2W^i{}_{,x^j}\varphi^j{}_{,x^k}\dot{\rho}^k + \varphi^i{}_{,x^jx^k}\dot{\rho}^j\dot{\rho}^k + \varphi^i{}_{,x^j}\ddot{\rho}^j\,.$

2.2. **Implications of the homothety hypothesis.** First, the differential $d\varphi$ of $\varphi(t,\cdot) : U \to M$ scales vectors by a conformal factor of $e^{\sigma t/2}$; $|d\varphi\, V|^2 = e^{\sigma t}|V|^2$. In particular, $|d\varphi(\dot\rho)|^2 = e^{\sigma t}|\dot\rho(t)|^2 = 1$. Lemma 1 implies $F(\dot{\mathcal{P}}) = 1$, and (i) is established.

Second, since $\varphi(t_o, \cdot)$ is a homothety, $\rho_o(t) := \varphi(t_o, \rho(t))$ is a geodesic of h with speed $|\dot\rho_o| = e^{\sigma(t_o - t)/2}$. The geodesic equation (2) implies

$$(6) \qquad \ddot\rho_o{}^i + 2\,\mathcal{G}^i(\rho_o, \dot\rho_o) = \frac{d}{dt}\left[\tfrac{1}{2}\sigma(t_o - t)\right]\dot\rho_o{}^i.$$

Differentiating by t yields $\dot\rho_o{}^i = \varphi^i{}_{,x^j}\,\dot\rho^j$ and $\ddot\rho_o{}^i = \varphi^i{}_{,x^j x^k}\,\dot\rho^j\dot\rho^k + \varphi^i{}_{,x^j}\,\ddot\rho^j$, where the partial derivatives of φ are evaluated at $(t_o, \rho(t))$. Observe that $\rho_o(t_o) = \mathcal{P}(t_o)$ and $\dot\rho_o(t_o) = \dot{\mathcal{P}}(t_o) - W_{\mathcal{P}(t_o)}$. Consequently, at $t = t_o$, (6) reads

$$\varphi^i{}_{,x^j x^k}\,\dot\rho^j\dot\rho^k + \varphi^i{}_{,x^j}\,\ddot\rho^j + 2\,\mathcal{G}^i(\mathcal{P}, \dot{\mathcal{P}} - W) = -\tfrac{1}{2}\sigma(\dot{\mathcal{P}}^i - W^i).$$

Since t_o is arbitrary, this expression must hold for all t. Making use of the equality $\frac{1}{2}(\mathcal{L}^i{}_j + \mathcal{C}^i{}_j) = W^i{}_{:j} = W_{i,x^j} + W^i\,\gamma^i{}_{jk}$, and the fact that $\mathcal{G}^i(x, y) = \frac{1}{2}\,\gamma^i{}_{jk}(x)y^jy^k$ is quadratic in y, this updates the formula (5) for $\ddot{\mathcal{P}}$ to

$$\begin{aligned}
\ddot{\mathcal{P}}^i &= -2\,\mathcal{G}^i(\mathcal{P}, \dot{\mathcal{P}}) + \tfrac{1}{2}\sigma(W^i - \dot{\mathcal{P}}^i) + \mathcal{L}^i{}_0 + \mathcal{C}^i{}_0 - \tfrac{1}{2}\mathcal{S}^i + \tfrac{1}{2}\mathcal{T}^i \\
&= -2\,\mathcal{G}^i(\mathcal{P}, \dot{\mathcal{P}}) + \tfrac{1}{2}\sigma\dot{\mathcal{P}}^i + \tfrac{1}{2}\mathcal{T}^i + \mathcal{C}^i{}_0.
\end{aligned}$$

In the second equality we have made use of the relationships

$$\mathcal{L}^i{}_0 = \sigma\dot{\mathcal{P}} \quad\text{and}\quad \mathcal{S}^i = \sigma W^i,$$

both consequences of the facts that W is an infinitesimal homothety, (4), and that the components of the Lie derivative are $(\mathcal{L}_W h)_{ij} = \mathcal{L}_{ij}$.

2.3. **Conclusion.** Finally, we turn to the expression (3) for ζ in Subsection 1.3. Note that (4) reduces the term $(2F\mathcal{S}_0 - \mathcal{L}_{00} - F^2\mathcal{L}_{WW})$ to $-\sigma|\dot{\mathcal{P}} - W|^2$, which is equal to $-\sigma$ by (i) and Lemma 1. Hence,

$$\zeta^i(\mathcal{P}, \dot{\mathcal{P}}) = -\tfrac{1}{4}\sigma\dot{\mathcal{P}}^i - \tfrac{1}{4}\mathcal{T}^i - \tfrac{1}{2}\mathcal{C}^i{}_0.$$

A quick comparison with the final form of $\ddot{\mathcal{P}}$ above yields $\ddot{\mathcal{P}} = -2\,\mathcal{G} - 2\zeta = -2G$. Hence (ii) holds, and $\mathcal{P}$ is a geodesic of F. $\qquad\square$

2.4. **The general case.** Theorem 2 does not hold for arbitrary W. Consider, as a counterexample, the vector field $W = \varepsilon\partial_\phi$ ($0 < \varepsilon < 1$) defined on the dense open subset of S^2 covered by the coordinate map $(\theta, \phi) \mapsto (\sin\theta\sin\phi, \cos\theta\sin\phi, \cos\phi)$. Note that W is not an infinitesimal isometry of standard Riemannian metric h on the sphere, and there is no parameterization $\rho(t)$ of the equator $\{\phi = \frac{\pi}{2}\}$ (a geodesic of h) with the property that $\mathcal{P}(t) = \varphi(t, \rho(t))$ is a unit speed geodesic of the associated Randers metric.

3. Photo gallery for nonpositive curvature

This section contains a selection of graphics illustrating solutions to Zermelo's problem of navigation (i.e. geodesics of F) on subsets of Euclidean plane and the Poincaré disc. The subset M on which F is defined will be determined by the constraint $|W| < 1$. These metrics are of constant, nonpositive curvature.

3.1. **In the Euclidean plane.** We begin with the Euclidean plane $(\mathbb{R}^2, h)$. The infinitesimal homotheties are given by

$$W(u,v) = \begin{pmatrix} \frac{1}{2}\sigma u + kv + c_1 \\ \frac{1}{2}\sigma v - ku + c_2 \end{pmatrix}, \quad (u,v) \in \mathbb{R}^2\,.$$

Here k, c_1, $c_2 \in \mathbb{R}$ are constants. Theorem 3 tells us the Randers metric F solving Zermelo's problem of navigation for (h, W) is of constant flag curvature $K = -\frac{1}{16}\sigma^2 \leq 0$.

If $k = 0$, then F is projectively flat (see Section 7 of [BRS04]). These spaces admit coordinate systems in which the geodesics are straight lines. In the examples below we will restrict our attention to examples that are not projectively flat, i.e. infinitesimal homotheties with $k \neq 0$.

Case I. W an infinitesimal rotation. Set $\sigma = 0$. Then the Randers metric has constant flag curvature $K = 0$. Let us suppose $c_i = 0$ for simplicity. The infinitesimal rotation $W(u,v) = (v, -u)$ has norm less than 1 on the unit disc $M = \{(u,v) \in \mathbb{R}^2 \ : \ u^2 + v^2 < 1\}$.
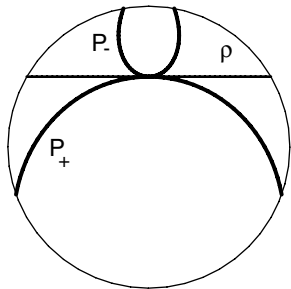
The geodesic characterization of Theorem 2 tells us to parameterize the Euclidean line ρ with unit speed. Take $\rho(t) = (t + u_o, v_o)$. Since $\rho(-t)$ is also a unit speed geodesic of h, both $\mathcal{P}_+ = \varphi(t, \rho(t))$ and $\mathcal{P}_-(t) = \varphi(t, \rho(-t))$ are geodesics of the Randers metric F.
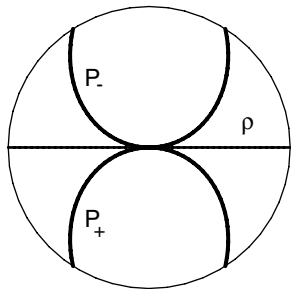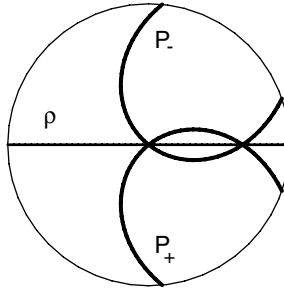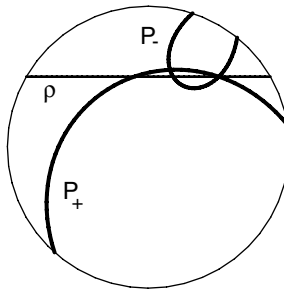

EXAMPLE 1. $u_o = 0 = v_o$ EXAMPLE 2. $u_o = \frac{2}{3}$, $v_o = 0$


EXAMPLE 3. $u_o = 0, v_o = \frac{1}{2}$ EXAMPLE 4. $u_o = \frac{1}{2} = v_o$

Note that the two F-geodesics $\mathcal{P}_+$ and $\mathcal{P}_-$ in Example 1 are tangent at the center of the disc. Their initial vectors at this point are $\dot{\mathcal{P}}_+(0) = (1,0)$ and $\dot{\mathcal{P}}_-(0) =$

$(-1,0)$. In Riemannian geometry two geodesics passing through a common point in opposite directions necessarily trace the same curve. In fact, all reversible Finsler metrics have this property. As Examples 1 and 3 indicate, the phenomenon does not extend to the nonreversible setting. In general, if $\mathcal{P}(t)$ is a geodesic, $\mathcal{P}(-t)$ will not be a geodesic.

Case II. W an infinitesimal homothety. When σ is nonzero the metric F has negative flag curvature $K = -\frac{1}{16}\sigma^2$. In this case a suitably chosen translation effects a change of coordinates for which $c_i = 0$. The infinitesimal homotheties are of the form $W(u,v) = -\frac{1}{2}\sigma(u,v) + k(v,-u)$. According to Theorem 2 we must parameterize the line ρ so that $|\rho(t)|^2 = e^{-\sigma t}$. To that end take $\rho_+(t) = [\frac{2}{\sigma}(e^{-\sigma t/2} - 1) + u_o, v_o]$. The curve $\rho_-(t) = [\frac{2}{\sigma}(1 - e^{-\sigma t/2}) + u_o, v_o]$ traverses the same path, again with Euclidean speed $e^{-\sigma t/2}$, but in the opposite direction. Hence, $\mathcal{P}_+(t) = \varphi(t, \rho_+(t))$ and $\mathcal{P}_-(t) = \varphi(t, \rho_-(t))$ are F-geodesics, both generated from the same Euclidean line.

For $\sigma = \sqrt{2}$ and $k = 1/\sqrt{2}$ we have $|W| < 1$ when $u^2 + v^2 < 1$. Hence $M \subset \mathbb{R}^2$ is again the unit disc.
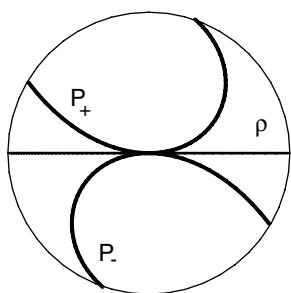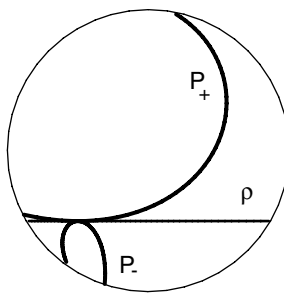

EXAMPLE 5. $u_o = 0 = v_o$ \qquad EXAMPLE 6. $u_o = -\frac{1}{2} = v_o$

3.2. **On the Poincaré disc.** Next let us turn our attention to the Poincaré model (D^2, h) of hyperbolic geometry on the unit disc with constant sectional curvature -1. In this case σ must vanish; the infinitesimal homotheties W are necessarily infinitesimal isometries, and the resulting Randers metric F is of constant flag curvature $K = -1$ (Theorem 3). The metric F will be projectively flat if and only if F is Riemannian ($W = 0$) [BRS04].

Since $\sigma = 0$ we parameterize the geodesics of h, arcs of circles intersecting ∂D orthogonally, with unit hyperbolic speed. If $\rho(t)$ is one such parameterization, then both $\mathcal{P}_+(t) = \varphi(t, \rho(t))$ and $\mathcal{P}_-(t) = \varphi(t, \rho(-t))$ are geodesics of F. In Examples 7 and 9, $\rho(t) = (0, \tanh(t/2))$. The geodesic ρ of Examples 8 and 10 is a hyperbolic translation of this vertical line $(0, \tanh(t/2))$ along the horizontal u-axis.

Case III. W is an infinitesimal rotation. In Examples 7 and 8 we take W to be an infinitesimal rotation about the origin, $W(u,v) = \frac{1}{2}(v,-u)$. The resulting Randers metric is defined on M as the disc of radius $(\sqrt{5} - 1)/2$, determined by the condition $|W| < 1$.

Case IV. W is an infinitesimal hyperbolic translation. Here the flow $\varphi(t, \cdot) : \mathbb{D} \to \mathbb{D}$ is a hyperbolic translation fixing the horizontal u-axis. The corresponding
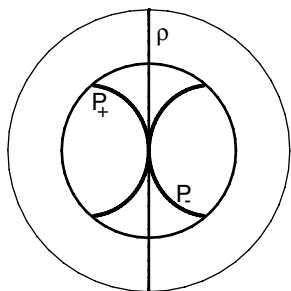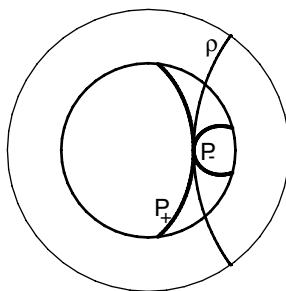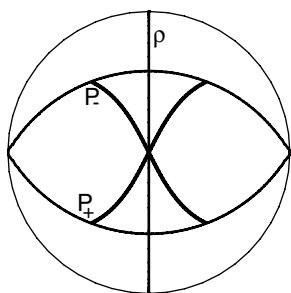
EXAMPLE 7.   EXAMPLE 8.
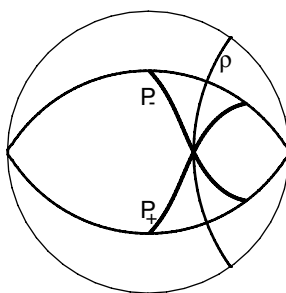
vector field W has norm less than one inside the smaller, eye-shaped region M indicated in Examples 9 and 10.

EXAMPLE 9.   EXAMPLE 10.

3.3. When (M, h) is complete.
In each of the examples above (M, h) fails to be complete.

Proposition 4. *Suppose is (M, F) of constant nonpositive flag curvature, and (M, h) is complete. Then F is either locally Minkowski, or a negatively curved Riemannian space form.*

Remark. Both these spaces are projectively flat.

Proof. Theorem 3 assures us that the Riemannian space (M, h) must be of constant sectional curvature $\kappa \leq 0$. Since (M, h) is complete it admits a universal cover $(\tilde{M}, \tilde{h})$ by Euclidean n-space $\mathbb{R}^n$ (when $\kappa = 0$) or the Poincaré hyperbolic ball D^n (when $\kappa < 0$). In both cases the covering map is a local isometry and W lifts to a infinitesimal homothety $\tilde{W}$ of $(\tilde{M}, \tilde{h})$. Since $|\tilde{W}| < 1$, $(\tilde{h}, \tilde{W})$ determines a globally defined Randers metric $\tilde{F}$ on $\tilde{M}$. The projection map is a local isometry of $(\tilde{M}, \tilde{F}) \to (M, F)$; see Lemma 2 of [BRS04].

The only infinitesimal homotheties of Euclidean space satisfying $|\tilde{W}| < 1$ globally are the translations. Hence $(\mathbb{R}^n, \tilde{F})$, and therefore (M, F), is locally Minkowski. (See [BCS00] for the definition of locally Minkowski metrics, and [BRS04] for examples of locally Minkowski Randers metrics.)

If $\kappa < 0$, then the universal cover is the Poincaré ball and σ must vanish. So W and $\tilde{W}$ are infinitesimal isometries. The length of any nonzero, globally defined

infinitesimal isometry is unbounded on the ball, whence the condition $|\tilde{W}| = |W| < 1$ forces $\tilde{W} = 0$. As a consequence $W = 0$, and $F = h$ is Riemannian. $\qquad\square$

4. S^n AND THE KATOK EXAMPLES

According to Theorem 3 the Randers metrics of constant flag curvature $K = 1$ on the sphere arise as solutions to Zermelo's problem of navigation on the canonical Riemannian sphere (S^n, h), under the influence of an infinitesimal isometry W. These metrics are projectively flat only in the case that W is identically zero [BRS04]; that is, $F = h$ is Riemannian. (Projectively flat Finsler metrics of constant flag curvature on the 2-sphere are analyzed in [Br97]. That paper is a sequel to [Br96], Bryant's analysis of arbitrary (not necessarily Randers) constant flag curvature metrics on S^2.)

Modulo an orthogonal transformation of $\mathbb{R}^{n+1}$, the vector field W is expressed by matrix multiplication $W(p) = p\,\Omega$, where $p = (p^0, \ldots, p^n) \in S^n \subset \mathbb{R}^{n+1}$, and Ω is an $(n+1) \times (n+1)$ constant matrix of the form

$$\begin{aligned} \Omega = a_1 J \oplus \cdots \oplus a_m J \oplus 0 &\quad \text{when } n = 2m \text{ is even,} \\ \Omega = a_1 J \oplus \cdots \oplus a_m J &\quad \text{when } n = 2m - 1 \text{ is odd.} \end{aligned}$$

The a_i are ordered, $a_1 \geq \cdots \geq a_m \geq 0$, and J denotes the 2×2 matrix

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Note that the globally defined W will have norm less than 1 if and only if $a_1 < 1$. Let us assume this is the case, so that $M = S^n$.

The flow of W is similarly given by $\varphi(t, p) = p\,\mathrm{Rot}(a_1 t, \ldots, a_m t)$, where $\mathrm{Rot}(\cdots)$ is the block diagonal matrix

$$\begin{aligned} \mathrm{Rot}(a_1 t, \ldots, a_m t) = R(a_1 t) \oplus \cdots \oplus R(a_m t) \oplus 1, &\quad \text{when } n = 2m \text{ is even,} \\ \mathrm{Rot}(a_1 t, \ldots, a_m t) = R(a_1 t) \oplus \cdots \oplus R(a_m t), &\quad \text{when } n = 2m - 1 \text{ is odd.} \end{aligned}$$

Here,

$$R(a_i t) = \begin{pmatrix} \cos a_i t & \sin a_i t \\ -\sin a_i t & \cos a_i t \end{pmatrix}.$$

The Randers metrics generated by the Euclidean (S^n, h) and the infinitesimal isometries W via Zermelo navigation were initially introduced by Katok [K73], and later studied by Ziller [Zi82], in the context of Hamiltonian systems. With a little thought, the following three observations may be made.

First, if each a_i is rational, then all geodesics of the Randers metric close.

Second, given one such Randers metric F, Theorem 3 implies it is of constant flag curvature $K = 1$. Hence, given any fixed point $p \in S^n$, the F-distance from p satisfies $d_F(p, q) \leq \pi$. (This is the Bonnet-Myers theorem for Finsler metrics. See [A55, BCS00].) Moreover, equality holds for a unique q (cf. Theorem 0.1 of [S96]). In the case that F is Riemannian (equivalently, $W = 0$) this unique q is the antipodal point $-p$. In general, $q = \varphi(\pi, -p)$. In analogy with the Riemannian case, q is the unique point conjugate to p with respect to the metric F (Section 5).

Last, denote by $\mathcal{S}_r^h(p) = \{q \in S^n \mid d_h(p, q) = r\}$ the *geodesic sphere* of radius r about p. Then $\mathcal{S}_r^h(p)$ is a Euclidean $(n-1)$-sphere of radius $\sin(r)$. The Randers geodesic sphere is simply a rotation of $\mathcal{S}_r^h(p)$. Explicitly, $\mathcal{S}_r^F(p) = \varphi(r, \mathcal{S}_r^h(p))$.

Suppose $\rho(t)$ is a geodesic of S^n, $n = 2m$ or $n = 2m - 1$, invariant under the flow $\varphi(t, \cdot)$. These include the m geodesics parameterizing the great circles $C_i = \{p \in S^n \mid p^j = 0, \ \forall j \neq 2i - 1, 2i\}$, $i = 1, \ldots, m$. The resulting F-geodesic $\mathcal{P}$

is simply a reparameterization of C_i. When $\rho(t)$ parameterizes C_i in the direction of the rotation $R(a_i t)$, $\mathcal{P}$ has F-length $2\pi/(1+a_i)$. Otherwise $\mathcal{P}$ has length $2\pi/(1-a_i)$. These two geodesics are considered distinct. In particular, there exist at least $2m$ closed geodesics of F on S^n, $n = 2m-1$, $2m$.

These geodesics are simple. In general, (S^n, F) will have geodesics that self-intersect. (This assumes of course that F is non-Riemannian: $W \neq 0$.) For example, suppose p is a fixed point of the flow φ. In the case that $n = 2m$ is even, the north and south poles $(0, \ldots, 0, \pm 1)$ are always fixed points. Let $\rho : (-\infty, \infty) \to S^{2m}$ be the unique h-geodesic passing through the north pole, $\rho(2\pi\ell) = (0, \ldots, 0, 1)$, $\ell \in \mathbb{Z}$, with $\dot{\rho}(2\pi\ell) = (1, 0, \ldots, 0)$. The resulting F-geodesic

$$\begin{aligned} \mathcal{P}(t) &= \varphi(t, \rho(t)) = (\sin t, 0, \ldots, 0, \cos t) \, \mathrm{Rot}(a_1 t, \ldots, a_m t) \\ &= (\sin t \cos a_1 t, \sin t \sin a_1 t, 0, \ldots, 0, \cos t) \end{aligned}$$

also passes through the north pole at $t = 2\pi\ell$, $\ell \in \mathbb{Z}$. However,

$$\dot{\mathcal{P}}(2\pi\ell) = (\cos 2\pi\ell a_1, \sin 2\pi\ell a_1, 0, \ldots, 0).$$

So $\dot{\mathcal{P}}(2\pi\ell) = \dot{\mathcal{P}}(0)$ if and only if $a_1\ell \in \mathbb{Z}$. Since $0 < a_1 < 1$, $\dot{\mathcal{P}}(2\pi) \neq \dot{\mathcal{P}}(0)$. Hence $\mathcal{P}$ self-intersects, but does not close, at $t = 2\pi$.

Implicit in this discussion is the fact that this $\mathcal{P}$ will close if and only if $\ell a_1 \in \mathbb{Z}$ for some integer ℓ. That is, a_1 must be rational. Indeed, Ziller observed that the a_i may be selected so that the only closed geodesics of F are, modulo parameterization, the m great circles C_i. (As above $n = 2m-1$, $2m$.) The following choice of a_i effects this scenario. Set $a_i = a/p_i$, where $0 < a < 1$ is irrational and the p_i are relatively prime integers with $1 < p_1 < p_2 < \cdots < p_m$. Note that there are infinitely many possible choices for the m-tuples $a = \{a_i\}_{i=1}^m$, each determining a Randers metric F_a. Proposition 8 of [BRS04] implies F_a and $F_{\tilde{a}}$ are locally isometric if and only if $a = \tilde{a}$.

Recollect that the F-length of C_i depends on the choice of orientation. We say each orientation produces a distinct geodesic. Hence (S^n, F) has precisely $2m$ closed geodesics. As a consequence we have the following.

Proposition 5 ([Zi82])**.** *There exist infinitely many, nonisometric Randers metrics of constant flag curvature $K = 1$ on S^n with only $2m$ closed geodesics. (Here $n = 2m$ or $2m - 1$.)*

We would like to make two comments here. First, the fact that Randers metrics are not reversible is key. R. Bryant has shown that any reversible Finsler metric of constant flag curvature $K = 1$ on the 2-sphere is necessarily the standard Riemannian metric of constant sectional curvature 1 [Br04]. In particular, *all* the geodesics close.

Second, the proposition implies that there exists a Finsler metric on S^2 with only 2 closed geodesics. This is a *non-Riemannian* counterexample to the Three Closed Geodesic Theorem of Lusternik and Schnirelmann [LS29, LS30]: *On S^2 with an arbitrary Riemannian metric, there exist three simple closed geodesics.* (See also [B78, G89].)

The phenomenon of Proposition 5 is most easily illustrated on the 2-sphere, where $W(p) = a(p^1, -p^0, p^2)$, $0 < a < 1$, is an infinitesimal rotation about the p^2-axis. The unique unit-speed geodesic of F through $q \in S^2$ in the direction $v \in T_q S^2$
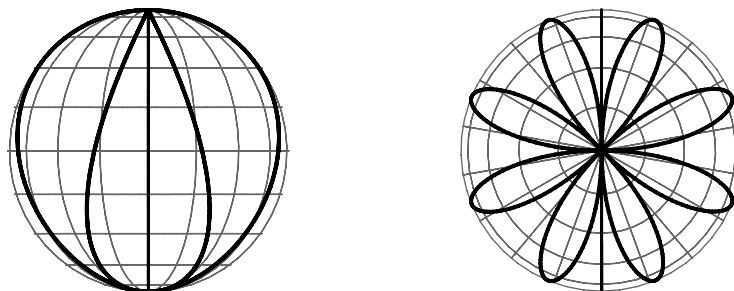
EXAMPLE 11. $a = \frac{1}{4}$. The Riemannian geodesic (here seen as the vertical line bisecting the sphere) passes through the North and South poles, the fixed points of the flow. The parameterization ρ traces the geodesic curve four times before the associated Randers geodesic closes.
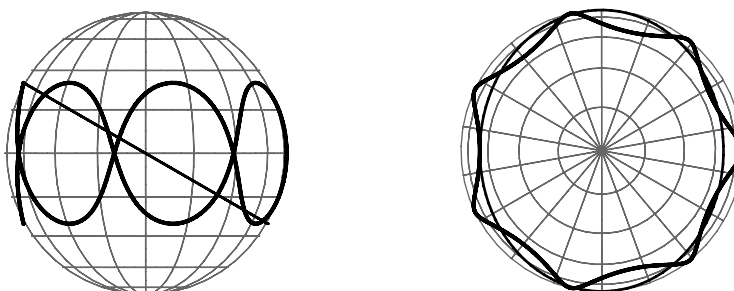
EXAMPLE 12. $a = \frac{5}{7}$. A Riemannian geodesic (the diagonal line on the left, and ellipse on the right) omitting the fixed points of the flow. The parameterization ρ circles the sphere seven times before the Randers geodesic closes.

is given by

$$\mathcal{P}(t) = \varphi(t, \rho(t)) = \begin{pmatrix} \rho^0(t)\cos at \ - \ \rho^1(t)\sin at \\ \rho^0(t)\sin at \ + \ \rho^1(t)\cos at \\ \rho^2(t) \end{pmatrix},$$

where $\rho(t) = (\rho^0(t), \rho^1(t), \rho^2(t)) = \cos(t)\, q + \sin(t)\,(v - W_q)$. In the event that a is irrational, it is not difficult to check that $\mathcal{P}(t)$ closes if and only if $\rho(t)$ parameterizes the equator $\{p^2 = 0\}$. As Ziller suggests, this great circle is the only geodesic of h invariant under the flow $\varphi(t, \cdot)$. The F-geodesic $\mathcal{P}$ also traces the equator. In the event that $\rho(t)$ parameterizes the equator in the direction of rotation by φ, then the F-length of $\mathcal{P}$ is $2\pi/(1 + a)$. Otherwise, $\mathcal{P}$ has length $2\pi/(1 - a)$. These two parameterizations of the equator are the only closed geodesics on (S^2, F).

Examples 11 and 12 are two examples of Randers geodesics on the 2-sphere. The pictures on the left are side-views of the sphere, on level with the equator. On the right we look down on the North pole.

We will close this discussion with two additional observations of Ziller [Zi82] and a new result of V. Bangert and Y. Long [BL04]: (1) Ziller has shown that
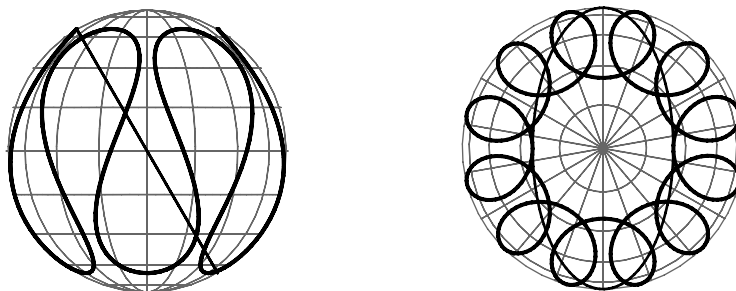
EXAMPLE 13. $a = \frac{5}{6}$. The Riemannian geodesic ρ (seen as a diagonal line on the left, and ellipse on the right) is traced six times before the Randers geodesic $\mathcal{P}_+(t) = \varphi(t, \rho(t))$ closes. Notice that $\mathcal{P}_+$ is a simple closed curve.
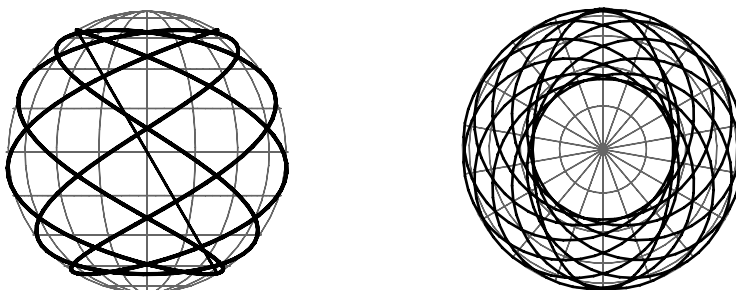
EXAMPLE 14. $a = \frac{5}{6}$. This Randers geodesic is generated by tracing the Riemannian geodesic in the opposite direction. That is, $\mathcal{P}_-(t) = \varphi(t, \rho(-t))$. The two Randers geodesics in Examples 13 and 14 are tangent at their initial point $\rho(0)$.

any Finsler metric on S^n sufficiently $\mathcal{C}^2$ close to h has *at least n* closed geodesics. The Randers examples above show that this lower bound is sharp in the case that $n = 2m$. Bangert–Long have established a stronger result in the case $n = 2$: *any* Finsler metric on the 2-sphere necessarily has two prime closed geodesics. (2) Ziller also points out that given a generic $\mathcal{C}^2$ Finsler metric on a compact manifold, the initial vectors of closed geodesics are dense in the unit tangent bundle. In particular, examples of Finsler metrics on S^n with only finitely many closed geodesics are rare.

It should be noted that Randers metrics of constant curvature on the sphere have been studied from several distinct perspectives. These include the Bao–Shen [BS02] one-parameter family of examples developed from the Hopf fibration on S^3, and the approach of Bejancu–Farran through Sasakian space forms [BF02, BF03].

The final two examples in this section, Examples 13 and 14, are generated from the same Riemannian geodesic, but with opposite orientations. Again the picture on the left is a side-view of the sphere, on level with the equator; on the right we look down on the North pole.

4.1. **A solution to the navigation problem.** We would like to sketch the means by which a pilot flying under windy conditions may determine the time-efficient paths. Let me emphasize that *this discussion applies only to those W generating*

rigid motions of the sphere. Imagine that the plane traverses the globe (represented by the 2-sphere) with unit speed. Introduce a mild wind (speed less than one) describing rotation about the z-axis with velocity vector field W. This wind is an isometry, and $|W| < 1$, so the hypotheses of Theorem 2 are met. The plane now travels with unit speed relative to the wind, not the earth. The goal is to determine a time-efficient flight path from p to q.

According to the geodesic classification these paths are of the form $\mathcal{P}(t) = \varphi(t, \rho(t))$. So, if $\mathcal{P} : [0, L] \to S^2$ joins p to q, $\rho(t) = \varphi(-t, \mathcal{P}(t))$ is a geodesic path from p to $\varphi(-L, q)$, a point in the pre-orbit of q under the wind φ_t. Note that L is the time it takes the wind to blow a particle from $\rho(L)$ to q. These observations suggest the following approach.

Define $q(\tau) = \varphi(-\tau, q)$. This curve traces the circle in S^2 parallel to the xy-plane, and passing through q. Intuitively, τ is the time it takes the wind to travel from $q(\tau)$ to q, and the pilot gauges the wind to determine a point $q(\tau)$ at distance τ from p. Such a point in the pre-orbit certainly exists: Let $0 \le L(\tau) \le \pi$ be the Riemannian distance from p to $q(\tau)$. If $p \ne q$, then $L(0) > 0$, and the intermediate value theorem gives us $\tau_o \in (0, \pi]$ such that $L(\tau_o) = \tau_o$. That is, the Riemannian distance from p to $q(\tau_o)$ is equal to the time it takes the wind to blow $q(\tau_o)$ to q. Take the smallest such τ_o.

Then, if $\rho : [0, L(\tau_o)] \to S^2$ is a unit speed geodesic joining p to $q(\tau_o)$, $\mathcal{P}(t) = \varphi(t, \rho(t))$ will be a time-minimizing flight path from p to q. Note that this scenario is akin to the swimmer at the edge of a river, wishing to reach the bank directly opposite, who aims not for her desired destination but slightly upstream.

The following is a sequence of pictures illustrating this construction for three values of τ: $0 < \tau_- < \tau_o < \tau_+$. Each image contains the reverse flow $\varphi(-t, q)$ from q to $q(\tau)$, the Riemannian geodesic $\rho(t) = \rho_\tau(t)$ from p toward $q(\tau) = \varphi(-\tau, q)$, and the resulting Randers geodesic $\mathcal{P}(t) = \varphi(t, \rho(t))$; $t \in [0, \tau]$.
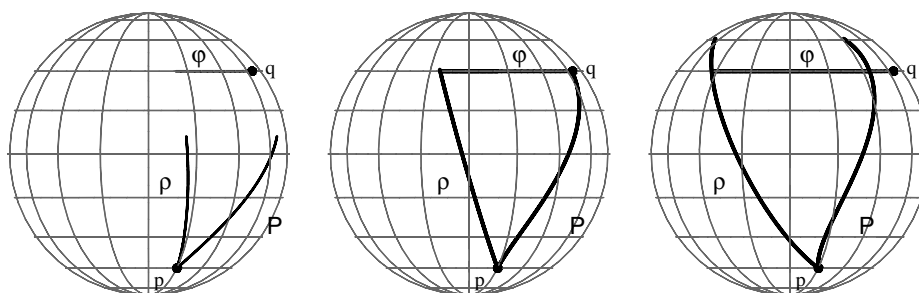
5. Conjugate points

Fix $p \in M$. This section develops the correspondence between points $\hat{q} \in \mathcal{M}$ conjugate to p with respect to the Riemannian metric h, and points $q \in M$ conjugate to p with respect to the Randers metric F. As before $\mathcal{P} : [0, \ell] \to M$ denotes a unit speed F geodesic joining p to q. For convenience, we will assume throughout that $\rho(t) = \varphi(-t, \mathcal{P}(t))$ is a well-defined curve mapping the interval into $\mathcal{M}$. (This is the case for all the examples in Sections 3 and 4.)

Begin by supposing that $\rho(\ell) = \hat{q}$ is conjugate to p along ρ. Let $\hat{\rho} : [0, \hat{\ell}] \to \mathcal{M}$ be a unit h-speed reparameterization of ρ. By assumption there exists a nonzero Jacobi

field $\mathcal{J}$ along $\hat{\rho}$ that vanishes at p and $\hat{q}$. Suppose $\hat{\rho}_s : [0, \hat{\ell}] \to \mathcal{M}$, $-\varepsilon < s < \varepsilon$, is a geodesic variation of $\hat{\rho} = \hat{\rho}_0$ with variational vector field

$$\left. \frac{\partial \hat{\rho}_s}{\partial s} \right|_{s=0} = \mathcal{J} \,.$$

We wish to use this Riemannian geodesic variation to construct a Randers geodesic variation. To that end, it is first necessary to reparameterize the $\hat{\rho}_s$ so that the composition with the flow φ is a Randers geodesic. Define $\tau : [0, \ell] \to [0, \hat{\ell}]$ by

$$\tau(t) = \begin{cases} \frac{2}{\sigma} \left(1 - e^{-\sigma t/2} \right), & \sigma \neq 0, \\ t, & \sigma = 0 \,. \end{cases}$$

Then $\rho_s(t) = \hat{\rho}_s(\tau(t))$ reparameterizes the geodesics with speed $|\dot{\rho}_s(t)| = v(s)e^{-\sigma t/2}$; here $v(s)$ denotes the constant h-speed of $\hat{\rho}_s$.

Now $\mathcal{P}_s(t) = \varphi(v(s)t, \rho_s(t))$ is an F-geodesic variation. Each geodesic $\mathcal{P}_s$ has constant F-speed $v(s)$. In particular, $\mathcal{P} = \mathcal{P}_0$ has speed $v(0) = 1$. It follows that

$$J := \left. \frac{\partial \mathcal{P}_s}{\partial s} \right|_{s=0} = v'(0)W_{\mathcal{P}} + (d\varphi)\mathcal{J}$$

is a Jacobi field of F. (See [BCS00] for a through treatment of Jacobi fields in Finsler geometry.) Making use of the facts that $\mathcal{J}(0) = \mathcal{J}_p = 0$ and $\mathcal{J}(\ell) = \mathcal{J}_{\hat{q}} = 0$, a standard Jacobi field argument yields $v'(0) = 0$. Hence $J = (d\varphi)\mathcal{J}$. Recollect that $d\varphi$ denotes the differential of $\varphi(t, \cdot)$. Since $\varphi(t, \cdot)$ is a local diffeomorphism, the Jacobian $d\varphi$ is nonsingular. Hence J vanishes if and only if $\mathcal{J}$ does. In particular, J is nonzero, and $J(0) = 0 = J(\ell)$. Equivalently, $q = \mathcal{P}(\ell)$ is conjugate to $p = \mathcal{P}(0)$ along $\mathcal{P}$ with respect to the metric F.

This establishes the following: Suppose $\mathcal{P} : [0, \ell] \to M$ is a unit-speed F geodesic with associated Riemannian geodesic ρ (see Theorem 2). Then $\mathcal{P}(\ell)$ is conjugate to $\mathcal{P}(0)$ with respect to F whenever $\rho(\ell)$ is conjugate to $\rho(0)$ with respect to h. A similar argument shows the converse holds as well, and we have the following.

Proposition 6. *Assume $(\mathcal{M}, h)$ is a Riemannian manifold equipped with a globally defined infinitesimal homothety W. Let F denote the Randers metric on $M = \{|W| < 1\} \subset \mathcal{M}$ with navigation data (h, W). Suppose that $\mathcal{P} : [0, \ell] \to M$ is a geodesic of (M, F), and that the associated Riemannian geodesic $\rho(t) = \varphi(-t, \mathcal{P}(t))$ (Theorem 2) is defined for all $t \in [0, \ell]$. Then $\mathcal{P}(\ell)$ is conjugate to $\mathcal{P}(0)$ along $\mathcal{P}$ (with respect to the F) if and only if $\rho(\ell)$ is conjugate to $\rho(0)$ (with respect to h).*

In Riemannian geometry p is conjugate to $\hat{q}$ if and only if $\hat{q}$ is conjugate to p. This is because $\rho(t)$ is a geodesic if and only if $\rho(-t)$ is as well. As indicated in the discussion of Case I in Subsection 3.1, the phenomenon does not hold for Randers metrics in general.

6. Globally minimal geodesics

In this section we will establish a relationship between the minimal geodesics of F and those of h. Assume throughout that $(\mathcal{M}, h)$ is complete, so that $\rho(t)$ and $\varphi(t, p)$ are defined for all $t \in \mathbb{R}$ and $p \in \mathcal{M}$. All the examples of Sections 3 and 4 meet this criterion. The Hopf–Rinow Theorem assures us that any two points of $\mathcal{M}$ may be joined by a globally h-length minimizing geodesic. We will see that, in the event that M satisfies a given condition (determined by σ and specified below), any two points of (M, F) can also be joined by a globally F-length minimizing geodesic.

Note that in each of the examples in Sections 3 and 4 $\{|W| < 1\}$ is connected, and *convex* in the following sense: any globally h-length minimizing geodesic joining $p, q \in \{|W| < 1\}$, *a priori* contained in $\mathcal{M}$, lies in $\{|W| < 1\}$. Moreover, in the example with nonzero σ (Case II), $d_s{}^2|W_{\eta(s)}|^2 \geq \frac{1}{2}\sigma^2$ for any unit speed geodesic $\eta(s)$ of h. These observations motivate the hypotheses below.

Proposition 7. *Suppose $(\mathcal{M}, h)$ is a complete, connected Riemannian manifold, and that F solves the navigation problem for an infinitesimal homothety W on a connected component M of the open sub-manifold $\{|W| < 1\} \subset \mathcal{M}$.*

- *If $\sigma = 0$, assume M is convex.*
- *If $\sigma \neq 0$, assume*

$$\frac{d^2}{ds^2}|W_{\eta(s)}|^2 \geq \frac{1}{2}\sigma^2 > 0$$

  for every unit speed h-geodesic η.

Then any two points p, q in the Randers space (M, F) may be joined by a globally length minimizing geodesic.

Remark. Note that every constant flag curvature Randers metric F meets the conditions of the proposition. This is a consequence of Theorem 3 which assures us that these metrics arise as solutions to Zermelo's problem on a Riemannian manifold $(\mathcal{M}, h)$ of constant curvature under an infinitesimal homothety. In the case that $\sigma = 0$ (so that W is an infinitesimal isometry) the connected components of $\{|W| < 1\}$ are convex in each of the three models of Riemannian space forms: the sphere, Euclidean space and the hyperbolic ball. In general, $(\mathcal{M}, h)$ will be locally isometric to one of these three manifolds so that every $p \in M$ admits a convex neighborhood and so the proposition will hold locally, at the very least, for these F, and globally when $\mathcal{M}$ is one of the three models.

(That the proposition holds locally is of little interest: short geodesics are always length minimizing on a Finsler manifold. See [BCS00]. Here the interesting content lies in the proof, which will establish (one-half of) a relationship between those ρ minimizing h-length and those $\mathcal{P}$ minimizing F-length. See Proposition 8.)

If, on the other hand $\sigma \neq 0$, then h must be flat and there exist coordinates in which h is the standard Euclidean metric and $W = \frac{1}{2}\sigma x + Qx + C$, for some skew-symmetric matrix Q and $C \in \mathbb{R}^n$ (cf. [BRS04]). Hence $d_s{}^2|W|^2 \geq \frac{1}{2}\sigma^2$, and the conditions of the proposition are again met. We will discuss at the end of this section the generality with which we may expect this inequality to hold.

Proof. Let $\ell = d_F(p, q)$ denote the distance from p to q in M. Then there are F-unit speed curves $C_i : [0, \ell_i] \to M$ such that $C_i(0) = p$, $C_i(\ell_i) = q$ and $\ell_i \to \ell$. Define $c_i : [0, \ell_i] \to \mathcal{M}$ by $c_i(t) = \varphi(-t, C_i(t))$. Note that

$$c_i(\ell_i) = \varphi(-\ell_i, C_i(\ell_i)) = \varphi(-\ell_i, q) \longrightarrow \varphi(-\ell, q) \,.$$

Additionally, Lemma 1, $F(\dot{C}_i) = 1$ and the fact that $\varphi(t, \cdot)$ is a homothety imply $|\dot{c}_i(t)|^2 = e^{-\sigma t}$. As a result

$$d_h(p, c_i(\ell_i)) \leq h\text{-length}(c_i) = \begin{cases} \frac{2}{\sigma}\left(1 - e^{-\sigma\ell_i/2}\right), & \sigma \neq 0, \\ \ell_i, & \sigma = 0. \end{cases}$$

The continuity of $x \mapsto d_h(p, x)$, in conjunction with these two observations, allows us to deduce

$$d_h(p, \varphi(-\ell, q)) \leq \begin{cases} \frac{2}{\sigma} \left(1 - e^{-\sigma\ell/2}\right), & \sigma \neq 0, \\ \ell, & \sigma = 0. \end{cases}$$

Define $q(\tau) := \varphi(-\tau, q)$. Intuitively, τ is the time it takes the flow to carry $q(\tau)$ onto q. Note that $q = q(0)$. Since $(\mathcal{M}, h)$ is complete, the Hopf-Rinow theorem guarantees a globally h-length minimizing geodesic $\rho_\tau : [0, L(\tau)] \to \mathcal{M}$ joining p to $q(\tau)$, and parameterized so that $|\dot\rho_\tau(t)|^2 = e^{-\sigma t}$. The expression above for $d_h(p, \varphi(-\ell, q)) = d_h(p, q(\ell))$ implies $L(\ell) \leq \ell$.

If $p = q$, the proposition is immediate. So suppose that $p \neq q = q(0)$. Then $L(0) > 0$. The Mean Value Theorem gives us $\ell_o \in (0, \ell]$ such that $L(\ell_o) = \ell_o$. Write $\rho_{\ell_o} = \rho$. It follows that $\mathcal{P}(t) = \varphi(t, \rho(t))$ is a unit speed F-geodesic joining $\mathcal{P}(0) = p$ to $\mathcal{P}(\ell_o) = \varphi(\ell_o, \rho(\ell_o)) = \varphi(\ell_o, q(\ell_o)) = q$.

Since $\mathcal{P}$ *a priori* maps $[0, \ell_o]$ into $\mathcal{M}$, it remains to confirm that the interval is mapped into M. To that end, consider the sets

$$M_t := \varphi(t, M) = \{x \in \mathcal{M} : |W_x| < e^{\sigma t/2}\}.$$

The second equality follows from the hypotheses that (a) $\mathcal{M}$ is connected; (b) $d\varphi W_x = W_{\varphi(t,x)}$ (since φ is the flow of W); and (c) W is an infinitesimal homothety, $|W_{\varphi(t,x)}| = |d\varphi W_x| = e^{\sigma t/2}|W_x|$. Recall $d\varphi$ is the differential of $\varphi(t, \cdot) : \mathcal{M} \to \mathcal{M}$. In general, $\mathcal{P}(t) = \varphi(t, \rho(t))$ will lie in M if and only if $\rho(t) \in M_{-t}$.

In the case that $\sigma = 0$, $M_t = M$. In particular, $\rho(0), \rho(\ell_o) \in M = M_{-\ell_o}$, and convexity implies $\rho(t) \in M$. Therefore $\mathcal{P}(t) = \varphi(t, \rho(t)) \in M_t = M$.

Let us now address the case $\sigma \neq 0$. In order to show that $P(t) \in M$, it suffices to see that $\rho(t) \in M_{-t}$. Equivalently,

$$|W_{\rho(t)}|^2 < e^{-\sigma t}.$$

Define $s = (2/\sigma)(1 - e^{-\sigma t/2})$. Then $\eta(s(t)) = \rho(t)$ defines a unit speed reparameterization $\eta(s)$ of $\rho(t)$. The inequality above is recast as

$$w(s) := |W_{\eta(s)}|^2 < \left(1 - \frac{\sigma}{2}s\right)^2 =: f(s).$$

Since $\eta(0) = \rho(0) = p \in M$ and $\eta(s(\ell_o)) = \rho(\ell_o) = \varphi(-\ell_o, q) \in M_{-\ell_o}$, the expression holds at the end points. Additionally, $0 < \frac{1}{2}\sigma^2 = f''(s) \leq w''(s)$. Therefore the inequality is satisfied for all $s \in [0, s(\ell_o)]$. We conclude that $\mathcal{P}(t) \in M$, whence $\mathcal{P}$ is a globally length minimizing geodesic of (M, F). Finally, $d_F(p, q) = \ell$ implies $\ell = \ell_o$. $\qquad\square$

Implicit in this proof is the fact that if $\mathcal{P}$ is a globally minimal geodesic of (M, F), then the associated Riemannian geodesic ρ is a global minimizer of $(\mathcal{M}, h)$. The converse holds as well.

Suppose that ρ is a length minimizing geodesic of $(\mathcal{M}, h)$ joining $\rho(t) = p$ to $\rho(\ell) = \hat{q}$. If $\mathcal{P}(t) = \varphi(t, \rho(t))$ lies in M for all t, so that it is a geodesic of (M, F), then $\mathcal{P}$ is also a global minimizer. To see this, we argue by contradiction, assuming that $d_F(p, q) = \ell_o < \ell$, where $q = \mathcal{P}(\ell)$. Let $\mathcal{P}_o : [0, \ell_o] \to M$ be a length minimizing curve joining $p = \mathcal{P}_o(0)$ to $q = \mathcal{P}_o(\ell_o)$, and consider the associated Riemannian geodesic $\rho_o(t) = \varphi(-t, \mathcal{P}_o(t))$ from p to $\rho_o(\ell_o)$. The path $\xi : [-\ell_o, -\ell] \to \mathcal{M}$ defined by $\xi(t) = \varphi(t, q)$ runs from $\rho_o(\ell_o)$ to $\rho(\ell)$. Consequently $\mathrm{length}_h(\rho) \leq \mathrm{length}_h(\rho_o) + \mathrm{length}_h(\xi)$. Computing $|\dot\xi(t)|^2 = |W_{\varphi(t,q)}|^2 = |d\varphi W_q|^2 = e^{\sigma t}|W_q|^2 <$

$e^{\sigma t}$, a straightforward computation produces $\text{length}_h(\xi) < \text{length}_h(\rho) - \text{length}_h(\rho_o)$. This is a contradiction; consequently $\mathcal{P}$ must be minimal. We have established the following.

Proposition 8. *Suppose $(\mathcal{M}, h)$ is a complete, connected Riemannian manifold, and that F solves the navigation problem for an infinitesimal homothety W on a connected component M of the open sub-manifold $\{|W| < 1\} \subset \mathcal{M}$.*

- *If the homothety constant σ is zero, assume M is convex.*
- *If $\sigma \neq 0$, assume that*

$$\frac{d^2}{ds^2}|W_{\eta(s)}|^2 \geq \frac{1}{2}\sigma^2 > 0$$

for every unit speed h geodesic η.

Then a geodesic $\mathcal{P}$ of F is a global minimizer of F-path-length from $\mathcal{P}(0)$ to $\mathcal{P}(\ell)$ if and only if the associated Riemannian geodesic ρ (Theorem 2) globally minimizes h-path-length between $\rho(0)$ and $\rho(\ell)$.

Corollary 9. *The point q is a cut point of p, with respect to F, if and only if $\hat{q} = \varphi(-\ell, q)$ is a cut point of p with respect to h. (Here ℓ is the F-distance from p to q.)*

As promised, we close this section by discussing the inequality $d_s{}^2|W|^2 \geq \frac{1}{2}\sigma^2$. First note that it does not hold for any infinitesimal homothety. As a counter-example take the standard 2-sphere parameterized by

$$(\theta, \phi) \mapsto (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi).$$

Note that $W = \partial_\theta$ is an infinitesimal isometry, and $\eta(s) = (\theta(s), \phi(s)) = (0, s)$ is a unit speed geodesic along which $d_s{}^2|W|^2 = 2\left(\sin^2(s) - \cos^2(s)\right)$ fails the inequality.

However, the inequality will hold whenever the Riemannian sectional curvature $\kappa(\dot{\eta} \wedge W)$ of the plane spanned by $\{\dot{\eta}, W\}$ is nonpositive. To see this note that

$$|\nabla_{\dot{\eta}}W| \ \geq \ h(\dot{\eta}, \nabla_{\dot{\eta}}W) \ = \ \tfrac{1}{2}\sigma,$$

a consequence of the homothety hypothesis $h(v, \nabla_v W) = \frac{1}{2}\sigma|v|^2$, for all $v \in T_x\mathcal{M}$. Whence

$$(7) \qquad d_s{}^2|W|^2 \ = \ 2|\nabla_{\dot{\eta}}W|^2 + 2h(W, \nabla_{\dot{\eta}}{}^2 W) \ \geq \ \tfrac{1}{2}\sigma^2 + 2h(W, \nabla_{\dot{\eta}}^2 W).$$

Since W is an infinitesimal homothety we have

$$(8) \qquad\qquad\qquad W_{i:j:k} + W_{j:i:k} = 0.$$

Consequently, the inner product appearing above may be re-expressed as

$$\begin{aligned} h(W, \nabla_{\dot{\eta}}^2 W) \ &= \ W^i W_{i:j:k}\dot{\eta}^j\dot{\eta}^k \ = \ -\dot{\eta}^i W_{i:j:k} W^j \dot{\eta}^k \\ &= \ -\kappa(\dot{\eta} \wedge W)\left(|W|^2 - h(\dot{\eta}, W)^2\right). \end{aligned}$$

Here, (8) provides the second equality, and the Ricci identity for W, followed by a second application of (8), provides the last. Since $|W|^2 - h(\dot{\eta}, W)^2 \geq 0$, we have $h(W, \nabla_{\dot{\eta}}^2 W) \geq 0$ whenever $\kappa(\dot{\eta} \wedge W) \leq 0$, and (7) implies

$$d_s{}^2|W|^2 \geq \tfrac{1}{2}\sigma^2.$$

Acknowledgments

The author thanks D. Bao and W. Ziller for many helpful comments.

References

[AZ88] H. Akbar-Zadeh, Sur les espaces de Finsler á courbures sectionnelles constantes. *Acad. Roy. Belg. Bull. Cl. Sci.* (5) **74** (1988), 281-322. MR1052466 (91f:53069)

[A55] L. Auslander, On curvature in Finsler geometry, *Trans. of Amer. Math. Soc.* **79** (1955), 378-388. MR0071833 (17:190d)

[B78] W. Ballmann, Der Satz von Lusternik und Schnirelmann (German), *Beiträge zur Differentialgeometrie, Heft 1*, 1-25. Bonner Math. Schriften, **102**, Univ. Bonn, Bonn, 1978. MR0520178 (80d:58017)

[BL04] V. Bangert and Y. Long, Multiple closed geodesics on Finsler 2-spheres and a conjecture of D. V. Anosov, in preparation.

[BCS00] D. Bao, S.S. Chern and Z. Shen, *An Introduction to Riemann–Finsler Geometry*, Graduate Texts in Mathematics, **200**, Springer, 2000. MR1747675 (2001g:53130)

[BR04] D. Bao and C. Robles, On Ricci and flag curvatures in Finsler geometry, in *Some perspectives in Finsler geometry*, MSRI Publications **50**, Cambridge University Press, 2004. MR2132660 (2005k:53124)

[BRS04] D. Bao, C. Robles and Z. Shen, Zermelo navigation on Riemannian manifolds, *J. Diff. Geom.* **66** (2004), 391-449. MR2106471 (2005k:58023)

[BS02] D. Bao and Z. Shen, Finsler metrics of constant positive curvature on the Lie group S^3, *J. London Math. Soc.* **66** (2002), 453–467. MR1920414 (2003g:53129)

[BF02] A. Bejancu and H. Farran, Finsler metrics of positive constant flag curvature on Sasakian space forms, *Hokkaido Math. J.* **31**(2) (2002), 459–468. MR1914971 (2003f:53133)

[BF03] A. Bejancu and H. Farran, Randers manifolds of positive constant curvature, *Int'l. J. Math. & Math'l. Sc.* **18** (2003), 1155–1165. MR1978413 (2004d:53090)

[Br96] R. Bryant, Finsler structures on the 2-sphere satisfying $K = 1$, *Cont. Math.* **196** (1996), 27–41. MR1403574 (97e:53128)

[Br97] R. Bryant, Projectively flat Finsler 2-spheres of constant curvature, *Selecta Math. (N.S.)* **3**(2) (1997), 161–203. MR1466165 (98i:53101)

[Br04] R. Bryant, Geodesically reversible Finsler 2-spheres of constant curvature, arXiv:math.DG/0407514.

[C99] C. Carathéodory: *Calculus of Variations and Partial Differential Equations of the First Order*, AMS Chelsea Publishing, 1999. MR1279593 (96e:49001)

[F29] P. Funk, Über Geometrien, bei denen die Geraden die Kürzesten sind, *Math. Ann.* **101** (1929), 226–237. MR1512527

[G89] M. Grayson, Shortening embedded curves, *Ann. of Math.* (2) **129** (1989), no. 1, 71-111. MR0979601 (90a:53050)

[K73] A. B. Katok, Ergodic perturbations of degenerate integrable Hamiltonian systems, *Math. USSR Izvestija* (3) **7** (1973), 535–571. MR0331425 (48:9758)

[LS29] L. Lusternik and L. Schnirelmann, Sur le problème de trois géodésiques fermées sur les surfaces de genre 0, *C. R. Acad. Sci. Paris* **189** (1929), 269-271.

[LS30] L. Lusternik and L. Schnirelmann, Topological methods in the calculus of variations, *Gosudarstv. Izdat. Tehn.-Teor. Lit.*, Moscow (1930) [Russian]; Méthodes topologiques dans les problèmes variationelles, *1re partie, Actualité Sci. Industr.* **188** (1934).

[O83] T. Okada, On models of projectively flat Finsler spaces of constant negative curvature, *Tensor, N.S.* **40** (1983), 117–124. MR0837784 (87c:53124)

[R41] G. Randers, On an asymmetric metric in the four-space of general relativity, *Phys. Rev.* **59** (1941), 195–199. MR0003371 (2:208a)

[Ra61] A. Rapcsák, Über die bahntreuen Abbidungen metrischer Räume, *Publ. Math., Debrecen* **8** (1961), 285–290. MR0138079 (25:1526)

[S96] Z. Shen, Finsler spaces of constant positive curvature, in *Finsler Geometry*, Contemporary Math. **196** (1996), 83-92. MR1403580 (97m:53120)

[S02] Z. Shen, Finsler metrics with $K = 0$ and $S = 0$, *Canadian J. Math.* **55** (2003), 112–132. MR1952328 (2004e:53112)

[Z31]   E. Zermelo, Über das Navigationsproblem bei ruhender oder veränderlicher Wind-
        verteilung, *Z. Angew. Math. Mech.* **11** (1931), 114–124.
[Zi82]   W. Ziller, Geometry of the Katok examples, *Ergod. Th. & Dynam. Sys.* **3** (1982), 135–
        157. MR0743032 (86g:58036)

Department of Mathematics, Texas A&M University, College Station, Texas 77843-
3368
    E-mail address: `robles@math.tamu.edu`