A Riemannian approach to Randers geodesics

Dorje C. Brody[a,b], Gary W. Gibbons[c], David M. Meier[a]

[a] *Department of Mathematics, Brunel University London, Uxbridge UB8 3PH, UK*
[b] *Department of Optical Physics and Modern Natural Science, St Petersburg National Research University of Information Technologies, Mechanics and Optics, 49 Kronverksky Avenue, St Petersburg 197101, Russia*
[c] *Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK*

Abstract

In certain circumstances tools of Riemannian geometry are sufficient to address questions arising in the more general Finslerian context. We show that one such instance presents itself in the characterisation of geodesics in Randers spaces of constant flag curvature. To achieve a simple, Riemannian derivation of this special family of curves, we exploit the connection between Randers spaces and the Zermelo problem of time-optimal navigation in the presence of background fields. The characterisation of geodesics is then proven by generalising an intuitive argument developed recently for the solution to the quantum Zermelo problem.

Key words: Finsler geometry, Zermelo navigation, Randers metric, control theory
2000 MSC: 53B40, 58J60, 53B50

Investigations of Finsler manifolds usually require tools more involved than those of Riemannian geometry [1]. For instance, whereas the Levi–Civita connection of Riemannian geometry is a linear connection on the tangent bundle of the underlying manifold, one of its generalisations in the Finslerian context, the so-called Chern connection, is a linear connection on a distinguished vector bundle over the projective sphere bundle [2]. Nevertheless, in certain situations Riemannian methods are sufficient to deal with aspects of Finsler geometry and the resulting simplifications, such as the ones reported below, can be substantial. Specifically, what we show is that the main result of [3], namely, the characterisation of the geodesics of a special class of Finsler spaces, can be proven using tools from Riemannian geometry only.

To begin, let us recall that a Finsler manifold $(\mathcal{M}, F)$ is a C^∞ manifold $\mathcal{M}$ together with a positive function $F(x,y)$ on the tangent bundle, called the Finsler function, which is required to be C^∞ and homogeneous of first degree, that is, $F(x,\lambda y) = \lambda F(x,y)$ for any $\lambda > 0$. Moreover, the Hessian of F^2 with respect to y:

$$g_{ij}(x,y) = \frac{1}{2}\frac{\partial^2}{\partial y^i y^j}\, F^2(x,y)$$

is assumed to be positive-definite outside the zero-section of $T\mathcal{M}$. It can be shown that $F(x,y) = \sqrt{g_{ij}(x,y)y^i y^j}$.

If F can be expressed in the form

$$F(x,y) = \sqrt{\alpha_{ij}y^i y^j} + \beta_i y^i,$$

where α is a Riemannian metric and β a one-form, then $\mathcal{M}$ is called a Randers space. The Finslerian metric on $\mathcal{M}$ of Randers type thus takes the form

$$g_{ij}(x,y) = \alpha_{ij} + \beta_i\beta_j + \frac{(\alpha_{ij}\beta_k + \alpha_{jk}\beta_i + \alpha_{ki}\beta_j)y^k}{(\alpha_{kl}y^k y^l)^{1/2}} + \frac{(\beta_k y^k)\alpha_{ik}\alpha_{jl}y^k y^l}{(\alpha_{kl}y^k y^l)^{3/2}}.$$

Randers spaces were first introduced in [4] in the context of a unified theory of gravitation and electromagnetism and arise in a wide range of physical applications such as the electron microscope [5], the propagation of sound and light rays in moving media [6, 7], and the time-optimal control in the presence of background fields [8]—the last point being of particular relevance for the present discussion.

To explain the connection between Randers spaces and time-optimal control, we start from a Riemannian manifold $\mathcal{M}$ with metric h, together with a vector field W that satisfies $|W| < 1$ and plays the role of background field, or 'wind'. The goal is to solve the Zermelo problem, that is, to navigate from one point on $\mathcal{M}$ to another along the path $q(s)$ in the shortest possible time under the influence of W, assuming a maximum attainable speed of $|\dot{q}| = 1$ if wind were absent. A problem of this kind was first posed and solved by Zermelo for the navigation of ships at sea (modelled as the Euclidean plane) for a general spacetime-dependent field W [9] (see also [10]). The general formulation on Riemannian manifolds under time-independent fields, and the connection to Randers spaces, was identified more recently by Shen [8]. The idea can be illustrated as follows. Supposing for a moment that one were able to travel for finite time in a tangent space $T_p\mathcal{M}$ for a fixed p, it is clear that the set of destinations reachable in one unit of time coincides with the unit circle, shifted by $W(p)$. Correspondingly, the minimum time $F(p,v)$ it takes to reach the tip of a given vector v in $T_p\mathcal{M}$ is given by the ratio $|v|/|\rho_v|$ of Euclidean norms, where ρ_v is the unique vector collinear with v that lies on the shifted unit circle. To put it differently, the vector $v/F(p,v) - W(p)$ has unit length. It follows that

$$F(p,v) = \frac{-h(v,W(p)) + \sqrt{h(v,W(p))^2 + |v|^2(1-|W(p)|^2)}}{1-|W(p)|^2}.$$

The function F defined in this manner is a Finsler function of Randers type. Specifically,

$$\alpha_{ij} = \frac{h_{ij}}{1-|W|^2} + \frac{W_i W_j}{(1-|W|^2)^2}, \qquad \beta_i = -\frac{W_i}{1-|W|^2},$$

where $W_i = h_{ij}W^j$. Conversely, it can be shown that for each Randers space there is a corresponding Zermelo problem [11]. We remark in passing that there is yet another equivalent perspective, whereby with each Randers space is associated a conformally stationary spacetime [12].

The preceding discussion implies that if a curve $q:[a,b]\to\mathcal{M}$ is traversed at maximum speed, then the time it takes to complete the journey is given by the Randers length

$$T = \int_a^b F(q(s),\dot q(s))\,\mathrm{d}s,$$

where we wrote $\dot q(s)$ for the derivative with respect to the curve parameter s. If the curve $q(s)$ has the physical parameterisation, that is, $q(s)$ corresponds to the location reached by the maximum speed trajectory at time $s-a$ after setting off from $q(a)$, then $F(q(s),\dot q(s)) = 1$ and $T = b-a$. In other words, curves in the physical parameterisation have unit Randers speed and the passage of time is measured by Randers length. As a consequence, Randers geodesics in the physical parameterisation correspond to solutions of the Zermelo problem. To make this statement more precise, recall that Randers geodesics are curves that locally minimise Randers length. That is, $q:[a,b]\to\mathcal{M}$ is a Randers geodesic if and only if for any $c\in[a,b]$ there exists an interval $I = [c - \varepsilon, c + \varepsilon]$ such that $q|_I$ minimises Randers length among all curves defined on I with the same endpoints. Hence, if endowed with the physical parameterisation, Randers geodesics are the same as curves that locally minimise travel time. Using this equivalence, we can reformulate Theorem 2 of [3] in the following equivalent manner. We write $\mathcal{L}$ for Lie derivative.

Theorem 1. *Assume that the wind vector field W in the Zermelo problem above is an infinitesimal homothety, that is, $\mathcal{L}_W h = \sigma h$ for a constant σ. Then, if $q : (-\varepsilon, \varepsilon) \to \mathcal{M}$ is a locally time-minimising curve, $p(t) = \varphi_t(t, q(t))$ is a Riemannian geodesic of $(\mathcal{M}, h)$, where φ is the flow of $-W$. Conversely, if $p : (-\varepsilon, \varepsilon) \to \mathcal{M}$ is a Riemannian geodesic, $q(t) = \varphi^{-1}(t, p(t))$ is a locally time-minimising curve, where φ^{-1} is the flow of W.*

Notice that the existence of the flow maps on neighbourhoods containing $q(t)$ and $p(t)$, respectively, can be ensured by scaling ε if necessary.

The proof of Theorem 2 of [3] (reformulated here as Theorem 1 above) relies on the geodesic equation for Randers spaces, derived for instance in [2, Chapter 11], which is then verified by explicit calculation—but as we saw above, the Randers geodesics on $(\mathcal{M}, F)$ correspond precisely to the locally time-minimising curves of the Zermelo problem. To exploit this fact, in the above formulation of the theorem we have made direct reference to the solution curves of the Zermelo problem, which suggests that a Riemannian proof, without the derivation of the equation characterising Randers geodesics as a prerequisite, should be possible. Before we proceed with this, let us remark first that the theorem applies in particular to Randers spaces of constant flag curvature, since their wind vector fields are homotheties [11].

To gain an intuition for our Riemannian derivation, it will be instructive to examine a concrete example. For this purpose let us consider a particular problem of time-optimal quantum control. In the quantum Zermelo navigation problem, introduced by Russell & Stepney in [13], one considers a quantum system under the influence of an ambient field characterised by a Hamiltonian operator $\hat{H}_0 \in \mathfrak{su}(N)$, whose Hilbert–Schmidt norm $[\mathrm{tr}(H_0^2)]^{1/2}$ is less than unity (in a suitable unit of energy). The goal is to find the time-dependent control Hamiltonian $\hat{H}_1(t)$ that satisfies the bound $\mathrm{tr}(\hat{H}_1(t)^2) \leq 1$ and achieves, in shortest possible time, the transformation $\hat{U}_I \to \hat{U}_F$ between specified initial and final unitary operators (quantum gates) in $\mathrm{SU}(N)$. It was shown in [14, 15] that the optimal control Hamiltonian takes the simple form

$$\hat{H}_1(t) = \mathrm{e}^{-\mathrm{i}\hat{H}_0 t} \hat{H}_1(0)\, \mathrm{e}^{\mathrm{i}\hat{H}_0 t}.$$

Moreover, the solution $\hat{U}(t)$ of the Schrödinger equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}(t) = -\mathrm{i}(\hat{H}_0 + \hat{H}_1(t))\, \hat{U}(t)$$

emanating from $\hat{U}_I$ is given by

$$\hat{U}(t) = \mathrm{e}^{-\mathrm{i}\hat{H}_0 t}\, \mathrm{e}^{-\mathrm{i}\hat{H}_1(0)t}\, \hat{U}_I$$

on account of the special form of $\hat{H}_1(t)$ above, together with standard results in the interaction-picture analysis of quantum mechanics [16]. Alternatively, one can verify by differentiation of $\hat{U}(t)$ that the relevant Schrödinger equation is indeed satisfied. We now develop an intuitive characterisation of the time-optimal solution $\hat{U}(t)$, which will be useful for later analysis. For this purpose, we first recast the solution in the form

$$\mathrm{e}^{\mathrm{i}\hat{H}_0 t}\, \hat{U}(t) = \mathrm{e}^{-\mathrm{i}\hat{H}_1(0)t}\, \hat{U}_I.$$

Let us abbreviate the expression on the left hand side by $\hat{Z}(t)$. That is, $\hat{Z}(t)$ represents the curve $\hat{U}(t)$ in a frame that is dragged along by the right-invariant vector field on $\mathrm{SU}(N)$ given by $\hat{W}(\hat{U}) = -\mathrm{i}\hat{H}_0\hat{U}$. Clearly, $\hat{Z}(t)$ starts at $\hat{U}_I$ and hits the 'moving target' $\mathrm{e}^{\mathrm{i}\hat{H}_0 t}\hat{U}_F$ at some optimal time T. Moreover, one can check by differentiation that

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{Z}(t) = -\mathrm{i}\, \mathrm{e}^{\mathrm{i}\hat{H}_0 t}\, \hat{H}_1(t)\, \mathrm{e}^{-\mathrm{i}\hat{H}_0 t}\, \hat{Z}(t).$$

This implies that the right-invariant velocity $(\partial_t\hat{Z}(t))\hat{Z}^{-1}(t)$ has unit length in the Hilbert–Schmidt norm no matter how the control $\hat{H}_1(t)$ is chosen, so long as it satisfies the full-throttle condition $\mathrm{tr}(\hat{H}_1(t)^2) = 1$.

To put the matter differently, let us introduce the bi-invariant Riemannian metric γ on $\mathrm{SU}(N)$ whose restriction to $\mathfrak{su}(N)$ coincides with the Hilbert–Schmidt inner product $\gamma(A, B) = \mathrm{tr}(A, B)$. Then the preceding velocity constraint can be expressed in the form $\gamma(\partial_t\hat{Z}(t), \partial_t\hat{Z}(t)) = 1$, or, more succinctly, $|\partial_t\hat{Z}(t)| = 1$. Clearly, any optimal control must meet the full-throttle condition at all times and will thus indeed satisfy $|\partial_t\hat{Z}(t)| = 1$. Hence, since the speed of $\hat{Z}(t)$ is fixed, the only strategy for shortening the time until the moving target is intercepted, is to shorten the path that is traversed in the interim. Therefore, $\hat{Z}(t)$ should be a Riemannian geodesic on $\mathrm{SU}(N)$. Alternatively stated, Randers geodesics in the Schrödinger picture should correspond to Riemannian geodesics in the interaction picture. This conclusion is indeed borne out by the right hand side of the relation $\mathrm{e}^{\mathrm{i}\hat{H}_0 t}\,\hat{U}(t) = \mathrm{e}^{-\mathrm{i}\hat{H}_1(0)t}\,\hat{U}_I$, and is in agreement with Theorem 1, upon noting that $\hat{W}$ induces an isometric flow and the parameter σ appearing in Theorem 1 thus vanishes. A similar intuitive argument can be developed for the time-optimal control of quantum states (rather than gates) in the presence of background fields [17].

A key ingredient in the foregoing example is the method, familiar from mechanics and optimal control theory [18, 19, 20], of switching to a *moving frame*. This strategy can be generalised to the generic case considered in Theorem 1. To see this, let $C(t)$ be a time-varying control, assumed to satisfy $h(C(t), C(t)) = |C(t)|^2 = 1$ at all times. By definition, any trajectory $q(t)$ produced by the control satisfies $\partial_t q(t) = W(q(t)) + C(t)$. Writing $\varphi(t, \cdot) = \varphi_t(\cdot)$ for the flow of $-W$ at time t, we introduce the curve $p(t) = \varphi_t(q(t))$. Intuitively speaking, $p(t)$ represents $q(t)$ in a coordinate frame that is pulled along by the wind. If $q(t)$ is time-optimal between $q(0) = q_I$ and $q(T) = q_F$, then $p(T) = \varphi_T(q_F)$. That is, $p(t)$ intercepts the moving target $\varphi_t(q_F)$ in the shortest possible time. Writing $D\varphi_t$ for the differential of φ_t, we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\,\varphi_t(q(t)) &= -W(p(t)) + D\varphi_t(\dot{q}(t)) \\ &= -W(p(t)) + D\varphi_t(W(\varphi_t^{-1}(p(t))) + C(t)) = D\varphi_t(C(t)), \end{aligned}$$

where in the last step we used the fact that (cf. [21, Proposition 9.41])

$$\frac{\mathrm{d}}{\mathrm{d}t}\,D\varphi_t(W \circ \varphi_t^{-1}) = D\varphi_t([W, W] \circ \varphi_t^{-1}) = 0$$

and therefore that $D\varphi_t(W \circ \varphi_t^{-1}) = W$ for all t. Since by assumption W is an infinitesimal homothety, we have for any vector v that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,h(D\varphi_t(v), D\varphi_t(v)) = -\sigma\, h(D\varphi_t(v), D\varphi_t(v))$$

and hence that $h(D\varphi_t(v), D\varphi_t(v)) = \mathrm{e}^{-\sigma t} h(v, v)$. As a consequence we deduce that

$$h(D\varphi_t(C(t)), D\varphi_t(C(t))) = \mathrm{e}^{-\sigma t}.$$

This shows, in particular, that the speed $|\partial_t p(t)| = \mathrm{e}^{-\sigma t/2}$ is independent of the chosen control. It follows that the time until the moving target is intercepted is a strictly increasing function of the distance traveled, and intuition thus dictates that if $q(t)$ is a time-optimal curve between q_I and q_F, then $p(t)$ should be a geodesic. To prove Theorem 1 we only need to make this intuition rigorous.

Proof. Let $q : [-\varepsilon, \varepsilon] \to \mathcal{M}$ be locally time-minimising. Then there exists for each $s \in (-\varepsilon, \varepsilon)$ a $\delta > 0$ such that $q_s(t) = q(s+t)$ solves the Zermelo problem between $q_I = q(s)$ and $q_F = q(s + \delta)$. Hence, the curve $p_s(t) = \varphi_t(q_s(t))$ defined on $[0, \delta]$ (existence of the flow map can be ensured by scaling δ, if necessary) has fixed speed, as in the discussion above, and intercepts the moving target $\varphi_t(q_F)$ at the shortest possible time δ. It must therefore be a geodesic on $[0, \delta]$, since otherwise one could reach the point $\varphi_\delta(q_F)$ before δ and then navigate towards the target to obtain an earlier rendezvous. But clearly, if $p_s(t)$ is a geodesic on $[0, \delta]$, then $p(t) = \varphi_t(q(t)) = \varphi_s(p_s(t - s))$ is a geodesic on $[s, s + \delta]$, bearing in mind that φ_s scales the metric by a constant. Since s was arbitrary, we conclude that $p(t)$ is a geodesic on $(-\varepsilon, \varepsilon)$, as required.

Conversely, suppose that $p(t) = \varphi_t(q(t))$ is a geodesic on $(-\varepsilon, \varepsilon)$. Just as before, one can take any $s \in (-\varepsilon, \varepsilon)$ and find a $\delta > 0$ such that $p_s(t) = p(s+t)$ is length-minimising between $p(s)$ and $p(s+\delta)$. The claim is that $q_s(t) = q(s+t)$ is a time-minimising trajectory between $q_s(0) = q(s)$ and $q_s(\delta) = q(s + \delta)$. Suppose this were not the case. That is, there exists a curve $Q(t)$ commencing at $q(s)$ and arriving at $q(s + \delta)$ at time $\delta' < \delta$. If we set $P(t) = \varphi_t(Q(t))$ for $t \leq \delta'$ then our earlier calculations show that this curve commences at $q(s)$, has speed $\mathrm{e}^{-\sigma t}$ and intercepts the moving target $\varphi_t(q(s + \delta))$ at time $\delta' < \delta$. For $t > \delta'$, define $P(t)$ to move from the interception point along the flow line of $-W$, its speed still satisfying the same constraint. By choosing δ small enough, one can ensure that $P(t)$ moves at above wind speed and thus arrives, still before time δ, at $\varphi_\delta(q(s + \delta))$. Observing that $P(t)$ and $\varphi_s^{-1} p_s(t) = \varphi_t(q(s + t))$ have the same speed and that the latter curve takes a longer time, δ, to arrive at $\varphi_\delta(q(s + \delta))$, one concludes that the distance travelled by $P(t)$ must be shorter. Upon applying φ_s to both paths, this implies that $\varphi_s(P(t))$, connecting $p_s(0) = p(s)$ and $\varphi_{s+\delta}(q(s + \delta)) = p_s(\delta)$, is shorter than $p_s(t)$, in contradiction to the length-minimising property of the latter curve. This completes the proof of Theorem 1. $\square$

To summarise, we have presented a novel treatment of geodesics in a certain class of Randers spaces including those of constant flag curvatures. Taking inspiration from a recent analysis of the quantum Zermelo problem in the interaction picture, we achieved a formulation that relied exclusively on standard Riemannian methods, thus bypassing a more involved Finslerian analysis of Randers spaces. Simplifications of this nature serve an important purpose in making results that have been formulated in a specialised mathematical language accessible to a wider audience. At the same time our analysis offers an instance whereby the strength of physical intuition contributes to transparency in abstract mathematical reasoning. As a final remark we mention that Randers geodesics in spaces of constant flag curvature correspond to null geodesics of conformally flat spacetimes [12]. It would be interesting to analyse Theorem 1 in this spacetime context.

References

[1] Rund, H. 1959 *The Differential Geometry of Finsler Spaces* (Berlin: Springer).

[2] Bao, D., Chern, S.-S. & Shen, Z. 2000 *Introduction to Riemann-Finsler geometry.* (New York: Springer)

[3] Robles, C. 2007 Geodesics in Randers spaces of constant curvature. *Trans. Am. Math. Soc.* **359**, 1633.

[4] Randers, G. 1941 On an asymmetrical metric in the four-space of general relativity. *Phys. Rev.* **59**, 195.

[5] Ingarden, R. S. 1957 On the geometrically absolute optical representation in the electron microscope. *Trav. Soc. Sci. Lett. Wrocaw* B**45**, 1.

[6] Gibbons, G. W. & Warnick, C. M. 2011 The geometry of sound rays in a wind. *Contemp. Phys.* **52**, 197.

[7] Luneburg, R. K. 1964 *Mathematical Theory of Optics.* (Berkeley: University of California Press)

[8] Shen, Z. 2003 Finsler metrics with K=0 and S=0. *Canad. J. Math.* **55**, 112-132.

[9] Zermelo, E. 1931 Über das Navigationsproblem bei ruhender oder veränderlicher Windverteilung. *Ztschr. f. angew. Math. und Mech.* **11**, 114.

[10] Carathéodory, C. 1935 *Variationsrechnung und Partielle Differentialgleichungen erster Ordnung.* (Berlin: B. G. Teubner).

[11] Bao, D., Robles, C. & Shen, Z. 2004 Zermelo navigation on Riemannian manifolds. *J. Diff. Geom.* **66**, 391.

[12] Gibbons, G. W., Herdeiro, C. A. R., Warnick, C. M. & Werner, M. C. 2009 Stationary metrics and optical Zermelo-Randers-Finsler geometry. *Phys. Rev. D.* **79**, 044022.

[13] Russell, B. & Stepney, S. 2014 Zermelo navigation and a speed limit to quantum information processing. *Phys. Rev.* A**90**, 012303.

[14] Russell, B. & Stepney, S. 2015 Zermelo navigation in the quantum brachistochrone. *J. Phys. A: Math. Theor.* **48**, 115303.

[15] Brody, D. C. & Meier, D. M. 2015 Solution to the quantum Zermelo navigation problem. *Phys. Rev. Lett.* **115**, 100502

[16] Sunakawa, S. 1991 *Quantum Mechanics.* (Tokyo: Iwanami)

[17] Brody, D. C., Gibbons, G. W. & Meier, D. M. 2015 Time-optimal navigation through quantum wind. *New J. Phys.* **17**, 033048.

[18] Jurdjevic, V. 1999 Optimal control, geometry, and mechanics. In *Mathematical Control Theory*, pp. 227–267. (New York: Springer)

[19] Jurdjevic, V. 1997 *Geometric control theory.* (Cambridge: Cambridge University Press)

[20] Agrachev, A. A., and Sachkov, Y. 2004 *Control theory from the geometric viewpoint.* (Berlin: Springer)

[21] Lee, J. M. 2013 *Introduction to Smooth Manifolds*, Second Edition. (New York: Springer)