DRAFT VERSION. Part III: Tracking in Orientation Scores Optimal Paths for Variants of the 2D and 3D Reeds-Shepp Car

R. Duits · E.J. Bekkers · J.M. Portegies

Abstract We consider a PDE-based approach for finding minimal paths for the Reeds-Shepp car. In the model we minimize a (data-driven) functional involving both curvature and length penalization, with several generalizations. Our approach encompasses the two and three dimensional variants of this model, state dependent costs (modeling mobility), and moreover, the possibility of removing the reverse gear of the vehicle.

Study objectives on the minimal path model:

- We study exact solutions of the model in the case of uniform cost (i.e. constant mobility) for the 2D model. For details on how this extends to the 3D model see [28].
- We prove both global and local controllability results of the models.
- We compute distance maps w.r.t. highly anisotropic Finsler functions, which approximate the singular (pseudo)-metrics underlying the model. The computation is done via eikonal equations on the manifold $\mathbb{R}^d \times \mathbb{S}^{d-1}$. We justify the use of our approximating metrics by proving convergence results.
- We compute solutions (minimizing geodesics) of the data-driven variant of the model by a Fast-Marching (FM) method, building on Mirebeau [56,55]. The FM method is based on specific discretization stencils which are adapted to the preferred directions of the metric and obey a generalized acuteness property. The shortest paths can be found with a gradient descent method on the distance map, which we formalize in a theorem.
- We compare, for uniform costs, numeric solutions to exact solutions by Duits & Sachkov [28,24,57].
- We apply our curve optimization model in $\mathbb{R}^d \times \mathbb{S}^{d-1}$ with data-driven cost for the extraction of complex tubular structures from medical images, e.g. crossings, and incomplete data due to occlusions or low contrast. Numerical experiments show the high potential of our method in two applications: vessel tracking in retinal images for the case d = 2 by Bekkers et al. [9], and brain connectivity measures from diffusion weighted MRI-data for the case d = 3.

Literature: This section of the lecture notes is primarily based on the following scientific works [9,26, 24,27,28,56,57,66,67].

Learning objectives:

- Understand the procedure of geodesic tracking in orientation scores.
- Know the procedure for globally optimal path optimization in position and orientation space. This procedure is an intrinsic gradient descent on distance maps obtained via the eikonal equation on $\mathbb{R}^d \times S^{d-1}$.
- Understand the concept of local and global controllability.
- Know how to derive exact solutions for the uniform cost and d = 2, via the Pontryagin maximum principle.
- Know how to set up the Hamiltonian system for the model.
- Know how to derive all preservations laws of the Hamiltonian system.
- Understand the notion of topological interest points such as keypoints and cusps, and know how to compute them.

- Understand the notion of Maxwell points (where global optimality is lost) and conjugate points (where local optimality is lost).
- Understand the general numerical procedures for solving the eikonal equations.
- Be able to apply vessel tracking in orientation scores via a prepared *Mathematica notebook* and understand the connection of the code to the theory.

NB:

Technical parts and exercises with extra material are indicated in blue and may be skipped: The black text forms the core of the course and its exam. Regular exercises are indicated in red. The black text and the red exercises nor do not rely on the blue text nor do they rely on the blue exercises.

1 Introduction

Shortest paths in position and orientation space are central in this paper. Dubins describes in [23] the problem of finding shortest paths for a car in the plane between initial and final points and direction, with a penalization on the radius of curvature, for a car that has no reverse gear. Reeds and Shepp consider in [62] the same problem, but then for a car that does have the possibility for backward motion. In both papers, the focus lies on describing and proving the general shape of the optimal paths, without giving explicit solutions for the shortest paths.

This can be considered a curve optimization problem in the space $\mathbb{R}^2 \times (\mathbb{R}/2\pi\mathbb{Z})$, equipped with the natural Euclidean metric but only among curves $\gamma(t) = (x(t), y(t), \theta(t))$ subject to the constraint that $(\dot{x}(t), \dot{y}(t))$ is proportional to $(\cos \theta(t), \sin \theta(t))$. Formulating the problem this way, it becomes one of the simplest examples of sub-Riemannian (SR) geometry: the tangent vector $\dot{\gamma}(t)$ is constrained to remain in the span of $(\cos \theta(t), \sin \theta(t), 0)$ and (0, 0, 1). The SR curve optimization problem and the properties of its geodesics in $\mathbb{R}^2 \times \mathbb{S}^1$ have been studied and applied in image analysis by [60, 18, 24, 12, 57, 52, 2, 11], whereas a complete and optimal synthesis for the geometric control problem on $\mathbb{R}^2 \times \mathbb{S}^1$ with uniform cost is provided by Sachkov [66]. Properties of SR geodesics in $\mathbb{R}^d \times \mathbb{S}^{d-1}$ with d = 3 have been studied in [28] and for general d in [27].

On the left in Fig. 1, we show an example of an optimal path between two points in $\mathbb{R}^2 \times \mathbb{S}^1$. The projection on \mathbb{R}^2 of this curve shows a cusp, after which the car continues in reverse (the red part of the line). From the perspective of image analysis applications this is undesirable and it is a valid question what the optimal paths are if cusps and reverse gear are not allowed. In this paper, similar to the difference between the Dubins car and the Reeds-Shepp car, we also consider this variant: it can be accounted for by requiring that the spatial propagation is forward. This variant falls outside the SR framework and requires asymmetric Finsler geometry instead.

Furthermore, we would like to extend the metric using two data-driven factors that can vary with position and orientation. This can be used to compute shortest paths for a car, where for example road conditions and obstacles are taken into account. In [9] it is shown that this approach is useful for tracking vessels in retinal images. Likewise, the 3D variant of the problem provides a basis for algorithms for blood vessel detection in 3D MRA data, or detection of shortest paths and quantification of structural connectivity in 5D diffusion weighted MRI data of the brain.

1.0.1 A distance function and the corresponding shortest paths on $\mathbb{R}^d \times \mathbb{S}^{d-1}$

We fix the dimension $d \geq 2$, and let $\mathbb{M} := \mathbb{R}^d \times \mathbb{S}^{d-1}$ be the 2d-1 dimensional manifold of positions and orientations. We use a Finsler metric¹ on the tangent bundle of \mathbb{M} , $\mathcal{F} : T(\mathbb{M}) \to [0, +\infty]$, of which specific properties are discussed later, to define a geometry on \mathbb{M} . Any such Finsler metric \mathcal{F} induces a measure of length Length_{\mathcal{F}} on the class of paths with Lipschitz regularity, defined as²

$$\operatorname{Length}_{\mathcal{F}}(\gamma) := \int_0^1 \mathcal{F}(\gamma(t), \dot{\gamma}(t)) \, \mathrm{d}t,$$

¹ We use the common terminology of Finsler metric, although \mathcal{F} is also called 'Finsler function', since formally it is not a metric.

² In contrast to previous works [28,12,9,51,24] we parameterize such that the time integration stays on [0, 1], and t > 0 is *not* a priori reserved (unless explicitly stated otherwise) for arclength parametrization (which satisfies $\mathcal{F}_{\gamma(t)}(\dot{\gamma}(t)) = 1$).



Fig. 1 Top: Example of a shortest path with (left) and without (right) reverse gear in $\mathbb{R}^2 \times S$ and its projection on \mathbb{R}^2 . The black arrows indicate the begin and end condition in the plane. In the left figure, the projection of the path has a cusp, after which the line is covered backwards (the red part). On the right, backward motion is not possible. Instead, according to our model, the shortest path is a concatenation of an in-place rotation (green), a SR geodesic, and again an in-place rotation. Bottom: corresponding control sets as defined in (7) for the allowed velocities at each position and orientation, with $B_{\mathcal{F}_0}$ on the left and $B_{\mathcal{F}_0^+}$ on the right.

with the convention $\dot{\gamma}(t) := \frac{d}{dt}\gamma(t)$. The path is said to be *normalized* w.r.t. \mathcal{F} iff $\mathcal{F}(\gamma(t), \dot{\gamma}(t)) = \text{Length}_{\mathcal{F}}(\gamma)$ for all $t \in [0, 1]$. Any Lipschitz continuous path of finite length can be normalized by a suitable reparametrization. Finally, the pseudo-distance $d_{\mathcal{F}} : \mathbb{M} \times \mathbb{M} \to [0, +\infty]$ is defined for all $\mathbf{p}, \mathbf{q} \in \mathbb{M}$ by

$$d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) := \inf\{ \operatorname{Length}_{\mathcal{F}}(\gamma) \mid \gamma \in \Gamma, \, \gamma(0) = \mathbf{p}, \, \gamma(1) = \mathbf{q} \},$$
(1)

with $\Gamma := \text{Lip}([0, 1], \mathbb{M})$. Normalized minimizers of (1) are called minimal geodesics from **p** to **q** w.r.t. \mathcal{F} . For certain pairs (**p**, **q**) these minimizers may not be unique, and these points are often of interest.

Definition 1 (Maxwell point) Let $\mathbf{p}_S \in \mathbb{M}$ be a fixed point source and $\gamma \in \Gamma$ a geodesic connecting \mathbf{p}_S with $\mathbf{q} \in \mathbb{M}$, $\mathbf{q} \neq \mathbf{p}_S$. Then \mathbf{q} is a Maxwell point if there exists another extremal path $\tilde{\gamma} \in \Gamma$ connecting \mathbf{p}_S and \mathbf{q} , with Length_{\mathcal{F}}(γ) = Length_{\mathcal{F}}($\tilde{\gamma}$). If \mathbf{q} is the first point (distinct from \mathbf{p}_S) on γ where such $\tilde{\gamma}$ exists, then \mathbf{q} is called the first Maxwell point. The curves $\gamma, \tilde{\gamma}$ lose global optimality after the first Maxwell point.

1.1 Geometry of the Reeds-Shepp model

We introduce the metric \mathcal{F}_0 underlying the Reeds-Shepp car model, and the metric \mathcal{F}_0^+ corresponding to the variant without reverse gear. Let $(\mathbf{p}, \dot{\mathbf{p}}) \in T(\mathbb{M})$ be a pair consisting of a point $\mathbf{p} \in \mathbb{M}$ and a tangent vector $\dot{\mathbf{p}} \in T_{\mathbf{p}}(\mathbb{M})$ at this point. The physical and angular components of a point $\mathbf{p} \in \mathbb{M}$ are denoted by $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{n} \in \mathbb{S}^{d-1}$, and this convention carries over to the tangent:

$$\mathbf{p} = (\mathbf{x}, \mathbf{n}), \qquad \qquad \dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in T_{\mathbf{p}}(\mathbb{M}).$$



Fig. 2 left; lifting of planar curves into position and orientation space, right; lifting of 2D image data to data on position and orientation space (via an orientation score transform) yielding the mobility $\mathbf{C}(x, y, \theta)$ in orange. In this lifted data (orientation score), crossings are disentangled and tracking is performed.

We say that $\dot{\mathbf{x}}$ is proportional to \mathbf{n} , that we write as $\dot{\mathbf{x}} \propto \mathbf{n}$, if and only if there exists a $\lambda \in \mathbb{R}$ such that $\dot{\mathbf{x}} = \lambda \mathbf{n}$. We define

$$\mathcal{F}_{0}(\mathbf{p}, \dot{\mathbf{p}})^{2} := \begin{cases} \mathcal{C}_{1}^{2}(\mathbf{p}) |\dot{\mathbf{x}} \cdot \mathbf{n}|^{2} + \mathcal{C}_{2}^{2}(\mathbf{p}) |\|\dot{\mathbf{n}}\|^{2} & \text{if } \dot{\mathbf{x}} \propto \mathbf{n}, \\ +\infty & \text{otherwise.} \end{cases}$$

$$\mathcal{F}_{0}^{+}(\mathbf{p}, \dot{\mathbf{p}})^{2} := \begin{cases} \mathcal{C}_{1}^{2}(\mathbf{p}) |\dot{\mathbf{x}} \cdot \mathbf{n}|^{2} + \mathcal{C}_{2}^{2}(\mathbf{p}) |\|\dot{\mathbf{n}}\|^{2} & \text{if } \dot{\mathbf{x}} \propto \mathbf{n} \text{ and} \\ & \dot{\mathbf{x}} \cdot \mathbf{n} \ge 0, \\ +\infty & \text{otherwise.} \end{cases}$$

$$(2)$$

$$(3)$$

Here $\|\cdot\|$ denotes the norm and "." the dot product of the Euclidean space \mathbb{R}^d . The functions C_1 and C_2 are assumed to be continuous on \mathbb{M} , and uniformly bounded from below by a positive constant $\delta > 0$. In applications, C_1 and C_2 are chosen so as to favor paths which remain close to regions of interest, e.g. along blood vessels in retinal images. Note that their physical units are distinct: if one wishes $d_{\mathcal{F}}$ to have the dimension [T] of a travel time, then C_1^{-1} is a physical, (strictly) spatial velocity [Length] $[T]^{-1}$, and C_2^{-1} is an angular velocity $[\operatorname{Rad}][T]^{-1}$. For simplicity one often sets $C_1 = \xi C_2$, where $\xi^{-1} > 0$ is a unit of spatial length. The special case $C_1(\mathbf{p}) = \xi C_2(\mathbf{p}) = \xi$ for all $\mathbf{p} \in \mathbb{M}$ is referred to as the uniform cost case.

Often in applications, such as vessel tracking in orientation scores [9], we consider the case $C_1(\mathbf{p}) = \xi C(\mathbf{p})$ and $C_2(\mathbf{p}) = C(\mathbf{p})$. In this case $C(\mathbf{x}, \mathbf{n})$ is a general 'mobility' for the lifted Reeds-Shepp car to be at position $\mathbf{x} \in \mathbb{R}^d$ with orientation $\mathbf{n} \in S^{d-1}$ in the orientation score. If it is low the car can proceed with low local cost, if it is high the car can proceed with high local costs. For now, the precise choice of continuously differentiable $C : \mathbb{R}^d \times S^{d-1} \to [\delta, \infty]$ is not crucial, you can assume it to be given. To get an impression of how such a mobility looks like in practice, see the orange densities in Fig. 2.

1.2 The eikonal equation and the fast marching algorithm

We compute the distance map to a point source on a volume using the relation to eikonal equations. Let $\mathbf{p}_{\mathrm{S}} \in \mathbb{M}$ be an arbitrary source point, and let U be the associated distance function

$$U(\mathbf{p}) := d_{\mathcal{F}}(\mathbf{p}_{\mathrm{S}}, \mathbf{p}). \tag{4}$$

Then U is the unique viscosity solution [21,20] to the eikonal PDE:

$$\begin{cases} \mathcal{F}^*(\mathbf{p}, \mathrm{d}U(\mathbf{p})) = 1 & \text{for all } \mathbf{p} \in \mathbb{M} \setminus \{\mathbf{p}_{\mathrm{S}}\}, \\ U(\mathbf{p}_{\mathrm{S}}) = 0. \end{cases}$$
(5)

Here \mathcal{F}^* is the dual metric of \mathcal{F} and dU is the differential of the distance map U. However, for these relations to hold, and for numerical discretization to be practical, \mathcal{F} should be at least continuous³. We therefore propose in Section 2.3 for both \mathcal{F}_0 and \mathcal{F}_0^+ an approximating metric, that we denote by $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$, respectively, that are continuous and converge to \mathcal{F}_0 and \mathcal{F}_0^+ as $\varepsilon \to 0$. The approximating metrics correspond to a highly anisotropic Riemannian and Finslerian metric, rather than a sub-Riemannian or sub-Finslerian metric. The metric \mathcal{F}_{ϵ} is in line with previous approximations [18,9,67] for the case d = 2.

We will design (in Section 7) a monotone and causal discretization scheme for the static Hamilton-Jacobi PDE (5), which allows to apply efficient, single pass Fast-Marching Algorithm [74]. Let us emphasize that designing a causal discretization scheme for (5) is non-trivial, because its local connectivity needs to be obey an *acuteness property* [70,76] depending on the geometry defined by \mathcal{F} . We provide constructions for the metrics $\mathcal{F}_{\varepsilon}$ or $\mathcal{F}_{\varepsilon}^+$ of interest, based on the earlier works [56,55].

In all cases, the reader should keep in mind the key practical advantages of considering eikonal equations (and the induced wavefront propagation) on $\mathbb{M} = \mathbb{R}^d \times S^{d-1}$ instead of considering them on position space \mathbb{R}^d only:

- Wavefront propagations in position space leaks at crossings, whereas wavefront propagation in M works generically at crossings (regardless the type of crossing).
- Wavefront propagations in position space often requires cumbersome multi-valued solutions [69], whereas wavefront propagation does not require such post-fixes.
- Wavefront propagations in position space typically fails at bifurcations and wavefront propagation in \mathbb{M} via \mathcal{F}_0^+ deals with bifurcations.
- Wavefront propagations do not intrinsically encode alignment of local orientations requiring ad-hoc modeling, whereas wavefront propagation in position and orientation space include this⁴.

See Fig. 3 for a visual impression on these advantages of wavefront propagation. See Fig. 4 where our vessel tracking algorithm in metric space model $(\mathbb{M}, d^+_{\mathcal{F}_0})$ performs better than a state-of-the-art industrial method acting directly in the image domain.

1.3 Shortest Paths and Minimal Distances in Medical Images

The application of the proposed method for finding shortest paths has been shown to be useful for vesseltracking in retinal images [9], see Fig. 5 (top, right). A related approach using fast marching with elastica functionals can be found in [15,16]. The sub-Riemannian approach by Bekkers et al. [9] concerns the twodimensional Reeds-Shepp car model with reverse gear, where 2D gray-scale images are first lifted to an orientation score defined on the higher dimensional manifold $\mathbb{R}^2 \times \mathbb{S}^1$. There, the combination of the sub-Riemannian metric, the cost function derived from the orientation score, and the numerical fast-marching solver, provided a solid approach to accurately track vessels in challenging sets of images.

For this reason, we apply the same strategy to diffusion-weighted MRI data. For such images, a signal related to the amount of diffusion is measured, which in the case of neuroimages is considered to be related to the underlying cell-structure. The images can in a natural way be considered to have domain $\Omega \subset \mathbb{R}^3 \times \mathbb{S}^2$. Fig. 5 (bottom) illustrates such images. On the left we use a glyph visualization, that shows

³ From a theoretical standpoint, one may rely on the notion of discontinuous viscosity solution [7]. But this concept is outside of the scope of this paper, and in addition it forbids the use of a singleton $\{\mathbf{p}_S\}$ as the target set.

⁴ Akin to cortical models for line perception in the primary visual cortex of the human visual system [18,60].



Fig. 3 Top: An invertible orientation score is an image representation that provides a complete overview of how the image is decomposed out of local orientations. Botttom: Conventional wavefront propagation in images (in red) typically leaks at crossings, whereas the wavefront propagation in orientation scores (in green) does not suffer from this fundamental complication. A minimum intensity projection over orientation gives the optimal fronts in the image. The orange densities denote mobility in the score. This regulates the speed of the propagation of the green opaque spheres; left: propagation symmetric sub-Riemannian spheres (in $(\mathbb{M}, d_{\mathcal{F}_0})$, right: propagation asymmetric Finsler spheres (in $(\mathbb{M}, d_{\mathcal{F}_1})$.

a surface for each grid point, where the distance from the surface to the corresponding grid point \mathbf{x} is proportional to the data-value $U(\mathbf{x}, \mathbf{n})$ and the coloring is related to the orientation $\mathbf{n} \in \mathbb{S}^2$.

A large number of tractography methods exist, that are designed to estimate/approximate the fiber paths in the brain based on dMRI data. Most of these methods construct tracks that locally follow the structure of the data, see e.g. [73,22] or references in [45]. More related to our approach are geodesic methods, that have the advantage that they minimize a functional, and thereby are less sensitive to noise and provide a certain measure of connectivity between regions. These methods can be based on diffusion tensors in combination with Riemannian geometry on position space, e.g. [42,48,44]. One can also make use of the more general Finsler geodesic tracking to include directionality [53,54], and use high angular resolution data (HARDI), examples of which can be found in [69,5]. Recently, a promising method has been proposed, based on geodesics on the full position and orientation space using a data-adaptive Riemannian metric [59]. We also work on this joint space of positions and orientations, but use either Riemannian or asymmetric Finsler metrics that are highly anisotropic, that we solve by a numerical fast marching method that is able to deal with this high anisotropy. We show on artificial datasets how our method can be employed to give global minimizers between two regions w.r.t the imposed Finsler metric, and that these paths correctly follow the bundle structure.



Fig. 5 Challenges and applications. Top row: the case d = 2, with a toy problem for finding the shortest way with or without reverse gear (blue and red, respectively) to the exit in Centre Pompidou (top left) and a vessel tracking problem in a retinal image. Bottom row: the case d = 3, connectivity in (simulated) dMRI data. Left: visualization of a dataset with two crossing bundles without torsion, with a glyph visualization of the data in $\mathbb{R}^3 \times \mathbb{S}^2$ and a magnification of one such glyph, indicating two main fiber directions. Right: the spatial configuration in \mathbb{R}^3 of bundles with torsion in an artificial dataset on $\mathbb{R}^3 \times \mathbb{S}^2$.

1.4 Contributions and Outline

The extension to 3D of the Reeds-Shepp metric and the adaptation to model shortest paths for cars that cannot move backwards are new and provide an interesting collection of new theoretical and practical results:

 In Theorem 1 we show that the Reeds-Shepp model is globally and locally controllable, and that the Reeds-Shepp model without reverse gear is globally but not locally controllable. Hence the distance map loses continuity.

- We introduce regularizations $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$ of the Reeds-Shepp metrics \mathcal{F}_0 and \mathcal{F}_0^+ , which make our numerical discretization possible. We show that both the corresponding distances converge to $d_{\mathcal{F}_0}$ and $d_{\mathcal{F}_{\varepsilon}^+}$ as $\varepsilon \to 0$ and the minimizing curves converge to the ones for $\varepsilon = 0$, see Theorem 2.
- We present and prove for d = 2 and uniform cost a theorem that describes the occurrence of cusps for the SR Reeds-Shepp model \mathcal{F}_0 , and that using \mathcal{F}_0^+ leads to geodesics that are a concatenation of purely angular motion, a sub-Riemannian geodesic without cusps and again a purely angular motion. We call the positions where in-place rotation (or purely angular motion) takes place *keypoints*. For uniform cost, we show that the only possible keypoints are the begin and end point, and for many end conditions we can describe how this happens. The precise theoretical statement and proof are found in Theorem 3.
- Furthermore, we show in Theorem 4 how the geodesics can be obtained from the distance map, for a general Finsler metric, and in the more specific cases that we use in this paper. For our cases of interest, we show that backtracking of geodesics is either done via a single intrinsic gradient descent (for the models with reverse gear), or via two intrinsic gradient descents (for the model without reverse gear).
- For our numerical experiments we make use of a Fast-Marching implementation, for d = 2 introduced in [56]. In Section 6 we give a summary of the numerical approach for d = 3, but a detailed discussion of the implementation and an evaluation of the accuracy of the method is beyond the scope of this paper, and will follow in future work. For d = 2, we show an extensive comparison between the models with and without reverse gear for uniform cost, to illustrate the useful principle of the keypoints, and to show the qualitative difference between the two models. In examples with non-uniform cost, see for example the top row of Fig. 5, we show that the model places the keypoints optimally at corners/bifurcations in the data, where the in-place rotation forms a natural, *automatic* 're-initialization' of the tracking. For d = 3, we give several examples to show the influence of the model parameters, in particular the cost parameter. The examples indicate that the method adequately deals with crossing or kissing structures.

Outline In Section 2, we give a detailed overview of the theoretical results of the paper. The theorems 1, 3 and 4 are discussed and proven in Sections 4, 5 and 6, respectively. The reader who is primarily interested in the application of the methods may choose to skip these three sections. The proof of Theorem 2 is given in Appendix A. We discuss the numerics briefly in Section 7. Section 8 contains all experimental results. Conclusion and discussion follow in Section 9. For an overview of notations, Appendix F may be helpful.

2 Main results

In this section, we state formally the mathematical results announced in Section 1. Some preliminaries regarding the distance function are introduced in the Section below. Results regarding the exact Reeds-Shepp car models are gathered in Section 2.2. The description of the approximate models and the related convergence results appear in Section 2.3. Analysis of special interest points (cusps and keypoints) are done in Section 2.4. Results on the eikonal equation, and subsequent backtracking of optimal geodesics via intrinsic gradients is presented in Section 2.5.

2.1 Preliminaries on the (pseudo) Distance Function and Underlying Geometry

Geometries on the manifold of states $\mathbb{M} = \mathbb{R}^d \times \mathbb{S}^{d-1}$ are defined by means of Finsler metrics which are functions $\mathcal{F} : T(\mathbb{M}) \to [0, +\infty]$. On each tangent space, the metric should be 1-homogeneous, convex and quantitatively non-degenerate with a uniform constant $\delta > 0$: for all $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}$, $\dot{\mathbf{p}}, \dot{\mathbf{p}}_0, \dot{\mathbf{p}}_1 \in T_{\mathbf{p}}(\mathbb{M})$, and $\lambda \geq 0$:

$$\mathcal{F}(\mathbf{p}, \lambda \dot{\mathbf{p}}) = \lambda \mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}),$$

$$\mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}_0 + \dot{\mathbf{p}}_1) \le \mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}_0) + \mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}_1),$$

$$\mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}) \ge \delta \sqrt{\|\dot{\mathbf{x}}\|^2 + \|\dot{\mathbf{n}}\|^2}.$$
 (6)

A weak regularity property is required as well, see the next remark. The induced distance $d_{\mathcal{F}}$, defined in (1), obeys $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) = 0$ iff $\mathbf{p} = \mathbf{q}$, and obeys the triangle inequality. However, unlike a regular distance,

 $d_{\mathcal{F}}$ needs not be finite, or continuous, or symmetric in its arguments. Note that \mathcal{F}_0 and \mathcal{F}_0^+ as defined in (2) and (3), respectively, indeed satisfy the properties in (6).

Remark 1 In contrast to the more common definition of Finsler metrics, we will not assume the Finsler metric to be smooth on $T(\mathbb{M})$, but use a weaker condition instead. Following [14], we require that the sets

$$\mathcal{B}_{\mathcal{F}}(\mathbf{p}) := \{ \dot{\mathbf{p}} \in T_{\mathbf{p}} \mathbb{M} \, | \, \mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}) \le 1 \}$$
(7)

are closed and vary continuously with respect to the point $\mathbf{p} \in \mathbb{M}$ in the sense of the Hausdorff distance. The sets $\mathcal{B}_{\mathcal{F}}(\mathbf{p})$ are illustrated in Fig. 1 for the models of interest. The condition implies that a minimal path exists from \mathbf{p} to $\mathbf{q} \in \mathbb{M}$ whenever $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q})$ is finite, and is used to prove convergence results in Appendix A.

A common technique in optimal control theory is to reformulate the shortest path problem defining the distance $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q})$ into a time optimal control problem. That is, for $p \in [1, \infty]$ one has by Hölder's (in)equality, time re-parametrization, and by 1-homogeneity of \mathcal{F} in its 2nd entry, that:

$$d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) = \inf\{\int_{0}^{1} \mathcal{F}(\gamma(t), \dot{\gamma}(t)) dt \mid \gamma \in \Gamma_{1}, \ \gamma(0) = \mathbf{p}, \gamma(1) = \mathbf{q}\}$$

$$= \inf\{(\int_{0}^{1} |\mathcal{F}(\gamma(t), \dot{\gamma}(t))|^{p} dt)^{\frac{1}{p}} | \gamma \in \Gamma, \gamma(0) = \mathbf{p}, \gamma(1) = \mathbf{q}\}$$

$$= \inf\{T \ge 0 \mid \exists \gamma \in \Gamma_{T}, \ \gamma(0) = \mathbf{p}, \ \gamma(T) = \mathbf{q}, \forall_{t \in [0, T]} \dot{\gamma}(t) \in \mathcal{B}_{\mathcal{F}}(\gamma(t))\},$$

$$(8)$$

where $\Gamma_T := \operatorname{Lip}([0, T], \mathbb{M})$, and with $\mathcal{B}_{\mathcal{F}}(\mathbf{p})$ as defined in (7).

Exercise 1 Show that the (pseudo)-distance $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q})$ in (8) does not depend on the choice of (monotonic) parameterization of the curve $\gamma(\cdot)$.

Exercise 2 Prove both equalities in (8).

The latter reformulation is used in Appendix A to prove convergence results via closedness of controllable paths and Arzela-Ascoli's theorem, based on a general result originally applied to Euler elastica curves in [14].

In the special case $\mathcal{F} = \mathcal{F}_0$ the geodesics are SR geodesics. If $\mathcal{F} = \mathcal{F}_0$ the Finsler function is obtained by the square root of quadratic form associated to a SR metric $\mathcal{G}_0|_{\mathbf{p}}(\cdot, \cdot) = \mathcal{F}_0(\mathbf{p}, \cdot)^2$ on a SR manifold $(\mathbb{M}, \Delta, \mathcal{G}_0)$, where $\Delta \subset T(\mathbb{M})$ is a strict subset of allowable tangent vectors that comes along with the horizontality constraint

$$\dot{\mathbf{x}}(t) = (\dot{\mathbf{x}}(t) \cdot \mathbf{n}(t))\mathbf{n}(t), \qquad \forall t \in [0, 1],$$
(9)

that arises from (2). For details on the case d = 2 see [12,66], for d = 3 see [28].

Finally, we note that for the uniform cost case $(\xi^{-1}C_1 = C_2 = 1)$, the problem is covariant with respect to rotations and translations. For the data-driven case, such covariance is only obtained when simultaneously rotating the data-driven cost factors C_1, C_2 . Therefore, only in the uniform cost case, for d = 2, 3, we shall use a reference point ('the origin') $\mathbf{e} \in \mathbb{R}^d \times \mathbb{S}^{d-1}$. To adhere to common conventions we use

$$\mathbf{e} = (\mathbf{0}, \mathbf{a}) \in \mathbb{R}^{d} \times \mathbb{S}^{d-1}, \text{ with}$$
$$\mathbf{a} := (1, 0)^{T} \quad \text{if } d = 2 \text{ and}$$
$$\mathbf{a} := (0, 0, 1)^{T} \quad \text{if } d = 3.$$
(10)

2.2 Controllability of the Reeds-Shepp model

A model $(\mathbb{M}, \mathcal{F})$ is globally controllable if the distance $d_{\mathcal{F}}$ takes finite values on $\mathbb{M} \times \mathbb{M}$. In Theorem 1 we show that this is indeed the case for $\mathcal{F} = \mathcal{F}_0$ and $\mathcal{F} = \mathcal{F}_0^+$, given in (2) and (3). Local controllability is satisfied when $d_{\mathcal{F}}$ satisfies a certain continuity requirement: if $\mathbf{p} \to \mathbf{q} \in (\mathbb{M}, \|\cdot\|)$, with $\|\cdot\|$ denoting the standard (flat) Euclidean norm on $\mathbb{M} = \mathbb{R}^d \times \mathbb{S}^{d-1}$, we must have $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) \to 0$. We prove in Theorem 1 that $(\mathbb{M}, \mathcal{F}_0)$ is locally controllable, but $(\mathbb{M}, \mathcal{F}_0^+)$ is not. Indeed the SR Reeds-Shepp car can achieve sideways motions by alternating the forward and reverse gear with slight direction changes, whereas the

model without reverse gear lacks this possibility. For completeness, the theorem contains a standard (rough) estimate of the distance near the source (due to well-known estimates [43,72,18,61]).

Furthermore, we prove existence of minimizers for the Reeds-Shepp model without reverse gear. Existence results of minimizers of the model with reverse gear (the SR model) already exist, by the Chow-Rashevski theorem and Fillipov theorems [2].

Theorem 1 ((Local) controllability properties) Minimizers exist for both the classical Reeds-Shepp model, and for the Reeds-Shepp model without reverse gear. Both models are globally controllable.

- The Reeds-Shepp model without reverse gear is not locally controllable, since

$$\limsup_{\mathbf{p}' \to \mathbf{p}} d_{\mathcal{F}_0^+}(\mathbf{p}, \mathbf{p}') \ge 2\pi\delta, \text{ for all } \mathbf{p} \in \mathbb{M}.$$
(11)

If the cost $C_2 = \delta$ is constant on \mathbb{M} , then this inequality is sharp:

$$\limsup_{\mathbf{p}' \to \mathbf{p}} d_{\mathcal{F}_0^+}(\mathbf{p}, \mathbf{p}') = \lim_{\mu \downarrow 0} d_{\mathcal{F}_0^+}((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mu \mathbf{n}, \mathbf{n})) = 2\pi\delta.$$
(12)

- The sub-Riemannian Reeds-Shepp model is locally controllable, since

$$d_{\mathcal{F}_0}(\mathbf{p}, \mathbf{p}') = \mathcal{O}\left(\mathcal{C}_2(\mathbf{p}) \|\mathbf{n} - \mathbf{n}'\| + \sqrt{\mathcal{C}_2(\mathbf{p})\mathcal{C}_1(\mathbf{p})} \|\mathbf{x} - \mathbf{x}'\|\right)$$

$$as \mathbf{p}' = (\mathbf{x}', \mathbf{n}') \to \mathbf{p} = (\mathbf{x}, \mathbf{n}).$$
(13)

For a proof see Section 4.

Exercise 3 Use the result of the previous theorem to show that for the uniform cost case $\xi^{-1}C_1 = C_2 = 1$ with $\xi = 1$ one has that

- the quasi distance mapping $d_{\mathcal{F}_0^+}(\mathbf{p}_S, \cdot)$, given by (8) and (3), is not continuous on \mathbb{M} w.r.t. metric topology of the standard distance given by

$$d(\mathbf{p}_{1}, \mathbf{p}_{2}) = d((\mathbf{x}_{1}, \mathbf{n}_{1}), (\mathbf{x}_{2}, \mathbf{n}_{2})) = \sqrt{\|\mathbf{x}_{1} - \mathbf{x}_{2}\|^{2} + |\arccos(\mathbf{n}_{1} \cdot \mathbf{n}_{2})|^{2}},$$

for all $\mathbf{p}_{1} = (\mathbf{x}_{1}, \mathbf{n}_{1}), \mathbf{p}_{2} = (\mathbf{x}_{2}, \mathbf{n}_{2}) \in \mathbb{M}.$ (14)

- the symmetric distance mapping $d_{\mathcal{F}_0}(\mathbf{p}_S, \cdot)$, given by (8) and (2), is continuous on \mathbb{M} w.r.t. metric topology of the standard distance given by (14).
- the symmetric distance spheres

$$\mathcal{S}_{R}^{\mathcal{F}}(\mathbf{p}_{S}) := \{ \mathbf{p} \in \mathbb{M} \mid d_{\mathcal{F}}(\mathbf{p}_{S}, \mathbf{p}) = R \} \text{ with } \mathcal{F} = \mathcal{F}_{0},$$
(15)

are compact.

Exercise 4 (the scaling homothety)

Consider $\mathbf{p}_{S} := (0, 0, 0)$ and d = 2.

Let $\xi > 0$ denote a stiffness parameter $\xi^{-1} > 0$ with physical dimension [Length].

Recall \mathcal{F}_0 was given by Eq. (2) which includes relative cost-functions \mathcal{C}_1 and \mathcal{C}_2 . Let $\mathcal{F}_{0,\xi}$ denote the special case where the Finsler function is induced by constant relative costs: $\mathcal{C}_1 = \xi > 0$ and $\mathcal{C}_2 = 1$. Let $d_{\mathcal{F}_{0,\xi}}$ be the metric induced by $\mathcal{F} = \mathcal{F}_{0,\xi}$ given by (8). Consider the corresponding spheres $\mathcal{S}_R^{\mathcal{F}_{0,\xi}}(\mathbf{p}_S)$ given by (15). Define a scaling operator on M:

$$\Phi_{\xi}(x, y, \theta) = (\xi x, \xi y, \theta).$$

Let us write $\Gamma(\cdot) = (X(\cdot), Y(\cdot), \Theta(\cdot)) = \Phi_{\xi}(\gamma(\cdot))$ for the scaled output curve.

a Show that

$$\mathcal{F}_{0,\xi}(\gamma(t),\dot{\gamma}(t)) = \mathcal{F}_{0,1}(\Gamma(t),\dot{\Gamma}(t)),$$

for all $t \in (0, 1)$, and check that the physical dimensions are consistent in this identity. b Show that

$$\mathcal{S}_{R}^{\mathcal{F}_{0,\xi}}(\mathbf{p}_{S}) = \Phi_{\xi}^{-1}(\mathcal{S}_{R}^{\mathcal{F}_{0,1}}(\mathbf{p}_{S}))$$

for all R > 0.



Fig. 6 Levelsets for d = 2 of the (approximating) metrics $\mathcal{F}_{\varepsilon}(\mathbf{0}, (\dot{x}, \dot{y}, \dot{\theta})) = 1$ (left) and $\mathcal{F}_{\varepsilon}^{+}(\mathbf{0}, (\dot{x}, \dot{y}, \dot{\theta})) = 1$ (right), with $\varepsilon = 0.2$ (top) and $\varepsilon = 0$ (bottom). In this example, $\mathcal{C}_{2}(\mathbf{0}) = 2\mathcal{C}_{1}(\mathbf{0})$.

2.3 A Continuous Approximation for the Reeds-Shepp geometry

We introduce approximations $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$ of the Reeds-Shepp metrics \mathcal{F}_0 and \mathcal{F}_0^+ , depending on a small parameter $0 < \varepsilon \leq 1$, which are continuous and in particular take only finite values. This is a prerequisite for our numerical methods. Both approximations penalize the deviation from the constraints of collinearity $\dot{\mathbf{x}} \propto \mathbf{n}$, and in addition, $\mathcal{F}_{\varepsilon}^+$ penalizes negativity of the scalar product $\dot{\mathbf{x}} \cdot \mathbf{n}$, appearing in (2) and (3). For that purpose, we introduce some additional notation: for $\dot{\mathbf{x}} \in \mathbb{R}^d$ and $\mathbf{n} \in \mathbb{S}^{d-1}$ we define

$$\begin{aligned} |\dot{\mathbf{x}} \wedge \mathbf{n}||^2 &:= ||\dot{\mathbf{x}}||^2 - |\dot{\mathbf{x}} \cdot \mathbf{n}|^2, \\ (\dot{\mathbf{x}} \cdot \mathbf{n})_- &:= \min\{0, \dot{\mathbf{x}} \cdot \mathbf{n}\}, \quad (\dot{\mathbf{x}} \cdot \mathbf{n})_+ := \max\{\dot{\mathbf{x}} \cdot \mathbf{n}, 0\}. \end{aligned}$$
(16)

These are respectively the norm of the orthogonal projection⁵ of $\dot{\mathbf{x}}$ onto the plane orthogonal to \mathbf{n} , and the negative and positive parts of their scalar product. The two metrics $\mathcal{F}_{\varepsilon}, \mathcal{F}_{\varepsilon}^{+}: T(\mathbb{M}) \to \mathbb{R}_{+}$ are defined for each $0 < \varepsilon \leq 1$, as follows: for $(\mathbf{p}, \dot{\mathbf{p}}) \in T(\mathbb{M})$ with components $\mathbf{p} = (\mathbf{x}, \mathbf{n})$ and $\dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}})$ we define

$$\mathcal{F}_{\varepsilon}(\mathbf{p}, \dot{\mathbf{p}})^{2} := \mathcal{C}_{1}(\mathbf{p})^{2} (|\dot{\mathbf{x}} \cdot \mathbf{n}|^{2} + \varepsilon^{-2} \|\dot{\mathbf{x}} \wedge \mathbf{n}\|^{2}) + \mathcal{C}_{2}(\mathbf{p})^{2} \|\dot{\mathbf{n}}\|^{2},$$
(17)

$$\mathcal{F}_{\varepsilon}^{+}(\mathbf{p},\dot{\mathbf{p}})^{2} := \mathcal{C}_{1}(\mathbf{p})^{2} (|\dot{\mathbf{x}}\cdot\mathbf{n}|^{2} + \varepsilon^{-2} ||\dot{\mathbf{x}}\wedge\mathbf{n}|^{2} + (\varepsilon^{-2} - 1)(\dot{\mathbf{x}}\cdot\mathbf{n})^{2}) + \mathcal{C}_{2}(\mathbf{p})^{2} ||\dot{\mathbf{n}}||^{2}$$
(18)

$$= \mathcal{C}_1(\mathbf{p})^2 ((\dot{\mathbf{x}} \cdot \mathbf{n})_+^2 + \varepsilon^{-2} \|\dot{\mathbf{x}} \wedge \mathbf{n}\|^2 + \varepsilon^{-2} (\dot{\mathbf{x}} \cdot \mathbf{n})_-^2) + \mathcal{C}_2(\mathbf{p})^2 \|\dot{\mathbf{n}}\|^2.$$
(19)

See Fig. 6 for a visualization of a levelset of both metrics in $\mathbb{R}^2 \times \mathbb{S}^1$. Note that $\mathcal{F}_{\varepsilon}$ is a Riemannian metric on \mathbb{M} (with the same smoothness as the cost functions $\mathcal{C}_2, \mathcal{C}_1$), and that $\mathcal{F}_{\varepsilon}^+$ is neither Riemannian nor smooth due to the term $(\dot{\mathbf{x}} \cdot \mathbf{n})_-$. One clearly has the pointwise 'convergence' $\mathcal{F}_{\varepsilon}(\mathbf{p}, \dot{\mathbf{p}}) \to \mathcal{F}_0(\mathbf{p}, \dot{\mathbf{p}})$ on $\mathbb{R} \cup \{\infty\}$ as $\varepsilon \to 0$, and likewise $\mathcal{F}_{\varepsilon}^+(\mathbf{p}, \dot{\mathbf{p}}) \to \mathcal{F}_0^+(\mathbf{p}, \dot{\mathbf{p}})$. The use of $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$ is further justified by the following convergence result.

Theorem 2 (Convergence of the Approximative Models to the Exact Models) One has the pointwise convergence: for any $\mathbf{p}, \mathbf{q} \in \mathbb{M}$

$$\begin{aligned} & d_{\mathcal{F}_{\varepsilon}}(\mathbf{p},\mathbf{q}) \to d_{\mathcal{F}_{0}}(\mathbf{p},\mathbf{q}), \\ & d_{\mathcal{F}_{\varepsilon}^{+}}(\mathbf{p},\mathbf{q}) \to d_{\mathcal{F}_{\varepsilon}^{+}}(\mathbf{p},\mathbf{q}), \end{aligned} \quad as \ \varepsilon \to 0. \end{aligned}$$

⁵ The quantity $\|\dot{\mathbf{x}} \wedge \mathbf{n}\|$ is also the norm of the wedge product of $\dot{\mathbf{x}}$ and \mathbf{n} , but defining it this way would require introducing some algebra which is not needed in the rest of this paper.



Fig. 7 Illustration of cusps in symmetric distance minimizers in $\mathbb{M} = \mathbb{R}^d \times \mathbb{S}^{d-1}$. Left: cusps in spatial projections $\mathbf{x}(\cdot)$ of symmetric SR distance ($\varepsilon = 0$) minimizers $\gamma(\cdot) = (\mathbf{x}(\cdot), \mathbf{n}(\cdot))$ for d = 2, right: cusps (red dots) appearing in spatial projections of symmetric SR distance geodesics for d = 3. In the 3D case we indicate the corresponding rotations \mathbf{R}_{n_1} via a local 3D frame.

Consider for each $\varepsilon > 0$ a minimizing path γ_{ε}^* from **p** to **q**, with respect to the metric $\mathcal{F}_{\varepsilon}$, parametrized at constant speed

$$\mathcal{F}_{\varepsilon}(\gamma_{\varepsilon}^{*}(t), \dot{\gamma}_{\varepsilon}^{*}(t)) = d_{\mathcal{F}_{\varepsilon}}(\mathbf{p}, \mathbf{q}), \qquad \forall t \in [0, 1].$$

Assume that there is a unique minimal path γ^* from **p** to **q** with respect to the sub-Riemannian distance $d_{\mathcal{F}_0}$ (in other words **q** is not within the cut locus of **p**), parametrized at constant speed:

$$\mathcal{F}_0(\gamma^*(t), \dot{\gamma}^*(t)) = d_{\mathcal{F}_0}(\mathbf{p}, \mathbf{q}), \qquad \forall t \in [0, 1].$$

Then $\gamma_{\varepsilon}^* \to \gamma^*$ as $\varepsilon \to 0$, uniformly on [0,1]. Likewise replacing $\mathcal{F}_{\varepsilon}$ with $\mathcal{F}_{\varepsilon}^+$ for all $\varepsilon \ge 0$.

The proof, presented in Appendix A is based on a general result originally applied to the Euler elastica curves in [14]. Combining Theorem 2 with the local controllability properties established in Theorem 1, one obtains that $d_{\mathcal{F}_{\varepsilon}} \to d_{\mathcal{F}_{0}}$ locally uniformly on $\mathbb{M} \times \mathbb{M}$, and that the convergence $d_{\mathcal{F}_{\varepsilon}^{+}} \to d_{\mathcal{F}_{0}^{+}}$ is only pointwise.

Remark 2 If there exists a family of minimal geodesics $(\gamma_i^*)_{i \in I}$ from **p** to **q** with respect to \mathcal{F}_0 (resp. \mathcal{F}_0^+), then one can show that for any sequence $\varepsilon_n \to 0$ one can find a subsequence and an index $i \in I$ such that $\gamma_{\varepsilon_{\omega(n)}}^* \to \gamma_i^*$ uniformly as $n \to \infty$.

2.4 Points of Interest in Spatial Projections of Geodesics for the Uniform Cost Case: Cusps vs. Keypoints

Next we provide a theorem that tells us in each of the models/metric spaces $(\mathbb{M}, d_{\mathcal{F}_0})$, $(\mathbb{M}, d_{\mathcal{F}_\varepsilon})$ and $(\mathbb{M}, d_{\mathcal{F}_0^+})$, $(\mathbb{M}, d_{\mathcal{F}_\varepsilon^+})$, with $\mathcal{C}_1 = \mathcal{C}_2 = 1$ and d = 2 where cusps occur in spatial projections of geodesics or where keypoints with in-place rotations take place.

Definition 2 (Cusp) A cusp point $\mathbf{x}(t_0)$ on a spatial projection of a (SR-)geodesic $t \mapsto (\mathbf{x}(t), \mathbf{n}(t))$ in \mathbb{M} is a point where

$$\tilde{u}(t_0) = 0$$
, and $\tilde{u}(t_0) \neq 0$,
where $\tilde{u}(t) := \mathbf{n}(t) \cdot \dot{\mathbf{x}}(t)$ for all t . (20)

I.e. a cusp point is a point where the spatial control aligned with $\mathbf{n}(t_0)$ vanishes and switches sign locally.

Although this definition explains the notion of a cusp geometrically (as can be observed in Fig. 1 and Fig. 7), it contains a redundant part for the relevant case of interest: the second condition automatically follows when considering the SR geodesics in $(\mathbb{M}, d_{\mathcal{F}_0})$. The following lemma gives a characterization of a cusp point in terms of the distance function along a curve.

Lemma 1 Consider a SR-geodesic $\gamma = (\mathbf{x}, \mathbf{n}) : [0, 1] \to (\mathbb{M}, d_{\mathcal{F}_0})$, parametrized at constant speed, and which physical position $\mathbf{x}(\cdot)$ is not identically constant. Denote $\mathbf{p}_S := \gamma(0)$ and $U(\cdot) := d_{\mathcal{F}_0}(\mathbf{p}_S, \cdot)$. Let $t_0 \in (0,1)$ be such that U is differentiable at $\gamma(t_0) = (\mathbf{x}(t_0), \mathbf{n}(t_0))$. Then

$$\mathbf{x}(t_0) \text{ is a cusp point } \Leftrightarrow \mathbf{n}(t_0) \cdot \dot{\mathbf{x}}(t_0) = 0$$

$$\Leftrightarrow \mathbf{n}(t_0) \cdot \nabla_{\mathbb{R}^d} U(\mathbf{x}(t_0), \mathbf{n}(t_0)) = 0.$$
(21)

The proof can be found in Appendix C.

Definition 3 (Keypoint) A point $\tilde{\mathbf{x}}$ on the spatial projection of a geodesic $\gamma(\cdot) = (\mathbf{x}(\cdot), \mathbf{n}(\cdot))$ in \mathbb{M} is a keypoint of γ if there exist $t_0 < t_1$, such that $\mathbf{x}(t) = \tilde{\mathbf{x}}$ and $\dot{\mathbf{n}}(t) \neq 0$ for all $t \in [t_0, t_1]$, i.e., a point where an in-place rotation takes place.

Definition 4 We define the set $\mathfrak{R} \subset \mathbb{M}$ to be all endpoints that can be reached with a geodesic γ^* : $[0,1] \to \mathbb{M}$ in $(\mathbb{M}, d_{\mathcal{F}_0})$ whose spatial control $\tilde{u}(t)$ stays positive for all $t \in [0,1]$.

Remark 3 The word 'geodesic' in this definition can (in the case d = 2) be replaced by 'globally minimizing geodesic' [12]. For a definition in terms of the exponential map of a geometrical control problem \mathbf{P}_{curve} , see e.g. [24,27], in which the same positivity condition for \tilde{u} is imposed. Fig. 8 shows more precisely what this set looks like for d = 2 [24], in particular that it is contained in the half-space $\mathbf{a} \cdot \mathbf{x} \ge 0$, and for d = 3[27]. We extend these results with the following theorem.

Theorem 3 (Cusps and Keypoints) Let $\varepsilon > 0$, d = 2, $C_1 = C_2 = 1$. Then,

- in $(\mathbb{M}, d_{\mathcal{F}_0})$ cusps are present in spatial projections of almost every optimal SR geodesics when their times t are extended on the real line until they reach their cut-time. In fact, the straight-lines connecting specific boundary points $\mathbf{p} = (\mathbf{x}, \mathbf{n})$ and $\mathbf{q} = (\mathbf{x} + \lambda \mathbf{n}, \mathbf{n})$ with $\lambda \in \mathbb{R}$ are the only exceptions.
- in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ and $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ and $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ no cusps appear in spatial projections of geodesics.

Furthermore,

- $-in(\mathbb{M}, d_{\mathcal{F}_0}), (\mathbb{M}, d_{\mathcal{F}_{\varepsilon}}) and(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+}) keypoints only occur with vertical geodesics (moving only angularly).$ - in $(\mathbb{M}, d_{\mathcal{F}^+})$ keypoints only occur at the endpoints of minimal paths.
- An optimal geodesic γ_+ in $(\mathbb{M}, d_{\mathcal{F}^+_0})$ departing from $\mathbf{e} = (0, 0, 0)$ and ending in $\mathbf{p} = (x, y, \theta)$ has

A) no keypoint if $\mathbf{p} \in \overline{\mathfrak{R}}$,

- B) a keypoint in (0,0) if x < 0,

- C) a keypoint only in (x, y) if C1) $\mathbf{p} \in \overline{\mathfrak{R}}^c$ and $x \ge 2$, C2) $\mathbf{p} \in \overline{\mathfrak{R}}^c$ and $0 \le x < 2$ and
 - $|y| \leq -ix E\left(iarcsinh\left(\frac{x}{\sqrt{4-x^2}}\right), \frac{x^2-4}{x^2}\right), where E(z,m)$ denotes the Elliptic integral of the second kind.

Remark 4 In case A, γ_+ is an optimal geodesic in $(\mathbb{M}, d_{\mathcal{F}_0})$ as well. In case B, γ_+ departs from a cusp. In case C, γ^+ is a concatenation of an optimal geodesic in $(\mathbb{M}, d_{\mathcal{F}_0})$ and an in-place rotation. For other endpoints (x, y, θ) for geodesics departing from **e** with $0 \le x < 2$, other than the ones reported in C2 it is not immediately clear what happens, due to [24, Thm.9]. Also points with x < 0 may have keypoints at the end as well. See Fig. 9 where various cases of minimizing geodesics in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ are depicted.

Remark 5 See [38, Fig. 6] to see the smoothing effect of taking ε small but nonzero on the cusps of non-optimal geodesics in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ and keypoints in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$.

⁶ Here $\overline{\mathfrak{R}}^c = \mathbb{M} \setminus \overline{\mathfrak{R}}$ denotes the complement of the closure $\overline{\mathfrak{R}}$ of \mathfrak{R} , and $E(z,m) = \int_0^z \sqrt{1 - m \sin^2 v} \, \mathrm{d}v$.



Fig. 8 The set \Re of endpoints reachable from the origin **e** (recall (10)) via SR geodesics whose spatial projections do not exhibit cusps has been studied for the case d = 2 (left), and for the case d = 3 (right). For d = 2 it is contained in $x \ge 0$ and for d = 3 it is contained in $z \ge 0$. The boundary of this set contains of endpoints of geodesics departing at a cusp (in red) or of endpoints of geodesics ending in a cusp (in blue). If an endpoint (**x**, **n**) is placed outside \Re (e.g. the green points above) then following the approach in Theorem 4, depending on its initial spatial location it first connects to a blue point (**x**, **n**_{new}) via a spherical geodesic end then connects to the origin **e** via a SR geodesic. Then it has a keypoint at the endpoint. For other locations spatial locations (orange points), the geodesic has the keypoint in the origin, or even at both boundaries, cf. Fig. 9.

2.5 The Eikonal PDE Formalism

As briefly discussed in Section 1.2, continuous metrics like $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$ for any $\varepsilon > 0$, allow to use the standard theory of viscosity solutions of eikonal PDEs, and thus to design provable and efficient numerical schemes for the computation of distance maps and minimal geodesics. More precisely, consider a continuous Finsler function $\mathcal{F} \in C^0(T(\mathbb{M}), \mathbb{R}^+)$, and define the dual Finsler function \mathcal{F}^* on the co-tangent bundle as follows: for all $(\mathbf{p}, \hat{\mathbf{p}}) \in T^*(\mathbb{M})$

$$\mathcal{F}^*(\mathbf{p}, \hat{\mathbf{p}}) := \sup_{\dot{\mathbf{p}} \in T_{\mathbf{p}} \mathbb{M} \setminus \{0\}} \frac{\langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle}{\mathcal{F}(\mathbf{p}, \dot{\mathbf{p}})}.$$
(22)

The distance map $U = d_{\mathcal{F}}(\mathbf{p}_{\mathrm{S}}, \cdot)$ from a given source point $\mathbf{p}_{\mathrm{S}} \in \mathbb{M}$ is the unique solution, in the sense of viscosity solutions, of the static Hamilton Jacobi equation: $U(\mathbf{p}_{\mathrm{S}}) = 0$, and for all $p \in \mathbb{M}$

$$\mathcal{F}^*(\mathbf{p}, \mathrm{d}U(\mathbf{p})) = 1. \tag{23}$$

Furthermore, if γ is a minimal geodesic from \mathbf{p}_{S} to some $\mathbf{p} \in \mathbb{M}$, then it obeys the ordinary differential equation (ODE):

$$\begin{cases} \dot{\gamma}(t) = -L \, \mathrm{d}_{\hat{\mathbf{p}}} \mathcal{F}^*(\gamma(t), \mathrm{d}U(\gamma(t))), \ L := d_{\mathcal{F}}(\mathbf{p}_{\mathrm{S}}, \mathbf{p}) \\ \gamma(0) = \mathbf{p}, \qquad \gamma(1) = \mathbf{p}_{\mathrm{S}}. \end{cases}$$
(24)

for any $t \in [0, 1]$ such that the differentiability of U and \mathcal{F}^* holds at the required points. The proof of the ODE (24) is for completeness derived in Proposition 4 of Appendix B, where we also discuss in Remark 16 the common alternative formalism based on the Hamiltonian. We denoted by $d_{\hat{\mathbf{p}}}\mathcal{F}^*$ the differential of



Fig. 9 Shortest paths for d = 2 using the Finsler functions \mathcal{F}_0 (blue) and \mathcal{F}_0^+ (red), with point source $\mathbf{p}_S = (0,0,0)$ and varying end conditions. Row A: $\mathbf{p} = (0,0.8,\pi n/4)$. Row B: $\mathbf{p} = (0.8,0.8,\pi n/4)$. Row C: $\mathbf{p} = (-0.8,0,\pi n/4)$. Here $n = 1,\ldots,8$, corresponding to the columns. When there are two optimal geodesics, both are drawn. Circles around the begin or end point indicate in-place rotation of the red curve at that point. We see that whenever the blue geodesic has a cusp, the red geodesic has at least one in-place rotation (keypoint). This numerically supports our statements in Theorem 3 considering cusps and keypoints. For high accuracy we applied the relatively slow iterative PDE approach [9] on a 101 × 101 × 64-grid in \mathbb{M} to compute $d_{\mathcal{F}_0}(\mathbf{p}, \mathbf{p}_S)$ and $d_{\mathcal{F}_0^+}(\mathbf{p}, \mathbf{p}_S)$, see [38, App. B].

the dual metric \mathcal{F}^* with respect to the second variable $\hat{\mathbf{p}}$, hence $d_{\hat{\mathbf{p}}}\mathcal{F}^*(\mathbf{p}, \hat{\mathbf{p}}) \in T^{**}_{\mathbf{p}}(\mathbb{M}) \cong T_{\mathbf{p}}(\mathbb{M})$ is indeed a tangent vector to \mathbb{M} , for all $(\mathbf{p}, \hat{\mathbf{p}}) \in T^*\mathbb{M}$.

Remark 6 (Steepest Descent)

The ODE (24) boils down to a steepest descent on the distance map. It is sometimes re-formulated as a steepest ascent⁷ where one interchanges the role of \mathbf{p}_S and \mathbf{p} above and where one applies $t \mapsto 1 - t$. Note that this time inversion yields a plus sign instead of a minus sign in the righthand side of the PDE in (24).

In the rest of this section, we specialize (23) and (24) to the metrics $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$. Our first result provides explicit expressions for the dual metrics (required for the eikonal equation).

Proposition 1 For any $0 < \varepsilon \leq 1$, the duals to the approximate Reeds-Shepp metrics $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$ are: for all $(\mathbf{p}, \hat{\mathbf{p}}) \in T^*(\mathbb{M})$, with $\mathbf{p} = (\mathbf{x}, \mathbf{n})$ and $\hat{\mathbf{p}} = (\hat{\mathbf{x}}, \hat{\mathbf{n}})$

$$\mathcal{F}_{\varepsilon}^{*}(\mathbf{p}, \hat{\mathbf{p}})^{2} = (\mathcal{C}_{2}(\mathbf{p}))^{-2} \|\hat{\mathbf{n}}\|^{2} + (\mathcal{C}_{1}(\mathbf{p}))^{-2} (|\hat{\mathbf{x}} \cdot \mathbf{n}|^{2} + \varepsilon^{2} \|\hat{\mathbf{x}} \wedge \mathbf{n}\|^{2})$$

$$\mathcal{F}_{\varepsilon}^{+*}(\mathbf{p}, \hat{\mathbf{p}})^{2} = (\mathcal{C}_{2}(\mathbf{p}))^{-2} \|\hat{\mathbf{n}}\|^{2} + (\mathcal{C}_{1}(\mathbf{p}))^{-2} (|\hat{\mathbf{x}} \cdot \mathbf{n}|^{2} + \varepsilon^{2} \|\hat{\mathbf{x}} \wedge \mathbf{n}\|^{2} - (1 - \varepsilon^{2})(\hat{\mathbf{x}} \cdot \mathbf{n})_{-}^{2})$$

$$= (\mathcal{C}_{2}(\mathbf{p}))^{-2} \|\hat{\mathbf{n}}\|^{2} + (\mathcal{C}_{1}(\mathbf{p}))^{-2} ((\hat{\mathbf{x}} \cdot \mathbf{n})_{+}^{2} + \varepsilon^{2} (\hat{\mathbf{x}} \cdot \mathbf{n})_{-}^{2} + \varepsilon^{2} \|\hat{\mathbf{x}} \wedge \mathbf{n}\|^{2})$$
(25)

In order to relate the Finslerian HJB equation (23) and backtracking equation (24) to some more classical Riemannian counterparts, we introduce two Riemannian metric tensor fields on \mathbb{M} . The first is defined as the polarisation of the norm $\mathcal{F}_{\varepsilon}(\mathbf{p}, \cdot)$

$$\mathcal{G}_{\mathbf{p};\varepsilon}(\dot{\mathbf{p}},\dot{\mathbf{p}}) = |\mathcal{F}_{\varepsilon}(\mathbf{p},\dot{\mathbf{p}})|^2 = \mathcal{C}_1^2(\mathbf{p})((\dot{\mathbf{x}}\cdot\mathbf{n})^2 + \varepsilon^{-2} \|\dot{\mathbf{x}}\wedge\mathbf{n}\|^2) + \mathcal{C}_2^2(\mathbf{p})\|\dot{\mathbf{n}}\|^2 , \qquad (26)$$

where $\dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}})$, and then one can also rely on gradient fields $\mathbf{p} \mapsto \mathcal{G}_{\mathbf{p};\varepsilon}^{-1} dU(\mathbf{p})$ relative to this metric. This has benefits if it comes to geometric understanding of the eikonal equation and its tracking. Even in the analysis of the non-symmetric case –where one does not have a single metric tensor– this notion plays

 $^{^{7}\,}$ In both the ascent and descent formulations one must keep in mind that the distance function is not differentiable at the origin.

a role, as we will see in the next main theorem. To this end, in the non-symmetric case, we shall rely on a second spatially isotropic metric tensor given by:

$$\widetilde{\mathcal{G}}_{\mathbf{p};\varepsilon}(\dot{\mathbf{p}},\dot{\mathbf{p}}) := \mathcal{C}_1^2(\mathbf{p})\,\varepsilon^{-2}\,\|\dot{\mathbf{x}}\|^2 + \mathcal{C}_2^2(\mathbf{p})\|\dot{\mathbf{n}}\|^2.$$
(27)

We denote by $\nabla_{\mathbb{S}^{d-1}}$ the gradient operator on \mathbb{S}^{d-1} with respect to the inner product induced by the embedding $\mathbb{S}^{d-1} \subset \mathbb{R}^d$, and by $\nabla_{\mathbb{R}^d}$ the canonical gradient operator on \mathbb{R}^d .

Corollary 1 Let $\varepsilon \geq 0$. Then the eikonal PDE (5) for the case $(\mathbb{M}, \mathcal{F}_{\varepsilon})$ now takes the form

$$\begin{split} \sqrt{\frac{\|\nabla_{\mathbb{S}^{d-1}}U(\mathbf{p})\|^2}{\mathcal{C}_2^2(\mathbf{p})}} + \frac{\varepsilon^2 \|\nabla_{\mathbb{R}^d}U(\mathbf{p})\|^2 + (1-\varepsilon^2)\|\mathbf{n}\cdot\nabla_{\mathbb{R}^d}U(\mathbf{p})\|^2}{\mathcal{C}_1^2(\mathbf{p})} = 1, \\ \Leftrightarrow \\ \mathcal{G}_{\mathbf{p};\varepsilon}|_{\mathbf{p}} \left(\mathcal{G}_{\mathbf{p};\varepsilon}^{-1} \mathrm{d}U(\mathbf{p}), \mathcal{G}_{\mathbf{p};\varepsilon}^{-1} \mathrm{d}U(\mathbf{p}) \right) = 1. \end{split}$$

The eikonal PDE (5) for the case $(\mathbb{M}, \mathcal{F}_{\varepsilon}^+)$ now takes the explicit form:

$$\begin{split} \sqrt{\frac{\|\nabla_{\mathbb{S}^{d-1}}U^{+}(\mathbf{p})\|^{2}}{\mathcal{C}_{2}^{2}(\mathbf{p})}} + \frac{\varepsilon^{2}\|\nabla_{\mathbb{R}^{d}}U^{+}(\mathbf{p})\|^{2} + (1-\varepsilon^{2})|(\mathbf{n}\cdot\nabla_{\mathbb{R}^{d}}U^{+}(\mathbf{p}))_{+}|^{2}}{\mathcal{C}_{1}^{2}(\mathbf{p})} & \Leftrightarrow \\ \begin{cases} \mathcal{G}_{\mathbf{p};\varepsilon}|_{\mathbf{p}} \left(\mathcal{G}_{\mathbf{p};\varepsilon}^{-1}\mathrm{d}U^{+}(\mathbf{p}), \mathcal{G}_{\mathbf{p};\varepsilon}^{-1}\mathrm{d}U^{+}(\mathbf{p})\right) = 1, \\ & \text{if } \mathbf{p} \in \mathbb{M}_{+} := \{\mathbf{p} \in \mathbb{M} \mid \langle \mathrm{d}U^{+}(\mathbf{p}), \mathbf{n} \rangle > 0\}, \\ & \widetilde{\mathcal{G}}_{\mathbf{p};\varepsilon}|_{\mathbf{p}} \left(\widetilde{\mathcal{G}}_{\mathbf{p};\varepsilon}^{-1}\mathrm{d}U^{+}(\mathbf{p}), \widetilde{\mathcal{G}}_{\mathbf{p};\varepsilon}^{-1}\mathrm{d}U^{+}(\mathbf{p})\right) = 1, \\ & \text{if } \mathbf{p} \in \mathbb{M}_{-} := \{\mathbf{p} \in \mathbb{M} \mid \langle \mathrm{d}U^{+}(\mathbf{p}), \mathbf{n} \rangle < 0\}. \end{split}$$

for those $\mathbf{p} \in \mathbb{M}_+ \cup \mathbb{M}_-$ where U^+ is differentiable⁸.

The proof of Proposition 1 and Corollary 1 can be found in Section 6.

We finally specialize the geodesic ODE (24) to the models of interest. Note that in the case of $\mathcal{F}_{\varepsilon}^+$, the backtracing switches between qualitatively distinct modes, respectively almost sub-Riemannian and almost purely angular, in the spirit of Theorem 3. Given $\varepsilon > 0$ and $\mathbf{n} \in \mathbb{S}^{d-1}$ let $D_{\mathbf{n}}^{\varepsilon}$ denote the $d \times d$ symmetric positive definite matrix with eigenvalue 1 in the direction \mathbf{n} , and eigenvalue ε^2 in the orthogonal directions :

$$D_{\mathbf{n}}^{\varepsilon} := \mathbf{n} \otimes \mathbf{n} + \varepsilon^{2} (\mathrm{Id} - \mathbf{n} \otimes \mathbf{n}).$$
(28)

Theorem 4 (Back-tracking) Let $0 < \varepsilon < 1$. Let $\mathbf{p}_{\mathrm{S}} \in \mathbb{M}$ be a source point. Let $U(\mathbf{p}) := d_{\mathcal{F}_{\varepsilon}}(\mathbf{p}, \mathbf{p}_{s})$, $U^{+}(\mathbf{p}) := d_{\mathcal{F}_{\varepsilon}^{+}}(\mathbf{p}, \mathbf{p}_{s})$ be distance maps from \mathbf{p}_{s} , w.r.t. Finster function $\mathcal{F}_{\varepsilon}$, and $\mathcal{F}_{\varepsilon}^{+}$. Let $\gamma, \gamma^{+} : [0, 1] \to \mathbb{M}$ be normalized geodesics of length L starting at \mathbf{p}_{s} in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ resp. $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^{+}})$. Let time $t \in [0, 1]$.

For the Riemannian approximation paths of the Reeds-Shepp car we have, provided that U is differentiable at $\gamma(t) = (\mathbf{x}(t), \mathbf{n}(t))$, that

$$\dot{\gamma}(t) = L \mathcal{G}_{\gamma(t);\varepsilon}^{-1} \mathrm{d}U(\gamma(t)) \Leftrightarrow (\dot{\mathbf{n}}(t) = L \mathcal{C}_2(\gamma(t))^{-1} \nabla_{\mathbb{S}^{d-1}} U(\gamma(t)), (\dot{\mathbf{x}}(t) = L \mathcal{C}_1(\gamma(t))^{-1} D_{\mathbf{n}(t)}^{\varepsilon} \nabla_{\mathbb{R}^d} U(\gamma(t)).$$

$$(29)$$

For the approximation paths of the car without reverse gear we have, provided that U^+ is differentiable at $\gamma^+(t) = (\mathbf{x}^+(t), \mathbf{n}^+(t))$, that

$$\dot{\gamma}^{+}(t) = L \begin{cases} \mathcal{G}_{\gamma^{+}(t);\varepsilon}^{-1} \mathrm{d}U^{+}(\gamma^{+}(t)) & \text{if } \gamma^{+}(t) \in \mathbb{M}_{+}, \\ \widetilde{\mathcal{G}}_{\gamma^{+}(t);\varepsilon}^{-1} \mathrm{d}U^{+}(\gamma^{+}(t)) & \text{if } \gamma^{+}(t) \in \mathbb{M}_{-}, \end{cases}$$
(30)

with $\widetilde{\mathcal{G}}_{\mathbf{p};\varepsilon}(\dot{\mathbf{p}},\dot{\mathbf{p}})$ given by (27), with disjoint Riemannian manifold splitting $\mathbb{M} = \mathbb{M}_+ \cup \mathbb{M}_- \cup \partial \mathbb{M}_{\pm}$. Manifold \mathbb{M}_+ is equipped with metric tensor $\mathcal{G}_{\varepsilon}$, \mathbb{M}_- is equipped with metric tensor $\widetilde{\mathcal{G}}_{\varepsilon}$ and

$$\partial \mathbb{M}_{\pm} := \overline{\mathbb{M}_{+}} \setminus \mathbb{M}_{+} = \overline{\mathbb{M}_{-}} \setminus \mathbb{M}_{-} \tag{31}$$

denotes the transition surface (surface of keypoints).

⁸ On ∂M_{\pm} distance function U^+ is not differentiable

Remark 7 The general abstract formula (30) reflects that the backtracking in $(\mathbb{M}, \mathcal{F}^+)$ is a combined gradient descent flow on the distance map U^+ on a splitting of \mathbb{M} into two (symmetric) Riemannian manifolds. Its explicit form (likewise (29)) is

$$\begin{cases} \dot{\mathbf{n}}^{+}(t) = L C_{2}(\gamma^{+}(t))^{-1} \nabla_{\mathbb{S}^{d-1}} U^{+}(\gamma^{+}(t)), \\ \dot{\mathbf{x}}^{+}(t) = L \begin{cases} C_{1}(\gamma^{+}(t))^{-1} D_{\mathbf{n}(t)}^{\varepsilon} \nabla_{\mathbb{R}^{d}} U^{+}(\gamma^{+}(t)) \\ & \text{if } \gamma^{+}(t) \in \mathbb{M}_{+}, \end{cases}$$
(32)
$$\varepsilon^{2} C_{1}(\gamma^{+}(t))^{-1} \nabla_{\mathbb{R}^{d}} U^{+}(\gamma^{+}(t)) \\ & \text{if } \gamma^{+}(t) \in \mathbb{M}_{-}, \end{cases}$$

Note that for the (less useful) isotropic case $\varepsilon = 1$, models \mathcal{F}_1 and \mathcal{F}_1^+ coincide and geodesics consist of straight lines $\mathbf{x}(\cdot)$ in \mathbb{R}^d and great circles $\mathbf{n}(\cdot)$ in \mathbb{S}^d that do not influence each other.

Remark 8 In Theorem 4, we assumed distance maps U and U^+ to be differentiable along the path, which is not always the case. In points where the distance map is not differentiable, one can take any sub-gradient in the sub-differential $\partial U(\mathbf{p})$ in order to identify Maxwell points (and Maxwell strata). In particular, in SR geometry, the set of points where the squared distance function $(d_{\mathcal{F}_0}(\cdot, e))^2$ is smooth is open and dense in any compact subset of M, see [1, Thm. 11.15]. The points where it is non-smooth are rare and meaningful: they are either first Maxwell points, conjugate points or abnormal points. The last type does not appear here, because we have a 2-bracket generating distribution, see e.g. [28, Remark 4] and [1, Ch. 20.5.1.]. At points in the closure of the first Maxwell set, two geodesically equidistant wavefronts collide for the first time, see for example [9, Fig.3, Thm 3.2] for the case d = 2 and $\mathcal{C} = \mathcal{C}_1 = \mathcal{C}_2 = 1$. See also Fig. 9, where for some end conditions 2 optimally back-tracked geodesics end with the same length in such a first Maxwell point. The conjugate points are points where local optimality is lost, for a precise definition see e.g. [1, Def. 8.43].

Remark 9 Recall the convergence result from Theorem 2, and the non-local controllability for the model $(\mathbb{M}, d_{\mathcal{F}_0^+})$. From this we see that the convergence holds pointwise but not uniformly (otherwise the limit distance $d_{\mathcal{F}_0^+}$ was continuous). Nevertheless the minimal paths converge strongly as $\varepsilon \downarrow 0$, and we see that the spatial velocity tends to 0 in (32) if $\varepsilon \downarrow 0$ if $\gamma_{\varepsilon}^*(t) \in \mathbb{M}_-$. In the SR case $\varepsilon = 0$, the gradient flows themselves fit continuously and the interface $\partial \mathbb{M}_{\pm}$ is reached with $\dot{\mathbf{x}} \cdot \mathbf{n} = 0$ (and $\dot{\mathbf{x}} = 0$).

Theorem 4 can be extended to the SR case:

Corollary 2 (SR Backtracking) Let the cost C_1, C_2 be smooth, let the source $\mathbf{p}_S \in \mathbb{M}$ and $\mathbf{p} \neq \mathbf{p}_S \in \mathbb{M}$ be such that they can be connected by a unique smooth minimizer γ_{ε}^* in $(\mathbb{M}, \mathcal{F}_{\varepsilon})$ and γ_0^* in $(\mathbb{M}, \mathcal{F}_0)$, such that $\gamma_{\varepsilon}^*(t)$ is not a conjugate point for all $t \in [0, 1]$ and all sufficiently small $\varepsilon > 0$, say $\varepsilon < \varepsilon_0$, for some $\varepsilon_0 > 0$. Then defining $U_0 : \mathbf{q} \in \mathbb{M} \mapsto d_{\mathcal{F}_{\varepsilon}}(\mathbf{p}_s, \mathbf{q})$ one has

$$\dot{\gamma}_0^*(t) = U_0(\mathbf{p})\mathcal{G}_{\gamma_0^*(t);0}^{-1} \,\mathrm{d}U_0(\gamma_0^*(t)), \qquad t \in [0,1]$$

assuming U_0 is differentiable at $\gamma_0^*(t)$. In addition U_0 satisfies the SR eikonal equation:

$$\sqrt{\mathcal{G}_{\mathbf{p};0}\left(\mathcal{G}_{\mathbf{p};0}^{-1}\mathrm{d}U_0(\mathbf{p}),\mathcal{G}_{\mathbf{p};0}^{-1}\mathrm{d}U_0(\mathbf{p})\right)}=1,$$

with respect to sub-Riemannian metric tensor field $\mathbf{p} \mapsto \mathcal{G}_{\mathbf{p};0}$ given by

$$\mathcal{G}_{\mathbf{p}}(\dot{\mathbf{p}}, \dot{\mathbf{p}}) = (\mathcal{C}_1(\mathbf{p}))^2 \left(\dot{\mathbf{x}} \cdot \mathbf{n} \right)^2 + (\mathcal{C}_2(\mathbf{p}))^2 \| \dot{\mathbf{n}} \|^2,$$
(33)

for all $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}$ and all $\dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in T_{\mathbf{p}}(\mathbb{M})$ with constraint $\mathbf{n} = \pm \frac{\dot{\mathbf{x}}}{\|\dot{\mathbf{x}}\|}$ recall Fig. 2.

Proof. From our assumptions on **p** and $\gamma_{\varepsilon}^{*}(t)$ for $\varepsilon < \varepsilon_{0}$, we have, recall Remark 8, that $(U_{\varepsilon}(\cdot))^{2}$ is differentiable at $\gamma_{\varepsilon}^{*}(t)$ for all $0 \le t \le 1$ and $0 \le \varepsilon < \varepsilon_{0}$. This implies that U_{ε} is differentiable at $\{\gamma_{\varepsilon}^{*}(t) \mid 0 < t \le 1\}$, for all $0 < \varepsilon < \varepsilon_{0}$.

From Theorem 2 we have pointwise convergence $U_{\varepsilon}(\mathbf{p}) \to U_0(\mathbf{p})$ and uniform convergence $\gamma_{\varepsilon}^* \to \gamma_0^*$ as $\varepsilon \downarrow 0$. Moreover, as γ_{ε}^* and γ_0^* are solutions of the canonical ODEs of Pontryagin's Maximum Principle,

the trajectories are continuously depending on $\varepsilon > 0$, and so are the derivatives $\dot{\gamma}_{\varepsilon}^*$. As a result, we can apply the backtracking Theorem 4 for $\varepsilon > 0$ and take the limits:

$$\dot{\gamma}_{0}^{*}(t) = \lim_{\varepsilon \downarrow 0} \dot{\gamma}_{\varepsilon}^{*}(t)$$

$$\overset{\text{Thm. 4}}{=} \lim_{\varepsilon \downarrow 0} U_{\varepsilon}(\mathbf{p}) \left(\mathcal{G}_{\gamma_{\varepsilon}^{*}(t);\varepsilon}^{-1} \mathrm{d}U_{\varepsilon})(\gamma_{\varepsilon}^{*}(t)) \right)$$

$$= U_{0}(\mathbf{p}) \left(\lim_{\varepsilon \downarrow 0} \mathcal{G}_{\gamma_{\varepsilon}^{*}(t);\varepsilon}^{-1} \right) \left(\lim_{\varepsilon \downarrow 0} (\mathrm{d}U_{\varepsilon}(\gamma_{\varepsilon}^{*}(t))) \right)$$

$$\overset{\text{Thm. 2}}{=} U_{0}(\mathbf{p}) \mathcal{G}_{\gamma_{\varepsilon}^{*}(t);0}^{-1} (\mathrm{d}U_{0})(\gamma_{0}^{*}(t)).$$

$$(34)$$

Furthermore,

$$1 = \lim_{\varepsilon \downarrow 0} \sqrt{\mathcal{G}_{\mathbf{p};\varepsilon} \left(\mathcal{G}_{\mathbf{p};\varepsilon}^{-1} \mathrm{d}U_{\varepsilon}(\mathbf{p}), \mathcal{G}_{\mathbf{p};\varepsilon}^{-1} \mathrm{d}U_{\varepsilon}(\mathbf{p}) \right)} = \sqrt{\mathcal{G}_{\mathbf{p};0} \left(\mathcal{G}_{\mathbf{p};0}^{-1} \mathrm{d}U_{0}(\mathbf{p}), \mathcal{G}_{\mathbf{p};0}^{-1} \mathrm{d}U_{0}(\mathbf{p}) \right)}$$

where we recall Corollary 1. Here due to our assumptions, U_{ε} and U_0 are both differentiable at **p**. Note that the limit for the inverse metric $\mathcal{G}_{\mathbf{p},\varepsilon}^{-1}$ as $\varepsilon \downarrow 0$ exists, recall Cor. 1.

Now that we stated our 4 main theoretical results we will prove them in the subsequent sections (and Appendix A). Before that, we would like to study the optimal geodesic procedure:

3 Computing Optimal Geodesics in \mathbb{M}

3.1 The Procedure for Computing Optimal Geodesics in $\mathbb M$

Let us consider the procedure for computing optimal geodesics in metric space $(\mathbb{M}, d_{\mathcal{F}})$:

- Step 1: Compute the dual Finsler function \mathcal{F}^* .
- **Step 2:** Derive the eikonal system (5).
- Step 3: Obtain the distance map as the viscosity solution of the eikonal system.
- Step 4: Compute the derivative $d\mathcal{F}^*$ of the dual Finsler function (w.r.t. 2nd entry).
- Step 5: Apply backtracking via an intrinsic gradient descent (24) on the distance map.

for two special (uniform cost) cases for which we can derive exact analytic solutions:

- 1. The isotropic Riemannian case $(\mathbb{M}, d_{\mathcal{F}_1})$ with $\xi^{-1}\mathcal{C}_1 = \alpha^{-1}\mathcal{C}_2 = 1$, with $\xi > 0$, $\alpha > 0$ and $d \in \{2, 3\}$ and $\epsilon = 1$.
- 2. The sub-Riemannian case $(\mathbb{M}, d_{\mathcal{F}_0})$ with $\xi^{-1}\mathcal{C}_1 = \mathcal{C}_2 = 1$, with $\xi > 0$ and $d \in \{2, 3\}$ and $\epsilon = 0$.

3.1.1 Intermezzo: The General Procedure for setting up Geodesic Equations in Riemannian manifolds

Let (M, \mathcal{G}) be a smooth Riemannian manifold of dimension n with a given smooth parametrization

$$\mathbf{p} = X(x^1, \dots, x^n)$$

and with a metric tensor field $\mathbf{p} \mapsto \mathcal{G}|_{\mathbf{p}} = \sum_{ij=1}^{n} g_{ij}(\mathbf{p}) dx^{i}|_{\mathbf{p}} \otimes dx^{j}|_{\mathbf{p}}$ that is expressed in the dual basis dx^{i} such that

$$\langle \mathrm{d}x^i \Big|_{\mathbf{p}}, \, \partial_{x^j} \Big|_{\mathbf{p}} \rangle = \delta^i_j,$$

for all $\mathbf{p} \in M$. Here we briefly recall the standard ODE for computing the geodesics on the smooth Riemannian manifold (M, \mathcal{G}) : The ODE's (in the given coordinate system) that describe the geodesics

$$t \mapsto \gamma(t) := X(x^1(t), \dots, x^n(t))$$

are given by :

$$\ddot{x}^i + \sum_{k,j=1}^n \Gamma_{kj}^i \dot{x}^k \dot{x}^j = 0 \text{ for all } i = 1..., n,$$

where the coordinate-dependent(!) Christoffel symbols⁹ Γ_{ij}^k are given by

$$\partial_{x^i}\partial_{x^j}X = \sum_{k=1}^n \Gamma_{ij}^k(X(\cdot)) \ \partial_{x^k}X.$$

It can be shown that

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{l=1}^n g^{kl} (\partial_{x^i} g_{jl} + \partial_{x^j} g_{li} - \partial_{x^l} g_{ij}).$$

where g^{ij} denotes the components of the inverse matrix of matrix $[g_{ij}]$.

Exercise 5 Let $\mathbf{n}_0 \in S^2$ and $\mathbf{n} \in S^2$. Consider the curve optimization problem on S^2 :

$$d_{S^{2}}(\mathbf{n}_{0}, \mathbf{n}) = \inf_{\substack{\mathbf{n}(\cdot) \in \operatorname{Lip}([0, 1], S^{2}) \\ \mathbf{n}(0) = \mathbf{n}_{0}, \\ \mathbf{n}(1) = \mathbf{n}}} \int_{0}^{1} \|\mathbf{n}'(s)\| \, \mathrm{d}s = \arccos(\mathbf{n}_{0} \cdot \mathbf{n}).$$
(35)

- a Derive the equations for geodesic back-tracking (24) without relying on Euler angles.
- b Express the geodesic equations for the specific case $\mathbf{n}_0 = \mathbf{e}_z = (0, 0, 1)$, and only for geodesics that do not pass $(0, 0, \pm 1)$, in standard Euler angles $\mathbf{n}(\beta, \gamma) = (\cos \gamma \sin \beta, \sin \gamma \sin \beta, \cos \beta)^T$, $\beta \in (0, \pi)$, $\gamma \in [0, 2\pi)$.
- c Show that the geodesics at b. lay on great circles.
- d What is the 1st Maxwell time for the geodesics on S^2 departing from a fixed point \mathbf{n}_0 ?
- e What is the 1st conjugate time for the geodesics on S^2 departing from a fixed point \mathbf{n}_0 ?
- f What is the cut-locus for problem (35)?

3.2 Two examples with analytic solutions

3.2.1 The isotropic Riemannian uniform cost case

Set $\xi^{-1}C_1 = \alpha^{-1}C_2 = 1$, with $\xi > 0$, $\alpha > 0$ and $d \in \{2, 3\}$ and $\epsilon = 1$. Then due to the isotropy $\epsilon = 1$ the Finsler function is no longer truly dependent on **p**:

$$\mathcal{F}_1(\mathbf{p}, \dot{\mathbf{p}}) = \sqrt{\xi^2 \|\dot{\mathbf{x}}\|^2 + \alpha^2 \|\dot{\mathbf{n}}\|^2}.$$

Next we follow the 5 steps of the backtracking procedure and derive the exact solutions. **Step 1:** Compute the dual Finsler function:

$$\mathcal{F}_{1}^{*}(\mathbf{p}, \hat{\mathbf{p}}) = \sqrt{\xi^{-2} \|\hat{\mathbf{x}}\|^{2} + \alpha^{-2} \|\hat{\mathbf{n}}\|^{2}} \text{ with } \hat{\mathbf{p}} = (\hat{\mathbf{x}}, \hat{\mathbf{n}}) \in T_{\mathbf{p}}^{*}(\mathbb{M}).$$
(36)

where $T^*_{\mathbf{p}}(\mathbb{M})$ is the dual space to $T_{\mathbf{p}}(\mathbb{M})$. Step 2: Derive the eikonal system:

$$\begin{cases} \sqrt{\xi^{-2} \|\nabla_{\mathbb{R}^d} U(\mathbf{p})\|^2 + \alpha^{-2} \|\nabla_{S^{d-1}} U(\mathbf{p})\|^2} = 1 & \text{for all } \mathbf{p} \in \mathbb{M} \setminus \{\mathbf{p}_S\}, \\ U(\mathbf{p}_S) = 0. \end{cases}$$
(37)

Step 3: Derive the viscosity solution:

$$U(\mathbf{p}) = \sqrt{\xi^2 \|\mathbf{x} - \mathbf{x}_0\|^2 + \alpha^2 |d_{S^{d-1}}(\mathbf{n}_0, \mathbf{n})|^2}, \text{ with } \mathbf{p} = (\mathbf{x}, \mathbf{n}), \mathbf{p}_S = (\mathbf{x}_0, \mathbf{n}_0),$$
(38)

where $d_{S^{d-1}}$ denotes the usual distance on S^{d-1} . **Step 4:** Compute the derivative of the dual Finsler function (w.r.t. 2nd entry only):

$$d\mathcal{F}_{1}^{*}(\mathbf{p}, \hat{\mathbf{p}}) = \frac{1}{\sqrt{\xi^{-2} \|\hat{\mathbf{x}}\|^{2} + \alpha^{-2} \|\hat{\mathbf{n}}\|^{2}}} (\xi^{-2} \hat{\mathbf{x}}, \alpha^{-2} \hat{\mathbf{n}}), \text{ with } \hat{\mathbf{p}} = (\hat{\mathbf{x}}, \hat{\mathbf{n}}).$$

⁹ Associated to the usual torsion-free, metric compatible Levi-Civita connection.

Step 5: Apply backtracking. In addition to (24) we reverse time $t \mapsto 1-t$ as this is how it is implemented (i.e. backtracking from end-point back to source point $\mathbf{p}_{S} = (\mathbf{x}_{0}, \mathbf{n}_{0})$:

$$\begin{cases} \dot{\mathbf{x}}(t) = -d_{\mathcal{F}_{1}}(\mathbf{p}, \mathbf{p}_{S}) \, \xi^{-2} \nabla_{\mathbb{R}^{d}} U(\mathbf{x}(t), \mathbf{n}(t)), & t \in [0, 1] \\ \dot{\mathbf{n}}(t) = -d_{\mathcal{F}_{1}}(\mathbf{p}, \mathbf{p}_{S}) \, \alpha^{-2} \nabla_{S^{d-1}} U(\mathbf{x}(t), \mathbf{n}(t)), \\ \mathbf{x}(0) = \mathbf{x}, \\ \mathbf{n}(0) = \mathbf{n} \end{cases}$$

$$(39)$$

which has the following exact solutions

$$\mathbf{x}(t) = \mathbf{x}_0 + (1-t)(\mathbf{x} - \mathbf{x}_0),$$

$$\mathbf{n}(t) = R_{\mathbf{n}_0 \to \mathbf{n}, (1-t) d_{S^{d-1}}(\mathbf{n}_0, \mathbf{n})} (\mathbf{n}_0),$$
(40)

where $R_{\mathbf{n}_0 \to \mathbf{n}, \psi}$ denotes the in-planar counter-clockwise rotation in the plane spanned by \mathbf{n}_0 and \mathbf{n} about angle ψ . So we conclude that the spatial part of the geodesics describe straight-line motions, whereas the angular part of the geodesics describe great circle motions. In this case the minimal paths are not dependent on ξ and α , in contrast to the distance. We also see that for $\xi > 0$ and $\alpha > 0$ there are no Maxwell-points.

Exercise 6 a Verify all steps of the above procedure (and verify in particular Eq. (36), (39) and (40)). Here you may restrict yourself to the case d = 2.

- b What happens if $\xi > 0$ stays fixed and $\alpha \to \infty$?
- c What happens to the equations when considering non-uniform cost: $1 \neq \xi^{-1} \mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C} \in C^1([\delta, \infty)) ?$
- d Why is it important to have a lower bound $\mathcal{C} \geq \delta > 0?$
- e Explicitly compute $\mathbf{R}_{\mathbf{n}_0 \to \mathbf{n}_1, td_{S^{d-1}}(\mathbf{n}_0, \mathbf{n}_1)}$ for d = 2.
- f Explicitly compute $\mathbf{R}_{\mathbf{n}_0 \to \mathbf{n}_1, td_{S^{d-1}}(\mathbf{n}_0, \mathbf{n}_1)}$ for d = 3 and $\mathbf{n}_0 \cdot \mathbf{n}_1 \ge 0$. g Show that if $\alpha > 0$ stays fixed and $\xi \to \infty$ the isotropic Riemannian curve optimization problem with uniform cost on \mathbb{M} is no longer controllable.

3.2.2 The sub-Riemannian uniform cost case

Let us consider the special case $(\mathbb{M}, d_{\mathcal{F}_0})$ with $\xi^{-1}\mathcal{C}_1 = \mathcal{C}_2 = 1$, with $\xi > 0$ and $d \in \{2, 3\}$ and $\epsilon = 0$. Intuitively, one considers a Reeds-Shepp vehicle with two degrees of freedom:

- Hit the gas via spatial control $\tilde{u}(t) := \dot{\mathbf{x}}(t) \cdot \mathbf{n}(t)$.
- Turn the wheel via angular velocity control $\dot{\mathbf{n}}(t)$.

The fact that $\epsilon = 0$ corresponds to infinite costs for the vehicle to move sidewards, recall Figure 2. Recall also that \mathcal{F}_0 was given by (2).

Next we follow the 5 steps of the backtracking procedure and we derive the exact solutions afterwards. For simplicity

Step 1: Compute the dual Finsler function:

$$\mathcal{F}_{0}^{*}(\mathbf{p}, \hat{\mathbf{p}}) = \sqrt{\xi^{-2} |\mathbf{n} \cdot \hat{\mathbf{x}}|^{2} + \alpha^{-2} \|\hat{\mathbf{n}}\|^{2}} \text{ with } \hat{\mathbf{p}} = (\hat{\mathbf{x}}, \hat{\mathbf{n}}) \in T_{\mathbf{p}}(\mathbb{M}).$$
(41)

Step 2: Derive the eikonal system:

$$\begin{cases} \sqrt{\xi^{-2} |\mathbf{n} \cdot \nabla_{\mathbb{R}^d} U(\mathbf{p})|^2 + \alpha^{-2} \|\nabla_{S^{d-1}} U(\mathbf{p})\|^2} = 1 & \text{for all } \mathbf{p} \in \mathbb{M} \setminus \{\mathbf{p}_S\}, \\ U(\mathbf{p}_S) = 0. \end{cases}$$
(42)

Step 3: Derive the viscosity solution. This must be done numerically [67, 56, 9]. We return to this in Section 7. For now we just write $U(\mathbf{p}) = d_{\mathcal{F}_0}(\mathbf{p}, \mathbf{p}_0)$.

Step 4: Compute the derivative of the dual Finsler function (w.r.t. 2nd entry only):

$$\mathrm{d}\mathcal{F}_0^*(\mathbf{p},\hat{\mathbf{p}}) = \frac{(\xi^{-2}\mathbf{n}\,\hat{\mathbf{x}}\cdot\mathbf{n},\hat{\mathbf{n}})}{\mathcal{F}_0^*(\mathbf{p},\hat{\mathbf{p}})}, \text{ with } \hat{\mathbf{p}} = (\hat{\mathbf{x}},\hat{\mathbf{n}}).$$

Step 5: Apply backtracking. In addition to Eq. (24) we reverse time $t \mapsto 1 - t$ as this is how it is implemented (i.e. backtracking from end-point $\mathbf{p} = (\mathbf{x}, \mathbf{n})$ back to source point $\mathbf{p}_{\rm S} = (\mathbf{x}_0, \mathbf{n}_0)$:

$$\begin{cases} \dot{\mathbf{x}}(t) = -U(\mathbf{p}) \ \frac{1}{\xi^2} \mathbf{n}(t) \ (\mathbf{n}(t) \cdot \nabla_{\mathbb{R}^d} U)(\mathbf{x}(t), \mathbf{n}(t)), & t \in [0, 1], \\ \dot{\mathbf{n}}(t) = -U(\mathbf{p}) \ \nabla_{S^{d-1}} U(\mathbf{x}(t), \mathbf{n}(t)), & \\ \mathbf{x}(0) = \mathbf{x}, & \\ \mathbf{n}(0) = \mathbf{n}. \end{cases}$$
(43)

Exercise 7 Verify each of the above steps (except for step 3).

Exercise 8 Indicate after a careful study of the next paragraph (where application of the Pontryagin Maximum Principle (PMP) provides exact solutions for the sub-Riemannian geodesics) for each step the relation to the Pontryagin Maximum Principle.

Exercise-Mathematica 1. Identify the 5 steps for computing optimal sub-Riemannian geodesics in SE(2) in the notebook "SE2_SRGeodesicsCeq1UpwindPDE.nb"

(downloadable at www.LieAnalysis.nl > Education > Lecture 5-Tracking > Eikonal Solver Upwind).In this notebook the Eikonal equation is solved using an iterative upwind scheme. Perform the tasks at the end of the notebook:

- Identify different types of geodesics, i.e. U-curves, S-curves, sub-Riemannian geodesics with and without cusps, Asymmetric Finsler geodesics with and without key-points.
- Locate Maxwell-points.

Remark: In case you do not have the latest version of Mathematica installed and still would like to see how the notebooks work in practice you can look at 'HTML Renders from the Notebooks' at http://www.lieanalysis.nl/education/,

where notebooks including output can be viewed (without running Mathematica).

Exercise-Mathematica 2. Reproduce the exercises of Exercise 1 using the notebook "SE2_SRGeodesicsCeq1FastMarching.nb"

(downloadable at www.LieAnalysis.nl > Education > Lecture 5-Tracking > Fast marching). In this notebook the Eikonal equation is solved using a fast marching solver.

3.2.3 Derivation of Exact Cuspless sub-Riemannian Geodesics via PMP

Let us derive the geodesics in the case $(\mathbb{M}, d_{\mathcal{F}_0})$ with $\xi^{-1}C_1 = C_2 = 1$, with $\xi > 0$, $\epsilon = 0$. In this case the geodesics are sub-Riemannian geodesics. But now let us, for the sake of simplicity, constrain ourselves to d = 2. In this section, for simplicity we set $\mathbf{p}_S = \mathbf{p}_0 = \mathbf{p}(0)$, (and do not apply backtracking like in (43).

The Pontryagin Maximum Principle (PMP) [1,2] starts with the derivation of the Hamiltonian. In our settings (of a 1-homogeneous Finsler function) the Hamiltonian for $\mathcal{F} = \mathcal{F}_0$ is by definition given by

$$\begin{split} \mathfrak{H}(\mathbf{p}, \hat{\mathbf{p}}) &:= \sup_{\dot{\mathbf{p}} \in T_{\mathbf{p}}(\mathbb{M})} \left\{ \langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle - (\mathcal{F}_{0}(\mathbf{p}, \dot{\mathbf{p}}))^{2} \right\} \\ &= \sup_{\dot{\mathbf{p}} \in T_{\mathbf{p}}(\mathbb{M})} \left\{ \langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle - \mathcal{G}|_{\mathbf{p};0} \left(\dot{\mathbf{p}}, \dot{\mathbf{p}} \right) \right\}. \end{split}$$
(44)

Now we can compute this in a simple way by introducing a left-invariant frame and corresponding leftinvariant coordinate system, as this is a moving frame of reference intrinsically attached to the Reeds-Shepp car. We write

$$\dot{\mathbf{p}} = u^1 \,\mathcal{A}_1|_{\mathbf{p}} + u^2 \,\mathcal{A}_2|_{\mathbf{p}} + u^3 \,\mathcal{A}_3|_{\mathbf{p}} \text{ and } \hat{\mathbf{p}} = \lambda_1 \,\omega^1 \Big|_{\mathbf{p}} + \lambda_2 \,\omega^2 \Big|_{\mathbf{p}} + \lambda_3 \,\omega^3 \Big|_{\mathbf{p}} \tag{45}$$

with left-invariant vector fields and duals

$$\mathcal{A}_{1} = \cos\theta \,\partial_{x} + \sin\theta \,\partial_{y}, \ \mathcal{A}_{2} = -\sin\theta \,\partial_{x} + \cos\theta \,\partial_{y}, \ \mathcal{A}_{3} = \partial_{\theta}, \omega^{1} = \cos\theta \,dx + \sin\theta \,dy, \ \omega^{2} = -\sin\theta \,dx + \cos\theta \,dy, \ \omega^{3} = d\theta.$$
(46)

where we use the common short notation $\partial_x := \frac{\partial}{\partial x}$. Recall that

$$\langle \omega^i \Big|_{\mathbf{p}}, \mathcal{A}_j \Big|_{\mathbf{p}} \rangle = \delta^i_j, \tag{47}$$

with δ_j^i equals 1 if i = j and 0 if $i \neq j$.

Definition 5 We use the following short notation:

$$\dot{\gamma}^{i}(t) := \left\langle \left. \omega^{i} \right|_{\gamma(t)}, \dot{\gamma}(t) \right\rangle \text{ for } i = 1, 2, 3 \text{ and all } t \in [0, 1].$$

Exercise 9 Use (45) and (47) to verify that

$$u^{i}(t) = \dot{\gamma}^{i}(t) \text{ for } i = 1, 2, 3.$$
 (48)

Set distribution $\Delta := \text{span}\{\mathcal{A}_1, \mathcal{A}_3\}$ and consider sub-Riemannian manifold $(SE(2), \Delta, \mathcal{G})$. Then (44) reduces to a Fenchel transform on \mathbb{R}^2 :

$$\mathfrak{H} = \sup_{(u^1, 0, u^3)^T \in \mathbb{R}^3} \{ u^1 \lambda_1 + u^3 \lambda_3 - \xi^2 |u^1|^2 - |u^3|^2 \} = \frac{1}{2} \sum_{i \in \{1, 3\}} \frac{|u^i|^2}{|\xi^i|^2} = \frac{1}{2} \left(\xi^{-2} \lambda_1^2 + \lambda_3^2 \right) = \frac{1}{2}$$
(49)

with $\xi_1 = \xi, \, \xi_3 = 1.$

Exercise 10 In case d = 2 we have $\mathbb{M} \equiv SE(2)$ the roto-translation group on the plane that acts transitively on \mathbb{M} by

$$(\mathbf{x}, R) \odot (\mathbf{x}', \mathbf{n}') = (R\mathbf{x}' + \mathbf{x}, R\mathbf{n}').$$

with group product

$$(\mathbf{x}, R)(\mathbf{x}', R') = (R\mathbf{x}' + \mathbf{x}, RR').$$
(50)

Identify θ with R_{θ} the counter clockwise rotation about angle θ . Let $e = (0, 0, \theta = 0) \in \mathbb{M}$ be the unity element.

- Verify that the left-invariant vector fields on SE(2) are given by push-forward of the left-multiplication $L_gh = gh$

$$\mathcal{A}_i|_g = (L_g)_* A_i$$
 with $A_1 = \partial_x|_e$, $A_2 = \partial_y|_e$, $A_3 = \partial_\theta|_e$

- Verify that they can also be obtained by infinitesimal generators of the right-regular representation $\mathcal{R}_g U(h) = U(hg)$ for all $g, h \in SE(2)$:

$$\mathcal{A}_i = \mathrm{d}\mathcal{R}(A_i) \text{ with } (\mathrm{d}\mathcal{R}(A_i)U)(g) = \lim_{h \downarrow 0} \frac{U(ge^{hA_i}) - U(g)}{h} \ i = 1, 2, 3 \text{ for all } U : SE(2) \to \mathbb{R} \text{ smooth } .$$

- Explain why the above procedures to compute a basis for the Lie-algebra $\mathcal{L}(G)$ of left-invariant vector fields work on Lie groups G in general, and verify that $d\mathcal{R}: T_e(G) \to \mathcal{L}(G)$ is a Lie algebra isomorphism.

Exercise 11 The definition of the Hamiltonian is a coordinate free definition. Verify this explicitly with a similar computation in the fixed frame of reference with

$$\dot{p} = (\dot{x}, \dot{y}, \dot{ heta}) = \dot{x} \, \partial_x + \dot{y} \, \partial_y + \dot{ heta} \, \partial_{ heta} ext{ and } \hat{p} = p_x \, \mathrm{d}x + p_y \, \mathrm{d}y + p_{ heta} \, \mathrm{d} heta$$

where you must obtain the same Hamiltonian (expressed in a different coordinate system).

Now that the Hamiltonian \mathfrak{H} is explicitly computed (49) we compute its lift (the corresponding Hamiltonian vector field) $\overrightarrow{\mathfrak{H}}$ on the cotangent bundle given by

$$\forall_{\dot{\mathbf{p}}\in T_{\mathbf{p}}(\mathbb{M})}\forall_{\dot{\lambda}\in T(T^*_{\mathbf{p}}(M))} : \sigma(\vec{\mathfrak{H}}(\mathbf{p},\hat{\mathbf{p}}), (\dot{\mathbf{p}},\dot{\lambda})) = \langle \mathrm{d}_{\dot{\mathbf{p}}}\mathfrak{H}(\mathbf{p},\hat{\mathbf{p}}), (\dot{\mathbf{p}},\dot{\lambda})) \rangle \text{ with canonical volume form } \sigma = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\hat{\mathbf{p}}$$

or shortly $\sigma(\vec{\mathfrak{H}}(\mathbf{p}, \hat{\mathbf{p}}), \cdot) = \langle d_{\hat{\mathbf{p}}} \mathfrak{H}(\mathbf{p}, \hat{\mathbf{p}}), \cdot \rangle$, and write the so-called **Hamiltonian flow** on the cotangent bundle $T^*(\mathbb{M}) = \{(\mathbf{p}_0, \hat{\mathbf{p}}_0) \mid \hat{\mathbf{p}}_0 \in T^*_{\mathbf{p}_0}(\mathbb{M})\}$ for SR-geodesics:

$$\begin{cases} \frac{d}{d\tau} \begin{pmatrix} \mathbf{p}(\tau) \\ \hat{\mathbf{p}}(\tau) \end{pmatrix} = \overrightarrow{\mathfrak{H}}(\mathbf{p}(\tau)), \\ \mathbf{p}(0) = \mathbf{p}_0, \\ \hat{\mathbf{p}}(0) = \hat{\mathbf{p}}_0 \text{ subject to } \mathfrak{H}(\mathbf{p}_0, \hat{\mathbf{p}}_0) = \frac{1}{2} \end{cases}$$
(51)

Remark 10 Throughout the remainder of this section we will use sub-Riemannian arclength parameter

$$\tau = t d_{\mathcal{F}_0}(\mathbf{p}_S, \mathbf{p}) \in [0, d_{\mathcal{F}_0}(\mathbf{p}_S, \mathbf{p})] \Leftrightarrow t \in [0, 1],$$

as this is more convenient for analytic (rather than numeric) computations.

The first component of this vector ODE gives the flow on the base manifold \mathbb{M} for the state variables and is therefore called "the horizontal part", whereas the second component gives the flow of the momentum variables and is therefore called "the vertical part". In explicit left-invariant coordinates (45) this geometric flow can be written (where we write $\gamma(\tau) = \mathbf{p}(\tau)$ for the geodesics) in a relatively simple form:

$$\begin{cases} \dot{\gamma}^{1}(\tau) = u^{1}(\tau) = \xi^{-2}\lambda_{1}(\tau), \\ \dot{\gamma}^{2}(\tau) = u^{2}(\tau) = 0, \\ \dot{\gamma}^{3}(\tau) = u^{3}(\tau) = \lambda_{3}(\tau) \\ \dot{\lambda}_{1}(\tau) = \lambda_{2}(\tau) u^{3}(\tau) = \lambda_{2}(\tau) \lambda_{3}(\tau), \\ \dot{\lambda}_{2}(\tau) = -\lambda_{1}(\tau) u^{3}(\tau) = -\lambda_{1}(\tau) \lambda_{3}(\tau), \\ \dot{\lambda}_{3}(\tau) = \lambda_{2}(\tau) u^{1}(\tau) = -\xi^{-2} \lambda_{2}(\tau) \lambda_{1}(\tau), \\ \lambda_{1}(0) = \lambda_{10}, \\ \lambda_{2}(0) = \lambda_{20}, \\ \lambda_{3}(0) = \lambda_{30}, \\ \text{with } \xi^{-2}\lambda_{10}^{2} + \lambda_{30}^{2} = 1. \\ \gamma(0) = \mathbf{p}_{0}, \end{cases}$$
(52)

Remark 11 The Hamiltonian flow is a linear operator and therefore it is denoted by $e^{t \cdot \hat{y}}$:

$$\begin{aligned} (\gamma(\tau), \lambda(\tau)) &= e^{\tau \, \vec{\mathfrak{S}}} \, (\gamma(0), \lambda(0)), \\ \text{where we use the notation} \\ (\gamma(\tau), \lambda(\tau)) &= (\mathbf{p}(\tau), \hat{\mathbf{p}}(\tau)). \end{aligned}$$

However, in practice (e.g. when solving a boundary value problem) one wants to solve for the geodesic $\gamma(\cdot)$ separately without knowledge of the momentum $\lambda(\cdot)$. Then a serious challenge arises as the Hamiltonian flow couples momentum λ and velocity $\dot{\gamma}$ whose components are given by (45).

To tackle this problem, the general idea is to first isolate the momentum in a nonlinear ODE that only involves momentum (like we did in the middle three equations of (52)), then solve this ODE, and then substitute these momentum solutions in the first 3 equations of (52). This gives the velocity of the geodesics. Finally, the geodesics themselves follow by integration of these velocities.

The Exponential Map of the Hamiltonian Flow and a Fundamental Symmetry

In the next few exercises we are going to

- 1. derive and understand the PDE system (52),
- 2. give an important geometric interpretation for PDE system (52),
- 3. analytically solve PDE system (52) for 'cuspless' geodesics.

Before that we define 2 typical tools from geometric control theory:

1.) The solver that solves $(51) \Leftrightarrow (52)$ is typically denoted by the exponential map¹⁰ of the geometrical control problem. as follows

$$\gamma(\tau) = (x(\tau), y(\tau), \theta(\tau)) = EXP_{\mathbf{p}_0}(\tau, \hat{\mathbf{p}}_0) = \Pi \circ e^{\tau \vec{\mathfrak{H}}} \begin{pmatrix} \mathbf{p}_0 \\ \hat{\mathbf{p}}_0 \end{pmatrix},$$

 $^{^{10}}$ This exp map should not be mistaken for the usual exponential map from Lie algebra to Lie group! For an intuitive geometric visualization of the exponential map (with a domain restriction such that only endpoints of 'cuspfree' geodesics appear in the range) see [24, Fig.11,Thm.6].

with manifold projection $\Pi(\mathbf{p}, \hat{\mathbf{p}}) = \mathbf{p}$.

2.) It can be shown that

$$\forall_{h_0 \in SE(2)} \forall_{\tau \in \mathbb{R}} : EXP_{h_0 \mathbf{p}_0}(\tau, \operatorname{CoAd}(h_0)\hat{\mathbf{p}}_0) = h_0 EXP_{\mathbf{p}_0}(\tau, \hat{\mathbf{p}}_0),$$
(53)

where the co-adjoint action is given by $\operatorname{CoAd}(h_0) = (\operatorname{Ad}(h_0^{-1}))^*$. Here the adjoint-representation is given by the push-forward (recall Part I) of the conjugation operator, i.e. $\operatorname{Ad}(g) = (L_g R_{g^{-1}})^*$, where $L_g \mathbf{p} = g\mathbf{p}$ and $R_g \mathbf{p} = \mathbf{p}g^{-1}$ where we identify \mathbb{M} with the Lie group SE(2) equipped with group product (50). This invariance property can be used to select a single element on the co-adjoint orbit

$$\{\operatorname{CoAd}(h_0)\hat{\mathbf{p}}_0 \mid h_0 \in SE(2)\}$$

that allows for simple computation of $\tilde{\gamma}_{min}(\tau)$ after which the true minimizer is found by¹¹

$$\gamma_{min}(\tau) = h_0^{-1} \tilde{\gamma}(\tau).$$

Exercise 12 Two questions on the Hamiltonian system (52) structure. Let us restrict ourselves to $\xi = 1$. a) Compute the structure constants c_{ij}^k of the Lie algebra which are given by

$$[\mathcal{A}_i, \mathcal{A}_j] := \mathcal{A}_i \mathcal{A}_j - \mathcal{A}_j \mathcal{A}_i = \sum_{k=1}^3 c_{ij}^k \mathcal{A}_k,$$
(54)

b) Express the (left-invariant) sub-Riemannian metric tensor field (33) as

$$\mathcal{G}_0 = \sum_{i,j \in \{1,3\}} g_{ij} \; \omega^i \otimes \omega^j,$$

where we recall (46), and compute the constants components g_{ij} w.r.t. the left-invariant frame. c) Show that the ODE in (52) is given by¹²

$$\begin{split} \dot{\gamma}^i &= \lambda^i := \lambda_i \text{ for } i \in \{1,3\} \\ \dot{\gamma}^2 &= 0, \\ \dot{\lambda}_i &= -\sum_{k=1}^3 \sum_{j \in \{1,3\}} c^k_{ij} \lambda_k \lambda^j \text{ for } i \in \{1,2,3\} \end{split}$$

Exercise 13 Questions on different fundamental approaches to the Hamiltonian system (52):

a) Poisson brackets are given by

$$\{f,g\} = \sum_{i=1}^{3} \left(\frac{\partial f}{\partial \hat{p}_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial \hat{p}_i} \right)$$

Show that

$$\{f,g\} = \sum_{i=1}^{3} \left(\frac{\partial f}{\partial \lambda_i} \mathcal{A}_i g - \frac{\partial g}{\partial \lambda_i} \mathcal{A}_i f \right)$$

and show that the ODE in (52) can be computed by¹³

$$\begin{cases} u^{i} = \{\mathfrak{H}, u^{i}\}, & \text{(horizontal part)}\\ \dot{\lambda}_{i} = \{\mathfrak{H}, \lambda_{i}\}, & \text{(vertical part)}, \end{cases}$$

for i = 1..., 3.

¹¹ Note that the Hamiltonian (49) is affected when such a left shift is applied. ¹² with $\lambda^j = \sum_{j' \in \{1,3\}} g^{jj'} \lambda_{j'} = (\xi^j)^{-1}$ with $\xi^1 = \xi = 1$ and $\xi^2 = 1$ since we have a diagonalized inverse SR-metric on the dual Δ^* of the distribution Δ .

 $^{^{13}\,}$ In fact this is the common way to derive the canonical equations.

b) Study the left Cartan connection $\overline{\nabla}$ in Appendix D. Show that the Hamiltonian system of the Pontryagin maximum principle (52) is equivalent to

$$\begin{cases} \dot{\gamma} = \mathcal{G}_0^{-1} P_{\Delta *} \lambda & \text{(horizontal part)}, \\ \overline{\nabla}^* \Big|_{\dot{\gamma}} \lambda = 0 & \text{(vertical part)}. \end{cases}$$

where λ denotes a covector field, i.e. a section in the cotangent bundle

$$T^*(\mathbb{M}) = \{ (\mathbf{p}, \hat{\mathbf{p}}) \mid \mathbf{p} \in \mathbb{M}, \hat{\mathbf{p}} \in T^*_{\mathbf{p}}(\mathbb{M}) \}.$$

- c) Compute the auto-parallels $\overline{\nabla}_{\dot{\gamma}}\dot{\gamma}=0$ in the sub-Riemannian manifold.
- d) Give a geometric interpretation of the above two exercises (hint: see Theorem 7 and Figure 20 in Appendix D).

Exercise 14 Derive a complete set of functionally independent first integrals (preservation laws) of the Hamiltonian system (52).

Exercise 15 The Hamiltonian system (52) can be solved analytically [57]. The elegant general formulas are technical and involve many Jacobi functions. It boils down to phase portrait analysis of the mathematical pendulum. In this exercise we set $\xi = 1$.

a.) Show that the vertical part of the Hamiltonian equations boils down to the well-known mathematical pendulum ODE given by

$$\ddot{\nu}(\tau) = -\sin\nu(\tau), \quad \nu(\tau) \in (-\pi, 3\pi).$$

hint: apply the substitution $\nu = 2 \cdot \arg\{-\lambda_3 + i\lambda_1\}$ and $c = \dot{\nu} = -2 \cdot \lambda_2$.

b.) Show that under the above substitution the horizontal part of the Hamiltonian equations is given by

$$\begin{aligned} \dot{x} &= \sin \frac{\nu}{2} \, \cos \theta \, , \\ \dot{y} &= \sin \frac{\nu}{2} \, \sin \theta \, , \\ \dot{\theta} &= -\cos \frac{\nu}{2} . \end{aligned}$$

c.)] Plot the phase portrait of the mathematical pendulum ODE above. What can be said about the shape of the spatial projections of the sub-Riemannian geodesics?

Exercise 16 For the specific case of sub-Riemannian geodesics γ whose spatial projections do not exhibit cusps one can obtain a simpler analysis of the solutions [24,28] as only for these 'cuspfree' geodesics one can use spatial arclength parametrization:

$$s \mapsto \gamma(\tau(s)) = (\mathbf{x}(s), \theta(s)), \text{ with } \|\frac{d}{ds}\mathbf{x}(s)\| = 1.$$

In this exercise we will derive these solutions of the 'cuspless' sub-Riemannian geodesics step by step. Recall that the 'cuspless' sub-Riemannian geodesics in $(\mathbb{M} \equiv SE(2), d_{\mathcal{F}_0})$ (with forward spatial control) are minimizers of the following problem

$$d_{\mathcal{F}_0}(\mathbf{p}_0, \mathbf{p}) = \min_{\substack{\gamma = (\mathbf{x}(\cdot), \mathbf{n}(\cdot)) \in \operatorname{Lip}([0, 1], \mathbb{M}) \\ \gamma(0) = \mathbf{p}_0 \\ \gamma(1) = \mathbf{p} = (x, y, \theta), \\ \dot{\mathbf{x}} \propto \mathbf{n}_i} \int_0^1 \mathcal{F}_{0,\xi}(\gamma(t), \dot{\gamma}(t)) \, \mathrm{d}t \;,$$

with Finsler function 14 :

$$\left. \mathcal{F}_{0,\xi} \right|_{\mathbf{p}} \left(\dot{\mathbf{p}}, \dot{\mathbf{p}} \right) = \begin{cases} \sqrt{\xi^2 |\dot{\mathbf{x}} \cdot \mathbf{n}|^2 + \|\dot{\mathbf{n}}\|^2} \text{ if } \dot{\mathbf{x}} \propto \mathbf{n} \\ \infty & \text{else.} \end{cases}$$

for all $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M} \equiv SE(2), \, \dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in T_{\mathbf{p}}(\mathbb{M})$. We assume that \mathbf{p}_0 is chosen such¹⁵ that the minimizer $\gamma(\cdot) = (\mathbf{x}(\cdot), \mathbf{n}(\cdot) = (\cos \theta(\cdot), \sin \theta(\cdot)))$ has a strictly forward spatial control i.e. $\mathbf{n}(\cdot) = + \|\dot{\mathbf{x}}\|^{-1} \dot{\mathbf{x}}(\cdot)$.

By rotation and translation covariance of the curve optimization problem we will restrict ourselves to the case $\mathbf{p}_0 = \mathbf{0} := (0, 0, 0)$.

a) Show that the minimizing geodesic solution $[0,1] \ni t \mapsto \gamma_{\min}^{\xi}(t)$ connecting **0** to $\mathbf{p} = (x, y, \theta)$ for $\xi > 0$ relates to the minimizing geodesic solution $[0,1] \ni t \mapsto \gamma_{\min}^{\xi=1}(t)$ connecting **0** and $\mathbf{p} = (\xi x, \xi y, \theta)$ for $\xi = 1$ as follows:

$$\gamma_{\min}^{\xi}(t) = (x_{\min}^{\xi}(t), y_{\min}^{\xi}(t), \theta_{\min}^{\xi}(t)) \text{ and } \gamma_{\min}^{\xi=1}(t) = (x_{\min}^{1}(t), y_{\min}^{1}(t), \theta_{\min}^{1}(t)), \\
\text{with} \\
(x_{\min}^{\xi}(\cdot), y_{\min}^{\xi}(\cdot), \theta_{\min}^{\xi}(\cdot)) = (\xi^{-1}x_{\min}^{1}(\cdot), \xi^{-1}y_{\min}^{1}(\cdot), \theta_{\min}^{1}(\cdot)).$$
(55)

and check for physical dimensions. Note that in this item a), in contrast to the subsequent items, it is convenient to parameterize by rescaled sub-Riemannian arclength $t = \tau/d_{\mathcal{F}_0}((x, y, \theta), (0, 0, 0)) \in [0, 1]$ in order to precisely match the parameterizations in (55).

Therefore we will restrict ourselves to $\xi = 1$ and sub-Riemannian arclength parametrization $\tau \mapsto \gamma(\tau)$ in the remainder of this exercise.

b) Show that (for 'cuspfree' geodesics¹⁶)

$$\tau(s) = \int_{0}^{s} \sqrt{\kappa^{2}(\tilde{s}) + 1} \,\mathrm{d}\tilde{s} \text{ and } \tau'(s) = \lambda_{1}^{-1}(\tau(s))$$
(56)

where s > 0 denotes the spatial arc-length, and $\kappa(\cdot)$ denotes the curvature, of the spatially projected geodesic $\mathbf{x}(\cdot)$.

hint: Use the ODE system (52) and recall (49).

- c) Use (56) to express the vertical part of the Hamiltonian system in s parametrization.
- d) Plot the phase portrait of the orbits $(\lambda_3(s), \lambda_2(s))$ in the space $|\lambda_3| < 1$.

Let the initial momentum $\hat{\mathbf{p}}(0) = \hat{\mathbf{p}}_0 := \sum_{i=1}^3 \lambda_i(0) \omega^i \big|_{\mathbf{p}(0)}$ be given such that $\mathfrak{H}(\hat{\mathbf{p}}_0) = \frac{1}{2}$, and recall $\mathbf{p}_0 = \mathbf{0}$.

¹⁴ So we set constant relative costs: $C_1 = \xi > 0$ and $C_2 = 1$ in (2).

 $^{^{15}\,}$ The allowable set \Re for \mathbf{p}_0 is fully characterized in [24, Thm.9] and part of it is depicted in Figure 8.

 $^{^{16}}$ We write "cuspfree" geodesics, as the sub-Riemannian geodesics themselves do not exhibit cusps, their spatial projections do, recall Fig. 7. For those geodesics spatial arclength parametrization does not break down.

- e) Compute the total spatial length s_{max}(p̂₀) towards the (first) cusp.
 hint: for ξ = 1 one has by the Hamiltonian (49) that λ₁(τ) = 0 ⇔ λ₃(τ) = 1
- f) For which $\hat{\mathbf{p}}_0$ do we get S-shaped curves and for which $\hat{\mathbf{p}}_0$ do we get U-shaped curves? hints: check out the sign of the signed curvature $\kappa(s)$, use $\lambda_3(s) = \frac{\kappa(s)}{\sqrt{(\kappa(s))^2 + \xi^2}}$ for $\xi = 1$, and rely one the first integral $|\lambda_2(s)|^2 + |\lambda_1(s)|^2 = |\lambda_2(0)|^2 + |\lambda_1(0)|^2 = \mathfrak{c}^2$.
- g) Show that $\operatorname{Ad}(g_1g_2) = \operatorname{Ad}(g_1) \circ \operatorname{Ad}(g_2)$ for all $g_1, g_2 \in SE(2)$.
- h) Show that $\operatorname{CoAd}(g_1g_2) = \operatorname{CoAd}(g_1) \circ \operatorname{CoAd}(g_2)$ for all $g_1, g_2 \in SE(2)$.
- i) Show that for all $g_0 = (x_0, y_0, \theta_0) \in SE(2)$ one has:

$$\begin{aligned} \operatorname{CoAd}(g_0) \, \hat{\mathbf{p}}_0 &= (\lambda_1(0) \, \cos\theta_0 - \lambda_2(0) \, \sin\theta_0) \, \omega^1 \big|_{\mathbf{p}_0} \, + \, (\lambda_2(0) \, \cos\theta_0 + \lambda_1(0) \, \sin\theta_0) \, \omega^2 \big|_{\mathbf{p}_0} \\ &+ (\lambda_3(0) - y_0(\lambda_1(0) \cos\theta_0 - \lambda_2(0) \sin\theta_0) + x_0(\lambda_1(0) \sin\theta_0 + \lambda_2(0) \cos\theta_0)) \, \omega^3 \big|_{\mathbf{p}_0} \,, \end{aligned}$$

and show that $\omega^1|_{\mathbf{p}_0} = dx|_{\mathbf{p}_0}, \omega^2|_{\mathbf{p}_0} = dy|_{\mathbf{p}_0}, \omega^3|_{\mathbf{p}_0} = d\theta|_{\mathbf{p}_0}$ for $\mathbf{p}_0 = (0, 0, 0)$. j) Show that

$$\{\operatorname{CoAd}(g_0)\,\hat{\mathbf{p}}_0 \mid g_0 \in SE(2)\} = \{\hat{\mathbf{p}} = \sum_{i=1}^3 \lambda_i \,\omega^i \Big|_{\mathbf{p}_0} \in T^*_{\mathbf{p}_0}(SE(2)) \mid |\lambda_1|^2 + |\lambda_2|^2 = \mathfrak{c}^2\},\$$

with co-adoint orbit radius $\mathfrak{c} > 0$.

k) Show that there exists $h_0 \in SE(2)$ such that

$$\operatorname{CoAd}(h_0)\hat{\mathbf{p}}_0 = -\mathfrak{c} \left.\omega^2\right|_{\mathbf{p}_0} = -\mathfrak{c} \left.\mathrm{d}y\right|_{(0,0,0)}$$

1) Now let us store the left-invariant covector components into a modified row vector

$$\boldsymbol{\lambda} = (-\lambda_2, \lambda_1, \lambda_3)$$

Show that the vertical part of the Hamiltonian equations can be expressed as

$$\frac{d}{d\tau}\boldsymbol{\lambda}(\tau) = \boldsymbol{\lambda}(\tau)(m(\boldsymbol{\gamma}(\tau)))^{-1}\frac{d}{d\tau}m(\boldsymbol{\gamma}(\tau)) \Rightarrow$$

$$\frac{d}{ds}\boldsymbol{\lambda}(s) = \boldsymbol{\lambda}(s)(m(\boldsymbol{\gamma}(s)))^{-1}\frac{d}{ds}m(\boldsymbol{\gamma}(s))$$
(57)

with SE(2) matrix-representation:

$$m(x, y, \theta) = \begin{pmatrix} \cos \theta - \sin \theta x \\ \sin \theta & \cos \theta y \\ 0 & 0 & 1 \end{pmatrix}.$$

m) Show that (57) implies that

$$\forall_{0 \leq s \leq s_{max}(\boldsymbol{\lambda}(0))} : \boldsymbol{\lambda}(s) (m(\gamma(s)))^{-1} = \boldsymbol{\lambda}(0) (m(\gamma(0)))^{-1}.$$

n) Compute $h_0 \in SE(2)$ s.t.

 $\lambda(0)(m(h_0))^{-1} = (\mathfrak{c}, 0, 0)$

with $\mathfrak{c} = \sqrt{|\lambda_1(0)|^2 + |\lambda_2(0)|^2}$ the co-adjoint orbit radius.

hint: Why is it the same answer as your answer at 16k)?

- o) Compute $\tilde{\gamma}(\tau(s)) = EXP_{h_0}(\tau(s), -\mathfrak{c} \, \mathrm{d}y|_{(0,0,0)})$ for $s \le s_{MAX}(-\mathfrak{c} \, \mathrm{d}y|_{(0,0,0)})$.
- p) Compute all solutions $(x(s), y(s), \theta(s)) = \gamma(\tau(s)) = EXP_{\mathbf{p}_0}(\tau(s), \hat{\mathbf{p}}_0)$ for $s \leq s_{MAX}(\hat{\mathbf{p}}_0)$. hint: use (53).

Exercise 17 Set the origin as initial point, i.e. $\mathbf{p}_0 = \mathbf{0}$. 1st Maxwell-points (relative to the origin) are typically induced by reflectional symmetries (which provide new geodesics with the same length and same boundary conditions. Such reflectional symmetries can be considered in the group SE(2) or in the phase-portrait of orbits in momentum space, cf. [57].

- a) Show that $\epsilon^1(\nu_{\tau'}, c_{\tau'}, \tau') = (\nu_{\tau-\tau'}, -c_{\tau-\tau'}, \tau-\tau')$ is one reflectional symmetry of the mathematical pendulum ODE in Exercise 15.
- b) Provide all 8 reflectional symmetries ϵ_i , i = 1, ..., 8, of the mathematical pendulum ODE in Exercise 15.
- c) Provide the corresponding reflectional symmetries ε_i , $i = 1, \ldots, 8$ in the group SE(2), such that

$$Exp_{\mathbf{p}_{0}}(\epsilon^{i}(\hat{\mathbf{p}}_{0},\tau)) = \varepsilon_{i}Exp_{\mathbf{p}_{0}}(\hat{\mathbf{p}}_{0},\tau).$$

hint: $\epsilon^1(x, y, \theta) = (x \cos \theta + y \sin \theta, x \sin \theta - y \cos \theta, \theta)$ and yields a reflection of a geodesic in the median line between (x_0, y_0) and (x_τ, y_τ) .

- d) Some 1st Maxwell-points can be continuously contracted to the origin, this is called the local part \mathcal{M}^{LOC} of the 1st Maxwell-set. How do they appear in the sub-Riemannian spheres depicted in item A of Figure 11?
- e) Which reflectional symmetry is responsible for \mathcal{M}^{LOC} ?
- f) Some 1st Maxwell-points cannot be continuously contracted to the origin, this is called the global part \mathcal{M}^{GLOB} of the 1st Maxwell-set and is contained within $|\theta| = \pi$. How do they appear in the sub-Riemannian spheres depicted in item A of Figure 11?
- g) Which reflectional symmetries are responsible for \mathcal{M}^{GLOB} ?

Exercise 18 In many applications (with d = 2) one cannot distinguish between directions on a sphere, and one needs to identify antipodal points and consider $\mathbb{R}^2 \times P^1$ instead of $\mathbb{R}^2 \times S^1$, with $P^1 = S^1/_{\sim}$ with equivalence relation $\mathbf{n}_1 \sim \mathbf{n}_2 \Leftrightarrow \mathbf{n}_1 = \pm \mathbf{n}_2$. The effect of this identification on the 1st Maxwell set is quite intruiging, cf. [10]. The intersection \mathcal{M}^R of the first Maxwell set $\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \mathcal{M}_3$ with growing sub-Riemannian spheres with radius R in SE(2) is depicted in Figure 10 (taken from [10]). Now compare \mathbf{A} and \mathbf{B} in Figure 11.

Can you explain what the 2 main causes are for the highly different development of \mathcal{M}_R compared to the normal sub-Riemannian case where antipodal points are not identified?

4 Controllability Properties: Proof of Theorem 1, and Maxwell-points in $(\mathbb{M}, d_{\mathcal{F}^+_{\alpha}})$

(Global controllability) The two considered Reeds-Shepp models (or metric spaces) $(\mathbb{M}, d_{\mathcal{F}_0})$ and $(\mathbb{M}, d_{\mathcal{F}_0^+})$ are globally controllable, in the sense that the distance maps $d_{\mathcal{F}_0}$ and $d_{\mathcal{F}_0^+}$ take finite values on $\mathbb{M} \times \mathbb{M}$. This easily follows from the observation that any path $\mathbf{x} : [0, 1] \to \mathbb{R}^d$, which time derivative $\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt}$ is Lipschitz and non-vanishing, can be lifted into a path $\gamma : [0, 1] \to \mathbb{M}$ of finite length w.r.t. \mathcal{F}_0 and \mathcal{F}_0^+ , defined by $\gamma(t) := (\mathbf{x}(t), \dot{\mathbf{x}}(t)/||\dot{\mathbf{x}}(t)||)$ for all $t \in [0, 1]$. The fact that the infimum in (1) is actually a minimum for $\mathcal{F} = \mathcal{F}_0^+$ follows by Corollary 3 in App. A and (8), and the fact that the pseudo-distances take finite values.

(Local controllability) In order to show that the model $(\mathbb{M}, d_{\mathcal{F}_0^+})$ is *not* locally controllable, we need the following lemma.

Lemma 2 Let $\mathbf{n} : [0, \pi] \to \mathbb{S}^{d-1}$ be strictly 1-Lipschitz. Then $\int_0^{\pi} \mathbf{n}(0) \cdot \mathbf{n}(t) dt > 0$. Let $\mathbf{n} : \mathbb{R} \to \mathbb{S}^{d-1}$ be strictly 1-Lipschitz and 2π -periodic. Then all points $\mathbf{n}(t)$ lay in a common strict hemisphere. In particular $\mathbf{0} \notin Hull\{\mathbf{n}(t) \mid t \in [0, 2\pi]\}$.

Proof. The Lipschitzness assumption implies $\mathbf{n}(0) \cdot \mathbf{n}(t) > \cos(t)$ for all $t \in (0, \pi]$ so $\int_0^{\pi} \mathbf{n}(0) \cdot \mathbf{n}(t) dt > 0$.



Fig. 10 Top: The 1st Maxwell set \mathcal{M} in $\mathbb{R}^2 \rtimes P^1$ consists of 3 parts. Middle: the intersection \mathcal{M}^R of \mathcal{M} with the SR-spheres with growing radii R > 0 reveals itself by folds. (NB. ν denotes the multiplicity of the Maxwell-points, \overline{R} and \tilde{R} are critical radii in this respect, for details see [10]). Bottom: visualization of specific 1st Maxwell points where distinct geodesics with equal length first meet.

Let $\mathbf{n} : \mathbb{R} \to \mathbb{S}^{d-1}$ be strictly 1-Lipschitz and 2π -periodic. Set $\mathbf{M} := \int_0^{2\pi} \mathbf{n}(t) \, dt$. Then for any $t_0 \in [0, 2\pi]$ one has by the two assumptions

$$\mathbf{n}(t_0) \cdot \mathbf{M} = \int_0^{\pi} \mathbf{n}(t_0) \cdot \mathbf{n}(t_0 + t) \, \mathrm{d}t + \int_0^{\pi} \mathbf{n}(t_0) \cdot \mathbf{n}(t_0 - t) \, \mathrm{d}t > 0,$$

so for all t_0 , $\mathbf{n}(t_0) \in {\mathbf{n} \in \mathbb{S}^{d-1} | \mathbf{n} \cdot \mathbf{M} > 0}.$

Now the statements (11) and (12) on the non-local controllability of $(\mathbb{M}, d_{\mathcal{F}_0^+})$ are shown in two steps. Step 1: we show in the case of a constant cost function $\mathcal{C}_2 = \delta$ one has $\limsup_{\mathbf{p}' \to \mathbf{p}} d_{\mathcal{F}_0^+}(\mathbf{p}, \mathbf{p}') \leq 2\pi\delta$, for any $\mathbf{p} \in \mathbb{M}$. Indeed, one can design an admissible curve in $(\mathbb{M}, \mathcal{F}_0^+)$ as the concatenation of an in-place rotation, a straight line, and an in-place rotation. The length of the straight-line is $\mathcal{O}(||\mathbf{p}' - \mathbf{p}||)$ and vanishes when

 $\mathbf{p}' \to \mathbf{p}$, and the in-place rotations each have maximum $\cot \pi \delta$. Step 2: we prove the lower bound $\lim_{\mu \downarrow 0} d_{\mathcal{F}_0^+}((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mu \mathbf{n}, \mathbf{n})) \ge 2\pi\delta$, for any $(\mathbf{x}, \mathbf{n}) \in \mathbb{M}$. This and the above established upper bound implies the required result. As $C_1, C_2 \ge \delta$, we can restrict ourselves to the case of uniform $\cot C_1 = C_2 = \delta = 1$ and just show equality (12), as the estimate (11) follows by scaling with δ .

Consider a Lipschitz regular path $\gamma(t) = (\mathbf{x}(t), \mathbf{n}(t))$, with $\dot{\mathbf{x}} \propto \mathbf{n}$ and $\dot{\mathbf{x}} \cdot \mathbf{n} \ge 0$, from (\mathbf{x}, \mathbf{n}) to $(\mathbf{x} - \mu \mathbf{n}, \mathbf{n})$. Then

$$\mathbf{0} = \mu \mathbf{n} + \int_0^1 \dot{\mathbf{x}}(t) dt = \mu \mathbf{n}(0) + \int_0^1 \|\dot{\mathbf{x}}(t)\| \mathbf{n}(t) dt,$$

so $\mathbf{0} \in \operatorname{Hull}\{\mathbf{n}(t); 0 \leq t \leq 1\}$. Let $\mathbf{m} : [0,1] \to \mathbb{S}^{d-1}$ be a constant speed parametrization of \mathbf{n} . Let $\tilde{\mathbf{m}} : \mathbb{R} \to \mathbb{S}^{d-1}$ be defined by $\tilde{\mathbf{m}}(2\pi t) = \mathbf{m}(t)$ for all $t \in [0, 2\pi]$, and extended by 2π -periodicity. If $\tilde{\mathbf{m}}(\cdot)$ were strictly 1-Lipschitz then by Lemma 2 we would get $\mathbf{0} \notin \operatorname{Hull}\{\tilde{\mathbf{m}}(t) \mid t \in [0, 2\pi]\} = \operatorname{Hull}\{\mathbf{n}(t) \mid t \in [0, 1]\}$ and a contradiction. Hence there exists a $t_0 \in \mathbb{R}$ such that $\|\tilde{\mathbf{m}}(t_0)\| \geq 1$ and via the constant speed parametrization assumption we get the required coercivity:

$$1 \le \|\dot{\tilde{\mathbf{m}}}(t_0)\| = \frac{1}{2\pi} \int_0^1 \|\dot{\mathbf{n}}(t)\| \, \mathrm{d}t \Rightarrow \\ \int_0^1 \mathcal{F}_0^+(\gamma(t), \dot{\gamma}(t)) \, \mathrm{d}t \ge \int_0^1 \mathcal{C}_2(\gamma(t)) \, \|\dot{\mathbf{n}}(t)\| \, \mathrm{d}t \ge 2\pi\delta.$$

To prove local controllability of the model $(\mathbb{M}, d_{\mathcal{F}_0})$, we apply the logarithmic approximation for weighted sub-coercive operators on Lie groups, cf. [72] applied to the Lie group $SE(d) = \mathbb{R}^d \rtimes SO(d)$, in which the space of positions and orientations is placed via a Lie group quotient $SE(d)/(\{0\} \times SO(d-1))$. One obtains a sharp estimate¹⁷, where the weights of allowable (horizontal) vector fields is 1, whereas the remaining spatial vector fields orthogonal to $\mathbf{n} \cdot \nabla_{\mathbb{R}^d}$ get weight 2, as they follow by a single commutator of allowable vector fields, see e.g. [28,27]. Relaxing all spatial weights to 2 and continuity of costs $\mathcal{C}_1, \mathcal{C}_2$, yields (13). \Box

Remark 12 In view of the above one might expect that the point $(\mathbf{x} - \mu \mathbf{n}, \mathbf{n})$ is reached by a geodesic that consists of a concatenation of 1. an in-place rotation by π , 2. a straight line, 3. an in-place rotation by π . However, this is not the case as can be observed in the very lower left corner in Fig. 9, where the *two* minimizing red curves show a very different behavior. This is explained by the next lemma.

Lemma 3 Let $\mu > 0$, and $C_1 = C_2 = \delta$. Let \mathbf{R}_{θ} denote the (counter-clockwise) rotation matrix about the origin by angle θ . The endpoint $(\mathbf{x} - \mu \mathbf{n}, \mathbf{n})$ for each $\mu \ge 0$ is a Maxwell point w.r.t. (\mathbf{x}, \mathbf{n}) , since there are two minimizing geodesics in $(\mathbb{M}, d_{\mathcal{F}_{\alpha}^+})$ that are a concatenation

- 1. an in-place rotation from (\mathbf{x}, \mathbf{n}) to $(\mathbf{x}, \mathbf{R}_{\pm \frac{\pi}{2}} \mathbf{n})$,
- 2. a full U-curve, see [57], departing from and ending in a cusp from $(\mathbf{x}, \mathbf{R}_{\pm \frac{\pi}{2}}\mathbf{n})$ to $(\mathbf{x} \mu \mathbf{n}, \mathbf{R}_{\pm \frac{\pi}{2}}\mathbf{n})$,
- 3. an in-place rotation from $(\mathbf{x} \mu \mathbf{n}, \mathbf{R}_{\mp \frac{\pi}{2}} \mathbf{n})$ to
- $(\mathbf{x} \mu \mathbf{n}, \mathbf{n}).$

We have the limit $\lim_{\mu \downarrow 0} d_{\mathcal{F}_0^+}((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mu \mathbf{n}, \mathbf{n})) = 2\pi \delta.$

Proof. See [38].

Remark 13 Consider the case d = 2, $C_1 = C_2 = \delta$, and source point $\mathbf{p}_S = (\mathbf{x}, \mathbf{n}) = \mathbf{e} = (0, 0, \theta = 0)$. The end-points $(\mathbf{x} - \mu \mathbf{n}, \mathbf{n}) = (-\mu, 0, 0)$, with $\mu > 0$ sufficiently small, are 1st Maxwell-points in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ where geodesically equidistant wavefronts departing from the source point collide for the first time, see Fig. 11C. The distance mapping $d^+_{\mathcal{F}_0}(\mathbf{p}_S, \cdot)$ is not continuous, but the asymmetric distance spheres $\mathcal{S}_R := \{\mathbf{p} \in \mathbb{M} \mid d^+_{\mathcal{F}_0}(\mathbf{p}_S, \mathbf{p}) = R\}$ are connected and compact, and they collide at $R = 2\pi$ in such a way that the origin \mathbf{p}_s becomes an interior point in the asymmetric balls of radius $R > 2\pi$.

5 Cusps and Keypoints: Proof of Theorem 3

In this section we provide a proof of Theorem 3 on the occurrence of cusps and keypoints. For the uniform cost case $C_1 = C_2 = 1$ for d = 2, our curve-optimization problem (1) ($\mathbb{M}, d_{\mathcal{F}_0}$) in consideration, boils down to a standard left-invariant curve optimization in the roto-translation group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$. As we will apply tools from previous works [24, 12, 11, 66], we will make use of the following notations for

¹⁷ For specific sharp estimates for d = 3, in the context of heat-kernels estimation, see [61, ch.5.1].



Fig. 11 The development of spheres centered around $\mathbf{e} = (0, 0, 0)$ with increasing radius R. A: the normal SR spheres on \mathbb{M} given by $\{\mathbf{p} \in \mathbb{M} \mid d_{\mathcal{F}_0}(\mathbf{p}, \mathbf{e}) = R\}$ where the folds reflect the 1st Maxwell sets [9,66]. B: the SR spheres with identification of antipodal points given by $\{\mathbf{p} \in \mathbb{M} \mid \min\{ d_{\mathcal{F}_0}(\mathbf{p}, \mathbf{e}), d_{\mathcal{F}_0}(\mathbf{p} + (0, 0, \pi), \mathbf{e}) \} = R\}$ with additional folds (1st Maxwell sets) due to π -symmetry. C: the asymmetric Finsler norm pseudo spheres given by $\{\mathbf{p} \in \mathbb{M} \mid d_{\mathcal{F}_0^+}(\mathbf{p}, \mathbf{e}) = R\}$ visualized from two perspectives with extra folds (1st Maxwell sets) at the back $(-\mu, 0, 0)$. The black dots indicate points with two folds. In the case of B, this is a Maxwell-point with 4 geodesics merging. In the case of C, this is just the origin itself reached from behind at $R = 2\pi$, recall Lemma 3. Although not depicted here, if the radius $R > 2\pi$ the origin becomes an interior point of the corresponding pseudo ball.

expansion¹⁸ of velocity and momentum in the left-invariant (co)-frame:

$$\begin{cases} \mathcal{A}_{1} := \cos \theta \, \partial_{x} + \sin \theta \, \partial_{y}, \\ \mathcal{A}_{2} := -\sin \theta \, \partial_{x} + \cos \theta \, \partial_{y}, \\ \mathcal{A}_{3} := \partial_{\theta}, \end{cases}$$

$$\begin{cases} \omega^{1} := \cos \theta \, dx + \sin \theta \, dy, \\ \omega^{2} := -\sin \theta \, dx + \cos \theta \, dy, \\ \omega^{3} := d\theta, \end{cases}$$

$$\dot{\gamma}(t) = \sum_{i=1}^{3} u^{i}(t) |\mathcal{A}_{i}|_{\gamma(t)} \in T_{\gamma(t)}(\mathbb{M}),$$

$$\hat{\mathbf{p}}(t) = \sum_{i=1}^{3} \hat{p}_{i}(t) |\omega^{i}|_{\gamma(t)} \in T_{\gamma(t)}^{*}(\mathbb{M}), \end{cases}$$
(58)

where the indexing of the left-invariant frame is different here, in order to stick to the ordering (x, y, θ) applied in this article. Note that for the case $\varepsilon = 0$ admissible smooth curves γ in $(\mathbb{M}, d_{\mathcal{F}_0})$ satisfy the horizontality constraint $\dot{\gamma}(t) \in \text{Span}\{\mathcal{A}_1|_{\gamma(t)}, \mathcal{A}_3|_{\gamma(t)}\}$.

Proof of the statements regarding cusps:

- We can describe our curve optimization problem (1) using a Hamiltonian formalism, with Hamiltonian $H(\hat{\mathbf{p}}) = \frac{1}{2} (\hat{p}_1^2 + \hat{p}_3^2) = \frac{1}{2}$ [57]. By Pontryagin's Maximum Principle, geodesics adhere to the following Hamilton equations:

$$\begin{cases} \dot{p}_1 = u^1 = \hat{p}_1, \\ \dot{p}_2 = u^2 = 0, \\ \dot{p}_3 = u^3 = \hat{p}_3, \end{cases} \qquad \begin{cases} \frac{dp_1}{dt} = \hat{p}_2 \hat{p}_3, \\ \frac{dp_2}{dt} = -\hat{p}_1 \hat{p}_3, \\ \frac{dp_3}{dt} = -\hat{p}_1 \hat{p}_2. \end{cases}$$
(59)

 $^{^{18}\,}$ Note that we use upper-indices for the control's (velocity components) as they are contra-variant.

For fixed initial momentum $\hat{\mathbf{p}}(0)$, this uniquely determines a SR geodesic. Moreover, SR geodesics are contained within the (co-adjoint) orbits

$$(\hat{p}_1(t))^2 + (\hat{p}_2(t))^2 = (\hat{p}_1(0))^2 + (\hat{p}_2(0))^2.$$
(60)

The parameter t in the system (59) is SR arc length, but by reparametrizing (possible as long as u^1 does not change sign) to spatial arc length parameter s, with $\frac{ds}{dt} = \hat{p}_1$, we get a partially linear system. Combining (59) and (60), we find orbits in the (hyperbolic) phase portrait induced by

$$\begin{cases} \hat{p}'_2(s) = -\hat{p}_3 \\ \hat{p}'_3(s) = -\hat{p}_2 \end{cases} \begin{cases} \hat{p}_2(s) = \hat{p}_2(0)\cosh s - \hat{p}_3(0)\sinh s \\ \hat{p}_3(s) = -\hat{p}_2(0)\sinh s + \hat{p}_3(0)\cosh s. \end{cases}$$

Hence $|\hat{p}_3(s)| = 1$ always has a solution for some finite (possibly negative) s, except when $\hat{p}^2(0) = \hat{p}_3(0) = 0$, in which case the solutions are straight lines. Preservation of the Hamiltonian then implies $\hat{p}_1(s) = u^1(s) = \tilde{u}(s) = 0$. We conclude that every SR geodesic (with unconstrained time $t \in \mathbb{R}$) in $(\mathbb{M}, d_{\mathcal{F}_0})$ which is not a straight line admits a cusp.

- We now consider $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}}), \varepsilon > 0$. To have a cusp, we need $\hat{p}_1(t) = \hat{p}_2(t) = 0$ for some $t \in \mathbb{R}$. The co-adjoint orbit condition (60) then implies that $\hat{p}_1(t) = \hat{p}_2(t) = 0$ for all t, corresponding to a vertical geodesic that has purely angular momentum and no cusp. The same argument holds for $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$. In $(\mathbb{M}, d_{\mathcal{F}_0^+})$ we have the condition that $u^1 \geq 0$, hence by definition it can never switch sign and all geodesics are cuspless.

Proof of the statements regarding keypoints:

- For the cases $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ and $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ with $\varepsilon > 0$ we can use the same line of arguments as above. Also here both spatial controls have to vanish, resulting in vertical geodesics. The spatial projection of such curves is a single keypoint. For $(\mathbb{M}, d_{\mathcal{F}_0})$ we rely on the result that SR geodesics are analytical, and therefore if the control $u^1(t) = 0$ for some open time interval (t_0, t_1) , then $u^1(t) = 0$ for all $t \in \mathbb{R}$, again corresponding to purely angular motion.
- Geodesics in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ can have key points only at the boundaries. Suppose a geodesic $\gamma : [0, 1] \to \mathbb{M}$ in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ has an internal key point, with a corner of angle $\delta > 0$, at internal time $T_1 \in (0, 1)$. Then one can create a local short-cut with a straight-line segment connecting two sufficiently close points before and after the corner with two in-place rotations whose angles add up to δ . With a suitable mollifier this short-cut can be approximated by a curve in Γ . For details see similar arguments in [12].

Next we explain the cases A), B) and C), where we fix initial point $\gamma(0) = \mathbf{e} = (0, 0, 0)$.

- A) Suppose that the endpoint $\mathbf{p} = (x, y, \theta) \in \overline{\mathfrak{R}}$ and $x \ge 0$. Then \mathbf{p} can already be reached by a geodesic in $(\mathbb{M}, d_{\mathcal{F}_0})$ and the positivity constraint (i.e. no reverse gear), which can only increase length, becomes obsolete.
- B) Now suppose the endpoint $\mathbf{p} = (x, y, \theta)$ lays in the half-space x < 0. Then by the half-space property of geodesics in $(\mathbb{M}, d_{\mathcal{F}_0})$, cf.[24, Thm.7], the geodesic in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ must have a keypoint. By the preceding keypoints can only be located at the boundaries. If it takes place at the endpoint only, then still the constraint x < 0 is not satisfied, thereby it must take place at the origin.
- C) In those cases the endpoint **p** lays outside the connected cone of reachable angles, which are by [24, Thm.9] bounded (for those endpoints) by geodesics ending in a cusp (so not endpoints of geodesics starting at a cusp). So for those points optimal geodesics will first move by an in-place rotation (along a spherical geodesic) until it hits the cusp surface $\partial \Re$, after which it is traced back to the origin by a regular geodesic with strictly positive spatial control inside the volume \Re .

6 Eikonal equations and backtracking: Proof of Prop. 1, Corr. 1 and Thm. 4

First we shall prove Proposition 1, regarding the duals $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$, and Corollary 1, providing explicit expressions for the corresponding eikonal equations. To this end we need a basic lemma on computing dual norms on \mathbb{R}^n , where later we will set $n = 2d - 1 = \dim(\mathbb{M})$.

Lemma 4 Let $\mathbf{w} \in \mathbb{R}^n$ and let $M \in \mathbb{R}^{n \times n}$ be symmetric, positive definite. Define the norm $F_{M,\mathbf{w}} : \mathbb{R}^n \to \mathbb{R}^+$ by

$$F_{M,\mathbf{w}}(\mathbf{v}) = \sqrt{(M\mathbf{v},\mathbf{v}) + (\mathbf{w},\mathbf{v})^2_-}.$$

Then its dual norm $F_{M,\mathbf{w}}^* : (\mathbb{R}^n)^* \to \mathbb{R}^+$ equals

$$F_{M,\mathbf{w}}^*(\hat{\mathbf{v}}) = \sqrt{(\hat{\mathbf{v}}, \hat{M}\hat{\mathbf{v}}) + (\hat{\mathbf{v}}, \hat{\mathbf{w}})_+^2},\tag{61}$$

with $\hat{M} = (M + \mathbf{w} \otimes \mathbf{w})^{-1}$ and $\hat{\mathbf{w}} = \frac{M^{-1}\mathbf{w}}{\sqrt{1 + (\mathbf{w}, M^{-1}\mathbf{w})}}$.

Proof. For n = 1 the result is readily verified, and for $\mathbf{w} = \mathbf{0}$ the result is classical. We next turn to the special case M = Id, and $\mathbf{w} = (w_1, \mathbf{0}_{\mathbb{R}^{n-1}})$ is zero except maybe for its first coordinate w_1 . Thus for any $\mathbf{v} = (v_1, \mathbf{v}_2) \in \mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$ one has the splitting

$$F_{M,\mathbf{w}}(v_1,\mathbf{v}_2)^2 = \left(|v_1|^2 + (w_1v_1)_-^2\right) + \|\mathbf{v}_2\|^2$$

$$:= F_1(v_1)^2 + F_2(\mathbf{v}_2)^2.$$
 (62)

Using the compatibility of norm duality with such splittings, and the special cases n = 1 and $\mathbf{w} = 0$ mentioned above, we obtain

$$(F_{M,\mathbf{w}}^{*}(\hat{v}_{1},\hat{\mathbf{v}}_{2}))^{2} = (F_{1}^{*}(\hat{v}_{1}))^{2} + (F_{2}^{*}(\hat{\mathbf{v}}_{2}))^{2}$$
$$= \frac{|\hat{v}_{1}|^{2} + (w_{1}\hat{v}_{1})^{2}_{+}}{1 + |w_{1}|^{2}} + \|\hat{\mathbf{v}}_{2}\|^{2},$$

which is exactly of the form (61). The general case for arbitrary \mathbf{w} and symmetric positive definite M follows from affine invariance. Indeed let A be an invertible $n \times n$ matrix, and let $M' = A^{\mathrm{T}}MA$ and $\mathbf{w}' = A^{\mathrm{T}}\mathbf{w}$. Let $F = F_{M,\mathbf{w}}$ and $F' = F_{M',\mathbf{w}'}$, so that $F'(\mathbf{v}) = F(A\mathbf{v})$ for all $\mathbf{v} \in \mathbb{R}^n$. Let F^* , \hat{M} , $\hat{\mathbf{w}}$, and F'^* , \hat{M}' , $\hat{\mathbf{w}}'$, be respectively the dual norms and the matrices defined by the explicit formulas above. Then denoting $B := (A^{\mathrm{T}})^{-1}$ one has by the definition of dual norms that $F'^*(\hat{\mathbf{v}}) = F^*(B\hat{\mathbf{v}})$ for all $\hat{\mathbf{v}} \in \mathbb{R}^n$, and by the explicit formulas $\hat{M}' = B^{\mathrm{T}}\hat{M}B$, $\mathbf{w}' = B^{\mathrm{T}}\mathbf{w}$. Thus, $F^* = F^*_{M,\mathbf{w}}$ holds if and only if $F'^* = F^*_{M',\mathbf{w}'}$. Since for any M, \mathbf{w} , there exists a linear change of variables A such that $M' = \mathrm{Id}$ and \mathbf{w}' is zero except maybe for its first coordinate, the proof is complete.

Now Proposition 1 follows from Lemma 4 by writing out the dual norm, using for each $\mathbf{p} \in \mathbb{M}$:

$$M_{\mathbf{p}} = (\mathcal{C}_{1}(\mathbf{p}))^{2} (D_{\mathbf{n}}^{\varepsilon})^{-1} \oplus (\mathcal{C}_{2}(\mathbf{p}))^{2} I_{d} \quad \text{and} \mathbf{w}_{\mathbf{p}} = \begin{cases} \mathcal{C}_{1}(\mathbf{p}) \sqrt{\varepsilon^{-2} - 1} (\mathbf{n}, \mathbf{0}), & \text{for } \mathcal{F}_{\varepsilon}^{+}, \\ \mathbf{0}, & \text{for } \mathcal{F}_{\varepsilon}, \end{cases}$$
(63)

with $D_{\mathbf{n}}^{\varepsilon}$ as in (28). Corollary 1 then follows by setting the momentum covector $\hat{\mathbf{p}} = dU(\mathbf{p})$ equal to the derivative of the value function evaluated at \mathbf{p} .

Now that we have derived the eikonal equations, we obtain the backtracking Theorem 4 by Proposition 4 in App. B, which shows us that level sets of solutions of the eikonal equations are geodesically equidistant surfaces and that geodesics are found by an instrinsic gradient descent.

However, to obtain the explicit backtracking formulas we differentiate the Hamiltonian, rather than the dual metric, which is equivalent thanks to (82) (in Remark 16 in App. B). We focus below on the model $\mathcal{F}_{\varepsilon}^+$ without reverse gear, since the other case is similar. Let $\mathbf{p} \in \mathbb{M}$, let $F := \mathcal{F}_{\varepsilon}^+(\mathbf{p}, \cdot)$, and let $\hat{\mathbf{p}} = (\hat{\mathbf{x}}, \hat{\mathbf{n}}) \in T_{\mathbf{p}}^*(\mathbb{M})$. Then differentiating w.r.t. $\hat{\mathbf{n}}$ we obtain

$$\mathrm{d}_{\hat{\mathbf{n}}}F^*(\hat{\mathbf{x}},\hat{\mathbf{n}})^2 = \mathcal{C}_2(\mathbf{p})^{-2}\,\mathrm{d}_{\hat{\mathbf{n}}}\|\hat{\mathbf{n}}\|^2 = 2\,\mathcal{C}_2(\mathbf{p})^{-2}\hat{\mathbf{n}},$$

where $\|\cdot\|$ is the Riemannian metric induced by the embedding $\mathbb{S}^{d-1} \subset \mathbb{R}^d$. Differentiating w.r.t. $\hat{\mathbf{x}}$ we obtain

$$d_{\hat{\mathbf{x}}}F^{*}(\hat{\mathbf{x}},\hat{\mathbf{n}})^{2} = \mathcal{C}_{1}(\mathbf{p})^{-2} d_{\hat{\mathbf{x}}}(\hat{\mathbf{x}} \cdot D_{\mathbf{n}}^{\varepsilon}\hat{\mathbf{x}} - (1 - \varepsilon^{2})(\hat{\mathbf{x}} \cdot \mathbf{n})_{-}^{2})$$

$$= 2\mathcal{C}_{1}(\mathbf{p})^{-2} \begin{cases} D_{\mathbf{n}}^{\varepsilon}\hat{\mathbf{x}} & \text{if } \hat{\mathbf{x}} \cdot \mathbf{n} \ge 0, \\ \varepsilon^{2} \operatorname{Id} \hat{\mathbf{x}} & \text{if } \hat{\mathbf{x}} \cdot \mathbf{n} \le 0. \end{cases}$$
(64)

The announced result (32), which is equivalent to its more concise abstract form (29), follows by choosing $\hat{\mathbf{x}} := \nabla_{\mathbb{R}^d} U(\gamma(t))$ and $\hat{\mathbf{n}} := \nabla_{\mathbb{S}^{d-1}} U(\gamma(t))$ and a basic re-scaling $[0, L] \in t \mapsto t/L \in [0, 1]$.

Remark 14 The computation of the dual norms can be simplified by expressing velocity (entering the Finsler metric) and momentum (entering the dual metric) in a (left-invariant) local, orthogonal, moving frame of reference, attached to the point $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}$:

$$\dot{\mathbf{p}} = \sum_{i=1}^{2d-1} u^i \left. \mathcal{A}_i \right|_{\mathbf{p}}, \quad \hat{\mathbf{p}} = \sum_{i=1}^{2d-1} \hat{p}_i \left. \omega^i \right|_{\mathbf{p}}$$
(65)

where a moving frame of reference is chosen such that

$$\begin{cases} u^{d} = \tilde{u} = \mathbf{n} \cdot \dot{\mathbf{x}}, \\ \sum_{i=1}^{d-1} (u^{i})^{2} = \|\dot{\mathbf{x}}\|^{2} - (\mathbf{n} \cdot \dot{\mathbf{x}})^{2} \\ \sum_{i=1}^{d-1} (u^{d+i})^{2} = \|\dot{\mathbf{n}}\|^{2}, \end{cases}$$

inducing a corresponding dual frame $\{\omega^i|_{\mathbf{p}}\}$ via

$$\langle \omega^i \Big|_{\mathbf{p}}, \mathcal{A}_j \Big|_{\mathbf{p}} \rangle = \delta^i_j, \text{ for all } i, j = 1, \dots, 2d - 1.$$
 (66)

W.r.t. the left-invariant frame the matrices $D_{\mathbf{n}}^{\varepsilon}$, $M_{\mathbf{p}}$ as in (63) and $\hat{M}_{\mathbf{p}}$ all become diagonal matrices, and the dual can be computed straightforwardly. Furthermore, in this formulation we can see from the expression for the dual $(\mathcal{F}_0^+)^*$, i.e. in the limit $\varepsilon \downarrow 0$, that the positive spatial control u^d constraint results in a positive momentum \hat{p}_d constraint:

$$\left(\mathcal{F}_{0}^{+}\right)^{*}(\mathbf{p},\hat{\mathbf{p}}) = \sqrt{\frac{\left(\hat{p}_{d}\right)_{+}^{2}}{\mathcal{C}_{1}^{2}(\mathbf{p})}} + \frac{1}{\mathcal{C}_{2}^{2}(\mathbf{p})} \sum_{i=d+1}^{2d-1} (\hat{p}_{i})^{2}.$$
(67)

Therefore the eikonal equation in the positive control model $(\mathbb{M}, d_{\mathcal{F}_0^+})$ is simply given by

$$\sqrt{\frac{\|\nabla_{\mathbb{S}^{d-1}}U(\mathbf{p})\|^2}{\mathcal{C}_2^2(\mathbf{p})}} + \frac{((\mathbf{n}\cdot\nabla_{\mathbb{R}^d}U(\mathbf{p}))_+)^2}{\mathcal{C}_1^2(\mathbf{p})} = 1$$
(68)

7 Discretization of the Eikonal PDEs

7.1 Causal operators and the fast marching algorithm

The fast marching algorithm is an efficient numerical method [74] for numerically solving the static first order Hamilton-Jacobi-Bellman (or simply eikonal) PDE (5) which characterizes the distance map Uto a fixed source point \mathbf{p}_S . Fast marching is tightly connected with Dijkstra's algorithm on graphs, an in particular it shares the $\mathcal{O}(KN \ln N)$ complexity, where N = #(X) is the cardinality of the discrete domain $X \subset \mathbb{M}, X \ni \mathbf{p}_S$, and K is the average number of neighbors for each point. Both fast marching and Dijkstra's algorithms can be regarded as specialized solvers of non-linear fixed point systems of equations Au = u, where the unknown $u \in \mathbb{R}^X$ is a discrete map representing the front arrival times, which rely on the a-priori assumption that the operator $A : \mathbb{R}^X \to \mathbb{R}^X$ is causal (and monotone, but this second assumption is not discussed here). Causality informally means that the estimated front arrival time $Au(\mathbf{p})$ at a point $\mathbf{p} \in X$ depends on the given arrival times $u(\mathbf{q}), \mathbf{q} \in X$, prior to $Au(\mathbf{p})$, but not on the simultaneous or the future ones. Formally, one requires that for any $u, v \in \mathbb{R}^X, t \in \mathbb{R}$:

If
$$u^{ then $(\Lambda u)^{\le t} = (\Lambda v)^{\le t}$,
where $u^{ (69)$$$

and $v^{<t}$, $(\Lambda u)^{\leq t}$ and $(\Lambda v)^{\leq t}$ are defined similarly.

A semi-Lagrangian scheme. We implemented two discretizations of the eikonal equation (5) which benefit from the causality property. The first one is a semi-lagrangian scheme, inspired by *Bellman's optimality principle* which informally states that any sub-policy of an optimal policy is an optimal policy. Formally, let \mathcal{F} be a Finsler metric, and let $U := d_{\mathcal{F}}(\cdot, \mathbf{p}_S)$ be defined as the distance to a given source point \mathbf{p}_S . Then for any $\mathbf{p} \in \mathbb{M}$ and any neighborhood V of \mathbf{p} not containing \mathbf{p}_S one has the property

$$U(\mathbf{p}) := \min_{\mathbf{q} \in \partial V} d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) + U(\mathbf{q}).$$
(70)

In the spirit of [74,70] we discretize (70) by introducing for each interior $\mathbf{p} \in X \setminus {\mathbf{p}_{\mathrm{S}}}$ a small polygonal neighborhood $V(\mathbf{p})$, which vertices belong to the discrete point set X. The nonlinear operator Λ is defined as

$$\Lambda u(\mathbf{p}) := \min_{\substack{\{\mathbf{q}_1, \cdots, \mathbf{q}_n\}\\\text{facet of } \partial V(\mathbf{p})}} \min_{\xi \in \Xi} \mathcal{F}\left(\mathbf{p}, \sum_{i=1}^n \xi_i \mathbf{q}_i - \mathbf{p}\right) + \sum_{i=1}^n \xi_i u(\mathbf{q}_i), \tag{71}$$

where $\Xi = \{\xi \in \mathbb{R}^n_+; \sum_{i=1}^n \xi_i = 1\}$. In other words, the boundary point $\mathbf{q} \in \partial V(\mathbf{p})$ in (70) is represented in (71) by the barycentric sum $\mathbf{q} = \sum_{i=1}^n \xi_i \mathbf{q}_i$, the distance $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q})$ is approximated with the norm $\mathcal{F}(\mathbf{p}, \mathbf{q} - \mathbf{p})$, and the value $U(\mathbf{q})$ is approximated with the interpolation $\sum_{i=1}^n \xi_i u(\mathbf{q}_i)$.

We refer to [70,76] for proofs of convergence, and for the following essential property: the operator (71) obeys the causality property (69) iff the chosen stencil $V(\mathbf{p})$ obeys the following generalized acuteness property: for any \mathbf{q}, \mathbf{q}' in a common facet of $V(\mathbf{p})$, one has

$$\langle d_{\hat{\mathbf{p}}} \mathcal{F}(\mathbf{p}, \mathbf{q} - \mathbf{p}), \mathbf{q'} - \mathbf{p} \rangle \ge 0$$

For the construction of such stencils $V(\mathbf{p}), \mathbf{p} \in X$, we rely on the previous works [56,55] and on the following observation: the metrics $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$ associated to the Reeds-Shepp car models can be decomposed as

$$\mathcal{F}(\mathbf{p}, (\dot{\mathbf{x}}, \dot{\mathbf{n}}))^2 = \mathcal{F}_1(\mathbf{p}, \dot{\mathbf{x}})^2 + \mathcal{F}_2(\mathbf{p}, \dot{\mathbf{n}})^2, \tag{72}$$

which allows to build the stencils $V(\mathbf{p})$ for \mathcal{F} by combining, as discussed in [56, p. 9], some lower dimensional stencils $V_1(\mathbf{p})$ and $V_2(\mathbf{p})$ built independently for for the spatial $\mathbf{x} \in \mathbb{R}^d$ and spherical $\mathbf{n} \in \mathbb{S}^{d-1}$ variables.

We discretize S^1 uniformly, with the standard choice of stencil. We discretize S^2 by refining uniformly the faces of an icosahedron and projecting their vertices onto the sphere (as performed by the Mathematica[®] Geodesate function). The resulting triangulation only features acute interior angles, in the classical Euclidean sense, and thus provides adequate stencils since in our applications $\mathcal{F}_2(\mathbf{p}, \dot{\mathbf{n}}) = \mathcal{C}_2(\mathbf{p}) ||\dot{\mathbf{n}}||$ is proportional to the Euclidean norm, see Fig. 12. We typically use 60 discretization points for S^1 , and from 200 to 2000 points for S^2 .

We discretize \mathbb{R}^d using the cartesian grid $h\mathbb{Z}^d$, where h > 0 is the discretization scale. The norm $\mathcal{F}_{\varepsilon,1}(\mathbf{p}, \dot{\mathbf{x}}) = C_1(\mathbf{p})\sqrt{\dot{\mathbf{x}}^T(D_{\mathbf{n}}^\varepsilon)^{-1}\dot{\mathbf{x}}}$, recall the notation in (72), induced by the approximate Reeds-Shepp model $\mathcal{F}_{\varepsilon}$ on the physical variables in \mathbb{R}^d , is of Riemannian type and strongly anisotropic. In dimension $d \leq 3$, this is the adequate setting for the adaptive stencils of [56], built using discrete geometry tools known as lattice basis reduction. The norm $\mathcal{F}_{\varepsilon,1}^+(\mathbf{p},\dot{\mathbf{x}}) = C_1(\mathbf{p})\sqrt{\dot{\mathbf{x}}^T(D_{\mathbf{n}}^\varepsilon)^{-1}\dot{\mathbf{x}} + (\varepsilon^{-2} - 1)(\mathbf{n},\mathbf{x})_-^2}$ induced by $\mathcal{F}_{\varepsilon}^+$ on \mathbb{R}^d is Finslerian (i.e. non-Riemannian) and strongly anisotropic. In dimension d = 2, this is the adequate setting for the adaptive stencils of [55], built using an arithmetic object known as the Stern-Brocot tree.

Direct approximation of the Hamiltonian. A new approach, not semi-Lagrangian, had to be developed for the model $\mathcal{F}_{\varepsilon}^+$ in dimension d = 3 due to our failure to construct viable (i.e. with a reasonably small number of reasonably small vertices) stencils obeying the generalized acuteness property in this case, see Fig. 13. For manuscript size reasons, we only describe it informally, and postpone proofs of convergence for future work.

Let $\mathbf{n} \in \mathbb{S}^2$ and let $\varepsilon > 0$ be fixed. Then one can find *non-negative* weights and *integral* vectors $(\rho_i, \mathbf{w}_i) \in (\mathbb{R}_+ \times \mathbb{Z}^3)^6$, such that for all $\mathbf{v} \in \mathbb{R}^3$

$$\sum_{1 \le i \le 6} \rho_i (\mathbf{w}_i \cdot \mathbf{v})^2 = (\mathbf{n} \cdot \mathbf{v})^2 + \varepsilon^2 \|\mathbf{n} \times \mathbf{v}\|^2.$$
(73)



Fig. 12 Left: Stencil used for the metric $\mathcal{F}_{\varepsilon}$ on $\mathbb{R}^2 \times \mathbb{S}^1$, $\varepsilon = 0.1$, obeying the generalized acuteness property required for the Bellman type discretization (71). See also the control sets in Fig. 1. Center: likewise with $\mathcal{F}_{\varepsilon}^+$, $\varepsilon = 0.1$. Right: Coarse discretization of \mathbb{S}^2 with 162 vertices, used in some experiments posed on $\mathbb{R}^3 \times \mathbb{S}^2$. Some acute stencils (in the classical Euclidean sense) shown in color.



Fig. 13 Left: Slice in \mathbb{R}^3 of the control sets (7) for $\mathcal{F}_{\varepsilon}$ on $\mathbb{R}^3 \times \mathbb{S}^2$, $\varepsilon = 0.2$, for different orientations of **n**. Stencils obeying the generalized acuteness property required for Bellman type discretizations (71). Right: Slice in \mathbb{R}^3 of the control sets for $\mathcal{F}_{\varepsilon}^+$, $\varepsilon = 0.2$. Offsets used for the finite differences discretization (74), for four distinct orientations **n**.

A simple and efficient construction of $(\rho_i, \mathbf{w}_i)_{i=1}^6$, relying on the concept of obtuse superbase of a lattice, is in [41] described and used to discretize anisotropic diffusion PDEs. One may furthermore assume that $(\mathbf{n}, \mathbf{w}_i) \ge 0$ for all $1 \le i \le 6$, up to replacing \mathbf{w}_i with its opposite. Then

$$\sum_{\substack{1 \le i \le 6 \\ \mathbf{n} \cdot \nabla_{\mathbb{R}^3} U(\mathbf{p}))_+^2 \approx} (\mathbf{n} \cdot \mathbf{v})_+^2,$$

$$(\mathbf{n} \cdot \nabla_{\mathbb{R}^3} U(\mathbf{p}))_+^2 \approx$$

$$\frac{1}{h^2} \sum_{i=1}^6 \rho_i (U(\mathbf{x}, \mathbf{n}) - U(\mathbf{x} - h\mathbf{w}_i, \mathbf{n}))_+^2,$$
(74)

up to respectively an $\mathcal{O}(\varepsilon^2) \|\mathbf{v}\|^2$ and $\mathcal{O}(\varepsilon^2 + h)$ error. Following [63], we design a similar upwind discretization of the angular part of the metric

$$\|\nabla_{\mathbb{S}^2} U(\mathbf{p})\|^2 \approx (\delta_\theta U(\mathbf{p}))^2 + \frac{1}{\sin^2 \theta} (\delta_\varphi U(\mathbf{p}))^2, \tag{75}$$

where $\delta_{\theta} U(\mathbf{p})$, and likewise $\delta_{\varphi} U(\mathbf{p})$, is defined as

$$\delta_{\theta} U(\mathbf{p}) := \frac{1}{h} \max\{0, U(\mathbf{x}, \mathbf{n}) - U(\mathbf{x}, \mathbf{n}(\theta + h, \varphi)), \\ U(\mathbf{x}, \mathbf{n}) - U(\mathbf{x}, \mathbf{n}(\theta - h, \varphi))\}.$$

We denoted by $\mathbf{n}(\theta, \varphi) := (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ the parametrization of \mathbb{S}^2 by Euler angles $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi]$. Combining (74) and (75), one obtains an approximation of $\mathcal{F}_0^{+*}(\mathbf{p}, \mathrm{d}U(\mathbf{p}))^2$, within $\mathcal{O}(\varepsilon^2 + r(\varepsilon)h)$ error for smooth U, denoted $\mathfrak{F}_{\varepsilon}U(\mathbf{p})$. We denoted by $r(\varepsilon) := \max_{i=1}^{\delta} |\mathbf{w}_i|$ the norm of the largest offset appearing in (73), since these clearly depend ε . Importantly, $\mathfrak{F}_{\varepsilon}U(\mathbf{p})$ only depends on positive parts of finite differences $(U(\mathbf{p}) - U(\mathbf{q}))_+$, hence the system $\mathfrak{F}_{\varepsilon}U(\mathbf{p}) = 1$ can be solved using the fast-marching algorithm, as shown in [63]. The convergence analysis of this discretization, as the grid scale h and tolerance ε tend to zero suitably, is postponed for future work.



Fig. 14 Comparison of exact geodesics (black curves) and their numerical approximation (colored curves), with $\xi = 1/64$ and $\varepsilon = .1$, for five different end conditions $(a = ((0, 0, 60), (0, 0, 1)), b = ((6.4, 6.4, 60), (0, 0, 1)), c = ((-60, 0, 60), (-1, 0, 0)), d = ((0, 60, 60), 1/\sqrt{6}(-1, 2, 1)), e = ((60, 60, 10), (0, 0, -1)).$ The color indicates the error with the exact sub-Riemmannian geodesics [28].

Note that this approach could also be applied in dimension d = 2, and to the symmetric model $(\mathbb{M}, \mathcal{F}_{\varepsilon})$ featuring a reverse gear. We present only a single assessment of the numerical performance of our method, see Fig. 14. We compare numerically obtained shortest paths with exact SR geodesics for a small number of end points, that correspond to various types of curves. For fair end conditions (a, b, c) the numerical curves are close the exact curves. For very challenging end conditions inducing torsion (d) or extreme curvature (e) the curves are further from the exact SR geodesics. An extensive evaluation of the performance of the numerics is left for future work.

8 Applications

To show the potential of anisotropic fast marching for path-tracing in 2D and 3D (medical) images we performed experiments on each of the datasets in Fig. 5:

- a 2D toy example using a map of Centre Pompidou,
- a 2D retinal image,
- two synthetic Diffusion-weighted Magnetic Resonance Imaging (dMRI) datasets, with different bundle configurations.

We use the 2D datasets to point out the difference in results when using the metric $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$, and to explain the role of the keypoints when using $\mathcal{F}_{\varepsilon}^+$, that occur instead of (possibly unwanted) cusps.

On the synthetic dMRI datasets we present the first application of our methods to this type of data. We present how a cost function can be extracted from the data, and how this leads to correct tracking of bundles, similar to the 2D case. The benefits of anisotropic metrics compared to isotropic metrics are demonstrated by performing backtracking for various model parameter variations.

The experiments were performed using an anisotropic FM implementation written in C++, for d = 2 described in [56]. Implementation details for d = 3 will be described in future work. Mathematica 11.0 (Wolfram Research, Inc., Champaign, IL) was used for further data analysis, applying Wolfram LibraryLink (Wolfram Research, Inc., Champaign, IL) to interface with the FM library.



Fig. 15 Comparison between the shortest paths from end points (black) to one of the exits (green) in a model map of Centre Pompidou, for cars with (left, blue lines) and without (right, red lines) reverse gear. The yellow arrows indicate the orientation of the curve. The background colors show the distances at each position, minimized over the orientation. White points left indicate the cusps, white points right indicate the (automatically placed) key points where in-place rotations take place.



Fig. 16 Left: SR geodesics (in blue) in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ with given boundary conditions (both forward and backward). Right: SR geodesics (in red) in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ with the same boundary conditions. We recognize one end-condition case where on the left we get a cusp, whereas on the right we have a key-point (with in-place rotation) precisely at the bifurcation.

8.1 Applications in 2D

8.1.1 Shortest Path to the Exit in Centre Pompidou

To illustrate the difference between the models with and without reverse gear and to show the role of the keypoints for non-uniform cost, we use a map of Centre Pompidou as a 2D image, see Fig. 15. The walls (in black) have infinite cost, everywhere else the cost is 1. We place end points (black dots) in various places of the museum and look for the shortest path from those points to one of the two exits, regardless of the end orientation. Since there are now two exits, say at \mathbf{p}_0 and \mathbf{p}_1 , the distance $U(\mathbf{p})$ of any point $\mathbf{p} \in \mathbb{M}$ to one of the exits is given by

$$U_{\mathcal{F}}(\mathbf{p}) = \min\{d_{\mathcal{F}}(\mathbf{p}_0, \mathbf{p}), d_{\mathcal{F}}(\mathbf{p}_1, \mathbf{p})\}.$$
(76)

We use a resolution of $N_x \times N_y \times N_o = 706 \times 441 \times 60$. The cost in this example is only dependent on position, but constant in the orientation. Moreover, we use $C_1 = C_2$ and $\varepsilon = 1/10$.

On the left of Fig. 15 we see optimal paths (in blue) obtained using the metric $\mathcal{F} = \mathcal{F}_{\varepsilon}$. The fast marching algorithm successfully connects all end points to one of the exits. Some of the geodesics have cusps, indicated with white points, resulting in backward motion on (a part of) the curve. The colors show the distance $U_{\mathcal{F}_{\varepsilon}}$ as above, at each position minimized over the orientations.

On the right, the optimal paths using the asymmetric metric $\mathcal{F} = \mathcal{F}_{\varepsilon}^+$ are shown in red. The curves no longer exhibit cusps, but have in-place rotations (white dots) instead. These keypoints occur in this example on corners of walls. (Due to the fact that ε is small but nonzero, there can still be small sideways motion.) The shortest paths for this model are successions of sub-Riemannian geodesics and of in place rotations, which can be regarded as reinitializations of the former: the orientation is adapted until an orientation is found from which the path can continue in an optimal sub-Riemannian way. We stress that the fast marching algorithm has no special treatment for keypoints, which are only detected in a postprocessing step. We observe that keypoints are automatically positioned at positions where it makes sense to have an in-place rotation. Small differences in the distance maps between $U_{\mathcal{F}_{\varepsilon}}$ left and $U_{\mathcal{F}_{\varepsilon}^+}$ right can be observed: the constrained model usually has a slightly higher cost right around corners.

8.1.2 Vessel Tracking in Retinal Images

Another application is vessel tracking in retinal images, for which the model with reverse gear and the fast-marching algorithm have shown to be useful in [9,67]. Although the algorithm works fast and led to successful vessel segmentation in many cases, in some cases, in particular bifurcations of vessels, cusps occur. Fig. 16 shows one such example on the left. The image has resolution $N_x \times N_y \times N_o = 121 \times 114 \times 64$. The cost is constructed as in [9]: the image is first lifted using cake wavelets [25], resulting in an image on $\mathbb{R}^2 \times S^1$. For the lifting and for the computation of the cost function from the lifted image, we rely on their parameter settings. We use $C_1 = \xi C_2$, with $\xi = 0.02$ (top) and $\xi = 0.04$, and $\varepsilon = 0.1$. The orientations of the forward and the backward case. The vessel with end condition C is particularly challenging, since it comes across a bifurcation. For the tracking of this vessel, we indicated the orientation with yellow arrows.

The unconstrained model $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$, corresponding to the blue tracks on the left half of Fig. 16, gives a correct vessel tracking for the forward end conditions of A and B, for both values of ξ . This is obviously the better choice than the backward cases. However, for end condition C, neither the forward or backward with neither values of ξ gives a vessel tracking without cusps. On the other hand, if we use the constrained model $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$, we obtain an in-place rotation or keypoint in the neighborhood of the bifurcation. Typically a higher value of ξ brings these points closer to the bifurcation. Taking the backward end conditions in combination with this model, we see in some cases that end locations are first passed by the vessel tracking algorithm, until it reaches a point where in-place rotation is cheaper, and then returns to the end position.

Exercise-Mathematica 3. Perform retinal vessel tracking using the notebook "SE2_RetinaTracking.nb" (downloadable at www.LieAnalysis.nl > Education > Lecture 5-Tracking > Retinal Vessel Tracking). Perform the tasks at the end of the notebook:

- Experiment with parameters (ξ : flexibility of the curve, λ : influence of the cost, p: contrast of the cost).
- Experiment with different geodesic models ("ReedsSheppForward2", "ReedsShepp2", "ReedsShepp2"projective).
- Run the experiment with different end conditions (try to find curves with cusps).

8.2 Application to Diffusion-Weighted MRI Data

DW-MRI is a magnetic resonance technique for non-invasive measurement of water diffusion in fibrous tissues [58]. In the brain, diffusion is less constrained parallel to white matter fibers (or axons) than perpendicular to them, allowing us to infer the paths of these fibers. The diffusion measurements are distributions $(\mathbf{y}, \mathbf{n}) \mapsto U(\mathbf{y}, \mathbf{n})$ within the manifold \mathbb{M} for d = 3. From these measurements a fiber orientation distribution (FOD) can be created, yielding a probability of finding a fiber at a certain position and orientation [75].

Backtracking is performed through forward Euler integration of the backtracking PDE involving the intrinsic gradient, following Theorem 4 and Eq. (29) and Eq. (32). The spatial derivative was implemented as a first-order Gaussian derivative. The angular derivatives are implemented by a first-order spherical



Fig. 17 Comparison of the results of backtracking on a 2D plane in a synthetic dMRI dataset on $\mathbb{M} = \mathbb{R}^3 \times \mathbb{S}^2$. In case A the default parameters for σ , ξ and ε are applied resulting in a global minimizing geodesic (left) and its corresponding distance map (right). Case B reflects the influence of the data-term σ . Case C reflects the isotropic Riemannian case. Case D reflects a high cost for moving spatially and results in curves that resemble a piecewise linear curve. The distance map is illustrated using a glyph visualization in which the size of the glyph corresponds to $exp(-d_{\mathcal{F}_{\varepsilon}}(\mathbf{p}_s, \mathbf{p}_e)/s)^p$ where \mathbf{p}_s is the seed location, \mathbf{p}_e is a location on a glyph, and s and p are chosen based on visualization clarity.

harmonic derivative. The latter has the key advantage that in a spherical harmonic basis exact analytic computations can be done. Here, one must rely on two-fold recursions in [40, Lemma 2 & 4], so that the poles due to a standard Euler angle parametrization of S^2 do not appear in exact recursions of Legendre polynomials!

If data-driven factors C_1 and C_2 come in a spherical sampling or if one wants to work in a spherical sampling (e.g. higher order tesselation of the icosahedron) in a fast-marching method, than one can easily perform the pseudo-inverse of the discrete inverse spherical harmonic transform, where one typically keeps the number of spherical harmonics very close to the number of spherical sampling points, so that maximum accuracy order is maintained for computing angular derivatives in the intrinsic gradient descent of Theorem 4.

8.2.1 Construction of the Cost Function

The synthetic dMRI data is created by generating/simulating a Fiber Orientation Density (FOD) of a desired structure. There are sophisticated methods for this, e.g. [19,13], but evaluation on phantom data constructed with these tools is left for future work. Here we use a basic but practical method on two simple configurations of bundles in \mathbb{R}^3 , the ones on the bottom row in Fig. 5. In each voxel inside a bundle, we place a spherical δ -distribution, with the peak in the orientation of the bundle. We convolve each δ -distribution with an FOD kernel that was extracted from real dMRI data and is related to the dMRI signal measured in a voxel with just a single orientation of fibers. Spherical rotation of the FOD kernel is done in the spherical harmonics domain by use of the Wigner D-matrix to prevent interpolation issues. We compose from all distributions an FOD function $W : \mathbb{M} \to \mathbb{R}^+$. This function evaluates to high values in positions/orientations that are inside and aligned with the bundle structure.

We use the FOD W to define the cost function $\frac{1}{1+\sigma} \leq C \leq 1$ via

$$\mathcal{C}(\mathbf{p}) = \frac{1}{1 + \sigma \left| \frac{W_{+}(\mathbf{p})}{\|W_{+}\|_{\infty}} \right|^{p}}$$

where $\sigma \geq 0$, $p \in \mathbb{N}$, with $\|\cdot\|_{\infty}$ the sup-norm and $W_+(\mathbf{p}) = \max\{0, W(\mathbf{p})\}$. The cost function \mathcal{C} induces the following spatial and angular cost functions $(\mathcal{C}_1, \mathcal{C}_2)$:

$$\mathcal{C}_1(\mathbf{p}) = \xi \mathcal{C}(\mathbf{p}), \qquad \mathcal{C}_2(\mathbf{p}) = \mathcal{C}(\mathbf{p}).$$

The implementation of nonuniform cost is comparable to the application of vessel tracking in retinal images in d = 2 by Bekkers et al. [9].

8.2.2 Influence of model parameters

The first synthetic dataset consists of a curved and a straight vessel, which cross at two locations as shown in Fig. 17. The experiments using metric $\mathcal{F}_{\varepsilon}$ demonstrate the effect of the model parameters on the geodesic back-traced from the bottom-left to the seed location at the bottom-right of the curved vessel. A distance map is computed for parameter configuration A (Fig. 17, right) in which suitable values are used for the data-term σ , and the fast-marching parameters ξ and ε . Furthermore, fixed values are used for data sharpening p = 3, spatial smoothing $\sigma s = 0.5$, forward-Euler integration step size $\delta t = 0.04$, and a gridscale of 1. By use of these parameters the global minimizing geodesic (Fig. 17.A, left) is shown to take the longer, curved route. In parameter configuration B the data-term σ is lowered, which creates a geodesic that is primarily steered by internal curve-dependent costs and is shown to take the shortcut route (Fig. 17.B). Setting $\varepsilon = 1$ in configuration C leads to a Riemannian case where the geodesic resembles a piecewise linear curve. In configuration D the relative cost of spatial movement relative to angular movement is high, leading to geodesics with shortcuts.

We conclude that configuration **A** with a relatively strong data term, large bending stiffness ($\xi^{-1} = 10$), and a nearly SR geometry ($\varepsilon = 0.1$) avoids unwanted shortcuts.

8.2.3 Positive control constraint

For the application of FM in dMRI data it is desirable that the resulting geodesic is not overly sensitive to the boundary conditions, i.e. the placement and orientation of the geodesic tip. Furthermore, since neural fibers do not form cusps, these are undesirable in the backtracking results. In Fig. 18 the backtracking results are shown for the cases without reverse gear $\mathcal{F}_{\varepsilon}^+$ (top) and the model with reverse gear $\mathcal{F}_{\varepsilon}$ (bottom). The distance map for $\mathcal{F}_{\varepsilon}^+$ was computed by the iterative method implementing the forward Reeds-Shepp car, while for $\mathcal{F}_{\varepsilon}$ the FM method was used.

We conclude that without the positive control constraint, small changes in tip orientation cause large variations in the traced geodesic in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$, whereas the traced geodesic in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ is both more stable and more reasonable.

8.2.4 Robustness to neighboring structures

A pitfall of methods that provide globally optimal curves using a dataterm is that dominant structures in the data attract many of the curves, much like the highway usually has the preference for cars rather than local roads. This phenomenon is to a certain extent unwanted in our applications, and we illustrate with the following example that it can be circumvented using a sub-Riemannian instead of Riemannian metric. We use the dataset as introduced in Fig. 5. It consists of one bundle that has torsion (green), that crosses with another bundle (blue), and a third bundle (red) that is parallel with the first in one part. The cost in these bundles is constructed in the same way as above, but now the cost in the red bundle is twice as low as in the other bundles. A small part of the data is visualized on the left of Fig. 19. This data is used to construct the cost function as explained above.

The resolution of the data is $N_x \times N_y \times N_z \times N_o = 32 \times 32 \times 32 \times 162$. Again we use $C_1 = \xi C_2 = C$, with $\xi = 0.1$. To have comparable parameters as in the previous experiment, despite increasing the amplitude in one of the bundles by a factor 2, we choose to construct the cost using parameter p = 3, and $\sigma = 3 \cdot 2^p = 24$.



Fig. 18 Backtracking of optimal geodesics of the model $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ without reverse gear (top) and the model with reverse gear $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ (bottom) using the model parameters of configuration A ($\sigma = 3.0, \xi = 0.1$ and $\varepsilon = 0.1$) for various end conditions.



Fig. 19 Left: 3D configuration of bundles and a visualization of part of the synthetic dMRI data. Middle: backtracking of geodesics in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ from several points inside the curves to end points of the bundle is successful when using $\varepsilon = 0.1$. Right: when using $\varepsilon = 1$, the dominant red bundle can cause the paths from the green bundle to deviate from the correct structure.

From various positions inside the green, blue and red bundle, the shortest paths to the end of the bundles computed by the FM algorithm nicely follow the shape of the actual bundles, when we choose $\varepsilon = .1$ small, corresponding to an almost SR geodesic. This is precisely what prevents the geodesic in the green bundle to drift into the (much cheaper) red bundle. We show on the right in Fig. 19 that choosing $\varepsilon = 1$, corresponding to having an isotropic Riemannian metric, this unwanted behavior can easily occur.

We conclude that the SR geodesics in $(\mathbb{M} = \mathbb{R}^3 \times S^2, d_{\mathcal{F}_{\varepsilon}})$ with $\varepsilon \ll 1$, are less attracted to parallel, dominant structures than isotropic Riemannian geodesics.

Exercise-Mathematica 4. Perform geodesic tracking on synthetic diffusion MRI volumes using the notebook "SRGeodesics_3Dexample.nb"

(downloadable at www.LieAnalysis.nl > Education > Mathematica Exercises > Part III, where within the zip file you can open

 $Part_III_Exercise_4_Bundle_Tracking_SE(3).nb$).

Try to reconstruct the 3 different bundles in the volume with different settings and tracking models.

9 Conclusion and Discussion

We have extended the existing methodology for modelling and solving the problem of finding optimal paths for a Reeds-Shepp car to 3D and to a case without reverse gear. We have shown that the use of the constrained model leads to more meaningful shortest paths in some cases and that the extension to 3D has opened up the possibility for tractography in dMRI data.

Instead of using a hard constraint on the curvature as in the original paper by Reeds and Shepp [62], we used Riemannian/Finslerian metrics. We have introduced these metrics, \mathcal{F}_0 and \mathcal{F}_0^+ , for d = 2, 3, such that they allow for curves that have a spatial displacement proportional to the orientation, with a positive proportionality constant in the case of \mathcal{F}_0^+ .

We have captured theoretically some of the nature of the distance maps and geodesics following from the new constrained model. We have shown in Thm. 1 that both models are globally controllable, but only the unconstrained model is also locally controllable.

The sub-Riemannian and sub-Finslerian nature is difficult to capture numerically. To this end, we introduced approximating metrics $\mathcal{F}_{\varepsilon}$ and $\mathcal{F}_{\varepsilon}^+$, that do allow for numerical approaches. We have shown in Thm. 2 that as $\varepsilon \to 0$, the distance map converges pointwise and the geodesics converge uniformly, implying that for sufficiently small ε we indeed have a reasonable approximation of the $\varepsilon = 0$ case.

We have analyzed cusps in $(\mathbb{M}, d_{\mathcal{F}_0})$ and keypoints in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ which occur on the interface surface $\partial \mathbb{M}_{\pm}$ given by (31). The analysis, for uniform costs, is summarized in Thm. 3. We have shown that cusps are absent in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ for $\varepsilon > 0$, that keypoints in $(\mathbb{M}, d_{\mathcal{F}_0^+})$ occur only on the boundary, and we provided analysis on how this happens. In Thm. 4 we have shown how optimal geodesics in $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ and $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ can be obtained from the distance maps with an intrinsic gradient descent method.

To obtain solutions for the distance maps and optimal paths, we used a Fast-Marching method. By formulating an equivalent problem to the minimization problem for optimal paths in the form of an eikonal equation, the FM method can be used using specific discretization schemes. We briefly compared the numerical solutions using $\mathcal{F}_{\varepsilon}$ with $\varepsilon \ll 1$ with the exact sub-Riemannian geodesics in SE(2) with uniform cost, which showed sufficient accuracy for not too extreme begin and end conditions.

To show the use of our method in image analysis, we have tested it on two 2D problems and two 3D problems. All four experiments confirm that the combination of the eikonal PDE formulation, the Fast-Marching method and the construction of the non-uniform cost from the images, results in geodesics that follow the desired paths. From the experiment on an image of Centre Pompidou, with constant, finite cost everywhere except for the walls, it followed that instead of having cusps when using the $\mathcal{F}_{\varepsilon}$ metric, we get keypoints (in-place rotations) when using $\mathcal{F}_{\varepsilon}^+$. These keypoints turn out to be located on logical places in the image. On the 2D retinal image we showed that the metric $\mathcal{F}_{\varepsilon}^+$ gives a new tool for tackling vessel tracking through bifurcations. We see that keypoints appear close to the bifurcation, leading to paths that more correctly follow the data.

The basic experiments on 3D show advantages of the model $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ with $0 < \varepsilon \ll 1$ over the model $(\mathbb{M}, d_{\mathcal{F}_1})$ in the sense that the optimal geodesics better follow the curvilinear structure and deal with crossings and nearby parallel bundles (even if torsion is present). Furthermore, we have shown the advantage of model $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}^+})$ with $0 < \varepsilon \ll 1$, compared to $(\mathbb{M}, d_{\mathcal{F}_{\varepsilon}})$ in terms of stability, with keypoints instead of cusps.

The strong performance of the Reeds-Shepp car model in 2D vessel tracking and positive first results on artificial dMRI data, encourages us to pursue a more quantitative assessment of the performance in both 3D vessel tracking problems and in actual dMRI data. Such 3D vessel tracking problems are encountered in for example Magnetic Resonance Angiography. In future work we will elaborate on the implementation and evaluation of the fast-marching and the iterative PDE implementation of [38, App. B]. Furthermore, we aim to integrate locally adaptive frames [37] into the Finsler functions $\mathcal{F}_{\varepsilon}$, $\mathcal{F}_{\varepsilon}^+$, for a more adaptive vessel/fiber tracking.

10 Acknowledgements

This section of the lecture notes is primarily based on the following scientific works [9,26,24,27,28,56,57, 66,67], including important contributions by Jean-Marie Mirebeau, Yuri Sachkov, Ugo Boscain, Francesco Rossi, Igor Moiseev, Alexey Mashtakov, Stephan Meesters, Gonzalo Sanguinetti building on many other pioneering works, e.g. by Giovanna Citti, Alessandro Sarti, Jean Petitot and many others.

We thank Wouter Scharpach MsC student at TU/e for feedback on the document. The research leading to results presented in this lecture notes has received funding from the European Research Council under the European Community's 7th Framework Programme (FP7/20072014)/ERC grant agreement No. 335555 (Lie Analysis).

A Well-posedness and convergence of the Reeds-Shepp models

We introduce in A.1 some general elements of control theory, which are specialized in A.2 to the Reeds-Shepp models and their approximations.

A.1 Closedness of controllable paths

In this section, we introduce the notion of an admissible path γ with respect to some controls \mathfrak{B} . We state in Theorem 5 a closedness result, slightly generalizing the one from [14], from which we deduce in Corollaries 3 and 4 an existence and a convergence result for a minimum time optimal control problem. The first ingredient of this approach is the notion of Hausdorff distance on a metric space.

Definition 6 Given a metric space \mathbb{E} , we let $\mathcal{K}(\mathbb{E})$ be the collection of non-empty compact subsets of \mathbb{E} . The distance function $d_A : \mathbb{E} \to \mathbb{R}_+$ and the Hausdorff distance $\mathcal{H}(A, B)$, where $A, B \in \mathcal{K}(\mathbb{E})$, are defined respectively by

$$d_A(x) := \inf_{y \in A} d(x, y), \ \mathcal{H}(A, B) := \max\{\sup_{x \in B} d_A(x), \sup_{x \in A} d_B(x)\}.$$

In the following, we fix a closed set \mathbb{X} , contained in an Euclidean vector space \mathbb{E} , or in a complete Riemannian manifold \mathbb{M} . In the applications considered in this paper, \mathbb{X} is of the form $\mathbb{X}_0 \times \mathbb{S}^{d-1}$, where $\mathbb{X}_0 \subset \mathbb{R}^d$ is some image domain, see Fig. 16, or the set of accessible points in a map (which excludes the walls), see Fig. 15. The embedding space can be the vector space $\mathbb{E} = \mathbb{R}^d \times \mathbb{R}^d$, which is an acceptable but rather extrinsic point of view, or the Riemannian manifold $\mathbb{M} = \mathbb{R}^d \times \mathbb{S}^{d-1}$, equipped with the metric $\mathcal{G}_{\varepsilon}$ for some arbitrary but fixed $\varepsilon > 0$, see (26).

We equip the collection of all Lipschitz paths $\Gamma := \text{Lip}([0,1], \mathbb{X})$ with the topology of uniform convergence. We will make use of Ascoli's lemma [4,3], which states that any uniformly bounded and equicontinuous sequence of paths admits a converging sub-sequence. In our case the paths are Lipschitz with a common Lipschitz constant.

Definition 7 Given a normed vector space V, we denote by $\mathfrak{C}(V) \subset \mathcal{K}(V)$ the collection of non-empty compact subsets of V, which are convex and contained in the unit ball.

Remark 15 The restriction to convex subsets is essential. For a uniformly converging sequence of Lipschitz functions $\gamma_n : [0,1] \to \mathbb{M}$ with limit γ_* , with $\dot{\gamma}_n(t) \in K$ for a.e. $t \in [0,1]$ and K a compact set, we can deduce that $\dot{\gamma}_* \in \text{Hull}(K)$, for a.e. $t \in [0,1]$. The convexity then guarantees that $\dot{\gamma}_* \in K = \text{Hull}(K)$.

Definition 8 A family of controls \mathcal{B} on the set \mathbb{X} is an element of the set \mathfrak{B} defined by

- If $\mathbb{X} \subset \mathbb{E}$ an Euclidean vector space, then $\mathfrak{B} := C^0(\mathbb{X}, \mathfrak{C}(\mathbb{E})).$

 $- \text{ If } \mathbb{X} \subset \mathbb{M} \text{ a Riemannian manifold, then } \mathfrak{B} := \{ \mathcal{B} \in C^0(\mathbb{X}, \mathcal{K}(T\mathbb{M})) \, \big| \, \forall \mathbf{p} \in \mathbb{X}, \, \mathcal{B}(\mathbf{p}) \in \mathfrak{C}(T_\mathbf{p}\mathbb{M}) \}.$

In both cases, ${\mathfrak B}$ is equipped with the topology of locally uniform convergence.

Definition 9 A path γ is \mathcal{TB} -admissible, where $\gamma \in \Gamma$, $T \in \mathbb{R}_+$ and $\mathcal{B} \in \mathfrak{B}$, iff for almost every $t \in [0, 1]$

 $\dot{\gamma}(t) \in T\mathcal{B}(\gamma(t)).$

We denoted $TB := \{T\mathbf{v} \mid \mathbf{v} \in B\}$, where $T \in \mathbb{R}_+$ and B is a subset of a vector space. Note the potential conflict of notation with the tangent space $T\mathbb{M}$ to the embedding manifold \mathbb{M} , which should be clear from context. If a path γ is $T\mathcal{B}$ -admissible for some controls $\mathcal{B} \in \mathfrak{B}$, then it must be T-Lipschitz. The following result slightly extends, for our convenience, Corollary A.5 in [14].

Theorem 5 The set $\{(\gamma, T, \mathcal{B}) \in \Gamma \times \mathbb{R}_+ \times \mathfrak{B} \mid \gamma \text{ is } T\mathcal{B}\text{-admissible}\}$ is closed.

Proof. Let $(\gamma_n, T_n, \mathcal{B}_n)$ be sequences of paths, times and controls converging to $(\gamma_{\infty}, T_{\infty}, \mathcal{B}_{\infty})$, and such that γ_n is $T_n \mathcal{B}_n$ -admissible for all $n \geq 0$. Since the paths γ_n are converging as $n \to \infty$, they lay in a common compact subset \mathbb{X}' of the closed domain \mathbb{X} , recall Remark 15. As a result, the restricted controls $\mathcal{B}'_n := (\mathcal{B}_n|_{\mathbb{X}'})$ are uniformly converging as $n \to \infty$. In the case where $\mathbb{X} \subset \mathbb{E}$ a Euclidean space, applying Corollary A.5 in [14] to the sequence $(\gamma_n, T_n \mathcal{B}'_n)$ we obtain that γ_∞ is $T_\infty \mathcal{B}_\infty$ -admissible as announced.

In the case where $\mathbb{X} \subset \mathbb{M}$ a Riemannian manifold, an additional proof ingredient is required. Let \mathbb{M}' be an open neighborhood of \mathbb{X}' with compact closure in \mathbb{M} , and let $\mathcal{I} : \mathbb{M}' \to \mathbb{E}$ be an embedding (i.e. an injective immersion) with bounded distortion of the manifold \mathbb{M}' into a Euclidean space \mathbb{E} of sufficiently high dimension, which by Whitney's embedding theorem is known to exist. Define the set $\mathbb{X}'' := \mathcal{I}(\mathbb{X}')$, the paths $\gamma''_n := \mathcal{I} \circ \gamma_n$, and controls $\mathcal{B}''_n(\mathcal{I}(\mathbf{p})) :=$ $d\mathcal{I}(\mathbf{p}, \mathcal{B}_n(\mathbf{p}))$ for all $\mathbf{p} \in \mathbb{X}'$ and $n \in \mathbb{N} \cup \{\infty\}$. Applying again Corollary A.5 in [14] we obtain that γ''_{∞} is $T_{\infty} \mathcal{B}''_{\infty}$ admissible, hence that γ_{∞} is $T_{\infty} \mathcal{B}_{\infty}$ -admissible as announced. In line with the 3rd identity in (8), we rely on the following definition where we rescale the time interval to [0, 1].

Definition 10 For any $\mathcal{B} \in \mathfrak{B}$, $\mathbf{p}, \mathbf{q} \in \mathbb{X}$, we let

$$T_{\mathcal{B}}(\mathbf{p}, \mathbf{q}) := \inf\{T \ge 0 \mid \exists \gamma \in \Gamma, \gamma(0) = \mathbf{p}, \gamma(1) = \mathbf{q}, \\ \text{and } \gamma \text{ is } T\mathcal{B}\text{-admissible}\}.$$
(77)

Corollary 3 If $\mathcal{B} \in \mathfrak{B}$, $\mathbf{p}, \mathbf{q} \in \mathbb{X}$ are such that $T_{\mathcal{B}}(\mathbf{p}, \mathbf{q}) < \infty$, then the inf. (77) is attained.

Proof. Let $T := T_{\mathcal{B}}(\mathbf{p}, \mathbf{q})$, and for each $0 < \varepsilon \leq 1$ let γ_{ε} be a $(T + \varepsilon)\mathcal{B}$ -admissible path from \mathbf{p} to \mathbf{q} , which is thus (T + 1)-Lipschitz. By Arzela-Ascoli's lemma [3,4] there exists a converging sequence of paths $\gamma_{\varepsilon_n} \to \gamma_0$ as $n \to \infty$. The limit path γ_0 is $T\mathcal{B}$ -admissible by Theorem 5, and the result follows.

Corollary 4 For all $\varepsilon \in [0,1]$ let $\mathcal{B}_{\varepsilon} \in \mathfrak{B}$. Assume that $\mathcal{B}_{\varepsilon} \to \mathcal{B}_0$ as $\varepsilon \to 0$, and that $\mathcal{B}_{\varepsilon}(\mathbf{p}) \supset \mathcal{B}_0(\mathbf{p})$ for all $\varepsilon \ge 0$, $\mathbf{p} \in \mathbb{X}$. Then

$$T_{\mathcal{B}_{\varepsilon}}(\mathbf{p},\mathbf{q}) \to T_{\mathcal{B}_0}(\mathbf{p},\mathbf{q}), \quad as \ \varepsilon \to 0.$$

Let $T_{\varepsilon} := T_{\mathcal{B}_{\varepsilon}}(\mathbf{p}, \mathbf{q})$ for each $\varepsilon \geq 0$. Assume in addition that there exists a unique $T_0\mathcal{B}_0$ -admissible path γ_0 from \mathbf{p} to \mathbf{q} , and for each $\varepsilon > 0$ denote by γ_{ε} an arbitrary path from \mathbf{p} to \mathbf{q} which is $(\varepsilon + T_{\varepsilon})\mathcal{B}_{\varepsilon}$ admissible. Then $\gamma_{\varepsilon} \to \gamma_0$ as $\varepsilon \to 0$.

Proof. The inclusion $\mathcal{B}_{\varepsilon}(\mathbf{p}) \subset \mathcal{B}_{0}(\mathbf{p}), \forall \mathbf{p} \in \mathbb{M}$, implies the inequality $T_{\varepsilon} \leq T_{0}$, for all $\varepsilon \geq 0$. Denoting $T_{*} := \limsup T_{\varepsilon}$ as $\varepsilon \to 0$, we thus observe that $T_{*} \leq T_{0}$. For the reverse inequality $T_{*} \geq T_{0}$, we apply Arzela-Ascoli lemma to the family of paths $(\gamma_{\varepsilon})_{0<\varepsilon\leq 1}$ which are $(T_{0}+1)$ -Liptschitz by construction, and obtain a converging subsequence of paths $\gamma_{\varepsilon_{n}} \to \gamma_{*}$. Theorem 5 implies the admissibility of γ_{*} with respect to the controls $T_{*}\mathcal{B}_{0}$. Thus $T_{*} \geq T_{0}$ but since $T_{*} \leq T_{0}$, we must have $T_{*} = T_{0}$, and $\gamma_{*} = \gamma_{0}$ by the uniqueness assumption. The result follows.

More generally, if the infimum (77) is realized by a family $(\gamma_i)_{i \in I}$ of paths, then for any sequence $\varepsilon_n \to 0$ one can find a subsequence such that $\gamma_{\varepsilon_{\varphi(n)}} \to \gamma_i$ as $n \to \infty$ for some $i \in I$.

A.2 Specialization to the Reeds-Shepp models

We begin this section by recalling, and slightly generalizing, the notion of Finsler function introduced in §2.2. We then prove that the Reeds-Shepp metrics \mathcal{F}_0 and \mathcal{F}_0^+ are indeed Finsler functions in this sense.

Definition 11 A metric on a complete Riemannian manifold \mathbb{M} is a map $\mathcal{F} : T\mathbb{M} \to [0, +\infty]$. With respect to the second variable, it must be 1-homogeneous, convex, and bounded below by $\delta \| \cdot \|$, where δ is a positive constant. In terms of regularity, the sets $\mathcal{B}_{\mathcal{F}}(\mathbf{p}) := \{\dot{\mathbf{p}} \in T_{\mathbf{p}}\mathbb{M} \mid \mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}) \leq 1\}$ must be closed and depend continuously on $\mathbf{p} \in \mathbb{M}$ with respect to the Hausdorff distance on $T\mathbb{M}$.

The next proposition is due to (8).

Proposition 2 With the notations of Definition 11, the sets $\mathbf{p} \in \mathbb{M} \mapsto \mathcal{B}_{\mathcal{F}}(\mathbf{p})$ form a family of controls on $(\mathbb{M}, \delta \| \cdot \|)$. In addition for all $\mathbf{p}, \mathbf{q} \in \mathbb{M}$

$$d_{\mathcal{F}}(\mathbf{p},\mathbf{q}) = T_{\mathcal{B}_{\mathcal{F}}}(\mathbf{p},\mathbf{q}).$$

Proposition 3 The Reeds-Shepp metrics $(\mathcal{F}_{\varepsilon})_{0 \le \varepsilon \le 1}$ and $(\mathcal{F}_{\varepsilon}^+)_{0 \le \varepsilon \le 1}$ are indeed metrics in the sense of Definition 11, for any $\varepsilon \in [0,1]$. The associated controls $\mathcal{B}_{\varepsilon} := \mathcal{B}_{\mathcal{F}_{\varepsilon}}$, $\mathcal{B}_{\varepsilon}^+ := \mathcal{B}_{\mathcal{F}_{\varepsilon}^+}$ depend continuously on the parameter $\varepsilon \in [0,1]$, and satisfy the inclusions $\mathcal{B}_{\varepsilon}(\mathbf{p}) \subset \mathcal{B}_{\varepsilon'}(\mathbf{p})$ and $\mathcal{B}_{\varepsilon}^+(\mathbf{p}) \subset \mathcal{B}_{\varepsilon'}(\mathbf{p})$ for any $\mathbf{p} \in \mathbb{M}$ and $0 \le \varepsilon \le \varepsilon' \le 1$.

Proposition 3 allows to apply the results of §A.1 to the Reeds-Shepp metrics. Theorem 2 then directly follows from Corollary 4. The only remaining non-trivial claim in Proposition 3 is the continuity of the controls on \mathbb{M} , recall Definitions 8, and their convergence $\mathcal{B}_{\varepsilon} \to \mathcal{B}_0$ as $\varepsilon \to 0$, as required in Corollary 4. These two properties are implied by the continuity on $[0, 1] \times \mathbb{M}$, that we next prove, of the following maps

$$[0,1] \times \mathbb{M} \ni (\varepsilon, \mathbf{p}) \to \mathcal{B}_{\varepsilon}(\mathbf{p}) \in \mathfrak{C}(T_{\mathbf{p}}\mathbb{M}), [0,1] \times \mathbb{M} \ni (\varepsilon, \mathbf{p}) \to \mathcal{B}_{\varepsilon}^{+}(\mathbf{p}) \in \mathfrak{C}(T_{\mathbf{p}}\mathbb{M}),$$

$$(78)$$

with $\mathfrak{C}(T_{\mathbf{p}}\mathbb{M})$ defined in Definition 7 and equipped with the Hausdorff distance.

Lemma 5 Let B be a compact subset of a metric space \mathbb{E} , and let $\varphi \in C^0(B, \mathbb{E})$. Then

$$\mathcal{H}(B,\varphi(B)) \le \sup_{x \in B} d(x,\varphi(x)).$$

This basic lemma, stated without proof, is used in the next lemma to obtain an explicit estimate of the Hausdorff distance between the controls sets of the Reeds-Shepp models.

Lemma 6 Let $\mathbf{n}_1, \mathbf{n}_2 \in \mathbb{S}^{d-1}$, let $a_1, a_2, b_1, b_2 \ge 1$, and let $\varepsilon_1, \varepsilon_2 \in [0, 1]$. For each $i \in \{1, 2\}$, let B_i be the collection of all $(\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in \mathbb{R}^d \times \mathbb{R}^d$ obeying

$$\dot{\mathbf{h}} \cdot \mathbf{n}_{i} = 0,$$

$$\begin{cases}
a_{i}^{2} \|\dot{\mathbf{n}}\|^{2} + b_{i}^{2} \left(|\dot{\mathbf{x}} \cdot \mathbf{n}_{i}|^{2} + \varepsilon_{i}^{-2} \|\dot{\mathbf{x}} \wedge \mathbf{n}_{i}\|^{2} \right) \leq 1, \quad \varepsilon_{i} > 0 \\
a_{i}^{2} \|\dot{\mathbf{n}}\|^{2} + b_{i}^{2} |\dot{\mathbf{x}} \cdot \mathbf{n}_{i}|^{2} \leq 1 \quad and \quad \dot{\mathbf{x}} \wedge \mathbf{n}_{i} = 0, \quad \varepsilon_{i} = 0.
\end{cases}$$

$$Then \, \mathcal{H}(B_{1}, B_{2}) \leq |a_{1}^{-1} - a_{2}^{-1}| + |b_{1}^{-1} - b_{2}^{-1}| \\
+ \sqrt{2(1 - \mathbf{n}_{1} \cdot \mathbf{n}_{2})} + |\varepsilon_{1} - \varepsilon_{2}|.$$
(79)

The same estimate holds for the sets B_i^+ , $i \in \{1, 2\}$, defined by the inequalities

$$\begin{cases} \dot{\mathbf{n}} \cdot \mathbf{n}_i = 0, \\ a_i^2 \|\dot{\mathbf{n}}\|^2 + b_i^2 \left((\dot{\mathbf{x}} \cdot \mathbf{n}_i)_+^2 + \varepsilon_i^{-2} (\|\dot{\mathbf{x}} \wedge \mathbf{n}_i\|^2 + (\dot{\mathbf{x}} \cdot \mathbf{n}_i)_-^2) \right) \leq 1, \\ if \varepsilon_i > 0, \\ a_i^2 \|\dot{\mathbf{n}}\|^2 + b_i^2 (\dot{\mathbf{x}} \cdot \mathbf{n}_i)_+^2 \leq 1 \quad and \quad \dot{\mathbf{x}} \wedge \mathbf{n}_i = 0, \quad \dot{\mathbf{x}} \cdot \mathbf{n}_i \geq 0, \\ if \varepsilon_i = 0. \end{cases}$$

Proof. It suffices to establish the announced estimate (79) when the tuples $(a_i, b_i, \mathbf{n}_i, \varepsilon_i), i \in \{1, 2\}$, differ by a *single* element of the four, and then to use the sub-additivity of the Hausdorff distance. In each case we apply Lemma 5 to a well chosen surjective map $\varphi: B_1 \to B_2 \text{ (resp } \varphi^+: B_1^+ \to B_2^+\text{)}.$

- Case $a_1 \neq a_2$. Assume w.l.o.g. that $a_1 < \infty$, and observe that for all $(\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in B_1$ one has $a_1 ||\dot{\mathbf{x}}|| \leq 1$, hence $||a_1\dot{\mathbf{x}}/a_2 \dot{\mathbf{x}}|| \leq |a_1^{-1} a_2^{-1}|$. Choose $\varphi(\dot{\mathbf{x}}, \dot{\mathbf{n}}) := (a_1\dot{\mathbf{x}}/a_2, \dot{\mathbf{n}})$.
- Case $\mathbf{h}_1 \neq \mathbf{h}_2$. As above, with $\varphi(\dot{\mathbf{x}}, \dot{\mathbf{n}}) := (\dot{\mathbf{x}}, \mathbf{h}_1 \dot{\mathbf{n}}/b_2)$, yielding upper bound $|b_1^{-1} b_2^{-1}|$. Case $\mathbf{n}_1 \neq \mathbf{n}_2$. Let R be the rotation of \mathbb{R}^d which maps \mathbf{n}_1 onto \mathbf{n}_2 , in such a way that it maps the space orthogonal to the plane Span($\mathbf{n}_1, \mathbf{n}_2$) onto itself. A simple calculation yields $||R \mathrm{Id}|| = 2 \sin[\frac{1}{2}\cos^{-1}(\mathbf{n}_1 \cdot \mathbf{n}_2)] = \sqrt{2(1 \mathbf{n}_1 \cdot \mathbf{n}_2)}$. The result follows by choosing $\varphi(\dot{\mathbf{x}}, \dot{\mathbf{n}}) := (R\dot{\mathbf{x}}, R\dot{\mathbf{n}})$, so that $\|\varphi(\dot{\mathbf{x}}, \dot{\mathbf{n}}) - (\dot{\mathbf{x}}, \dot{\mathbf{n}})\| \le \|R - \operatorname{Id}\|\sqrt{\|\dot{\mathbf{n}}\|^2 + \|\dot{\mathbf{x}}\|^2} \le \|R - \operatorname{Id}\|$ for all $(\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in B_1$ as announced.
- Case $\varepsilon_1 \neq \varepsilon_2$. Assume w.l.o.g. that $\varepsilon_1 > 0$, and consider the orthogonal projections

$$P_1(\dot{\mathbf{x}}) := (\dot{\mathbf{x}} \cdot \mathbf{n}_1)\mathbf{n}_1 \qquad \qquad P_1^{\perp}(\dot{\mathbf{x}}) := (\mathrm{Id} - P_1)(\dot{\mathbf{x}}).$$

Note that $P_1^{\perp}(\dot{\mathbf{x}}) \leq \varepsilon_1$ if $(\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in B_1$, and that $\|\dot{\mathbf{x}}\| \leq \varepsilon_1$ if $(\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in B_1^+$ and $\dot{\mathbf{x}} \cdot \mathbf{n}_1 \leq 0$. The result follows by choosing

$$\begin{split} \varphi(\dot{\mathbf{x}}, \dot{\mathbf{n}}) &:= \left(P_1(\dot{\mathbf{x}}) + \frac{\varepsilon_2}{\varepsilon_1} P_1^{\perp}(\dot{\mathbf{x}}), \dot{\mathbf{n}} \right), \\ \varphi^+(\dot{\mathbf{x}}, \dot{\mathbf{n}}) &:= \begin{cases} \varphi(\dot{\mathbf{x}}, \dot{\mathbf{n}}) & \text{if } \dot{\mathbf{x}} \cdot \mathbf{n}_1 \ge 0, \\ \left(\frac{\varepsilon_2}{\varepsilon_1} \dot{\mathbf{x}}, \dot{\mathbf{n}}\right) & \text{otherwise.} \end{cases} \end{split}$$

Proof of Proposition 3. Since working with Hausdorff distances on the abstract tangent bundle $T\mathbb{M}$ is not very practical, we make use of the canonical embedding $\mathcal{I}: \mathbb{R}^d \times \mathbb{S}^{d-1} \to \mathbb{R}^d \times \mathbb{R}^d$ of the manifold \mathbb{M} into the Euclidean vector space \mathbb{R}^{2d} given by $(\mathbf{x}, \mathbf{n}) \mapsto (\mathbf{x}, \mathbf{n})$, which has bounded distortion. It suffices to prove the continuity of the image of the control sets $(\varepsilon, \mathbf{p}) \to d\mathcal{I}(\mathbf{p}, \mathcal{B}_{\mathcal{F}_{\varepsilon}}(\mathbf{p}))$ (resp. likewise with $\mathcal{F}_{\varepsilon}^+$) by the tangent maps to this embedding, which follows by Lemma 6. Indeed the lemma shows that

$$((\varepsilon_1, \mathbf{p}_1) \to (\varepsilon_2, \mathbf{p}_2)) \implies (\mathcal{H}(B_{\mathcal{F}_{\varepsilon_1}}, B_{\mathcal{F}_{\varepsilon_2}}) \to 0),$$

and it includes the spherical constraint via the velocity constraint $\dot{\mathbf{n}} \cdot \mathbf{n}_i = \frac{d}{dt} (\mathbf{n}(t) \cdot \mathbf{n}(t))|_{t=0} = 0$ for a smooth curve $\gamma(t) = (\mathbf{x}(t), \mathbf{n}(t))$ passing through $\gamma(0) = (\mathbf{x}_i, \mathbf{n}_i)$.

B Backtracking of Geodesics in $(\mathbb{M}, d_{\mathcal{F}})$

This section is devoted to a generic ingredient in the proof of Theorem 4, regarding backtracking of Geodesics in the (pseudo)-Metric Space $(\mathbb{M}, d_{\mathcal{F}})$ in general. Although, these results are standard in Finsler Geometry, we aim to provide a concise overview.

Lemma 7 Let F be an asymmetric norm on a vector space \mathbb{E} , and assume that F^* is differentiable at $\hat{\mathbf{p}} \in \mathbb{E}^*$. Then

$$F(\mathrm{d}F^*(\hat{\mathbf{p}})) = 1, \qquad \langle \hat{\mathbf{p}}, \mathrm{d}F^*(\hat{\mathbf{p}}) \rangle = F^*(\hat{\mathbf{p}}).$$

Proof. The 1st claim follows by differentiation of F^*

$$F^*(\hat{\mathbf{p}}) = \sup_{\dot{\mathbf{p}} \in \mathbb{E} \setminus \{0\}} \frac{\langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle}{F(\dot{\mathbf{p}})} = \max_{F(\dot{\mathbf{p}})=1} \langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle.$$

The 2nd claim is Euler's formula for homogeneous functions.

Proposition 4 Let $\mathbf{p}_{\mathrm{S}}, \mathbf{p}_{\mathrm{T}} \in \mathbb{M}$, let γ be a minimal geodesic from \mathbf{p}_{S} to \mathbf{p}_{T} w.r.t. a continuous metric \mathcal{F} , and let $t \in [0, 1]$. Assume that the distance map U from \mathbf{p}_{S} is differentiable at $\gamma(t)$, and that the dual metric \mathcal{F}^* is differentiable w.r.t. the second variable at $(\gamma(t), \mathrm{d}U(\gamma(t)))$. Then γ is differentiable at time t and with $L := d_{\mathcal{F}}(\mathbf{p}_{\mathrm{S}}, \mathbf{p}_{\mathrm{T}})$

$$\dot{\gamma}(t) = L \,\mathrm{d}_{\hat{\mathbf{p}}} \mathcal{F}^*(\gamma(t), \mathrm{d}U(\gamma(t))), \ \gamma(0) = \mathbf{p}_S, \gamma(1) = \mathbf{p}_T.$$

$$\tag{80}$$

Proof. The path γ has constant speed L, and $t \mapsto U(\gamma(t))$ increases linearly from 0 to L on it. Let $t \in [0, 1]$ be as in the statement of the proposition, and let

$$\Gamma(t) := \lim_{n \to \infty} (\gamma(t + \varepsilon_n) - \gamma(t)) / \varepsilon_n$$

for some sequence $\varepsilon_n \to 0$. Then

$$\mathcal{F}(\gamma(t), \dot{\Gamma}(t)) = L$$
 and $\langle dU(\gamma(t)), \dot{\Gamma}(t) \rangle = L$.

For typographic simplicity let us denote $\mathbf{p} := \gamma(t)$, $\dot{\mathbf{p}} := \dot{\Gamma}(t)$, $F = \mathcal{F}(\mathbf{p}, \cdot)$ and $F^* := \mathcal{F}^*(\mathbf{p}, \cdot)$. By Lemma 7 and the eikonal equation (5), the vector $\dot{\mathbf{q}} = \mathrm{d}F^*(\mathrm{d}U(\mathbf{p}))$ obeys

$$F(\dot{\mathbf{q}}) = F(\mathrm{d}F^*(\mathrm{d}U(\mathbf{p}))) = 1,$$

$$\langle \mathrm{d}U(\mathbf{p}), \dot{\mathbf{q}} \rangle = \langle \mathrm{d}U(\mathbf{p}), \mathrm{d}F^*(\mathrm{d}U(\mathbf{p})) \rangle = F^*(\mathrm{d}U(\mathbf{p})) = 1.$$

Note that the duality-bracket/norm inequality is saturated by $\langle dU(\mathbf{p}), \dot{\mathbf{q}} \rangle = 1 = F^*(dU(\mathbf{p}))F(\dot{\mathbf{q}})$, and that the assumed differentiability of the dual norm F^* at the point $\hat{\mathbf{p}} = dU(\mathbf{p})$ implies the strict convexity of the primal norm F (up to 1-homogeneity) at the point $dF^*(\hat{\mathbf{p}}) = \dot{\mathbf{q}}$. Hence $\dot{\mathbf{q}}$ is the unique solution to the system " $F^*(\dot{\mathbf{q}}) = 1$ and $\langle dU(\mathbf{p}), \dot{\mathbf{q}} \rangle = 1$ ", and therefore $\dot{I} = L\dot{\mathbf{q}}$. This implies the differentiability of γ at time t, and the announced equality (80).

Remark 16 (Lagrangians and Hamiltonians) Given an arbitrary metric \mathcal{F} on \mathbb{M} , its half-square $\mathfrak{L} := \frac{1}{2}\mathcal{F}^2 : T(\mathbb{M}) \to [0, +\infty]$ is usually called the Lagrangian. The shortest path problem (1) can be reformulated in terms of the Lagrangian, thanks to the Cauchy-Schwartz's inequality which gives

$$d_{\mathcal{F}}(\mathbf{p}, \mathbf{q})^2 = \inf\{\int_0^1 \mathcal{F}(\gamma(t), \dot{\gamma}(t))^2 dt \mid \gamma \in \operatorname{Lip}([0, 1], \mathbb{M}) , \gamma(0) = \mathbf{p}, \gamma(1) = \mathbf{q}\}.$$
(81)

A path γ is a minimizer of (81) iff it is simultaneously normalized and a minimizer of (1). The Hamiltonian \mathfrak{H} is the Lengendre-Fenchel transform of its Lagrangian \mathfrak{L} w.r.t. the second variable, hence $\mathfrak{H} = \frac{1}{2}(\mathcal{F}^*)^2$ (for details see [6, ch.14.8]) The eikonal equation can thus be rephrased in terms of the Hamiltonian:

$$\mathcal{F}^*(\mathbf{p}, \mathrm{d}U(\mathbf{p})) = 1 \iff \mathfrak{H}(\mathbf{p}, \mathrm{d}U(\mathbf{p})) = \frac{1}{2}.$$

The Hamiltonian can also be used to reformulate the backtracing ODE of geodesics, thanks to the following identity which follows from the eikonal equation: for any $\mathbf{p} \in \mathbb{M}$

$$d_{\hat{\mathbf{p}}}\mathfrak{H}(\mathbf{p}, \mathrm{d}U(\mathbf{p})) = \mathcal{F}^{*}(\mathbf{p}, \mathrm{d}U(\mathbf{p})) d_{\hat{\mathbf{p}}}\mathcal{F}^{*}(\mathbf{p}, \mathrm{d}U(\mathbf{p})) = d_{\hat{\mathbf{p}}}\mathcal{F}^{*}(\mathbf{p}, \mathrm{d}U(\mathbf{p})).$$
(82)

In geometric control theory this Hamiltonian is often referred to the 'fixed time Hamiltonian of the action functional', cf. [2,9,66], and is typically used [57] in the Pontryagin maximum principle [2] for (sub-)Riemannian geodesics.

C Characterization of Cusps: Proof of Lemma 1

Consider Lemma 1. The structure of this lemma is $a \Leftrightarrow b \Leftrightarrow c$. The implication $a \Rightarrow b$ is trivial. The equivalence $b \Leftrightarrow c$ follows by Theorems 4, 2. The implication $b \Rightarrow a$ remains.

Suppose the *d*-th spatial control aligned with $\mathbf{n}(t_0)$, recall (20), vanishes: $\tilde{u}(t_0) = 0$. Now we show by contradiction that in this case $\dot{\tilde{u}}(t_0) \neq 0$. Suppose $\tilde{u}(t_0) = \dot{\tilde{u}}(t_0) = 0$.

Then by application of the PMP (Pontryagin Maximum Principle), similar to [9, App.A], [27]) and coercivity/invertibility of the SR-metric tensor $\mathcal{G}_0|_{\gamma(t_0)}$, recall (26), constrained to the horizontal part of the tangent space $\Delta|_{\gamma(t)} = \{(\mathbf{p}_0 = (\mathbf{x}_0, \mathbf{n}_0), \dot{\mathbf{p}}_0 = (\dot{\mathbf{x}}_0, \dot{\mathbf{n}}_0)) \in T(\mathbb{M}) \mid \mathbf{n}_0 \equiv \dot{\mathbf{x}}_0\}$, that the (analytic) spatial control variable $\tilde{u} = \mathcal{C}_1^{-2} \tilde{\lambda}$ vanishes for all times (for d = 2 this is directly deduced from the pendulum phase portrait [57] in momentum space). This leaves only purely angular momentum and motion, contradicting $\dot{\mathbf{x}}(\cdot) \neq \mathbf{0}$ in Lemma 1.

Next we verify $\tilde{u}(t_0) = \dot{\tilde{u}}(t_0) = 0 \Rightarrow \tilde{\lambda}(t_0) = 0 = \tilde{\lambda}(t_0)$. By the chain rule for differentiation (applied to the *d*-th spatial momentum component $\tilde{\lambda}(t) = \langle \lambda(t), (\mathbf{n}(t), \mathbf{0}) \rangle$):

$$\frac{d}{dt}\tilde{\lambda}(t)\Big|_{t=t_0} = \frac{d}{dt}(\mathcal{C}_1(\gamma(t)))^{-2}\tilde{u}(t)\Big|_{t=t_0} = \frac{d}{dt}(\mathcal{C}_1(\gamma(t)))^{-2}\Big|_{t=t_0}\tilde{u}(t_0) + \frac{d}{dt}(\mathcal{C}_1(\gamma(t)))^{-2}\Big|_{t=t_0}\dot{u}(t_0) = 0.$$

We deduce from PMP's Hamiltonian equations (cf. [27]) that

$$ilde{\lambda}(t_0) = ilde{\lambda}(t_0) = 0 \Rightarrow ilde{\lambda}(\cdot) = 0 \Rightarrow ilde{u}(\cdot) = 0.$$

D Short \neq Straight: A Special Cartan Connection on $SE(2) \equiv \mathbb{R}^2 \rtimes S^1$

D.1 Connections

With the left-invariant vector fields and covector fields defined, we can now define rotation and translation covariant derivatives on images and orientation scores. If we would like to define second order differential operators (such as the Hessian), we need to know how to take (directional) derivatives of such vector fields. The way this is done is prescribed by a choice of **connection** ∇ and the induced **covariant derivatives** ∇_X for each choice of tangent vector field X on a smooth manifold M with tangent bundle T(M) given by

$$T(M) = \{ (p, \dot{p}) \mid p \in M, \dot{p} \in T_p(M) \}.$$

To indicate the directional derivative of a smooth vector field Y, along a smooth vector field X as defined via the connection ∇ we write $\nabla_X Y$.

Given two smooth vector fields X and Y on a smooth Riemannian manifold (M, \mathcal{G}) expressed in a basis ∂_{x^i} , i.e.,

$$X|_{m} = \sum_{i=1}^{\dim(M)} x^{i}(m) \ \partial_{x^{i}}|_{m}, \quad \text{and} \quad Y|_{m} = \sum_{i=1}^{\dim(M)} y^{i}(m) \ \partial_{x^{i}}|_{m}.$$
(83)

Expressed in this basis connection takes the following form

$$\nabla_X Y := \sum_{k=1}^{\dim(M)} \left(\dot{y}^k + \sum_{i,j=1}^{\dim(M)} \Gamma_{ij}^k x^i y^i \right) \partial_{x^k}, \tag{84}$$

with \dot{y}^k the directional derivative along direction vector $X|_m$ of the k^{th} component of the vector $Y|_k$ given by

$$\dot{y}^k\Big|_p := \left.\mathrm{d} y^k\Big|_p \left(\left.X\right|_p\right),$$

and with the derivative of y^k given by

$$\left\langle \, \mathrm{d} y^k \right|_p, \dot{p}
ight
angle = \sum_{i=1}^{\dim(M)} \left(\left. \partial_{x^i} \right|_p y^k
ight) \dot{p}^i.$$

for all $p \in M$ and all tangents $\dot{p} = \sum_{i=1}^{\dim(M)} \dot{p}^i \partial_{x^i} \Big|_p \in T_p(M)$. In (84), the first term \dot{y}^k is thus the directional derivative (along X) of the separate components of the vector field Y. The second term compensates for motion of the basis (serving as a reference frame) $\{\partial_{x^i}\}_{i=1}^{\dim(M)}$ and is characterized by the **Christoffel symbols** Γ_{ij}^k . The Christoffel symbols are determined by a connection and a choice of basis for the tangent bundle. They directly depend on the choice of basis, and do not have a generic geometric basis independent meaning.

In the subsequent subsections we consider two types of connections: the Levi-Civita connection and the left Cartan connection, of which we will see in Subsec. D.4 that the latter is of most interest to us.

Remark 17 In the above we introduced a connection on the tangent bundle T(M) via a basis on the tangent bundle. In general a connection on a tangent bundle is defined without the use of a basis. See the exercise below.

Remark 18 Connections can be generalized to fiber bundles [46], where sections need not be (tangent) vector fields, and where fibers need not be tangent spaces, or need not even be vector spaces! In this course we avoid such generalizations, even though many imaging applications do require such generalizations [29–31].

Exercise 19 Consider a smooth manifold M and its tangent bundle T(M). Consider the space of vector fields above M and denote this by $\chi(M)$. Let X be such a tangent vector field on M. Then the covariant derivative ∇_X a linear map $\nabla : \chi(M) \to \chi(M)$, and a Koszul connection ∇ is defined must satisfy (regardless coordinate systems):

$$\begin{cases} \nabla_X(Y_1+Y_2) = \nabla_X Y_1 + \nabla_X Y_2, \\ \nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla X_2 Y, \\ \nabla X(fY) = f \nabla_X Y + X(f) Y, \\ \nabla_{fX} Y = f \nabla_X Y, \end{cases}$$
(85)

for all vector fields $X, X_1, X_2, Y, Y_1, Y_2 \in \chi(M)$, and all $f \in C^{\infty}(M, \mathbb{R})$. Show that (85) indeed implies the standard form (84) in a basis $\{\partial_{x^i}\}_{i=1}^{\dim(M)}$.

D.2 The Levi-Civita Connection

Usually in Riemannian geometry people work with the Levi-Civita connection ∇^{LC} , which is by definition the unique torsion free metric compatible¹⁹ connection on the Riemannian manifold (M, \mathcal{G}) . The requirement that the connection should be torsion free means that

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] = 0,$$

which holds for $\nabla = \nabla_{LC}$.

Under these conditions the Levi-Civita connection is uniquely defined (fundamental theorem of Riemannian geometry), and the Christoffel symbols are then given by

$$(\Gamma^{LC})_{ij}^{k} = \frac{1}{2} \sum_{m=1}^{\dim M} g^{im} \left(\frac{\partial g_{mk}}{\partial x^{l}} + \frac{\partial g_{ml}}{\partial x^{k}} - \frac{\partial g_{kl}}{\partial x^{m}} \right)$$
(86)

and with matrix $[g_{ij}]=[\mathcal{G}(\partial_{x^i},\partial_{x^j})]$ and $[g^{ij}]$ the corresponding inverse matrix.

Exercise 20 Let
$$\nabla$$
 be a connection on a Riemannian manifold (M, \mathcal{G}) . Show that

$$\begin{split} T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y] = 0, \\ \nabla_X \mathcal{G}(Y,Z) &= \mathcal{G}(\nabla_X Y,Z) + \mathcal{G}(Y,\nabla_X Z), \end{split}$$

indeed leads to the unique Levi-Civita connection $\nabla = \nabla_{LC}$ given by (86).

For the Levi-Civita connection we have that stationary curves of the Riemannian distance

$$d_M(m_0, m_1) := \min_{\substack{\gamma \in \text{Lip}([0, 1], M) \\ \gamma(0) = m_1 \\ \gamma(1) = m_1}} \int_0^1 \sqrt{\mathcal{G}|_{\gamma(t)} \left(\dot{\gamma}(t), \dot{\gamma}(t)\right)} dt,$$
(87)

satisfy the following geodesic equations

$$\nabla^{LC}_{\dot{\gamma}}\dot{\gamma} = 0, \tag{88}$$

see e.g. [46, Lemma 5.1.1]. Expressed in holonomic coordinates $\{x^i\}_{i=1}^{\dim(M)}$ this gives

$$\ddot{\gamma}^{k} + \sum_{i,j=1}^{\dim(M)} (\Gamma^{LC})^{k}_{ij} \dot{\gamma}^{i} \dot{\gamma}^{j} = 0, \quad \text{with} \quad \dot{\gamma}^{k} = \langle \, \mathrm{d}x^{k} \Big|_{\gamma} \,, \dot{\gamma} \rangle, \quad \ddot{\gamma}^{k} = \frac{d}{d\tau} \dot{\gamma}^{k}, \tag{89}$$

Exercise 21 Two assignments regarding the geodesic equation and the Levi-Civita connection:

- a) Derive (89) from (88) and (84).
- b) Verify the derivations of the geodesic equations in [46, Lemma 5.1.1], and indicate in which equalities metric compatibility of ∇^{LC} is used, and in which equalities the vanishing of torsion of ∇^{LC} is used.

In general, curves γ for which $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$ are called **auto parallel** curves with respect to the connection ∇ . The auto parallel curves with respect to the Levi-Civita connection ∇^{LC} are the geodesics, which in the standard Euclidean case are straight lines as we show in the next example.

Example 1 Consider the standard 2D Euclidean setting with $M = \mathbb{R}^2$, with the metric tensor components $g_{ij} = \delta_{ij}$, and metric tensor $dx^1 \otimes dx^1 + dx^2 \otimes dx^2$, with duals $\{dx^1, dx^2\}$ relative to the standard basis $\{\partial_{x1}, \partial_{x2}\}$. Since the metric tensor components g_{ij} are constant over M, the Christoffel symbols (computed via (86)) are given by

$$\Gamma_{ij}^k = 0.$$

The geodesic equations (88), (89), then reduce to

$$\begin{aligned} \ddot{\gamma}^k(\tau) &= 0, \\ \dot{\gamma}^k(\tau) &= c^k. \end{aligned}$$

The geodesics in the Riemannian manifold $(M, dx^1 \otimes dx^1 + dx^2 \otimes dx^2)$ are thus straight lines.

¹⁹ Metric compatible means $\nabla_X \mathcal{G}(Y, Z) = \mathcal{G}(\nabla_X Y, Z) + \mathcal{G}(Y, \nabla_X Z)$

In the 2D Euclidean case the straight curves and shortest curves (geodesics) coincide, and both are auto parallel curves with respect to the Levi-Civita connection. In our sub-Riemannian geometry on $SE(2) \equiv \mathbb{R}^2 \rtimes S^1$ this is not the case! Here we deal with a curved and torqued geometry (with $\Gamma_{ij}^k \neq 0$ relative to the left-invariant basis) and the "straight curves" and shortest curves no longer coincide. As we will see in the next Subsection, the straight curves (the exponential curves) are instead the auto parallel curves with respect to the Cartan connection ∇ .

D.3 A Special Cartan Connection on SE(2)

In the remainder of this section we consider the special case where the Riemannian manifold equals $M = \mathbb{M} = (SE(2), \mathcal{G}_{\epsilon})$ with \mathcal{G}_{ϵ} given by (26). Recall also (from Eq. (46)) that the left-invariant vector fields \mathcal{A}_i and their duals ω^i were given by

 $\begin{aligned} \mathcal{A}_1 &= \cos\theta\,\partial_x + \sin\theta\,\partial_y, \ \mathcal{A}_2 &= -\sin\theta\,\partial_x + \cos\theta\,\partial_y, \ \mathcal{A}_3 &= \partial_\theta, \\ \omega^1 &= \cos\theta\,\mathrm{d}x + \sin\theta\,\mathrm{d}y, \ \omega^2 &= -\sin\theta\,\mathrm{d}x + \cos\theta\,\mathrm{d}y, \ \omega^3 &= \mathrm{d}\theta. \end{aligned}$

The duals are defined by the property

 $\langle \omega^i, \mathcal{A}_j \rangle = \delta^i_j,$

where δ_{i}^{i} denotes the usual Kronecker delta, i.e. 1 if i = j and 0 otherwise.

Then the metric tensor \mathcal{G}_{ϵ} for d = 2 can be expressed as

$$\mathcal{G}_{\epsilon} = \xi^2 \omega^1 \otimes \omega^1 + \xi^2 \epsilon^{-2} \omega^2 \otimes \omega^2 + \omega^3 \otimes \omega^3 \tag{90}$$

and we have

$$\mathcal{G}_{\epsilon}|_{\gamma(\tau)}\left(\dot{\gamma}(\tau),\dot{\gamma}(\tau)\right) = +\xi^{2}|\dot{x}(\tau)\cos\theta(\tau) + \dot{y}(\tau)\sin\theta(\tau)|^{2} + \frac{\xi^{2}}{\epsilon^{2}}|-\dot{y}(\tau)\sin\theta(\tau) + \dot{x}(\tau)\cos\theta(\tau)|^{2} + |\dot{\theta}(\tau)|^{2}, \quad (91)$$

with tangent $\dot{\gamma}(\tau) = (\dot{x}(\tau), \dot{y}(\tau), \dot{\theta}(\tau))$ along curve $t \mapsto \gamma(\tau) = (x(\tau), y(\tau), \theta(\tau))$ in SE(2).

Exercise 22 Verify by a few computations that (46) for d = 2 indeed simplifies to (119) and (91).

Recall that our left-invariant vector fields \mathcal{A}_i on SE(2) were obtained from the basis $\{\partial_x|_e, \partial_y|_e, \partial_\theta|_e\}$ in $T_e(SE(2))$ by the push-forward $(L_g)_*$ of the left-multiplication L_g . Conversely, we can use the so-called Maurer-Cartan form $(L_g^{-1})_*$ to map all $T_g(SE(2))$ back to $T_e(SE(2))$. This allows us to 'connect' all tangent spaces in a principle fiber bundle structure [71]. When using the adjoint representation for the associated vector bundle [36, App.B] one obtains a special Cartan connection on the tangent bundle.

Here we avoid these technicalities and just provide the resulting Cartan connection ∇ on the tangent bundle $T(\mathbb{M})$ (and provide some intuitive explanations):

$$\nabla_{\dot{\gamma}}Y := \sum_{k=1}^{3} \left(\dot{y}^{k} + \mathrm{d}\omega^{k} \Big|_{\gamma(\tau)} \left(\dot{\gamma}(\tau), Y \right) \right) \mathcal{A}_{k}, \tag{92}$$

with $d\omega^k |_{\gamma(\tau)} (\dot{\gamma}(\tau), Y) \in \mathbb{R}$, and with

$$\dot{y}^k(\tau) := \frac{d}{d\tau} y^k(\tau) := \frac{d}{d\tau} \left. \omega^k \right|_{\gamma(\tau)} (Y|_{\gamma(\tau)})$$

which produces a vector field $\nabla_{\dot{\gamma}} Y$ that indicates the covariant derivative of vector field Y along the flow-field of γ .

Remark 19 When expressing the derivative in the moving dual frame one must keep in mind that the moving dual frame is not an exact frame. An exact differential k-form is a differential form that is the exterior derivative of another differential (k-1)-form. In the \mathbb{R}^2 case the co-vectors dx^i (recall Example 1) are exact 1-forms, and therefore satisfy

$$\mathrm{d}\mathrm{d}x^i = 0.$$

In SE(2) this is not the case for the local dual frame $\{\omega^i\}_{i=1}^3$ defined in (66) due to the structural formulas of Cartan:

$$d\omega^{k} = \sum_{i,j=1}^{3} \frac{1}{2} c_{ij}^{k} (\omega^{i} \wedge \omega^{j}) = \sum_{i,j \in \{1,2,3\}, i < j} c_{ij}^{k} (\omega^{i} \wedge \omega^{j}),$$
(93)

where c_{ij}^k denote the structure constants of the Lie algebra of SE(2), recall (54).

Exercise 23 Regarding the structure constants and Cartan structural formula on SE(2):

- a) Show that the only non-zero structure constants are $c_{13}^2 = -c_{31}^2 = -1$.
- b) Express each of the covectors $d\omega^1$, $d\omega^2$ in terms of dx, dy.
- c) Show that $\omega^1 \neq d(x \cos \theta + y \sin \theta)$.

d) Show that $d\omega^3 = 0$.

e) Verify (103) for k = 1, 2, 3.

The intuition is that this covariant derivative (92) takes into account that, when taking the covariant derivative of vector field $Y = \sum_{k=1}^{3} y^{k} \mathcal{A}_{k}$, one differentiates its components

$$y^k(\tau) := \omega^k \Big|_{\gamma(\tau)} (Y|_{\gamma(\tau)}),$$

and simultaneously accounts for the movement of the (dual) left-invariant frame (recall Cartan's structural formula (103)). Expressed in components of the left-invariant frame and coframe this becomes:

$$\nabla_{\dot{\gamma}}Y := \sum_{k=1}^{3} \left(\dot{y}^k + \sum_{i,j=1}^{3} c^k_{ij} \dot{\gamma}^i y^j \right) \mathcal{A}_k.$$

$$\tag{94}$$

So now the Christoffels in (86) are again the structure constants $\Gamma_{ij}^k = c_{ij}^k$. Since $c_{ij}^k = -c_{ji}^k$ we have that $\sum_{i,j=1}^3 c_{ij}^k \dot{\gamma}^i y^j = 0$ and see that the auto parallel curves are the exponential curves:

$$\nabla_{\dot{\gamma}}\dot{\gamma} = 0 \quad \Leftrightarrow \quad \begin{cases} \ddot{\gamma}^k(\tau) = 0, \\ \dot{\gamma}^k(\tau) = c^k. \end{cases}$$

By duality (cf. the exercise below), connection ∇ on the tangent bundle induces $T(\mathbb{M})$ also a connection ∇^* on the cotangent bundle $T^*(\mathbb{M})$:

$$\nabla_{\dot{\gamma}}^* \lambda = \sum_{i=1}^3 \left(\dot{\lambda}_i + \sum_{j,k=1}^3 c_{ij}^k \lambda_k \dot{\gamma}^j \right) \omega^i, \tag{95}$$

with a sign change before the Christoffels, and with $\dot{\lambda}_i(\tau) = \frac{d}{d\tau} \langle \lambda |_{\gamma(\tau)}, \mathcal{A}_i |_{\gamma(\tau)} \rangle$ and $\lambda = \lambda_i \omega^i |_{\mathbf{p}}$. Here we recall that the cotangent bundle is given by

 $T^*(\mathbb{M}) = \{ (\mathbf{p}, \hat{\mathbf{p}}) \mid \mathbf{p} \in \mathbb{M}, \hat{\mathbf{p}} \in T^*_{\mathbf{p}}(\mathbb{M}) \},\$

with $T^*_{\mathbf{p}}(\mathbb{M})$ the topological dual space to $T_{\mathbf{p}}(\mathbb{M})$, and that λ is a cotangent field (i.e. a section in the cotangent bundle).

Exercise 24 Verify that the formula (95), follows by $c_{ij}^k = -c_{ji}^k$, (92), and

$$0 = \frac{d}{d\tau} \delta^{i}_{j} = \frac{d}{d\tau} \langle \omega^{i} |_{\gamma(\tau)}, \mathcal{A}_{j} |_{\gamma(\tau)} \rangle = \langle \nabla^{*}_{\gamma(\tau)} \omega^{i} |_{\gamma(\tau)}, \mathcal{A}_{j} |_{\gamma(\tau)} \rangle + \langle \omega^{i} |_{\gamma(\tau)}, \nabla_{\gamma(\tau)} \mathcal{A}_{j} |_{\gamma(\tau)} \rangle$$

where brackets denote functional evaluation (not inner-products).

For the sub-Riemannian case $(\epsilon \downarrow 0)$ one has metric tensor

$$\mathcal{G}_0 = \xi^2 \,\omega^1 \otimes \omega^1 + \omega^3 \otimes \omega^3,\tag{96}$$

and a direct hard constraint

$$\mathbf{p} \in \Delta|_{\mathbf{p}} := \operatorname{span}\{\mathcal{A}_1|_{\mathbf{p}}, \mathcal{A}_3|_{\mathbf{p}}\},\tag{97}$$

on the tangent bundle (but not on the co-tangent bundle) and one has to rely on a partial connection $\overline{\nabla}$ instead:

$$\overline{\nabla}_{\dot{\gamma}}Y := \sum_{k \in \{1,3\}} \left(\dot{y}^k + \sum_{i,j \in \{1,3\}} c^k_{ij} \dot{\gamma}^i y^j \right) \mathcal{A}_k,$$

$$\overline{\nabla}^*_{\dot{\gamma}}\lambda := \sum_{i=1}^3 \left(\dot{\lambda}_i + \sum_{j \in \{1,3\}} \sum_{k=1}^3 c^k_{ij} \lambda_k \dot{\gamma}^j \right) \omega^i.$$
(98)

D.4 Why Do We Use the special Cartan Connection (for Image Processing)?

ġ

Now that we introduced the special Cartan connection by (94), let us see why it is a useful connection for our purposes. Here we list its nice properties:

- 'Straight curves', i.e. solutions to $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$, are the exponential curves. Left-invariant PDE's on SE(2) and wavefront propagation can be expressed in covariant derivatives and left-invariant flow; transport takes place along these curves.
- 'Shortest curves', have instead parallel momentum $\nabla_{\dot{\gamma}}^* \lambda = 0$. This can be observed by the nested sub-Riemannian spheres, where we note that covectors may be geometrically represented by local parallel planes, see for example Fig. 20.
- When considering second order best exponential curve fits for determining locally adaptive frames [37], which can be employed for both vessel enhancement and vessel curvature measurements, curve fitting boils down to eigenvector analysis of Hessians induced by this very connection [37, App.4].



Fig. 20 A: Geodesically equidistant surfaces $S_t^{\epsilon} = \{g \in SE(2) | d_{\epsilon}(0,g) = t\}$ and geodesic (in green) for the sub-Riemannian case: $\epsilon = 0$ and $\mathcal{C} = 1$. B: Geodesically equidistant surfaces S_t^{ϵ} and geodesic for the isotropic Riemannian case: $\epsilon = 1$ and $\mathcal{C} = 1$. Now the geodesics are straight lines. C: A set of horizontal exponential curves for which $\dot{\gamma}(\tau) = c^1 \mathcal{A}_1|_{\gamma(\tau)} + c^3 \mathcal{A}_3|_{\gamma(\tau)} \in \Delta$, with constant tangent vector components c^1 and c^3 .

 It correctly accounts for the motion of the moving frame of reference when taking derivatives of vector fields and covector fields.

Further details on this special choice can be found in Appendix E where we consider a parameterized class of Cartan connections $\nabla^{[\nu]}$ and where the case above corresponds to $\nu = 1$.

There is a fundamental relation between a special Cartan connection, and 'straight' and 'shortest' curves in the metric space $(\mathbb{M}, d_{\mathcal{F}_{\epsilon}})$ with:

$$d_{\mathcal{F}_{\epsilon}}(\mathbf{p}_{1},\mathbf{p}_{2}) = \inf \left\{ \int_{0}^{1} \mathcal{G}_{\epsilon}(\dot{\gamma}(t),\dot{\gamma}(t)) \,\mathrm{d}t \mid \gamma \in \mathrm{Lip}([0,1],\mathbb{M}), \ \gamma(0) = \mathbf{p}, \gamma(1) = \mathbf{q} \right\}.$$

In summary, we have the following equivalences

Theorem 6 In a Riemannian manifold $(SE(2), T(SE(2)), \mathcal{G}_{\epsilon})$, with the tangent bundle T(SE(2)) and metric tensor \mathcal{G}_{ϵ} defined in (119), the metric $d_{\mathcal{F}_e}$ defined in (125), and the Cartan connection defined in (92), we have the following relations for "straight" curves:

$$\gamma \text{ is a } \nabla \text{-straight curve} \Leftrightarrow \begin{array}{c} \gamma \text{ is an exponential} \\ curve \end{array} \Leftrightarrow \begin{array}{c} \nabla_{\dot{\gamma}} \dot{\gamma} = 0 \\ \psi \text{ locity} \end{array} \Leftrightarrow \begin{array}{c} \gamma \text{ has } \nabla \text{-auto parallel} \\ \psi \text{ velocity} \end{array}$$

and the following for "shortest" curves:

 $\gamma_{\epsilon} \text{ is a shortest curve } \Leftrightarrow \begin{array}{c} \gamma_{\epsilon} \text{ is a minimizing} \\ \text{curve in } d_{\epsilon} \end{array} \Leftrightarrow \begin{array}{c} \left\{ \nabla^{*}_{\dot{\gamma}_{\epsilon}} \lambda_{\epsilon} = 0 \\ \dot{\gamma}_{\epsilon} = \mathcal{G}_{\epsilon}^{-1} \lambda_{\epsilon} \end{array} \Leftrightarrow \begin{array}{c} \gamma_{\epsilon} \text{ has } \nabla^{*} \text{-parallel} \\ \text{momentum} \end{array} \right\}$

In a sub-Riemannian manifold $(SE(2), \Delta, \mathcal{G}_0)$ with tangent bundle Δ defined in (121), the sub-Riemannian metric tensor \mathcal{G}_0 defined in (122), and the partial Cartan connection defined in (98) we have the following relations for "straight" curves

$$\gamma \text{ is a } \overline{\nabla} \text{-straight curve} \Leftrightarrow \qquad \begin{array}{c} \gamma \text{ is a horizontal} \\ exponential curve} \Leftrightarrow \qquad \overline{\nabla}_{\dot{\gamma}} \dot{\gamma} = 0 \qquad \Leftrightarrow \begin{array}{c} \gamma \text{ has } \overline{\nabla} \text{-auto parallel} \\ velocity \end{array}$$

and the following for "shortest" curves

$$\gamma_0 \text{ is a shortest curve } \Leftrightarrow \begin{array}{c} \gamma_0 \text{ is a minimizing} \\ curve \text{ in } d_0 \end{array} \Leftrightarrow \begin{cases} \overline{\nabla}^*_{\dot{\gamma}_0} \lambda_0 = 0 \\ \dot{\gamma}_0 = \mathcal{G}_0^{-1} \mathbb{P}^*_\Delta \lambda_0 \end{cases} \Leftrightarrow \begin{array}{c} \gamma_0 \text{ has } \overline{\nabla}^* \text{-parallel} \\ momentum \end{cases}$$

in which $\mathbb{P}^*_{\Lambda}(\lambda_1\omega^1 + \lambda_2\omega^2 + \lambda_3\omega^3) = \lambda_1\omega^1 + \lambda^3\omega^3$.

Remark 20 Although beyond the scope of this course, the above considerations generalize to the case d > 2 and to the Lie group quotient R

$$\mathbb{R}^d \rtimes S^{d-1} := SE(d)/(\{\mathbf{0}\} \times SO(d-1)),$$

considered in detail in [28,27]. Moreover, they generalize to a wide class of Lie groups (non-commutative and 2-step nilpotent) and sub-Riemannian manifolds within, as explained in the next section

Remark 21 A generalization of the above theorem to a wide class of Lie groups is explained in Theorem 7 (that strongly builds on Theorem 8) in Appendix E.

E Cartan-connections $\nabla^{[\nu]}$ on Lie Groups

This section is quite technical. Therefore we first provide a brief paragraph of what is coming:

Applications of geometric flows to multi-orientation image processing require the choice of an (affine) connection on the Lie group G of roto-translations. Typical choices of such connections are called the (-), (0) and (+) connection. As the constructions of these connections in standard references is quite involved, we provide an overview. We show that these connections are members of a larger, one-parameter class of connections, and we motivate that the (+) connection is most suited for our image-analysis applications. The class $\nabla^{[\nu]}$, with $\nu \in \mathbb{R}$, is given by $\nabla^{[\nu]}_X Y = \nu[X, Y]$ for all left-invariant vector fields X, Y on G. Their auto-parallel curves are the exponential curves. Their torsion is $T[X, Y] = (2\nu - 1)[X, Y]$, and the (-), (0), and (+) connections arise for $\nu = 0, \frac{1}{2}, 1$.

We propose the case $\nu = 1$, as then Hamiltonian flows on $T^*(G)$ for Riemannian distance minimizers on G (induced by left-invariant metric tensor field \mathcal{G}) reduce to $\nabla_{\dot{\gamma}}^{[1]}\lambda = 0$ and $\dot{\gamma} = \mathcal{G}|_{\gamma}^{-1}\lambda$, where $\dot{\gamma}$ is velocity and λ is momentum. So now 'shortest curves' have parallel momentum, whereas 'straight curves' have auto-parallel velocity. We extend this idea also to sub-Riemannian geometry via a partial connection.

Remark 22 The connection underlies PDE-flows for crossing-preserving geodesic wavefront propagation and denoising in multi-orientation image processing, where in our applications [?] we use

- 1. the 'shortest curves' for tracking in multi-orientation image representations,
- 2. the 'straight curve fits' for locally adaptive frames in PDEs for crossing-preserving image denoising and enhancement.

E.1 A Parameterized Class of Cartan Connections and their Duals

Let G be a Lie group of dimension n. Let $\mathbb{L}_2(G)$ denote the space of square integrable functions on G endowed with the left-invariant Haar measure. Let $T_e(G)$ be the tangent space at unity element e. Let G be a Lie group such that the exponential map exp : $T_e(G) \to G$ is surjective. Then $T_e(G)$ is a Lie Algebra with Lie-Bracket

$$[A,B] = -\frac{d}{dt}\Big|_{t=0} (\gamma^{-B}(\sqrt{t}) \gamma^{-A}(\sqrt{t}) \gamma^{B}(\sqrt{t}) \gamma^{A}(\sqrt{t})) \in T_{e}(G),$$

$$= -\frac{1}{2} \frac{d^{2}}{dt^{2}}\Big|_{t=0} (\gamma^{-B}(t) \gamma^{-A}(t) \gamma^{B}(t) \gamma^{A}(t)),$$
(99)

where $t \mapsto \gamma^X(t) = e^{tX}$ is a differentiable curve in G with $\gamma^X(0) = e$ and $(\gamma^X)'(0) = X$ for X = A, B. For details on Lie brackets see [?].

Let the right-regular representation $\mathcal{R}: G \to BL(\mathbb{L}_2(G))$ be given by $\mathcal{R}_g V(h) = V(hg)$. Then $d\mathcal{R}$ is given by

$$(\mathrm{d}\mathcal{R}(A))V(g) = \lim_{t\downarrow 0} \frac{(\mathcal{R}_{e^{tA}} - I)V(g)}{t}, \text{ for } V \in \mathcal{D}(\mathrm{d}\mathcal{R}(A))$$

and the domain $\mathcal{D}(\mathrm{d}\mathcal{R}(A))$ of this unbounded operator $\mathcal{R}(A)$ is the subset of $\mathbb{L}_2(G)$ for which the above limit exists in \mathbb{L}_2 -sense.

Let $L_g: G \to G$ denote the left-multiplication given by $L_gh = gh$. Let us choose a basis $\{A_1, \ldots, A_n\}$ in $T_e(G)$ and let us define the corresponding vector fields

$$\mathcal{A}_i|_q = (L_g)_* A_i, \text{ for } i = 1, \dots, n.$$

Let us define the corresponding dual basis ('left-invariant co-frame') in $T_q^*(G)$ by

$$\left\langle \left. \omega^{i} \right|_{g}, \mathcal{A}_{j} \right|_{g} \right\rangle = \delta^{i}_{j} \tag{100}$$

with δ_j^i denoting the usual Kronecker delta. Then one has $\mathcal{A}_i = d\mathcal{R}(A_i)$ and the structure constants c_{ij}^k of the Lie algebra relate via

$$[A_i, A_j] = \sum_{k=1}^n c_{i,j}^k A_k \Leftrightarrow [\mathcal{A}_i, \mathcal{A}_j] = \mathcal{A}_i \circ \mathcal{A}_j - \mathcal{A}_j \circ \mathcal{A}_i = \sum_{k=1}^n c_{i,j}^k \mathcal{A}_k.$$
 (101)

If one imposes a left-invariant metric tensor field $g \mapsto \mathcal{G}_g(\cdot, \cdot) : T_g(G) \times T_g(G) \to \mathbb{R}$ to form a Riemannian manifold (M, \mathcal{G}) then there exists a unique *constant* matrix $[g_{ij}] \in \mathbb{R}^{n \times n}$ such that

$$\mathcal{G}_g = \sum_{i,j=1}^n g_{ij} \omega^i \big|_g \otimes \omega^j \big|_g.$$

for all $g \in G$. We restrict ourselves to the diagonal case

$$g_{ij} = \xi_i \,\,\delta_{ij} \tag{102}$$

with $\xi_i > 0$ for i = 1, ..., n and the Kronecker δ_{ij} . As a result for all $g \in G$ the mapping $(L_{g^{-1}})_* : T_g(G) \to T_e(G)$ is unitary. The mapping is known as the Cartan-Maurer form and 'connects' tangent spaces in a left-invariant way. See Fig. 21 where the Maurer-Cartan form is illustrated for the group SE(2) of roto-translations in the plane with group product (50). The associated 'Cartan – connection' [47] is given by:

$$\begin{split} \nabla^{-} &:= \sum_{i,k=1}^{n} \omega^{i} \otimes \left(\mathcal{A}_{i} \circ \omega^{k}(\cdot) \right) \ \mathcal{A}_{k}, \, \text{inducing covariant derivative:} \\ \nabla_{X}^{-} Y &:= \sum_{i,k=1}^{n} \omega^{i}(X) \left(\mathcal{A}_{i}(\omega^{k}(Y)) \right) \mathcal{A}_{k}. \end{split}$$

More precisely, for two arbitrary vector fields $X = \sum_{i=1}^{n} x^{i} A_{i}$ and $Y = \sum_{j=1}^{n} y^{j} A_{j}$, possibly non-left-invariant, (i.e. x^{i} and y^{j} need not be constant) one has

$$abla_X^- Y = \sum_{k=1}^n \left(\sum_{i=1}^n x^i \mathcal{A}_i y^k \right) \mathcal{A}_k.$$

This connection ∇^- has vanishing Christoffel symbols $\Gamma_{ij}^k = 0$ relative to the left-invariant frame (and co-frame) of reference, since

$$\Gamma_{ij}^{k} = \left\langle \omega^{k}, \nabla_{\mathcal{A}_{i}} \mathcal{A}_{j} \right\rangle.$$
(103)

This has big limitations and is not always the right choice for a connection on a Lie group G. Therefore, we consider a more general class of connections on the Lie group G, the so-called Lie-Cartan connections, as we define next. Then in particular we consider a 1-parameter class of Cartan connections. We will call these connections 'Lie-Cartan connections' as they are directly induced by the Lie-bracket.



Fig. 21 The Maurer-Cartan form (in red) 'connects' tangent space $T_g(G)$ to $T_e(G)$ in a left-invariant way. It underlies the Lie-Cartan connection with $\nu = 0$ as can be seen in Lemma 8. Right we depict the Lie group case $SE(2) = \mathbb{R}^2 \rtimes S^1$ and left we show spatial projections $\mathbf{x}(t)$ of the curves $\gamma(t) = (\mathbf{x}(t), \theta(t)) \in SE(2)$.

Definition 12 Per [34, section 5.2], [33], a Cartan (or canonical) connection on a Lie group is a vector bundle connection with the following additional properties:

1. left invariance:

X, Y are left invariant vector fields $\Rightarrow \nabla_X Y$ is a left invariant vector field, (104)

2. for any $\mathbf{a} \in T_e(G)$ the exponential curve and auto-parallel curve coincide:

$$\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0$$
 where $\gamma(t) = \gamma(0) \exp(t\mathbf{a}).$ (105)

We now look at a specific set of Cartan connections that relate to the Lie bracket.

Definition 13 (Lie-Cartan Connection) Consider a Lie group with Lie brackets $[\cdot, \cdot]$ and structure constants $c_{ij}^k \in \mathbb{R}$ s.t. $[\mathcal{A}_i, \mathcal{A}_j] = \sum_k c_{ij}^k \mathcal{A}_k$. Then the Lie-Cartan connection indexed with $\nu \in \mathbb{R}$ equals:

$$\nabla^{[\nu]} := \sum_{i,k=1}^{n} \omega^{i} \otimes \left(\mathcal{A}_{i} \circ \omega^{k}(\cdot)\right) \ \mathcal{A}_{k} + \sum_{i,j,k=1}^{n} \omega^{i} \otimes \omega^{j} \ \nu \ c_{ij}^{k} \ \mathcal{A}_{k}$$
(106)

Remark 23 Left-invariant vector field X can be written as $X = \sum_{i=1}^{d} x^{i} \mathcal{A}_{i}$ with constant coefficients $x^{i} \in \mathbb{R}$. As a result we have that for left-invariant vector fields X, Y the first term vanishes in (106) and we have that $\nabla_{X}^{[\nu]}Y = \nu[X, Y]$.

have that for fert-invariant vector fields X, T the first term vanishes in (100) and we have that $\nabla_X T = \nu[X, T]$.

Remark 24 The Christoffel symbols Γ_{ij}^k (103) relative to the left-invariant moving frame of reference equal $\Gamma_{ij}^k = \nu c_{ij}^k$ and vanish iff $\nu = 0$ and indeed one has for the classical 'minus' Cartan connection $\nabla^- = \nabla^{[0]}$. It is common [34,47,33, ?] to index the Cartan connections in terms of their Torsion $T_{\nabla^{[\nu]}}$ given by

$$T_{\nabla^{[\nu]}}(X,Y) := \nabla^{[\nu]}_X Y - \nabla^{[\nu]}_Y X - [X,Y] = (2\nu - 1)[X,Y],$$

for left-invariant vector fields X, Y, but we prefer to index the Lie-Cartan Connections $\nabla^{[\nu]}$ with the parameter ν arising in the commutator rather than with the parameter $2\nu - 1$ in the torsion of the connection:

$$\nabla^{[\nu]} = \nabla^{2\nu-1}$$
 and thus $\nabla^{[0]} = \nabla^-, \ \nabla^{[1]} = \nabla^+.$

Remark 25 Lie-Cartan connections $\nabla^{[\nu]}$ are clearly connections on a vector bundle (satisfying the standard 4 requirements for Koszul connections).

Furthermore, they are indeed Cartan connections (Def. 12). The first item follows by Remark 23. The second item follows by anti-symmetry of the Christoffel symbols relative to the left-invariant frame of vector fields. We will show this later in (133).

Lemma 8 For arbitrary smooth vector fields X, Y on G we have

$$(\nabla_{\dot{\gamma}}^{[0]}Y)(g) = \lim_{t \to 0} \frac{\left(L_{g(\gamma(t))^{-1}}\right)_* Y(\gamma(t)) - Y(g)}{t}.$$
(107)

For left-invariant vector fields X, Y on G, we have

$$\nu = 0: \quad \nabla^{[0]}Y = 0,$$

$$\nu \in \mathbb{R}: (\nabla^{[\nu]}_{\dot{\gamma}}Y)(g) = \lim_{t \to 0} \frac{(\widetilde{Ad}(\gamma(\nu t))Y)(g) - Y(g)}{t} \ i.e.$$
(108)

$$\nabla^{[\nu]}_XY = \nu[X, Y],$$

with $\gamma(t)$ is an integral curve of left-invariant vector field X with $\gamma(0) = g \in G$, and

$$Ad(q) = (L_g)_* \circ Ad(q) \circ (L_{g^{-1}})_*, \tag{109}$$

with $Ad(g) = (L_g \circ R_{g^{-1}})_* : T_e(G) \to T_e(G)$, where $(\cdot)_*$ denotes the push-forward, so that $(Ad)_* = ad$ with $ad(X_e)(Y_e) = ad$ $[X_e, Y_e]$, and the transferred adjoint representation given by $\widetilde{Ad}(g) = (L_g \circ R_{g^{-1}})_* : T_g(G) \to T_g(G)$ that satisfies

$$(Ad)_*(X_g)(Y_g) = [X_g, Y_g] \text{ for all } g \in G.$$
(110)

Proof. The proof follows by direct computations as we show next. Let X and Y be vector fields on G and γ the integral curve of X with $\gamma(0) = g$. We write $X = \sum_{i=1}^{n} x^{i} A_{i}$ and $Y = \sum_{j=1}^{n} y^{j} \mathcal{A}_{j}$. By the definition of $\nabla^{[0]}$ we have that

$$\begin{split} \left(\nabla_{\dot{\gamma}}^{[0]}Y\right)(g) &= \sum_{i,k=1}^{n} x^{i} \mathcal{A}_{i}|_{g}\left(y^{k}\right) \mathcal{A}_{k}|_{g} = \sum_{k=1}^{n} X|_{g}\left(y^{k}\right) \mathcal{A}_{k}|_{g} \\ &= \sum_{k=1}^{n} \left(\lim_{t \to 0} \frac{y^{k}(\gamma(t)) - y^{k}(g)}{t}\right) \mathcal{A}_{k}|_{g} \\ &= \lim_{t \to 0} \frac{\sum_{k=1}^{n} y^{k}(\gamma(t)) \mathcal{A}_{k}|_{g} - \sum_{k=1}^{n} y^{k}(g) \mathcal{A}_{k}|_{g}}{t} \\ &= \lim_{t \to 0} \frac{\sum_{k=1}^{n} y^{k}(\gamma(t)) \left(L_{g\gamma(t)-1}\right)_{*} \mathcal{A}_{k}|_{\gamma(t)} - Y(g)}{t} \\ &= \lim_{t \to 0} \frac{\left(L_{g\gamma(t)^{-1}}\right)_{*} Y(\gamma(t)) - Y(g)}{t} \;. \end{split}$$

This proves (107).

Now let X, Y be left-invariant. Note that $\nabla^{[0]}Y = 0$ because $\left(L_{g(\gamma(t))^{-1}*}\right)Y(\gamma(t)) = Y(g)$ in (107) regardless of γ . Then the alternative formula (108) for general Lie-Cartan Connection $\nabla^{[\nu]}$ follows, as the structure constants $c_{ij}^k \in \mathbb{R}$ satisfy $\sum_{i,j} c_{ij}^k \mathcal{A}_k = [\mathcal{A}_i, \mathcal{A}_j] \text{ and the Lie bracket is bilinear for left-invariant vector fields, and we find <math>\nabla_X^{[\nu]} Y = \nabla_X^{[0]} Y + \nu[X, Y] = \sum_{i,j} c_{ij}^k \mathcal{A}_k$ $\nu[X,Y]$

For our reformulation in (108) we used $(110):(\widehat{Ad})_*(X_q)(Y_q) = [X_q, Y_q]$ that we show next. By the derivation in [46, Lemma 5.4.2] one has $(Ad)_*(X_e)(Y_e) = [X_e, Y_e]$. Now the Cartan-Maurer form is a Lie algebra isomorphism, and we get (110):(100)

$$\widetilde{\mathrm{Ad}}_{*}(X_{g})(Y_{g}) = \widetilde{\mathrm{Ad}}_{*}((L_{g})_{*}X_{e})((L_{g})_{*}Y_{e}) \stackrel{(109)}{=} (L_{g})_{*}\mathrm{Ad}_{*}(X_{e}, Y_{e}) \\ = [(L_{g})_{*}X_{e}, (L_{g})_{*}Y_{e}] = [X_{g}, Y_{g}].$$

Lemma 9 (Properties of the Lie-Cartan connections) Let X, Y, Z be left-invariant vector fields. The torsion tensor gives

$$T_{\nabla^{[\nu]}}(X,Y) = (2\nu - 1)[X,Y]. \tag{111}$$

The curvature tensor gives

$$R_{\nabla[\nu]}(X,Y)Z = \nu(1-\nu)[Z,[X,Y]].$$
(112)

Relative to the left-invariant frame on G we have the following components:

 $T_{jk}^i = (2\nu - 1)c_{jk}^i$ and $R_{k,ij}^l = \nu(1 - \nu) \sum_{q=1}^n c_{kq}^l c_{ij}^q$. The Lie-Cartan connections satisfy the following identity:

$$(\nabla^{[\nu]}\mathcal{G})(X,Y,Z) = -\nu\left(\mathcal{G}([X,Y],Z) + \mathcal{G}([X,Z],Y)\right).$$
(113)

Proof. Let X, Y, Z be left-invariant vector fields. For all computations, we use the characterisation of Lie-Cartan connections (108) from Lemma 8.

Torsion of $\nabla^{[\nu]}$: We have

$$T_{\nabla^{[\nu]}}(X,Y) = \nabla^{[\nu]}_X Y - \nabla^{[\nu]}_Y X - [X,Y]$$

= $\nu[X,Y] - \nu[Y,X] - [X,Y] = (2\nu - 1)[X,Y].$

Curvature of $\nabla^{[\nu]}$: By the Jacobi identity for Lie brackets we have

$$\begin{split} R_{\nabla^{[\nu]}}(X,Y)Z &= \nabla_X^{[\nu]} \nabla_Y^{[\nu]} Z - \nabla_Y^{[\nu]} \nabla_X^{[\nu]} Z - \nabla_{[X,Y]}^{[\nu]} Z \\ &= \nu^2 \left([X,[Y,Z]] - [Y,[X,Z]] \right) - \nu \ [[X,Y],Z] \\ &= \nu^2 \ [[X,Y],Z] - \nu \ [[X,Y],Z] = \nu(\nu-1)[[X,Y],Z] \end{split}$$

Metric compatibility of $\nabla^{[\nu]}$: We have

$$\begin{split} \nabla^{[\nu]} \mathcal{G}(X,Y,Z) &= X(\mathcal{G}(Y,Z)) - \mathcal{G}\left(Y, \nabla^{[\nu]}_X Z\right) - \mathcal{G}\left(\nabla^{[\nu]}_X Y, Z\right) \\ &= X(\mathcal{G}(Y,Z)) - \nu \ \mathcal{G}(Y,[X,Z]) - \nu \ \mathcal{G}([X,Y],Z) \\ &= -\nu \left(\mathcal{G}(Y,[X,Z]) + \mathcal{G}([X,Y],Z)\right) \ , \end{split}$$

where we note that $X(\mathcal{G}(Y, Z)) = 0$ because \mathcal{G} is also left invariant.

 $Remark \ 26$ (from left-invariant vector fields to general vector fields) The formulas above in Lemma 9 only hold for left-invariant vector fields. For example, the general formula for the torsion is

$$T_{\nabla[\nu]} = (2\nu - 1) \sum_{i,j,k=1}^{n} \omega^{i} \otimes \omega^{j} c_{ij}^{k} \mathcal{A}_{k}.$$
 (114)

so only for left-invariant vector fields we have $T_{\nabla}[\nu](X,Y) = (2\nu - 1)[X,Y]$. It is not a coincidence that vanishing torsion for arbitrary non-commuting vector fields gives $\nu = \frac{1}{2}$, whereas the same conclusion can be drawn from left-invariant non-commuting vector fields. In general, the torsion T_{∇} and curvature R_{∇} of a connection, and the covariant derivative $\nabla \mathcal{G}$ of the metric tensor fields, are *tensor fields*, so e.g.

$$T_{\nabla^{[\nu]}}(f_1X_1 + f_2X_2, g_1Y_1 + g_2Y_2) = f_{1g_1}T_{\nabla^{[\nu]}}(X_1, Y_1) + f_{2g_1}T_{\nabla^{[\nu]}}(X_2, Y_1) + f_{1g_2}T_{\nabla^{[\nu]}}(X_1, Y_2) + f_{2g_2}T_{\nabla^{[\nu]}}(X_2, Y_2)$$
(115)

for all $f_i, g_i \in C^{\infty}(G)$ and all vector fields X_i, Y_i on G, i = 1, 2.

Corollary 5 Let G be a non-commutative Lie group and assume G is not 2-step nilpotent. The Lie-Cartan connection $\nabla^{[\nu]}$ is

- 1. torsion free iff $\nu = \frac{1}{2}$,
- 2. curvature free iff $\nu \in \{0, 1\}$
- 3. metric compatible w.r.t. left-invariant metric \mathcal{G} if $\nu = 0$.

Proof. By Remark 26 we may as well restrict our Lie-Cartan connection $\nabla^{[\nu]}$ to left-invariant vector fields, since T_{∇} , R_{∇} and $\nabla \mathcal{G}$ are all tensor fields. Therefore they have C^{∞} - linearity (such as in (115)) in all of their entries. This C^{∞} linearity allows us to turn arbitrary vector fields into left-invariant vector fields by linear combinations.

The first item now follows by (111) and G being non-commutative (i.e. there exist left-invariant vector fields X, Y s.t. $[X, Y] \neq 0$ as $2\nu - 1 = 0 \Leftrightarrow \nu = \frac{1}{2}$. Note that it also follows by (114). The second item follows by (112) and by the assumptions on G there exist (left-invariant) X, Y, Z s.t. $[Z, [X, Y]] \neq 0$, and thereby $\nu(1 - \nu) \Leftrightarrow \nu \in \{0, 1\}$. The third item follows by (113) as for metric compatibility the covariant derivative of the metric tensor should vanish.

The above properties explain why the choices $\nu \in \{0, \frac{1}{2}, 1\}$ are the most common choices for Cartan connections. Our application (recall Fig. 2) will require torsion and metric-incompatibility of connections on G. Metric incompatibility allows us to distinguish between 'straight curves' (auto-parallel curves with parallel velocity) and 'shortest curves' (distance minimizing geodesics with parallel momentum), as we will see in Theorem 7.

Remark 27 For Lie group SE(2), recall its product (50), the only metric compatible Lie-Cartan connection is the case with $\nu = 0$, since (113) becomes $\nabla^{[\nu]}\mathcal{G}(X,Y,Z) = -1\nu$ if we take $X = \cos\theta\partial_x + \sin\theta\partial_y$, $Y = -\sin\theta\partial_x + \cos\theta\partial_y$, $Z = \partial_\theta$. This corresponds to $X = \mathcal{A}_1$, $Y = \mathcal{A}_2$ and $Z = \mathcal{A}_3$ in Figure 21.

E.1.1 Expressing the Lie-Cartan connection (and its Dual) in Left-invariant Coordinates

Next we express the Lie-Cartan connection explicitly in left-invariant coordinates. That is, the covariant derivative of a field $Y = \sum_{k=1}^{n} y^k \mathcal{A}_k$, along a smooth vector field $X = \sum_{i=1}^{n} x^i \mathcal{A}_i$ is given by (for details see Remark 28 below)

$$\nabla_X^{[\nu]} Y = \sum_{k=1}^n \left(\dot{y}^k + \sum_{i,j=1}^n \nu c_{ij}^k x^i y^j \right) \mathcal{A}_k.$$
(116)

where we use common short notation [46, (3.1.6)] $\dot{y}^k(t) = \frac{d}{dt}y^k(\gamma(t))$ which equals $\dot{y}^k(t) = (Xy^k)(\gamma(t))$ and $x^i = \dot{\gamma}^i(t)$ where $x^i|_{\gamma(t)} := \gamma^i(t) = \langle \omega^i|_{\gamma(t)}, \dot{\gamma}(t) \rangle$ along all flow-lines γ of smooth vector field X. A 'flowline' is a smooth curve γ satisfying $\dot{\gamma}(t) = X_{\gamma(t)}$). With slight abuse of notation we write

$$\nabla_{\dot{\gamma}}^{[\nu]}Y = \sum_{k=1}^{n} \left(\dot{y}^{k} + \sum_{i,j=1}^{n} \nu c_{ij}^{k} \dot{\gamma}^{i} y^{j} \right) \mathcal{A}_{k}.$$
(117)

The corresponding dual connection on the co-tangent bundle is given by

$$\nabla_{\dot{\gamma}}^{[\nu],*}\lambda = \sum_{i=1}^{n} \left(\dot{\lambda}_i + \sum_{k,j=1}^{n} \nu c_{ij}^k \lambda_k \dot{\gamma}^j \right) \omega^i, \tag{118}$$

where $\lambda = \sum_{i=1}^{n} \lambda_i \omega^i \in T^*(G)$. Note that $\langle \nabla_X^{[\nu],*} \lambda, Y \rangle = X \langle \lambda, Y \rangle - \langle \lambda, \nabla_X^{[\nu]} Y \rangle$ and from this formula we see how (118) follows from (117). The fact that both formulas involve a plus sign for the summation reflects that the Christoffel symbols [46] of the connection and dual connection (in the left-invariant frame) are each others inverse:

$$0 = \nu(c_{ji}^k + c_{ij}^k) = \langle \nabla_{\mathcal{A}_i}^{[\nu],*} \omega^k, \mathcal{A}_j \rangle + \langle \omega^k, \nabla_{\mathcal{A}_i}^{[\nu]} \mathcal{A}_j \rangle.$$

Remark 28 Next we explain how (116) follows by the coordinate free formulation (106):

$$\begin{aligned} \nabla_X^{[\nu]}(Y) &:= \nabla^{[\nu]}(X,Y) = \sum_{i,j,k=1}^n \left(\left(\omega^i \otimes \left(\mathcal{A}_i \circ \omega^k(\cdot) \right) \right) (X,Y) + (\omega^i \otimes \omega^j)(X,Y) \ \nu c_{ij}^k \right) \mathcal{A}_k \\ &= \sum_{i,k=1}^n x^i \left(\mathcal{A}_i y^k \right) \mathcal{A}_k + \sum_{i,j,k=1}^n \nu c_{ij}^k x^i y^j \ \mathcal{A}_k, \end{aligned}$$

with $X|_{\gamma} = \sum_{i=1}^{n} x^{i} |A_{i}|_{\gamma(\cdot)} = \dot{\gamma} = \sum_{i=1}^{n} \dot{\gamma}^{i} |A_{i}|_{\gamma(\cdot)}$ and $Y = \sum_{k=1}^{n} y^{k} A_{k}$, and $\dot{y}^{k}(t) = \frac{d}{dt} y^{k}(\gamma(t)) = \sum_{i=1}^{n} x^{i} |(A_{i}y^{k})(\gamma(t)) = X(y^{k})(\gamma(t))$ via the chain-law.

E.1.2 (Partial) Lie-Cartan connections for (Sub)-Riemannian Geometry

The Lie-Cartan connections introduced will be in support of understanding Riemannian geometry when the Lie group is considered as a Riemannian manifold (G, \mathcal{G}) with a left-invariant Riemannian metric tensor field given by

$$\mathcal{G} = \sum_{i,j=1}^{n} g_{ij} \,\omega^i \otimes \omega^j, \tag{119}$$

with g_{ij} constant relative to the left-invariant co-frame ω^i given by (100) s.t. matrix $[g_{ij}] \in \mathbb{R}^{n \times n}$ is symmetric positive definite. Recall we restricted ourselves to the diagonal case (102). The linear map associated to metric tensor field \mathcal{G} is written as

$$\tilde{\mathcal{G}}(X) = \mathcal{G}(X, \cdot) \tag{120}$$

In many applications (robotics [17,64], image analysis [9], cortical vision [68,60]) it is useful to rely on sub-Riemannian geometry [65] where certain direction in the tangent bundle are forbidden as they go with infinite cost. This means that tangents of connecting curves are prescribed to be in a sub-bundle Δ (also known as 'distribution') of the tangent bundle T(G), i.e.

$$\dot{\gamma}(t) \in \Delta_{\gamma(t)} \subset T_{\gamma(t)}(G) \text{ for all } t \in \text{Dom}(\gamma) \subset \mathbb{R}.$$

Typically for a controllable system, Δ and its commutators should fill the full tangent space, in view of Hörmander's theorem. Here we will constrain ourselves to the case that the Lie-algebra is 2-bracket generating $\Delta + [\Delta, \Delta] = T(G)$, as this will be the case for our image analysis application later on.

Remark 29 For instance, let us consider the car in Fig. 2 that needs to move in Lie group SE(2). As the car can proceed forward (by giving gas) and change its orientation (by turning the wheel), it cannot move sidewards. Optimal paths for the car boil down to sub-Riemannian geodesic problems in which the partial Cartan connection $\overline{\nabla}^{[1]}$ will turn out to play a major role, as we will show in the next subsection.

Now let us assume that we label the Lie algebra in such a way that

$$\Delta = \operatorname{Span}\{\mathcal{A}_i\}_{i \in I} \tag{121}$$

for some index set $I \subset \{1, \ldots, n\}$. Let us now consider *partial Cartan connections* on G that will play a major role on sub-Riemannian problems on sub-Riemannian manifolds $(G, \Delta, \mathcal{G}_0)$ with

$$\mathcal{G}_0 = \sum_{i,j \in I} g_{ij} \,\omega^i \otimes \omega^j,\tag{122}$$

as we will see later. Again restrict ourselves to the diagonal case $g_{ij} = \xi_i \delta_{ij}$.

Definition 14 (Partial Lie-Cartan Connection)

Consider a Lie group with Lie brackets $[\cdot, \cdot]$ and structure constants $c_{ij}^k \in \mathbb{R}$ so that $[\mathcal{A}_i, \mathcal{A}_j] = \sum_{k=1}^n c_{ij}^k \mathcal{A}_k$. Consider the distribution given by (121). Then the partial Lie-Cartan connection with parameter $\nu \in \mathbb{R}$ (defined only on vector fields that map into the distribution) is given by:

$$\overline{\nabla}^{[\nu]} := \sum_{i,k \in I} \omega^i \otimes \left(\mathcal{A}_i \circ \omega^k(\cdot) \right) \ \mathcal{A}_k + \sum_{i,j,k \in I} \omega^i \otimes \omega^j \ \nu c_{ij}^k \ \mathcal{A}_k,$$
(123)

So from this definition we deduce that

$$\overline{\nabla}_{\dot{\gamma}}^{[\nu]}Y = \sum_{i,j,k\in I} \left(\dot{y}^k + \nu c_{ij}^k \dot{\gamma}^i y^j\right) \mathcal{A}_k ,$$

$$\overline{\nabla}_X^{[\nu],*}\lambda = \sum_{i=1}^n \left(\dot{\lambda}_i + \nu \sum_{k=1}^n \sum_{j\in I} c_{ij}^k \lambda_k \dot{\gamma}^j\right) \omega^i,$$
(124)

where we highlighted the difference with the full Lie-Cartan connection in red, compare to Def 13, (117), (118). Again X, Y are vector fields and λ a dual vector field and γ is an integral curve of X, and $\dot{y}^k(t) := \frac{d}{dt}y^k(\gamma(t)), \dot{\lambda}_k(t) := \frac{d}{dt}\lambda_k(\gamma(t)).$

E.2 The Special Case of Interest $\nu=1$ and Hamiltonian Flows for the Riemannian Geodesic Problem on G

Let $\mathcal{C}: G \to \mathbb{R}^+$ be an a priori smooth cost (or mobility) for moving through Lie group G that is bounded from below. For the moment it can be considered as constant, but later on in the application sections it will play an important role.

Then the Riemannian metric tensor field ${\mathcal G}$ induces a Riemannian metric on G:

$$d_{\mathcal{G}}(g_{0},g_{1}) := \min_{\substack{\gamma \in \operatorname{Lip}([0,1], G) \\ \gamma(0) = g_{0}, \\ \gamma(1) = g_{1}, \\ \gamma(1) = g_{1}, \\ \gamma(1) = g_{1}, \\ (125)$$

for all $g_0, g_1 \in G$. The sub-Riemannian metric tensor field \mathcal{G}_0 induces a sub-Riemannian metric on G by $d_{\mathcal{G}_0} : G \times G \to \mathbb{R}^+$ on G:

$$d_{\mathcal{G}_{0}}(g_{0},g_{1}) := \min_{\substack{\gamma \in \operatorname{Lip}([0,1],G) \\ \gamma(0) = g_{0} \in G, \\ \gamma(1) = g_{1} \in G, \\ \forall_{t \in [0,1]} : \dot{\gamma}(t) \in \Delta|_{\gamma(t)}}} \int_{0}^{1} \mathcal{C}(\gamma(t)) \sqrt{\mathcal{G}_{0}|_{\gamma(t)} (\dot{\gamma}(t), \dot{\gamma}(t))} \, \mathrm{d}t.$$
(126)

The next theorem motivates the choice $\nu = 1$ for the Lie-Cartan connection $\nabla^{[1]}$, that is underlying the Hamiltonian flow associated to (125). Recall from geometric control theory [65,32] that the Pontryagin Maximum Principle describes the Hamiltonian flow. It allows us to simultaneously analyze all lifted geodesics $(\gamma(\cdot), \lambda(\cdot))$ in the co-tangent bundle $T^*(G)$, where $\lambda(\cdot)$ denotes the momentum along the geodesic.

Theorem 7 In a Riemannian manifold $(G, T(G), \mathcal{G})$, with the tangent bundle T(G) and metric tensor field \mathcal{G} defined in (119) and the induced metric $d_{\mathcal{G}}$ defined in (125), and the Lie-Cartan connection $\nabla^{[\nu]}$ for $\nu = 1$ defined in (106), we have the following relations for "straight" curves:

$$\Rightarrow \qquad \begin{array}{ccc} \gamma \text{ is a } \nabla^{[1]-\text{straight curve}} & \Leftrightarrow & \gamma \text{ is an exponential curve} \\ \Leftrightarrow & \nabla^{[1]}_{\dot{\gamma}}\dot{\gamma} = 0 & \Leftrightarrow & \gamma \text{ has } \nabla^{[1]}\text{-auto parallel velocity} \quad , \end{array}$$
(127)

and the following for "shortest" curves (minimizers in (125)), recall also (120):

 γ is a shortest curve $\Leftrightarrow \gamma$ is a minimizing curve in $d_{\mathcal{G}}$

$$\Rightarrow \begin{cases} \nabla_{\dot{\gamma}}^{[1],*} \lambda = 0 \\ \dot{\gamma} = \tilde{\mathcal{G}}^{-1} \lambda \end{cases} \Leftrightarrow \gamma \text{ has } \nabla^{[1],*}\text{-parallel momentum} \qquad (128)$$

In a sub-Riemannian (SR) manifold (SE(2), Δ , \mathcal{G}_0) with sub-bundle Δ defined in (123), the sub-Riemannian metric tensor \mathcal{G}_0 (122) and distance (126), and partial Cartan connection (123), we have the following relations for "straight" curves:

$$\begin{array}{ll} \gamma \ is \ a \ \overline{\nabla}^{[1]} \text{-straight curve} & \Leftrightarrow & \gamma \ is \ a \ horizontal \ exponential \ curve} \\ \Leftrightarrow & \overline{\nabla}^{[1]}_{\dot{\gamma}} \dot{\gamma} = 0 & \Leftrightarrow & \gamma \ has \ \overline{\nabla}^{[1]} \text{-auto parallel velocity} &, \end{array}$$
(129)

and the following for "shortest" curves (minimizers in (126)):

14.1

 γ_0 is a shortest curve $\Leftrightarrow \gamma_0$ is a minimizing curve in $d_{\mathcal{G}_0}$

$$\Rightarrow \begin{cases} \overline{\nabla}_{\dot{\gamma}_{0}}^{[1],*} \lambda = 0 \\ \dot{\gamma}_{0} = \tilde{\mathcal{G}}_{0}^{-1} \mathbb{P}_{\Delta}^{*} \lambda \end{cases} \Leftrightarrow \gamma_{0} \text{ has } \overline{\nabla}^{[1],*} \text{-parallel momentum} \quad , \qquad (130)$$

in which \mathbb{P}^*_{Δ} is the projection

$$\mathbb{P}^*_{\Delta}\left(\sum_{i=1}^n \lambda_i \omega^i\right) = \sum_{i \in I} \lambda_i \omega^i$$

For the reverse in (128) and (130); for a minimizing curve between $g_1 = \gamma(0)$ and $g_2 = \gamma(t)$ one must have $0 \le t \le t_{cut} = \min\{t_{conj}(\lambda(0)), t_{Max,1}(\lambda(0)\}, cf.$ [65, ?] for details. They are found by steepest descent:

$$\gamma(t) = \gamma(0) + \int_0^t \operatorname{grad}_{\mathcal{G}} W(\gamma(s)) \, \mathrm{d}s, \tag{131}$$

on distance maps $W(g) = d_{\mathcal{G}}(g, e)$ that are viscosity solutions of eikonal PDE:

$$\begin{cases} \|grad_{\mathcal{G}}W(g)\| = \sqrt{\mathcal{G}|_g} \left(grad_{\mathcal{G}}W(g), grad_{\mathcal{G}}W(g)\right) = 1, \\ W(e) = 0, \end{cases}$$
(132)

with (metric-intrinsic) gradient $\operatorname{grad}_{\mathcal{G}}W(g) = \tilde{\mathcal{G}}^{-1} \mathrm{d}W(g)$, as this only gives global minimizing curves, even in the SR setting $\mathcal{G} \to \mathcal{G}_0$.

Proof. First we address the 'shortest curves' part of the theorem. The items (128) and (130) follow by the Pontryagin Maximum Principle [65] and Theorem 8 in the subsequent subsection. Theorem 8 proves the actual fundamental relation between the (partial) Lie-Cartan connection $\nabla^{[1]}$ to the Hamiltonian flow, for the (sub)-Riemannian setting. Here we stress that PMP provides only *local* optimality of geodesics.

The geodesics are found by the exponential map that integrates the Hamiltonian flow $(\lambda(0), t) \mapsto (\gamma(t), \lambda(t)) = e^{t\mathfrak{h}}(\lambda_0)$. Optimality of $t \mapsto \gamma(t)$, requires t to be less than the cut-time. Such a cut-time is the minimum of the conjugate time $t_{conj}(\lambda(0)) \in \mathbb{R} \cup \{\infty\}$ where local optimality is lost, and the first Maxwell time $t_{Max,1}$, where two equidistant geodesics meet for the first time and where global optimality is lost. Now $t \leq t_{cut}(\lambda(0))$ is guaranteed by steepest descent (131) on the distance maps W which are obtained as viscosity solutions [35, ?] to the eikonal PDE. This is well-known for the Riemannian case [49,35], but also applies to the sub-Riemannian²⁰ case [50,9] and holds even in more general Finsler geometrical settings [26].

Secondly, regarding the 'straight curves' (127) one has (by (117)) and anti-symmetry of the structure constants (101) that:

$$\nabla_{\dot{\gamma}}^{[1]} \dot{\gamma} = 0 \Leftrightarrow \forall_{k \in \{1,...,n\}} : \ddot{\gamma}^{k} - \sum_{i,j=1}^{n} c_{ij}^{k} \dot{\gamma}^{i} \dot{\gamma}^{j} = \ddot{\gamma}^{k} = 0$$

$$\Leftrightarrow \forall_{k \in \{1,...,n\}} : \langle \omega^{k} |_{\gamma}, \dot{\gamma} \rangle =: \dot{\gamma}^{k} = c^{k} = \text{ constant}$$

$$\Leftrightarrow \gamma(t) = \gamma(0) e^{t} \sum_{k=1}^{n} c^{k} A_{k}, \qquad (133)$$

with $\ddot{\gamma}^k(t) = \frac{d}{dt}\dot{\gamma}^k(t)$. Note that the first, third and fourth statement in (127) are just tautological, so that (133) proves the remaining second equivalence. The SR-case (129) follows similarly by (124) taking into account the restriction to (121) via projection P^*_{Δ} which means we constrain the summations to index set I and set $\dot{\gamma}^i = 0$ if $i \notin I$.

 $^{^{20}}$ For an intuitive illustration that in the viscosity solutions of the PDEs non-optimal wavefronts are cut (at the 1st Maxwell set) in the sub-Riemannian setting see [9, Fig.3].

E.2.1 Hamiltonian Flow of the left-invariant (sub)Riemannian Geodesic Problem on Lie group G

The family of all geodesics $\gamma(t)$ augmented to $\mathbf{v}(t) = (\gamma(t), \lambda(t))$ with their momentum representation $\lambda(t) = \sum_{i=1}^{n} \lambda_i(t) \omega^i |_{\gamma(t)}$ along the geodesic are flow lines of a so-called 'Hamiltonian flow' on the cotangent bundle $T^*(G)$. Controlling the Hamiltonian flow means controlling the complete family of all geodesics (minimal distance curves) together. Next we explain the concept of Hamiltonian flows, and derive the canonical Hamiltonian equations associated to the left-invariant Riemannian and sub-Riemannian problem of interest.

To a Hamiltonian function \mathfrak{h}

$$T^*(G) \ni (g,\lambda) \mapsto \mathfrak{h}(g,\lambda) \in \mathbb{R}^+$$

one associates a Hamiltonian vector field $\vec{\mathbf{h}}$ (or 'Hamiltonian lift') in the co-tangent bundle. It is determined via the fundamental symplectic form that is given by

$$\sigma = \sum_{i=1}^{n} \omega^i \wedge \mathrm{d}\lambda_i,$$

where $d\lambda_i$ is defined by $\langle d\lambda_i, \partial_{\lambda_j} \rangle = \delta_j^i$, by means of

$$\forall_{V=(\dot{g},\dot{\lambda})\in T_g(G)\times T(T_g^*(G))} : \sigma(\overrightarrow{\mathfrak{h}}(g,\lambda),V) = \langle \mathrm{d}\mathfrak{h}(g,\lambda),V\rangle.$$
(134)

Remark 30 (background on Hamiltonian lifts)

A direct consequence of (134) is that along the flowlines of the Hamiltonian flow the Hamiltonian is preserved (take $Y = \vec{\mathfrak{h}}$) and

$$\frac{d}{dt}\mathfrak{h}(\mathbf{v}(t)) = \sigma(\overrightarrow{\mathfrak{h}}(\mathbf{v}(t)), \overrightarrow{\mathfrak{h}}(\mathbf{v}(t))) = 0, \text{ with } \mathbf{v}(t) = (\gamma(t), \lambda(t)),$$

Furthermore the lifting of a Hamiltionian \mathfrak{h} to its Hamiltonian lift $\overrightarrow{\mathfrak{h}}$ is a Lie algebra isomorphism [65]:

$$\overline{\{h_1, h_2\}} = [\overrightarrow{h}_1, \overrightarrow{h}_2] \tag{135}$$

where $\{\cdot, \cdot\}$ denote Poisson brackets and $[\cdot, \cdot]$ denotes the usual Lie bracket of vector fields. In the left-invariant (co)-frames Poisson brackets are expressed as

$$\{g,f\} = \sum_{i=1}^{n} (\mathcal{A}_{i}f) \ \frac{\partial g}{\partial \lambda_{i}} - \frac{\partial f}{\partial \lambda_{i}} \ (\mathcal{A}_{i}g), \tag{136}$$

but this may also be expressed in canonical coordinates [65, eq.11.21].

Remark 31 (simple example of Hamiltonian lifts on $T^*(\mathbb{R})$ We set $\vec{\sigma} = dx \wedge d\lambda$. We set $\vec{\mathfrak{h}} = h^1 \partial_x + h^2 \partial_\lambda$. Then from (134) one can deduce the following standard canonical equations:

$$\vec{\mathbf{h}} = \frac{\partial \mathbf{h}}{\partial \lambda} \partial_x - \frac{\partial \mathbf{h}}{\partial x} \partial_\lambda + \Rightarrow \dot{x} \partial_x + \dot{\lambda} \partial_\lambda = \dot{\mathbf{v}} = \vec{\mathbf{h}} (\mathbf{v})$$
$$\Leftrightarrow \begin{cases} \dot{x} = \frac{\partial \mathbf{h}}{\partial \lambda} \text{ (horizontal part)}, \\ \dot{\lambda} = -\frac{\partial \mathbf{h}}{\partial x} \text{ (vertical part)}. \end{cases}$$

Generalizing the above example, the next theorem provides the Hamiltonian flows for the left-invariant Riemannian and sub-Riemannian problem on G.

Theorem 8 The Hamiltonian on Riemannian manifold (G, \mathcal{G}) , with left-invariant metric tensor field \mathcal{G} given by (119), equals

$$\mathfrak{h} = \frac{1}{2} \sum_{i=1}^{n} \lambda^i \lambda_i = \frac{1}{2} \sum_{i,j=1}^{n} g^{ij} \lambda_i \lambda_j \tag{137}$$

and the corresponding Hamiltonian flow (generated by the Hamiltonian vector field $\overrightarrow{\mathfrak{h}}$) can be written as (recall the definition of linear map $\tilde{\mathcal{G}}$ (120))

$$\dot{\mathbf{v}} = \overrightarrow{\mathfrak{h}}(\mathbf{v}) \Leftrightarrow \begin{cases} \tilde{\mathcal{G}}^{-1}\lambda = \dot{\gamma} & (\text{horizontal part}) \\ \nabla_{\dot{\gamma}}^{[1],*}\lambda = 0 & (\text{vertical part}) \end{cases}$$

$$\Leftrightarrow \begin{cases} \dot{\gamma}^{i} = u^{i} = \lambda^{i} := \sum_{j=1}^{n} g^{ij}\lambda_{j} & (\text{horizontal part}) \\ \dot{\lambda}_{i} = \{\mathfrak{h}, \lambda_{i}\} = -\sum_{j,k=1}^{n} c_{ij}^{k}\lambda_{k}u^{j} & (\text{vertical part}) \end{cases}$$
(138)

with velocity controls $u^i := \dot{\gamma}^i = \langle \omega^i |_{\gamma(\cdot)}, \dot{\gamma} \rangle$ and $\mathbf{v}(t) = (\gamma(t), \lambda(t))$ a curve in the co-tangent bundle $T^*(G)$ where the geodesic $\gamma(t) \in G$ and the momentum along the geodesic $\lambda(t) \in T^*_{\gamma(t)}(G)$, and with $\{\cdot, \cdot\}$ denoting Poisson brackets, recall (136).

The Hamiltonian on sub-Riemannian manifold $(G, \Delta = span\{A_j\}_{j \in I}, \mathcal{G}_0)$ equals

$$\mathfrak{h} = \frac{1}{2} \sum_{i \in I} \lambda^i \lambda_i = \frac{1}{2} \sum_{i,j \in I} g^{ij} \lambda_j \lambda_i \tag{139}$$

and the Hamiltonian flow can be written as

$$\dot{\mathbf{v}} = \overrightarrow{\mathfrak{h}}(\mathbf{v}) \Leftrightarrow \begin{cases} \widetilde{\mathcal{G}}_{0}^{-1} P_{\Delta^{*}} \lambda = \dot{\gamma} \text{ (horizontal part)} \\ \overline{\nabla}_{\gamma}^{[1],*} \lambda = 0 \text{ (vertical part)} \end{cases} \Leftrightarrow \\ \begin{cases} \dot{\gamma}^{i} = u^{i} = \lambda^{i} \text{ for } i \in I \text{ and } u^{j} = 0 \text{ if } j \notin I \text{ (horizontal part)} \\ \dot{\lambda}_{i} = \{\mathfrak{h}, \lambda_{i}\} = -\sum_{k=1}^{n} \sum_{i \in U} c_{ij}^{k} \lambda_{k} u^{j} \text{ (vertical part)} \end{cases}$$
(140)

where P_{Δ^*} denotes the projection onto the dual Δ^* of Δ , as given in Theorem 7.

Proof. The results (138) and (140) follow from standard application of the Pontryagin Maximum Principle (PMP [65]) to respectively the Riemannian and sub-Riemannian geodesic problem. First of all we note that regarding the Hamiltonian in the Riemannian case (137) we have that it is computed by applying the Fenchel transform on the integrand of the action functional (i.e. squared Lagrangian)

$$\begin{split} \mathfrak{h}(g,\lambda) &= \sup_{\dot{\gamma}\in T_g(G)} \left\{ \langle \lambda, \dot{\gamma} \rangle - \mathcal{L}^2(g,\dot{\gamma}) \right\} \text{ with } \lambda = \sum_{i=1}^n \lambda_i \, \omega^i \Big|_g \in T_g^*(G), \\ \text{hence we get the Hamiltonian } \mathfrak{h} : T^*(G) \to \mathbb{R}^+ \text{ given by} \\ \mathfrak{h} &= \max_{(v^1,\dots,v^n)} \left\{ \sum_{i=1}^n \lambda_i v^i - \frac{1}{2} \sum_{i,j=1}^n v^i v^j g_{ij} \right\} = \frac{1}{2} \sum_{i,j=1}^n \lambda^i g_{ij} \lambda^j = \frac{1}{2} \sum_{i=1}^n \lambda^i \lambda_i, \end{split}$$
(141)

with $\lambda^i = \sum_{j=1}^n g^{ij}\lambda_j$. The Hamiltonian in the SR-case (139) comes with the constraint $\dot{\gamma} \in \Delta$ (i.e. $\dot{\gamma}^i = 0$ if $i \notin I$) and then with a similar type of reasoning above (but then with $v^i = 0$ if $i \notin I$) we get $\mathfrak{h} = \frac{1}{2} \sum_{i \in I} \lambda^i \lambda_i$ with $\lambda^i = \sum_{j \in I} g^{ij}\lambda_j$, and we find the 'ortropole controls' [65]: $u^i = u^i = \lambda^i$

find the 'extremal controls' [65]: $v_{max}^i = u^i = \lambda^i$. Note that (138) and (140) are of the form $a \Leftrightarrow b \Leftrightarrow c$. We first comment on $a \Leftrightarrow c$ and then show $b \Leftrightarrow c$.

 $a \Leftrightarrow c$ follows by direct computation as we show next. By computing We have the following relation in Poisson brackets:

$$[\mathcal{A}_i, \mathcal{A}_j] = \sum_{k=1}^n c_{ij}^k \mathcal{A}_k \Leftrightarrow \{\lambda_i, \lambda_j\} = \mathcal{A}_i \lambda_j - \mathcal{A}_j \lambda_i = \sum_{k=1}^n c_{ij}^k \lambda_k,$$

as the 'conjugate momentum mapping' gives rise to a Lie-algebra morphism, see [65, p.164]. Thereby (via (136), (141)) we find (with Liouville's theorem and $c_{ij}^k = -c_{ji}^k$):

$$\begin{aligned} \ddot{\gamma}^i &= \dot{u}^i = \{\mathfrak{h}, u^i\} = \dot{\lambda}^i \Rightarrow u^i = \lambda^i, \\ \dot{\lambda}_i &= \{\mathfrak{h}, \lambda_i\} = \sum_{j \in J} \frac{2}{2} \{\lambda_j, \lambda_i\} \lambda^j = -\sum_{k=1}^n \sum_{j \in J} c^k_{ij} \lambda_k u^j, \end{aligned}$$
(142)

which hold for i = 1, ..., n in the Riemannian case and for $i \in I$ in the sub-Riemannian case. In the above expression one must set $J = \{1, ..., n\}$ in the Riemannian case, and J = I in the sub-Riemannian case.

 $b \Leftrightarrow c$ follows by (142), and the expression (118) for the Lie-Cartan connection (with $\nu = 1$), respectively expression (124) for the partial Lie-Cartan connection (again with $\nu = 1$) expressed in left-invariant coordinates.

F Table of Notations

Symbol	Explanation	Reference
\mathbb{R}^d . x	Position space with vectors $\mathbf{x} = (x^1, \dots, x^d)^T$.	Sect. 1.0.1. Sect. 1.1
$\mathbb{S}^{d-1},\mathbf{n}$	Angular space, $\mathbb{S}^{d-1} = \{\mathbf{n} \in \mathbb{R}^d \mid \mathbf{n} = 1\}.$	Sect. 1.0.1, Sect. 1.1,
a	Reference axis. For $d = 2$, $\mathbf{a} = (1,0)^T$, for $d = 3$,	Eq. (10), Remark 3
	$\mathbf{a} = (0, 0, 1)^T.$	
\mathbb{M}, \mathbf{p}	Manifold $\mathbb{M} = \mathbb{R}^d \times \mathbb{S}^{d-1}$, with $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}$	Sect. 1.0.1
$T(\mathbb{M}), T^*(\mathbb{M}), T_{\mathbf{p}}(\mathbb{M})$	Tangent bundle $T(\mathbb{M}) = \{(\mathbf{p}, \dot{\mathbf{p}}) \mid \mathbf{p} \in \mathbb{M}, \dot{\mathbf{p}} \in T_{\mathbf{p}}(\mathbb{M})\},\$	Sect. 1.0.1, Sect. 2.5, Sect. 1.1,
	and cotangent bundle $T^*(\mathbb{M})$, with tangent space	
Γ.γ	Space $\Gamma = \text{Lip}([0, 1], \mathbb{M})$ of admissible curves, with	Eq. (1)
- , ,	$t \mapsto \gamma(t) = (\mathbf{x}(t), \mathbf{n}(t)).$	-4. (1),
$\mathcal{F}, \mathcal{F}^*, \mathcal{F}_0, \mathcal{F}_0^+, \mathcal{F}_\varepsilon, \mathcal{F}_\varepsilon^+,$	Finsler function \mathcal{F} defined on \mathbb{M} , its dual \mathcal{F}^* :	Sect. 1.0.1, Eqs. (2) , (3) , (5) , (6) , (17) ,
$(\mathcal{F}_{\varepsilon})^*, (\mathcal{F}_{\varepsilon}^+)^*$	$T^*(\mathbb{M}) \to \mathbb{R}$ the models with and without reverse gear	(18), Prop. 1,
	$\mathcal{F}_0, \mathcal{F}_0^+$, their approximations $\mathcal{F}_{\varepsilon}, \mathcal{F}_{\varepsilon}^+$ and their duals.	
$d_{\mathcal{F}}, U$	Distance function $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q})$ for $\mathbf{p}, \mathbf{q} \in \mathbb{M}$, and $U(\mathbf{p}) =$	Eqs. $(1), (4), \ldots$
	$d_{\mathcal{F}}(\mathbf{p}_S, \mathbf{p})$ for a fixed source $\mathbf{p}_S \in \mathbb{M}$	\mathbf{P}_{1} (16) \mathbf{P}_{2} (
ε	Anisotropy parameter in the metric, $\varepsilon = 0$ corresponds to the sub Diamannian manifold area	Eqs. (17) , (18) , Fig. 6,
~	to the sub-ritemannan mannoid case. We write $\dot{\mathbf{x}} \propto \mathbf{n}$ when $\dot{\mathbf{x}} = \lambda \mathbf{n}$ for some $\lambda \in \mathbb{R}$	$E_{ac}(2)(3)$ Sect 2.3 Thm 1
$C_1 C_2 E$	External cost $\mathcal{C} : \mathbb{M} \to \mathbb{R}^+$ analytic and strictly	Sect. 1.1
$[\mathcal{C}_1, \mathcal{C}_2, \zeta]$	bounded from below, and $\xi > 0$ to balance the cost	5000. 1.1,
	of spatial motion relative to angular motion, when we	
	choose $C_1 = \xi C_2$	
$\mathfrak{B}, \mathcal{B}_{\mathcal{F}}$	Set of controls \mathfrak{B} , and the set of admissible controls	Fig. 1, Eq. (7), Appendix A
	$\mathcal{B}_{\mathcal{F}}(\mathbf{p}) = \{ \dot{\mathbf{p}} \in T_{\mathbf{p}}(\mathbb{M}) \mathcal{F}(\mathbf{p}, \dot{\mathbf{p}}) \le 1 \}$	
а	Reference axis. For $d = 2$, $\mathbf{a} = (1,0)^T$, for $d = 3$,	Eq. (10) , Remark 3
	$\mathbf{a} = (0, 0, 1)^{T}$.	\mathbf{F} (1c)
$(\cdot)_{-}, (\cdot)_{+}$	$(\cdot)_{-} = \min(\cdot, 0), (\cdot)_{+} = \max(\cdot, 0)$	Eqs. (16),
n, m, m	Subset $\mathfrak{R} \in \mathbb{M}$ of end-points that are reached by cusp- less geodesics, the closure \mathfrak{R} and its complement \mathfrak{R}^c	Der. 4, Remark 3, 1 nm. 3, Sect. 5.
\mathcal{A}_i, ω^i	Left-invariant frame \mathcal{A}_i and the dual frame ω^i .	Sect. 5, Eqs. (65), (66), Remark 14.
$u^i, \hat{p}_i, ilde{u}$	Controls (velocity components) u^i , momentum com-	Def. $2, (58), \ldots$
~	ponents \hat{p}_i and the special spatial \tilde{u}	
$\mathcal{G}_{\mathbf{p},\varepsilon}, \mathcal{G}_{\mathbf{p},\varepsilon}$	Metric tensors $\mathcal{G}_{\mathbf{p},\varepsilon}, \mathcal{G}_{\mathbf{p},\varepsilon} : T_{\mathbf{p}}(\mathbb{M}) \times T_{\mathbf{p}}(\mathbb{M}) \to \mathbb{R}^+$	Eq. (26), (27)
$\nabla, \mathcal{G}_{\mathbf{p},\varepsilon}^{-1} \mathrm{d}, \mathcal{G}_{\mathbf{p},\varepsilon}^{-1} \mathrm{d}$	Standard gradient $\nabla = (\nabla_{\mathbb{R}^d}, \nabla_{\mathbb{S}^{d-1}})$, the intrinsic	Cor. 1, Thm. 4, Remark 32,
	gradient $\mathcal{G}_{\mathbf{p},\varepsilon}^{-1}\mathbf{d}$ of the manifold $(\mathbb{M}_+, d_{\mathcal{F}_{\varepsilon}})$ and $\tilde{\mathcal{G}}_{\mathbf{p},\varepsilon}^{-1}\mathbf{d}$	
	the intrinsic gradient of $(\mathbb{M}_{-}, d_{\mathcal{F}_{\varepsilon}^{+}})$	
$F_{M,\mathbf{w}}, F^*_{\hat{M},\hat{\mathbf{w}}}$	Norm $F_{M,\mathbf{w}}$: $\mathbb{R}^n \to \mathbb{R}^+$ and dual norm $F^*_{\hat{M},\hat{w}}$:	Lemma 4
	$(\mathbb{R}^n)^* \to \mathbb{R}^+$	
X, X	Discrete subset X of \mathbb{R}^d , and image support $\mathbb{X} \subset \mathbb{M}$.	Sect. 7, Appendix A
$D_{\mathbf{n}}^{\varepsilon}$	Symmetric positive definite matrix $D_{\mathbf{n}}^{\varepsilon} = \mathbf{n} \otimes \mathbf{n} + \varepsilon^{2} (\mathrm{Id} - \mathbf{n} \otimes \mathbf{n})$	Eq. (28), (29),
$\mathbb{M}_+, \mathbb{M}, \partial \mathbb{M}_\pm$	$\mathbb{M}_{+} = \{ \mathbf{p} \in \mathbb{M} \mid \langle \mathrm{d}U^{+}(\mathbf{p}), \mathbf{n} \rangle > 0 \}, \ \mathbb{M}_{-} = \{ \mathbf{p} \in \mathbb{N} \}$	Cor. 1, Thm. 4
	$\mathbb{M} \mid \langle \mathrm{d}U^+(\mathbf{p}), \mathbf{n} \rangle < 0 \}$ and their boundary	
$N_x, N_y, N_z, \overline{N_o}$	Resolution in spatial/angular coordinates	Sect. 8
σ, p	Parameters $\sigma > 0, p \in \mathbb{N}$ of the cost function \mathcal{C}	Sect. 8

Table 1 Symbols used throughout the paper, their brief explanation and references to where they appear, or where they are defined/first appear. The dots in the Reference column indicate that they are used frequently.

References

- 1. A.A. Agrachev, D. Barilari, and U. Boscain. Introduction to Riemannian and Sub-Riemannian geometry. https://webusers.imj-prg.fr/~davide.barilari/2016-11-21-ABB.pdf, 2016.
- 2. A.A. Agrachev and Y.L. Sachkov. Control Theory from the Geometric Viewpoint, volume 87 of Encyclopaedia of Mathematical Sciences. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.
 C. Arzela. Sulle funzioni di linee. Mem. Accad. Sci. lst. Bologna Cl. Sci. Fis. Mat., 5(5):55-74, 1895.
- 4. G. Ascoli. Le curve limiti di una varieta data di curve. Atti della R. Accad. Dei Lincei Memorie della Cl. Sci. Fis. Mat. Nat., 18(3):521-586, 1883.
- 5. L. Astola and L.M.J. Florack. Finsler geometry on higher order tensor fields and applications to high angular resolution diffusion imaging. International Journal of Computer Vision, 92(3):325-336, 2011.
- 6. D. Bao, S.-S. Chern, and Z. Shen. An Introduction to Riemann-Finsler Geometry. Graduate Texts in Mathematics. Springer New York, 2000.

- 7. M Bardi and I Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Birkhauser, 1997.
- E. Bekkers, Retinal Image Analysis using Sub-Riemannian Geometry in SE(2). PhD thesis, Eindhoven University of Technology, 2017. cum laude (≤ 5% best at TU/e), URL: https://pure.tue.nl/ws/files/52750592/20170123 _ Bekkers.pdf.
- 9. E.J. Bekkers, R. Duits, A. Mashtakov, and G. Sanguinetti. A PDE Approach to Data-Driven Sub-Riemannian Geodesics in SE(2). SIAM J. Imaging Sci., 8(4):2740–2770, 2015.
- 10. E.J. Bekkers, R. Duits, A. Mashtakov, and Y. L. Sachkov. Vessel Tracking via Sub-Riemannian Geodesics on $\mathbb{R}^2 \times P^1$. available on arxiv, accepted for publication in proc. Geometric Science of Information GSI 2017.
- U. Boscain, G. Charlot, and F. Rossi. Existence of planar curves minimizing length and curvature. Proc. Steklov Inst. Math., 270(1):43–56, October 2010.
- U. Boscain, R. Duits, F. Rossi, and Y. Sachkov. Curve cuspless reconstruction via sub-Riemannian geometry. ESAIM: Control, Optimisation and Calculus of Variations, 20(3):748–770, July 2014.
- 13. E. Caruyer, A. Daducci, M. Descoteaux, J-C. Houde, J-P. Thiran, and R. Verma. Phantomas: a flexible software library to simulate diffusion MR phantoms. In *ISMRM*, Milan, Italy, May 2014.
- 14. D. Chen. New Minimal Paths Models for Tubular Structure Extraction and Image Segmentation. PhD thesis, University Paris-Dauphine, 2016.
- D. Chen, J-M. Mirebeau, and L. D. Cohen. Vessel tree extraction using radius-lifted keypoints searching scheme and anisotropic fast marching method. Journal of Algorithms & Computational Technology, 10(4):224–234, December 2016.
- D. Chen, J-M. Mirebeau, and L.D. Cohen. Global Minimum for a Finsler Elastica Minimal Path Approach. Int J Comput Vis, pages 1–26, December 2016.
- 17. G. S. Chirikjian and A. B. Kyatkin, Engineering applications of noncommutitative harmonic analysis: with emphasis on rotation and motion groups. Boca Raton CRC Press, 2001.
- G. Citti and A. Sarti. A Cortical Based Model of Perceptual Completion in the Roto-Translation Space. J Math Imaging Vis, 24(3):307–326, February 2006.
- T. G. Close, J-D. Tournier, F. Calamante, L. A. Johnston, I. Mareels, and A. Connelly. A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. *NeuroImage*, 47(4):1288–1300, October 2009.
- M. G. Crandall, H. Ishii, and P-L Lions. User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc., 27(1):1–67, 1992.
- M. G. Crandall and P-L Lions. Viscosity solutions of Hamilton-Jacobi equations. Trans. Amer. Math. Soc., 277(1):1– 42, 1983.
- M. Descoteaux, R. Deriche, T. R. Knosche, and A. Anwander. Deterministic and Probabilistic Tractography Based on Complex Fibre Orientation Distributions. *IEEE Transactions on Medical Imaging*, 28(2):269–286, February 2009.
 L. E. Dubins. On Curves of Minimal Length with a Constraint on Average Curvature, and with Prescribed Initial
- and Terminal Positions and Tangents. American Journal of Mathematics, 79(3):497–516, 1957.
- 24. R. Duits, U. Boscain, F. Rossi, and Y. Sachkov. Association Fields via Cuspless Sub-Riemannian Geodesics in SE(2). J Math Imaging Vis, 49(2):384–417, December 2013.
- R. Duits, M. Felsberg, G. Granlund, and B. ter Haar Romeny. Image Analysis and Reconstruction using a Wavelet Transform Constructed from a Reducible Representation of the Euclidean Motion Group. Int J Comput Vision, 72(1):79–102, July 2006.
- 26. R. Duits, S. Meesters, J. Mirebeau, J. Portegies, Optimal paths for variants of the 2d and 3d Reeds-Shepp car with applications in image analysis, (arXiv:1612.06137), submitted to JMIV 2017.
- R. Duits, A. Ghosh, T. Dela Haije, and Y. Sachkov. Cuspless Sub-Riemannian Geodesics within the Euclidean Motion Group SE(d). In *Neuromathematics of Vision*, Lecture Notes in Morphogenesis, pages 173–215. 2014. DOI: 10.1007/978-3-642-34444-2_5.
- R. Duits, A. Ghosh, T. C. J. Dela Haije, and A. Mashtakov. On Sub-Riemannian Geodesics in SE(3) Whose Spatial Projections do not Have Cusps. J Dyn Control Syst, 22(4):771–805, October 2016.
- R. Duits, B.J. Janssen, A. Becciu, and H.van Assen. A Variational Approach to Cardiac Motion Estimation based on Covariant Derivatives and Multi-scale Helmholtz Decomposition. *Quarterly of Applied Mathematics, AMS*, (71):1–36, 2012.
- 30. T. Georgiev. Relighting, retinex theory, and perceived gradients. Proceedings of Mirage, INRIA Rocquencourt, 2005.
- T. Batard and M.Bertalmio. On Covariant Derivatives and their Applications to Image Regularization. SIAM Journal on Imaging Sciences 7(4):2393-2422, 2014.
- 32. A. Agrachev, D. Barilari, and U. Boscain, A Comprehensive Introduction to Sub-Riemannian Geometry. CUP Cambridge studies in advanced mathematics, 2020.
- 33. Élie Cartan. Sur une classe remarquable d'espaces de Riemann. Bulletin de la Société Mathématique de France 54:214–264, 1926.
- A. Cogliati and P. Mastrola. Cartan, Schouten and the search for connection. Historia Mathematica 45(1):39-74, 2018.
- 35. M. Crandall and P.-L. Lions, "Viscosity solutions of hamilton-jacobi equations," Trans. A.M.S., vol. 277, no. 1, pp. 1–42, 1983.
- 36. R.Duits, E.Franken. Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores. part II: Non-linear left-invariant diffusions on invertible orientation score. Q APPL MATH, 68, 293–331, 2010.
- 37. R. Duits, M. H. J. Janssen, J. Hannink, and G. R. Sanguinetti. Locally Adaptive Frames in the Roto-Translation Group and Their Applications in Medical Imaging. J Math Imaging Vis, 56(3):367–402, March 2016.
- R. Duits, S.P.L. Meesters, J-M. Mirebeau, and J.M. Portegies. Optimal Paths for Variants of the 2d and 3d Reeds-Shepp Car with Applications in Image Analysis. arXiv:1612.06137 [math], December 2016. arXiv: 1612.06137.
 39.

- 40. M. Eshagh. Alternative expressions for gravity gradients in local north-oriented frame and tensor spherical harmonics. Acta Geophysica, 58(2):215-243, 2010.
- 41. J. Fehrenbach and J-M. Mirebeau. Sparse Non-negative Stencils for Anisotropic Diffusion. Journal of Mathematical Imaging and Vision, pages 1-25, 2013.
- 42. P.T. Fletcher and S. Joshi. Riemannian geometry for the statistical analysis of diffusion tensor data. Signal Processing, 87(2):250-262, 2007.
- 43. M. Gromov. Carnot-Carathéodory spaces seen from within. In Sub-Riemannian Geometry, number 144 in Progress in Mathematics, pages 79–323. Birkhäuser Basel, 1996. DOI: 10.1007/978-3-0348-9210-0_2.
- 44. S. Jbabdi, P. Bellec, R. Toro, J. Daunizeau, M. Pélégrini-Issac, and H. Benali. Accurate Anisotropic Fast Marching for Diffusion-Based Geodesic Tractography. Int J Biomed Imaging, 2008, 2008.
- 45. S. Jbabdi and H. Johansen-Berg. Tractography: Where Do We Go from Here? Brain Connectivity, 1(3):169-183, August 2011.
- 46. J. Jost. Riemannian Geometry and Geometric Analysis, Sixth Edition. Universitext, Springer Berlin Heidelberg, 2011.
- 47. S. Kobayashi and K. Nomizu. Foundations of differential geometry, volume 1. New York, 1963.
- 48. C. Lenglet, E. Prados, J. Pons, R. Deriche, and O. Faugeras. Brain Connectivity Mapping Using Riemannian Geometry, Control Theory, and PDEs. SIAM J. Imaging Sci., 2(2):285-322, January 2009.
- 49. C. Mantegazza and A. Mennucci, "Hamilton-jacobi equations and distance functions on Riemannian manifolds," App. Math. and Optim., vol. 47, no. 1, pp. 1-25, 2002.
- 50. R. Monti and F. Cassano, "Surface measures in Carnot-carathéody spaces," *Calc. Var.*, vol. 13, pp. 339–376, 2001. 51. A. Mashtakov, R. Duits, Yu Sachkov, E. J. Bekkers, and I. Beschastnyi. Tracking of Lines in Spherical Images via Sub-Riemannian Geodesics in SO(3). J Math Imaging Vis, pages 1–26, February 2017.
- 52. A. P. Mashtakov, A. A. Ardentov, and Y. L. Sachkov. Parallel Algorithm and Software for Image Inpainting via Sub-Riemannian Minimizers on the Group of Rototranslations. Numerical Mathematics: Theory, Methods and Applications, 6(01):95-115, February 2013.
- J. Melonakos, V. Mohan, M. Niethammer, K. Smith, M. Kubicki, and A. Tannenbaum. Finsler Tractography for White 53.Matter Connectivity Analysis of the Cingulum Bundle. Med Image Comput Comput Assist Interv, 10(0 1):36-43, 2007.
- 54. J. Melonakos, E. Pichon, S. Angenent, and A. Tannenbaum. Finsler active contours. IEEE Trans Pattern Anal Mach Intell, 30(3):412-423, March 2008.
- 55. J-M. Mirebeau. Efficient fast marching with Finsler metrics. Numer. Math., 126(3):515-557, July 2013.
- 56. J-M. Mirebeau. Anisotropic Fast-Marching on Cartesian Grids Using Lattice Basis Reduction. SIAM J. Numer. Anal., 52(4):1573–1599, January 2014.
- 57. I. Moiseev and Y. L. Sachkov. Maxwell strata in sub-Riemannian problem on the group of motions of a plane. ESAIM: Control, Optimisation and Calculus of Variations, 16(02):380–399, April 2010.
- 58. S. Mori. Introduction to Diffusion Tensor Imaging. Elsevier Science, 2007.
- 59. M. Péchaud, M. Descoteaux, and R. Keriven. Brain connectivity using geodesics in HARDI. Med Image Comput Comput Assist Interv, 12(Pt 2):482-489, 2009.
- 60. J. Petitot. The neurogeometry of pinwheels as a sub-Riemannian contact structure. Journal of Physiology-Paris, 97(2-3):265-309, March 2003.
- 61. J.M. Portegies and R. Duits. New exact and numerical solutions of the (convection-)diffusion kernels on SE(3). arxiv, 2016. https://arxiv.org/pdf/1604.03843.pdf.
- 62. J. A. Reeds and L. A. Shepp. Optimal paths for a car that goes both forwards and backwards. Pacific J. Math., 145(2):367-393, 1990.
- 63. E. Rouy and A. Tourin. A Viscosity Solutions Approach to Shape-From-Shading. SIAM Journal on Numerical Analysis, 29(3):867-884, jul 1992.
- 64. A. Saccon, A. P. Aguiar, A. J. Hausler, J. Hauser, and A. M. Pascoal, "Constrained motion planning for multiple vehicles on se(3)," in 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), pp. 5637-5642, Dec 2012.
- 65. A. A. Agrachev and Y. L. Sachkov, Control Theory from the Geometrical Viewpoint, vol. 87. Springer-Verlag, 2004. Y. L. Sachkov. Cut locus and optimal synthesis in the sub-Riemannian problem on the group of motions of a plane. 66. ESAIM: Control, Optimisation and Calculus of Variations, 17(2):293-321, April 2011.
- 67. G.R. Sanguinetti, E.J. Bekkers, R. Duits, M.H.J. Janssen, A. Mashtakov, and J-M. Mirebeau. Sub-Riemannian Fast Marching in SE(2). In Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, number 9423 in Lecture Notes in Computer Science, pages 366–374. Springer International Publishing, 2015.
- 68. G. Citti and A. Sarti, Models of the Visual Cortex in Lie Groups, pp. 1–55. Basel: Springer Basel, 2015.
- 69. N. Sepasian. Multi-Valued Geodesic Tractography for Diffusion Weghted Imaging. PhD thesis, Eindhoven University of Technology, Dep. of BME, 2011.
- Ordered upwind methods for static Hamilton-Jacobi equations. PNAS, 70. J. A. Sethian and A. Vladimirsky. 98(20):11069-11074, September 2001.
- 71. M.D. Spivak, A Comprehensible Introduction to Differential Geometry, Vol. 2. chap. Connections in Principal Bundles, 305-349. Publish or Perish, 3rd edition, 1999.
- 72. A.F.M. ter Elst and D.W. Robinson. Weighted subcoercive operators on lie groups. Journal of Functional Analysis, 157:88-163, 1998
- 73. J-D. Tournier, F. Calamante, and A. Connelly. MRtrix: Diffusion tractography in crossing fiber regions. Int. J. Imaging Syst. Technol., 22(1):53-66, March 2012.
- 74. J. N. Tsitsiklis. Efficient algorithms for globally optimal trajectories. IEEE Transactions on Automatic Control, 40(9):1528–1538, September 1995.
- 75. D. S. Tuch. Q-ball imaging. Magn Reson Med, 52(6):1358-1372, Dec 2004.
- 76. A. Vladimirsky. Static PDEs for time-dependent control problems. Interfaces and Free Boundaries, pages 281-300, 2006.