

# BICYCLE PATHS, ELASTICAE AND SUB-RIEMANNIAN GEOMETRY

ANDREY ARDENTOV, GIL BOR, ENRICO LE DONNE, RICHARD MONTGOMERY,  
AND YURI SACHKOV

ABSTRACT. We relate the sub-Riemannian geometry on the group of rigid motions of the plane to ‘bicycling mathematics’. We show that this geometry’s geodesics correspond to bike paths whose front tracks are either non-inflectional Euler elasticae or straight lines, and that its infinite minimizing geodesics (or ‘metric lines’) correspond to bike paths whose front tracks are either straight lines or ‘Euler’s solitons’ (also known as Syntractrix or Convicts’ curves).

## CONTENTS

|   |    |
|---|----|
| 1. Introduction   | 2  |
| 2. Wider Context  | 5  |
| 3. Concepts building to the proofs                        | 6  |
| 3.1. Elasticae  | 6  |
| 3.2. Configuration space. Metric concepts.                | 7  |
| 4. The proof of Theorem 1.1 (and some more)               | 9  |
| 4.1. Generalities on geodesics in sub-Riemannian geometry | 10 |
| 4.2. The bicycling geodesic equations                     | 11 |
| 4.3. Proving the 1st claim (Theorem 1.1)                  | 12 |
| 4.4. Proving the 2nd claim                                | 13 |
| 4.5. Proving the 3rd claim                                | 13 |
| 4.6. Proving the 4th claim                                | 14 |
| 5. The proof of Theorem 1.2.                              | 15 |
| 6. Loose Ends and Scattered Wheels                        | 17 |
| 6.1. Bicycling Correspondence                             | 17 |
| 6.2. Not of bundle type                                   | 18 |
| 6.3. Other models for bicycling geometry                  | 18 |
| 6.4. Bicycle parallel transport and hyperbolic geometry   | 19 |
| 6.5. A heuristic proof of Theorem 1.1.                    | 20 |
| 6.6. Open Questions                                       | 20 |
| Appendix A. Proof of Theorem 3.3 on Isometries            | 21 |
| References  | 23 |

## 1. INTRODUCTION

An oriented line segment of fixed length  $\ell$  moves in the Euclidean plane. We think of the segment as a bicycle so that its end points mark the points of contact of the front and back wheels with the ground. As the segment moves, its end points trace a pair of curves, the *front* and *back tracks*. We impose the ‘no-skid’ condition on the motion: the line segment must be tangent to the back track at each instant. Any such motion of a line segment will be called a *bicycle path*. See Figure 1. We define the *length* of a bicycle path to be the ordinary Euclidean length of its front track.

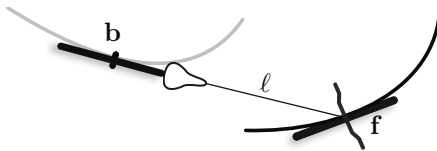


FIGURE 1. The front and back tracks of a bicycle path (the dark and light curves, respectively).

*What are the minimizing bike paths?* These are bike paths whose length minimizes the length among all competing bike paths which connect two given placements of the line segment.

We will say that two curves in the plane *have the same shape* if one curve can be taken onto the other by a homothety, that is, a composition of an isometry and a dilation. The *width* of a plane curve is the infimum of the distances between two parallel lines which bound a strip containing that curve.

**Theorem 1.1.** *The front track of a minimizing bicycle path is a straight line or an arc of a non-inflectional elastic curve of width twice the bicycle length or less. Every possible shape of non-inflectional elastic curve arises in this way.*

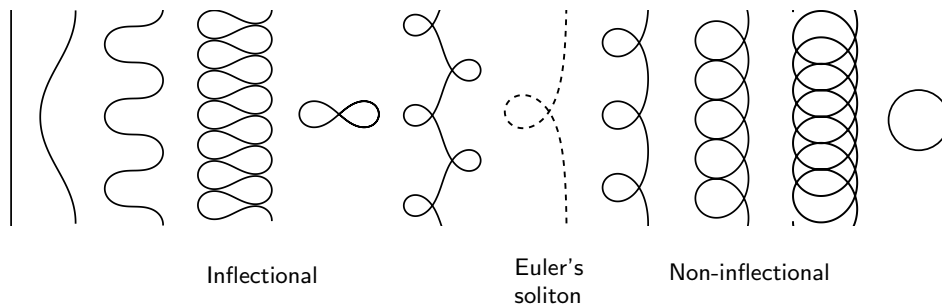


FIGURE 2. The family of elastic curves.

See Figure 2 for some examples of elastic curves, also known as *elasticae*, a remarkable family of plane curves studied by Jacques Bernoulli (1691), Euler (1744) and many others. (We recommend [22] for a nice historical review.) Elasticae can be parameterized by elliptic functions. They are the planar curves having critical total

curvature squared, among all curves with fixed length connecting two given points. They are defined by the differential equation (1) below. Another characterization of elasticae is as curves whose curvature varies linearly with the (signed) distance to some fixed line, the *directrix* of the elastica. (Can you see this line for each of the curves in Figure 2?). Theorem 1.1 provides yet another characterization of elasticae, apparently new. In Figure 2 the Euler soliton and all the curves to its right are ‘non-inflectional’: they have no points with null curvature. All the elasticae to the left of the Euler soliton are inflectional. See Section 3.1 below for more information on elasticae.

In Section 4 we derive relationships between the shapes and widths of the elasticae of Theorem 1.1. In general, a bike path *is not determined* by its front track. That is, for a given front track, there is a circle’s worth of corresponding back tracks, each of which determined by the bicycle frame orientation at some fixed point of the front track. However, for each of the minimizing bike paths of Theorem 1.1, except those whose front track is a line segment, its front track, combined with the condition that the bike path minimizes, *does determine* the back track. For a given shape of a non-inflectional elastica there are two distinct types of minimizing bike paths: one whose front track has width  $2\ell$  and another of certain lesser width (depending on the shape). We call them ‘wide’ and ‘narrow’ front paths. (Exception: Euler solitons appear only in width  $2\ell$ .) The shapes of the back tracks of these two types are quite different. See Figure 7 and Proposition 4.5 for the full details.

Let us emphasize that Theorem 1.1 does *not* state that arbitrary subsegments of a given non-inflectional elastica occur as front tracks of minimizing bike paths. In fact, typically, the opposite is true. Consider for example Figure 3. It depicts a geodesic bike path connecting two horizontal placements of the bike. Clearly, this is not a minimizing path; a straightforward eastward ride will be much shorter. Theorem 1.1 only states that short enough subsegments of this path are minimizing between their endpoints. We do not address here how short is ‘short enough’. For comprehensive results in that direction see [25, 26, 27]. According to our next theorem, the fact that geodesics eventually fail to minimize, as depicted in Figure 3, is typical, with two exceptions.

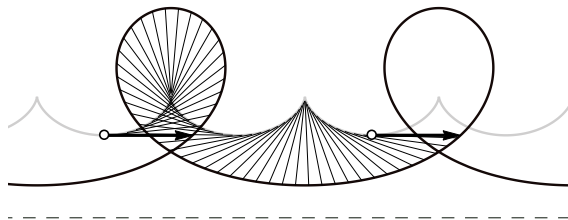


FIGURE 3. A non-minimizing geodesic segment.

**Theorem 1.2.** *An infinitely long bike path is a global minimizer, that is, all of its compact subsegments minimize length between their end points, if and only if it is one of the following two types:*

- (1) *its front track is a straight line and its back track is a tractrix or a straight line, or*

- (2) *its front track is an Euler soliton of width twice the bike length and its back track is a tractrix.*

See Figure 4. Furthermore, there is an isometric involution of the bicycle configuration space which takes paths of one type to paths of the other, provided the back track of the path is a tractrix and not a line. See Lemma 3.2.

In the soliton case, at the ‘highest point’ of the soliton curve, that is, at its point of maximum curvature, the bike frame is oriented perpendicular to the directrix, pointing away from it. For an explicit parametrization of the soliton and tractrix, see Lemma 5.1 below.

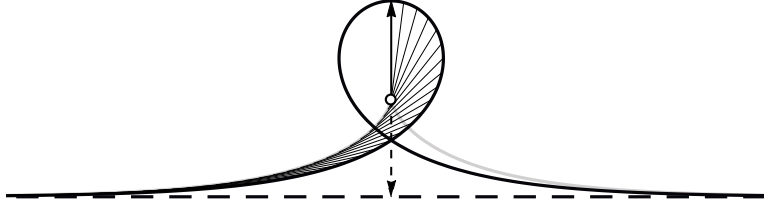


FIGURE 4. Two infinite minimizing bike paths share the tractrix (light curve) as a common back track; the two front tracks are a straight line (dashed dark horizontal line) and an ‘Euler’s soliton’ (solid dark curve).

**About the proofs.** With one notable exception, the proofs of the two theorems above, once set up in the appropriate language, reduce to standard calculations with the geodesic equations of sub-Riemannian geometry. Such a calculation yields Theorem 1.1 and ‘one half’ of Theorem 1.2; namely, that all geodesics, except the two types mentioned in Theorem 1.2, are not globally minimizing (the argument for the last statement is essentially contained in Figure 3). That bike paths whose front track is a straight line are global minimizers follows directly from the definition of bike path length. What remains to show is that geodesics of Theorem 1.1 whose front tracks are Euler solitons are global minimizers. Here, the notable exception mentioned above, we found a surprisingly simple proof, inspired by ‘bicycle mathematics’. The so called ‘bicycle transformation’ (or Darboux transformation or Bäcklund transformation or flip) consists of rotating a bicycle by  $180^\circ$  about its rear end. It is easy to check that this transformation is an isometric involution on the bicycle configuration space, so takes global minimizers to global minimizers. Applying it to a (generic) global minimizer whose front track is a line, we obtain a global minimizer whose front track is an Euler soliton, as depicted in Figure 4.

**Comparison with Previous Works.** One of us has published a series of works [25, 26, 27] on the geodesics and their minimality (or ‘optimal synthesis’ in the language of control theory) for this same subRiemannian geometry. These earlier works focused only on the back wheel projection. The front wheel was not present. What is new in our work is the focus on the front wheel projection and the realization that the front wheel traces out elasticae. We could have derived our minimality results by translating the earlier results from the back wheel over to the front wheel but we have found it simpler and more illuminating to directly study the geodesics from the front wheel point of view.

Our other new contributions are the subRiemannian involution taking straight line tracks to Euler solitons (Lemma 3.2, Theorem 3.3) and the relations sketched

in subsection 6.5 between the geodesics here and those occurring when rolling the hyperbolic plane along the Euclidean plane as investigated by Jurdjevic [17, 18].

**Computer graphics and animations.** Most figures in this article were made using the computer program Mathematica. They are complemented with some ‘bicycle mathematics’ animations, found on the web page <https://www.cimat.mx/~gil/bicycling/>.

**Acknowledgments.** AA and YS were supported by the Russian Science Foundation under grant 17-11-01387-P and performed in Ailamazyan Program Systems Institute of Russian Academy of Sciences. E.L.D. was partially supported by the Academy of Finland (grant 288501 ‘*Geometry of sub-Riemannian groups*’ and by grant 322898 ‘*Sub-Riemannian Geometry via Metric-geometry and Lie-group Theory*’) and by the European Research Council (ERC Starting Grant 713998 GeoMeG ‘*Geometry of Metric Groups*’). GB was supported by CONACYT Grant A1-S-4588.

## 2. WIDER CONTEXT

For a number of surprising theorems around bike paths, and their relations to integrable systems, see [8].

The bicycling configuration space is diffeomorphic to the three-dimensional Lie group  $SE_2$  of rigid motions of the plane (orientation preserving isometries). Its length structure comes from a left-invariant sub-Riemannian metric on this group. See Section 3.2 below for details. Such a structure is unique up to scale, [1, 9], and that scale can be interpreted as the length of the bicycle frame. This structure, from the perspective of the *back wheel* track, has been investigated by many authors [13, 25, 26, 27, 15] and used to understand aspects of mammalian vision. In that latter context the group  $SE_2$  is typically referred to as the “roto-translational group” and the orientation of the bicycle frame is the crucial object, as optical processing in the brain involves cells whose function is to perceive orientations of line segments.

Gershkovich and Vershik gave a general description and classification of left invariant sub-Riemannian structures on three-dimensional Lie groups in [30], see also [1]. In all cases the geodesic equations are those of ‘generalized elastica’.

On any metric space we can speak of ‘globally minimizing geodesics’ or, synonymously, ‘metric lines’: isometric embeddings of the real line into the metric space. See [10]. What are the metric lines for a given sub-Riemannian structure? Theorem 1.2 answers this question for the bicycling case.

Hakavuouri and LeDonne [19] prove a number of powerful general theorems regarding metric lines in sub-Riemannian geometries by implementing the operation of “blowing down” a geodesic. Sufficient iterations of blow-down yield a line in a Euclidean space. As a corollary, they prove that if a sub-Riemannian geometry  $Q$  comes, like ours, with a sub-Riemannian submersion  $\pi$  to the Euclidean plane  $\mathbb{R}^2$ , then (1) the projection of any metric line in  $Q$  must lie a bounded distance from a line in the plane, and (2) if that planar line is given by  $x = 0$  and if we write the projected geodesic as  $(x(t), y(t))$  then  $x(t)$  cannot be a non-constant periodic function. Item (2) excludes all the elasticae of Theorem 1.2 besides the line and the soliton from being metric lines.

We know five other rank 2 sub-Riemannian geometries besides our  $SE_2$  geometry whose geodesics project to elasticae under a sub-Riemannian submersion onto the Euclidean plane. (See the third paragraph of Section 3.2 for the definition of a ‘sub-Riemannian submersion’.) Two are Carnot geometries, one being the Engel

group, whose growth vector is  $(2, 3, 4)$  (see [4, 3]), and the other, sometimes called the Cartan group, being the unique Carnot group with growth vector  $(2, 3, 5)$  (see [28]). (The growth vector of a Carnot group is its basic numerical invariant and encapsulates the graded dimensions of its Lie algebra.) Another is the flat Martinet geometry, see [2]. The remaining two are five-dimensional, arising from rolling a constant curvature surface along the Euclidean plane, and have state spaces  $SO_3 \times \mathbb{R}^2$  and  $PSL_2(\mathbb{R}) \times \mathbb{R}^2$ . See [17, 18] for a derivation of elasticae as their geodesics. In all five geometries the geodesics projecting to Euclidean lines are metric lines and in the flat Martinet geometry these exhaust the set of metric lines. In the other four geometries some of the geodesics which project onto solitons are also metric lines. In the  $SO_3 \times \mathbb{R}^2$  case all elasticae, and in particular all Euler solitons, arise as projections of geodesics onto  $\mathbb{R}^2$ , but only some of the solitons, namely those whose kink is ‘small enough’, arise as projections of metric lines.

*Are all these occurrences of elasticae and Euler solitons in sub-Riemannian geometries related?* There is a sub-Riemannian submersion from the Cartan group onto the Engel group, so that the space of Engel geodesics embed into the space of Cartan geodesics by horizontal lift. Similarly, the space of flat Martinet geodesics embed into the space of Engel geodesics. The bike configuration space  $Q = SE_2$  can be constructed as a circle bundle associated to the hyperbolic rolling space  $PSL_2(\mathbb{R}) \times \mathbb{R}^2$ , viewed as a principal  $PSL_2(\mathbb{R})$  bundle, and this fact and its related geometry allows us to embed the bike geodesics into the hyperbolic rolling geodesics. See the last paragraph of Section 6.4 below. We leave the possibility of uncovering relations between the other pairs of geometries and of some deeper reason underlying the ubiquity of elasticae in sub-Riemannian geometry to future researchers.

### 3. CONCEPTS BUILDING TO THE PROOFS

**3.1. Elasticae.** An immersed plane curve is an *elastica* if its curvature  $\kappa(t)$ , as a function of arc length  $t$ , satisfies the 2nd order ODE

$$\ddot{\kappa} + \frac{1}{2}\kappa^3 + A\kappa = 0$$

for some constant  $A$ . See, for example, [31]. This is an equation of Newton’s type, with potential  $\frac{1}{8}\kappa^4 + \frac{1}{2}A\kappa^2$ . Consequently, there is a constant ‘energy’  $B \in \mathbb{R}$  such that

$$(1) \quad \frac{1}{2}(\dot{\kappa})^2 + \frac{1}{8}\kappa^4 + \frac{A}{2}\kappa^2 = B.$$

We call the latter equation the ‘energy form’ of the elastica equation. If  $\kappa(t_0) = 0$  at some point then the energy equation asserts that  $B \geq 0$ . Consequently, if  $B < 0$  we must have that  $\kappa$  never vanishes along the curve. Since  $\kappa(t_0) = 0$  corresponds to an inflection point of the curve, we call such elasticae ‘non-inflectional.’ Elasticae for which  $B = 0$ ,  $A < 0$  are also non-inflectional and consist of the Euler solitons. All non-inflectional elasticae, except the Euler solitons, have periodic curvature.

Equation (1) can be rewritten (by ‘completing the square’) as

$$(2) \quad \dot{\kappa}^2 + \left( \frac{\kappa^2}{2} + A \right)^2 = 2B + A^2.$$

Thus the parameters must satisfy  $2B + A^2 \geq 0$ . The set of elasticae is invariant under dilations. To dilate an immersed plane curve  $c(t)$  parameterized by arclength  $t$  by a

factor  $\lambda > 0$  we form  $\tilde{c}(t) = \lambda c(t/\lambda)$ . The dilated curve  $\tilde{c}(t)$  is still parameterized by arclength and has curvature  $\tilde{\kappa}(t) = \frac{1}{\lambda} \kappa(\frac{t}{\lambda})$ . It follows by a direct computation that the  $\lambda$ -dilate of an elastica satisfying equation (1) with parameters  $A, B$  satisfies a new equation (1), now with rescaled parameters  $\tilde{A} = A/\lambda^2, \tilde{B} = B/\lambda^4$ . Thus

$$(3) \quad \mu := -2B/A^2 \leq 1$$

is scale invariant and can be thought of as a ‘shape parameter’. Non-inflectional elasticae correspond to  $B \leq 0$ , that is,  $0 \leq \mu \leq 1$ , in which case  $A \leq 0$  as well. There are 3 types of ‘exceptional’ elasticae, lines, circles and solitons. Lines correspond to solutions of equation (1) with  $\kappa = B = 0$ , solitons to  $B = 0, A < 0, \kappa \neq 0$ , and circles to  $\mu = 1$ , that is,  $B = -A^2, \dot{\kappa} = 0$ . See Figure 5.

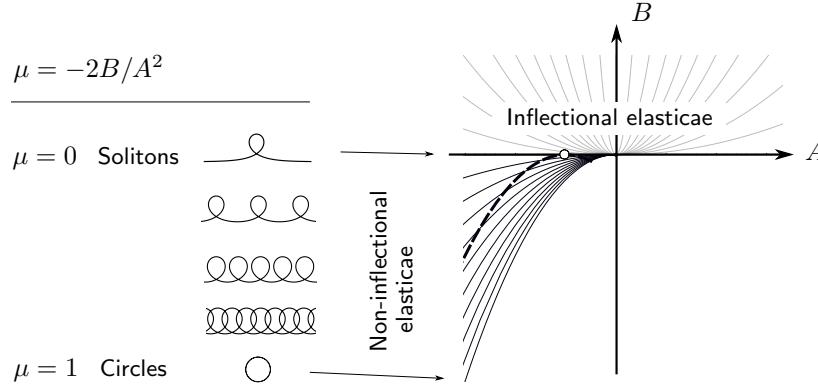


FIGURE 5. Elasticae parameter space. The light and dark solid curves parametrize inflectional ( $B > 0$ ) and non-inflectional ( $B < 0$ ) elasticae (respectively) of ‘constant shape’, level curves of the shape parameter  $\mu = -2B/A^2$  of equation (3). The dashed heavy curve in the third quadrant corresponds to the elasticae which appear as front tracks of geodesic bike paths for fixed bike length  $\ell = 1$  (see Proposition 4.3, where this curve is parametrized by  $a$ ). Its intersection with the  $A$ -axis (marked with a white dot) stands for the Euler soliton ( $a = 1$  in equation (11)). Each non-zero level curve of  $\mu$  in the third quadrant intersects the dashed curve at 2 points, corresponding to the two sizes of non-inflectional elasticae appearing as front tracks of bicycle geodesics, ‘wide’ and ‘narrow’ (to the left and right of the white dot, respectively).

**3.2. Configuration space. Metric concepts.** We begin by reformulating our theorems in the language of sub-Riemannian and metric geometry.

Let  $\ell > 0$  be the bicycle length. Then the configuration space for bike motions can be expressed as  $Q = \{(\mathbf{b}, \mathbf{f}) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid \|\mathbf{f} - \mathbf{b}\| = \ell\} \subset \mathbb{R}^2 \times \mathbb{R}^2$ , where  $\|\cdot\|$  is the standard Euclidean norm on  $\mathbb{R}^2$ . It is easy to see that  $Q$  is a smooth manifold diffeomorphic to  $\mathbb{R}^2 \times S^1$ . The no-skid condition defines a rank 2 distribution  $D \subset TQ$  on  $Q$  by saying that a vector  $(\dot{\mathbf{b}}, \dot{\mathbf{f}}) \in T_{(\mathbf{b}, \mathbf{f})}Q$  belongs to  $D_{(\mathbf{b}, \mathbf{f})}$  if and only if  $\dot{\mathbf{b}}$  is a multiple of  $\mathbf{f} - \mathbf{b}$ . *Bike paths are the integral curves of  $D$ .*

$D$  is a *contact distribution*. We prove this in Lemma 4.1 below. Alternatively, in Section 6.3 we show how to identify  $Q$  with the space of (oriented) tangent lines to the plane, also known as “contact elements” since they represent 1st order contact

of curves. In this context  $D$  is the canonical contact distribution on this space of contact elements, one of the first examples of a contact manifold. See for example Appendix 4 of Arnol'd's famous book, [5].

Let  $\pi_f : Q \rightarrow \mathbb{R}^2$  be the front wheel projection,  $(\mathbf{b}, \mathbf{f}) \mapsto \mathbf{f}$ .  $D$  is transverse to the fibers of  $\pi_f$ , hence one can equip  $D$  with an inner product by pulling back the Euclidean metric on  $\mathbb{R}^2$  to  $Q$  by  $\pi_f$ , then restricting to  $D$ . The 3-manifold  $Q$ , together with the distribution  $D$  and the inner product on it, is an example of a *sub-Riemannian manifold*. We constructed the inner product on  $D$  in such a way that the front wheel projection is a *sub-Riemannian submersion*: for each  $\mathbf{q} \in Q$  the differential  $d\pi_f(\mathbf{q})$  maps the 2-plane  $D_{\mathbf{q}}$  isometrically onto  $T_{\pi_f(\mathbf{q})}\mathbb{R}^2 = \mathbb{R}^2$ . This sub-Riemannian structure is isometric to the standard sub-Riemannian structure on the group  $\text{SE}_2$  studied in [13, 15, 25, 26, 27]. Up to scaling and isometries, it is the unique left-invariant sub-Riemannian structure on  $\text{SE}_2$  of contact type.

Since  $Q$  is connected and  $D$  is contact, the Chow-Rashevskii Theorem [23] implies that any two points in  $Q$  are connected by a bicycle path. The length of such a path is defined using the inner product on  $D$ , just as in Riemannian geometry. In view of our construction of the inner product, the length of a bike path equals the length of its front wheel projection to  $\mathbb{R}^2$ , as asserted in the introduction.

Defining the distance between two points of  $Q$  to be the infimum of the lengths of the bike paths connecting them turns  $Q$  into a metric space. A *minimizing geodesic* in  $Q$  is a bike path  $\gamma : I \rightarrow Q$ , where  $I \subset \mathbb{R}$  is a compact interval, realizing the distance between its end points. A *geodesic* is a bike path  $\gamma : I \rightarrow Q$ , where  $I \subset \mathbb{R}$  is an interval (possibly non-compact), such that every  $t_0 \in I$  is contained in a compact subinterval  $I' \subset I$  for which  $\gamma|_{I'}$  is a minimizing geodesic. In addition, we require that if  $t_0$  is an interior point of  $I$  then  $t_0$  is also an interior point of  $I'$ . (This last condition excludes arbitrary concatenations of minimizing geodesics from being geodesics.)

Theorem 1.1 states that the  $\pi_f$ -image of any geodesic is a non-inflectional elastic curve or a straight line. A *metric line* in  $Q$  is an infinite geodesic all of whose compact subsegments are minimizing geodesics. Theorem 1.2 states that the  $\pi_f$ -image of any metric line is either a Euclidean line or an Euler soliton. (The ‘width’ of this soliton is twice the length of the bike frame.)

A *sub-Riemannian isometry* of  $Q$  is a diffeomorphism that preserves  $D$  and the inner product on it.

**Remark 3.1.** Clearly, a sub-Riemannian isometry is a distance preserving homeomorphism. The latter can be taken as a weaker ‘metric’ definition of isometry. For a general sub-Riemannian manifold, the equivalence of the two definitions is an open problem. For an *equi-regular* sub-Riemannian structure, such as our case (or any homogeneous sub-Riemannian manifold), the two notions are equivalent [11, 21].

By construction, the action of the group  $E_2$  of isometries of the plane  $\mathbb{R}^2$  lifts to an action on  $Q$  by sub-Riemannian isometries. An element  $g \in E_2$  acts on  $Q$  sending  $(\mathbf{b}, \mathbf{f})$  to  $(g\mathbf{b}, g\mathbf{f})$  so that our sub-Riemannian submersion  $\pi_f$  intertwines the  $E_2$ -action on  $Q$  with the standard action of  $E_2$  on  $\mathbb{R}^2$ . But these are not all the sub-Riemannian isometries of  $Q$ . There is one extra symmetry that plays an important role in our proof of Theorem 1.2.



**Lemma 3.2.** *The map  $\Phi : Q \rightarrow Q$ ,  $(\mathbf{b}, \mathbf{f}) \mapsto (\mathbf{b}, 2\mathbf{b} - \mathbf{f})$ , which ‘flips’ the bike frame about the back wheel is a sub-Riemannian isometry of  $Q$ . See Figure 6.*

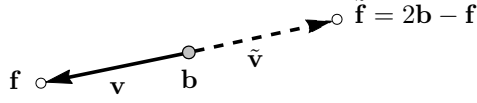


FIGURE 6. Lemma 3.2: Flipping a bike about its back wheel.

*Proof.*  $\Phi$  is the restriction of a linear map to  $Q \subset \mathbb{R}^2 \times \mathbb{R}^2$ . Thus its derivative is given by the same formula,  $(\dot{\mathbf{b}}, \dot{\mathbf{f}}) \mapsto (\dot{\mathbf{b}}, 2\dot{\mathbf{b}} - \dot{\mathbf{f}})$ . It clearly preserves the no-skid condition hence it leaves  $D$  invariant. It remains to show that  $\|\dot{\mathbf{f}}\| = \|2\dot{\mathbf{b}} - \dot{\mathbf{f}}\|$ . Now decompose orthogonally  $\dot{\mathbf{f}} = \dot{\mathbf{f}}_{\parallel} + \dot{\mathbf{f}}_{\perp}$ ,  $\dot{\mathbf{b}} = \dot{\mathbf{b}}_{\parallel} + \dot{\mathbf{b}}_{\perp}$ , where  $\dot{\mathbf{f}}_{\parallel}, \dot{\mathbf{b}}_{\parallel}$  are the orthogonal projections along  $\mathbf{b} - \mathbf{f}$ . The bicycling no-skid condition implies  $\dot{\mathbf{b}}_{\perp} = 0$  and  $\|\mathbf{b} - \mathbf{f}\| = \text{const}$  implies  $\dot{\mathbf{f}}_{\parallel} = \dot{\mathbf{b}}_{\parallel}$ , hence  $\dot{\mathbf{f}}_{\parallel} = \dot{\mathbf{b}}_{\parallel}$ . Thus  $2\dot{\mathbf{b}} - \dot{\mathbf{f}} = 2\dot{\mathbf{f}}_{\parallel} - (\dot{\mathbf{f}}_{\parallel} + \dot{\mathbf{f}}_{\perp}) = \dot{\mathbf{f}}_{\parallel} - \dot{\mathbf{f}}_{\perp}$ . That is,  $2\dot{\mathbf{b}} - \dot{\mathbf{f}}$  is the reflection of  $\dot{\mathbf{f}}$  about  $\mathbf{b} - \mathbf{f}$ . It follows that  $\|2\dot{\mathbf{b}} - \dot{\mathbf{f}}\| = \|\dot{\mathbf{f}}\|$ .  $\square$

For completeness we describe the full group of isometries of  $Q$ .

**Theorem 3.3.** *The group  $\text{Isom}(Q)$  of all sub-Riemannian isometries of  $Q$  is an extension of  $\text{E}_2$  by the two-element group  $\mathbb{Z}/2\mathbb{Z}$ . This two-element group is generated by the isometric involution  $\Phi$  which ‘flips the bike frame’, as described in Lemma 3.2 above. Thus*

$$\text{Isom}(Q) \simeq \text{E}_2 \rtimes \mathbb{Z}/2\mathbb{Z} \simeq \text{SE}_2 \rtimes (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}).$$

*The identity component of  $\text{Isom}(Q)$  is  $\text{SE}_2$ , acting freely and transitively on  $Q$  and so induces a sub-Riemannian isometry between  $Q$  and a left-invariant sub-Riemannian metric on  $\text{SE}_2$ .*

We prove this theorem in the appendix. Hladky [16], in his final section, computes that the Lie algebra of  $\text{Isom}(Q)$  is that of  $\text{SE}_2$ . The same conclusion can be drawn from the asphericity of the associated CR structure, as in [9, §7]. But calculating the Lie algebra of  $\text{Isom}(Q)$  only describes the identity component of  $\text{Isom}(Q)$ , missing the ‘discrete part’ (or ‘isotropy representation’) of the isometry group, as we do in the appendix.

#### 4. THE PROOF OF THEOREM 1.1 (AND SOME MORE)

We prove a more detailed version of Theorem 1.1, subdividing the assertions into 4 claims. Most of these claims do not hold for the ‘exceptional’ elasticae (line, circle, soliton). We first describe the non-exceptional situation, then correct for the exceptional elasticae.

**Claim 1** (Theorem 1.1). The front track of each bicycle geodesic is a NIE (non inflectional elastica) or a straight line.

**Claim 2** (Wide and narrow). For a fixed bicycle frame of length  $\ell$ , each shape of NIE appears as a front track in two different sizes, ‘wide’ and ‘narrow’: The wide front tracks are NIE of width  $2\ell$ . A narrow front track can have any width

in  $(0, 2\ell)$ , depending on its shape: the more circular is a narrow front track, the narrower it is. See Figure 7.

Exceptions: the circle and the soliton appear as front tracks only with width  $2\ell$ .

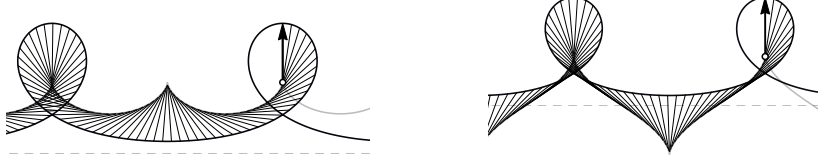


FIGURE 7. Two bicycling geodesics with front tracks (dark solid curve) which are non-inflectional elasticae of the same shape but of different size: ‘wide’ (left) and ‘narrow’ (right). In each of the two figures: the horizontal dashed light line is the directrix, the light solid curve is the back track, the arrow depicts the bicycle frame, pointing towards the front wheel, at the moment of going through a point of maximum curvature of the front track. See Proposition 4.7.

**Claim 3** (Unique horizontal lift). Each non-linear NIE front track, wide or narrow, has a unique horizontal lift to a bicycle geodesic. This lift is determined by the back track found by the following rule: at points of maximum curvature of the front track the bicycle frame is perpendicular to the front track, pointing ‘outside’ the front track (that is, in the direction opposite to the acceleration vector of the front track). See Figure 7 and our web animations [7].

The bicycle frame is also perpendicular to the front track at the points of minimum curvature. For the wide NIE, the frame at this point also points outside the front track. For the narrow NIE the frame points inside.

Exception: all horizontal lifts of a Euclidean line are globally minimizing bike paths. Two of the lifts correspond to riding along the line, either forward or backwards, with the bike frame aligned with the line. The rest of the lifts correspond to the back wheel tracing a *tractrix* of width  $\ell$  (the light solid curve of Figure 4).

**Claim 4** (Flipping a front track). There is a sub-Riemannian involution  $\Phi : Q \rightarrow Q$  on the bicycling configuration space, rotating the bicycle frame by  $180^\circ$  about its rear end. It acts on the space of bicycle geodesics, as well as their front tracks, preserving the ‘narrow’ and ‘wide’ subclasses. Each NIE has its ‘length’  $L$ : the distance between two successive points along the curve of maximum (or minimum) curvature, see Figure 11. The flip of a wide NIE is obtained by translating it by  $L/2$  along its directrix. The flip of a narrow NIE is obtained by a ‘glide reflection’: translation by  $L/2$  along the directrix followed by a reflection about it. See Figure 8.

Exceptions: the flip of the circle is the circle itself, the flip of the line is the soliton, of width  $2\ell$ , and vice versa.

**4.1. Generalities on geodesics in sub-Riemannian geometry.** To prove the above 4 claims we review some general facts from sub-Riemannian geometry. For more details see Chapter 1 of [23].

Let  $M$  be a smooth manifold. We can turn a smooth vector field  $X$  on  $M$  into a fiber-linear function  $P_X : T^*M \rightarrow \mathbb{R}$  by the rule  $P_X(\mathbf{q}, \mathbf{p}) = \mathbf{p}(X(\mathbf{q}))$ , where  $\mathbf{q} \in M$



FIGURE 8. The flips of bicycle geodesics with ‘wide’ (left) and ‘narrow’ (right) front tracks of the same shape. In each of the two figures: the dashed dark curve is the ‘flip’ of the solid dark curve and vice versa, the light solid curve is their common back track, the light dashed horizontal line is their common directrix, the solid arrow indicates the bicycle frame at a point of maximum curvature of the solid front track and the dashed arrow indicates the bicycle frame at a point of minimum curvature of the dashed front track. See Proposition 4.11.

and  $\mathbf{p} \in T_{\mathbf{q}}^*M$ . Consider a general rank  $r$  distribution  $D \subset TM$ , equipped with a sub-Riemannian metric on  $D$  and an orthonormal frame  $X_1, \dots, X_r \in \Gamma(D)$ . Form the corresponding fiber-linear functions  $P_i := P_{X_i}$ . Then the *normal geodesics* of this sub-Riemannian structure are, by definition, the projections onto  $M$  of the solutions to the standard Hamiltonian equation on  $T^*M$ ,

$$(4) \quad \dot{\mathbf{q}} = \partial_{\mathbf{p}} H, \quad \dot{\mathbf{p}} = -\partial_{\mathbf{q}} H, \quad \text{where } H = \frac{1}{2} \sum_i (P_i)^2.$$

See Theorem 1.14 on page 9 of [23] for the full statement and later, a proof.

Normal geodesics parametrized by arc length correspond to solutions of equation (4) with energy  $H = 1/2$ . Short enough segments of normal geodesics are length minimizers, but the converse is not true, in general, due to the existence of *singular* (or *abnormal*) geodesics. See [23], particularly Chapters 3 and 5. However, a basic result of the theory is: *if  $D$  is a contact distribution then all length minimizing  $D$ -horizontal curves are normal geodesics*. See the example at the top of page 59 in [23].

**4.2. The bicycling geodesic equations.** Let  $Q = \{(\mathbf{b}, \mathbf{f}) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid \|\mathbf{b} - \mathbf{f}\| = 1\}$  be the bicycling configuration space, equipped with the coordinates  $(x, y, \theta)$ , where  $\mathbf{f} = (x, y)$ ,  $\mathbf{b} = \mathbf{f} - (\cos \theta, \sin \theta)$ , with associated global coordinate vector field framing  $\partial_x, \partial_y, \partial_\theta$ . (We take, without loss of generality the bike length  $\ell = 1$ . The general case reduces to this case by an easy rescaling argument.) The conjugate fiber coordinates on  $T^*Q$  are  $p_x := P_{\partial_x}$ ,  $p_y := P_{\partial_y}$ ,  $p_\theta := P_{\partial_\theta}$ .

**Lemma 4.1.** *The no-skid condition defines on  $Q$  a rank 2 distribution  $D \subset TQ$ , the kernel of the 1-form*

$$(5) \quad \eta := d\theta - \mathbf{c}dy + \mathbf{s}dx, \quad \text{where } \mathbf{c} = \cos \theta, \quad \mathbf{s} = \sin \theta.$$

*It follows that  $\eta \wedge d\eta = -dx \wedge dy \wedge d\theta$  is non-vanishing, hence  $D$  is a contact distribution.*

*Proof.* Let  $\mathbf{q}(t) = (\mathbf{b}(t), \mathbf{f}(t))$  be a curve in  $Q$  satisfying the non-skid condition. Let  $\mathbf{v} := \mathbf{f} - \mathbf{b} = (\mathbf{c}, \mathbf{s})$  and decompose orthogonally  $\dot{\mathbf{f}} = \dot{\mathbf{f}}_{\parallel} + \dot{\mathbf{f}}_{\perp}$ , where  $\dot{\mathbf{f}}_{\parallel}, \dot{\mathbf{f}}_{\perp}$  are the

orthogonal projections of  $\dot{\mathbf{f}}$  on  $\mathbf{v}, \mathbf{v}^\perp$ , respectively. The condition  $\|\mathbf{f} - \mathbf{b}\| = \text{const.}$  and the no-skid condition  $\dot{\mathbf{b}}\|\mathbf{v}$  are equivalent to  $\dot{\mathbf{f}}_\parallel = \dot{\mathbf{b}}$ . From  $\mathbf{f} = \mathbf{b} + \mathbf{v}$  follows  $\dot{\mathbf{f}} = \dot{\mathbf{b}} + \dot{\mathbf{v}}$ , hence  $\dot{\mathbf{f}}_\parallel = \dot{\mathbf{b}}$  is equivalent to  $\dot{\mathbf{v}} = \dot{\mathbf{f}}_\perp = \dot{\mathbf{f}} - \dot{\mathbf{f}}_\parallel = \dot{\mathbf{f}} - \langle \dot{\mathbf{f}}, \mathbf{v} \rangle \mathbf{v}$ . In coordinates, this is  $\dot{\theta} - c\dot{y} + s\dot{x} = 0$ . That is,  $\dot{\mathbf{q}} \in \text{Ker}(\eta)$ .  $\square$

Thus minimizing bike paths are arcs of normal geodesics. An orthonormal framing for  $D = \text{Ker}(\eta)$  is

$$(6) \quad X_1 := \partial_x - s\partial_\theta, \quad X_2 := \partial_y + c\partial_\theta,$$

with the associated

$$(7) \quad P_1 := P_{X_1} = p_x - sp_\theta, \quad P_2 := P_{X_2} = p_y + cp_\theta.$$

The Hamiltonian equations associated to  $H = [(P_1)^2 + (P_2)^2]/2$  are

$$(8) \quad \begin{aligned} \dot{x} &= \partial_{p_x} H = P_1 = p_x - sp_\theta, & \dot{p}_x &= -\partial_x H = 0, \\ \dot{y} &= \partial_{p_y} H = P_2 = p_y + cp_\theta, & \dot{p}_y &= -\partial_y H = 0, \\ \dot{\theta} &= \partial_{p_\theta} H = p_\theta + cp_y - sp_x, & \dot{p}_\theta &= -\partial_\theta H = p_\theta(cp_x + sp_y). \end{aligned}$$

So  $p_x, p_y$  are constant, as well as  $H = (\dot{x}^2 + \dot{y}^2)/2$ . Fixing  $H = 1/2$  thus means that  $\mathbf{f}(t)$  is parametrized by arc length. Rotations act on the space of solutions of equations (8) by rotating  $(x, y)$ ,  $(p_x, p_y)$  and  $(c, s)$  simultaneously and shifting  $\theta$ , leaving  $p_\theta$  unchanged. So we can assume without loss of generality say  $p_y = 0$  and  $a := p_x \geq 0$ . Equations (8) and  $H = 1/2$  now become

$$(9) \quad \dot{x} = a - sp_\theta, \quad \dot{y} = cp_\theta, \quad \dot{\theta} = p_\theta - as, \quad \dot{p}_\theta = acp_\theta, \quad \dot{\theta}^2 + a^2c^2 = 1.$$

**Lemma 4.2.** *Let  $\kappa$  be the geodesic curvature of the front track  $\mathbf{f}(t) = (x(t), y(t))$  of a solution to equations (9). Then  $\kappa = p_\theta$ .*

*Proof.* We calculate:  $\kappa = \dot{x}\ddot{y} - \dot{y}\ddot{x} = \dot{\theta} + as = p_\theta$ .  $\square$

We can thus rewrite the unit speed geodesic equations (9) as

$$(10) \quad \dot{x} = a - s\kappa, \quad \dot{y} = c\kappa, \quad \dot{\theta} = \kappa - as, \quad \dot{\kappa} = ac\kappa, \quad \dot{\theta}^2 + a^2c^2 = 1.$$

We are now ready to prove the 4 claims.

#### 4.3. Proving the 1st claim (Theorem 1.1).

**Proposition 4.3.** *The curvature  $\kappa$  of the front track of a bicycle geodesic (solution to equations (8)), as a function of arc length  $t$ , satisfies the ‘energy form’ of the elastica equation (1),*

$$\frac{\dot{\kappa}^2}{2} + \frac{\kappa^4}{8} + \frac{A\kappa^2}{2} = B,$$

with

$$(11) \quad A = -\frac{a^2 + 1}{2}, \quad B = -\frac{(a^2 - 1)^2}{8}.$$

*That is, the front track is a non-inflectional elastica or a straight line.*

*Proof.* The statement is invariant under rigid motions, so we can use instead equations (10). Then  $\dot{\kappa} = ac\kappa$  and  $\dot{\theta}^2 = (\kappa - as)^2 = 1 - a^2c^2$ , which simplifies to  $2as\kappa = \kappa^2 + a^2 - 1$ . Thus  $4\dot{\kappa}^2 + (\kappa^2 + a^2 - 1)^2 = 4a^2\kappa^2$ , which gives the stated formula.  $\square$

**Remark 4.4.** In the last paragraph of Section 6.4 below we sketch an alternative proof of Claim 1. This alternative proof uses a relation between hyperbolic rolling geodesics and bicycling geodesics and the fact that the hyperbolic rolling geodesics had been already computed and shown to correspond to elasticae [17], [18].

**4.4. Proving the 2nd claim.** The ‘shape parameter’ of the front track is  $\mu = -2B/A^2 = (a^2 - 1)^2/(a^2 + 1)^2 \in [0, 1]$ . Each  $\mu \in (0, 1)$  has 2 preimages,  $a$  and  $1/a$ , one in  $(0, 1)$  the other in  $(1, \infty)$ . It follows that each NIE shape appears as a front track for two values of  $a$ . Let us determine the widths of these front tracks. Let  $\kappa_{max}, \kappa_{min} > 0$  be the maximum and minimum value of  $\kappa$  along the front track.

**Proposition 4.5.** *The front track of a solution to equations (10) with  $\kappa > 0$  has  $\kappa_{max} = 1 + a$  and  $\kappa_{min} = |1 - a|$ . It follows that the width of the front track is*

- 2 if  $0 < a \leq 1$  (a ‘wide’ front track);
- $2/a$  if  $a > 1$  (a ‘narrow’ front track).

See Figure 7.

*Proof.* Since  $\dot{\kappa} = 0$  at  $\kappa_{max}, \kappa_{min}$ , these critical values must satisfy  $(\kappa^2 + a^2 - 1)^2 = 4a^2\kappa^2$  (see Proposition 4.3 and its proof). The solutions of this equation are  $\kappa = \pm 1 \pm a$ . For  $0 < a < 1$  the positive solutions are  $1 \pm a$ , hence  $\Delta\kappa = \kappa_{max} - \kappa_{min} = 2a$ . For  $a > 1$ , the positive solutions are  $a \pm 1$ , hence  $\Delta\kappa = 2$ . From  $\dot{\kappa} = a\dot{y}$  it follows that  $\Delta y = 2$  in the 1st case and  $2/a$  in the 2nd case, as needed.  $\square$

**Remark 4.6.** One can also use equations (10) to find the widths of the respective back tracks:  $(1 - \sqrt{1 - a^2})/a$  for a ‘wide’ front track, and  $2/a$  for a ‘narrow’ front track (same as the width of the front track).

#### 4.5. Proving the 3rd claim.

**Proposition 4.7.** *Consider a solution of equations (10) with  $a > 0$  and  $\kappa > 0$  (this can always be arranged for a non-linear front track by appropriate reflections about the  $x$  and  $y$  axes). Then*

- (1)  $\theta = \pi/2$  at a point where  $\kappa_{max}$  occurs.
- (2)  $\theta = -\pi/2$  at a point where  $\kappa_{min}$  occurs and  $0 < a < 1$  (a ‘wide’ front track).
- (3)  $\theta = \pi/2$  at a point where  $\kappa_{min}$  occurs and  $a > 1$  (a ‘narrow’ front track).

See Figure 7.

**Remark 4.8.** The  $a = 1$  case is either a soliton, where  $\kappa$  does not have a minimum, or a straight line. The  $a = 0$  case is that of the unit circle and is safely left to the reader.

*Proof.* By equations (10),  $\dot{\kappa} = a\mathbf{c}\kappa = a\dot{y}$ . Thus in all 3 cases,  $\dot{\kappa} = 0$  implies  $\mathbf{c} = \dot{y} = 0$ , which implies  $\mathbf{s} = \pm 1$  and  $\dot{x} = \pm 1$ . We shall also use the formulas  $\kappa_{max} = 1 + a$  and  $\kappa_{min} = |1 - a|$  from Proposition 4.5, and  $\dot{x} = a - \mathbf{s}\kappa$  of equations (10).

- (1) Substitute  $\kappa = 1 + a$  in  $\dot{x} = a - \mathbf{s}\kappa$  and get  $\dot{x} + \mathbf{s} = a(1 - \mathbf{s})$ . If  $\mathbf{s} = -1$  then  $\dot{x} = 2a + 1 > 1$ , which is impossible, hence  $\mathbf{s} = 1, \dot{x} = -1$  and  $\theta = \pi/2$ .
- (2) If  $0 < a < 1$  then  $\kappa_{min} = 1 - a$ . Substitute this in  $\dot{x} = a - \mathbf{s}\kappa$  and get  $\dot{x} + \mathbf{s} = a(1 + \mathbf{s})$ . If  $\mathbf{s} = 1$  then  $\dot{x} = 2a - 1$ . Together with  $0 < a < 1$  this implies  $-1 < \dot{x} < 1$  which contradicts  $\dot{x} = \pm 1$ . Hence  $\mathbf{s} = -1, \dot{x} = 1$  and  $\theta = -\pi/2$  at a point where  $\kappa_{min}$  occurs.

- (3) If  $a > 1$  then  $\kappa_{min} = a - 1$ . Substitute this in  $\dot{x} = a - s\kappa$  and get  $\dot{x} - s = a(1 - s)$ . If  $s = -1$  then  $\dot{x} = 2a - 1$ . Together with  $a > 1$  this implies  $\dot{x} > 1$ , which contradicts  $\dot{x} = \pm 1$ . Hence  $s = 1$ ,  $\dot{x} = 1$  and  $\theta = \pi/2$  at a point where  $\kappa_{min}$  occurs.  $\square$

**4.6. Proving the 4th claim.** Let  $\gamma$  be a bicycle geodesic. A *vertex* of  $\gamma$  is a point on it where an extremum of the curvature of the front track occurs ( $\dot{\kappa} = 0$ ). Our involution  $\Phi : Q \rightarrow Q$  is a sub-Riemannian isometry, hence  $\Phi \circ \gamma$  is a geodesic as well.

**Lemma 4.9.** *If  $\gamma$  is a bicycle geodesic with  $a \neq 1$  (that is, its front track is not a line or soliton) then  $\Phi$  maps vertices of  $\gamma$  to vertices of  $\Phi \circ \gamma$ .*

*Proof.* Let  $\gamma(t) = (x(t), y(t), \theta(t))$ ,  $\mathbf{f}(t) = (x(t), y(t))$  its front track and  $\mathbf{v}(t) = (\cos \theta(t), \sin \theta(t))$  the frame direction. By Proposition 4.7, the vertices of  $\gamma$  are the points where the frame is perpendicular to the front track,  $\langle \dot{\mathbf{f}}, \mathbf{v} \rangle = 0$ . Let  $\tilde{\gamma} = \Phi \circ \gamma$ . Then  $\tilde{\mathbf{f}} = \mathbf{f} - 2\mathbf{v}$  and  $\tilde{\mathbf{v}} = -\mathbf{v}$ , hence  $\langle \dot{\tilde{\mathbf{f}}}, \tilde{\mathbf{v}} \rangle = -\langle \dot{\mathbf{f}} - 2\dot{\mathbf{v}}, \mathbf{v} \rangle = -\langle \dot{\mathbf{f}}, \mathbf{v} \rangle$ , since  $\langle \mathbf{v}, \mathbf{v} \rangle = 1$  implies  $\langle \mathbf{v}, \dot{\mathbf{v}} \rangle = 0$ . It follows that vertices of  $\gamma$  and  $\tilde{\gamma}$  occur simultaneously.  $\square$

The statement of Claim 4 is invariant under rigid motions and time reparametrizations, so we can assume, without loss of generality, that  $\gamma(t) = (x(t), y(t), \theta(t))$  satisfies equations (10) with  $a > 0$ ,  $a \neq 1$ ,  $\kappa > 0$  and  $\mathbf{f}_0 = \mathbf{f}(0)$  is a point where  $\kappa_{max}$  occurs. According to Proposition 4.7 and its proof we then have  $\theta_0 = \pi/2$ ,  $\mathbf{v}_0 = (0, 1)$ ,  $\kappa_0 = 1 + a$ ,  $\dot{\mathbf{f}}_0 = (-1, 0)$  and  $\ddot{\mathbf{f}}_0 = (0, -\kappa_0) = -(0, 1 + a)$ .

Now let  $\tilde{\mathbf{f}}(t)$  be the front track of  $\tilde{\gamma} = \Phi \circ \gamma$ . That is,  $\tilde{\mathbf{f}}(t) = \mathbf{f}(t) - 2\mathbf{v}(t)$ .

**Lemma 4.10.** (1)  $\dot{\tilde{\mathbf{f}}}_0 = -\dot{\mathbf{f}}_0 = (1, 0)$ , (2)  $\ddot{\tilde{\mathbf{f}}}_0 = (0, 1 - a)$ .

*Proof.* (1) From equation (10),  $\dot{\theta} = \kappa - as$ . At  $t = 0$ ,  $\kappa_0 = 1 + a$ ,  $\theta_0 = \pi/2$ , hence  $\dot{\theta}_0 = 1$ . Now  $\dot{\mathbf{v}} = \dot{\theta}(-s, c)$ , hence  $\dot{\mathbf{v}}_0 = (-1, 0)$ . Thus  $\dot{\tilde{\mathbf{f}}}_0 = \dot{\mathbf{f}}_0 - 2\dot{\mathbf{v}}_0 = (-1, 0) - 2(-1, 0) = (1, 0)$ .

(2) From equations (10),  $\ddot{\theta} = \dot{\kappa} - a\dot{\theta}c = ac\kappa - ac(\kappa - as) = acs$ , hence  $\ddot{\theta}_0 = 0$ . Thus  $\ddot{\mathbf{v}} = \ddot{\theta}(-s, c) - \dot{\theta}^2(c, s)$  implies  $\ddot{\mathbf{v}}_0 = (0, -1)$ . It follows that  $\ddot{\tilde{\mathbf{f}}}_0 = \ddot{\mathbf{f}}_0 - 2\ddot{\mathbf{v}}_0 = (0, -1 - a) - 2(0, -1) = (0, 1 + a)$ .  $\square$



FIGURE 9. The proof of claim 4 for ‘wide’ (left) and ‘narrow’ (right) front tracks.

We can conclude from the last lemma:

**Proposition 4.11.** *For any bicycle geodesic  $\gamma$  with  $a \neq 0, 1$ , let  $\tilde{\gamma} = \Phi \circ \gamma$ . Then*

- *if  $0 < a < 1$  (wide front track) then the front track of  $\tilde{\gamma}$  is the result of translating the front track of  $\gamma$  along its directrix for half its length;*
- *if  $a > 1$  (narrow front track) then the front track of  $\tilde{\gamma}$  is the result of translating the front track of  $\gamma$  along its directrix for half its length, then reflecting about the directrix.*

See Figure 9.

## 5. THE PROOF OF THEOREM 1.2.

We first prove half of Theorem 1.2, the ‘if’ part, without invoking Theorem 1.1. Doing so illustrates some amusing bicycling mathematics.

By the ‘horizontal lift’ of a front track  $\mathbf{f}(t)$  we mean any bicycle path  $(\mathbf{b}(t), \mathbf{f}(t))$  whose front track projection is the given curve  $\mathbf{f}(t)$ . Since a Euclidean line is a metric line in  $\mathbb{R}^2$ , and since  $\pi_f$  preserves lengths when applied to bicycle paths, every horizontal lift of a Euclidean straight line is a metric line in  $Q$ . However, just because the front wheel moves in a straight line does not mean that the back wheel moves along the same straight line. Indeed, the back wheel typically traces a *tractrix* of width  $\ell$  associated to the linear front track (unless at some moment the back wheel lies on the straight line, in which case the back wheel also travels along the same straight line.) See Figure 10.

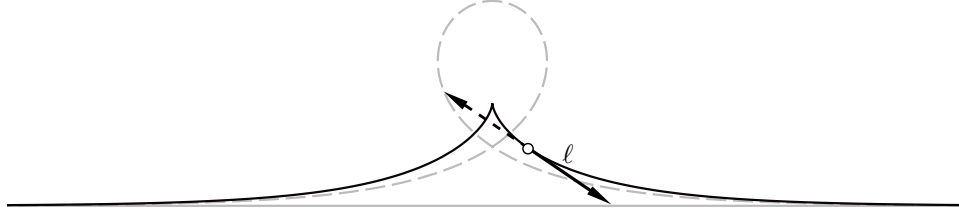


FIGURE 10. A tractrix: the back wheel track (dark solid curve) when the front wheel travels along a straight line (light solid horizontal line). The ‘flipped’ front track is Euler’s soliton (light dashed curve).

**Lemma 5.1.** *Let  $\gamma(t) = (\mathbf{b}(t), \mathbf{f}(t)) \in Q$  be a horizontal lift of the straight line  $\mathbf{f}(t) = (t, 0)$ . Then either  $\mathbf{b}(t) = (t \pm \ell, 0)$  or  $\mathbf{b}(t)$  is a tractrix of width  $\ell$ . Explicitly,*

$$\mathbf{b}(t) = (t - \ell \tanh[(t - t_0)/\ell], \ell \operatorname{sech}[(t - t_0)/\ell]), \quad t_0 \in \mathbb{R}.$$

*The associated Euler soliton, obtained via the involution  $\Phi$ , is*

$$\tilde{\mathbf{f}}(t) = 2\mathbf{b}(t) - \mathbf{f}(t) = (t - 2\ell \tanh[(t - t_0)/\ell], 2\ell \operatorname{sech}[(t - t_0)/\ell]).$$

*Proof.* From Lemma 4.1, the horizontality condition on  $\gamma$  is  $\ell\dot{\theta} + \sin\theta = 0$ . The general solution of this ODE is  $\theta(t) = -2\cot^{-1}(e^{(t-t_0)/\ell})$ , provided that  $-\pi < \theta(0) < 0$ , from which the statement follows. (To get the solutions with  $0 < \theta(0) < \pi$  note that the equation is invariant under  $\theta \rightarrow -\theta$ .)  $\square$

**Remark 5.2.** Note that the tractrix  $\mathbf{b}(t)$  of Lemma 5.1 tends to the earlier solutions  $(t \pm \ell, 0)$ , as the ‘phase parameter’  $t_0 \rightarrow \pm\infty$ .

We continue with the proof of the ‘if’ part of Theorem 1.2. Let  $c : \mathbb{R} \rightarrow \mathbb{R}^2$  be an arc length parametrization of a straight line and  $\gamma : \mathbb{R} \rightarrow Q$  any horizontal lift of  $c$ . As discussed immediately above, any horizontal lift of a metric line must be a metric line, hence  $\gamma$  is a metric line in  $Q$ . The back track of such a lift is either a straight line or a tractrix of width  $\ell$ . In case the back track is a tractrix apply  $\Phi$  to  $\gamma$  and project back to arrive at  $\tilde{c} = \pi_f \circ \Phi \circ \gamma$ . By the last lemma any such  $\tilde{c}$  is an Euler soliton of width  $2\ell$ . Since isometries map metric lines to metric lines the curve  $\tilde{\gamma} = \Phi \circ \gamma$  is a metric line and so the Euler soliton  $\tilde{c}$  is the projection of a metric line.  $\square$

Note that we can construct any Euler soliton of width  $2\ell$  in this way.

**The rest of the proof of Theorem 1.2 (the ‘only if’ part).** We have just proven that all infinite geodesics in  $Q$  whose front tracks are straight lines or Euler solitons of width  $2\ell$  are metric lines. To prove that there are no other metric lines in  $Q$  we invoke Theorem 1.1. According to this theorem it suffices to eliminate all the non-inflectional elasticae, other than the Euler soliton, as front tracks of metric lines in  $Q$ . Our proof follows the idea suggested by Figure 3. Given any non-inflectional elasticae  $\mathbf{f}(t)$ , other than the Euler soliton, where  $t$  is arc-length, we can rigidly rotate it so that its directrix is horizontal, that is,  $\mathbf{f}(t) = (x(t), y(t))$ , where  $y(t)$  is periodic of some period  $T > 0$  and  $x(t+T) = x(t) + L$  for some  $L > 0$ . By translations in  $x, y$  and  $t$  we can further assume that  $\mathbf{f}(0) = (0, 0)$  is a vertex of maximum curvature, so that  $x(0) = y(0) = 0$  and  $\theta(0) = \pi/2$ . (There are explicit expressions for  $x(t)$  and  $y(t)$  in terms of elliptic functions but we will not need these.)

Thus, after one period we have  $\mathbf{f}(T) = (L, 0)$ . Because  $t$  is arc-length, the length of the elastica segment between  $\mathbf{f}(0)$  and  $\mathbf{f}(T)$  is  $T$ . But a straight horizontal line segment is the shortest curve connecting  $\mathbf{f}(0)$  to  $\mathbf{f}(T)$  and its length is  $L$ , so we must have that  $L < T$ .

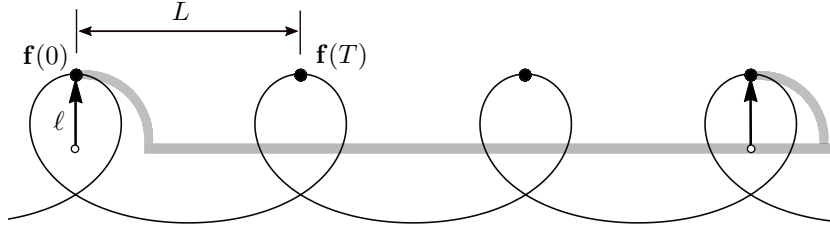


FIGURE 11. A shortcut.

After  $N$  periods the length of the elastica segment between  $\mathbf{f}(0)$  and  $\mathbf{f}(NT)$  is  $NT$ . As shown in Figure 11, we can find a shorter bike path between  $\gamma(0)$  and  $\gamma(NT)$  for  $N$  large enough, as follows: ride along a quarter circle of radius  $\ell$  clockwise without moving the back wheel; then ride along a straight line eastwards a distance of  $NL$ , then a quarter turn counterclockwise. The total length of this path is  $\pi\ell + NL$ . For  $N > \pi\ell/(T - L)$  this is shorter than  $NT$ .  $\square$

**Remark 5.3.** Assuming Theorem 1.1, Theorem 1.2 follows from the results described in [27] where the sub-Riemannian geodesics for a sub-Riemannian metric on  $SE_2$  isometric to our metric on  $Q$  were studied and characterized as solutions



$\theta(t)$  to a family of pendulum equations. Back in our problem, that angle  $\theta$  is the angle the bike frame makes with the  $x$ -axis. In [27] it was proved that a geodesic minimizes for all time if and only if  $\theta(t) \equiv \text{const}$  or  $\theta(t)$  is a non-periodic homoclinic solution of the pendulum problem. These conditions mean that the front track of the bike moves along a straight line or is an Euler soliton.

## 6. LOOSE ENDS AND SCATTERED WHEELS

**6.1. Bicycling Correspondence.** In the first part the proof of Theorem 1.2 (the “if” part) we build the Euler soliton out of a Euclidean line by using the tractrix back-wheel curve as an intermediary step. In the language of [8], the line and the Euler soliton are in “bicycle correspondence” with each other, the tractrix mediating the correspondence. Take any sufficiently smooth front wheel curve  $\mathbf{f}(t)$ . Choose any one of its horizontal lifts  $\gamma(t) = (\mathbf{b}(t), \mathbf{f}(t))$ . There are a circle’s worth of such lifts, corresponding to an initial choice of point  $\mathbf{b}(t_0)$  on the circle of radius  $\ell$  about  $\mathbf{f}(t_0)$ . Apply the ‘flip’ isometry  $\Phi$  of Lemma 3.2 to  $\gamma$ . Project  $\Phi \circ \gamma$  back to the plane to arrive at the new front wheel curve  $\tilde{\mathbf{f}}(t) = 2\mathbf{b}(t) - \mathbf{f}(t)$ , which shares its back wheel track  $\mathbf{b}(t)$  with  $\gamma(t)$ . Then the two front wheel curves  $\mathbf{f}(t)$  and  $\tilde{\mathbf{f}}(t)$  are said to be in bicycle correspondence. There are thus a circle’s worth of bicycle correspondents to  $\mathbf{f}(t)$ , corresponding to the circle’s worth of choices for  $\mathbf{b}(t)$ .

**Question.** Is a bicycle correspondent to a projected geodesic always a projected geodesic?

**No.** The circle of radius  $\ell$  is the projection of a geodesic corresponding to a back track fixed at this circle’s center. Most bicycle correspondents of the circle are not elastica and hence not projections of sub-Riemannian geodesics. It is interesting to note that these correspondents to the circle are, instead, *pressurized elasticae* which means their curvature  $\kappa$  satisfies the ODE  $\ddot{\kappa} + \frac{1}{2}\kappa^3 + A\kappa = C$  with a nonzero constant  $C$ . See Figure 12.

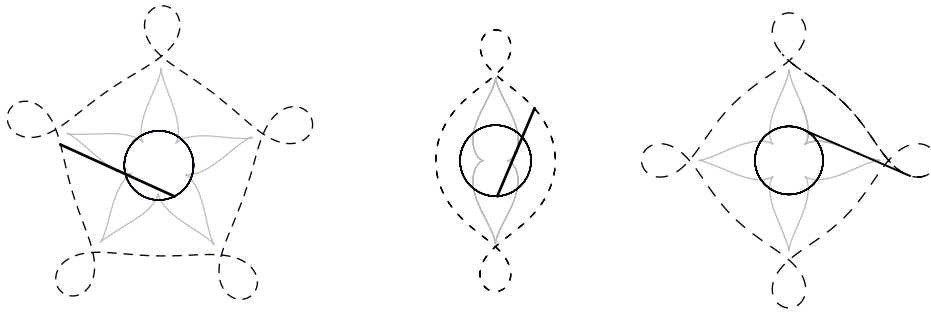


FIGURE 12. Pressurized elasticae (dashed curves), in bicycle correspondence with a circle, sharing a common back track (light curve).

**Question.** Is every horizontal lift of a projected geodesic a geodesic?

**No.** We just saw this above with the case of the circle. Alternatively, see Claim 3 of Section 4.

**6.2. Not of bundle type.** For the most familiar sub-Riemannian submersions  $M \rightarrow B$  the answer to the preceding question is **yes**: every horizontal lift of every projected geodesic is a geodesic. Examples include the Heisenberg group, Carnot groups  $G$  with  $B = G/[G, G]$ , the Hopf fibration examples  $S^3 \rightarrow S^2$  and the various principal bundle examples in [23]. What makes these geometries different from bicycling geometry, group-theoretically speaking, is that for them the group of sub-Riemannian isometries acts transitively on each fiber.

**Definition 6.1.** A sub-Riemannian manifold  $M$  is of *bundle type* if it admits a sub-Riemannian submersion  $\pi : M \rightarrow B$  and a Lie subgroup  $H \subset \text{Isom}(M)$  such that the fibers of  $\pi$  are orbits of  $H$ .

If  $M \rightarrow B$  is of bundle type then, necessarily, every horizontal lift of a projected geodesic is a geodesic. So, our bicycling sub-Riemannian geometry with its front track projection  $\pi_f : Q \rightarrow \mathbb{R}^2$  cannot be of bundle type.

The front track submersion is a principal  $S^1$ -fibration, so that its fibers are the orbits of a free  $S^1$ -action on  $Q$ , but this action cannot be an action by isometries, as we have just seen. To see this fact directly, fix a base point  $\mathbf{q}_0 \in Q$ . Identify  $\text{SE}_2 \simeq Q$ ,  $\mathbf{q}_0 \mapsto g\mathbf{q}_0$ . Then the induced sub-Riemannian structure on  $\text{SE}_2$  is invariant under *left*-translations by  $\text{SE}_2$ , while  $\pi_f : \text{SE}_2 \rightarrow \mathbb{R}^2$  is the quotient by *right*-translations by  $S^1 \subset \text{SE}_2$ , the rotations about  $\pi_f(\mathbf{q}_0) \in \mathbb{R}^2$ .

**Remark 6.2.** In fact, this  $S^1$ -action is not even by contact symmetries: right translation  $R_g$  by an element  $g \in S^1$  defines a map of  $\text{SE}_2$  which does not preserve the contact distribution  $D$ . This failure is easily seen by observing that  $R_g$  acts on a bike path in  $Q$  (a  $D$ -horizontal curve) by rotating the bike frame along the path by a fixed angle, without changing the front track, producing ‘skidding’ of the back wheel.

**6.3. Other models for bicycling geometry.** The bicycling configuration space  $Q$  can be identified,  $\text{SE}_2$ -equivariantly, with  $ST\mathbb{R}^2$ , the space of unit tangent vectors to the plane. Write elements of  $ST\mathbb{R}^2$  as pairs  $(\mathbf{b}, \mathbf{v})$  where  $\mathbf{b} \in \mathbb{R}^2$  and  $\mathbf{v} \in \mathbb{R}^2$  is a unit vector attached at the point  $\mathbf{b}$ . Identify  $\mathbf{b}$  with the location of the back wheel and  $\mathbf{v}$  with the direction of the frame. Then the isomorphism  $ST\mathbb{R}^2 \rightarrow Q$  is

$$(12) \quad (\mathbf{b}, \mathbf{v}) \mapsto (\mathbf{b}, \mathbf{f}), \text{ with } \mathbf{f} = \mathbf{b} + \ell \mathbf{v}$$

The induced contact distribution on  $ST\mathbb{R}^2$ , also denoted as  $D$ , can be described by the condition that its smooth integral curves  $(\mathbf{b}(t), \mathbf{v}(t))$  satisfy  $\dot{\mathbf{b}}(t) \in \mathbb{R}\mathbf{v}(t)$ . Write  $\mathbf{b} = (x_b, y_b)$  and  $\mathbf{v} = (\cos \theta, \sin \theta)$ , to define global coordinates  $(x_b, y_b, \theta)$  on  $ST\mathbb{R}^2$ . In these coordinates a smooth curve  $(x_b(t), y_b(t), \theta(t))$  is horizontal if and only if there is a smooth scalar function  $\lambda(t)$  such that  $\dot{x}_b = \lambda \cos \theta$ ,  $\dot{y}_b = \lambda \sin \theta$ . Eliminating  $\lambda$ , the contact distribution  $D$  is given by the vanishing of the contact form

$$(13) \quad (\sin \theta) dx_b - (\cos \theta) dy_b.$$

The vector fields

$$(14) \quad \mathbf{S} = \text{‘straight ahead’} = (\cos \theta) \partial_{x_b} + (\sin \theta) \partial_{y_b}$$

and

$$(15) \quad \mathbf{T} = \text{‘turn’} = \partial_\theta$$

clearly frame  $D$ .<sup>1</sup> An integral curve of  $\mathbf{S}$  corresponds to the bicycle moving along a straight line passing through the bike frame. An integral curve of  $\mathbf{T}$  corresponds to a circus trick: the back wheel is stationary, marking the center of a circle about which the front wheel traces a circle of radius  $\ell$ . To do this trick, the front wheel must be turned at 90 degrees to the frame. The front wheel tracks of this line and a circle are orthogonal. By basic geometric considerations we see that

$$\langle \mathbf{S}, \mathbf{S} \rangle = 1, \quad \langle \mathbf{S}, \mathbf{T} \rangle = 0, \quad \langle \mathbf{T}, \mathbf{T} \rangle = \ell^2.$$

**A second proof of Lemma 3.2.** Consider the map  $\tilde{\Phi}(x_b, y_b, \theta) = (x_b, y_b, \theta + \pi)$  on  $ST\mathbb{R}^2$ . One computes  $\tilde{\Phi}_*\mathbf{S} = -\mathbf{S}$ ,  $\tilde{\Phi}_*\mathbf{T} = \mathbf{T}$  which shows that  $\tilde{\Phi}$  is a sub-Riemannian *isometry*. In terms of our  $(\mathbf{b}, \mathbf{v})$  representation of  $ST\mathbb{R}^2$  we have

$$(16) \quad \tilde{\Phi}((\mathbf{b}, \mathbf{v})) = (\mathbf{b}, -\mathbf{v}).$$

Rewritten, using the isomorphism (12), the map (16) becomes the map  $\Phi : Q \rightarrow Q$  of the lemma.  $\square$

**6.4. Bicycle parallel transport and hyperbolic geometry.** Associated to a sub-Riemannian submersion  $\pi : M \rightarrow B$  and a path  $c : I \rightarrow B$  we have a parallel transport map. If the initial and final endpoints of  $c$  are  $f_0$  and  $f_1$  then this is a map  $\Psi : \pi^{-1}(f_0) \rightarrow \pi^{-1}(f_1)$ .

**Question.** Is parallel transport for bicycling an isometry between fibers?

**No.** One way to see this is via the following theorem.

**Theorem 6.3** (Foote [14]). *The parallel transport map for bicycling is a linear fractional transformation of  $S^1$ . Every linear fractional transformation can be obtained by parallel transport along some closed curve.*

There is no metric on the circle for which the the group  $\mathrm{PSL}_2(\mathbb{R})$  of linear fractional transformations acts by isometries, so Foote's theorem implies the '**no**' answer above. Again, if  $M$  were of bundle type then the answer to the above question would be **yes**: parallel transport would be an  $H$ -map and hence an isometry.

Let us say a few words about what parallel transport involves for bicycling. Fix a front path  $c$  joining two front wheel locations  $\mathbf{f}_0, \mathbf{f}_1$  in the plane. The fiber  $\pi_f^{-1}(\mathbf{f}_0)$  is the circle of radius  $\ell$  centered at  $\mathbf{f}_0$ . The points of this circle represent all ways of placing the back wheel before bicycling the front wheel along the path  $c$ . Choosing one such placement  $\mathbf{b}_0 \in \pi_f^{-1}(\mathbf{f}_0)$  leads to a unique horizontal lift  $\gamma$  of  $c$  starting at  $\gamma(0) = (\mathbf{b}_0, \mathbf{f}_0)$ . Here we assume that  $c$  is parametrized by the unit interval  $[0, 1]$ . Writing  $\gamma(t) = (\mathbf{b}(t), c(t))$ , we have that the parallel transport of  $\mathbf{b}_0$  along  $c$  is  $\mathbf{b}(1)$ , which is an element in the circle of radius  $\ell$  about  $\mathbf{f}_1$ . So parallel transport, or holonomy, along  $c$  is a diffeomorphism between two circles, one centered at  $\mathbf{f}_0$ , the other at  $\mathbf{f}_1$ . Use translation and scaling to identify each circle with the standard unit circle so that this parallel transport becomes a map of the standard unit circle to itself. Foote's Theorem above might be called the 'first theorem' of 'bicycling mathematics'. See [8].

Although the group  $\mathrm{PSL}_2(\mathbb{R})$  does not act isometrically on the circle, it does act isometrically on the hyperbolic plane. In fact  $\mathrm{PSL}_2(\mathbb{R})$  equals the group of rigid motions of that plane. The configuration space for rolling a hyperbolic plane on

<sup>1</sup>The notation  $\partial_\theta$  stands here for a different vector field from the one in equations (6), because we used there different coordinates,  $(x, y, \theta)$ .

the Euclidean plane can be identified with  $M := \mathrm{PSL}_2(\mathbb{R}) \times \mathbb{R}^2$  and inherits, in a canonical way, a rank 2 sub-Riemannian geometry such that the projection onto  $\mathbb{R}^2$  is a sub-Riemannian submersion of bundle type. [17, 18] prove that its geodesics project to planar elasticae, both inflectional and non-inflectional.

**6.5. A heuristic proof of Theorem 1.1.** Our bicycle configuration space  $Q$  can be identified with the circle bundle associated to  $M \rightarrow \mathbb{R}^2$ , where the structure group  $\mathrm{PSL}_2(\mathbb{R})$  acts on the circle by fractional linear transformations, as per Foote’s theorem. Associated with any plane curve  $c : I \rightarrow \mathbb{R}^2$  we have its hyperbolic rolling parallel transport, an element  $k = k(c) \in \mathrm{PSL}_2(\mathbb{R})$  acting by left multiplication on the fibers. The projection to  $\mathbb{R}^2$  of a sub-Riemannian geodesic on  $M$  solves the following *iso-holonomic problem* (see [24] and [23], Chapter 11, especially Theorem 11.8): among all plane curves  $c$  connecting given points  $\mathbf{f}_0$  to  $\mathbf{f}_1$  and having a fixed hyperbolic transport  $k = k(c) \in \mathrm{PSL}_2(\mathbb{R})$ , find the shortest. Now imagine fixing the bicycle placement as well as the front wheel locations, which is to say, let us fix back wheel locations  $\mathbf{b}_0, \mathbf{b}_1$ , writing them as  $\mathbf{b}_i = \mathbf{f}_i + \ell \mathbf{v}_i$ . Recall that  $k \in \mathrm{PSL}_2(\mathbb{R})$  acts on the unit circle. Now, it may or may not be true that  $k(\mathbf{v}_0) = \mathbf{v}_1$ . If not, let  $k$  vary. Consider all  $k \in \mathrm{PSL}_2(\mathbb{R})$  satisfying  $k(\mathbf{v}_0) = \mathbf{v}_1$ . For any such  $k$  form the corresponding hyperbolic rolling geodesic  $c$  for which  $k(c) = k$ . Now, minimize the lengths of all such  $c$ ’s over all of the  $k$ ’s satisfying the condition that they take  $\mathbf{v}_0$  to  $\mathbf{v}_1$ . The curve achieving this minimizer will be the front wheel projection of a bike geodesic minimizing the length between  $(\mathbf{b}_0, \mathbf{f}_0)$  and  $(\mathbf{b}_1, \mathbf{f}_1)$ , and will also be itself a particular type of hyperbolic rolling geodesic. Since we know by [17, 18] that hyperbolic rolling geodesics project to elasticae, we’re done!

What makes this proof heuristic? For one thing, the set of  $k$ ’s over which we’re minimizing is a non-compact set, so we have no guarantee that the minimum exists. For another thing, the proof does not single out the non-inflectional elasticae from all elasticae.

**6.6. Open Questions.** The bicycle correspondents of a curve  $c$  are the result of the compositions  $c \rightarrow \pi_f \circ \Phi \circ hc$ , where  $hc$  indicates any of the circle’s worth of horizontal lifts of the front track  $c$  and where  $\Phi$  is the flipping isometry of Lemma 3.2. There are a number of hints in [8] that the ‘transformation’ of forming bicycling correspondents shares much in common with the Bäcklund transformations arising in the theory of integrable PDE.

What is the family of curves that we get by forming the bicycle correspondents of elastica? Repeat and form all the bicycle correspondents of all the curves in this new family. What do we get now? Let us call this set the ‘2nd generation’ of correspondents to elastica. Keep going. Does the procedure eventually close up, or, do we get new curves at each generation? Do the curves at the  $n$ -th generation satisfy some “nice” ODE? Are they projections of sub-Riemannian geodesics on some sub-Riemannian geometry constructed iteratively from  $Q$  or  $\mathrm{PSL}_2(\mathbb{R}) \times \mathbb{R}^2$ ?

We could also ride our bicycle on a sphere or hyperbolic plane. This change of bicycling arena corresponds to investigating a left-invariant sub-Riemannian structure on either  $\mathrm{SO}_3$  or  $\mathrm{PSL}_2(\mathbb{R})$ , these being the unit tangent bundles and also the group of rigid motions of the sphere or hyperbolic plane, respectively. (When bicycling on the sphere of radius  $R$  one may need to insist that the frame’s length is not equal to an integer multiple of  $R\pi/2$  to avoid various pathologies. See [6] for interesting relations that might arise between front and back wheel curves when

the spherical frame length is  $R\pi/2$ .) How would our two main theorems change? Are the front wheel projections of sub-Riemannian geodesics still elastica, meaning curves whose geodesic curvatures satisfy equation 1? We guess so, but have not checked and are open to surprises. Would bicycling on these non-Euclidean geometries add to our understanding of how (or if) these different occurrences of elastica in sub-Riemannian geometry are related? Perhaps.

#### APPENDIX A. PROOF OF THEOREM 3.3 ON ISOMETRIES

The group  $\text{SE}_2$  of orientation preserving isometries of the euclidean plane acts freely and transitively by sub-Riemannian isometries on the bicycling configuration space  $Q$ . Fixing a point  $\mathbf{q}_0 \in Q$ , we identify  $\text{SE}_2 \simeq Q$ ,  $g \mapsto g \cdot \mathbf{q}_0$ . This identification is  $\text{SE}_2$ -equivariant, hence induces a left-invariant sub-Riemannian structure on  $\text{SE}_2$ , given by its value at the identity  $e \in \text{SE}_2$ , a 2-dimensional subspace  $D_e \subset \mathfrak{se}_2$ , equipped with an inner product.

To determine the isometry group  $\text{Isom}(\text{SE}_2)$  of this sub-Riemannian structure we use two ingredients: (1) Cartan's equivalence method, applied to the local classification of 3-dimensional sub-Riemannian manifolds of contact type; (2) A calculation of  $\text{Aut}(\mathfrak{se}_2, D_e)$ , the group of automorphism of the Lie algebra of  $\text{SE}_2$  preserving the contact plane at  $e \in \text{SE}_2$  and the inner product.

(1) Let  $M$  be a 3-dimensional sub-Riemannian manifold of contact type (that is,  $D \subset TM$  is bracket generating). Similar to the Riemannian case, one can use the Cartan method of equivalence to construct a canonical connection on  $TM$  and associated curvature tensor, whose vanishing is equivalent to  $M$  being 'flat', that is, locally isometric to the maximally symmetric case, the sub-Riemannian structure induced on  $S^3$  from  $S^2$  via the Hopf fibration  $S^3 \rightarrow S^2$ , admitting a 4-dimensional isometry group (the standard action of  $U_2$  on  $\mathbb{C}^2 \supset S^3$ ). In the non-flat case, such as ours, the equivalence method shows that the isometry group, even the local one, is at most 3-dimensional. It follows that the space  $\mathbf{isom}(M)$  of sub-Riemannian Killing fields (vector fields whose flow acts by sub-Riemannian isometries) is at most 3-dimensional. A good reference for this circle of ideas is [20].

Now let  $G$  be a 3-dimensional connected Lie group with a left-invariant non-flat sub-Riemannian structure of contact type. Let  $\mathfrak{g} = T_e G$  be its Lie algebra, equipped with the Lie bracket coming from the commutator of left-invariant vector fields. Let  $L(G) \subset \text{Isom}(G)$  be the (isomorphic) image of the action of  $G$  on itself by left translations. Then, by part (1) above,  $\dim[\text{Isom}(G)] = 3$ . Let  $\mathfrak{R}(G)$  be the right-invariant vector fields on  $G$ . They generate left translations, hence  $\mathfrak{R}(G) \subset \mathbf{isom}(G)$ . But  $\dim[\mathbf{isom}(G)] \leq 3$ , so  $\mathfrak{R}(G) = \mathbf{isom}(G)$ . Let  $\text{Isom}_e(G)$  be the stabilizer of  $e$  in  $\text{Isom}(G)$ , a discrete subgroup.

**Lemma A.1.**  $\text{Isom}(G) = L(G) \rtimes \text{Isom}_e(G)$ . That is,  $L(G)$  is a normal subgroup of  $\text{Isom}(G)$ ,  $L(G) \cap \text{Isom}_e(G) = \{e\}$  and  $\text{Isom}(G) = L(G)\text{Isom}_e(G)$ .

*Proof.* By our assumptions on  $G$  and dimensionality,  $L(G)$  is the identity component of  $\text{Isom}(G)$  hence is a normal subgroup. If  $L_g \in L(G) \cap \text{Isom}_e(G)$  then  $e = L_g(e) = ge = g$ , hence  $g = e$ . Let  $f \in \text{Isom}(G)$  and  $g = f(e)$ . Then  $L_{g^{-1}} \circ f \in \text{Isom}_e(G)$ , hence  $f \in L(G)\text{Isom}_e(G)$ .  $\square$

**Lemma A.2.** *The map  $\text{Isom}_e(G) \rightarrow \text{GL}(\mathfrak{g})$ ,  $f \mapsto df_e$ , is (a) injective, (b) its image is contained in  $\text{Aut}(\mathfrak{g}, D_e)$ , the group of Lie algebra automorphisms of  $\mathfrak{g}$  preserving  $D_e$  and its inner product.*

*Proof.* (a) An isometry of sub-Riemannian connected manifolds of contact type is determined by its derivative at a single point (one can deduce it from the existence of a canonical Riemannian metric on such a manifold). Hence  $f \mapsto df_e$  is injective.

(b) An isometry of a sub-Riemannian manifold  $M$  acts on its algebra of Killing vector fields  $\mathbf{isom}(M)$  as an automorphism of Lie algebras. In our case,  $\mathbf{isom}(G) = \mathfrak{A}(G)$  and the evaluation map  $\mathfrak{A}(G) \rightarrow \mathfrak{g}$  is a Lie algebra anti-isomorphism, hence  $df_e$  preserves the negative of the Lie bracket on  $\mathfrak{g}$ , and thus the Lie bracket itself, that is,  $df_e \in \text{Aut}(\mathfrak{g})$ . Since  $f$  is a sub-Riemannian isometry and fixes  $e$ , it leaves  $D_e$  invariant, acting on it by isometries.  $\square$

(2) After all these preliminaries, it remains to make some calculations in our case of  $G = \text{SE}_2$ , equipped with a left-invariant sub-Riemannian structure induced by its action on the bicycling configuration space  $Q$ .

First, to show that such a sub-Riemannian structure is non-flat, we note that it is of contact type (see Lemma 4.1) and that an even stronger statement is known to hold; namely, that the CR structure associated to such a sub-Riemannian structure on  $\text{SE}_2$  (they are all equivalent) is not flat (the CR structure associated to a sub-Riemannian structure is obtained by keeping only the conformal structure on  $D$ , ‘forgetting scale’). See for example the calculation in §7 of [9]. This statement was already known to É. Cartan, who classified all homogeneous 3-dimensional CR structures [12]. We conclude that the group of sub-Riemannian isometries  $\text{Isom}(\text{SE}_2)$  is 3-dimensional, where the identity component is generated by left translations of  $\text{SE}_2$  on itself. Alternatively, one can use the last section of [15] to arrive at the same conclusion.

Next, we fix a basis of  $\mathfrak{se}_2$ , given by the following Killing vector fields on  $\mathbb{R}^2$ ,

$$\partial_x, \partial_y, \partial_\theta = x\partial_y - y\partial_x,$$

satisfying

$$[\partial_x, \partial_y] = 0, [\partial_\theta, \partial_x] = -\partial_y, [\partial_\theta, \partial_y] = \partial_x.$$

Next fix  $\mathbf{q}_0 = (\mathbf{b}_0, \mathbf{f}_0) \in Q$ , where  $\mathbf{b}_0 = (-1, 0)$ ,  $\mathbf{f}_0 = (0, 0)$  (we assume  $\ell = 1$ , the general case follows easily from this case by a rescaling argument). In the coordinates  $(x, y, \theta)$  of Section 4.2,  $\mathbf{q}_0$  is given by  $x_0 = y_0 = \theta_0 = 0$ . The actions of  $\partial_x, \partial_y, \partial_\theta$  at  $\mathbf{b}_0$  are  $\partial_x, \partial_y, -\partial_y$ , respectively and the no-skid condition at  $\mathbf{q}_0$  is  $\dot{\mathbf{b}} \parallel \partial_x$ . It follows that  $a\partial_x + b\partial_y + c\partial_\theta \in D_e \subset \mathfrak{se}_2$  if and only if  $b = c$ . Thus  $D_e$  is span by  $X_1 := \partial_x$ ,  $X_2 := \partial_y + \partial_\theta$ . They act at  $\mathbf{f}_0 = (0, 0)$  by  $\partial_x, \partial_y$ , respectively, hence they form an orthonormal basis for  $D_e$ . Let  $X_3 = [X_1, X_2] = \partial_y$ .

**Lemma A.3.**  $\text{Aut}(\mathfrak{se}_2, D_e) = \{id, \varphi_1, \varphi_2, \varphi_1\varphi_2\} = \{id, \varphi_1\} \cdot \{id, \varphi_2\} \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ , where  $\varphi_1, \varphi_2$  are given in the basis  $X_1, X_2, X_3$  by  $\text{diag}(1, -1, -1)$ ,  $\text{diag}(-1, -1, 1)$ , respectively.

*Proof.* One verifies easily that  $\varphi_1, \varphi_2 \in \text{Aut}(\mathfrak{se}_2, D_e)$  and that they generate a group  $\{id, \varphi_1, \varphi_2, \varphi_1\varphi_2\} = \{id, \varphi_1\} \cdot \{id, \varphi_2\} \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ . It remains to show that any element  $\varphi \in \text{Aut}(\mathfrak{se}_2, D_e)$  is in this group. The  $X_1X_3$ -plane (the linear span of  $X_1, X_3$ ) is the only 2-dimensional abelian ideal in  $\mathfrak{se}_2$ , hence is  $\varphi$ -invariant. It follows that the  $X_1$ -axis, the intersection of the  $X_1X_3$ -plane and  $D_e$ , is  $\varphi$ -invariant.

Being an isometry of  $D_e$ ,  $\varphi(X_1) = \varepsilon_1 X_1$ ,  $\varphi(X_2) = \varepsilon_2 X_2$ , with  $\varepsilon_1, \varepsilon_2 \in \{1, -1\}$ . Being an automorphism,  $\varphi(X_3) = \varphi([X_1, X_2]) = [\varphi(X_1), \varphi(X_2)] = \varepsilon_1 \varepsilon_2 X_3$ .  $\square$

Next, we realize  $\text{Aut}(\mathfrak{se}_2, D_e)$  by elements of  $\text{Isom}_e(\text{SE}_2)$ . With each element  $f \in \text{Isom}(Q)$  is associated an element  $\tilde{f} \in \text{Isom}(\text{SE}_2)$  via the identification  $\text{SE}_2 \simeq Q$ ,  $g \mapsto g\mathbf{q}_0$ . For  $g \in \text{SE}_2 \subset \text{Isom}(Q)$ ,  $\tilde{g} = L_g$ . Let  $\rho \in E_2$  be reflection about the  $x$  axis. Using complex notation,  $\rho(\mathbf{z}) = \bar{\mathbf{z}}$ . Let  $g_{\mathbf{u}, \mathbf{w}} \in \text{SE}_2$ ,  $\mathbf{z} \mapsto \mathbf{u}\mathbf{z} + \mathbf{w}$ , where  $\mathbf{u}, \mathbf{w} \in \mathbb{C}$  and  $|\mathbf{u}| = 1$ . Then  $\rho g_{\mathbf{u}, \mathbf{z}_0} = g_{\bar{\mathbf{u}}, \bar{\mathbf{z}}_0} \rho$ . Hence  $\tilde{\rho} \cdot g_{\mathbf{u}, \mathbf{w}} = g_{\bar{\mathbf{u}}, \bar{\mathbf{w}}}$ . Similarly,  $\tilde{\Phi} \cdot g_{\mathbf{u}, \mathbf{w}} = g_{-\mathbf{u}, \mathbf{w} - 2\mathbf{u}}$ .

**Lemma A.4.**  $\varphi_1 = d\tilde{\rho}_e$ ,  $\varphi_2 = d\tilde{f}_e$ , where  $\rho$  is reflection about the  $x$  axis,  $f = \rho'\Phi$ , and  $\rho'$  is the reflexion about the line  $x = -1$ .

*Proof.* This is a routine verification. 1st verify that both  $\rho, f$  leave  $\mathbf{q}_0$  fixed, so  $\tilde{\rho}, \tilde{f} \in \text{Isom}_e(\text{SE}_2)$ . Next check that  $d\tilde{\rho}_e : \partial_x \mapsto \partial_x$ ,  $\partial_y \mapsto -\partial_y$ ,  $\partial_\theta \mapsto -\partial_\theta$ . It follows that  $d\tilde{\rho}_e = \varphi_1$ . Next check that  $d\tilde{f}_e : \partial_x \mapsto -\partial_x$ ,  $\partial_y \mapsto -\partial_y$ ,  $\partial_\theta \mapsto -2\partial_y - \partial_\theta$ . It follows that  $d\tilde{f}_e = \varphi_2$ .  $\square$

**Corollary A.5.** Let  $\Gamma = \{id, \rho, \Phi, \rho\Phi\} \subset \text{Isom}(Q)$ , where  $\rho \in E_2 \setminus \text{SE}_2$  (a reflection about a line). Then  $\Gamma = \{id, \rho\} \cdot \{id, \Phi\} \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$  and  $\text{Isom}(Q) = \text{SE}_2 \rtimes \Gamma$ .

*Proof.* Clearly,  $\text{SE}_2 \cap \Gamma = \{id\}$ , so it remains to show that  $\text{SE}_2 \cdot \Gamma = \text{Isom}(Q)$ . We can assume, by conjugating by an element of  $\text{SE}_2$  that maps the fixed line of  $\rho$  to the  $x$ -axis, that  $\rho$  is the reflection about the  $x$ -axis (the same  $\rho$  as in the last lemma). By the previous lemmas,  $\text{Isom}(Q) = \text{SE}_2 \cdot \Gamma_0$ , where  $\Gamma_0 = \text{Isom}_{\mathbf{q}_0}(Q) = \{id, \rho, \rho'\Phi, \rho\rho'\Phi\}$ . Now  $\rho\Phi \equiv \rho'\Phi$  and  $\Phi \equiv \rho\rho'\Phi \pmod{\text{SE}_2}$ , hence  $\text{SE}_2 \cdot \Gamma = \text{SE}_2 \cdot \Gamma_0 = \text{Isom}(Q)$ .  $\square$

## REFERENCES

- [1] A. Agrachev, D. Barilari, *Sub-Riemannian structures on 3D Lie groups*, J. Dynamical and Control Systems **18** (2012), 21–44.
- [2] A. Agrachev, B. Bonnard, M. Chyba, I. Kupka, *Sub-Riemannian sphere in Martinet flat case*. J. ESAIM: Control, Optimisation and Calculus of Variations (1997), vol. 2, 377–448.
- [3] A. A. Ardentov, Yu. L. Sachkov, *Extremal trajectories in a nilpotent sub-Riemannian problem on the Engel group*, Sbornik: Mathematics **202** (2011), no. 11, 1593–1616.
- [4] A. A. Ardentov, Yu. L. Sachkov, *Conjugate points in Nilpotent Sub-Riemannian Problem on the Engel Group*, Journal of Mathematical Sciences **195** (2013), no. 3, 369–390.
- [5] V. I. Arnol'd, *Mathematical Methods in Classical Mechanics*, 2nd ed., Springer-Verlag, New York, 1989.
- [6] V. I. Arnol'd, *The geometry of spherical curves and quaternion algebra*, Russian Mathematical Surveys, **50** (1995), no. 1, 1–68.
- [7] G. Bor, *Bicycling mathematics* (web animations). <https://www.cimat.mx/~gil/bicycling/>
- [8] G. Bor, M. Levi, R. Perline, S. Tabachnikov, *Tire tracks and integrable curve evolution*, Int. Math. Res. Notes, published online, 2018. <https://academic.oup.com/imrn/article/2020/9/2698/4998816>
- [9] G. Bor, H. Jacobowitz, *Left-Invariant CR Structures on 3-dimensional Lie Groups*, 2019 preprint. <https://arxiv.org/abs/1909.08160>
- [10] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Math. Society, Providence, Rhode Island, 2001.
- [11] L. Capogna, E. Le Donne, *Smoothness of sub-Riemannian isometries*, Amer. J. Math. **138** (2016), no. 5, 1439–1454.
- [12] É. Cartan, *Sur la geometrie pseudo-conforme des hypersurfaces de deux variables complexes*. Part I: Ann. Math. Pura Appl. **11** (1932), 17–90. Part II: Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 2e serie, **1** (1932), 333–354.

- [13] G. Citti, A. Sarti, *A cortical based model of perceptual completion in the roto-translational space*, J. of Math. Imaging and Vision **24** (2006), no. 3, 307–326.
- [14] R. Foote, *Geometry of the Prytz planimeter*, Reports Math. Physics **42** (1998), 249–271.
- [15] R. Hladky, S. Pauls, *Minimal surfaces in the roto-translation group with applications to a neuro-biological image completion model*, J. Math. Imaging Vision **36** (2010), no. 1, 1–27.
- [16] R. Hladky, *Isometries of complemented sub-Riemannian manifolds*, Adv. Geom. **14** (2014), no. 2, 319–352.
- [17] V. Jurdjevic, J. Zimmerman, *Rolling problems on spaces of constant curvature*, in: Lagrangian and Hamiltonian methods for nonlinear control 2006, series *Lect. Notes Control Inf. Sci.* **366** (2007), 221–231, Springer-Verlag, Berlin.
- [18] V. Jurdjevic and J. Zimmerman, *Rolling sphere problems on spaces of constant curvature*, Math. Proc. Cambridge Philos. Soc. **144** (2008), no. 3, 729–747.
- [19] E. Hakavuori, E. Le Donne, *Blowups and blowdowns of geodesics in Carnot groups*, 2018 preprint. <https://arxiv.org/abs/1806.09375>
- [20] K. Huguen, *The Geometry of Sub Riemannian Three-Manifolds*. PhD thesis, Duke, 1995. <https://people.ucsc.edu/~rmont/papers/HughenThesis>
- [21] E. Le Donne, A. Ottazzi, *Isometries of Carnot Groups and Sub-Finsler Homogeneous Manifolds*, J. Geom. Anal. **26** (2016), no. 1, 330–345.
- [22] R. Levien, *The elastica: a mathematical history*. [https://www.levien.com/phd/elastica\\_hist](https://www.levien.com/phd/elastica_hist)
- [23] R. Montgomery, *A Tour of SubRiemannian Geometry*, Mathematical Surveys and Monographs **91**, American Math. Society, Providence, Rhode Island, 2006.
- [24] R. Montgomery, *Isoholonomic Problems and Some Applications*, Comm. Math. Phys. **128** (1990), 565–592.
- [25] I. Moiseev, Yu. Sachkov, *Maxwell strata in sub-Riemannian problem on the group of motions of a plane*, ESAIM: COCV **16** (2010), 380–399.
- [26] Yu. Sachkov, *Conjugate and cut time in the sub-Riemannian problem on the group of motions of a plane*, ESAIM: COCV **16** (2010), 1018–1039.
- [27] Yu. Sachkov, *Cut locus and optimal synthesis in the sub-Riemannian problem on the group of motions of a plane*, ESAIM: COCV **17** (2011), 293–321.
- [28] Yu. Sachkov, *Exponential map in the generalized Dido problem*, Sb. Math. **194** (2003), 1331–1359.
- [29] S. Tabachnikov, *On the bicycle transformation and the filament equation: results and conjectures*, J. Geom. and Phys. **115** (2017), 116–123.
- [30] A. Vershik, V. Gershkovich, *The geometry of the nonholonomic sphere for three-dimensional Lie groups*, in *Global analysis – studies and applications, III*, Springer Lecture Notes in Math. **1334**, 309–331, Springer-Verlag, Berlin, 1988.
- [31] D.A. Singer, *Lectures on elastic curves and Rods*, AIP Conference Proceedings **1002** (2008), no. 3, 3–32. <https://case.edu/artsci/math/singer/publish/elaslecs>

(Ardentov) AILAMAZYAN PROGRAM SYSTEMS INSTITUTE, RUSSIAN ACADEMY OF SCIENCES, PERESLAVL-ZALESSKY, RUSSIA

Email address: [aaa@pereslavl.ru](mailto:aaa@pereslavl.ru)

(Bor) CIMAT, A.P. 402, GUANAJUATO, GTO. 36000, MEXICO

Email address: [gil@cimat.mx](mailto:gil@cimat.mx)

(Le Donne) DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PISA, LARGO B. PONTECORVO 5, 56127 PISA, ITALY, &, UNIVERSITY OF JYVÄSKYLÄ, DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O. BOX (MAD), FI-40014, FINLAND

Email address: [ledonne@msri.org](mailto:ledonne@msri.org)

(Montgomery) MATHEMATICS DEPARTMENT, UNIVERSITY OF CALIFORNIA, SANTA CRUZ, SANTA CRUZ CA 95064

Email address: [rmont@ucsc.edu](mailto:rmont@ucsc.edu)

(Sachkov) AILAMAZYAN PROGRAM SYSTEMS INSTITUTE, RUSSIAN ACADEMY OF SCIENCES, PERESLAVL-ZALESSKY, RUSSIA

Email address: [yusachkov@gmail.com](mailto:yusachkov@gmail.com)