

Imperial College London
Department of Mathematics

Invariant higher-order variational problems: Reduction, geometry and applications

David Meier

7 October 2013

Supervised by Prof. Darryl D. Holm

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy
in Mathematics of Imperial College London and the Diploma of Imperial College London

Abstract

This thesis is centred around higher-order invariant variational problems defined on Lie groups. We are mainly motivated by applications in computational anatomy and quantum control, but the general framework is relevant in many other contexts as well. We first develop a higher-order analog of Euler–Poincaré reduction theory for variational problems with symmetry and discuss the important examples of Riemannian cubics and their higher-order generalisations. The theory is then applied to higher-order template matching and the optimal curves on the Lie group of transformations are shown to satisfy higher-order Euler–Poincaré equations. Motivated by questions of model selection in interpolation problems of computational anatomy, we then study the relationship between Riemannian cubics on manifolds with a group action (‘object manifolds’) and Riemannian cubics on the corresponding group itself. It is shown, for example, that in Type I symmetric spaces only those Riemannian cubics can be lifted horizontally that lie in flat, totally geodesic submanifolds. We then return to higher-order template matching and provide an alternative derivation of the Euler–Lagrange equations using Lagrange multipliers, which leads to a geometric interpretation of the equations in terms of higher-order Legendre–Ostrogradsky momenta. Building on this approach, we develop a variational integrator that respects the geometric properties of continuous-time solution curves. We also derive the corresponding adjoint equations. The remainder of the thesis is concerned with an application to quantum control, namely, to the problem of experimentally steering a quantum system through a series of target states at prescribed times. We show that the Euler–Lagrange equations lead to Riemannian cubic splines on the special unitary group, under whose action the system evolves optimally. Finally, we perform numerical experiments for two-level quantum systems and extend the formalism to the control of coherent states in bosonic multi-particle systems.

Declaration

Originality. I herewith certify that all material in this thesis which is not my own work has been duly acknowledged. Selected results from this thesis have been disseminated in scientific publications as detailed in Chapter 1.4.

Copyright. The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work

To my parents

Acknowledgments

In the process of writing this thesis I have been very fortunate to have the support, mathematical and otherwise, of many people who were colleagues, friends and often both. Without them this thesis would not have been possible.

I would like to thank my supervisor Darryl Holm for his enlightening advice and guidance throughout. I am grateful to François-Xavier Vialard for his kindness and patience during many hours of invaluable discussion. I would like to thank Dorje Brody, Chris Burnett, François Gay-Balmaz and Tudor Ratiu for interesting, enjoyable and successful collaborations. I gratefully acknowledge discussions with Lyle Noakes and a visit to Peter Michor in Vienna. It was a great pleasure to be working alongside, and learning from, the other members of our research group as well as regular visitors: Martins Bruveris, Chris Cantwell, Sehun Chun, Leo Colombo, Jaap Eldering, David Ellis, Henry Jacobs, Laurent Risser, Cesare Tronci and Joris Vankerschaver. I am grateful to Benjamin Doyon, who initially inspired me to pursue a PhD and who has been a wonderful mentor over the last few years. Dorje Brody and Colin Cotter kindly agreed to read my transfer report and examine me. I thank David Martín de Diego and John Gibbons for taking the time to read this thesis and carry out the PhD viva. I would like to extend my gratitude to the Department of Mathematics at Imperial College London for providing a great research environment and for funding my research through a Roth Award. I also thank the Department of Mathematical Sciences at Brunel University, where I put the finishing touches on this thesis.

During the course of my PhD I had the great fortune to learn the technique of Vipassana meditation. I am grateful to all those who helped run the courses I attended at Dhamma Dipa, Dhamma Padhana and Dhamma Sumeru. Many friends, both at home and in England, provided encouragement and welcome distraction. Thank you Andrew, Anton, Bongi, Christian, Daniel, Grace, Max, Nico, Priska and Stefan.

Finally, there are no words to thank my family and Emma for all their love and understanding. They have been my rock throughout this period. I dedicate this thesis to my parents, who have always, always been there for me.

Contents

Abstract	3
Declarations	5
Acknowledgments	9
1 Introduction	15
1.1 Variational principles in physics and applied mathematics	15
1.2 Content of this work	21
1.3 Contributions of this work	23
1.4 Publications	24
2 Invariant higher-order variational problems, Part I	25
2.1 Main content of the chapter	25
2.2 Geometric setting	27
2.2.1 k^{th} -order tangent bundles	27
2.2.2 k^{th} -order Euler–Lagrange equations	28
2.2.3 Examples: Riemannian cubic polynomials and generalisations . . .	29
2.2.4 Quotient space and reduced Lagrangian	33
2.3 Higher-order Euler–Poincaré reduction	34
2.3.1 Quotient map, variations and k^{th} -order Euler–Poincaré equations .	34
2.4 Riemannian cubics and geometric k -splines	38
2.4.1 Computing the second-order Euler–Poincaré equations for Riemannian cubics	41
2.4.2 Bi-invariant metrics and the NHP equation	42
2.5 Higher-order template matching problems	43

2.5.1	Previous work on longitudinal data interpolation in computational anatomy	43
2.5.2	Euler–Lagrange equations for higher-order template matching . . .	44
2.5.3	Two examples of interest for computational anatomy	49
2.5.4	Template matching on the sphere	50
2.6	Final remarks	52
3	Invariant higher-order variational problems, Part II	56
3.1	Main content of the chapter	57
3.2	Geometric setting	59
3.2.1	Group actions	59
3.2.2	Normal metrics	60
3.2.3	The connector of a Riemannian metric	61
3.2.4	Riemannian submersion property	62
3.3	Covariant derivatives	63
3.3.1	Covariant derivatives for normal metrics	64
3.4	Cubics for normal metrics	66
3.4.1	Preparations	66
3.4.2	A generalised variational problem	68
3.4.3	Cubics for normal metrics: Euler–Lagrange equations	68
3.4.4	Cubics on Lie groups revisited	70
3.4.5	Cubics on symmetric spaces	72
3.4.6	Curvature from cubics	77
3.5	Horizontal lifts	78
3.5.1	Symmetric spaces	79
3.5.2	Riemannian submersions	81
3.6	Extended analysis: Reduction by isotropy subgroup	85

3.6.1	Setting	86
3.6.2	First-order Lagrange–Poincaré reduction	88
3.6.3	Example: $\mathbf{G} = \mathbf{SO}(\mathbf{3})$, $\mathbf{Q} = \mathbf{S}^2$	92
3.6.4	Cubics and ballistic curves	93
3.6.5	Second-order Lagrange–Poincaré reduction	95
3.7	Final remarks	99
4	Inexact trajectory planning and inverse problems in the Hamilton–Pontryagin framework	101
4.1	Background and problem formulation	101
4.2	Main content of the chapter	102
4.3	Geometry of the inexact trajectory planning problem	103
4.3.1	Euler–Lagrange equations via Lagrange multipliers	104
4.3.2	Euler–Poincaré equations	108
4.3.3	Geometry of multipliers	109
4.3.4	Momentum conservation and Noether’s theorem	112
4.3.5	Node equations	113
4.3.6	Residual conservation law after partial symmetry breaking	114
4.4	Applications	115
4.4.1	Riemannian cubics revisited	115
4.4.2	Rigid body splines	117
4.4.3	Quantum splines	120
4.4.4	Macromolecular configurations	120
4.5	Geometric discretisation	122
4.5.1	A geometric integrator	123
4.5.2	Geometric properties	126
4.6	Practicalities	128

4.7	Final remarks	130
5	Quantum splines	132
5.1	Some elements of quantum mechanics	132
5.2	Motivation and problem statement	133
5.3	Quantum spline equations	135
5.4	Geometry	138
5.5	Two-level systems	139
5.6	Coherent states	140
5.7	Final remarks	143
6	Conclusions and Outlook	144
A	Probabilistic interpretation of Riemannian cubics	164
B	Gradient calculation via adjoint equations	167
C	Permission documents	170

1 Introduction

We begin with some background on variational principles and their importance in a wide range of applications. We will also give an overview of what is contained in this thesis and summarise its main contributions.

1.1 Variational principles in physics and applied mathematics

First-order variational principles. Much of classical mechanics revolves around the analysis of *Euler–Lagrange equations*, which characterise the time evolution of many physical systems. They arise from *Hamilton’s principle*, which states that the time evolution of the system is a stationary point of an action functional defined as the time integral of a real-valued function. This function, called the *Lagrangian*, is defined on the tangent bundle of configuration space. That is, the curve in configuration space itself and its first time derivative serve as arguments in the Lagrangian. Hamilton’s principle is thus a *first-order variational principle*.

For a classical particle of mass m moving freely in space the Lagrangian equals kinetic energy, $L = \frac{1}{2}m \|\dot{\mathbf{x}}\|^2 = \frac{1}{2}m \sum_{i=1}^3 \dot{x}_i^2$. Hamilton’s principle states that the trajectory of a physical system during the time interval $[t_i, t_f]$ is a stationary point of the action integral $S = \int_{t_i}^{t_f} L dt$ with respect to variations with fixed end points; in short, $\delta S = 0$. This leads to the Euler–Lagrange equations $m\ddot{\mathbf{x}} = 0$. This is Newton’s first law as set out in Volume I of his *Philosophiae Naturalis Principia Mathematica* [1]. Namely, a freely moving particle travels along straight lines at constant speed. For what follows it is important to note that such trajectories are *geodesics* in Euclidean space.

A wide range of closed physical systems can be described in a similar manner. Namely, suppose the configuration space is a Riemannian manifold whose metric γ measures the kinetic energy multiplied by two. That is, the Lagrangian is $L(q, \dot{q}) = \frac{1}{2}\gamma(q)(\dot{q}, \dot{q})$. As is well-known, Hamilton’s principle leads to the geodesic equation for curves on configuration space.¹

The dynamics of a freely moving rigid body (the *Euler top*) can be viewed in this

¹In the proof of Proposition 2.8 later on we will recall the derivation of this important result.

way. Another famous example is ideal incompressible fluid flow [2]. In fact, both of these examples have additional structure. Not only are their configuration spaces Riemannian manifolds, they are also groups: The group of rotations $SO(3)$ in the case of the rigid body, and the group of volume-preserving diffeomorphisms in the case of ideal incompressible fluid flow.² The Riemannian metric in these examples, and hence the Lagrangian L , exhibit certain symmetries with respect to group operations. By consequence, L can be expressed in an equivalent *reduced* form; that is, as a function on the Lie algebra rather than the tangent bundle.

The kinetic energy Lagrangian $L : TSO(3) \rightarrow \mathbb{R}$ of a freely moving rigid body with diagonal inertia tensor I , for example, is given as

$$L(g, \dot{g}) = \frac{1}{2} \gamma(g)(\dot{g}, \dot{g}) = \frac{1}{2} \operatorname{tr} \left((g^{-1} \dot{g})^T K (g^{-1} \dot{g}) \right),$$

where $(g, \dot{g})$ is a tangent vector at $g \in SO(3)$ and K is the diagonal matrix that satisfies $K_{11} + K_{22} = I_{33}$, $K_{11} + K_{33} = I_{22}$ and $K_{22} + K_{33} = I_{11}$ [4]. Evidently, L can be expressed in an equivalent reduced form $\ell : \mathfrak{so}(3) \rightarrow \mathbb{R}$,

$$\ell(\Omega) = \frac{1}{2} \operatorname{tr} (\Omega^T K \Omega),$$

where $\Omega := g^{-1} \dot{g}$ is the left-trivialisation of the tangent vector $(g, \dot{g})$. In this situation we call the Lagrangian L (or the metric γ) *left-invariant*.

In the example of ideal incompressible fluid flow [2, 5] one considers a bounded region D of a Riemannian manifold and its group of volume preserving diffeomorphisms, which we denote by $\operatorname{Diff}_{\operatorname{vol}}(D)$. Its Lie algebra $\mathfrak{X}_{\operatorname{div}}(D)$ consists of all vector fields with zero divergence on D that are tangent to the boundary. The kinetic energy Lagrangian $L :$

² $SO(3)$ is a (finite-dimensional) *Lie group*, that is, its group structure is compatible with the differentiable structure of the underlying finite-dimensional smooth manifold. For diffeomorphism groups the underlying manifold is a so-called *Fréchet manifold*, whose coordinate functions take values in an infinite-dimensional vector space. Accordingly, one speaks of *Fréchet Lie groups* or *infinite dimensional Lie groups*. In this thesis we will restrict ourselves almost exclusively to finite-dimensional systems. The one place where we do encounter diffeomorphism groups, in Section 2.5, the calculations are carried out at the formal level. For a detailed account of the geometry of infinite dimensional Lie groups we refer to [3].

$T\text{Diff}_{\text{vol}}(D) \rightarrow \mathbb{R}$ is

$$L(\phi, \partial_t \phi) = \frac{1}{2} \gamma(\phi)(\partial_t \phi, \partial_t \phi) = \frac{1}{2} \int_D \|(\partial_t \phi) \circ \phi^{-1}\|^2 dx,$$

where dx is the Riemannian volume element and $\|\cdot\|$ is the norm of the Riemannian metric on D . In this case the reduced Lagrangian $\ell : \mathfrak{X}_{\text{div}}(D) \rightarrow \mathbb{R}$ is

$$\ell(u) = \frac{1}{2} \int_D \|u\|^2 dx,$$

where $u := (\partial_t \phi) \circ \phi^{-1}$. Here the trivialisation is by multiplication from the right; hence, we call L (or the metric γ) *right-invariant*.

Whenever a variational principle on a Lie group exhibits this type of symmetry one can carry out *Euler–Poincaré reduction*; this involves taking constrained variations of the reduced Lagrangian and leads to a reduced form of the Euler–Lagrange equations in terms of objects in the *Lie algebra* and its dual space. In geometric mechanics these reduced equations are often called *Euler–Poincaré equations*. In the examples above, the kinetic energy Lagrangians correspond to Riemannian metrics on the respective groups; hence, Hamilton’s principle leads to geodesic equations. Their reduced forms are well-known in physics as *Euler’s equations* for rigid body dynamics and ideal incompressible fluid flow, respectively. The time evolution of many other systems is similarly characterised by geodesic curves on Lie groups with invariant metrics, finite or infinite dimensional: The equations that describe the motion of an ellipsoidal underwater vehicle (with coincident centres of gravity and buoyancy) fall into this framework, as do the equations of magnetohydrodynamics, the periodic Korteweg–de Vries equation or the Camassa–Holm equation. These examples and many more are discussed in [3] and [5] (see in particular the table on page 34 of the former book for an overview).

In some situations the Lie group acts on another manifold, which we will call *object manifold*, and the group action is built into the variational principle in an application-specific manner. If object manifolds are part of the variational description of a physical system, one cannot in general expect a full group symmetry. This effect is called *symmetry breaking*. However, sometimes a residual symmetry with respect to a subgroup persists that is related to isotropy properties of the object manifold.

Consider, for example, a rigid body in a gravitational field (the *heavy top*) [6]. The presence of gravity means that a full rotational symmetry is no longer present in the Lagrangian. This is expressed mathematically by introducing a unit vector in $\mathbb{R}^3$ (the object manifold) pointing in the direction of the gravitational field. Indeed, $L : TSO(3) \rightarrow \mathbb{R}$ is given by

$$L_{\mathbf{e}_z}(g, \dot{g}) = \frac{1}{2} \text{tr} \left((g^{-1} \dot{g})^T K (g^{-1} \dot{g}) \right) - m \lambda (g \boldsymbol{\chi}_0) \cdot \mathbf{e}_z,$$

where m is the mass of the rigid body; λ is the gravitational constant; and $\boldsymbol{\chi}_0 \in \mathbb{R}^3$ is the vector originating at the point of support and ending at the body's centre of mass, as seen in spatial coordinates when the body is in the configuration corresponding to $g = e$. We denote by $g \boldsymbol{\chi}_0$ the standard action of g on $\boldsymbol{\chi}_0$ (via matrix multiplication). The Lagrangian remains invariant with respect to left-multiplication by elements of the isotropy subgroup of $\mathbf{e}_z$. As a consequence of this S^1 symmetry, the z -component of spatial angular momentum is conserved.

The redundant degree of freedom associated with this residual symmetry can be removed by the method of *Lagrange–Poincaré reduction* [7], which will play an important role in this thesis and will be discussed in detail later. There is another, equivalent, reduction method, which is often called *Euler–Poincaré reduction with advected quantities* [6]. The idea is the following: If we define $\mathcal{L}(g, \dot{g}, \mathbf{x}) := L_{\mathbf{x}}(g, \dot{g})$, then $\mathcal{L}$ is $SO(3)$ -invariant in the sense that $\mathcal{L}(hg, h\dot{g}, h\mathbf{x}) = \mathcal{L}(g, \dot{g}, \mathbf{x})$, for all $h \in SO(3)$. This means that a variant of Euler–Poincaré reduction can be carried out, whereby both the object manifold $\mathbb{R}^3$ and the associated group action are crucial ingredients.

Many other physical systems allow variational descriptions that involve object manifolds in a very similar manner. A number of them are discussed in [6], where the object manifolds are vector spaces representing quantities advected by fluid flow. The analogous formalism with general object manifolds is treated in [8], which is motivated by the theory of liquid crystals. This reference also discusses the relation between the two approaches mentioned above: Euler–Poincaré reduction on the one hand and Lagrange–Poincaré reduction on the other.

Object manifolds are also central in image matching problems in *computational anatomy*, where they appear in the form of *shape spaces*. Computational anatomy is concerned with

quantitative comparisons of shape, in particular the shapes of organs in the human body [9]. In computational anatomy, shapes are defined by spatial distributions of various types of geometric data structures, such as points (landmarks), spatially embedded curves or surfaces (boundaries), or tensors that encode local orientation of muscle fibres, etc. A fruitful approach in this burgeoning field applies the large deformation matching (LDM) method; see [10] for an excellent introduction. In the LDM method shapes are compared by measuring the relative deformation required to optimally match one shape to another under the action of a diffeomorphism group [11, 12]. The optimality condition is formulated in terms of a first-order variational problem involving both the diffeomorphism group and the object manifold of data structures (or shapes). Specifically, typical cost functionals are of the type

$$E[u] = \frac{1}{2} \int_0^1 \|u\|^2 dt + d^2 \left(g(1)S_0, S_{\text{target}} \right), \quad (1.1)$$

where S_0 and S_{target} are initial and target shapes; u is a time-dependent vector field whose flow at time 1 is denoted by $g(1)$; $\|\cdot\|$ is a norm on the space of vector fields associated with a right-invariant Riemannian metric on the diffeomorphism group; and d is a distance function on shape space. Therefore, the first part of the cost functional is the integral of kinetic energy $\frac{1}{2}\|u\|^2$, while the second part measures the ‘mismatch’ between the transformed shape $g(1)S_0$ and the target shape S_{target} . The resulting optimal flow $g(t)$ is a geodesic on the diffeomorphism group with respect to the right-invariant metric, whereas the corresponding curve $g(t)S_0$ on shape space is a geodesic relative to a metric induced by the group action. We will return to this point shortly.

Higher-order variational principles. The Lagrangians of *higher-order variational principles* depend not just on positions and velocities, but also on acceleration and sometimes higher-order derivatives of the curve in configuration space. A natural arena for such higher-order variational principles is in *optimal control theory*, where one studies the evolution of physical systems under the influence of external controls. The task is to select optimal controls, that is, controls that optimise a given cost functional, under a set of constraints such as end point constraints on the trajectory. Both cost functional and constraints may depend on higher-order derivatives; we refer to [13, Chapter 7] for more information.

Our main motivation to study higher-order variational problems lies with potential applications to *longitudinal studies* in computational anatomy. Longitudinal studies in computational anatomy seek, among other goals, to determine a path that interpolates optimally through a *time-ordered series* of images, or shapes. Depending on the specific application, the interpolant will be required to have a certain degree of spatiotemporal smoothness. For example, the pairwise geodesic matching procedure described above can be extended to piecewise-geodesic interpolation through several shapes, as in [14, 15]. If a higher degree of smoothness is required, a natural approach is to investigate higher-order variational formulations of the interpolation problem [16].

As in the first-order case, systems whose configuration spaces are Riemannian manifolds are of special interest. In particular, the study of *Riemannian cubics* and their higher-order generalisations originated in [17], [18] and [19]. Riemannian cubics are solutions of Euler–Lagrange equations for a certain second-order variational problem in a finite-dimensional Riemannian manifold, to find a curve that interpolates between two points with given initial and final velocities, subject to minimal mean-square covariant acceleration. Riemannian cubics appear naturally in the solution of interpolation problems on Riemannian manifolds. In [19], for example, several of them are joined together in such a way that the resulting curve, which is then called a *Riemannian cubic spline*, is twice continuously differentiable. The mathematical theory of Riemannian cubics was subsequently developed in a series of papers including [20, 21, 22, 23, 24, 25, 26]. Engineering applications are discussed in [27, 28, 29, 30, 31], amongst others. Group structures, besides the Riemannian one, play an important role in these applications. For instance, [28] works with the group of special Euclidean transformations $SE(3)$ to develop an interpolation method for the motion of a rigid body, taking into account both spatial and rotational displacements.

Just as for first order, object manifolds (manifolds acted on by Lie groups) can encode important aspects of higher-order variational problems. For example, the constraints of control theoretic applications may be defined in terms of object manifolds. As an illustration, consider the problem of rotating a sphere in such a way that the north pole passes through a number of prescribed locations. The constraints are given in terms of points on an object manifold (the sphere) acted on by the group of rotations.

In the presence of object manifolds one may ask about the relationship between curves $g(t)$ on the Lie group and corresponding curves $g(t)q$ on the object manifold. For first-order variational problems this relationship is well understood. As we mentioned above, in the LDM method of computational anatomy, for instance, one computes an optimal path in the group of diffeomorphisms that carries an initial shape into a target shape. This path is a (horizontal) geodesic in the diffeomorphism group with respect to a right-invariant metric defined by the norm $\|\cdot\|$ used in the cost functional (1.1). The corresponding path in the object manifold (shape space) is a geodesic as well, with respect to a metric that is induced by the action of the diffeomorphism group. This *horizontal lifting property of geodesics* has been crucial in the understanding and the numerics of LDM. In particular, the geodesic flows on the diffeomorphism group are encoded by their initial momenta, and horizontality means that only momenta of a specific form are permitted. For instance, landmark-based geodesic image matching naturally summons the singular momenta that were introduced as solitons for shallow water waves on the real line in [32] and then characterised as singular momentum maps in any number of dimensions in [33]. We refer to [34, 35, 10, 36] for further details.

As we will see, the connection between horizontality and momentum maps is a fundamental one and will play a key role in this thesis.

1.2 Content of this work

At the start of each subsequent chapter is a detailed outline of contents. At this stage we only give a high-level overview over the main topics.

Chapter 2 begins our treatment of higher-order variational principles with a focus on those defined on Lie groups. For cases with symmetry we establish a higher-order reduction theory of the Euler–Poincaré type. We treat in detail the example of Riemannian cubics and their higher-order generalisations. The rest of the chapter is devoted to an application of these ideas to higher-order template matching, where we first encounter object manifolds. We then present numerical simulations in a finite-dimensional example, namely, fitting a curve to a series of points on the sphere.

In Chapter 3 we dive much more deeply into some of the questions that arise in the

presence of object manifolds. For a particular class of variational curves, Riemannian cubics, we ask how those on Lie groups are related to those on object manifolds. To begin with we introduce a natural metric on object manifolds, the so-called *normal metric*, and derive a new form of the equations for Riemannian cubics. The aforementioned connection between horizontality and momentum maps will be at the heart of the derivation. The form of the equations naturally lends itself to the analysis of horizontal lifts of cubics. For example, we show that in Type I symmetric spaces the presence of curvature is prohibitive to horizontal lifts in that precisely those Riemannian cubics can be lifted horizontally to the group of isometries that lie in a flat, totally geodesic submanifold. We also consider horizontal lifts of cubics in the more general framework of Riemannian submersions. In the rest of the chapter we include non-horizontal curves into our considerations and derive an expression of the obstruction for a Riemannian cubic on the group to project to a cubic on the object manifold. For symmetric spaces we also show that certain non-horizontal geodesics on the group of isometries project to Riemannian cubics.

In Chapter 4 we return in more generality to the higher-order template matching problem first encountered in Chapter 2. We give an alternative, more direct, derivation of the equations of motion using Lagrange multipliers. This has the added advantage of providing a geometric interpretation of the equations in the framework of higher-order Hamiltonian mechanics. We re-interpret the template matching problem as a type of inverse problem and discuss some examples. We then use the Lagrange multiplier approach to derive a discrete variational integrator for the template matching problem, which respects the geometric features of continuous-time solutions. In the numerical solution of the associated minimisation problem it is crucial to have an efficient method for computing gradients. To this end we derive the adjoint equations for the integrator (in Appendix B).

In Chapter 5 we discuss an application to time-dependent quantum control, namely, the problem of steering a finite-dimensional quantum system through a series of target states at prescribed times. We derive the Euler–Lagrange equations and show that they produce Riemannian cubic splines on the special unitary group, under the action of which the initial state evolves optimally. We carry out numerical experiments for two-level quantum systems. As another application we consider quantum control of coherent states

in bosonic multi-particle systems. The chapter is self-contained in that it can be read before or after Chapter 4.

Finally, in Chapter 6 we summarise the main results of the thesis and provide some perspectives on possible future research directions.

1.3 Contributions of this work

The main contributions of this thesis are the following.

- The higher-order Euler–Poincaré reduction theory is developed and shown to lead to a streamlined derivation of the reduced equations for Riemannian cubics on Lie groups.
- The higher-order template matching problem is formulated and Euler–Lagrange equations are derived. This generalises the work of [37] in two directions. Namely, we interpolate through *several* data points, and we allow for cost functionals that depend on higher-order derivatives of the control vector field.
- The relationship between Riemannian cubics on object manifolds and those on Lie groups is studied. Four main results are derived in this context: It is shown that in Type I symmetric space only cubics in flat totally geodesic submanifolds can be lifted horizontally to the group of isometries; a necessary and sufficient condition for horizontal lifts of cubics is given for Riemannian submersions; for symmetric spaces it is shown that certain non-horizontal geodesics on the group of isometries project to Riemannian cubics; and an expression is derived that reveals the obstruction for a Riemannian cubic on the group to project to a cubic on the object manifold.
- An alternative derivation is given of the Euler–Lagrange equations for the higher-order template matching problem using Lagrange multipliers. The approach includes in particular a variational derivation of the well-known higher-order Legendre–Ostrogradsky transform of [38].
- A geometric discretisation of the Euler–Lagrange equations for the higher-order template matching problem is derived from a variational perspective, together with

the associated adjoint equations.

- The quantum spline problem is formulated. The Euler–Lagrange equations are computed, simulated for two-level systems and applied to coherent state control in bosonic multi-particle systems.

1.4 Publications

Some of the results of this thesis have appeared in the following papers, referenced in the rest of the thesis as [39], [40], [41] and [42], respectively.

- F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu, and F.-X. Vialard. Invariant higher-order variational problems. *Communications in Mathematical Physics*, **309**(2):413–458, 2012
- F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu, and F.-X. Vialard. Invariant higher-order variational problems II. *Journal of Nonlinear Science*, **22**(4):553–597, 2012.
- D. C. Brody, D. D. Holm, and D. M. Meier. Quantum splines. *Physical Review Letters*, **109**:100501, 2012.
- C. L. Burnett, D. D. Holm, and D. M. Meier. Inexact trajectory planning and inverse problems in the Hamilton–Pontryagin framework. *Proceedings of the Royal Society A*, **469**:20130249, 2013.

2 Invariant higher-order variational problems, Part I

In this chapter we begin our investigation of invariant higher-order variational problems on Lie groups. After some preliminary work we derive the higher-order generalisation of Euler–Poincaré equations with a particular focus on Riemannian cubics, which are an important example. We then formulate a *higher-order template matching problem* motivated by potential applications to longitudinal studies in computational anatomy. We treat the problem in a general framework that highlights the importance of momentum maps and will serve as an anchor point in subsequent chapters. The results presented here have been published in [39].

The next section gives a more detailed overview of the present chapter.

2.1 Main content of the chapter

The main content is outlined as follows:

Section 2.2 discusses the geometric setting for the present investigation of extensions of group-invariant variational principles to higher order. In particular, Section 2.2.1 summarises the definition of higher-order tangent bundles mainly following the treatment in [7] for the geometric formulation of Lagrangian reduction.

Section 2.3 explains the quotient map for higher-order Lagrangian reduction by symmetry and uses it to derive the k^{th} -order Euler–Poincaré equations.

Section 2.4 takes advantage of the k^{th} -order Euler–Poincaré equations to derive the equations for Riemannian cubics on a Lie group and their higher-order generalisations.

Section 2.5 addresses theoretical and numerical results for our main motivation, longitudinal data interpolation. That is, interpolation through a sequence of data points. After a brief account of the previous work done in computational anatomy, we derive the equations that generalise the equations for geodesic template matching [37] to the case of *higher-order* cost functionals and sequences of *several* data points. We recover in particular the higher-order Euler–Poincaré equations. For a particular choice of cost functionals one can therefore think of the higher-order template

matching approach as *template matching by Riemannian cubics*. We discuss the gain in smoothness afforded by the higher-order approach, then we provide a qualitative discussion of two Lagrangians that are of interest for applications in computational anatomy. Finally, we close the section by demonstrating the higher-order approach to template matching in the finite dimensional case by interpolating a sequence of points on the sphere S^2 , using $SO(3)$ as the Lie group of transformations. This yields the template-matching analog of the so-called NHP equation of Noakes, Heinzinger and Paden [18], which we first encounter in (2.44) below. The results are shown as curves on the sphere in Figures 2.3. A sample figure is shown below to explain the type of results we obtain.

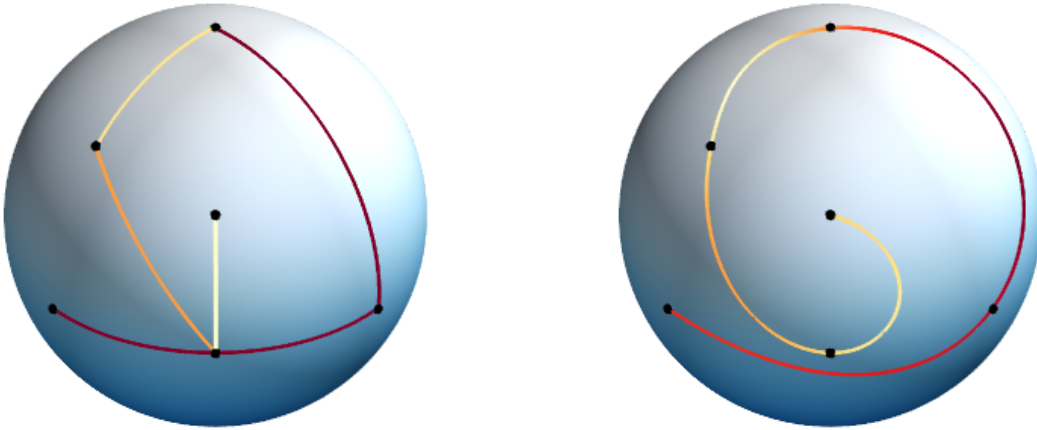


Fig. 2.1: First-order *vs.* second-order template matching results interpolating a sequence of evenly time-separated points on the sphere, using a bi-invariant metric on the rotation group $SO(3)$. The colours show the local speed along the curves on the sphere (white smaller, red larger). The motion slows as the curve tightens. *This figure appears in [39] – reproduced with kind permission from Springer Science and Business Media.*

Section 2.6 is used to summarise the results of the chapter and give an outlook. In particular we motivate some of the developments in the remainder of the thesis.

2.2 Geometric setting

Here we review the definition of higher-order tangent bundles. For more details and explanations of the geometric setting for higher-order variational principles one may consult one of the following references: [7, Chap. 3.2], [43, Chap. 1.3] or [44].

2.2.1 k^{th} -order tangent bundles

Let us define the k^{th} -order tangent bundle $\tau_Q^{(k)} : T^{(k)}Q \rightarrow Q$ of a smooth manifold Q . First, the fibre $T_{q_0}^{(k)}Q$ for any $q_0 \in Q$ is defined as a set of equivalence classes of curves originating at q_0 , as follows: Two curves $q_1(t), q_2(t)$ (both originating at q_0) are *equivalent*, if and only if their time derivatives at $t = 0$ up to order k coincide in any local chart. That is, $q_1^{(l)}(0) = q_2^{(l)}(0)$, for $l = 0, \dots, k$. The equivalence class of a given curve $q(t)$ is denoted by $[q]_{q_0}^{(k)}$. One then defines $T^{(k)}Q := \bigcup_{q_0 \in Q} T_{q_0}^{(k)}Q$, with projection

$$\tau_Q^{(k)} : T^{(k)}Q \rightarrow Q, \quad [q]_{q_0}^{(k)} \mapsto q_0. \quad (2.1)$$

Note that $T^{(0)}Q = Q$ and $T^{(1)}Q = TQ$. Of the bundles $T^{(k)}Q$ only the tangent bundle (i.e., $k = 1$) is a vector bundle.

Given a curve $q(t)$ one defines the k^{th} -order tangent element at time t to be

$$[q]_{q(t)}^{(k)} := [h]_{h(0)}^{(k)}, \quad \text{where } h : \tau \mapsto q(t + \tau). \quad (2.2)$$

We will sometimes use the coordinate notation $(q(t), q^{(1)}(t), \dots, q^{(k)}(t))$ to denote $[q]_{q(t)}^{(k)}$. Often, one writes $q^{(1)}(t) = \dot{q}(t)$ and $q^{(2)}(t) = \ddot{q}(t)$. For more information on higher-order tangent bundles see [7].

A smooth map $f : M \rightarrow N$ induces a map between k^{th} -order tangent bundles,

$$T^{(k)}f : T^{(k)}M \rightarrow T^{(k)}N, \quad [q]_{q_0}^{(k)} \mapsto [f \circ q]_{f(q_0)}^{(k)}. \quad (2.3)$$

Therefore, a group action $\Phi : G \times Q \rightarrow Q$ on the base manifold lifts to a group action on the k^{th} -order tangent bundle,

$$\Phi^{(k)} : G \times T^{(k)}Q \rightarrow T^{(k)}Q, \quad \Phi_g^{(k)} : [q]_{q_0}^{(k)} \mapsto [\Phi_g \circ q]_{\Phi_g(q_0)}^{(k)}. \quad (2.4)$$

The case of a Lie group. The k^{th} -order tangent bundle $T^{(k)}G$ of a Lie group G carries a natural Lie group structure: If $[g]_{g_0}^{(k)}$, and $[h]_{h_0}^{(k)}$ are classes of curves g and h in G , define $[g]_{g_0}^{(k)}[h]_{h_0}^{(k)} := [gh]_{g_0h_0}^{(k)}$. The Lie algebra $T_e T^{(k)}G$ of $T^{(k)}G$ can be naturally identified, as a vector space, with $(k+1)\mathfrak{g}$ (that is, the direct sum of $k+1$ copies of $\mathfrak{g}$) which, therefore, carries a unique Lie algebra structure such that this identification becomes a Lie algebra isomorphism.

To avoid confusion let us state explicitly that throughout this thesis we will use both $\frac{d}{dt}$ and, interchangeably, the more compact notation ∂_t to denote time derivatives. A similar remarks holds for covariant derivatives, which will be important later, where we write $\frac{D}{Dt}$ or, equivalently, D_t .

2.2.2 k^{th} -order Euler–Lagrange equations

A k^{th} -order *Lagrangian* is a smooth function $L : T^{(k)}Q \rightarrow \mathbb{R}$. In the higher-order generalisation of Hamilton’s principle, one seeks a critical point of the functional

$$\mathcal{J}[q] := \int_0^1 L(q(t), \dot{q}(t), \dots, q^{(k)}(t)) dt \quad (2.5)$$

with respect to deformations $q(t, \varepsilon)$ of the curve $q(t)$ fixing the endpoints up to order $k-1$; i.e., $q^{(l)}(0, \varepsilon) = q^{(l)}(0) = q_0^{(l)}$, $q^{(l)}(1, \varepsilon) = q^{(l)}(1) = q_1^{(l)}$ for all ε in an open interval of $\mathbb{R}$ containing the origin and all $l = 0, \dots, k-1$; where $(q_0, q_0^{(1)}, \dots, q_0^{(k-1)}) = [q]_{q_0}^{(k-1)} \in T_{q_0}^{(k-1)}Q$ and $(q_1, q_1^{(1)}, \dots, q_1^{(k-1)}) = [q]_{q_1}^{(k-1)} \in T_{q_1}^{(k-1)}Q$ are the given boundary conditions. The boundary conditions imply that the variations $\delta q^{(l)}(0) = 0$, $\delta q^{(l)}(1) = 0$, for all $l = 0, \dots, k-1$, where we shall use from now on the usual δ -notation for variations, i.e., if $\varepsilon \mapsto f_\varepsilon$ is a deformation of a quantity $f = f_0$, define

$$\delta f := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} f_\varepsilon. \quad (2.6)$$

With this notation, we have the following formulation of Hamilton’s Principle on k^{th} -order tangent bundles. Its proof is a straightforward verification that can be carried out in coordinates.

Theorem 2.1. *There is a unique bundle map*

$$\mathcal{EL}(L) : T^{(2k)}Q \rightarrow T^*Q$$

covering Q such that, for any deformation $q(t, \varepsilon)$, keeping the endpoints fixed, we have

$$\delta \mathcal{J}[q] = \int_0^1 \langle \mathcal{EL}(L)(q(t), \dots, q^{(2k)}(t)), \delta q(t) \rangle dt,$$

where

$$\delta \mathcal{J}[q] := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_0^1 L(q(t, \varepsilon), \dot{q}(t, \varepsilon), \dots, q^{(k)}(t, \varepsilon)) dt$$

and we wrote $\langle \cdot, \cdot \rangle$ for the duality pairing. The cotangent bundle-valued map $\mathcal{EL}(L)$ is called the k^{th} -order Euler–Lagrange operator and has the following expression in standard local charts

$$\langle \mathcal{EL}(L)(q, \dots, q^{(2k)}), \delta q \rangle = \left\langle \sum_{j=0}^k (-1)^j \frac{d^j}{dt^j} \frac{\partial L}{\partial q^{(j)}}, \delta q \right\rangle,$$

where, on the right hand side, it is understood that one formally takes the time derivatives and then replaces expressions of the form $\frac{d^j}{dt^j} q^{(l)}$ by $q^{(l+j)}$ for $l = 0, \dots, k$ and $j = 0, \dots, k$.

Thus, a curve $q(t)$ respecting the endpoint conditions satisfies Hamilton’s principle if and only if it satisfies the k^{th} -order Euler–Lagrange equations $\mathcal{EL}(L)(q(t), \dots, q^{(2k)}(t)) = 0$, which in standard local coordinates of $T^{(2k)}Q$ are

$$\sum_{j=0}^k (-1)^j \frac{d^j}{dt^j} \frac{\partial L}{\partial q^{(j)}} = 0. \quad (2.7)$$

2.2.3 Examples: Riemannian cubic polynomials and generalisations

Riemannian cubics, as introduced in [17, 18] and [19], generalise cubic polynomials in Euclidean space to Riemannian manifolds. Let (Q, γ) be a Riemannian manifold and denote by $\frac{D}{Dt}$ or, interchangeably, D_t , the covariant derivative with respect to the Levi–Civita connection ∇ for the metric γ . Also, let $\|\cdot\|$ be the norm induced by γ .

We recall the bundle isomorphisms $\flat : TQ \rightarrow T^*Q$ and $\sharp := \flat^{-1}$. The definition of $\flat$ is given, on each fibre $T_q Q$, by the identity

$$\langle v_q^\flat, w_q \rangle = \gamma_Q(q)(v_q, w_q), \quad \text{for arbitrary } v_q, w_q \in T_q Q.$$

Consider Hamilton's principle (2.5) for $k = 2$ with the Lagrangian $L : T^{(2)}Q \rightarrow \mathbb{R}$ given by

$$L(q, \dot{q}, \ddot{q}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} \right\|_q^2. \quad (2.8)$$

This Lagrangian is indeed well-defined on the second-order tangent bundle $T^{(2)}Q$, since in coordinates

$$\frac{D}{Dt} \dot{q}^k = \ddot{q}^k + \Gamma_{ij}^k(q) \dot{q}^i \dot{q}^j, \quad (2.9)$$

where $\Gamma_{ij}^k(q)$ are the Christoffel symbols at the point q . Denoting by R the curvature tensor defined by $R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$ for any vector fields $X, Y, Z \in \mathfrak{X}(Q)$, the Euler–Lagrange equation is ([18])

$$\frac{D^3}{Dt^3} \dot{q}(t) + R \left(\frac{D}{Dt} \dot{q}(t), \dot{q}(t) \right) \dot{q}(t) = 0. \quad (2.10)$$

A solution of this equation is called a *Riemannian cubic*, or *cubic* for short.

For completeness, let us derive (2.10). We have

$$\begin{aligned} \delta \int_0^1 L(q, \dot{q}, \ddot{q}) dt &= \delta \int_0^1 \frac{1}{2} \gamma(q) \left(\frac{D}{Dt} \dot{q}, \frac{D}{Dt} \dot{q} \right) dt \\ &= \int_0^1 \gamma(q) \left(\frac{D}{D\varepsilon} \frac{D}{Dt} \dot{q}, \frac{D}{Dt} \dot{q} \right) dt = \int_0^1 \gamma(q) \left(\frac{D}{Dt} \frac{D}{D\varepsilon} \dot{q} + R(\delta q, \dot{q}) \dot{q}, \frac{D}{Dt} \dot{q} \right) dt \\ &= \int_0^1 \gamma(q) \left(\frac{D^2}{Dt^2} \delta q + R(\delta q, \dot{q}) \dot{q}, \frac{D}{Dt} \dot{q} \right) dt = \int_0^1 \gamma(q) \left(\frac{D^3}{Dt^3} \dot{q} + R \left(\frac{D}{Dt} \dot{q}, \dot{q} \right) \dot{q}, \delta q \right) dt, \end{aligned}$$

where we used several standard identities of Riemannian geometry; they can be found, for example, in [45, Part II] or the more recent book [46]. Specifically, for the second equality we used the compatibility of the Levi–Civita connection with the metric³; for the fourth the symmetry of the connection⁴; and for the fifth a symmetry property of the curvature

³Compatibility implies that for any two vector fields V, W along a curve $q(t)$ one has

$$\frac{d}{dt} \gamma(q)(V, W) = \gamma(q)(D_t V, W) + \gamma(q)(V, D_t W).$$

See, for example, [46, Lemma 5.2].

⁴Symmetry of the connection implies that for a variation $q(t, \varepsilon)$ one has [46, Lemma 6.3]

$$\frac{D}{D\varepsilon} \frac{d}{dt} q(t, \varepsilon) = \frac{D}{Dt} \frac{d}{d\varepsilon} q(t, \varepsilon).$$

tensor⁵. Therefore, the Euler–Lagrange operator $\mathcal{EL}(L) : T^{(4)}Q \rightarrow T^*Q$ is given by

$$\mathcal{EL}(L)(q, \dot{q}, \dots, q^{(4)}) = \left[\frac{D^3}{Dt^3} \dot{q} + R \left(\frac{D}{Dt} \dot{q}, \dot{q} \right) \dot{q} \right]^b, \quad (2.11)$$

which implies (2.10).

Remark 2.2. *Riemannian cubics appear naturally in typical interpolation problems on Riemannian manifolds. A problem discussed in [19], for example, seeks a curve that has prescribed initial and final velocities and interpolates given points on a Riemannian manifold, subject to minimal mean-square covariant acceleration. The solution curve is called a Riemannian cubic spline and consists of several Riemannian cubics joined together, to form a curve that is twice continuously differentiable.*

We also mention two generalisations of Riemannian cubics. The first one consists of the class of *geometric k-splines* [20] for $k \geq 2$ with Lagrangian $L : T^{(k)}Q \rightarrow \mathbb{R}$,

$$L(q, \dot{q}, \dots, q^{(k)}) = \frac{1}{2} \left\| \frac{D^{k-1}}{Dt^{k-1}} \dot{q} \right\|_q^2. \quad (2.12)$$

Note that the case $k = 2$ recovers the Riemannian cubics. The Euler–Lagrange equations are [20]

$$\frac{D^{2k-1}}{Dt^{2k-1}} \dot{q}(t) + \sum_{j=2}^k (-1)^j R \left(\frac{D^{2k-j-1}}{Dt^{2k-j-1}} \dot{q}(t), \frac{D^{j-2}}{Dt^{j-2}} \dot{q}(t) \right) \dot{q}(t) = 0. \quad (2.13)$$

The second generalisation comprises the class of *cubics in tension* that arise from the following class of Lagrangians

$$L_\tau(q, \dot{q}, \ddot{q}) := \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} \right\|_q^2 + \frac{\tau^2}{2} \|\dot{q}\|_q^2, \quad (2.14)$$

where τ is a real constant. The Euler–Lagrange equations are

$$\frac{D^3}{Dt^3} \dot{q}(t) + R \left(\frac{D}{Dt} \dot{q}(t), \dot{q}(t) \right) \dot{q}(t) = \tau^2 \frac{D}{Dt} \dot{q}(t), \quad (2.15)$$

as proven in [30], where an application to space-based interferometric imaging was discussed. More precisely, the authors consider the following interpolation problem:

⁵Namely, for vector fields W, X, Y, Z , we have $\gamma(R(X, Y)Z, W) = \gamma(R(W, Z)Y, X)$ [46, Proposition 7.4].

Given $N + 1$ points $q_i \in Q$, $i = 0, \dots, N$ and tangent vectors $v_j \in T_{q_j}Q$, $j = 0, N$, minimise

$$\mathcal{J}[q] := \frac{1}{2} \int_{t_0}^{t_N} \left(\gamma_{q(t)} \left(\frac{D}{Dt} \dot{q}(t), \frac{D}{Dt} \dot{q}(t) \right) + \tau^2 \gamma_{q(t)} (\dot{q}(t), \dot{q}(t)) \right) dt, \quad (2.16)$$

among curves $t \mapsto q(t) \in Q$ that are C^1 on $[t_0, t_N]$, smooth on $[t_i, t_{i+1}]$, $t_0 \leq t_1 \leq \dots \leq t_N$, and subject to the interpolation constraints

$$q(t_i) = q_i, \quad \text{for all } i = 1, \dots, N - 1$$

and the boundary conditions

$$q(t_0) = q_0, \quad \dot{q}(t_0) = v_0, \quad \text{and} \quad q(t_N) = q_N, \quad \dot{q}(t_N) = v_N.$$

This problem is a generalisation of the one mentioned above in Remark 2.2. As shown in [30, Theorem 3.2] the minimiser is a C^2 curve and can accordingly be called a *Riemannian cubic spline in tension*.

Another more general problem is treated in Chapter 7 of [13]: The authors consider the second-order Lagrangian

$$L(q, \dot{q}, \ddot{q}) = \frac{1}{2} \gamma(q) \left(\frac{D}{Dt} \dot{q}, I \left(\frac{D}{Dt} \dot{q} \right) \right), \quad (2.17)$$

where $I : TQ \rightarrow TQ$ is a vector bundle isomorphism covering the identity. Evidently, if I is the identity mapping, one recovers the Lagrangian for Riemannian cubics. In particular, the derivation of the Euler–Lagrange equations for (2.17) that is given in [13] simplifies to the derivation of (2.10) we presented above.

In the context of group actions on object manifolds, we refer the reader to Section 2.5 of the present chapter for an example of higher-order interpolation particularly relevant for computational anatomy.

Remark 2.3. We also mention an alternative, non-variational, generalisation of cubic polynomials to Riemannian manifolds. Instead of (2.15), one considers solutions $q(t)$ of

$$\frac{D^3}{Dt^3} \dot{q}(t) = 0.$$

Higher-order polynomials can be defined similarly, by

$$\frac{D^k}{Dt^k} \dot{q}(t) = 0.$$

We refer to [47, 48] and references therein for more details and the recent work [49], which discusses polynomial regression in shape spaces using this type of curves.

2.2.4 Quotient space and reduced Lagrangian

When one deals with a Lagrangian $L : T^{(k)}Q \rightarrow \mathbb{R}$ that is invariant with respect to the lift $\Phi^{(k)} : G \times T^{(k)}Q \rightarrow T^{(k)}Q$ of a group action $\Phi : G \times Q \rightarrow Q$, then the invariance can be exploited to define a new function called the *reduced Lagrangian* on the quotient space $(T^{(k)}Q)/G$. We review this procedure here in the special case where $Q = G$. For the general case we refer to [50].

Let G be a Lie group and $h \in G$. The right-, respectively left-actions by h on G ,

$$R_h : G \rightarrow G, \quad g \mapsto gh, \quad \text{and} \quad L_h : G \rightarrow G, \quad g \mapsto hg,$$

can be naturally lifted to actions on the k^{th} -order tangent bundle $T^{(k)}G$ (see (2.4)). We will denote these lifted actions by concatenation, as in

$$\begin{aligned} R_h^{(k)} : T^{(k)}G &\rightarrow T^{(k)}G, \quad [g]_{g_0}^{(k)} \mapsto R_h^{(k)}([g]_{g_0}^{(k)}) =: [g]_{g_0}^{(k)}h, \quad \text{and} \\ L_h^{(k)} : T^{(k)}G &\rightarrow T^{(k)}G, \quad [g]_{g_0}^{(k)} \mapsto L_h^{(k)}([g]_{g_0}^{(k)}) =: h[g]_{g_0}^{(k)}. \end{aligned}$$

Consider a Lagrangian $L : T^{(k)}G \rightarrow \mathbb{R}$ that is right-, or left-invariant, i.e., invariant with respect to the lifted right-, or left-actions of G on itself. For any $[g]_{g_0}^{(k)} \in T^{(k)}G$ we then get

$$L([g]_{g_0}^{(k)}) = L|_{T_e^{(k)}G}([g]_{g_0}^{(k)}g_0^{-1}), \quad \text{or} \quad L([g]_{g_0}^{(k)}) = L|_{T_e^{(k)}G}(g_0^{-1}[g]_{g_0}^{(k)}), \quad (2.18)$$

respectively. The restriction $L|_{T_e^{(k)}G}$ of the Lagrangian to the k^{th} -order tangent space at the identity e therefore fully specifies the Lagrangian L . Moreover, there are natural identifications $\alpha_k : T_e^{(k)}G \rightarrow k\mathfrak{g}$ given by

$$\alpha_k([g]_e^{(k)}) := \left(\dot{g}(0), \frac{d}{dt} \Big|_{t=0} \dot{g}(t)g(t)^{-1}, \dots, \frac{d^{k-1}}{dt^{k-1}} \Big|_{t=0} \dot{g}(t)g(t)^{-1} \right), \quad (2.19)$$

or

$$\alpha_k([g]_e^{(k)}) := \left(\dot{g}(0), \frac{d}{dt} \Big|_{t=0} g(t)^{-1}\dot{g}(t), \dots, \frac{d^{k-1}}{dt^{k-1}} \Big|_{t=0} g(t)^{-1}\dot{g}(t) \right), \quad (2.20)$$

respectively, where $t \mapsto g(t)$ is an arbitrary representative of $[g]_e^{(k)}$.

The reduced Lagrangian $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$ is then defined as

$$\ell := L|_{T_e^{(k)}G} \circ \alpha_k^{-1}, \quad (2.21)$$

where one uses the choice for α_k that is appropriate, namely (2.19) for a right-invariant Lagrangian and (2.20) for a left-invariant Lagrangian. Let $t \mapsto g(t) \in G$ be a curve on the Lie group. For every t this curve defines an element in $T_{g(t)}^{(k)}G$, namely

$$[g]_{g(t)}^{(k)} := [h]_{g(t)}^{(k)}, \quad \text{where } h \text{ is the curve } \tau \mapsto h(\tau) := g(t + \tau). \quad (2.22)$$

Note that for the case $k = 1$ we write, as usual, $\dot{g}(t) := [g]_{g(t)}^{(1)}$. The following lemma is a direct consequence of the definitions:

Lemma 2.4. *Let $t \mapsto g(t)$ be a curve in G and $L : T^{(k)}G \rightarrow \mathbb{R}$ a right-, or left-invariant Lagrangian. Then the following equation holds for any time t_0 ,*

$$L\left([g]_{g(t_0)}^{(k)}\right) = \ell\left(\xi(t_0), \dot{\xi}(t_0), \dots, \xi^{(k-1)}(t_0)\right), \quad (2.23)$$

where $\xi(t) := \dot{g}(t)g^{-1}(t)$, or $\xi(t) := g^{-1}(t)\dot{g}(t)$, respectively.

This last equation will play a key role in the higher-order Euler–Poincaré reduction discussed in the next section.

2.3 Higher-order Euler–Poincaré reduction

In this section we derive the k^{th} -order Euler–Poincaré equations by reducing the variational principle associated to the Euler–Lagrange equations on $T^{(k)}G$. The equations adopt a factorised form, in which the Euler–Poincaré operator at $k = 1$ is applied to the Euler–Lagrange operator acting on the reduced Lagrangian $\ell(\xi, \dot{\xi}, \ddot{\xi}, \dots, \xi^{(k-1)}) : k\mathfrak{g} \rightarrow \mathbb{R}$ at the given order, k . We then derive the k^{th} -order Euler–Poincaré equations for Riemannian cubics and, more generally, geometric k -splines.

2.3.1 Quotient map, variations and k^{th} -order Euler–Poincaré equations

Let $L : T^{(k)}G \rightarrow \mathbb{R}$ be a right-, or left-invariant Lagrangian. Recall from Section 2.2.2 that the Euler–Lagrange equations are equivalent to the higher-order Hamilton’s principle, as

follows:

For given $h_i \in G$ and $[h]_{h_i}^{(k-1)} \in T_{h_i}^{(k-1)}G$, $i = 1, 2$, find a critical curve of the functional

$$\mathcal{J}[g] = \int_{t_1}^{t_2} L\left([g]_{g(t)}^{(k)}\right) dt$$

among all curves $g : t \in [t_1, t_2] \mapsto g(t) \in G$ satisfying the endpoint condition

$$[g]_{g(t_i)}^{(k-1)} = [h]_{h_i}^{(k-1)}, \quad i = 1, 2. \quad (2.24)$$

The time derivatives of up to order $k-1$ are therefore fixed at the endpoints, i.e., $[g]_{g(t_i)}^{(j)} = [h]_{h_i}^{(j)}$, $j = 0, \dots, k-1$, are automatically verified. Let $g : t \mapsto g(t) \in G$ be a curve and $(\varepsilon, t) \mapsto g_\varepsilon(t) \in G$ a variation of g respecting (2.24). We recall from Lemma 2.4 that, for any ε and any t_0 ,

$$L\left([g_\varepsilon]_{g_\varepsilon(t_0)}^{(k)}\right) = \ell\left(\xi_\varepsilon(t_0), \dots, \xi_\varepsilon^{(k-1)}(t_0)\right), \quad (2.25)$$

where $\xi_\varepsilon := g_\varepsilon g_\varepsilon^{-1}$, or $\xi_\varepsilon := g_\varepsilon^{-1} g_\varepsilon$ respectively for the right-, or left-invariant Lagrangian L . As is well-known from first-order Euler–Poincaré reduction, the variation $\delta\xi$ induced by the variation δg is given by ([5, Theorem 13.5.3])

$$\delta\xi = \dot{\eta} \mp [\xi, \eta], \quad (2.26)$$

where $\eta := (\delta g)g^{-1}$, or $\eta := g^{-1}(\delta g)$, respectively. It follows from the endpoint conditions (2.24) that $\eta(t_i) = \dot{\eta}(t_i) = \dots = \eta^{(k-1)}(t_i) = 0$ and therefore $\delta\xi(t_i) = \dots = \partial_t^{k-2}\delta\xi(t_i) = 0$, for $i = 1, 2$.

Remark 2.5 (Ad, ad and their duals). For any $\nu \in \mathfrak{g}$ we define $\text{ad}_\nu^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ to be the dual of the map

$$\text{ad}_\nu : \mathfrak{g} \rightarrow \mathfrak{g}, \quad \xi \mapsto [\nu, \xi].$$

We will use ad^* in the calculations that follow. For later reference we also define, for any $g \in G$, the Lie algebra automorphism

$$\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}, \quad \xi \mapsto TL_g TR_{g^{-1}} \xi$$

and its dual Ad^* . We recall the relation ([5, Proposition 9.1.5])

$$\text{ad}_\nu \xi = \left. \frac{d}{ds} \right|_{s=0} \text{Ad}_{g(s)} \xi,$$

where $g(s)$ is any curve in the group that originates at the identity with initial tangent vector ν .

We are now ready to compute the variation of $\mathcal{J}$:

$$\begin{aligned}
\delta \int_{t_1}^{t_2} L([g]_{g(t)}^{(k)}) dt &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{t_1}^{t_2} L([g_\varepsilon]_{g_\varepsilon(t)}^{(k)}) dt \stackrel{(2.25)}{=} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{t_1}^{t_2} \ell(\xi_\varepsilon, \dots, \xi_\varepsilon^{(k-1)}) dt \\
&= \sum_{j=0}^{k-1} \int_{t_1}^{t_2} \left\langle \frac{\delta \ell}{\delta \xi^{(j)}}, \delta \xi^{(j)} \right\rangle dt = \sum_{j=0}^{k-1} \int_{t_1}^{t_2} \left\langle \frac{\delta \ell}{\delta \xi^{(j)}}, \partial_t^j \delta \xi \right\rangle dt \\
&= \int_{t_1}^{t_2} \left\langle \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}}, \delta \xi \right\rangle dt \\
&= \int_{t_1}^{t_2} \left\langle \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}}, \partial_t \eta \mp [\xi, \eta] \right\rangle dt \\
&= \int_{t_1}^{t_2} \left\langle (-\partial_t \mp \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}}, \eta \right\rangle dt,
\end{aligned}$$

where we used the vanishing endpoint conditions $\delta \xi(t_i) = \dots = \partial_t^{k-2} \delta \xi(t_i) = 0$ and $\eta(t_i) = 0$, for $i = 1, 2$, when integrating by parts. Therefore, the stationarity condition $\delta \mathcal{J} = 0$ implies the k^{th} -order Euler–Poincaré equation,

$$\boxed{(\partial_t \pm \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = 0.} \tag{2.27}$$

Formula (2.27) takes the following forms for various choices of $k = 1, 2, 3$:

If $k = 1$:

$$(\partial_t \pm \text{ad}_\xi^*) \frac{\delta \ell}{\delta \xi} = 0,$$

If $k = 2$:

$$(\partial_t \pm \text{ad}_\xi^*) \left(\frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} \right) = 0, \tag{2.28}$$

If $k = 3$:

$$(\partial_t \pm \text{ad}_\xi^*) \left(\frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} + \partial_t^2 \frac{\delta \ell}{\delta \ddot{\xi}} \right) = 0.$$

The first of these is the usual Euler–Poincaré equation. The others adopt a *factorised* form in which the Euler–Poincaré operator $(\partial_t \pm \text{ad}_\xi^*)$ is applied to the Euler–Lagrange operation on the reduced Lagrangian $\ell(\xi, \dot{\xi}, \ddot{\xi}, \dots)$ at the given order.

The results obtained above are summarised in the following theorem.

Theorem 2.6. [k^{th} -order Euler–Poincaré reduction] Let $L : T^{(k)}G \rightarrow \mathbb{R}$ be a G -invariant Lagrangian and let $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$ be the associated reduced Lagrangian. Let $g(t)$ be a curve

in G and $\xi(t) = \dot{g}(t)g(t)^{-1}$, resp. $\xi(t) = g(t)^{-1}\dot{g}(t)$ be the reduced curve in the Lie algebra $\mathfrak{g}$. Then the following assertions are equivalent.

(i) The curve $g(t)$ is a solution of the k^{th} -order Euler–Lagrange equations for $L : T^{(k)}G \rightarrow \mathbb{R}$.

(ii) Hamilton’s variational principle

$$\delta \int_{t_1}^{t_2} L(g, \dot{g}, \dots, g^{(k)}) dt = 0$$

holds upon using variations δg such that $\delta g^{(j)}$ vanish at the endpoints for $j = 0, \dots, k-1$.

(iii) The k^{th} -order Euler–Poincaré equations are satisfied:

$$(\partial_t \pm \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = 0. \quad (2.29)$$

(iv) The constrained variational principle

$$\delta \int_{t_1}^{t_2} \ell(\xi, \dot{\xi}, \dots, \xi^{(k)}) = 0$$

holds for constrained variations of the form $\delta \xi = \partial_t \eta \mp [\xi, \eta]$, where η is an arbitrary curve in $\mathfrak{g}$ such that the $\eta^{(j)}$ vanish at the endpoints, for all $j = 0, \dots, k-1$.

Remark 2.7. As we have seen earlier, the k^{th} -order tangent bundle $T^{(k)}G$ is also a Lie group in a natural way. However it is worth mentioning that the group structure of $T^{(k)}G$ is not involved in the higher-order Euler–Poincaré reduction. The first-order Euler–Poincaré reduction for $T^{(k)}G$, viewed as a group, deals with Lagrangians defined on the vector bundle $T(T^{(k)}G)$ that are invariant with respect to the right or left group multiplication in $T^{(k)}G$. This situation is distinct from the one we discussed above, where the Lagrangian was defined on the fibre bundle $T^{(k)}G$. For more details we refer to [39, Remark 3.3].

Our next step is the treatment, in the framework of higher-order Euler–Poincaré reduction, of Riemannian cubics and their generalisations on Lie groups with invariant metrics.

2.4 Riemannian cubics and geometric k -splines

In this section we apply the k^{th} -order Euler–Poincaré reduction to the particular case of Riemannian cubics, or geometric 2-splines, on Lie groups. In Remark 2.9 below we deal with the general case of geometric k -splines. Fix a right-, respectively left-invariant Riemannian metric γ on the Lie group G . We denote by

$$\|v_g\|_g^2 := \gamma_g(v_g, v_g)$$

the corresponding squared norm of a vector $v_g \in T_g G$. The inner product induced on the Lie algebra $\mathfrak{g}$ is also denoted by $\gamma : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$ and its squared norm by

$$\|\xi\|_{\mathfrak{g}}^2 := \gamma(\xi, \xi).$$

We have the following proposition, which we will prove in two different ways. The first proof is elementary and only requires knowledge of the geodesic equation in Euler–Poincaré reduced form. The method of the second proof is more general and can in particular be applied to geometric k -splines with $k > 2$.

Proposition 2.8. *Consider the Lagrangian $L : T^{(2)}G \rightarrow \mathbb{R}$ for Riemannian cubics, given by*

$$L(g, \dot{g}, \ddot{g}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2, \quad (2.30)$$

where $\|\cdot\|$ is the norm of a right-, respectively left-invariant metric on G . Then L is right-, respectively left-invariant and induces the reduced Lagrangian $\ell : 2\mathfrak{g} \rightarrow \mathbb{R}$ given by

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \left\| \dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi \right\|_{\mathfrak{g}}^2, \quad (2.31)$$

where $\text{ad}^{\dagger}$ is defined by $\text{ad}_{\xi}^{\dagger} \eta := (\text{ad}_{\xi}^*(\eta^{\flat}))^{\sharp}$, for any $\xi, \eta \in \mathfrak{g}$.

First proof. Let us compute the Euler–Lagrange operator for the kinetic energy Lagrangian $L_{\text{KE}} : TG \rightarrow \mathbb{R}$, given by $L_{\text{KE}}(\dot{g}) = \frac{1}{2} \|\dot{g}\|_g^2$. We obtain variations

$$\begin{aligned} \delta \int_{t_1}^{t_2} L_{\text{KE}}(\dot{g}) dt &= \delta \int_{t_1}^{t_2} \frac{1}{2} \|\dot{g}\|_g^2 dt = \int_{t_1}^{t_2} \gamma(g) \left(\dot{g}, \frac{D}{D\varepsilon} \dot{g} \right) dt \\ &= \int_{t_1}^{t_2} \gamma(g) \left(\dot{g}, \frac{D}{Dt} \delta g \right) dt = - \int_{t_1}^{t_2} \gamma(g) \left(\frac{D}{Dt} \dot{g}, \delta g \right) dt. \end{aligned}$$

Therefore, the Euler–Lagrange operator $\mathcal{EL}(L_{\text{KE}}) : T^{(2)}G \rightarrow \mathbb{R}$ is given by

$$\mathcal{EL}(L_{\text{KE}})(g, \dot{g}, \ddot{g}) = - \left(\frac{D}{Dt} \dot{g} \right)^{\flat}. \quad (2.32)$$

Notice that this implies in particular that the Euler–Lagrange equation for the kinetic energy Lagrangian is the geodesic equation, as we mentioned in the introduction, Section 1.1. On the other hand, we can use the invariance of the kinetic energy Lagrangian to rewrite the variation in terms of the reduced velocity vector (just as we did for higher order in Section 2.3.1) to obtain, for right-invariance,

$$\begin{aligned} \delta \int_{t_1}^{t_2} L_{\text{KE}}(\dot{g}) dt &= \int_{t_1}^{t_2} \left\langle -\dot{\xi}^{\flat} - \text{ad}_{\xi}^* \xi^{\flat}, TR_{g^{-1}} \delta g \right\rangle dt \\ &= \int_{t_1}^{t_2} \gamma(e) \left(-\dot{\xi} - \text{ad}_{\xi}^{\dagger} \xi, TR_{g^{-1}} \delta g \right) dt = \int_{t_1}^{t_2} \gamma(g) \left(TR_g \left(-\dot{\xi} - \text{ad}_{\xi}^{\dagger} \xi \right), \delta g \right) dt, \end{aligned}$$

setting $\xi := \dot{g}g^{-1}$, and for left-invariance

$$\delta \int_{t_1}^{t_2} L_{\text{KE}}(\dot{g}) dt = \int_{t_1}^{t_2} \gamma(g) \left(TL_g \left(-\dot{\xi} + \text{ad}_{\xi}^{\dagger} \xi \right), \delta g \right) dt,$$

setting $\xi := g^{-1}\dot{g}$. This gives equivalent expressions

$$\begin{aligned} \mathcal{EL}(L_{\text{KE}})(g, \dot{g}, \ddot{g}) &= \left(TR_g \left(-\dot{\xi} - \text{ad}_{\xi}^{\dagger} \xi \right) \right)^{\flat} \quad \text{and} \\ \mathcal{EL}(L_{\text{KE}})(g, \dot{g}, \ddot{g}) &= \left(TL_g \left(-\dot{\xi} + \text{ad}_{\xi}^{\dagger} \xi \right) \right)^{\flat} \end{aligned}$$

for the Euler–Lagrange operator (2.32). We have shown in particular that

$$\frac{D}{Dt} \dot{g} = TR_g \left(\dot{\xi} + \text{ad}_{\xi}^{\dagger} \xi \right), \quad \text{and} \quad \frac{D}{Dt} \dot{g} = TL_g \left(\dot{\xi} - \text{ad}_{\xi}^{\dagger} \xi \right),$$

for a right-, or left-invariant metric γ , respectively. Using the invariance of the metric once more, we obtain

$$L(g, \dot{g}, \ddot{g}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2 = \frac{1}{2} \left\| \dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi \right\|_{\mathfrak{g}}^2, \quad (2.33)$$

which depends only on the right-invariant quantity $\xi = \dot{g}g^{-1}$, respectively the left-invariant quantity $\xi = g^{-1}\dot{g}$. Accordingly, L is right-, or left-invariant, and the group-reduced Lagrangian is

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \left\| \dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi \right\|_{\mathfrak{g}}^2.$$

This completes the proof. □

Second proof. Let us recall the expression of the Levi–Civita covariant derivative associated to a right-, respectively left-invariant Riemannian metric on G . For $X \in \mathfrak{X}(G)$ and $v_g \in T_g G$, we have (e.g., [51, Chap. 46.5])

$$\nabla_{v_g} X(g) = TR_g \left(df(v_g) + \frac{1}{2} \text{ad}_v^\dagger f(g) + \frac{1}{2} \text{ad}_{f(g)}^\dagger v - \frac{1}{2} [v, f(g)] \right), \quad v := v_g g^{-1} \quad (2.34)$$

$$\text{resp. } \nabla_{v_g} X(g) = TL_g \left(df(v_g) - \frac{1}{2} \text{ad}_v^\dagger f(g) - \frac{1}{2} \text{ad}_{f(g)}^\dagger v + \frac{1}{2} [v, f(g)] \right), \quad v := g^{-1} v_g \quad (2.35)$$

where $f \in \mathcal{F}(G; \mathfrak{g})$ is uniquely determined by the condition $X(g) = TR_g(f(g))$ for right-invariance and $X(g) = TL_g(f(g))$ for left-invariance; and we introduced the *exterior derivative* d , which satisfies $df(v_g) = \partial_{\varepsilon=0} f(g(\varepsilon))$ for any curve $g(\varepsilon)$ with $\partial_{\varepsilon=0} g(\varepsilon) = v_g$. Therefore, we have

$$\frac{D}{Dt} \dot{g}(t) = \nabla_{\dot{g}} \dot{g} = TR_g \left(\dot{\xi} + \frac{1}{2} \text{ad}_\xi^\dagger \xi + \frac{1}{2} \text{ad}_\xi^\dagger \xi - \frac{1}{2} [\xi, \xi] \right) = TR_g \left(\dot{\xi} + \text{ad}_\xi^\dagger \xi \right),$$

respectively

$$\frac{D}{Dt} \dot{g}(t) = \nabla_{\dot{g}} \dot{g} = TL_g \left(\dot{\xi} - \frac{1}{2} \text{ad}_\xi^\dagger \xi - \frac{1}{2} \text{ad}_\xi^\dagger \xi + \frac{1}{2} [\xi, \xi] \right) = TL_g \left(\dot{\xi} - \text{ad}_\xi^\dagger \xi \right),$$

where we used $v_g = \dot{g}$ and $X(g) = \dot{g}$; so $f(g) = \dot{g} g^{-1} = \xi$ (respectively, $f(g) = g^{-1} \dot{g} = \xi$) and $df(v_g) = \dot{\xi}$. One concludes the proof in the same manner as above. □

Remark 2.9. *The above considerations generalise to geometric k -splines for $k > 2$. Indeed, iterated application of formulas (2.34), (2.35) yields*

$$\frac{D^k}{Dt^k} \dot{g} = TR_g(\eta_k), \quad \text{respectively} \quad \frac{D^k}{Dt^k} \dot{g} = TL_g(\eta_k),$$

where the quantities $\eta_k \in \mathfrak{g}$ are defined by the recursive formulae

$$\eta_1 = \dot{\xi} \pm \text{ad}_\xi^\dagger \xi, \quad \text{and} \quad \eta_k = \dot{\eta}_{k-1} \pm \frac{1}{2} \left(\text{ad}_\xi^\dagger \eta_{k-1} + \text{ad}_{\eta_{k-1}}^\dagger \xi + \text{ad}_{\eta_{k-1}} \xi \right), \quad (2.36)$$

for $\xi = \dot{g}g^{-1}$, respectively $\xi = g^{-1}\dot{g}$. Therefore, the Lagrangian (2.12) for geometric k -splines on a Lie group G with right-, respectively left-invariant Riemannian metric,

$$L(g, \dot{g}, \dots, g^{(k)}) = \frac{1}{2} \left\| \frac{D^{k-1}}{Dt^{k-1}} \dot{g} \right\|_g^2,$$

is right-, respectively left-invariant, and the reduced Lagrangian is

$$\ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}) = \frac{1}{2} \|\eta_{k-1}\|_{\mathfrak{g}}^2. \quad (2.37)$$

2.4.1 Computing the second-order Euler–Poincaré equations for Riemannian cubics

Let us compute the Euler–Poincaré equations for $k = 2$. Upon defining $\eta := \dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi$ the required variational derivatives of the reduced Lagrangian (2.31) can be calculated as follows:

$$\left\langle \frac{\delta \ell}{\delta \xi}, \delta \xi \right\rangle = \langle \pm \text{ad}_{\delta \xi}^* \xi^{\flat} \pm \text{ad}_{\xi}^* \delta \xi^{\flat}, \eta \rangle = \langle \mp (\text{ad}_{\eta}^* \xi^{\flat} + (\text{ad}_{\eta} \xi)^{\flat}), \delta \xi \rangle$$

and

$$\left\langle \frac{\delta \ell}{\delta \dot{\xi}}, \delta \dot{\xi} \right\rangle = \langle \eta^{\flat}, \delta \dot{\xi} \rangle.$$

Hence,

$$\frac{\delta \ell}{\delta \xi} = \mp (\text{ad}_{\eta}^* \xi^{\flat} + (\text{ad}_{\eta} \xi)^{\flat}) \quad \text{and} \quad \frac{\delta \ell}{\delta \dot{\xi}} = \eta^{\flat}. \quad (2.38)$$

From formula (2.28) with $k = 2$ one then finds the second-order Euler–Poincaré equation

$$(\partial_t \pm \text{ad}_{\xi}^*) (\partial_t \eta^{\flat} \pm \text{ad}_{\eta}^* \xi^{\flat} \pm (\text{ad}_{\eta} \xi)^{\flat}) = 0, \quad \text{with} \quad \eta^{\flat} := \dot{\xi}^{\flat} \pm \text{ad}_{\xi}^* \xi^{\flat}, \quad (2.39)$$

or, equivalently,

$$\boxed{(\partial_t \pm \text{ad}_{\xi}^{\dagger}) (\partial_t \eta \pm \text{ad}_{\eta}^{\dagger} \xi \pm \text{ad}_{\eta} \xi) = 0, \quad \text{with} \quad \eta := \dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi.} \quad (2.40)$$

These are the reduced equations for Riemannian cubics, or geometric 2-splines, associated to a left-, or right-invariant Riemannian metric on the Lie group G . In an analogous fashion one can derive the Euler–Poincaré equations for geometric k -splines, using the reduced Lagrangian (2.37). We remark that a version of equation (2.40) was derived in [52] for the case of a left-invariant metric on the particular Lie group $G = SO(3)$.

When the metric is left- and right-invariant (bi-invariant), certain simplifications arise, as we shall discuss next.

2.4.2 Bi-invariant metrics and the NHP equation

In the case of a *bi-invariant* Riemannian metric, we have $\text{ad}_\xi^\dagger \eta = -\text{ad}_\xi \eta$ and therefore the reduced Lagrangian (2.31) becomes $\ell(\xi, \dot{\xi}) = \frac{1}{2} \|\dot{\xi}\|_{\mathfrak{g}}^2$. Therefore,

$$\frac{\delta \ell}{\delta \dot{\xi}} = \dot{\xi}^\flat \quad \text{and} \quad \frac{\delta \ell}{\delta \xi} = 0. \quad (2.41)$$

The second-order Euler–Poincaré equations (2.28) in this case become

$$(\partial_t \pm \text{ad}_\xi^*) \dot{\xi}^\flat = 0 \quad \text{or} \quad \left(\partial_t \pm \text{ad}_\xi^\dagger \right) \ddot{\xi} = 0, \quad (2.42)$$

or

$$\boxed{\ddot{\xi} \mp [\xi, \dot{\xi}] = 0.} \quad (2.43)$$

Note that since the metric is bi-invariant, one may choose to reduce the system either on the right *or* on the left; this choice will determine the sign above. Equation (2.43) appears in [19]. We call it the **NHP equation** after Noakes, Heinzinger and Paden, who first derived it for $G = SO(3)$ in [18]:

$$\boxed{\ddot{\Omega} \mp \Omega \times \ddot{\Omega} = 0.} \quad (2.44)$$

In order to understand the vector notation in the previous equation, let us make the following remark about conventions.

Remark 2.10 (Conventions for $\mathfrak{so}(3)$ and $\mathfrak{so}(3)^*$). *In equation (2.44) and throughout this thesis we use vector notation for the Lie algebra $\mathfrak{so}(3)$ of the Lie group of rotations $SO(3)$, as well as for its dual $\mathfrak{so}(3)^*$. One identifies $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the familiar isomorphism*

$$\hat{\cdot}: \mathbb{R}^3 \rightarrow \mathfrak{so}(3), \quad \Omega = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \Omega := \hat{\Omega} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}, \quad (2.45)$$

called the hat map. This is a Lie algebra isomorphism when the vector cross product $\times$ is used as the Lie bracket operation on $\mathbb{R}^3$. The identification of $\mathfrak{so}(3)$ with $\mathbb{R}^3$ induces an isomorphism of the dual spaces $\mathfrak{so}(3)^* \cong (\mathbb{R}^3)^* \cong \mathbb{R}^3$, whereby the standard dot product is used as duality pairing.

In [18], the unreduced equations (2.10) are derived for general Riemannian manifolds, but the symmetry reduced equation is given only for $SO(3)$ with bi-invariant metric. That derivation takes (2.10) as starting point and hence follows a different route from the Euler–Poincaré reduction method we described above.

The NHP equation can be integrated once to yield

$$\ddot{\xi} \mp [\xi, \dot{\xi}] = \nu \quad (2.46)$$

for a constant $\nu \in \mathfrak{g}$. Solutions ξ were called *Lie quadratics* in [23, 24]. For $G = SO(3)$, long-term behaviour and internal symmetries of Lie quadratics were studied there, both in the *null* ($\nu = 0$) and the *non-null* ($\nu \neq 0$) case. Generalisations to cubics in tension can be found in [53].

As we mentioned in the introductory section 1.1, the solution curves of higher-order variational principles typically have added smoothness, which makes them an attractive choice for longitudinal data interpolation, in particular in computational anatomy. This is the topic of the next section.

2.5 Higher-order template matching problems

We first give a brief account of previous work done on longitudinal data interpolation in computational anatomy. Then we derive the equations that generalise the first-order methods of [37] to higher order. After making a few remarks concerning the gain in smoothness, we provide a qualitative discussion of two Lagrangians of interest for computational anatomy. Finally, we close the section by demonstrating the spline approach to template matching for the finite dimensional case of fitting a smooth curve through a sequence of target points on the sphere, using the natural action of the rotation group.

2.5.1 Previous work on longitudinal data interpolation in computational anatomy

Computational anatomy is concerned with modeling and quantifying diffeomorphic evolutions of shapes, as presented in [54, 55]. Usually one aims at finding a geodesic path, on the space of shapes, between given initial and final data. This approach can be

adapted for longitudinal data interpolation; that is, interpolation through a sequence of data points. One may interpolate between the given data points in such a way that the path is piecewise-geodesic, [14, 15]. It was, however, argued in [16] that higher-order models, i.e., models that provide more smoothness than the piecewise-geodesic one, are better suited as growth models for typical biological evolutions. As an example of such a higher-order model, spline interpolation on the Riemannian manifold of landmarks was studied there. In the next paragraph we will consider another class of models of interest for computational anatomy that are inspired by an *optimal control viewpoint*. Indeed, the time-dependent vector field is seen as a control variable acting on the template and the penalty on this control variable will be defined on the Lie algebra. This means that the action functional will be sensitive to the curve in the transformation group *directly*, rather than only through its orbit in some data vector space. This class of models is an interesting alternative to the shape splines model presented in [16].

2.5.2 Euler–Lagrange equations for higher-order template matching

Let G be a Lie group with Lie algebra $\mathfrak{g}$, and let

$$G \times V \rightarrow V, \quad (g, I) \mapsto gI \quad (2.47)$$

be a left representation of G on the vector space V . Let $\|\cdot\|_V$ be a norm on V . We consider minimisation problems of the following abstract form:

Given a Lagrangian $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$, a tolerance parameter $\sigma \in \mathbb{R}$, initial time $t_0 = 0$, target times $t_1, \dots, t_l \in \mathbb{R}$, template and target data $T_0, I_{t_1}, \dots, I_{t_l} \in V$, and $\xi_0^0, \dots, \xi_0^{k-2} \in \mathfrak{g}$, minimise the functional

$$\begin{aligned} E[\xi] &= \int_0^{t_l} \ell(\xi(t), \dots, \xi^{(k-1)}(t)) dt + \frac{1}{2\sigma^2} \sum_{i=1}^l \|g^\xi(t_i)T_0 - I_{t_i}\|_V^2 \\ &:= \sum_{j=0}^{l-1} \int_{t_j}^{t_{j+1}} \ell(\xi(t), \dots, \xi^{(k-1)}(t)) dt + \frac{1}{2\sigma^2} \sum_{i=1}^l \|g^\xi(t_i)T_0 - I_{t_i}\|_V^2 \end{aligned} \quad (2.48)$$

subject to the conditions $\xi^{(j)}(0) = \xi_0^j$, $j = 0, \dots, k-2$, where $g^\xi(t_i)$ is the flow, as defined below, of $\xi(t)$ evaluated at time t_i . The minimisation is carried out over a space $\mathcal{P}_{k-1}$ of

curves $\xi : [0, t_l] \rightarrow \mathfrak{g}$. We place a number of requirements on this space: Of course, the cost functional E has to be defined on $\mathcal{P}_{k-1}$. Moreover, the curves need to be $2k-1$ times continuously differentiable, with existing limits, on the open intervals $(0, t_1), \dots, (t_{l-1}, t_l)$ and $k-2$ times continuously differentiable on $[0, t_l]$. This definition of $\mathcal{P}_{k-1}$ reflects the requirements of the variational calculus that follows. In particular it ensures that when carrying out integration by parts below, all terms are well-defined.

Given a curve $\xi(t) \in \mathcal{P}_{k-1}$ in the Lie algebra $\mathfrak{g}$, its flow $g^\xi : t \mapsto g^\xi(t) \in G$ is a continuous curve defined by the conditions

$$g^\xi(0) = e, \quad \text{and} \quad \frac{d}{dt}g^\xi(t) = \xi(t)g^\xi(t), \quad (2.49)$$

whenever t is in one of the open intervals $(0, t_1), \dots, (t_{l-1}, t_l)$. Here we used the notation $\xi(t)g^\xi(t) := TR_{g^\xi(t)}\xi(t)$. In the special case when $\xi(t) \equiv \xi$ is a constant, the flow at time 1 is called the *Lie exponential* and denoted by $\exp(\xi) := g^\xi(1)$. We typically think of $(I_{t_1}, \dots, I_{t_l})$ as the time-sequence of data, indexed by time points t_j , $j = 1, \dots, l$, and T_0 is the template (the source image). Moreover, $\xi : t \in [0, t_l] \mapsto \xi(t) \in \mathfrak{g}$ is typically a time-dependent vector field (sufficiently smooth in time) that generates a flow of diffeomorphisms $g^\xi : t \in [0, t_l] \mapsto g^\xi(t) \in G$. Note that, in this case, the Lie group G is infinite dimensional and a rigorous framework to work in is the *large deformations by diffeomorphisms* setting thoroughly explained in [56]. We will informally refer to this case as the *diffeomorphism case* or *infinite dimensional case*. The expression $g^\xi(t_i)T_0$ represents the template at time t_i , as it is being deformed by the flow of diffeomorphisms.

Inspired by the second-order model presented in [16], this subsection thus generalises the work of [37] in two directions. First, we allow for a *higher-order* penalty on the time-dependent vector field given by the first term of the functional (2.48); second, the similarity measure (second term in (2.48)) takes into account *several* time points in order to compare the deformed template with the time-sequence target.

Staying at a general level, we will take the geometric viewpoint of [37] in order to derive the Euler–Lagrange equations, which are satisfied by any minimiser of E . We suppose that the norm on V is induced by an inner product $\langle \cdot, \cdot \rangle_V$ and denote by $\flat$ the isomorphism

$$\flat : V \rightarrow V^*, \quad \omega \mapsto \omega^\flat$$

that satisfies

$$\langle I, J \rangle_V = \langle I^\flat, J \rangle \quad \text{for all } I, J \in V,$$

where we wrote $\langle \cdot, \cdot \rangle$ for the duality pairing between V and its dual V^* . The action (2.47) of G on V induces an action on V^* ,

$$G \times V^* \rightarrow V^*, \quad (g, \omega) \mapsto g\omega = (g^{-1})^*\omega$$

that is defined by the identity

$$\langle g\omega, I \rangle = \langle \omega, g^{-1}I \rangle \quad \text{for all } I \in V, \omega \in V^*, g \in G. \quad (2.50)$$

The *cotangent lift momentum map* $\diamond : V \times V^* \rightarrow \mathfrak{g}^*$ for the action of G on V is defined by the identity

$$\langle I \diamond \omega, \xi \rangle = \langle \omega, \xi I \rangle, \quad \text{for all } I \in V, \omega \in V^*, \xi \in \mathfrak{g}, \quad (2.51)$$

where the brackets on both sides represent the duality pairings of the respective spaces $\mathfrak{g}$ and V , and where ξI denotes the infinitesimal action of $\mathfrak{g}$ on V , defined as $\xi I := \frac{d}{dt}\big|_{t=0} g(t)I \in V$ for any C^1 curve $g : [-\varepsilon, \varepsilon] \rightarrow G$ that satisfies $g(0) = e$ and $\frac{d}{dt}\big|_{t=0} g(t) = \xi \in \mathfrak{g}$. We remark in passing that general momentum maps are often denoted by the letter J , however the diamond notation above is customary for the special case of vector spaces. Later in this thesis we will encounter momentum maps in various more general settings. Note that equations (2.50) and (2.51) imply

$$\text{Ad}_{g^{-1}}^*(I \diamond \omega) = gI \diamond g\omega. \quad (2.52)$$

For the flow defined in (2.49), we also introduce the notation

$$g_{t,s}^\xi := g^\xi(t) (g^\xi(s))^{-1}. \quad (2.53)$$

Lemma 2.5 in [37], which is an adaptation from [57] and [36], gives the derivative of the flow at a given time with respect to a variation $(\varepsilon, t) \mapsto \xi_\varepsilon(t) = \xi(t) + \varepsilon \delta \xi(t) \in \mathfrak{g}$ of a smooth curve $\xi = \xi_0$. Namely,

$$\delta g_{t,s}^\xi := \frac{d}{d\varepsilon} \bigg|_{\varepsilon=0} g_{t,s}^{\xi_\varepsilon} = g_{t,s}^\xi \int_s^t \left(\text{Ad}_{g_{s,r}^\xi} \delta \xi(r) \right) dr \in T_{g_{t,s}^\xi} G. \quad (2.54)$$

For completeness we include the proof of this formula, following [37]. Recall from (2.49) and (2.53) that we have, for all ε ,

$$\frac{d}{dt}g_{t,s}^{\xi_\varepsilon} = \xi_\varepsilon(t)g_{t,s}^{\xi_\varepsilon}, \quad g_{s,s}^{\xi_\varepsilon} = e.$$

Taking a derivative with respect to ε we obtain

$$\left. \frac{d}{dt} \frac{d}{d\varepsilon} \right|_{\varepsilon=0} g_{t,s}^{\xi_\varepsilon} = \frac{d}{dt} \delta g_{t,s}^\xi = \delta \xi(t) g_{t,s}^\xi + \xi(t) \delta g_{t,s}^\xi.$$

Hence,

$$\begin{aligned} \frac{d}{dt}(g_{t,s}^\xi)^{-1} \delta g_{t,s}^\xi &= -(g_{t,s}^\xi)^{-1} \left(\frac{d}{dt} g_{t,s}^\xi \right) (g_{t,s}^\xi)^{-1} \delta g_{t,s}^\xi + (g_{t,s}^\xi)^{-1} \frac{d}{dt} \delta g_{t,s}^\xi \\ &= -(g_{t,s}^\xi)^{-1} \xi(t) g_{t,s}^\xi (g_{t,s}^\xi)^{-1} \delta g_{t,s}^\xi + (g_{t,s}^\xi)^{-1} \left[\delta \xi(t) g_{t,s}^\xi + \xi(t) \delta g_{t,s}^\xi \right] = \text{Ad}_{g_{s,t}^\xi} \delta \xi(t). \end{aligned}$$

To arrive at (2.54) we integrate both sides, noting that $\delta g_{s,s}^\xi = 0$; and multiply the result by $g_{t,s}^\xi$ from the left.

Importantly, (2.54) also holds for the diffeomorphism case in a non-smooth setting. For more information on this aspect we refer to [58, Chaps. 8 & 9], where the assumption is $\xi \in L^2([0, t_l], \mathcal{B})$, for some space $\mathcal{B}$ of sufficiently smooth (in space) vector fields.

Formula (2.54) and equation (2.52) are the key ingredients needed to take variations of the similarity measure in (2.48). With these preparations it is now straightforward to adapt the calculations done in the proof of Theorem 2.5 of [37] to our case, in order to show the following theorem.

Theorem 2.11. *A curve $\xi \in \mathcal{P}_{k-1}$ is a stationary point for the functional E , i.e., $\delta E = 0$ if and only if (I), (II), and (III) below hold:*

(I) *For t in any of the open intervals $(0, t_1), \dots, (t_{l-1}, t_l)$,*

$$\sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}} = - \sum_{i=1}^l 1_{t \leq t_i} \left(g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right), \quad (2.55)$$

where π^i is defined by

$$\pi^i := \frac{1}{\sigma^2} \left(g_{t_i,0}^\xi T_0 - I_{t_i} \right)^\flat \in V^*,$$

and we set $1_{a \leq b}$, for any $a, b \in \mathbb{R}$, to be equal to 1 if $a \leq b$ and 0 otherwise.

(II) For $i = 1, \dots, l-1$ and $r = 0, \dots, k-2$,

$$\lim_{t \rightarrow t_i^-} \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t) = \lim_{t \rightarrow t_i^+} \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t). \quad (2.56)$$

(III) For $r = 0, \dots, k-2$,

$$\sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_l) = 0. \quad (2.57)$$

Note that there is no condition at $t_0 = 0$ analogous to (III) because of the fixed end point conditions $\xi^{(j)}(0) = \xi_0^j$, for $j = 0, \dots, k-2$.

Proof. Set $t_0 = 0$ for convenience. We also introduce shorthand notation $f(t_i^\pm)$ for the value of a function $f(t)$ as t approaches t_i from above or below, respectively.

A series of partial integrations taking into account the fixed end point conditions $\xi^{(j)}(0) = \xi_0^j$, $j = 0, \dots, k-2$, leads to

$$\begin{aligned} \delta \int_0^{t_l} \ell dt &= \sum_{i=0}^{l-1} \int_{t_i}^{t_{i+1}} \sum_{j=0}^{k-1} \left\langle \frac{\delta \ell}{\delta \xi^{(j)}}, \delta \xi^{(j)} \right\rangle dt \\ &= \sum_{i=0}^{l-1} \int_{t_i}^{t_{i+1}} \left\langle \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}}(t), \delta \xi(t) \right\rangle dt \\ &\quad + \sum_{i=1}^{l-1} \sum_{r=0}^{k-2} \left(\left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \left(\frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^-) - \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^+) \right), \delta \xi^{(r)}(t_i) \right\rangle \right. \\ &\quad \left. + \left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_l), \delta \xi^{(r)}(t_l) \right\rangle \right). \end{aligned} \quad (2.58)$$

Note that the hypothesis $\xi \in \mathcal{P}_{k-1}$ is sufficient to give meaning to the previous formula.

On the other hand, using formula (2.54) and mimicking the computations done in [37], one finds for the variation of the similarity measure⁶

$$\delta \left(\frac{1}{2\sigma^2} \sum_{i=1}^l \|g^\xi(t_i)T_0 - I_{t_i}\|_V^2 \right) = \int_0^{t_l} \left\langle \sum_{i=1}^l 1_{t \leq t_i} \left(g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right), \delta \xi(t) \right\rangle dt. \quad (2.59)$$

⁶The detailed computation is as follows:

$$\begin{aligned} \delta \left(\frac{1}{2\sigma^2} \sum_{i=1}^l \|g^\xi(t_i)T_0 - I_{t_i}\|_V^2 \right) &= \sum_{i=1}^l \left\langle \left(g_{t_i,0}^\xi \int_0^{t_i} \text{Ad}_{g_{0,r}^\xi} \delta \xi(r) dr \right) T_0, \pi^i \right\rangle \\ &= \int_0^{t_l} \sum_{i=1}^l 1_{r \leq t_i} \left\langle \left(\text{Ad}_{g_{0,r}^\xi} \delta \xi(r) \right) T_0, g_{0,t_i}^\xi \pi^i \right\rangle dr = \int_0^{t_l} \sum_{i=1}^l 1_{r \leq t_i} \left\langle \delta \xi(r) \left(g_{r,0}^\xi T_0 \right), g_{r,0}^\xi g_{0,t_i}^\xi \pi^i \right\rangle dr \end{aligned}$$

Assembling the two contributions to δE , we arrive at

$$\begin{aligned} \delta E = & \sum_{s=0}^{l-1} \int_{t_s}^{t_{s+1}} \left\langle \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}}(t) + \sum_{i=1}^l 1_{t \leq t_i} \left(g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right), \delta \xi(t) \right\rangle dt \\ & + \sum_{i=1}^{l-1} \sum_{r=0}^{k-2} \left(\left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \left(\frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^-) - \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^+) \right), \delta \xi^{(r)}(t_i) \right\rangle \right. \\ & \left. + \sum_{r=0}^{k-2} \left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_l), \delta \xi^{(r)}(t_l) \right\rangle \right). \end{aligned} \quad (2.60)$$

Stationarity $\delta E = 0$ therefore leads to equations (2.55)–(2.57). $\square$

Remark 2.12. *The right-hand side of equation (2.55) follows coadjoint motion on every open interval $(0, t_1), \dots, (t_{l-1}, t_l)$. That is,*

$$\left(\frac{d}{dt} + \text{ad}_\xi^* \right) \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}} = 0, \quad (2.61)$$

in which we once again recognise the higher-order Euler–Poincaré equation (2.27).

2.5.3 Two examples of interest for computational anatomy

Regarding potential applications in computational anatomy, an interesting property of higher-order models is the gain in (temporal) smoothness of the optimal path $T : t \in [0, t_l] \mapsto g^\xi(t)T_0 \in V$, in comparison with first-order models. For instance, in the case of piecewise-geodesic (i.e., first-order) interpolation, where $\ell(\xi) := \frac{1}{2}\|\xi\|_{\mathfrak{g}}^2$, equation (2.55) reads

$$\xi(t) = - \sum_{i=1}^l 1_{t \leq t_i} \left(g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right)^\sharp. \quad (2.62)$$

In general therefore, ξ will be discontinuous at each time point t_i for $i < l$, which implies non-differentiability of T at these points. In contrast, for the Lagrangian $\ell_1(\dot{\xi}) := \frac{1}{2}\|\dot{\xi}\|_{\mathfrak{g}}^2$, equation (2.55) becomes

$$\ddot{\xi}(t) = \sum_{i=1}^l 1_{t \leq t_i} \left(g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right)^\sharp. \quad (2.63)$$

$$= \int_0^{t_l} \sum_{i=1}^l 1_{r \leq t_i} \left\langle g_{r,0}^\xi T_0 \diamond g_{r,t_i}^\xi \pi^i, \delta \xi(r) \right\rangle dr = \int_0^{t_l} \left\langle \sum_{i=1}^l 1_{r \leq t_i} \left(g_{r,0}^\xi T_0 \diamond g_{r,t_i}^\xi \pi^i \right), \delta \xi(r) \right\rangle dr,$$

where we used (2.54), (2.50) and (2.51).

Now the curves $\xi(t)$ and $T(t)$ are C^1 and C^2 on $[0, t_l]$, respectively. This situation is similar to the *exact* interpolation problems of [19, 31] discussed earlier in Section 2.2.3.⁷

Note also that the minimisation of the functional E for ℓ_1 when $l = 1$ (one target only) produces Lie-exponential solutions on G , as long as the initial conditions are chosen properly. More precisely, if the Lie-exponential map is surjective and the action of G on V is transitive, then there exists $\xi_0 \in \mathfrak{g}$ such that $\exp(t_1 \xi_0)T_0 = I_{t_1}$. Hence, the constant curve $\xi \equiv \xi_0$ is a minimiser of the functional E , with $E[\xi] = 0$. The Lie-exponential has been widely used in computational anatomy, for instance in [59, 60].

Another Lagrangian of interest for computational anatomy is $\ell_2(\xi, \dot{\xi}) := \frac{1}{2} \|\dot{\xi} + \text{ad}_\xi^\dagger \xi\|_{\mathfrak{g}}^2$, which measures the acceleration on the Lie group for the right-invariant metric induced by the norm $\|\cdot\|_{\mathfrak{g}}$. The Lagrangian ℓ_2 may therefore have more geometrical meaning than ℓ_1 . However, the definition of ℓ_2 requires special attention in the case of diffeomorphism groups. Roughly speaking, the difficulty arises because the reduced geodesic equation involves spatial derivatives of the trivialised velocity curve (which in this case is a time-dependent vector field), leading to a loss of regularity. In general then, ℓ_2 is not well-defined, since the norm $\|\cdot\|_{\mathfrak{g}}$ may not be defined on $\dot{\xi} + \text{ad}_\xi^\dagger \xi$. One way of dealing with this complication is to work with two different norms; one to define the covariant acceleration, the other to measure it [61]. A detailed investigation of these topics is beyond the scope of this thesis and will be left for future work.

2.5.4 Template matching on the sphere

Consider as a finite-dimensional example $G = SO(3)$ with norm $\|\Omega\|_{\mathfrak{so}(3)} = \sqrt{\Omega \cdot \mathbf{I} \Omega}$ on the Lie algebra $\mathfrak{so}(3)$, where $\mathbf{I}$ is a symmetric positive-definite matrix (the moment of inertia tensor). Let $V = \mathbb{R}^3$ with $\|\cdot\|_{\mathbb{R}^3}$ the Euclidean distance. We would like to interpolate a time sequence of points on the unit sphere $S^2 \subset \mathbb{R}^3$ starting from the

⁷In [39] it was stated that the exact second-order interpolation method of [31] lead to C^1 , as opposed to C^2 , solution curves. In fact, solution curves are C^2 .

template $T_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Choose the times to be $t_i = \frac{1}{5}i$ for $i = 1, \dots, 5$ and

$$I_{t_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad I_{t_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad I_{t_3} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad I_{t_4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad I_{t_5} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

The associated minimisation problem for a given Lagrangian $\ell(\boldsymbol{\Omega}, \dots, \boldsymbol{\Omega}^{(k-1)})$ is:

Minimise

$$E[\boldsymbol{\Omega}] := \int_0^1 \ell(\boldsymbol{\Omega}(t), \dots, \boldsymbol{\Omega}^{(k-1)}(t)) dt + \frac{1}{2\sigma^2} \sum_{i=1}^5 \|\Lambda^{\boldsymbol{\Omega}}(t_i)T_0 - I_{t_i}\|_{\mathbb{R}^3}^2, \quad (2.64)$$

subject to the conditions $\boldsymbol{\Omega}^{(j)}(0) = \boldsymbol{\Omega}_0^j$, $j = 0, \dots, k-2$, where $\Lambda^{\boldsymbol{\Omega}}(t)$ is a continuous curve defined by

$$\Lambda^{\boldsymbol{\Omega}}(0) = e, \quad \text{and} \quad \frac{d}{dt} \Lambda^{\boldsymbol{\Omega}}(t) = \Omega(t) \Lambda^{\boldsymbol{\Omega}}(t),$$

whenever t is in one of the open intervals $(0, t_1), \dots, (t_4, t_5)$. As we mentioned in Section 2.5.3, an important property of higher-order models is the increase in smoothness of the optimal path when compared with first-order models. We illustrate this behaviour in Figures 2.2 and 2.3 below.

Figure 2.2 shows the interpolation between the given points $I_{t_1}, \dots, I_{t_5}$ for the first-order Lagrangian

$$\ell(\boldsymbol{\Omega}) = \frac{1}{2} \boldsymbol{\Omega} \cdot \mathbf{I} \boldsymbol{\Omega}. \quad (2.65)$$

We contrast this with the second-order model

$$\ell(\boldsymbol{\Omega}, \dot{\boldsymbol{\Omega}}) = \frac{1}{2} \left(\dot{\boldsymbol{\Omega}} + \mathbf{I}^{-1}(\boldsymbol{\Omega} \times \mathbf{I} \boldsymbol{\Omega}) \right) \cdot \mathbf{I} \left(\dot{\boldsymbol{\Omega}} + \mathbf{I}^{-1}(\boldsymbol{\Omega} \times \mathbf{I} \boldsymbol{\Omega}) \right). \quad (2.66)$$

Note that this is the reduced Lagrangian for Riemannian cubics on $SO(3)$, as we discussed in Section 2.3, and for $\mathbf{I} = e$ we recognise equation (2.61) to be the NHP equation (2.44).

Figure 2.3 visualises the resulting interpolation for two different choices of the moment of inertia tensor $\mathbf{I}$, namely

$$\mathbf{I}_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{I}_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.67)$$

In order to compare the two cases we have normalised $\mathbf{I}_2$ in such a way that it has the same norm as $\mathbf{I}_1$ with respect to $\|\mathbf{I}\|^2 = \text{tr}(\mathbf{I}^\top \mathbf{I})$. The figures were obtained by minimising the functional E using the downhill simplex algorithm *fmin_tnc* that is included in the *optimize* package of SciPy, [62].

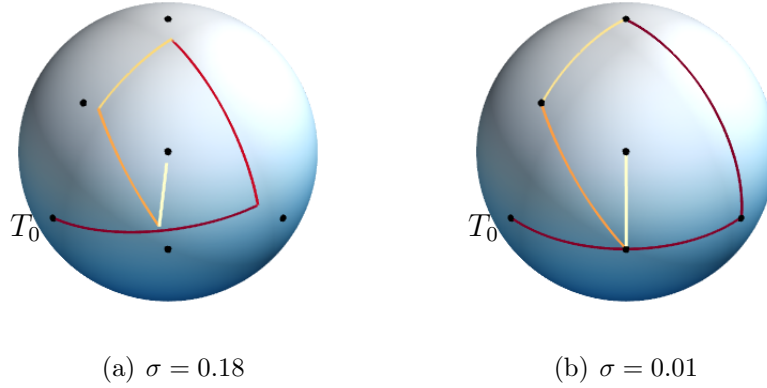


Fig. 2.2: *First-order template matching.* Results are shown for the Lagrangian (2.65) with $\mathbf{I} = e$, for two different values of tolerance σ . These values have been chosen so that the sum of the mismatch penalties is similar in size to the one obtained in the second-order template matching shown in Figure 2.3. As might be expected, when the tolerance is smaller, the first-order curves pass nearer their intended target points. These first-order curves possess jumps in tangent directions at the beginning of each new time interval. *This figure appears in [39] – reproduced with kind permission from Springer Science and Business Media.*

2.6 Final remarks

This chapter consisted of a first look at invariant higher-order variational problems on Lie groups. We introduced the higher-order generalisation of the Euler–Poincaré reduction method and derived corresponding Euler–Poincaré equations. Special attention was given to the important example of Riemannian cubics on Lie groups with right-, or left-invariant metrics. As an application we discussed higher-order template matching, where one seeks an optimal curve that interpolates given data points at prescribed target times. The formulation of the optimality condition involved both a Lie group acting on the data vector space and the data vector space itself.

Let us conclude the chapter with a number of remarks.

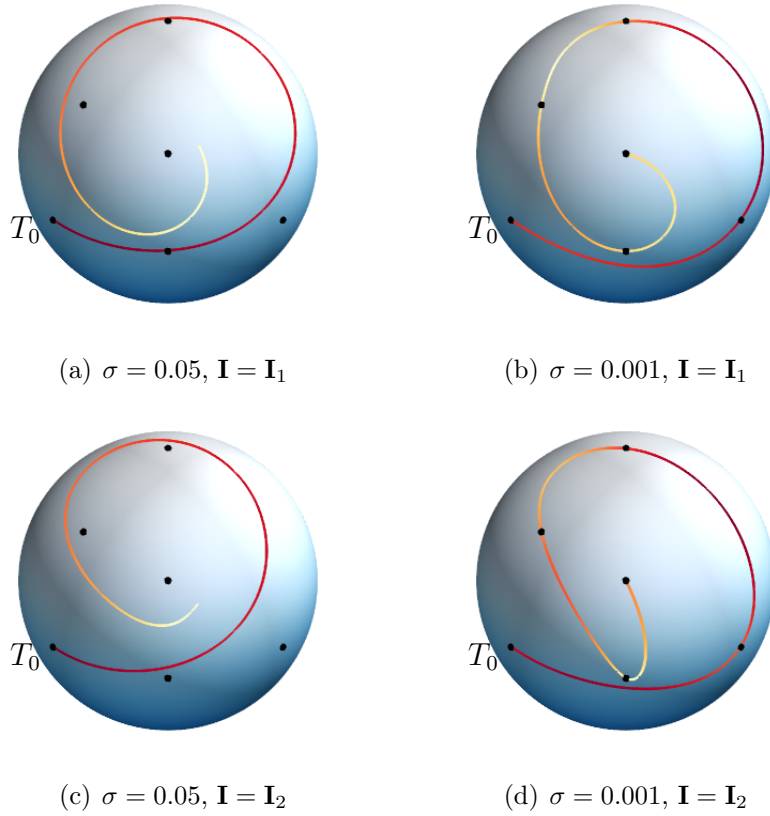


Fig. 2.3: *Second-order template matching.* The pictures in the top row show the template matching for the Lagrangian (2.65) with $\mathbf{I}_1$ for two different values of the tolerance σ . The bottom row represents the corresponding matching results for $\mathbf{I}_2$. One observes that the quality of matching increases as the tolerance decreases. This is due to the increased weight on the penalty term in (2.48). The color of the curves represents the magnitude of the velocity vector of the curve on the sphere (red is large, white is small). We fixed the initial angular velocity $\boldsymbol{\Omega}(0) = \frac{5\pi}{2}(0, 0, 1)$. On comparing these figures with those in the first-order case, one observes that the second-order method produces smoother curves. *This figure appears in [39] – reproduced with kind permission from Springer Science and Business Media.*

(1) One can verify directly that the higher-order Euler–Poincaré equations (2.27) imply

$$\frac{d}{dt} \text{Ad}_{g^{\pm 1}}^* \left[\sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} \right] = 0,$$

for right-, respectively left-invariant Lagrangians. In Chapter 4 we will show that this conservation arises, via Noether’s theorem, from the group-invariance of the higher-order Lagrangian.

- (2) One can similarly verify that

$$\frac{d}{dt} \left[\sum_{r=0}^{k-1} \langle \mu^r, \xi^{(r)} \rangle - \ell(\xi, \dots, \xi^{(k-1)}) \right] = 0,$$

where

$$\mu^r = \sum_{j=0}^{k-r-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(r+j)}} \quad (r = 0, \dots, k-1).$$

In Chapter 4 we will interpret the μ^r as higher-order Legendre momenta and the conserved quantity as the Hamiltonian of the higher-order system.

- (3) Our paper [39] contains the higher-order Euler–Poincaré equations and also discusses a number of other contexts. In particular, a higher-order version is given of the theory of parameter-dependent Lagrangians in the sense of [6]. Moreover, the reduced Hamiltonian formulation is given in terms of a higher-order Legendre mapping, also called the *Legendre–Ostrogradsky transform*. In this thesis we will instead develop the Hamiltonian side of the theory from the point of view of the so-called Hamilton–Pontryagin variational principle. This will be done in Chapter 4.
- (4) In the same chapter we will revisit in much more detail the higher-order template matching problem whose treatment we started above. We will recast it in a more general form using arbitrary object manifolds rather than just vector spaces. As we will see, this widens the range of potential applications of the theory. Moreover, we will provide a substantially simplified derivation of the Euler–Lagrange equations using Lagrange multipliers. In the process we will in particular re-discover, as a byproduct, the Legendre–Ostrogradsky transform of higher-order mechanics.
- (5) Finally, we mention two papers that are closely related to the developments in this chapter. The authors of [63] independently derived higher-order Euler–Poincaré equations taking the canonical Hamiltonian formulation as their starting point. Their paper also contains an extension to Lagrangians defined on constraint submanifolds of higher-order tangent bundles, with applications to underactuated control systems on Lie groups. The second paper we mention is [50], which develops higher-order Lagrange–Poincaré reduction and its Hamiltonian counterpart, Hamilton–Poincaré

reduction. We will take advantage of higher-order Lagrange–Poincaré reduction in the next chapter.

In Section 2.5 we derived necessary conditions for the optimal curve $g(t) \in G$ in the higher-order template matching problem. We found in Remark 2.12 that on the open intervals between target times $g(t)$ satisfies higher-order Euler–Poincaré equations. The presence of an object manifold, the vector space V in this case, naturally leads to a number of questions about the relationship between the optimal curve $g(t)$ and the corresponding trajectory $g(t)T_0 \in V$. This will be the topic of the next chapter.

3 Invariant higher-order variational problems, Part II

At the end of Section 1.1 we briefly mentioned an interesting area of investigation that presents itself when working with object manifolds. Let Q be an object manifold acted on by a Lie group of transformations G . Suppose some variational problem involving both quantities in Q and G leads to an optimal curve $g(t) \in G$ with $g(0) = e$, which induces a curve $g(t)q \in Q$ on the object manifold, for some initial point $q \in Q$. How are the two curves $g(t)$ and $g(t)q$ related? Of course, the answer depends on the particular situation one considers.

For example, as we pointed out in Section 1.1, the LDM method of computational anatomy, which is a first-order variational method, leads to a (horizontal) geodesic in the group of diffeomorphisms carrying some initial shape into a target shape. The corresponding path in the object manifold of shapes is also a geodesic, with respect to an induced metric that derives, by means of the group action, from the metric on the diffeomorphism group. In geometric mechanics these types of metrics are known as *normal metrics*, and they will play a central role in this chapter.

Here we shall start the investigation of such questions in the context of higher-order variational principles. More precisely, we are concerned with Riemannian cubics on Lie groups and object manifolds (endowed with normal metrics), focusing on their lifting and projection properties. First, let us give some more background on why such considerations are important.

In addressing an interpolation problem on the object manifold, two distinct strategies offer themselves. First, one may choose to define a variational principle on the Lie group, or indeed its Lie algebra, and find an optimal path $g(t)$ that transforms the initial shape q as $q(t) = g(t)q$, such that $q(t)$ passes through the prescribed configurations. This type of higher-order interpolation was proposed in regard to applications in computational anatomy in Section 2.5 of the previous chapter. Alternatively, one may define a variational principle on shape space itself, without making explicit reference to any group action, and find an optimal curve that interpolates the given shapes. This was the approach of [16], where interpolation by Riemannian cubics on shape space was proposed, and existence results for the shape space of landmarks were given.

The particular cost functionals that interest us here are, on the group,

$$S_G[g] = \int_0^1 \left\| \frac{D}{Dt} \dot{g} \right\|_g^2 dt,$$

and on the object manifold,

$$S_Q[q] = \int_0^1 \left\| \frac{D}{Dt} \dot{q} \right\|_q^2 dt.$$

Hamilton's principle, $\delta S = 0$, leads to Riemannian cubics in the respective manifolds G and Q . One is therefore naturally led to the question of how Riemannian cubics on the group of transformations are related to those on the object manifold. This is a new question in geometric mechanics and its answer is potentially important in applications of computational anatomy. We emphasise however that for the purposes of this chapter we consider finite-dimensional manifolds only.

We first analyse horizontal lifts of cubics on the object manifold to the group of transformations. In the context of so-called *Type I* symmetric spaces⁸, we completely characterise the class of cubics on the object manifold that can be lifted horizontally to cubics on the group of transformations. For rank-one symmetric spaces this selects geodesics reparametrised in time with cubic polynomials. We then study non-horizontal curves in G . We show that certain types of non-horizontal geodesics project to cubics in Q . Finally, we present the theory of second-order Lagrange–Poincaré reduction for Riemannian cubics in the group of transformations. The reduced form of the equations reveals the obstruction for such a cubic to project to a cubic on the object manifold. Most results of this chapter appear in [40].

3.1 Main content of the chapter

The main content of the chapter may be summarised as follows:

In Section 3.2 we outline the geometric setting for the present investigation of Riemannian cubics for normal metrics and their relation to Riemannian cubics on the Lie group of transformations. That is, we summarise the definition of normal metrics

⁸See the footnote on page 73.

and recall that the projection $G \rightarrow Q$ which maps an element of the group to the group transformation of a reference object is a Riemannian submersion.

In Section 3.3 we provide the key expressions for covariant derivatives of curves and vector fields along curves, both in Lie groups and in object manifolds with normal metric. The horizontal generator of a curve in the object manifold is introduced and expressed in terms of the momentum map of the cotangent lifted action.

In Section 3.4 we derive the equations of Riemannian cubics for normal metrics. For ease of exposition we first consider a more general context and then specialise to the case of Riemannian cubics. Here the horizontal generator plays a crucial role. Invariant metrics on Lie groups are a simple example of normal metrics, as are the metrics on Type I symmetric spaces. These examples are worked out in detail. For the convenience of the reader we will also briefly recall from the previous chapter how Riemannian cubics on Lie groups can be treated equivalently by Euler–Poincaré reduction. Our derivation of the Euler–Lagrange equations bypasses any mention of curvature. Therefore, these equations can also be used to *compute* curvatures by means of the general equation for Riemannian cubics (2.10). This is demonstrated by two simple examples, whose well-known curvatures we recover.

In Section 3.5 we study horizontal lifting properties of Riemannian cubics. Our form of the Euler–Lagrange equations is particularly well-suited for this task, due to the appearance of the horizontal generator of curves. We characterise the cubics in Type I symmetric spaces that can be lifted horizontally to cubics in the group of isometries. We then proceed to the more general situation of a Riemannian submersion and state necessary and sufficient conditions under which a cubic on the object manifold lifts horizontally to a cubic on the Lie group of transformations.

In Section 3.6 we extend the previous considerations to include non-horizontal curves on the Lie group. We show that certain non-horizontal geodesics on the group of transformations project to cubics on the object manifold. We then reduce the Riemannian cubic variational problem on the group by the isotropy subgroup of a reference object. To achieve this, we use higher-order Lagrange–Poincaré reduction

[7, 50]. The reduced Lagrangian couples horizontal and vertical parts of the motion, which explains the absence of a general horizontal lifting property for cubics. Namely, the reduced equations that describe Riemannian cubics on the Lie group contain the equation that characterises Riemannian cubics on the object manifold, plus extra terms. These extra terms represent the obstruction for a cubic on the Lie group to project to a cubic on the object manifold. In this sense, the reduced equations fully describe the relation between cubics on the Lie group and cubics on the object manifold. They also lend themselves to further study of the questions investigated in the present chapter.

This is a long chapter. The reader may find it useful to read the summary pages 146–148 in parallel, in order to track his or her progress.

3.2 Geometric setting

We begin with some preparations for the later developments in the chapter. In particular, we recall some background on group actions, normal metrics, connectors and Riemannian submersions.

3.2.1 Group actions

Let G be a Lie group with Lie algebra $\mathfrak{g}$, acting from the *left* on a smooth manifold Q (the object manifold). We denote the action by

$$\Phi : G \times Q \rightarrow Q, \quad (g, q) \mapsto gq := \Phi_g(q). \quad (3.1)$$

The *infinitesimal generator* of the action corresponding to $\xi \in \mathfrak{g}$ is the vector field on Q given by

$$\xi_Q(q) := \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)q. \quad (3.2)$$

In accordance with (2.4), the tangent lift of Φ is defined as the action of G on TQ ,

$$G \times TQ \rightarrow TQ, \quad (g, v_q) \mapsto gv_q := T\Phi_g(v_q), \quad (3.3)$$

with infinitesimal generator ξ_{TQ} corresponding to $\xi \in \mathfrak{g}$. Note that we have the relation

$$T\tau_Q(\xi_{TQ}(v_q)) = \xi_Q(q), \quad (3.4)$$

where $\tau_Q : TQ \rightarrow Q$ is the tangent bundle projection. Similarly, one defines the cotangent lifted action as

$$G \times T^*Q \rightarrow T^*Q, \quad (g, \alpha_q) \mapsto g\alpha_q := (T\Phi_{g^{-1}})^*(\alpha_q). \quad (3.5)$$

The momentum map $J : T^*Q \rightarrow \mathfrak{g}^*$ associated with the cotangent lift of Φ is given by

$$\langle J(\alpha_q), \xi \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = \langle \alpha_q, \xi_Q(q) \rangle_{T^*Q \times TQ}, \quad (3.6)$$

for arbitrary $\alpha_q \in T^*Q$ and $\xi \in \mathfrak{g}$. Notice that in (2.51) we already encountered a cotangent lift momentum map. There the manifold Q was a vector space, V , and we used the diamond notation customary in the vector case, that is, $J(\omega, I) = I \diamond \omega$ for $I \in V$ and $\omega \in V^*$.

3.2.2 Normal metrics

Let G be a Lie group acting *transitively* from the left on a smooth manifold Q . Transitivity of the action means in particular that Q is a G -orbit. Let γ_G be a right-invariant Riemannian metric on G . We will now use the action of G on Q in order to induce a metric γ_Q on Q . To do this, define a pointwise inner product on tangent spaces T_qQ by

$$\gamma_Q(v_q, v_q) := \min_{\{\xi \in \mathfrak{g} \mid \xi_Q(q) = v_q\}} \{\gamma_G(\xi, \xi)\}. \quad (3.7)$$

We refer to [58] for a rigorous treatment of the infinite dimensional case of diffeomorphism groups. We define the *vertical subspace of $\mathfrak{g}$ at q* as

$$\mathfrak{g}_q = \{\xi \in \mathfrak{g} \mid \xi_Q(q) = 0\} \quad (3.8)$$

and the *horizontal subspace* as the orthogonal complement $\mathfrak{g}_q^\perp$. Denote the orthogonal projection onto $\mathfrak{g}_q^\perp$ by $\xi \mapsto H_q(\xi)$. This projection operation depends smoothly on $q \in Q$ (see [40] for details). The vertical projection is similarly written as $\xi \mapsto V_q(\xi)$. Let

$\nu_1, \dots, \nu_k$ be an orthonormal basis of $\mathfrak{g}_q$. For v_q in $T_q Q$ and ξ any generator of v_q , i.e. $\xi_Q(q) = v_q$, we can write

$$\begin{aligned} \gamma_Q(v_q, v_q) &= \min_{\lambda^i \in \mathbb{R}} \left\{ \gamma_G \left(H_q(\xi) + \lambda^i \nu_i, H_q(\xi) + \lambda^j \nu_j \right) \right\} \\ &= \min_{\lambda^i \in \mathbb{R}} \left\{ \gamma_G(H_q(\xi), H_q(\xi)) + \sum_{i=1}^k (\lambda^i)^2 \right\} \\ &= \gamma_G(H_q(\xi), H_q(\xi)). \end{aligned} \tag{3.9}$$

The pair (Q, γ_Q) with γ_Q defined pointwise by (3.7) is a Riemannian manifold. The metric γ_Q is called a *normal metric* or *projected metric*. In particular, it coincides with the normal metric considered in [64].

3.2.3 The connector of a Riemannian metric

From (3.4), it follows that $T\tau_Q(\xi_{TQ}(v_q)) = 0_q$, for all $\xi \in \mathfrak{g}_q$. Let $T(TQ) \supset V_{v_q}(TQ) := \ker(T_{v_q}\tau_Q)$ be the vertical space at v_q . Intuitively, these are elements of $T_{v_q}TQ$ that are tangent to curves that remain in $\tau_Q^{-1}(q)$, to first order. Then $\xi_{TQ}(v_q) \in V_{v_q}(TQ)$ can be identified with an element of $T_q Q$ in the standard way: $\tau : V_{v_q}(TQ) \ni \xi_{TQ}(v_q) \xrightarrow{\sim} w_q \in T_q Q$, where w_q is the unique vector satisfying $\xi_{TQ}(v_q) = \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} (v_q + \varepsilon w_q)$.

Let γ_Q be a Riemannian metric on Q and define the *connector* of γ_Q to be the intrinsic map $K : TTQ \rightarrow TQ$ given in coordinates as

$$K_{\text{loc}}(x, w, u, v) := (x, v + \Gamma(x)(w, u)), \tag{3.10}$$

where Γ are the Christoffel symbols of the metric.

Recall that TTQ has two vector bundle structures with base TQ , namely the standard tangent bundle structure $\tau_{TQ} : TTQ \rightarrow TQ$ and $T\tau_Q : TTQ \rightarrow TQ$, where $\tau_Q : TQ \rightarrow Q$ is the tangent bundle projection. In coordinates, these maps are given by $\tau_{TQ}(x, w, u, v) = (x, w)$ and $T\tau_Q(x, w, u, v) = (x, u)$. In addition, TTQ has the canonical involution $\kappa_Q : TTQ \rightarrow TTQ$ which, in standard local charts, is given by $\kappa_Q(x, w, u, v) = (x, u, w, v)$. The connector is linear relative to both vector bundle structures of TTQ and it is symmetric, i.e., $K \circ \kappa_Q = K$, which is equivalent to the vanishing of the torsion of the Levi-Civita connection. A key property of the connector is given by the formula

$$\nabla_Y X = K \circ TX \circ Y$$

for any $X, Y \in \mathfrak{X}(Q)$. Moreover, the restriction of K to vertical spaces $V_{v_q} \subset TTQ$ coincides with $\tau : V_{v_q}(TQ) \xrightarrow{\sim} T_q Q$ at every $v_q \in T_q Q$, since $K_{\text{loc}}(x, u, 0, v) = (x, v) = \tau_{\text{loc}}(x, u, 0, v)$. For $\xi \in \mathfrak{g}_q$, one therefore obtains

$$\tau(\xi_{TQ}(v_q)) = K(\xi_{TQ}(v_q)) = K(T\xi_Q(q)(v_q)) = \nabla_{v_q} \xi_Q. \quad (3.11)$$

More information on the connector can be found, for example, in [65, Chap. 22.8, 22.9].

3.2.4 Riemannian submersion property

For (G, γ_G) and (Q, γ_Q) as above, fix $q_0 \in Q$, and consider the principal bundle projection

$$\Pi : G \rightarrow Q, \quad g \mapsto gq_0. \quad (3.12)$$

We decompose the tangent bundle of G into horizontal and vertical subbundles, $TG = HG \oplus VG$. The vertical space at $g \in G$ is defined as $V_g G := \ker T_g \Pi$ and the horizontal space as $H_g G := (V_g G)^\perp$, where the orthogonal complement is taken with respect to the right-invariant Riemannian metric γ_G . These spaces are translations of appropriate subspaces of $\mathfrak{g}$,

$$V_g G = T_e L_g(\mathfrak{g}_{q_0}) = T_e R_g(\mathfrak{g}_q), \quad \text{and} \quad H_g G = T_e R_g(\mathfrak{g}_q^\perp), \quad (3.13)$$

where $q = \Pi(g) = gq_0$. See Figure 3.1 for a schematic representation. Notice that

$$T_g \Pi(T_e R_g \xi) = \xi_Q(\Pi(g)), \quad \text{for all } g \in G, \xi \in \mathfrak{g}, \quad (3.14)$$

which can be written in shorter form as $\Pi_*(\xi_G(g)) = \xi_Q(\Pi(g))$. It is well-known that $\Pi : G \rightarrow Q$ is a Riemannian submersion; i.e., Π is a surjective submersion and

$$\gamma_G(g)(v_g, w_g) = \gamma_Q(q)(T_g \Pi(v_g), T_g \Pi(w_g)) \quad \text{for all } v_g, w_g \in H_g G. \quad (3.15)$$

This property will be useful when we compute covariant derivatives for normal metrics in the next section. Before we continue however, let us remark that if the metric on G is bi-invariant, then one has, in addition to (3.13), that

$$H_g G = T_e L_g(\mathfrak{g}_{q_0}^\perp). \quad (3.16)$$

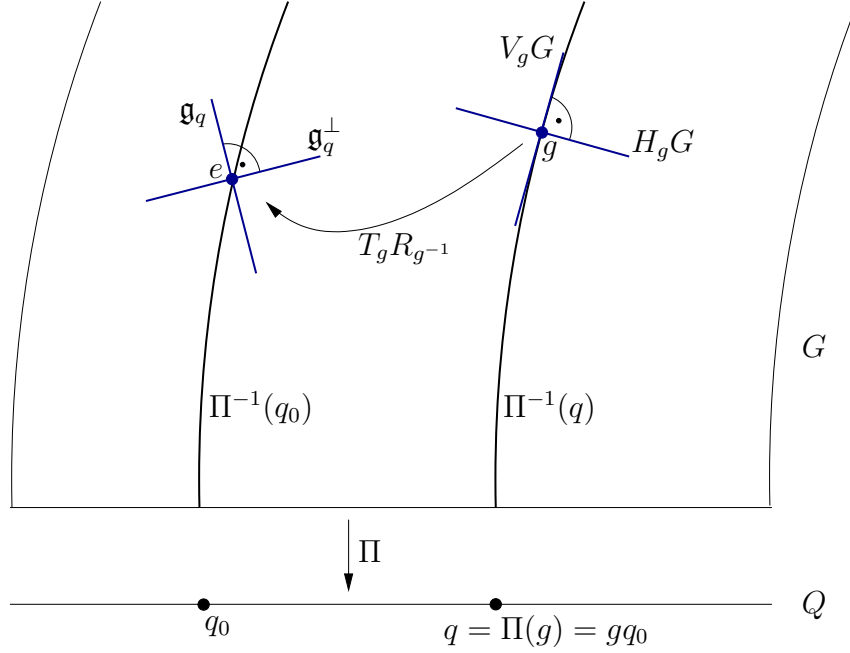


Fig. 3.1: *Schematic representation of horizontal and vertical subbundles.* The vertical bundle $H_g G$ at $g \in G$ consists of all vectors in $T_g G$ that point in the direction of the fibre $\Pi^{-1}(q)$, where $q = \Pi(g) = gq_0$. The horizontal space $H_g G$ is the orthogonal complement of $V_g G$. Applying $T_g R_{g^{-1}}$ to these spaces maps them to $\mathfrak{g}_q$ and $\mathfrak{g}_q^\perp$, respectively, as can be seen from equation (3.13).

3.3 Covariant derivatives

We recall and emphasise that in this chapter we consider *transitive* group actions only. That is, we assume that the object manifold Q consists of a single G -orbit, $\{gq_0 \mid g \in G\} = Q$. The main goal of the current section is to obtain expressions for the covariant derivative of a curve $q(t) \in Q$, where Q is equipped with a normal metric. The strategy is the following. First we compute covariant derivatives of curves in Lie groups with right-invariant metrics. Then we exploit the fact that the projection mapping Π introduced in Section 3.2.4 above is a Riemannian submersion.

For the second step, we recall a formula for the covariant derivative of *horizontal* vector fields for Riemannian submersions. Let $\Pi : (\tilde{Q}, \gamma_{\tilde{Q}}) \rightarrow (Q, \gamma_Q)$ be a Riemannian submersion and denote the covariant derivatives with respect to the Levi-Civita connections on $\tilde{Q}$ and Q by $\tilde{\nabla}$ and ∇ , respectively.

Define the vertical subbundle $V\tilde{Q} := \ker T\Pi \subset T\tilde{Q}$, whose fibre at $\tilde{q} \in \tilde{Q}$ is $V_{\tilde{q}}\tilde{Q} =$

$\ker T_{\tilde{q}}\Pi$. The horizontal subbundle $H\tilde{Q} \subset T\tilde{Q}$ consists of the fibres $H_{\tilde{q}}\tilde{Q} = (V_{\tilde{q}}\tilde{Q})^\perp$. Altogether one obtains a $\gamma_{\tilde{Q}}$ -orthogonal decomposition $T\tilde{Q} = V\tilde{Q} \oplus H\tilde{Q}$.

The *horizontal lift* of a vector field $X \in \mathfrak{X}(Q)$ is the unique horizontal vector field $\tilde{X} \in \mathfrak{X}(\tilde{Q})$ that satisfies $T_{\tilde{q}}\Pi(\tilde{X}(\tilde{q})) = X(\Pi(\tilde{q}))$, for all $\tilde{q} \in \tilde{Q}$. Similarly, a curve $\tilde{q}(t) \in \tilde{Q}$ is called a horizontal lift of a curve $q(t) \in Q$, if its tangent vectors lie in $H\tilde{Q}$ at all times, and $\Pi \circ \tilde{q} = q$.

As we pointed out in Section 3.2.4, the map $\Pi : G \rightarrow Q$ defined in (3.12) is a Riemannian submersion, and the resulting γ_G -orthogonal decomposition of TG was already given in (3.13).

Let $\tilde{X}, \tilde{Y} \in \mathfrak{X}(\tilde{Q})$ be the horizontal lifts of $X, Y \in \mathfrak{X}(Q)$, respectively. Then (see [66]),

$$\tilde{\nabla}_{\tilde{X}}\tilde{Y} = \widetilde{\nabla_X Y} + \frac{1}{2}[\tilde{X}, \tilde{Y}]^V, \quad (3.17)$$

where the superscript V denotes the vertical part. The *horizontal lifting property* of geodesics follows. Namely, if $\tilde{q}(t) \in \tilde{Q}$ is the horizontal lift of a geodesic $q(t) \in Q$, that is, $\nabla_{\dot{q}}\dot{q} = 0$, then $\tilde{q}(t)$ is a geodesic, since $\tilde{\nabla}_{\dot{\tilde{q}}}\dot{\tilde{q}} = \widetilde{\nabla_{\dot{q}}\dot{q}} = 0$. Note also that applying $T\Pi$ to both sides of (3.17) gives

$$T\Pi\left(\tilde{\nabla}_{\tilde{X}}\tilde{Y}\right) = \nabla_X Y. \quad (3.18)$$

3.3.1 Covariant derivatives for normal metrics

The following proposition is a compilation of well-known expressions that will be used extensively in the rest of the chapter. We showed (3.19) in Chapter 2 in the process of proving Proposition 2.8. The proofs of (3.20) and (3.21) can be found, for example, in [51]. The expression (3.24) below for the horizontal generator was also used in [16] and [64].

Proposition 3.1. *Let (G, γ_G) be a Lie group with right-invariant metric, acting transitively from the left on a manifold (Q, γ_Q) with normal metric γ_Q .*

(i) *Let $g(t)$ be a curve in G , and define $\xi(t) \in \mathfrak{g}$ by $\dot{g} = \xi_G(g)$. Then,*

$$\frac{D}{Dt}\dot{g} = \left(\dot{\xi} + \text{ad}_\xi^\dagger \xi\right)_G(g). \quad (3.19)$$

(ii) More generally, let $V(t) \in T_{g(t)}G$ be a vector field along a curve $g(t) \in G$. Define curves $\xi(t), \nu(t) \in \mathfrak{g}$ by $\dot{g} = \xi_G(g)$ and $V = \nu_G(g)$, respectively. Then,

$$\frac{D}{Dt}V = \left(\dot{\nu} + \frac{1}{2} \text{ad}_\xi^\dagger \nu + \frac{1}{2} \text{ad}_\nu^\dagger \xi - \frac{1}{2} [\xi, \nu] \right)_G(g). \quad (3.20)$$

Furthermore, let $M(t) \in T_{g(t)}^*G$ be a covector field along $g(t)$ and define $\mu(t) \in \mathfrak{g}^*$ by $\mu = (TR_g)^*M$. Then,

$$\frac{D}{Dt}M = (TR_{g^{-1}})^* \left(\dot{\mu} - \frac{1}{2} (\text{ad}_\xi(\mu^\sharp))^\flat + \frac{1}{2} \text{ad}_{\mu^\sharp}^* \xi^\flat + \frac{1}{2} \text{ad}_\xi^* \mu \right). \quad (3.21)$$

(iii) Let $q(t)$ be a curve in Q and let $\xi(t) \in \mathfrak{g}$ be a curve satisfying $\dot{q} = \xi_Q(q)$. Then

$$\frac{D}{Dt}\dot{q} = \left(\dot{\xi} + \text{ad}_{H_q(\xi)}^\dagger H_q(\xi) \right)_Q(q) + \nabla_{\dot{q}}(V_q(\xi))_Q, \quad (3.22)$$

In particular, if $\xi(t) \in \mathfrak{g}_{q(t)}^\perp$ is the unique horizontal generator of $q(t)$, then

$$\frac{D}{Dt}\dot{q} = \left(\dot{\xi} + \text{ad}_\xi^\dagger \xi \right)_Q(q). \quad (3.23)$$

(iv) Let $q(t)$ be a curve in Q . The unique horizontal generator of $q(t)$ is given by the Lie algebra element $\bar{J}(\dot{q})$ defined by

$$\bar{J}(\dot{q}) := (J(\dot{q}^\flat))^\sharp \in \mathfrak{g}_q^\perp, \quad (3.24)$$

where J is the cotangent lift momentum map defined in Section 3.2.1. In particular,

$$\frac{D}{Dt}\dot{q} = \left(\partial_t \bar{J}(\dot{q}) + \text{ad}_{\bar{J}(\dot{q})}^\dagger \bar{J}(\dot{q}) \right)_Q(q). \quad (3.25)$$

Proof. For the proof of (i) and (ii) we refer to [51], or the proof of Proposition 2.8 in Chapter 2 above. In order to show (iii) recall the projection mapping $\Pi : G \rightarrow Q$, $g \mapsto ga$, for a fixed $a \in Q$. Let $q(t)$ be a curve in Q and define the curve $\xi(t) \in \mathfrak{g}$ to be its horizontal generator, that is, $\dot{q} = \xi_Q(q)$ and $\xi \in \mathfrak{g}_q^\perp$. Choose $g_0 \in \Pi^{-1}(q(0))$ and define $g(t) \in G$ by $g(0) = g_0$ and $\dot{g} = \xi_G(g)$. Then $g(t)$ is the horizontal lift of $q(t)$ through g_0 . We apply $T\Pi$ to (3.17) and use (3.19) and (3.14) to find

$$\frac{D}{Dt}\dot{q} = \nabla_{\dot{q}}\dot{q} = T_g\Pi \left(\tilde{\nabla}_{\dot{g}}\dot{g} \right) = T_g\Pi \left((\dot{\xi} + \text{ad}_\xi^\dagger \xi)_G(g) \right) = (\dot{\xi} + \text{ad}_\xi^\dagger \xi)_Q(q). \quad (3.26)$$

This shows (3.23).

Suppose now $\xi(t)$ is a generator of $q(t)$, possibly non-horizontal. Denote the horizontal generator curve by $\eta := H_q(\xi)$. On an open set $U \subset Q$ introduce a coordinate map $\phi : U \rightarrow \mathbb{R}^n, q \mapsto (x^1, \dots, x^n)$. Taking two time derivatives of the coordinate curve $x(t) = \phi(q(t))$ one obtains

$$\ddot{x} = T_x \xi_{\mathbb{R}^n}(\dot{x}) + \dot{\xi}_{\mathbb{R}^n}(x) = T_x \eta_{\mathbb{R}^n}(\dot{x}) + \dot{\eta}_{\mathbb{R}^n}(x) \quad (3.27)$$

By consequence,

$$T_q \phi(D_t \dot{q}) = (\dot{\eta} + \text{ad}_\eta^\dagger \eta)_{\mathbb{R}^n}(x) = (\dot{\xi} + \text{ad}_{H_q(\xi)}^\dagger H_q(\xi))_{\mathbb{R}^n}(x) + T_x((V_q(\xi))_{\mathbb{R}^n})(\dot{x}). \quad (3.28)$$

This is the coordinate version of

$$\frac{D}{Dt} \dot{q} = \left(\dot{\xi} + \text{ad}_{H_q(\xi)}^\dagger H_q(\xi) \right)_Q(q) + \tau((V_q(\xi))_{TQ}(\dot{q})), \quad (3.29)$$

where τ was defined in Section 3.2.3. By (3.11), this is equivalent to (3.22).

The final step is to prove **(iv)**. For a fixed $q \in Q$, arbitrary $w_q, v_q \in T_q Q$ and $\xi \in \mathfrak{g}$ with $\xi_Q(q) = v_q$ we obtain the following chain of equalities; note that $(J(w_q^\flat))^\sharp \in \mathfrak{g}_q^\perp$ is horizontal because of (3.6).

$$\gamma_Q(((J(w_q^\flat))^\sharp)_Q(q), v_q) = \gamma_G((J(w_q^\flat))^\sharp, \xi) = \langle J(w_q^\flat), \xi \rangle_{\mathfrak{g}^* \times \mathfrak{g}} \quad (3.30)$$

$$= \langle w_q^\flat, \xi_Q(q) \rangle_{T^*Q \times TQ} = \gamma_Q(w_q, v_q). \quad (3.31)$$

Since v_q was arbitrary we conclude $w_q = ((J(w_q^\flat))^\sharp)_Q(q)$. This, together with (3.23), shows **(iv)**. $\square$

3.4 Cubics for normal metrics

In this section we derive the equations of Riemannian cubics for normal metrics. For ease of exposition we first consider a more general context and then particularise to the case of Riemannian cubics. The examples of Lie groups and of symmetric spaces are worked out in detail.

3.4.1 Preparations

Consider a manifold Q with a linear connection on its tangent bundle TQ . Denote the covariant derivative with respect to this linear connection by $\frac{D}{Dt}$. For $v_q, w_q \in T_q Q$ write

$(v_q)^H_{w_q} \in T_{w_q}TQ$ for the horizontal lift of v_q to w_q , i.e., in local coordinates,

$$(q, v)_{(q, w)}^H = (q, w, v, -\Gamma(q)(w, v)),$$

where Γ is the Christoffel map of the linear connection. The vertical lift of v_q to w_q is written $(v_q)^V_{w_q} := \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} (w_q + \varepsilon v_q)$.

For a variation $(t, s) \mapsto q(t, s)$ of a curve $q(t) = q(t, 0)$ the curve $\delta\dot{q}(t) := \frac{d}{ds}\Big|_{s=0} \dot{q}(t, s) \in T_{\dot{q}(t)}(TQ)$ splits into horizontal and vertical parts

$$\delta\dot{q} = (\delta q)_{(q, \dot{q})}^H + \left(\frac{D}{Ds} \Big|_{s=0} \dot{q} \right)_{(q, \dot{q})}^V. \quad (3.32)$$

For a function $\xi : TQ \rightarrow \mathfrak{g}$ and an arbitrary $v_q \in TQ$ define the $\mathfrak{g}$ -valued linear form $\frac{\delta\xi}{\delta q}\Big|_{v_q}$ by

$$\left\langle \frac{\delta\xi}{\delta q} \Big|_{v_q}, w_q \right\rangle := \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \xi(v(\varepsilon)), \quad \text{for any } w_q \in T_q Q, \quad (3.33)$$

where $v(\varepsilon)$ is any curve in TQ with $\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} v(\varepsilon) = (w_q)^H_{v_q}$. On the other hand we write $\frac{\delta\xi}{\delta\dot{q}}\Big|_{v_q}$ for the fibre derivative of ξ at v_q . That is,

$$\left\langle \frac{\delta\xi}{\delta\dot{q}} \Big|_{v_q}, w_q \right\rangle := \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \xi(v_q + \varepsilon w_q) \quad \text{for any } w_q \in T_q Q. \quad (3.34)$$

Note that if $q(t, s)$ is a variation of a curve $q(t) = q(t, 0)$, then using the splitting (3.32) we get

$$\delta\xi := \frac{d}{ds} \Big|_{s=0} \xi(\dot{q}(t, s)) = T_{\dot{q}(t)}\xi(\delta\dot{q}) = \left\langle \frac{\delta\xi}{\delta q}, \delta q \right\rangle + \left\langle \frac{\delta\xi}{\delta\dot{q}}, \frac{D}{Ds} \Big|_{s=0} \dot{q} \right\rangle. \quad (3.35)$$

We also define the dual operator $(\frac{\delta\xi}{\delta q}\Big|_{v_q})^* : \mathfrak{g}^* \rightarrow T_q^*Q$ by

$$\left\langle \left(\frac{\delta\xi}{\delta q} \Big|_{v_q} \right)^* \mu, w_q \right\rangle = \left\langle \mu, \left\langle \frac{\delta\xi}{\delta q} \Big|_{v_q}, w_q \right\rangle \right\rangle, \quad \text{for any } \mu \in \mathfrak{g}^*, w_q \in T_q Q, \quad (3.36)$$

and similarly the operator $(\frac{\delta\xi}{\delta\dot{q}}\Big|_{v_q})^* : \mathfrak{g}^* \rightarrow T_q^*Q$ by

$$\left\langle \left(\frac{\delta\xi}{\delta\dot{q}} \Big|_{v_q} \right)^* \mu, w_q \right\rangle = \left\langle \mu, \left\langle \frac{\delta\xi}{\delta\dot{q}} \Big|_{v_q}, w_q \right\rangle \right\rangle, \quad \text{for any } \mu \in \mathfrak{g}^*, w_q \in T_q Q. \quad (3.37)$$

3.4.2 A generalised variational problem

Given a Lagrangian $\ell : 2\mathfrak{g} \rightarrow \mathbb{R}$ and a smooth map $\xi : TQ \rightarrow \mathfrak{g}$, consider the action functional on the space of curves $q(t) : [0, 1] \rightarrow Q$ given by

$$\mathcal{J}[q] = \int_0^1 \ell(\xi(q, \dot{q}), \partial_t \xi(q, \dot{q})) dt, \quad (3.38)$$

and Hamilton's principle $\delta\mathcal{J} = 0$ with respect to variations satisfying $\delta q(0) = \delta q(1) = 0$ and $\delta \dot{q}(0) = \delta \dot{q}(1) = 0$. As we shall see below, the Lagrangian of Riemannian cubics for normal metrics fits into this framework.

In the following calculations we assume that there is a Riemannian metric on Q , and we work with the Levi-Civita connection. Taking variations of $\mathcal{J}$ we obtain, using (3.35),

$$\begin{aligned} \delta\mathcal{J} &= \int_0^1 \left\langle \frac{\delta\ell}{\delta\xi}, \delta\xi \right\rangle + \left\langle \frac{\delta\ell}{\delta\dot{\xi}}, \delta\dot{\xi} \right\rangle dt = \int_0^1 \left\langle \frac{\delta\ell}{\delta\xi} - \frac{d}{dt} \frac{\delta\ell}{\delta\dot{\xi}}, \delta\xi \right\rangle dt \\ &= \int_0^1 \left\langle \frac{\delta\ell}{\delta\xi} - \frac{d}{dt} \frac{\delta\ell}{\delta\dot{\xi}}, \left\langle \frac{\delta\xi}{\delta q}, \delta q \right\rangle + \left\langle \frac{\delta\xi}{\delta \dot{q}}, \frac{D}{Ds} \Big|_{s=0} \dot{q} \right\rangle \right\rangle dt \\ &= \int_0^1 \left\langle \left(\frac{\delta\xi}{\delta q} \right)^* \left(\frac{\delta\ell}{\delta\xi} - \frac{d}{dt} \frac{\delta\ell}{\delta\dot{\xi}} \right), \delta q \right\rangle + \left\langle \left(\frac{\delta\xi}{\delta \dot{q}} \right)^* \left(\frac{\delta\ell}{\delta\xi} - \frac{d}{dt} \frac{\delta\ell}{\delta\dot{\xi}} \right), \frac{D}{Dt} \delta q \right\rangle dt \\ &= \int_0^1 \left\langle \left[\left(\frac{\delta\xi}{\delta q} \right)^* - \frac{D}{Dt} \circ \left(\frac{\delta\xi}{\delta \dot{q}} \right)^* \right] \left(\frac{\delta\ell}{\delta\xi} - \frac{d}{dt} \frac{\delta\ell}{\delta\dot{\xi}} \right), \delta q \right\rangle dt, \end{aligned}$$

where $\frac{D}{Dt} \circ \left(\frac{\delta\xi}{\delta \dot{q}} \right)^*$ represents the operation of evaluating the function $\left(\frac{\delta\xi}{\delta \dot{q}} \right)^*$ before taking the covariant derivative; note that we used the symmetry of the connection to obtain the fourth equality. The Euler-Lagrange equation for Hamilton's principle $\delta\mathcal{J} = 0$ is therefore

$$\left[\left(\frac{\delta\xi}{\delta q} \right)^* - \frac{D}{Dt} \circ \left(\frac{\delta\xi}{\delta \dot{q}} \right)^* \right] \left(\frac{\delta\ell}{\delta\xi} - \frac{d}{dt} \frac{\delta\ell}{\delta\dot{\xi}} \right) = 0. \quad (3.39)$$

3.4.3 Cubics for normal metrics: Euler-Lagrange equations

Let (G, γ_G) be a Lie group with right-invariant metric γ_G , acting transitively from the left on a manifold (Q, γ_Q) with normal metric γ_Q . Let $q(t)$ be a curve in Q , originating at $q_0 = q(0)$. Recall from (3.24) that its horizontal generator curve is given by $\bar{J}(\dot{q}) = (J(\dot{q}^\flat))^\sharp \in \mathfrak{g}_q^\perp$.

Lemma 3.2. *For any curve $q(t) \in Q$, the curve $\partial_t(\bar{J}(\dot{q})) + \text{ad}_{\bar{J}(\dot{q})}^\dagger \bar{J}(\dot{q})$ is horizontal, that is, in $\mathfrak{g}_{q(t)}^\perp$.*

Proof. Fix any $g_0 \in \Pi^{-1}(q_0)$ and define the horizontal curve $g(t) \in G$ by $g(0) = g_0$ and $\dot{g} = (\bar{J}(\dot{q}))_G(g)$. By formula (3.17), $\frac{D}{Dt}\dot{g}$ is horizontal, that is, in $H_g G$. Moreover, by Proposition 3.1,

$$\frac{D}{Dt}\dot{g} = \left(\partial_t \bar{J}(\dot{q}) + \text{ad}_{\bar{J}(\dot{q})}^\dagger \bar{J}(\dot{q}) \right)_G(g) = TR_g \left(\partial_t \bar{J}(\dot{q}) + \text{ad}_{\bar{J}(\dot{q})}^\dagger \bar{J}(\dot{q}) \right).$$

The statement of the lemma now follows from (3.13). $\square$

This lemma enables us to rewrite the Lagrangian (2.8) of Riemannian cubics, evaluated along the curve $q(t)$, as follows,

$$L(q, \dot{q}, \ddot{q}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} \right\|_q^2 = \frac{1}{2} \left\| \left(\partial_t \bar{J}(\dot{q}) + \text{ad}_{\bar{J}(\dot{q})}^\dagger \bar{J}(\dot{q}) \right)_Q(q) \right\|_q^2 = \frac{1}{2} \left\| \partial_t \bar{J}(\dot{q}) + \text{ad}_{\bar{J}(\dot{q})}^\dagger \bar{J}(\dot{q}) \right\|_{\mathfrak{g}}^2,$$

where we used (3.25) in the second and the normal metric expression (3.9) in the third equality.

Therefore, Hamilton's principle (2.5) for Riemannian cubics is $\delta \mathcal{J} = 0$ with cost functional $\mathcal{J}$ of the form (3.38),

$$\mathcal{J}[q] = \int_0^1 \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} \right\|_q^2 dt = \int_0^1 \ell(\bar{J}(\dot{q}), \partial_t \bar{J}(\dot{q})) dt,$$

where

$$\ell : 2\mathfrak{g} \rightarrow \mathbb{R}, \quad (\xi_1, \xi_2) \mapsto \frac{1}{2} \left\| \xi_2 + \text{ad}_{\xi_1}^\dagger \xi_1 \right\|_{\mathfrak{g}}^2.$$

Remarkably, the function ℓ coincides with the reduced Lagrangian in the Euler–Poincaré reduction of Riemannian cubics on Lie groups, which we first encountered in Proposition 2.8. We recall from (2.38) the variational derivatives of ℓ ,

$$\frac{\delta \ell}{\delta \xi_1} = (\text{ad}_{\xi_1} \eta)^\flat - \text{ad}_\eta^* \xi_1^\flat, \quad \text{and} \quad \frac{\delta \ell}{\delta \xi_2} = \eta^\flat, \quad \text{where } \eta := \xi_2 + \text{ad}_{\xi_1}^\dagger \xi_1.$$

Hence, the Euler–Lagrange equation (3.39) becomes

$$\left[\left(\frac{\delta \bar{J}}{\delta q} \right)^* - \frac{D}{Dt} \circ \left(\frac{\delta \bar{J}}{\delta \dot{q}} \right)^* \right] (-\partial_t \eta^\flat + (\text{ad}_{\bar{J}} \eta)^\flat - \text{ad}_\eta^* \bar{J}^\flat) = 0, \quad \text{where } \eta := \dot{\bar{J}} + \text{ad}_{\bar{J}}^\dagger \bar{J}. \quad (3.40)$$

3.4.4 Cubics on Lie groups revisited

Riemannian cubics on Lie groups (G, γ_G) with right-invariant metrics γ_G were treated in the previous chapter by second-order Euler–Poincaré reduction. To illustrate the ideas above, we now revisit the problem from the point of view of normal metrics.

Equations of motion. We first observe that γ_G is a particularly simple case of a normal metric. Namely, let (G, γ_G) act on G by *left* multiplication. Then the normal metric induced on G is again γ_G . The generator of a curve $g(t)$ is the right-invariant velocity vector, $\bar{J}(g, \dot{g}) = TR_{g^{-1}}\dot{g}$. Relative to the Levi–Civita connection of γ_G one arrives, by Proposition 3.1 (ii), at

$$\begin{aligned} \left. \frac{\delta \bar{J}}{\delta g} \right|_{(g, \dot{g})} : \delta g &\mapsto -\frac{1}{2} \text{ad}_\eta^\dagger \xi - \frac{1}{2} \text{ad}_\xi^\dagger \eta - \frac{1}{2} [\xi, \eta], \quad \text{where } \xi := TR_{g^{-1}}\dot{g}, \eta := TR_{g^{-1}}\delta g, \\ \left(\left. \frac{\delta \bar{J}}{\delta g} \right|_{(g, \dot{g})} \right)^* : \mu &\mapsto (TR_{g^{-1}})^* \left(\frac{1}{2} \text{ad}_{\mu^\#}^* \xi^\flat - \frac{1}{2} (\text{ad}_\xi \mu^\#)^\flat - \frac{1}{2} \text{ad}_\xi^* \mu \right), \quad \text{where } \xi := TR_{g^{-1}}\dot{g}, \\ \left. \frac{\delta \bar{J}}{\delta \dot{g}} \right|_{(g, \dot{g})} : v_g &\mapsto TR_{g^{-1}}v_g \quad \text{for any } v_g \in T_g G, \quad \text{and} \\ \left(\left. \frac{\delta \bar{J}}{\delta \dot{g}} \right|_{(g, \dot{g})} \right)^* : \mu &\mapsto (TR_{g^{-1}})^* \mu. \end{aligned}$$

Using the expression for $\frac{D}{Dt}$ given in (3.21) we now get

$$\left[\left(\left. \frac{\delta \bar{J}}{\delta g} \right|_{(g, \dot{g})} \right)^* - \frac{D}{Dt} \circ \left(\left. \frac{\delta \bar{J}}{\delta \dot{g}} \right|_{(g, \dot{g})} \right)^* \right] \mu = (TR_{g^{-1}})^* (-\partial_t - \text{ad}_{\bar{J}}^*) \mu \quad (3.41)$$

for any curves $g(t) \in G$ and $\mu(t) \in \mathfrak{g}^*$. The Euler–Lagrange equations (3.40) are therefore

$$(TR_{g^{-1}})^* (\partial_t + \text{ad}_{\bar{J}}^*) [\partial_t \eta^\flat - (\text{ad}_{\bar{J}} \eta)^\flat + \text{ad}_\eta^* \bar{J}^\flat] = 0, \quad \text{where } \eta := \dot{\bar{J}} + \text{ad}_{\bar{J}}^\dagger \bar{J}. \quad (3.42)$$

This is equivalent to

$$(\partial_t + \text{ad}_{\bar{J}}^*) [\partial_t \eta^\flat - (\text{ad}_{\bar{J}} \eta)^\flat + \text{ad}_\eta^* \bar{J}^\flat] = 0, \quad \text{where } \eta := \dot{\bar{J}} + \text{ad}_{\bar{J}}^\dagger \bar{J}, \quad (3.43)$$

and $\bar{J} := \bar{J}(g, \dot{g}) = TR_{g^{-1}}\dot{g}$. As expected, we have recovered the second-order Euler–Poincaré equation (2.39).

Bi-invariance and the NHP equation. We recall from the previous chapter that if the metric γ_G is bi-invariant, then $\text{ad}^\dagger = -\text{ad}$, and the above expressions simplify. Namely, $\eta = \dot{\bar{J}}$ and (3.43) is equivalent to the NHP equation given in (2.43),

$$\ddot{\bar{J}} + [\ddot{\bar{J}}, \bar{J}] = 0. \quad (3.44)$$

Euler–Poincaré reduction. For the convenience of the reader we include here a brief reminder of second-order Euler–Poincaré reduction and, in particular, the Riemannian cubics in this context. These topics were covered in detail in the previous chapter, see Sections 2.3 and 2.4.

Start with a right-invariant Lagrangian $L : T^{(2)}G \rightarrow \mathbb{R}$ with reduced Lagrangian $\ell : 2\mathfrak{g} \rightarrow \mathbb{R}$. Consider Hamilton’s principle

$$\delta \int_0^1 L(g, \dot{g}, \ddot{g}) dt = 0 \quad (3.45)$$

with respect to variations of curves $g(t) : [0, 1] \rightarrow G$ respecting boundary conditions $\delta g(0) = \delta g(1) = 0$ and $\delta \dot{g}(0) = \delta \dot{g}(1) = 0$. The right-invariance of L leads to the equivalent reduced formulation

$$\delta \int_0^1 \ell(\xi, \dot{\xi}) dt = 0, \quad (3.46)$$

with respect to *constrained* variations of curves $\xi(t) : [0, 1] \rightarrow \mathfrak{g}$. Namely, one considers variations of the form $\delta \xi = \dot{\eta} - [\xi, \eta]$ with $\eta(t) : [0, 1] \rightarrow \mathfrak{g}$ arbitrary up to boundary conditions $\eta(0) = \eta(1) = 0$ and $\dot{\eta}(0) = \dot{\eta}(1) = 0$. Solutions $g(t)$ of (3.45) and solutions $\xi(t)$ of (3.46) are equivalent through the reconstruction relation $\xi = TR_{g^{-1}}\dot{g} = \bar{J}(g, \dot{g})$.

Taking constrained variations of (3.46) leads to the second-order Euler–Poincaré equation

$$\left(\frac{d}{dt} + \text{ad}_\xi^* \right) \left(\frac{\delta \ell}{\delta \xi} - \frac{d}{dt} \frac{\delta \ell}{\delta \dot{\xi}} \right) = 0. \quad (3.47)$$

For a curve $g(t) \in G$ with right-invariant velocity vector $\xi(t) = TR_{g^{-1}}\dot{g}$ we rewrite the Lagrangian of cubics,

$$L(g, \dot{g}, \ddot{g}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2 = \frac{1}{2} \left\| (\dot{\xi} + \text{ad}_\xi^\dagger \xi)_G(g) \right\|_g^2 = \frac{1}{2} \left\| \dot{\xi} + \text{ad}_\xi^\dagger \xi \right\|_\mathfrak{g}^2,$$

where in the second equality we used (3.19) and the third equality follows from right-invariance of γ_G . This demonstrates that the Lagrangian L for Riemannian cubics only depends on the right-invariant velocity ξ and its time-derivative $\dot{\xi}$ and is therefore right-invariant. The reduced Lagrangian can be read off as

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \left\| \dot{\xi} + \text{ad}_\xi^\dagger \xi \right\|_{\mathfrak{g}}^2. \quad (3.48)$$

The dynamics are governed by the second-order Euler–Poincaré equation (3.47), which becomes, for ℓ as above,

$$(\partial_t + \text{ad}_\xi^*)[\partial_t \eta^\flat - (\text{ad}_\xi \eta)^\flat + \text{ad}_\eta^* \xi^\flat] = 0, \quad \text{where} \quad \eta := \dot{\xi} + \text{ad}_\xi^\dagger \xi.$$

This coincides with (3.43), since $\xi = TR_{g^{-1}}\dot{g} = \bar{J}(g, \dot{g})$.

3.4.5 Cubics on symmetric spaces

We particularise the equation for Riemannian cubics (3.40) to symmetric spaces. Due to the appearance of the horizontal generator, this equation lends itself to the analysis of horizontal lifting properties to be addressed in Section 3.5 below. We also comment on how it is related to the equivalent equation derived in [19].

The horizontal generator. Recall that for any curve $g(t)$ in a Lie group G with $\dot{g} = TR_g \xi$ for $\xi(t) \in \mathfrak{g}$ and any curve $\nu(t) \in \mathfrak{g}$,

$$\frac{d}{dt} \text{Ad}_{g^{-1}} \nu = \text{Ad}_{g^{-1}} (\dot{\nu} + [\nu, \xi]). \quad (3.49)$$

Let G be a Lie group with a *bi-invariant* metric γ_G that acts transitively on a manifold Q equipped with the normal metric γ_Q , so that the action is by isometries.⁹ Denote by G_a the isotropy subgroup of a fixed element $a \in Q$, so that Q is diffeomorphic to G/G_a , the

⁹To see that G acts isometrically, let $v_q = \xi_Q(q)$ and $w_q = \nu_Q(q)$ be two arbitrary vectors in $T_q Q$. Then

$$\begin{aligned} \gamma_Q(gq)(gv_q, gw_q) &= \gamma_Q(gq)((\text{Ad}_g H_q(\xi))_Q(gq), (\text{Ad}_g H_q(\nu))_Q(gq)) \\ &= \gamma_G(e)(\text{Ad}_g H_q(\xi), \text{Ad}_g H_q(\nu)) = \gamma_G(e)(H_q(\xi), H_q(\nu)) = \gamma_Q(q)(v_q, w_q), \end{aligned}$$

where we used the relations (3.13) and (3.16) for the second equality and bi-invariance of γ_G for the third.

quotient being taken with respect to the right-action of G_a on G . Recall the Riemannian submersion $\Pi : G \rightarrow Q$ given by $g \mapsto ga$. Note that for any $g \in G$ with $ga = q$,

$$\text{Ad}_g \circ H_a \circ \text{Ad}_g^{-1} = H_q \quad \text{and} \quad \text{Ad}_g \circ V_a \circ \text{Ad}_g^{-1} = V_q. \quad (3.50)$$

Lemma 3.3. *Let $\bar{J} := \bar{J}(\dot{q})$ be the horizontal generator of a curve $q(t) \in Q$. Then,*

$$V_q(\bar{J}) = 0, \quad V_q(\dot{\bar{J}}) = 0, \quad V_q\left(\ddot{\bar{J}} + [\dot{\bar{J}}, \bar{J}]\right) = 0, \quad (3.51)$$

$$V_q\left(\ddot{\bar{J}} + 2[\ddot{\bar{J}}, \bar{J}] + [[\dot{\bar{J}}, \bar{J}], \bar{J}]\right) = 0. \quad (3.52)$$

Proof. Let $g(t) \in G$ be a horizontal lift of $q(t)$ relative to the Riemannian submersion Π . The first equation in (3.51) follows from horizontality of $\bar{J}$. By (3.50) it is equivalent to $V_a(\text{Ad}_{g^{-1}} \bar{J}) = 0$. The second equation follows from taking a time derivative of this latter relation while noting that $\dot{g} = TR_g \bar{J}$ and, by consequence, using (3.49) with ξ replaced by $\bar{J}$. That is,

$$\partial_t (V_a(\text{Ad}_{g^{-1}} \bar{J})) = V_a(\text{Ad}_{g^{-1}} \dot{\bar{J}}) = 0.$$

Therefore, $V_q(\dot{\bar{J}}) = 0$. The third and fourth equations follow from taking two more time derivatives. $\square$

Symmetric spaces. Assume in addition that there exists an involutive Lie algebra automorphism σ of $\mathfrak{g}$ such that $\mathfrak{g}_a$ and $\mathfrak{g}_a^\perp$ are, respectively, the $+1$ and -1 eigenspaces. Then (G, Q, σ) is a *symmetric space structure*.¹⁰ The following inclusions hold for all $q \in Q$,

$$[\mathfrak{g}_q, \mathfrak{g}_q] \subset \mathfrak{g}_q, \quad [\mathfrak{g}_q^\perp, \mathfrak{g}_q] \subset \mathfrak{g}_q^\perp, \quad [\mathfrak{g}_q^\perp, \mathfrak{g}_q^\perp] \subset \mathfrak{g}_q. \quad (3.53)$$

¹⁰We refer to [67] for a comprehensive treatment of symmetric spaces. One starts from a so-called symmetric space structure (G, Q, σ) , whereby initially no Riemannian metrics are specified. G is assumed to act transitively on Q and σ is an involutive Lie algebra automorphism of $\mathfrak{g}$ whose $+1$ eigenspace is the isotropy Lie algebra $\mathfrak{g}_a$ for some base point $a \in Q$. It turns out that any G -invariant Riemannian metric on the symmetric space Q induces the same Levi-Civita connection on TQ , the *canonical connection* ([67, Corollary 4.3] and [68, Theorem 3.1]). In particular, the geodesic curves in Q as well as the curvature are entirely defined by the data (G, Q, σ) . In this thesis we consider symmetric spaces with the following additional properties: (1) the canonical connection arises from a normal metric that is associated with a bi-invariant metric on the group of isometries; and (2) $\mathfrak{g}_a^\perp$ is the -1 eigenspace of σ . These properties are satisfied by the symmetric spaces of *Type I* in Cartan's classification of irreducible symmetric spaces, all of which are listed in [67, Chap. IX, Table II].

The first identity follows because $\mathfrak{g}_q$ is a Lie subalgebra of $\mathfrak{g}$. The second one is a consequence of the Ad-invariance of the metric inner product on $\mathfrak{g}$. The third one is characteristic of symmetric spaces. It follows from the eigenspace structure of σ described above. As a consequence of (3.53) we can see that $\left[\left[\dot{\bar{J}}, \bar{J}\right], \bar{J}\right]$ is horizontal, so that (3.52) becomes

$$V_q \left(\ddot{\bar{J}} + 2 \left[\ddot{\bar{J}}, \bar{J} \right] \right) = 0. \quad (3.54)$$

We now compute the Euler–Lagrange equation (3.40) for Riemannian cubics in the symmetric space Q . We will find that it is equivalent to

$$H_q \left(\ddot{\bar{J}} + 2 \left[\ddot{\bar{J}}, \bar{J} \right] \right) = 0. \quad (3.55)$$

In particular, a curve $q(t) \in Q$ is a Riemannian cubic if and only if its horizontal generator curve $\bar{J}(t)$ satisfies

$$\ddot{\bar{J}} + 2 \left[\ddot{\bar{J}}, \bar{J} \right] = 0. \quad (3.56)$$

We start with a lemma.

Lemma 3.4. *Relative to the Levi–Civita connection on TQ ,*

$$\left\langle \frac{\delta \bar{J}}{\delta q} \Big|_{(q, \dot{q})}, \delta q \right\rangle = [\bar{J}(\delta q), \bar{J}(\dot{q})], \quad \text{and} \quad \left\langle \frac{\delta \bar{J}}{\delta \dot{q}} \Big|_{(q, \dot{q})}, \delta q \right\rangle = \bar{J}(\delta q).$$

Let us reiterate what we need to verify. One, that there exists a bi-invariant metric on G whose normal metric induces the canonical connection on Q . And two, that $\mathfrak{g}_a^\perp$ is the -1 eigenspace of σ .

For the first point we recall from the general theory that the isometry groups G of Type I symmetric spaces are compact, simple Lie groups. Then the negative of the Killing form on $\mathfrak{g}$ is a positive definite, Ad-invariant inner product and can therefore be brought onto the whole of G by right-, or left-translation, thereby establishing a bi-invariant Riemannian metric on G . The corresponding normal metric on Q is G -invariant, hence induces the canonical connection.

Secondly, as we said above, the Lie algebra automorphism σ of a symmetric space is such that the isotropy Lie algebra $\mathfrak{g}_a$ at the base point $a \in Q$ is the $+1$ eigenspace. The -1 eigenspace is orthogonal to $\mathfrak{g}_a$, which follows from the invariance of the Killing form with respect to Lie algebra automorphisms, in particular σ . Hence, the -1 eigenspace is exactly $\mathfrak{g}_a^\perp$. Whenever we say ‘symmetric space’ in the following we mean a Type I symmetric space.

Proof. The second statement is due to the linearity of $\bar{J}$ on fibres of TQ . In order to prove the first equation, let $q(\varepsilon)$ be a curve with $q(0) = q$ and $\partial_{\varepsilon=0}q(\varepsilon) = \delta q$, and let $g(\varepsilon)$ with $g(0) = g$ be horizontal above $q(\varepsilon)$. We construct the parallel transport of $\dot{q}$ along $q(\varepsilon)$. Define

$$\omega(\varepsilon) := \bar{J}(\dot{q}) + \varepsilon [\bar{J}(\delta q), \bar{J}(\dot{q})], \quad X(\varepsilon) := TR_{g(\varepsilon)}\omega(\varepsilon). \quad (3.57)$$

To first order in ε , $X(\varepsilon)$ is a horizontal vector field along $g(\varepsilon)$ and $\omega(\varepsilon)$ lies in $\mathfrak{g}_{q(\varepsilon)}^\perp$. Now define a vector field along $q(\varepsilon)$ by $v(\varepsilon) := T_{g(\varepsilon)}\Pi(X(\varepsilon))$. Note that to first order in ε , $\bar{J}(v(\varepsilon)) = \omega(\varepsilon)$. Denoting by $\tilde{\nabla}$ the covariant derivative on G with respect to the Levi-Civita connection of γ , we use (3.20) to get

$$\tilde{\nabla}_{\delta q}X = \left(\frac{1}{2} [\bar{J}(\delta q), \bar{J}(\dot{q})] \right)_G (g) = \widetilde{\nabla_{\delta q}v} + \frac{1}{2} [\delta g, X]^V,$$

where we used (3.17) in the second step. Recall from (3.53) that $[\bar{J}(\delta q), \bar{J}(\dot{q})]$ is in $\mathfrak{g}_q$, so that applying $T\Pi$ to the above shows $\nabla_{\delta q}v = 0$. To first order in ε , $v(\varepsilon)$ is therefore the parallel transport of $\dot{q}$ along $q(\varepsilon)$, and

$$\left\langle \left. \frac{\delta \bar{J}}{\delta q} \right|_{(q, \dot{q})}, \delta q \right\rangle = \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \bar{J}(v(\varepsilon)) = \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \omega(\varepsilon) = [\bar{J}(\delta q), \bar{J}(\dot{q})].$$

□

It follows that

$$\left(\left. \frac{\delta \bar{J}}{\delta q} \right|_{(q, \dot{q})} \right)^* : \mu \mapsto \left(([\bar{J}(\dot{q}), \mu^\#])_Q(q) \right)^\flat, \quad \left(\left. \frac{\delta \bar{J}}{\delta \dot{q}} \right|_{(q, \dot{q})} \right)^* : \mu \mapsto \left((\mu^\#)_Q(q) \right)^\flat, \quad (3.58)$$

for any $\mu \in \mathfrak{g}^*$. Applying the map $\sharp$ to (3.40) therefore gives

$$\left([\bar{J}, \ddot{J}] \right)_Q(q) - \frac{D}{Dt} \left[\left(\ddot{J} \right)_Q(q) \right] = 0. \quad (3.59)$$

It follows from (3.51) and relations (3.53) that $\ddot{J} + [\dot{J}, \bar{J}]$ is the horizontal generator of $(\ddot{J})_Q(q)$ and therefore

$$\begin{aligned} \frac{D}{Dt} \left[\left(\ddot{J} \right)_Q(q) \right] &= \frac{D}{Dt} \left[\left(\ddot{J} + [\dot{J}, \bar{J}] \right)_Q(q) \right] \\ &= \left(\ddot{J} + \frac{3}{2} [\ddot{J}, \bar{J}] - \frac{1}{2} [\bar{J}, [\dot{J}, \bar{J}]] \right)_Q(q) = \left(\ddot{J} + [\ddot{J}, \bar{J}] \right)_Q(q), \end{aligned}$$

where for the second equality we used (3.18) together with (3.20) and for the third one we used (3.53), namely

$$([\bar{J}, [\dot{\bar{J}}, \bar{J}]])_Q(q) = -([\bar{J}, V_q(\ddot{\bar{J}})])_Q(q) = -([\bar{J}, \ddot{\bar{J}}])_Q(q). \quad (3.60)$$

Therefore, (3.59) is equivalent to what we announced in (3.55), namely

$$H_q \left(\ddot{\bar{J}} + 2 [\ddot{\bar{J}}, \bar{J}] \right) = 0. \quad (3.61)$$

We argued above that, as a consequence,

$$\ddot{\bar{J}} + 2 [\ddot{\bar{J}}, \bar{J}] = 0. \quad (3.62)$$

We will exploit the apparent similarity of this equation with the NHP equation (3.44) when we analyse horizontal lifts of cubics in the next section.

Remark 3.5. *The equations for Riemannian cubics in symmetric spaces were first given in [19]. We briefly remark on how those equations are related to (3.62). Let $g(t)$ be a horizontal lift of a Riemannian cubic $q(t)$, and define $V(t) \in \mathfrak{g}_a^\perp$ by*

$$V = \text{Ad}_{g^{-1}} \bar{J}. \quad (3.63)$$

One checks easily that (3.62) implies

$$\ddot{V} + [V, [\dot{V}, V]] = 0. \quad (3.64)$$

This coincides with equation (46) of [19].

Example: $\mathbf{G} = \mathbf{SO}(3)$, $\mathbf{Q} = \mathbf{S}^2$. Let $G = SO(3)$ and $Q = S^2 \subset \mathbb{R}^3$ and let $SO(3)$ act on S^2 through its action on vectors in $\mathbb{R}^3$.

Remark 3.6. [Conventions for $SO(3)$ and S^2]

We already outlined our notational conventions for $SO(3)$ in Remark 2.10. In addition, we represent tangent and cotangent spaces of S^2 as

$$T_{\mathbf{x}}S^2 = \{(\mathbf{x}, \mathbf{v}) \in S^2 \times \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{v} = 0\}, \quad T_{\mathbf{x}}^*S^2 = \{(\mathbf{x}, \mathbf{p}) \in S^2 \times \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{p} = 0\} \quad (3.65)$$

with duality pairing $\langle (\mathbf{x}, \mathbf{p}), (\mathbf{x}, \mathbf{v}) \rangle_{T^*S^2 \times TS^2} = \mathbf{p} \cdot \mathbf{v}$. Whenever admissible, we will drop the explicit mention of the base point $\mathbf{x}$ in what follows.

The infinitesimal generator is given by $(\mathbf{\Omega})_{S^2}(\mathbf{x}) = \mathbf{\Omega} \times \mathbf{x}$. We consider the bi-invariant extension $\gamma_{SO(3)}$ to $SO(3)$ of the identity moment of inertia inner product $\langle \mathbf{\Omega}, \mathbf{\Omega} \rangle_{\mathfrak{so}(3)} = \mathbf{\Omega} \cdot \mathbf{\Omega} = \|\mathbf{\Omega}\|^2$ on $\mathfrak{so}(3)$. The corresponding normal metric on S^2 is the round metric. The vertical and horizontal spaces are,

$$\mathfrak{so}(3)_{\mathbf{x}} = \{ \mathbf{\Omega} \in \mathfrak{so}(3) \mid \mathbf{\Omega} = \lambda \mathbf{x} \text{ for } \lambda \in \mathbb{R} \} \quad \text{and} \quad \mathfrak{so}(3)_{\mathbf{x}}^{\perp} = \{ \mathbf{\Omega} \in \mathfrak{so}(3) \mid \mathbf{\Omega} \cdot \mathbf{x} = 0 \}.$$

The map $\bar{\mathbf{J}}$ of (3.24) becomes $\bar{\mathbf{J}}(\mathbf{x}, \mathbf{v}) = \mathbf{x} \times \mathbf{v}$. Equation (3.61) is

$$\mathbf{x} \times (\ddot{\bar{\mathbf{J}}} + 2\dot{\bar{\mathbf{J}}} \times \bar{\mathbf{J}}) = 0, \quad \text{with} \quad \bar{\mathbf{J}} = \mathbf{x} \times \dot{\mathbf{x}}. \quad (3.66)$$

Equation (3.66) appears in [26], where it is derived from the general Euler–Lagrange equation for cubics (2.10). The similarity of (3.66) with the NHP equation (2.44) on $SO(3)$, $\ddot{\bar{\mathbf{J}}} + \dot{\bar{\mathbf{J}}} \times \bar{\mathbf{J}} = 0$, was already remarked upon there. We will take advantage of this similarity in Section 3.5 for the investigation of horizontal lifts of cubics.

3.4.6 Curvature from cubics

The general equation (2.10) for Riemannian cubics involves the curvature of the underlying manifold. On the other hand, in our derivation of (3.40) we made no explicit mention of curvature. It is interesting to note that, therefore, it is possible to back out an expression for curvature from the equation of cubics. This is similar in principle to the fact that one can infer the Christoffel symbols for a given metric if one knows the geodesic equation.

We illustrate the idea in the case of Lie groups and symmetric spaces, recovering well-known curvature formulas.

Lie groups. Consider a Lie group G with metric γ_G that we assume, for simplicity, to be bi-invariant. Analogous arguments apply in the case of one-sided invariance. Let $g(t)$ be a curve in G and $\bar{J} = TR_{g^{-1}}\dot{g}$ its right-invariant velocity. Recall that $\text{ad}^\dagger = -\text{ad}$ and from (3.19), (3.20) that

$$D_t \dot{g} = (\partial_t \bar{J})_G(g), \quad D_t^2 \dot{g} = \left(\ddot{\bar{J}} - \frac{1}{2}[\bar{J}, \dot{\bar{J}}] \right)_G(g) =: (\nu)_G(g), \quad D_t^3 \dot{g} = \left(\dot{\nu} - \frac{1}{2}[\bar{J}, \nu] \right)_G(g),$$

where we defined $\nu := \ddot{\bar{J}} - \frac{1}{2}[\bar{J}, \dot{\bar{J}}]$. The general Euler–Lagrange equation for cubics (2.10) becomes

$$\left(\dot{\nu} - \frac{1}{2}[\bar{J}, \nu] \right)_G(g) + R\left(\dot{\bar{J}}_G(g), \bar{J}_G(g)\right)(\bar{J}_G(g)) = 0. \quad (3.67)$$

On the other hand, if $g(t)$ is a cubic, then the NHP equation (3.44) is satisfied, and therefore $\dot{\nu} = \frac{1}{2} [\bar{J}, \ddot{J}]$. Plugging this into (3.67), yields

$$\frac{1}{4} \left([\bar{J}, [\bar{J}, \ddot{J}]] \right)_G (g) + R \left(\dot{J}_G(g), \bar{J}_G(g) \right) (\bar{J}_G(g)) = 0.$$

We conclude that for any $g \in G$ and $\xi, \eta \in \mathfrak{g}$,

$$R(\eta_G(g), \xi_G(g)) \xi_G(g) = -\frac{1}{4} ([\xi, [\xi, \eta]])_G(g). \quad (3.68)$$

Symmetric spaces. For symmetric spaces one derives in a similar fashion that for a cubic $q(t)$ with horizontal generator $\bar{J} = \bar{J}(\dot{q})$,

$$D_t^3 \dot{q} = \left(\ddot{\bar{J}} + [\ddot{\bar{J}}, \bar{J}] \right)_Q (q).$$

It follows from (2.10) and (3.61) that

$$R(D_t \dot{q}, \dot{q}) \dot{q} = \left([\ddot{\bar{J}}, \bar{J}] \right)_Q (q) = - \left([\bar{J}, [\bar{J}, \ddot{J}]] \right)_Q (q),$$

where in the last step we used the third equation of (3.51). We conclude that for any $q \in Q$ and $\eta, \xi \in \mathfrak{g}_q^\perp$,

$$R(\eta_Q(q), \xi_Q(q)) \xi_Q(q) = -([\xi, [\xi, \eta]])_Q(q).$$

3.5 Horizontal lifts

The horizontal lifting property of geodesics in the normal metrics context has been an important feature of the large deformation matching framework in computational anatomy. References [34, 35, 10], amongst others, explore this aspect in detail. This motivates us to ask which Riemannian cubics on the manifold Q lift to horizontal cubics on the Lie group G . In the case of symmetric spaces we arrive at a geometric characterisation of the cubics that can be lifted horizontally: It turns out that the presence of curvature is prohibitive to horizontal lifts of cubics, that is, precisely those cubics can be lifted that lie in flat totally geodesic submanifolds. We then consider the more general setting of Riemannian submersions and formulate necessary and sufficient conditions under which horizontal lifts are possible.

3.5.1 Symmetric spaces

Here we characterise the Riemannian cubics that can be lifted horizontally to the group G of isometries. Let (G, Q, σ) be a symmetric space structure, as defined in Section 3.4.5. In particular we recall the important relations (3.53),

$$[\mathfrak{g}_q, \mathfrak{g}_q] \subset \mathfrak{g}_q, \quad [\mathfrak{g}_q^\perp, \mathfrak{g}_q] \subset \mathfrak{g}_q^\perp, \quad [\mathfrak{g}_q^\perp, \mathfrak{g}_q^\perp] \subset \mathfrak{g}_q. \quad (3.69)$$

Theorem 3.7. *A curve $q(t) \in Q$ is a Riemannian cubic and can be lifted horizontally to a Riemannian cubic $g(t) \in G$ if and only if it satisfies $\dot{q}(t) = (\xi(t))_Q(q(t))$ for a curve $\xi(t) \in \mathfrak{g}$ of the form*

$$\xi(t) = \frac{ut^2}{2} + vt + w, \quad (3.70)$$

where u, v, w span an Abelian subalgebra¹¹ that lies in $\mathfrak{g}_{q(0)}^\perp$.

Proof. Suppose $q(t)$ is a cubic and can be lifted horizontally to a cubic $g(t)$. The horizontal generator curve $\bar{J}(\dot{q}(t))$ simultaneously satisfies equations (3.44) and (3.56). Therefore, $[\ddot{\bar{J}}, \bar{J}] = 0$. In particular

$$\gamma\left([\ddot{\bar{J}}, \bar{J}], \dot{\bar{J}}\right) = \gamma\left(\ddot{\bar{J}}, [\bar{J}, \dot{\bar{J}}]\right) = \left\|[\bar{J}, \dot{\bar{J}}]\right\|_{\mathfrak{g}}^2 = 0,$$

where we used the Ad-invariance of γ in the first and (3.51) together with (3.69) in the second step. We infer that $[\bar{J}, \dot{\bar{J}}] = 0$. Together with the NHP equation (2.46) this reveals that $\ddot{\bar{J}}$ is constant. Therefore,

$$\bar{J}(\dot{q}(t)) = \frac{ut^2}{2} + vt + w$$

with constants $u, v, w \in \mathfrak{g}$. These are mutually commuting because of $[\bar{J}, \dot{\bar{J}}] = 0$ and its time derivative $[\bar{J}, \ddot{\bar{J}}] = 0$. Moreover, we know from the first two equations of (3.51) that $\bar{J}$ and $\dot{\bar{J}}$ are horizontal. Due to $[\bar{J}, \dot{\bar{J}}] = 0$ and the third equation of (3.51), $\ddot{\bar{J}}$ is also horizontal. Hence, $u, v, w \in \mathfrak{g}_{q(0)}^\perp$. Setting $\xi(t) = \bar{J}(\dot{q}(t))$ therefore completes the first part of the proof.

For the reverse, let $q(t)$ be a curve that satisfies $\dot{q}(t) = (\xi(t))_Q(q(t))$ for $\xi(t) \in \mathfrak{g}$ of the form (3.70) with mutually commuting $u, v, w \in \mathfrak{g}_{q(0)}^\perp$. Fix an element $a \in Q$ and recall the

¹¹Notice that $\mathfrak{g}_{q(0)}^\perp$ is not a Lie algebra. By ‘subalgebra’ we mean a Lie algebra that is contained in $\mathfrak{g}_{q(0)}^\perp$.

Riemannian submersion $\Pi : G \rightarrow Q$ given by $g \mapsto ga$. Assume without loss of generality that $q(0) = a$. We first show that $\xi(t)$ is horizontal at all times. Define the curve $g(t) \in G$ by $g(0) = e$ and $\dot{g} = TR_g \xi$. This curve lies in the Abelian subgroup $\text{Exp}(\text{span}(u, v, w))$ of G with Lie algebra $\text{span}(u, v, w) \subset \mathfrak{g}_{q(0)}^\perp$. Hence, Ad_g is the identity map on $\text{span}(u, v, w)$, which means in particular that $\xi \in \text{span}(u, v, w) = \text{Ad}_g(\text{span}(u, v, w)) \subset \text{Ad}_g \mathfrak{g}_{q(0)}^\perp = \mathfrak{g}_{q(t)}^\perp$. Hence $\xi(t)$ is the horizontal generator of $q(t)$, and therefore $g(t)$ is a horizontal lift of $q(t)$, relative to the Riemannian submersion Π . Moreover, $g(t)$ is a cubic in G since $\xi(t)$ satisfies the NHP equation (3.44). It also clearly satisfies (3.62). Therefore, $q(t)$ is a cubic in Q . This concludes the proof. $\square$

This result can be combined with elements of the general theory of symmetric spaces to arrive at the following theorem. We need the notion of totally geodesic submanifolds. A *totally geodesic submanifold* S of a Riemannian manifold M is a connected submanifold of M with the property that if a geodesic in M is tangent to S at any point in time, then it is a geodesic in S .

Theorem 3.8. *Precisely those Riemannian cubics can be lifted horizontally that lie in a flat, totally geodesic submanifold of Q .*

Proof. We know from [67, Ch. IV, Theorem 7.2 and Ch. V, Proposition 6.1] that if S is a flat totally geodesic submanifold of Q that contains the base point $a \in Q$, then S is the orbit of an Abelian subgroup $G_S = \text{Exp}(\mathfrak{s}) \subset G$, where $\mathfrak{s}$ is an Abelian subalgebra of $\mathfrak{g}_a^\perp$. But G_S is a horizontal submanifold of G . In particular, if $q(t)$ is a Riemannian cubic that lies in S , then its horizontal generator $\bar{J}(t)$ is in $\mathfrak{s}$. It follows from (3.56) that $\ddot{\bar{J}} = 0$. Therefore, the horizontal generator is of the form (3.70), which shows that $q(t)$ can be lifted horizontally by Theorem 3.7. This argument can be generalised with minor modifications to the case when S does not contain the base point a .

For the reverse, we recall from the proof of Theorem 3.7 that if $q(t)$ can be horizontally lifted to a cubic then its horizontal generator is of the form (3.70) and the lift $g(t)$ is in the Abelian subgroup $\text{Exp}(\text{span}(u, v, w))$. Therefore, $q(t)$ lies in the set $\text{Exp}(\text{span}(u, v, w))q(0)$. By [67, Ch. IV, Theorem 7.2] this is a totally geodesic submanifold of Q . Moreover, it is flat by [67, Ch. V, Proposition 6.1], since $\text{span}(u, v, w)$ is an Abelian subalgebra of $\mathfrak{g}_{q(0)}^\perp$. $\square$

A *Cartan subalgebra (CSA)* based at $q \in Q$ is a maximal Abelian subalgebra contained in $\mathfrak{g}_q^\perp$. The *rank* of the symmetric space is the dimension of its CSAs. The greater the rank of a symmetric space, the larger the set of vectors u, v, w consistent with the requirements of Theorem 3.7.

Corollary 3.9. *In rank-one symmetric spaces the only Riemannian cubics that can be lifted horizontally to Riemannian cubics on the group of isometries are geodesics composed with a cubic polynomial in time.*

Proof. Let Q be a rank-one symmetric space. By definition, any CSA is one-dimensional. A curve $q(t)$ is therefore a cubic that can be lifted horizontally if and only if $\dot{q}(t) = (\xi(t))_Q(q(t))$ with $\xi(t) \in \mathfrak{g}$ of the form

$$\xi(t) = \left(\frac{at^2}{2} + bt + c \right) d, \quad (3.71)$$

where $d \in \mathfrak{g}_{q(0)}^\perp$ and $a, b, c \in \mathbb{R}$. Therefore,

$$q(t) = e^{\left(\frac{at^3}{6} + \frac{bt^2}{2} + ct \right) d} q(0),$$

which corresponds to the geodesic $y(t) = e^{td} q(0)$, composed with a cubic polynomial in time. $\square$

Remark 3.10. *Consider $G = SO(3)$ and the rank-one symmetric space $Q = S^2$. The special cubics appearing in Corollary 3.9 are considered in more detail in [27]. In particular, the case $b = c = 0$ corresponds to the so-called natural splines in computer aided design (CAD) applications.*

3.5.2 Riemannian submersions

In this section we generalise the question of horizontal lifts of cubics to the Riemannian submersion setting. We also show how the result implies Theorem 3.7 of the previous section.

Theorem 3.11. *Let $\Pi : \tilde{Q} \rightarrow Q$ be a Riemannian submersion, and let $q(t) \in Q$ be a Riemannian cubic. Moreover, let $\tilde{q}(t) \in \tilde{Q}$ be a horizontal lift of $q(t)$. The curve $\tilde{q}(t)$ is a*

Riemannian cubic if and only if

$$\left[\dot{\tilde{q}}, \frac{D}{Dt} \dot{\tilde{q}} \right]^V = 0, \quad \text{for all } t, \quad (3.72)$$

where the superscript V denotes the vertical part.

Remark 3.12. For any two horizontal vectors $\tilde{v}, \tilde{w} \in H_{\tilde{q}}\tilde{Q}$ (the vector subspace of $T_{\tilde{q}}\tilde{Q}$ consisting of horizontal vectors), the expression $[\tilde{v}, \tilde{w}]^V$ is defined as $[\tilde{v}, \tilde{w}]^V := [\tilde{V}, \tilde{W}]^V(\tilde{q})$, for horizontal extensions $\tilde{V}, \tilde{W}$ of $\tilde{v}, \tilde{w}$.

Remark 3.13. In (3.81) below we shall define the so-called A -tensor for Riemannian submersions. Then (3.72) can be expressed as

$$A_{\dot{\tilde{q}}} \left(\frac{D}{Dt} \dot{\tilde{q}} \right) = 0.$$

Since A is a tensor this expression is well-defined without the horizontal extensions of the previous remark.

Proof. We denote $v(t) := \dot{q}(t)$ and $\tilde{v}(t) := \dot{\tilde{q}}(t)$. The metrics on Q and $\tilde{Q}$ are denoted $\tilde{\gamma}$ and γ . The covariant derivatives with respect to the Levi-Civita connections are written ∇ and $\tilde{\nabla}$ respectively. Recall that by definition $\frac{D}{Dt}v = \nabla_v v$ and $\frac{D}{Dt}\tilde{v} = \tilde{\nabla}_{\tilde{v}}\tilde{v}$. Using the formula $\tilde{\nabla}_{\tilde{X}}\tilde{Y} = \widetilde{\nabla_X Y} + \frac{1}{2}[\tilde{X}, \tilde{Y}]^V$ for the covariant derivative induced by a Riemannian submersion of horizontally lifted vector fields $\tilde{X}$ and $\tilde{Y}$, one obtains

$$\tilde{\nabla}_{\tilde{v}}\tilde{v} = \widetilde{\nabla_v v}, \quad (\tilde{\nabla}_{\tilde{v}})^2\tilde{v} = \widetilde{(\nabla_v)^2 v} + \frac{1}{2}[\tilde{v}, \widetilde{\nabla_v v}]^V, \quad (3.73)$$

$$(\tilde{\nabla}_{\tilde{v}})^3\tilde{v} = \widetilde{(\nabla_v)^3 v} + \frac{1}{2}[\tilde{v}, \widetilde{(\nabla_v^2 v)}]^V + \frac{1}{2}\tilde{\nabla}_{\tilde{v}}([\tilde{v}, \widetilde{\nabla_v v}]^V). \quad (3.74)$$

Suppose $q(t)$ and $\tilde{q}(t)$ are as in the statement of the theorem and let $\tilde{q}(t)$ be a Riemannian cubic, i.e., the respective equations of Riemannian cubics are satisfied,

$$\begin{cases} (\nabla_v)^3 v + R(\nabla_v v, v)v = 0, \\ (\tilde{\nabla}_{\tilde{v}})^3\tilde{v} + \tilde{R}(\tilde{\nabla}_{\tilde{v}}\tilde{v}, \tilde{v})\tilde{v} = 0. \end{cases} \quad (3.75)$$

For the following manipulations we record that

$$\tilde{\gamma} \left(\tilde{\nabla}_{\tilde{v}}([\tilde{v}, \widetilde{\nabla_v v}]^V), \tilde{\nabla}_{\tilde{v}}\tilde{v} \right) = \frac{d}{dt} \tilde{\gamma} \left([\tilde{v}, \widetilde{\nabla_v v}]^V, \tilde{\nabla}_{\tilde{v}}\tilde{v} \right) - \tilde{\gamma} \left([\tilde{v}, \widetilde{\nabla_v v}]^V, \tilde{\nabla}_{\tilde{v}}^2\tilde{v} \right) = -\tilde{\gamma} \left([\tilde{v}, \widetilde{\nabla_v v}]^V, \tilde{\nabla}_{\tilde{v}}^2\tilde{v} \right),$$

where the second step follows since $\tilde{\nabla}_{\tilde{v}}\tilde{v}$ is horizontal. We use this equality as well as equations (3.73)–(3.74) to obtain

$$\begin{aligned}\tilde{\gamma}\left((\tilde{\nabla}_{\tilde{v}})^3\tilde{v}, \tilde{\nabla}_{\tilde{v}}\tilde{v}\right) &= \tilde{\gamma}\left(\widetilde{(\nabla_v)^3v}, \widetilde{\nabla_v v}\right) + \frac{1}{2}\tilde{\gamma}\left(\tilde{\nabla}_{\tilde{v}}([\tilde{v}, \widetilde{\nabla_v v}]^V), \widetilde{\nabla_v v}\right) \\ &= \tilde{\gamma}\left(\widetilde{(\nabla_v)^3v}, \widetilde{\nabla_v v}\right) - \frac{1}{2}\tilde{\gamma}\left([\tilde{v}, \widetilde{\nabla_v v}]^V, (\tilde{\nabla}_{\tilde{v}})^2\tilde{v}\right) \\ &= \tilde{\gamma}\left(\widetilde{(\nabla_v)^3v}, \widetilde{\nabla_v v}\right) - \frac{1}{4}\left\|[\tilde{v}, \widetilde{\nabla_v v}]^V\right\|^2.\end{aligned}$$

Hence,

$$\tilde{\gamma}\left((\tilde{\nabla}_{\tilde{v}})^3\tilde{v}, \tilde{\nabla}_{\tilde{v}}\tilde{v}\right) = \gamma\left((\nabla_v)^3v, \nabla_v v\right) - \frac{1}{4}\left\|[\tilde{v}, \widetilde{\nabla_v v}]^V\right\|^2. \quad (3.76)$$

On the other hand O'Neill's formula for sectional curvatures of Riemannian submersions [66] (Equation 3. in Corollary 1), implies that

$$\tilde{\gamma}\left(\tilde{R}(\tilde{\nabla}_{\tilde{v}}\tilde{v}, \tilde{v})\tilde{v}, \tilde{\nabla}_{\tilde{v}}\tilde{v}\right) = \gamma\left(R(\nabla_v v, v)v, \nabla_v v\right) - \frac{3}{4}\left\|[\tilde{v}, \widetilde{\nabla_v v}]^V\right\|^2. \quad (3.77)$$

Adding (3.76) and (3.77) and using the equations of cubics (3.75) we conclude that

$$[\tilde{v}, \widetilde{\nabla_v v}]^V = [\tilde{v}, \tilde{\nabla}_{\tilde{v}}\tilde{v}]^V = 0, \quad (3.78)$$

which is (3.72).

To show the reverse direction, we note that when (3.72) holds, then (3.73) and (3.74) take the simplified form

$$(\tilde{\nabla}_{\tilde{v}})^2\tilde{v} = \widetilde{(\nabla_v)^2v}, \quad (\tilde{\nabla}_{\tilde{v}})^3\tilde{v} = \widetilde{(\nabla_v)^3v} + \frac{1}{2}[\tilde{v}, (\tilde{\nabla}_{\tilde{v}})^2\tilde{v}]^V. \quad (3.79)$$

Hence, $(\tilde{\nabla}_{\tilde{v}})^2\tilde{v}$ is horizontal. Moreover, $(\tilde{\nabla}_{\tilde{v}})^3\tilde{v}$ splits naturally into horizontal and vertical parts. Therefore, checking that the second equation in (3.75) holds, amounts to verifying that for any choice of horizontal vector field $\tilde{h}(t)$ and any choice of vertical vector field $\tilde{w}(t)$ along $\tilde{q}(t)$,

$$\begin{cases} \tilde{\gamma}\left(\tilde{R}(\tilde{\nabla}_{\tilde{v}}\tilde{v}, \tilde{v})\tilde{v}, \tilde{h}\right) - \gamma\left(R(\nabla_v v, v)v, h\right) = 0 \\ \tilde{\gamma}\left(\tilde{R}(\tilde{\nabla}_{\tilde{v}}\tilde{v}, \tilde{v})\tilde{v}, \tilde{w}\right) + \frac{1}{2}\tilde{\gamma}\left([\tilde{v}, (\tilde{\nabla}_{\tilde{v}})^2\tilde{v}]^V, \tilde{w}\right) = 0, \end{cases} \quad (3.80)$$

where we denoted $h := \Pi_*\tilde{h}$. To proceed, we introduce the $(1, 2)$ -tensors A and T defined, for arbitrary vector fields E, F , by

$$A_E F = \left(\tilde{\nabla}_{E^H}(F^H)\right)^V + \left(\tilde{\nabla}_{E^H}(F^V)\right)^H \quad (3.81)$$

$$T_E F = \left(\tilde{\nabla}_{E^V}(F^V) \right)^H + \left(\tilde{\nabla}_{E^V}(F^H) \right)^V. \quad (3.82)$$

The superscripts H and V denote the horizontal and vertical parts, respectively. Definitions (3.81) and (3.82) coincide with the ones given in [66]. It is shown there (in Equations {3} and {4}) that if X, Y, Z, H are horizontal vector fields and W is a vertical vector field, then

$$\begin{aligned} \tilde{\gamma} \left(\tilde{R}(X, Y)Z, H \right) &= \gamma \left(R(\Pi_* X, \Pi_* Y)\Pi_* Z, \Pi_* H \right) + 2\tilde{\gamma} (A_X Y, A_Z H) \\ &\quad - \tilde{\gamma} (A_Y Z, A_X H) - \tilde{\gamma} (A_Z X, A_Y H). \end{aligned} \quad (3.83)$$

and

$$\begin{aligned} \tilde{\gamma} \left(\tilde{R}(X, Y)Z, W \right) &= -\tilde{\gamma} \left((\tilde{\nabla}_Z A)_X Y, W \right) - \tilde{\gamma} (A_X Y, T_W Z) \\ &\quad + \tilde{\gamma} (A_Y Z, T_W X) + \tilde{\gamma} (A_Z X, T_W Y). \end{aligned} \quad (3.84)$$

Note that we differ from [66] in our sign convention for the curvature tensor. It is also shown in [66] that for any two horizontal vector fields X and Y one has $A_X Y = \frac{1}{2}[X, Y]^V$. In particular, (3.72) can be written as $A_{\tilde{v}}(\tilde{\nabla}_{\tilde{v}}\tilde{v}) = 0$. This, together with (3.83), implies that

$$\tilde{\gamma} \left(\tilde{R}(\tilde{\nabla}_{\tilde{v}}\tilde{v}, \tilde{v})\tilde{v}, \tilde{h} \right) = \gamma \left(R(\nabla_v v, v)v, h \right),$$

which is equivalent to the first equation in (3.80). In order to show the second equation we take a covariant derivative of (3.72) written in the form $A_{\tilde{v}}(\tilde{\nabla}_{\tilde{v}}\tilde{v}) = 0$ to obtain

$$0 = \tilde{\nabla}_{\tilde{v}}(A_{\tilde{v}}(\tilde{\nabla}_{\tilde{v}}\tilde{v})) = (\tilde{\nabla}_{\tilde{v}}A)_{\tilde{v}}(\tilde{\nabla}_{\tilde{v}}\tilde{v}) + A_{\tilde{v}}((\tilde{\nabla}_{\tilde{v}})^2\tilde{v}).$$

It follows from this and (3.84) that

$$\begin{aligned} \tilde{\gamma} \left(\tilde{R}(\tilde{\nabla}_{\tilde{v}}\tilde{v}, \tilde{v})\tilde{v}, \tilde{w} \right) &= -\tilde{\gamma} \left(\tilde{R}(\tilde{v}, \tilde{\nabla}_{\tilde{v}}\tilde{v})\tilde{v}, \tilde{w} \right) = \tilde{\gamma} \left((\tilde{\nabla}_{\tilde{v}}A)_{\tilde{v}}(\tilde{\nabla}_{\tilde{v}}\tilde{v}), \tilde{w} \right) \\ &= -\tilde{\gamma} \left(A_{\tilde{v}}((\tilde{\nabla}_{\tilde{v}})^2\tilde{v}), \tilde{w} \right) = -\frac{1}{2}\tilde{\gamma} \left([\tilde{v}, (\tilde{\nabla}_{\tilde{v}})^2\tilde{v}]^V, \tilde{w} \right). \end{aligned}$$

Therefore the second equation of (3.80) is satisfied. This concludes the proof. $\square$

Example: Normal metrics in the bi-invariant case. We now show how this result relates to Theorem 3.7 of Section 3.5.1. Let G be a Lie group with a bi-invariant metric

γ that acts transitively on a manifold Q equipped with the normal metric γ_Q .¹² Recall that for a fixed element $a \in Q$ the map $\Pi : G \rightarrow Q$, $g \mapsto ga$ is a Riemannian submersion.

Lemma 3.14. *Let $g \in G$ with $\Pi(g) = q$ and let ξ and η be in $\mathfrak{g}_q^\perp$. Then*

$$[\xi_G(g), \eta_G(g)]^V = (V_q([\xi, \eta]))_G(g). \quad (3.85)$$

Proof. In order to compute the left hand side we extend $\xi_G(g)$ to a horizontal vector field H^ξ on G . Namely, $H^\xi(h) = l_{\text{Ad}_{g^{-1}}(\xi)}(h)$, where l_η denotes the left-invariant vector field $l_\nu(h) = T_e L_h \nu$ for any $\nu \in \mathfrak{g}$. Similarly $H^\eta(h) = l_{\text{Ad}_{g^{-1}}(\eta)}(h)$. Now

$$[\xi_G(g), \eta_G(g)]^V = [H^\xi, H^\eta]^V(g) = \left(l_{\text{Ad}_{g^{-1}}[\xi, \eta]}(g) \right)^V = ([\xi, \eta]_G(g))^V = (V_q([\xi, \eta]))_G(g).$$

□

Theorem 3.15. *Let $q(t) \in Q$ be a Riemannian cubic with horizontal generator $\bar{J}(\dot{q}(t)) \in \mathfrak{g}_{q(t)}^\perp$. A horizontal lift $g(t) \in G$ of $q(t)$ is a Riemannian cubic if and only if*

$$V_q([\bar{J}, \dot{\bar{J}}]) = 0, \quad \text{for all } t. \quad (3.86)$$

Proof. Since $\dot{g} = (\bar{J})_G(g)$ and $D_t \dot{g} = (\dot{\bar{J}})_G(g)$ with both $\bar{J}$ and $\dot{\bar{J}}$ in $\mathfrak{g}_q^\perp$, Lemma 3.14 gives

$$\left[\dot{g}, \frac{D}{Dt} \dot{g} \right]^V = \left(V_q([\bar{J}, \dot{\bar{J}}]) \right)_G(g). \quad (3.87)$$

This expression vanishes according to Proposition 3.72, and therefore $V_q([\bar{J}, \dot{\bar{J}}]) = 0$. □

Remark 3.16. *If Q is a symmetric space, then relations (3.69) hold. Therefore $V_q([\bar{J}, \dot{\bar{J}}]) = 0$ is equivalent to $[\bar{J}, \dot{\bar{J}}] = 0$. This property was the key stepping stone in the proof of Theorem 3.7.*

3.6 Extended analysis: Reduction by isotropy subgroup

In the previous section we analysed the relationship between Riemannian cubics on Q and *horizontal* curves on G . In the present section we include in our considerations the *non-horizontal* curves on G .

¹²In this case the action of G is by isometries, which means that Q is a *homogeneous space* (see also the footnote on page 72).

We first show that certain non-horizontal geodesics on G project to cubics. We then extend the analysis in the following way. We reduce the Riemannian cubic variational problem on G by the isotropy subgroup G_a of a point $a \in Q$. The reduced Lagrangian couples horizontal and vertical parts of the motion, which accounts for the absence of a general horizontal lifting property. The reduced form of the equations reveals the obstruction for a cubic on G to project to a cubic on Q . The main technical tool in this section is second-order Lagrange–Poincaré reduction. The main references are [7] for the first-order theory and [50] for the recent generalisation to higher order.

3.6.1 Setting

Let G be a Lie group with metric γ_G , acting transitively from the left on a manifold Q with the normal metric γ_Q . For simplicity we assume that γ_G is bi-invariant.¹³ Recall the Riemannian submersion $\Pi : G \rightarrow Q$ given by $g \mapsto ga$ for a fixed element $a \in Q$. Let $G_a \subset G$ be the stabiliser of a . Consider the right action

$$\psi : G \times G_a \rightarrow G, \quad (g, h) \mapsto gh.$$

This action is free and the projection from G onto the quotient manifold G/G_a is a submersion. Moreover the map $G/G_a \rightarrow Q$ given by $[g] \mapsto ga$ is a diffeomorphism. The ingredients $(G, Q \cong G/G_a, G_a, \Pi, \psi)$ therefore constitute a principal fibre bundle with total space G , base manifold Q , structure group G_a , projection Π , and action ψ . The Lie algebra of the structure group G_a is $\mathfrak{g}_a$. Recall the vertical and horizontal projections, $V_q : \mathfrak{g} \rightarrow \mathfrak{g}_q$ and $H_q : \mathfrak{g} \rightarrow \mathfrak{g}_q^\perp$, for any $q \in Q$. We also recall that for any $g \in G$ with $ga = q$,

$$\text{Ad}_g \circ H_a \circ \text{Ad}_g^{-1} = H_q \quad \text{and} \quad \text{Ad}_g \circ V_a \circ \text{Ad}_g^{-1} = V_q. \quad (3.88)$$

We equip G with the principal bundle connection $\mathcal{A}$,

$$\mathcal{A} : TG \rightarrow \mathfrak{g}_a, \quad v_g \mapsto \mathcal{A}_g(v_g) := V_a(TL_{g^{-1}}v_g). \quad (3.89)$$

Recall from Section 3.2.4 that γ_G , together with the projection Π , induces a splitting of TG into horizontal and vertical subbundles $TG = HG \oplus VG$. This splitting coincides

¹³This means that Q is a homogeneous space. See the footnote on page 85 for details.

with the one prescribed by the connection $\mathcal{A}$, that is, $H_g G = \ker \mathcal{A}_g$ and $V_g G = \ker T_g \Pi$. The *curvature* of $\mathcal{A}$ is the $\mathfrak{g}_a$ -valued two-form

$$\begin{aligned} \mathcal{B}(u_g, v_g) &= [\mathcal{A}(u_g), \mathcal{A}(v_g)] + d\mathcal{A}(u_g, v_g) \\ &= [V_a(TL_{g^{-1}}u_g), V_a(TL_{g^{-1}}v_g)] - V_a([TL_{g^{-1}}u_g, TL_{g^{-1}}v_g]) \\ &= -V_a([H_a(TL_{g^{-1}}u_g), H_a(TL_{g^{-1}}v_g)]) \end{aligned} \quad (3.90)$$

for $u_g, v_g \in T_g G$.

Consider the following right-action of G_a on $G \times \mathfrak{g}_a$,

$$(G \times \mathfrak{g}_a) \times G_a \rightarrow G \times \mathfrak{g}_a, \quad (g, \xi, h) \mapsto (gh, \text{Ad}_h^{-1} \xi). \quad (3.91)$$

We define the *associated adjoint vector bundle* over Q , $\tilde{\mathfrak{g}}_a := (G \times \mathfrak{g}_a)/G_a$. The equivalence class, i.e., orbit, of $(g, \xi) \in G \times \mathfrak{g}_a$ will be denoted by square brackets, $\sigma = [g, \xi] \in \tilde{\mathfrak{g}}_a$. The vector bundle projection $\pi : \tilde{\mathfrak{g}}_a \rightarrow Q$ is given by $\sigma = [g, \xi] \mapsto ga$. In particular, the fibres $\pi^{-1}(q)$ are isomorphic to $\mathfrak{g}_a$. Indeed, upon fixing $g \in \Pi^{-1}(q)$ one can construct the isomorphism $\xi \mapsto [g, \xi]$ from $\mathfrak{g}_a$ to the fibre $\pi^{-1}(q)$. Moreover, insisting that this map be a Lie algebra homomorphism turns $\tilde{\mathfrak{g}}_a$ into a Lie algebra bundle. That is, each fibre carries a natural Lie algebra structure given by

$$[\sigma, \tilde{\sigma}] = [[g, \xi], [g, \tilde{\xi}]] := [g, [\xi, \tilde{\xi}]]. \quad (3.92)$$

The principal connection $\mathcal{A}$ induces a linear connection on $\tilde{\mathfrak{g}}_a$ with covariant derivative

$$\frac{D^{\mathcal{A}}}{Dt}[g(t), \xi(t)] = \left[g(t), \dot{\xi}(t) + [\mathcal{A}(\dot{g}(t)), \xi(t)] \right]. \quad (3.93)$$

We will sometimes use the shorthand $\dot{\sigma} := \frac{D^{\mathcal{A}}}{Dt}\sigma$ where $\sigma(t)$ is a curve in $\tilde{\mathfrak{g}}_a$. Moreover, we define the map

$$i : \tilde{\mathfrak{g}}_a \rightarrow \mathfrak{g}, \quad [g, \eta] \mapsto i([g, \eta]) := \text{Ad}_g \eta, \quad (3.94)$$

and write $[g, \eta] =: \sigma \mapsto \bar{\sigma} := i([g, \eta])$, as shorthand. Note that $i([g, \eta]) \in \mathfrak{g}_{ga}$. We introduce the fibre-wise inner product $\bar{\gamma}$ on $\tilde{\mathfrak{g}}_a$ given by

$$\bar{\gamma}(\sigma, \rho) := \gamma_G(\bar{\sigma}, \bar{\rho})$$

and its corresponding norm is denoted by $\|\cdot\|_{\tilde{\mathfrak{g}}_a}$. We also define the $\tilde{\mathfrak{g}}_a$ -valued *reduced curvature 2-form* $\tilde{\mathcal{B}}$,

$$\tilde{\mathcal{B}}(u_q, v_q) := [g, \mathcal{B}(u_g, v_g)], \quad (3.95)$$

for $u_q, v_q \in T_q Q$, where $g \in G$ and $u_g, v_g \in T_g G$ are such that $\Pi(g) = q$ and $T_g \Pi(u_g) = u_q$, $T_g \Pi(v_g) = v_q$.

Before we continue, let us make the following remark: If $(g, \dot{g})$ is any vector in $T_g G$, then $i([g, \mathcal{A}(\dot{g})]) = V_{ga}(TR_{g^{-1}}\dot{g})$. Moreover, $H_{ga}(TR_{g^{-1}}\dot{g})$ is the unique horizontal generator of $T_g \Pi(g, \dot{g}) \in T_{ga} Q$. What this means is that $i([g, \mathcal{A}(\dot{g})])$ and $T_g \Pi(g, \dot{g})$ uniquely determine the vertical and horizontal parts of the right-trivialised velocity vector $TR_{g^{-1}}\dot{g}$. This idea lies at the heart of the Lagrange–Poincaré reduction map, which we will introduce in the next section.

3.6.2 First-order Lagrange–Poincaré reduction

We start by recalling first-order Lagrange–Poincaré reduction, which makes use of the bundle diffeomorphism

$$\alpha_{\mathcal{A}}^{(1)} : TG/G_a \rightarrow TQ \times_Q \tilde{\mathfrak{g}}_a, \quad [g, \dot{g}] \mapsto (q, \dot{q}) \times [g, \mathcal{A}(\dot{g})]. \quad (3.96)$$

Here we introduced the quotient TG/G_a of TG by the natural action of G_a , whose elements we denote by $[g, \dot{g}] \in TG/G_a$. We also defined $(q, \dot{q}) := T_g \Pi(g, \dot{g})$ for any representative $(g, \dot{g})$ of $[g, \dot{g}]$.

Let $L : TG \rightarrow \mathbb{R}$ be a G_a -invariant Lagrangian. By a slight abuse of notation we also denote by L the corresponding function $TG/G_a \rightarrow \mathbb{R}$. The reduced Lagrangian ℓ is defined by $L = \ell \circ \alpha_{\mathcal{A}}^{(1)}$. In order to compute the Lagrange–Poincaré equations one takes constrained variations in the reduced variable space. Take for instance the kinetic energy Lagrangian $L(g, \dot{g}) = \frac{1}{2} \|\dot{g}\|_g^2$. Hamilton’s principle, $\delta S = 0$, for $S = \int_0^1 L dt$ leads to the Euler–Lagrange equation $D_t \dot{g} = 0$. This is the geodesic equation on G . Let us derive the corresponding Lagrange–Poincaré equations. We have the decomposition $\dot{g} = (\dot{g} - TL_g \mathcal{A}(\dot{g})) + TL_g \mathcal{A}(\dot{g})$ into horizontal and vertical parts. Therefore,

$$L(g, \dot{g}) = \frac{1}{2} \|\dot{g}\|_g^2 = \frac{1}{2} \|T_g \Pi(\dot{g})\|_{\Pi(g)}^2 + \frac{1}{2} \|\mathcal{A}(\dot{g})\|_{\mathfrak{g}}^2 = \frac{1}{2} \|T_g \Pi(\dot{g})\|_{\Pi(g)}^2 + \frac{1}{2} \|[g, \mathcal{A}(\dot{g})]\|_{\tilde{\mathfrak{g}}_a}^2.$$

Hence, the reduced Lagrangian $\ell : TQ \times_Q \tilde{\mathfrak{g}}_a$ is given by

$$\ell(q, \dot{q}, \sigma) = \frac{1}{2} \|\dot{q}\|_q^2 + \frac{1}{2} \|\sigma\|_{\tilde{\mathfrak{g}}_a}^2 = \frac{1}{2} \|\dot{q}\|_q^2 + \frac{1}{2} \|\bar{\sigma}\|_{\mathfrak{g}}^2, \quad (3.97)$$

where we recall that $\bar{\sigma} := i(\sigma)$. We will need the following result concerning variations $\delta \bar{\sigma} = \delta(i(\sigma))$.

Lemma 3.17. *Consider the map $i : \tilde{\mathfrak{g}}_a \rightarrow \mathfrak{g}$ defined in (3.94) and a curve $\varepsilon \mapsto \sigma_\varepsilon = [g_\varepsilon, \xi_\varepsilon] \in \tilde{\mathfrak{g}}_a$, covering the curve $\varepsilon \mapsto q_\varepsilon \in Q$. Then, we have the formula*

$$\frac{d}{d\varepsilon} i(\sigma_\varepsilon) = i \left(\frac{D^{\mathcal{A}}}{D\varepsilon} \sigma_\varepsilon \right) + F \left(\frac{d}{d\varepsilon} q_\varepsilon, \sigma \right),$$

where the covariant derivative $D^{\mathcal{A}}$ was defined in (3.93) and the map $F : TQ \times_Q \tilde{\mathfrak{g}}_a \rightarrow \mathfrak{g}$ is defined by

$$F(v_q, \sigma_q) := H_q([TR_{g^{-1}}v_g, \text{Ad}_g \eta]) = [H_q(TR_{g^{-1}}v_g), \text{Ad}_g \eta], \quad (3.98)$$

where $v_g \in TG$ and $\eta \in \mathfrak{g}_a$ are such that $T_g \Pi(v_g) = v_q$ and $[g, \eta] = \sigma_q$.

Proof. We have (see Section 3.4.1)

$$\frac{d}{d\varepsilon} i(\sigma_\varepsilon) = T_\sigma i(\delta\sigma) = T_\sigma i \left(\left(\frac{D^{\mathcal{A}}}{D\varepsilon} \sigma \right)_\sigma^V + (\delta q)_\sigma^H \right),$$

where $\tilde{\sigma}_\sigma^V \in T_\sigma \tilde{\mathfrak{g}}_a$ denotes the vertical lift of an element $\tilde{\sigma} \in \pi^{-1}(q)$ to σ , i.e.

$$\tilde{\sigma}_\sigma^V := \left. \frac{d}{ds} \right|_{s=0} (\sigma + s\tilde{\sigma}).$$

Linearity of i on fibres of the adjoint bundle implies that

$$T_\sigma i \left(\left(\frac{D^{\mathcal{A}}}{D\varepsilon} \sigma \right)_\sigma^V \right) = i \left(\left(\frac{D^{\mathcal{A}}}{D\varepsilon} \sigma \right)_\sigma^V \right).$$

The notation $(\delta q)_\sigma^H \in T_\sigma \tilde{\mathfrak{g}}_a$ on the other hand denotes the horizontal lift of an element $\delta q \in T_q Q$ to σ . It is clear then that the map F defined in the statement of the lemma must be given by $F(v_q, \sigma_q) := T_{\sigma_q} i \left((v_q)_{\sigma_q}^H \right)$. Let us verify the explicit formula (3.98). Recalling (3.93) we have,

$$(\delta q)_\sigma^H = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} [g_\varepsilon, \xi - \varepsilon[\mathcal{A}(\delta g), \xi]].$$

In particular,

$$\begin{aligned} T_\sigma i((\delta q)_\sigma^H) &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} i([g_\varepsilon, \xi - \varepsilon[\mathcal{A}(\delta g), \xi]]) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \text{Ad}_{g_\varepsilon}(\xi - \varepsilon[\mathcal{A}(\delta g), \xi]) \\ &= [TR_{g^{-1}}\delta g, \text{Ad}_g \xi] - \text{Ad}_g[V_a(TL_{g^{-1}}\delta g), \xi] = [H_q(TR_{g^{-1}}\delta g), \text{Ad}_g \xi]. \end{aligned}$$

This concludes the proof. $\square$

In short, $\delta\bar{\sigma} = i(\delta\sigma) + F(\delta q, \sigma)$, where $\delta\sigma := \frac{D^{\mathcal{A}}}{D\varepsilon}\Big|_{\varepsilon=0} \sigma$. Note that $i(\delta\sigma) \in \mathfrak{g}_q$ and $F(\delta q, \sigma) \in \mathfrak{g}_q^\perp$. One computes

$$\delta S = \int_0^1 \gamma_Q(\dot{q}, D_\varepsilon \dot{q}_\varepsilon) + \gamma_G(\bar{\sigma}, \delta\bar{\sigma}) dt = \int_0^1 -\gamma_Q(D_t \dot{q}, \delta q) + \bar{\gamma}(\sigma, \delta\sigma) dt.$$

Using the constrained variations computed in [7],¹⁴

$$\delta\sigma = \frac{D^{\mathcal{A}}}{Dt} \eta - [\sigma, \eta] + \tilde{\mathcal{B}}(\delta q, \dot{q}) \in \tilde{\mathfrak{g}}_a,$$

gives the horizontal and vertical Lagrange–Poincaré equations

$$\frac{D}{Dt} \dot{q} = -\left\langle \sigma, \mathbf{i}_{\dot{q}} \tilde{\mathcal{B}} \right\rangle^\sharp, \quad \frac{D^{\mathcal{A}}}{Dt} \sigma = 0, \quad (3.99)$$

where we defined

$$\gamma_Q\left(\left\langle \rho, \mathbf{i}_{v_q} \tilde{\mathcal{B}} \right\rangle^\sharp, w_q\right) = \bar{\gamma}\left(\rho, \mathbf{i}_{v_q} \tilde{\mathcal{B}}(w_q)\right) = \bar{\gamma}\left(\rho, \tilde{\mathcal{B}}(v_q, w_q)\right),$$

for all $v_q, w_q \in T_q Q$ and $\rho \in \pi^{-1}(q) \subset \tilde{\mathfrak{g}}_a$. Using expression (3.90) one computes the first equality in

$$\left\langle \rho, \mathbf{i}_{v_q} \tilde{\mathcal{B}} \right\rangle^\sharp = ([\bar{J}(v_q), \bar{\rho}])_Q(q) = -\nabla_{v_q} \bar{\rho}_Q, \quad (3.100)$$

and the second equality is shown as follows. We have

$$([\bar{J}(v_q), \bar{\rho}])_Q(q) = [\bar{\rho}_Q, \bar{J}(v_q)_Q](q) = \nabla_{\bar{\rho}_Q(q)} \bar{J}(v_q)_Q - \nabla_{v_q} \bar{\rho}_Q. \quad (3.101)$$

The first term on the right hand side vanishes since $\bar{\rho}_Q(q) = 0$. Therefore, (3.99) becomes

$$\frac{D}{Dt} \dot{q} = \nabla_{\dot{q}} \bar{\sigma}_Q, \quad \frac{D^{\mathcal{A}}}{Dt} \sigma = 0. \quad (3.102)$$

¹⁴The computation goes as follows: For a variation $g_\varepsilon(t)$ of a curve $g(t) = g_0(t)$ we obtain

$$\begin{aligned} \delta\sigma &= \frac{D^{\mathcal{A}}}{D\varepsilon}\Big|_{\varepsilon=0} [g_\varepsilon, \mathcal{A}(\dot{g}_\varepsilon)] = [g, \delta(\mathcal{A}(\dot{g})) + [\mathcal{A}(\delta g), \mathcal{A}(\dot{g})]] \\ &= [g, d\mathcal{A}(\delta g, \dot{g}) + \partial_t(\mathcal{A}(\delta g)) + [\mathcal{A}(\delta g), \mathcal{A}(\dot{g})]] \\ &= [g, \mathcal{B}(\delta g, \dot{g})] + [g, \partial_t(\mathcal{A}(\delta g)) + [\mathcal{A}(\dot{g}), \mathcal{A}(\delta g)] - [\mathcal{A}(\dot{g}), \mathcal{A}(\delta g)]] \\ &= \tilde{\mathcal{B}}(\delta q, \dot{q}) + \frac{D^{\mathcal{A}}}{Dt} \eta - [\sigma, \eta], \end{aligned}$$

where we used a standard identity for the exterior derivative of a one-form, $d\mathcal{A}(\delta g, \dot{g}) = \delta\mathcal{A}(\dot{g}) - \partial_t\mathcal{A}(\delta g) - \mathcal{A}([\delta g, \dot{g}])$. Moreover, we used equations (3.90), (3.93), (3.95); and we set $\eta := [g, \mathcal{A}(\delta g)]$ in the last equality.

The first of these equations is the geodesic equation on Q up to a forcing term on the right hand side. One recognises in particular, as we already know, that horizontal geodesics ($\sigma = 0$) on G project to geodesics on Q . The forcing term encodes the obstruction for a geodesic on G to project to a geodesic on Q . In Section 3.6.5 below, we will obtain an analogous description for Riemannian cubics, using the technique of second-order Lagrange–Poincaré reduction.

Remark 3.18. *Once a solution curve to the Lagrange–Poincaré equations has been found, one needs to reconstruct the corresponding curve $g(t)$. This is done via the reconstruction relations. Suppose a curve $t \mapsto \sigma(t) \in \tilde{\mathfrak{g}}_a$ solves the Lagrange–Poincaré equations for given initial conditions at $t = 0$, and denote the base curve by $q(t) := \pi(\sigma(t))$, as above. Then it follows from the reduction map (3.96) that the reconstructed curve $g(t)$ satisfies*

$$\dot{g} = TR_g(\bar{J}(\dot{q}) + \bar{\sigma}). \quad (3.103)$$

Indeed,

$$TR_{g^{-1}}\dot{g} = H_q(TR_{g^{-1}}\dot{g}) + V_q(TR_{g^{-1}}\dot{g}) = \bar{J}(\dot{q}) + \text{Ad}_g \mathcal{A}(\dot{g}) = \bar{J}(\dot{q}) + \bar{\sigma}.$$

Remark 3.19. *The projections to Q of (possibly non-horizontal) geodesics in G are called ballistic curves in [69].*

Remark 3.20. *Note that since the Lagrangian $L(g, \dot{g}) = \frac{1}{2} \|\dot{g}\|_g^2$ is not only G_a -invariant but G -invariant, the Lagrange–Poincaré equations (3.102) can be further reduced to yield the Euler–Poincaré equations*

$$\frac{d}{dt}\xi = -\text{ad}_\xi^\dagger \xi = 0, \quad \xi = TR_{g^{-1}}\dot{g}.$$

We thus obtain that ξ is a constant. This form of the equation is however not adapted for our purpose since it does not involve the manifold Q .

Remark 3.21. *If Q is a symmetric space, then the first equation in (3.102) can be used together with the second equality of (3.100) to write*

$$\dot{\bar{J}}_Q(q) = \frac{D}{Dt}\dot{q} = -([\bar{J}, \bar{\sigma}])_Q(q). \quad (3.104)$$

Since both $\dot{\bar{J}}$ and $[\bar{J}, \bar{\sigma}]$ are in $\mathfrak{g}_q^\perp$ we conclude that $\dot{\bar{J}} = -[\bar{J}, \bar{\sigma}]$. The second equation in (3.102) together with Lemma 3.17 yields

$$\dot{\bar{\sigma}} = \frac{d}{dt}i(\sigma) = F(\dot{q}, \sigma) = [\bar{J}, \bar{\sigma}],$$

where we recall that $\bar{\sigma} := i(\sigma)$. For symmetric spaces, (3.102) is therefore equivalent to

$$\dot{\bar{J}} = [\bar{\sigma}, \bar{J}], \quad \dot{\bar{\sigma}} = [\bar{J}, \bar{\sigma}]. \quad (3.105)$$

In particular, $\xi = \bar{J} + \bar{\sigma}$ is a constant, as in Remark 3.20 above.

3.6.3 Example: $\mathbf{G} = \mathbf{SO}(3)$, $\mathbf{Q} = \mathbf{S}^2$

We work with the conventions of Remark 3.6. Choose as anchor point a the North pole $\mathbf{e}_z \in S^2$ and define the map

$$(\cdot)_3 : \mathfrak{so}(3) \rightarrow \mathbb{R}, \quad \widehat{\Omega} \mapsto (\widehat{\Omega})_3 := \Omega \cdot \mathbf{e}_z.$$

The isotropy subgroup $SO(3)_a \cong S^1 \subset SO(3)$ corresponds to rotations around $\mathbf{e}_z$, and $\mathfrak{so}(3)_a$ is identified with $\mathbb{R}$ via the map $\lambda \mathbf{e}_z \mapsto \lambda$. The vector bundle $\widetilde{\mathfrak{so}(3)}_a$ is isomorphic to $S^2 \times \mathbb{R}$ via

$$\widetilde{\mathfrak{so}(3)}_a \rightarrow S^2 \times \mathbb{R}, \quad [\Lambda, \lambda \widehat{\mathbf{e}}_z] \mapsto (\Lambda \mathbf{e}_z, \lambda). \quad (3.106)$$

In particular, it follows that the space of reduced variables $TS^2 \times_{S^2} \widetilde{\mathfrak{so}(3)}_a$ can be identified with $TS^2 \times \mathbb{R}$. The map $\alpha_{\mathcal{A}}^{(1)}$ defined in (3.96) becomes

$$\alpha_{\mathcal{A}}^{(1)} : TSO(3)/S^1 \rightarrow TS^2 \times \mathbb{R}, \quad [\Lambda, \dot{\Lambda}] \mapsto (\mathbf{x}, \dot{\mathbf{x}}, (\Lambda^{-1} \dot{\Lambda})_3), \quad (3.107)$$

where $(\mathbf{x}, \dot{\mathbf{x}}) = T_\Lambda \Pi(\Lambda, \dot{\Lambda}) = \frac{d}{d\varepsilon} \big|_{\varepsilon=0} \Lambda(\varepsilon) \mathbf{e}_z$ for any curve $\Lambda(\varepsilon)$ whose tangent vector at $\varepsilon = 0$ is $(\Lambda, \dot{\Lambda})$. Moreover, the mapping i in (3.94) is, through the correspondence (3.106),

$$i : S^2 \times \mathbb{R} \rightarrow \mathfrak{so}(3), \quad (\mathbf{x}, \lambda) \mapsto \lambda \widehat{\mathbf{x}}. \quad (3.108)$$

The geodesic equations (3.102), respectively (3.105), on $SO(3)$ become

$$\dot{\bar{\mathbf{J}}} = \bar{\boldsymbol{\sigma}} \times \bar{\mathbf{J}}, \quad \dot{\bar{\boldsymbol{\sigma}}} = \bar{\mathbf{J}} \times \bar{\boldsymbol{\sigma}},$$

together with the reconstruction relation $TR_{\Lambda^{-1}} \dot{\Lambda} = \widehat{\bar{\mathbf{J}}} + \widehat{\bar{\boldsymbol{\sigma}}}$.

3.6.4 Cubics and ballistic curves

In this section we show that certain types of ballistic curves in a symmetric space Q are Riemannian cubics. Recall from Remark 3.19 that we call a ballistic curve the projection $q(t) \in Q$ of a geodesic $g(t) \in G$. Geodesics in G are given by the Lagrange–Poincaré equations (3.102). As explained in Remark 3.21 these are equivalent to

$$\dot{\bar{J}} = [\bar{\sigma}, \bar{J}], \quad \dot{\bar{\sigma}} = [\bar{J}, \bar{\sigma}], \quad (3.109)$$

where $\bar{J} := \bar{J}(\dot{q})$. As we observed above, the (constant) right-invariant velocity is $TR_{g^{-1}}\dot{g} = \xi = \bar{J} + \bar{\sigma}$, where $\bar{\sigma} := i(\sigma)$ with i defined in (3.94). Moreover, $\bar{J}(t) = \text{Ad}_{g(t)} \bar{J}(0)$ and $\bar{\sigma}(t) = \text{Ad}_{g(t)} \bar{\sigma}(0)$, where we assumed for simplicity that $g(0) = e$. This last assertion follows since $\text{Ad}_{g(t)^{-1}}(\xi) = \xi$, hence

$$\bar{J}(t) = H_{q(t)}(\xi) = \text{Ad}_{g(t)} H_a(\text{Ad}_{g(t)^{-1}}(\xi)) = \text{Ad}_{g(t)} H_a(\xi) = \text{Ad}_{g(t)} \bar{J}(0),$$

and similarly for the vertical part of ξ .

It turns out that some geodesics on G project to Riemannian cubics on Q . This happens under a certain condition on the (constant) velocity vector ξ that is related to the decomposition into horizontal and vertical parts.

Theorem 3.22. *The projection $q(t) \in Q$ of a geodesic $g(t) \in G$ is a Riemannian cubic if and only if at time $t = 0$*

$$[\bar{\sigma}, [\bar{\sigma}, [\bar{\sigma}, \bar{J}]]] + [\bar{J}, [\bar{J}, [\bar{J}, \bar{\sigma}]]] = 0. \quad (3.110)$$

Proof. Since Ad is a Lie automorphism it follows from $\bar{J}(t) = \text{Ad}_{g(t)} \bar{J}(0)$ and $\bar{\sigma}(t) = \text{Ad}_{g(t)} \bar{\sigma}(0)$ that (3.110) holds at $t = 0$ if and only if it holds at all times. Furthermore we obtain from (3.109) that

$$\ddot{\bar{J}} = [\bar{J} + \bar{\sigma}, \dot{\bar{J}}], \quad \ddot{\bar{\sigma}} = [\bar{J} + \bar{\sigma}, \dot{\bar{\sigma}}].$$

Therefore, using also (3.53) and (3.51),

$$\begin{aligned} H_q \left(\ddot{\bar{J}} + 2 [\ddot{\bar{J}}, \bar{J}] \right) &= H_q \left([\bar{\sigma} - \bar{J}, \ddot{\bar{J}}] \right) = [\bar{\sigma}, H_q(\ddot{\bar{J}})] - [\bar{J}, V_q(\ddot{\bar{J}})] \\ &= [\bar{\sigma}, [\bar{\sigma}, \dot{\bar{J}}]] - [\bar{J}, [\bar{J}, \dot{\bar{J}}]] = [\bar{\sigma}, [\bar{\sigma}, [\bar{\sigma}, \bar{J}]]] + [\bar{J}, [\bar{J}, [\bar{J}, \bar{\sigma}]]]. \end{aligned}$$

The theorem now follows from equation (3.61) for cubics in symmetric spaces. $\square$

It is clear that (3.110) is satisfied if $\bar{\sigma} = 0$ at time $t = 0$. This leads to geodesics $q(t)$ since the projection of a horizontal geodesic is a geodesic. Another class of solutions is given by the following corollaries.

Corollary 3.23. *Let $q(t) \in Q$ be the projection of a geodesic $g(t) \in G$ that satisfies $[\bar{J}, \bar{\sigma}] = 0$ at time $t = 0$. Then, $q(t)$ is a geodesic, see equation (3.104), and in particular a Riemannian cubic.*

Corollary 3.24. *Let $q(t) \in Q$ be the projection of a geodesic $g(t) \in G$ that satisfies, at time $t = 0$,*

$$[\bar{J}, [\bar{J}, \bar{\sigma}]] = c\bar{\sigma}, \quad [\bar{\sigma}, [\bar{\sigma}, \bar{J}]] = c\bar{J}, \quad (3.111)$$

for some $c \in \mathbb{R}$. Then, $q(t)$ is a Riemannian cubic.

Remark 3.25. *For the example $G = SO(3)$, $Q = S^2$, the solutions to equation (3.110) are illustrated in Figure 3.2. These solutions are fully described by Corollaries 3.23 and 3.24. Namely, by means of the standard identity*

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$$

for vectors in $\mathbb{R}^3$, (3.110) is seen to be equivalent to

$$\left(\|\bar{\sigma}\|^2 - \|\bar{J}\|^2 \right) \bar{J} \times \bar{\sigma} = 0.$$

Possible solutions are given by $\bar{J} \times \bar{\sigma} = 0$, or by $\|\bar{\sigma}\|^2 = \|\bar{J}\|^2$, at $t = 0$. Therefore, solutions are either described by Corollary 3.23 or Corollary 3.24. The first case is equivalent to $\bar{J} = 0$ or $\bar{\sigma} = 0$ at $t = 0$, and therefore at all times. This corresponds to trivial projected curves $\mathbf{x}(t) = \mathbf{x}(0) \in S^2$, and to projections of horizontal geodesics, respectively.

We proceed to analyse the alternative solutions given by $\|\bar{\sigma}\|^2 = \|\bar{J}\|^2$. For a given initial velocity $\dot{\mathbf{x}}(0) = \mathbf{v} \in T_{\mathbf{x}(0)}S^2$, the projection $\mathbf{x}(t) \in S^2$ describes a constant speed rotation of $\mathbf{x}(0)$ around the axis

$$\Omega = \bar{J} + \bar{\sigma} = \mathbf{x} \times \dot{\mathbf{x}} \pm \|\dot{\mathbf{x}}\| \mathbf{x}.$$

Hence, the (constant) rotation axis lies in the plane spanned by $\bar{J}$ and $\mathbf{x}$, enclosing a 45° angle with both vectors. The curve $\mathbf{x}(t)$ moves with constant speed along a circle of radius $r = \frac{1}{\sqrt{2}}$.

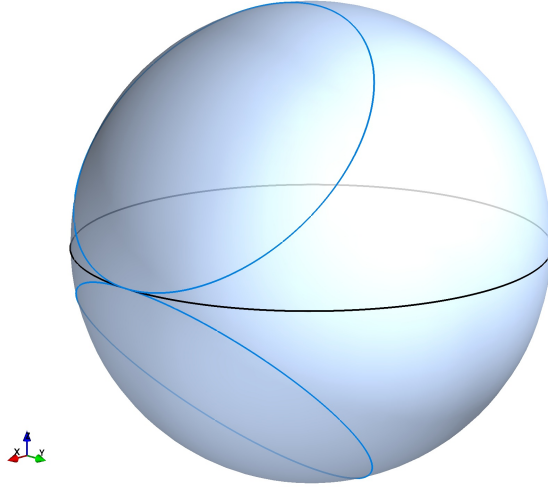


Fig. 3.2: *Ballistic curves and cubics on the sphere.* For given initial position and velocity, there are two types of trajectories that are, simultaneously, projections of geodesics on the rotation group (ballistic curves) and Riemannian cubics. The two trajectories are shown for initial position $(1, 0, 0)$ and initial velocity parallel to the y -axis. In black a constant-speed trajectory along the equator corresponding to the projection of a horizontal geodesic on the rotation group. The blue curves are the circular constant-speed trajectories of radius $\frac{1}{\sqrt{2}}$ described in Corollary 3.24 and Remark 3.25. *This figure appears in [40] – reproduced with kind permission from Springer Science and Business Media.*

3.6.5 Second-order Lagrange–Poincaré reduction

The goal of this section is to derive an equation similar to (3.102), but now for Riemannian cubics instead of geodesics. That is, we aim to exhibit the obstruction for a cubic on G to project to a cubic on Q . For this purpose we will employ the technique of second-order Lagrange–Poincaré reduction in a manner closely related to the first-order case.

To discuss second-order reduction, we introduce the quotient $T^{(2)}G/G_a$ of $T^{(2)}G$ by the natural action of G_a , whose elements will be denoted by $[g, \dot{g}, \ddot{g}] \in T^{(2)}G/G_a$. In the following let us write $2\tilde{\mathfrak{g}}_a$ to denote the Whitney sum $\tilde{\mathfrak{g}}_a \oplus \tilde{\mathfrak{g}}_a$. We will make use of the bundle diffeomorphism $\alpha_{\mathcal{A}}^{(2)} : T^{(2)}G/G_a \rightarrow T^{(2)}Q \times_Q 2\tilde{\mathfrak{g}}_a$,

$$[g, \dot{g}, \ddot{g}] \mapsto (q, \dot{q}, \ddot{q}) \times [g, \mathcal{A}(\dot{g})] \oplus \frac{D\mathcal{A}}{Dt} [g, \mathcal{A}(\dot{g})] = (q, \dot{q}, \ddot{q}) \times [g, \mathcal{A}(\dot{g})] \oplus [g, \partial_t \mathcal{A}(\dot{g})], \quad (3.112)$$

see [50]. Here, we defined $q(t) := \Pi(g(t))$, where $g(t)$ is a curve representing $[g, \dot{g}, \ddot{g}]$, that is, $(q, \dot{q}, \ddot{q}) = T_g^{(2)}\Pi(g, \dot{g}, \ddot{g})$. Denoting the right-invariant velocity by $\xi = TR_{g^{-1}}\dot{g}$ and

using the definition (3.89) of $\mathcal{A}$ together with (3.88), this becomes

$$[g, \dot{g}, \ddot{g}] \mapsto (q, \dot{q}, \ddot{q}) \times [g, \text{Ad}_{g^{-1}} V_q(\xi)] \oplus [g, \text{Ad}_{g^{-1}} V_q(\dot{\xi})]. \quad (3.113)$$

Remark 3.26. *In the example $G = SO(3)$, $Q = S^2$, working with the conventions laid out in Remark 3.6, the space of reduced variables $T^{(2)}S^2 \times_Q \widetilde{2\mathfrak{so}(3)}_a$ can be identified with $T^{(2)}S^2 \times \mathbb{R}^2$. The map $\alpha_{\mathcal{A}}^{(2)}$ is*

$$\alpha_{\mathcal{A}}^{(2)} : T^{(2)}SO(3)/S^1 \rightarrow T^{(2)}S^2 \times \mathbb{R}^2, \quad [\Lambda, \dot{\Lambda}, \ddot{\Lambda}] \mapsto (\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}, (\Lambda^{-1}\dot{\Lambda})_3, \partial_t(\Lambda^{-1}\dot{\Lambda})_3). \quad (3.114)$$

The reduced Lagrangian. For $L : T^{(2)}G \rightarrow \mathbb{R}$ a G_a -invariant Lagrangian we define the reduced Lagrangian ℓ by $L = \ell \circ \alpha_{\mathcal{A}}^{(2)}$. Consider the Lagrangian for Riemannian cubics $L = \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2$ and note the following equalities,

$$\begin{aligned} L(g, \dot{g}, \ddot{g}) &= \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2 = \frac{1}{2} \left\| \dot{\xi} \right\|_{\mathfrak{g}}^2 = \frac{1}{2} \left\| H_q(\dot{\xi}) \right\|_{\mathfrak{g}}^2 + \frac{1}{2} \left\| V_q(\dot{\xi}) \right\|_{\mathfrak{g}}^2 \\ &= \frac{1}{2} \left\| (\dot{\xi})_Q(q) \right\|_q^2 + \frac{1}{2} \left\| V_q(\dot{\xi}) \right\|_{\mathfrak{g}}^2 \\ &= \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} - \nabla_{\dot{q}}(V_q(\xi))_Q \right\|_q^2 + \frac{1}{2} \left\| V_q(\dot{\xi}) \right\|_{\mathfrak{g}}^2, \end{aligned} \quad (3.115)$$

where we used the right-invariance of L , the definition of the normal metric, and part (iii) of Proposition 3.1. It follows from (3.113) and (3.115) that the reduced Lagrangian $\ell : T^{(2)}Q \times_Q 2\tilde{\mathfrak{g}}_a \rightarrow \mathbb{R}$ reads

$$\ell(q, \dot{q}, \ddot{q}, \sigma, \dot{\sigma}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} - \nabla_{\dot{q}} \bar{\sigma}_Q \right\|_q^2 + \frac{1}{2} \|\dot{\sigma}\|_{\tilde{\mathfrak{g}}_a}^2, \quad (3.116)$$

where we recall $\dot{\sigma} := \frac{D^A}{Dt} \sigma$. The reduced Lagrangian therefore measures the deviations from the geodesic Lagrange–Poincaré equations (3.102).

Remark 3.27. *In the example $G = SO(3)$, $Q = S^2$, the reduced Lagrangian $\ell : T^{(2)}S^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is*

$$(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}, \sigma, \dot{\sigma}) \mapsto \frac{1}{2} \|D_t \dot{\mathbf{x}} - \sigma \mathbf{x} \times \dot{\mathbf{x}}\|_{\mathbf{x}}^2 + \frac{1}{2} \dot{\sigma}^2 = \frac{1}{2} \|D_t \dot{\mathbf{x}}\|_{\mathbf{x}}^2 - \sigma D_t \dot{\mathbf{x}} \cdot (\mathbf{x} \times \dot{\mathbf{x}}) + \frac{1}{2} \sigma^2 \|\dot{\mathbf{x}}\|_{\mathbf{x}}^2 + \frac{1}{2} \dot{\sigma}^2,$$

where the norm $\|\cdot\|_{\mathbf{x}}$ is evaluated as the standard Euclidean norm.

Coupling. The reduced Lagrangian couples the horizontal and vertical parts of the motion through the term $\nabla_{\dot{q}}\bar{\sigma}_Q$. This explains the absence of a general horizontal lifting property for Riemannian cubics studied in Section 3.5.2. Indeed, let us instead define the Lagrangian

$$L_{KK} : T^{(2)}G \rightarrow \mathbb{R}, \quad (g, \dot{g}, \ddot{g}) \mapsto \frac{1}{2} \left\| \frac{D}{Dt} T\Pi(\dot{g}) \right\|_{\Pi(g)}^2 + \frac{1}{2} \|\partial_t \mathcal{A}(\dot{g})\|_{\mathfrak{g}}^2 \quad (3.117)$$

with reduced Lagrangian

$$\ell_{KK} : T^{(2)}Q \times_Q 2\tilde{\mathfrak{g}}_a \rightarrow \mathbb{R}, \quad (q, \dot{q}, \ddot{q}, \sigma, \dot{\sigma}) \mapsto \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} \right\|_q^2 + \frac{1}{2} \|\dot{\sigma}\|_{\tilde{\mathfrak{g}}_a}^2. \quad (3.118)$$

The Lagrangian L_{KK} belongs to a class of Lagrangians that were studied in [50] as natural second-order generalisations of the Kaluza–Klein Lagrangian. The reduced Lagrangians ℓ in (3.116) and ℓ_{KK} in (3.118) differ by the coupling term $\nabla_{\dot{q}}\bar{\sigma}_Q$. The decoupled form of ℓ_{KK} leads to a general horizontal lifting theorem. Namely, any horizontal lift $g(t)$ to G of a cubic $q(t)$ on Q is a critical point of the action $\int L_{KK} dt$.

Lagrange–Poincaré equations. We now compute the Lagrange–Poincaré equations. Taking ε -variations and defining $V_q := \frac{D}{Dt}\dot{q} - \nabla_{\dot{q}}\bar{\sigma}_Q$, we have, for the first term of (3.116),

$$\delta \int_0^1 \frac{1}{2} \left\| \frac{D}{Dt} \dot{q} - \nabla_{\dot{q}} \bar{\sigma}_Q \right\|_q^2 dt = \int_0^1 \gamma_Q \left(V_q, \frac{D}{D\varepsilon} \frac{D}{Dt} \dot{q} - \frac{D}{D\varepsilon} \nabla_{\dot{q}} \bar{\sigma}_Q \right) dt.$$

We then compute

$$\frac{D}{D\varepsilon} \frac{D}{Dt} \dot{q} = \frac{D}{Dt} \frac{D}{D\varepsilon} \dot{q} + R(\delta q, \dot{q}) \dot{q}$$

and

$$\begin{aligned} \frac{D}{D\varepsilon} \nabla_{\dot{q}}(\bar{\sigma})_Q &= \frac{D}{D\varepsilon} \left(\frac{D}{Dt} \bar{\sigma}_Q(q) - (\partial_t \bar{\sigma})_Q(q) \right) \\ &= \frac{D}{Dt} \frac{D}{D\varepsilon} \bar{\sigma}_Q(q) + R(\delta q, \dot{q}) \bar{\sigma}_Q(q) - \frac{D}{D\varepsilon} (\partial_t \bar{\sigma})_Q(q) \\ &= \frac{D}{Dt} \left(\nabla_{\delta q} \bar{\sigma}_Q + (\delta \bar{\sigma})_Q(q) \right) - (\delta \partial_t \bar{\sigma})_Q(q) - \nabla_{\delta q} (\partial_t \bar{\sigma})_Q, \end{aligned}$$

Lemma 3.17 shows that

$$(\delta \bar{\sigma})_Q(q) = (i(\delta \sigma))_Q(q) + (F(\delta q, \sigma))_Q(q) = (F(\delta q, \sigma))_Q(q),$$

since $\delta\sigma := \left. \frac{D^{\mathcal{A}}}{D\varepsilon} \right|_{\varepsilon=0} \sigma_\varepsilon \in (\tilde{\mathfrak{g}}_a)_q$ and therefore $i(\delta\sigma) \in \mathfrak{g}_q$. So we have

$$\begin{aligned} \int_0^1 \gamma_Q \left(V_q, \frac{D}{Dt} (\delta\bar{\sigma})_Q(q) \right) dt &= - \int_0^1 \gamma_Q \left(\frac{D}{Dt} V_q, (F(\delta q, \sigma))_Q(q) \right) dt \\ &= - \int_0^1 \gamma \left(\bar{J} \left(\frac{D}{Dt} V_q \right), F(\delta q, \sigma) \right) dt = - \int_0^1 \gamma_Q \left(F_\sigma^T \bar{J} \left(\frac{D}{Dt} V_q \right), \delta q \right) dt \end{aligned}$$

and

$$\begin{aligned} \int_0^1 \gamma_Q (V_q, (\delta\partial_t \bar{\sigma})_Q(q)) dt &= \int_0^1 \gamma (\bar{J}(V_q), \partial_t \delta\bar{\sigma}) = - \int_0^1 \gamma (\partial_t \bar{J}(V_q), \delta\bar{\sigma}) \\ &= - \int_0^1 \gamma (\partial_t \bar{J}(V_q), i(\delta\sigma) + F(\delta q, \sigma)) = - \int_0^1 \bar{\gamma} (i_q^T \partial_t \bar{J}(V_q), \delta\sigma) + \gamma_Q (F_\sigma^T \partial_t \bar{J}(V_q), \delta q), \end{aligned}$$

where $F_{\sigma_q}^T : \mathfrak{g} \rightarrow T_q Q$ is the transpose of the map $F_{\sigma_q} : T_q Q \rightarrow \mathfrak{g}$, $F_{\sigma_q}(v_q) := F(v_q, \sigma_q)$, and $i_q^T : \mathfrak{g} \rightarrow (\tilde{\mathfrak{g}}_a)_q$ is the metric transpose of the map $i_q : (\tilde{\mathfrak{g}}_a)_q \rightarrow \mathfrak{g}$ (the restriction of the map i in (3.94) to the fibre $(\tilde{\mathfrak{g}}_a)_q$ of $\tilde{\mathfrak{g}}_a$ at q).

For the second term of (3.116), we have

$$\delta \int_0^1 \frac{1}{2} \|\dot{\sigma}\|_{\tilde{\mathfrak{g}}_a}^2 dt = \int_0^1 \bar{\gamma} \left(\dot{\sigma}, \frac{D^{\mathcal{A}}}{D\varepsilon} \dot{\sigma} \right) dt.$$

Using the variations

$$\delta\sigma = \frac{D^{\mathcal{A}}}{Dt} \eta - [\sigma, \eta] + \tilde{\mathcal{B}}(\delta q, \dot{q}) \in \tilde{\mathfrak{g}}_a, \quad \delta\dot{\sigma} = \frac{D^{\mathcal{A}}}{Dt} \delta\sigma - [\tilde{\mathcal{B}}(\dot{q}, \delta q), \sigma] \in \tilde{\mathfrak{g}}_a,$$

(see [50]) and the formula

$$\frac{d}{dt} J(\alpha(t)) - J \left(\frac{D}{Dt} \alpha(t) \right) = \mathcal{F}^\nabla(\alpha(t), \dot{q}(t)),$$

where $\langle \mathcal{F}^\nabla(\alpha_q, v_q), \eta \rangle := \langle \alpha_q, \nabla_{v_q} \eta_Q \rangle$, (see [70]) we get the equations

$$\begin{aligned} \frac{D^2}{Dt^2} V_q + \nabla \bar{\sigma}_Q^T \cdot \frac{D}{Dt} V_q + \nabla (\partial_t \bar{\sigma})_Q^T \cdot V_q + R(V_q, \dot{q}) \dot{q} \\ + \left\langle \frac{D^{\mathcal{A}}}{Dt} \dot{\sigma} + \text{ad}_\sigma^\dagger \dot{\sigma} + i_q^T \partial_t \bar{J}(V_q), \mathbf{i}_{\dot{q}} \tilde{\mathcal{B}} \right\rangle^\# = F_\sigma^T (\mathcal{F}^\nabla(V_q^\flat, \dot{q}))^\# \\ \left(\frac{D^{\mathcal{A}}}{Dt} + \text{ad}_\sigma^\dagger \right) \left(\frac{D^{\mathcal{A}}}{Dt} \dot{\sigma} + i_q^T \partial_t \bar{J}(V_q) \right) = 0, \end{aligned}$$

where we recall that $V_q := \frac{D}{Dt} \dot{q} - \nabla_{\dot{q}} \bar{\sigma}_Q \in TQ$.

Using (3.100) these equations can be rewritten as

$$\begin{aligned} \frac{D^3}{Dt^3}\dot{q} + R\left(\frac{D}{Dt}\dot{q}, \dot{q}\right)\dot{q} &= \frac{D^2}{Dt^2}\nabla_{\dot{q}}\bar{\sigma}_Q - \nabla\bar{\sigma}_Q^T \cdot \frac{D}{Dt}V_q - \nabla(\partial_t\bar{\sigma})_Q^T \cdot V_q + R(\nabla_{\dot{q}}\bar{\sigma}_Q, \dot{q})\dot{q} \\ &\quad + \nabla_{\dot{q}}\left(i\left(\ddot{\sigma} + \text{ad}_\sigma^\dagger\dot{\sigma} + i_q^T\partial_t\bar{J}(V_q)\right)\right)_Q + F_\sigma^T\left(\mathcal{F}^\nabla(V_q^\flat, \dot{q})\right)^\sharp \end{aligned} \quad (3.119)$$

$$\left(\frac{D^A}{Dt} + \text{ad}_\sigma^\dagger\right)\left(\ddot{\sigma} + i_q^T\partial_t\bar{J}(V_q)\right) = 0. \quad (3.120)$$

These equations are the second-order analogue of (3.102). If the left hand side of the first one equals zero, then $q(t) = \Pi(g(t))$ is a Riemannian cubic. Hence the right hand side of (3.119) is the obstruction for the projected curve to be a cubic. For symmetric spaces, solutions to (3.119)–(3.120) with vanishing obstruction include in particular the horizontal curves characterised in Theorem 3.7, but also the special geodesics on the group that project to the ballistic curves of Section 3.6.4.

Remark 3.28. For $G = SO(3)$, $Q = S^2$, the Lagrange–Poincaré equations are computed as

$$\begin{aligned} \frac{D^3}{Dt^3}\dot{\mathbf{x}} + R\left(\frac{D}{Dt}\dot{\mathbf{x}}, \dot{\mathbf{x}}\right)\dot{\mathbf{x}} &= \sigma^2\ddot{\mathbf{x}}^\perp + 2\sigma\dot{\sigma}\dot{\mathbf{x}} + \ddot{\sigma}(\mathbf{x} \times \dot{\mathbf{x}}) + 3\dot{\sigma}(\mathbf{x} \times \ddot{\mathbf{x}}) \\ &\quad + \sigma\left(2[\mathbf{x} \times \ddot{\mathbf{x}} + (\dot{\mathbf{x}} \times \ddot{\mathbf{x}})^\perp] + \|\dot{\mathbf{x}}\|^2(\mathbf{x} \times \dot{\mathbf{x}})\right) - \alpha(\mathbf{x} \times \dot{\mathbf{x}}) \\ &\quad - \mathbf{x} \cdot (\dot{\mathbf{x}} \times \ddot{\mathbf{x}}) + \sigma\|\dot{\mathbf{x}}\|^2 - \ddot{\sigma} = \alpha, \end{aligned}$$

for some constant $\alpha \in \mathbb{R}$. Here we denoted by $\mathbf{v}^\perp = \mathbf{v} - \mathbf{x}(\mathbf{x} \cdot \mathbf{v})$ the orthogonal projection of $\mathbf{v}$ onto the tangent plane to S^2 at $\mathbf{x}$.

3.7 Final remarks

In this chapter we investigated Riemannian cubics in object manifolds with normal metrics and in particular their relation to Riemannian cubics on the Lie group of transformations. Let us briefly summarise what we did.

Our starting point was the definition, in Section 3.2, of necessary concepts and a treatment of covariant derivatives for normal metrics in Section 3.3. The derivation of the Euler–Lagrange equation for cubics from the viewpoint of normal metrics followed in Section 3.4. The examples of Lie groups with invariant metrics and Type I symmetric

spaces were discussed in detail, and the relation with equations previously present in the literature was clarified. The new form of the equation was seen to lend itself to the analysis of horizontal lifts of cubics, due to the appearance of the horizontal generator of curves.

Section 3.5 proceeded with this line of investigation by deriving several results about horizontal lifting properties of cubics. For symmetric spaces a complete characterisation was achieved of the cubics that can be lifted horizontally to cubics on the group of isometries. In rank-one symmetric spaces this selects geodesics composed with cubic polynomials in time. The section continued with a treatment of the corresponding question in the context of Riemannian submersions. In Section 3.6 certain non-horizontal geodesics on the group were shown to project to cubics in the context symmetric spaces. A complete characterisation of such geodesics was given in the sense of Theorem 3.22. For the unit sphere acted on by the rotation group the corresponding projections were seen to be the circles of radius $\frac{1}{\sqrt{2}}$, traversed at constant speed. A discussion of Lagrange–Poincaré reduction of cubics led to reduced equations that identified the obstruction for projections of cubics to be cubics in the object manifold.

As we explained in the opening paragraphs, the investigations of this chapter were motivated by the need to select interpolation models, in particular in computational anatomy. A class of such models was discussed in Section 2.5. In the next chapter we will revisit this type of model in both more detail and more generality, allowing for a wider range of applications, one of which, to quantum control, will be the topic of Chapter 5.

4 Inexact trajectory planning and inverse problems in the Hamilton–Pontryagin framework

In this chapter we return, in a more general setting, to the inexact template matching application discussed in Section 2.5. We saw there that the Euler–Lagrange equations split into a set of Euler–Poincaré equations that hold on the open time intervals between the nodes, and a set of *node equations*, given in parts (II) and (III) of Theorem 2.11, which describe how to pass from one open interval to the next. The primary purpose of this chapter is to develop a new geometric understanding of the node equations in terms of conjugate momenta. Continuing from this, we develop a numerical algorithm for the trajectory planning problem that respects the geometric properties exhibited by the continuous-time solution. The results of this chapter were published in [42]. Chapter 5 discusses an application to quantum control. It can be read before or after the present one.

4.1 Background and problem formulation

Let us describe the mathematical setup we consider here. We look for an optimal curve $g(t)$ in a Lie group G that acts on a point Q_0 in an *object manifold* and generates a curve $q(t) = g(t)Q_0$, that passes through a sequence of target points at prescribed times. In some instances it is desirable to relax the target constraints in such a way that the optimal curve does not exactly pass through the target points, but still passes *near* them at the prescribed times. This may be achieved by including a *soft constraint* in the cost functional, that is, a term penalising the discrepancy between the trajectory $q(t)$ and the targets. This leads to cost functionals of the following type,

$$S = \int \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}) dt + \frac{1}{2\sigma^2} \sum_i d^2(q(t_i), q_i). \quad (4.1)$$

Here we recall that $\xi^{(j)}$ are the j -th time derivatives of a curve $\xi(t)$ in the Lie algebra that integrates to the curve $g(t)$, which in turn produces the trajectory in the object manifold according to $q(t) = g(t)Q_0$. The first part of the cost is the integral over a Lagrangian ℓ and is associated with the curve on the group. The second part sums up the squares of

the distances d between the curve $q(t)$ and the target points q_i , at the prescribed times t_i . The *tolerance parameter* σ may be adjusted to suitably weight the two parts. Recall from Section 2.5 the example of computational anatomy, where a diffeomorphism group acts by transforming a medical image (or sub-structures thereof, such as points of interest, fibres, or surfaces). See [54] for an overview. In this context the prescribed target images (or sub-structures) may not be diffeomorphically related to the initial one. That is, the initial and target configurations may lie on different group orbits. In general one expects therefore that exact matching may not be possible and works instead with a soft constraint, as we did in (2.48). In the next chapter we will study a problem of similar nature in quantum control, published in [41]. There one considers the group of unitary matrices acting on quantum state space, with the goal of finding the optimal experimental manipulation of the system such that the evolution of an initial state passes near a sequence of given target states. The cost functional is directly related to the required amount of change in the experimental apparatus over time. The introduction of a soft constraint is appropriate in this problem even though in this case the group action is transitive, since by increasing the tolerance parameter σ optimal trajectories may be achieved at a lower cost.

In a more general sense such trajectory planning problems can be thought of as *inverse problems*, where the data points $q_i \in Q$ have been determined experimentally at the times t_i and one seeks the corresponding curve $g(t)$ in configuration space G . In this context a natural choice of the tolerance parameter σ would be a measure of the uncertainty inherent in the experiment, such as standard deviation. The Lagrangian ℓ represents a modeling choice and is specific to the application one has in mind. As we will see in examples, a natural choice of Lagrangian ℓ leads to Riemannian cubics on the Lie group G .

4.2 Main content of the chapter

In Section 4.3 we shall rederive the Euler–Lagrange equations for the higher-order variational problem by using Lagrange multipliers in a generalisation of the symmetry reduced Hamilton–Pontryagin principle of geometric mechanics. In this approach, the derivation of the Euler–Lagrange equations simplifies considerably and a new geomet-

ric interpretation of the node equations emerges. Namely, they describe the evolution of Legendre–Ostrogradsky momenta across the nodes, in which the highest-order momentum experiences a discontinuous jump related to the distance between the curve in the object manifold and the target points. The discontinuity can be understood in terms of a momentarily broken symmetry at the node times. However, if the object manifold is isotropic with respect to a subgroup action then a residual symmetry remains. By Noether’s theorem, this residual symmetry leads to a conservation law across node times.

In Section 4.4 we discuss a number of applications, including rigid body splines, macromolecular configurations and quantum splines. Section 4.5 is concerned with the numerical solution of the inexact trajectory planning problem. More precisely, we describe a geometric discretisation of the higher-order Hamilton–Pontryagin principle for inexact trajectory planning, similar to the approach given in [71] for first-order systems. Our main motivation for the development of a geometric integrator is the exact momentum behaviour of the discrete solution. This leads in turn to a dimensionality reduction of the search space in the numerical optimisation.

4.3 Geometry of the inexact trajectory planning problem

We start with the statement of the problem considered here. One aims at steering from an initial point Q_0 in some *object manifold* Q along an optimal trajectory $q(t)$ that evolves via the action of a Lie group G . That is, $q(t) = g(t)Q_0$, where the right hand side denotes the action of $g(t)$ on Q_0 and the curve $q(t)$ lies in the G -orbit of Q_0 .

The optimality condition is given in terms of a function $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$ defined on the k -fold Cartesian product $k\mathfrak{g}$ of the Lie algebra $\mathfrak{g}$, which measures the cost of the transformation $g(t)$, and a distance function $d : Q \times Q \rightarrow \mathbb{R}$. As we shall see, the integer k determines the degree of smoothness of solution curves. The optimal curve $q(t)$ is required to pass near prescribed target points Q_{t_i} at prescribed *node times* t_i for $i = 1, \dots, l$. This is formalised by including a squared distance term $d^2(g(t_i)Q_0, Q_{t_i})$ in the cost functional, for each i . Thus, the cost functional $S : \mathcal{C}(\mathfrak{g}) \rightarrow \mathbb{R}$, where $\mathcal{C}(\mathfrak{g})$ is a

suitable space of $\mathfrak{g}$ -valued curves (see below) is defined by

$$S[\xi] := \int_0^{t_l} \ell(\xi, \dots, \xi^{(k-1)}) dt + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(g(t_i)Q_0, Q_{t_i}). \quad (4.2)$$

As usual, the notation $\xi^{(j)}$ is shorthand for $d^j\xi/dt^j$, the quantity σ is a *tolerance parameter*, and the curve $g(t)$ originates at the identity $g(0) = e$ and satisfies $\dot{g} = \frac{d}{d\varepsilon}\big|_{\varepsilon=0} \exp(\varepsilon\xi)g$. Let us also recall the notation R_g for multiplication by g from the right and TR_g for its differential, thus $\dot{g} = TR_g\xi$. In analogy to Section 2.5, variations are considered in a space $\mathcal{C}(\mathfrak{g})$ consisting of curves $\xi(t) : [0, t_l] \rightarrow \mathfrak{g}$ for which (4.2) exists and whose restrictions to open intervals (t_{i-1}, t_i) for $i = 1, \dots, l$ are C^{2k-2} and whose j -th derivatives $\xi^{(j)}$ are continuous on $[0, t_l]$ for $j = 0, \dots, k-2$. We also require the existence of one-sided limits of certain derivatives at the node times t_i . Moreover we assume that initial values of $\xi^{(j)}(0)$ for $j = 0, \dots, k-2$, are given.

As we discussed in Section 2.5, this type of trajectory planning problem is familiar, for example, from image registration in computational anatomy, where one typically thinks of Q_0 as a template shape being deformed by a curve of diffeomorphisms $g(t)$, in turn generated by the time-dependent vector field $\xi(t)$ [58]. At times t_i the resulting curve in shape space passes near the given target shapes Q_{t_i} , the parameter σ determining the proximity of the passage. In this case, the Lie group G of diffeomorphisms is infinite dimensional. However, in the present chapter we will restrict ourselves to the case of finite-dimensional Lie groups and object manifolds. As we will see in the next chapter, a finite-dimensional instance for illustrating these ideas arises in quantum control [41], where quantum state vectors evolve under the action of the unitary group. The generator curve $\xi(t)$ in this case corresponds to the Hamiltonian operator, which is controlled in experiments.

4.3.1 Euler–Lagrange equations via Lagrange multipliers

The Euler–Lagrange equations characterise solutions to Hamilton’s principle, $\delta S = 0$, and were derived in Section 2.5. As we noted in Theorem 2.11, the equations split into a set of higher-order Euler–Poincaré equations on the open time intervals between the node times and a number of *node equations* describing how the solution evolves across the nodes.

In our previous formulation of this problem, one must find the variation $\delta g(t_i)$ that is produced by a (time dependent) variation $\delta \xi(t)$. In relation to this task we quoted the result (2.54). The need for this non-trivial result can be removed by taking advantage of Lagrange multipliers in an equivalent variational formulation that we describe now. As this chapter demonstrates, the new approach also provides a geometric interpretation of the node equations and furthermore suggests a geometric numerical procedure for the solution of the problem, see Section 4.5 below.

The method of Lagrange multipliers involves enlarging the space on which the dynamics happen. We define the cost functional S on some space of curves $\mathcal{C}(G \times k\mathfrak{g} \times k\mathfrak{g}^*)$,

$$S[g, \xi^0, \dots, \xi^{k-1}, \mu^0, \dots, \mu^{k-1}] := \int_0^{t_l} \left[\ell(\xi^0, \xi^1, \dots, \xi^{k-1}) + \langle \mu^0, TR_{g^{-1}} \dot{g} - \xi^0 \rangle \right. \\ \left. + \sum_{r=1}^{k-1} \langle \mu^r, \dot{\xi}^{r-1} - \xi^r \rangle \right] dt + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(g(t_i)Q_0, Q_{t_i}). \quad (4.3)$$

Again some technical assumptions about the space of curves are needed, in order to carry out the variational calculus that follows. Namely, (4.3) exists, the curves are C^1 when restricted to the open intervals (t_{i-1}, t_i) for $i = 1, \dots, l$, and $g, \xi^0, \dots, \xi^{k-2}$ are continuous on $[0, t_l]$. Moreover, one-sided limits of the curves μ^r , $r = 0, \dots, k-1$, exist at the node times t_i , for $i = 0, \dots, l$. We also assume $g(0) = e$ and given initial values $\xi^j(0) = \xi_0^j$, for $j = 0, \dots, k-2$.

Before we take variations of S it is useful to recall the cotangent lift momentum map $J^Q : T^*Q \rightarrow \mathfrak{g}^*$ associated with the action of G on Q . It is defined via the equality

$$\langle \alpha_q, \xi_Q(q) \rangle = \langle J^Q(\alpha_q), \xi \rangle, \quad \text{for any } \alpha_q \in T^*Q, \xi \in \mathfrak{g}, \quad (4.4)$$

where we have used the notation $\xi_Q(q) := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} e^{\varepsilon \xi} q$ and $\langle \cdot, \cdot \rangle$ for the respective duality pairings. We have already encountered this map in several places in earlier chapters. In Section 2.5 it occurred in the special form of the diamond operator $\diamond$, and throughout Chapter 3 it played a crucial role due to its close relation with the horizontal generator of curves in a Riemannian setting, demonstrated in Theorem 3.1 **(iv)**. Notice that in the present chapter we superscribe the letter Q , in order to distinguish J^Q from a different momentum map, denoted J , that we will need later.

For convenience we introduce the shorthand $d_1 d(q_1, q_2) \in T_{q_1}^* Q$ to denote the exterior derivative of the distance function d with respect to the first entry, and $d_1 d(t_i) := d_1 d(g(t_i)Q_0, Q_{t_i})$.

Integrating by parts and using (4.4) we obtain

$$\begin{aligned}
\delta S = & \int_0^{t_l} \left[\langle -\dot{\mu}^0 - \text{ad}_{\xi^0}^* \mu^0, \eta \rangle + \sum_{r=0}^{k-2} \left\langle \frac{\delta \ell}{\delta \xi^r} - \mu^r - \dot{\mu}^{r+1}, \delta \xi^r \right\rangle + \left\langle \frac{\delta \ell}{\delta \xi^{k-1}} - \mu^{k-1}, \delta \xi^{k-1} \right\rangle \right. \\
& + \langle \delta \mu^0, TR_{g^{-1}} \dot{g} - \xi^0 \rangle + \sum_{r=1}^{k-1} \left\langle \delta \mu^r, \dot{\xi}^{r-1} - \xi^r \right\rangle \Big] dt \\
& + \left\langle \mu^0(t_l) + \frac{d(t_l)}{\sigma^2} J^Q(d_1 d(t_l)), \eta(t_l) \right\rangle + \sum_{r=0}^{k-2} \langle \mu^{r+1}(t_l), \delta \xi^r(t_l) \rangle \\
& + \sum_{s=1}^{l-1} \left[\left\langle \mu^0(t_s^-) - \mu^0(t_s^+) + \frac{d(t_s)}{\sigma^2} J^Q(d_1 d(t_s)), \eta(t_s) \right\rangle + \sum_{r=0}^{k-2} \langle \mu^{r+1}(t_s^-) - \mu^{r+1}(t_s^+), \delta \xi^r(t_s) \rangle \right] \\
& - \left\langle \mu^0(0) - \frac{d(0)}{\sigma^2} J^Q(d_1 d(0)), \eta(0) \right\rangle - \sum_{r=0}^{k-2} \langle \mu^{r+1}(0), \delta \xi^r(0) \rangle,
\end{aligned} \tag{4.5}$$

where we set $\eta := TR_{g^{-1}} \delta g$ and used the notation for limits $\mu^r(t_s^-) := \lim_{t \uparrow t_s} \mu^r(t)$ as well as $\mu^r(t_s^+) := \lim_{t \downarrow t_s} \mu^r(t)$. Note that the last line above could have been omitted since by assumption $\eta(0) = 0$ and $\delta \xi^j = 0$ for $j = 0, \dots, k-2$.

We can now read off the Euler–Lagrange equations. On the one hand, for t in any of the open intervals (t_i, t_{i+1}) , $i = 0, \dots, l-1$, we have

$$\dot{\mu}^0 + \text{ad}_{\xi^0}^* \mu^0 = 0, \tag{4.6}$$

$$TR_{g^{-1}} \dot{g} - \xi^0 = 0, \tag{4.7}$$

$$\dot{\xi}^{r-1} - \xi^r = 0, \quad (r = 1, \dots, k-1) \tag{4.8}$$

$$\dot{\mu}^r + \mu^{r-1} - \frac{\delta \ell}{\delta \xi^{r-1}} = 0, \quad (r = 1, \dots, k-1) \tag{4.9}$$

$$\mu^{k-1} - \frac{\delta \ell}{\delta \xi^{k-1}} = 0. \tag{4.10}$$

On the other hand, the node equations are given by

$$\mu^0(t_s^-) - \mu^0(t_s^+) + \frac{d(t_s)}{\sigma^2} J^Q(d_1 d(t_s)) = 0, \quad (s = 1, \dots, l-1) \tag{4.11}$$

$$\mu^r(t_s^-) - \mu^r(t_s^+) = 0, \quad (r = 1, \dots, k-1; s = 1, \dots, l-1) \tag{4.12}$$

$$\mu^0(t_l) + \frac{d(t_l)}{\sigma^2} J^Q(d_1 d(t_l)) = 0, \quad (4.13)$$

$$\mu^r(t_l) = 0. \quad (r = 1, \dots, k-1) \quad (4.14)$$

Remark 4.1. *There are 4 versions of the action functional. The one above can be called the left-action, right-reduction version since $g(t)$ acts on Q_0 from the left, while ξ^0 is the right-reduced velocity $\xi^0 = TR_{g^{-1}}\dot{g}$. There are the following three other cases.*

(1) *The right-action, right-reduction case with action functional*

$$\begin{aligned} S[g, \xi^0, \dots, \xi^{k-1}, \mu^0, \dots, \mu^{k-1}] := & \int_0^{t_l} \left[\ell(\xi^0, \xi^1, \dots, \xi^{k-1}) + \langle \mu^0, TR_{g^{-1}}\dot{g} - \xi^0 \rangle \right. \\ & \left. + \sum_{r=1}^{k-1} \langle \mu^r, \dot{\xi}^{r-1} - \xi^r \rangle \right] dt + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(g(t_i)^{-1}Q_0, Q_{t_i}). \end{aligned}$$

The variation of the penalty term in this case changes according to

$$\frac{1}{2} \delta d^2(g(t_i)^{-1}Q_0, Q_{t_i}) = -d(t_i) \langle \text{Ad}_{g(t_i)^{-1}}^* J(d_1 d(t_i)), \eta(t_i) \rangle.$$

This means that (4.11) and (4.13) are replaced by

$$\begin{aligned} \mu^0(t_s^-) - \mu^0(t_s^+) - \frac{d(t_s)}{\sigma^2} \text{Ad}_{g(t_s)^{-1}}^* J^Q(d_1 d(t_s)) &= 0, \quad (s = 1, \dots, l-1) \\ \mu^0(t_l) - \frac{d(t_l)}{\sigma^2} \text{Ad}_{g(t_l)^{-1}}^* J^Q(d_1 d(t_l)) &= 0. \end{aligned}$$

(2) *In the right-action, left-reduction case, the action functional is*

$$\begin{aligned} S[G, \Xi^0, \dots, \Xi^{k-1}, m^0, \dots, m^{k-1}] := & \int_0^{t_l} \left[\ell(\Xi^0, \Xi^1, \dots, \Xi^{k-1}) + \langle m^0, TL_{G^{-1}}\dot{G} - \Xi^0 \rangle \right. \\ & \left. + \sum_{r=1}^{k-1} \langle m^r, \dot{\Xi}^{r-1} - \Xi^r \rangle \right] dt + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(G(t_i)^{-1}Q_0, Q_{t_i}), \end{aligned}$$

where we wrote L_G for multiplication by G from the left and TL_G for its differential.

However, this is equivalent to the left-action, right-reduction case by identifying

$$G = g^{-1}, \quad \Xi^0 = -\xi^0, \dots, \Xi^{k-1} = -\xi^{k-1}, \quad m^0 = -\mu^0, \dots, m^{k-1} = -\mu^{k-1} \quad (4.15)$$

and setting $\ell = l \circ \kappa$, where $\kappa : k\mathfrak{g} \rightarrow k\mathfrak{g}$ is multiplication by -1 .

(3) By the same token the left-action, left-reduction case with action functional

$$S[G, \Xi^0, \dots, \Xi^{k-1}, m^0, \dots, m^{k-1}] := \int_0^{t_l} \left[l(\Xi^0, \Xi^1, \dots, \Xi^{k-1}) + \langle m^0, TL_{G^{-1}}\dot{G} - \Xi^0 \rangle \right. \\ \left. + \sum_{r=1}^{k-1} \langle m^r, \dot{\Xi}^{r-1} - \Xi^r \rangle \right] dt + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(G(t_i)Q_0, Q_{t_i})$$

can be mapped to the right-action, right-reduction case.

In the analysis that follows we will largely restrict ourselves to the left-action, right-reduction case. Anything we say can be transferred to the remaining three cases by applying the modifications listed above.

4.3.2 Euler–Poincaré equations

From (4.6)–(4.10) it follows that on open intervals (t_i, t_{i+1}) , $i = 0, \dots, l-1$,

$$\left(\frac{d}{dt} + \text{ad}_{\xi^0}^* \right) \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^j} = 0, \quad (4.16)$$

which we recognise as k -th order Euler–Poincaré equations (2.27). This is in line with our earlier Remark 2.12. It is interesting to note however that in the present chapter we have not yet said anything about invariant Lagrangians. Let us therefore briefly summarise why, from the viewpoint of invariant Lagrangians, the Euler–Poincaré equations must appear in the context of (4.2). We refer to Section 2.2.1 for definitions concerning higher-order tangent bundles and recall that a k -th order Lagrangian $L : T^{(k)}G \rightarrow \mathbb{R}$ is said to be *right-invariant* if $L([h]_{h(0)}^{(k)}) = L([h h(0)^{-1}]_e^{(k)})$, for all elements $[h]_{h(0)}^{(k)}$ of the k -th order tangent bundle $T^{(k)}G$. Let L be a right-invariant Lagrangian and consider Hamilton’s principle (Section 2.2.2), $\delta \mathcal{J} = 0$, for

$$\mathcal{J}[g] = \int_a^b L(g(t), \dot{g}(t), \dots, g^{(k)}(t)) dt, \quad (4.17)$$

where variations are taken with respect to fixed end points up to order $k-1$, that is, $\delta g^{(j)}(a) = \delta g^{(j)}(b) = 0$, for $j = 0, \dots, k-1$, in any local chart. As we derived in Section 2.3, the Euler–Lagrange equations can be written in terms of the right-reduced velocity vector $\xi = TR_{g^{-1}}\dot{g}$, which leads to the k -th order Euler–Poincaré equations

$$\left(\frac{d}{dt} + \text{ad}_{\xi}^* \right) \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}} = 0. \quad (4.18)$$

Here we denoted by ℓ the *reduced Lagrangian* $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$ given by $\ell := L|_{T_e^{(k)}G} \circ \alpha_k^{-1}$, with the reduction map α_k as defined in (2.19).

It is now straightforward to see from the viewpoint of invariant Lagrangians that the Euler–Poincaré equation (4.16) must characterise optimal curves in the trajectory planning problem on open time intervals. We can *un-reduce* (4.2) to write

$$S[g] = \int_0^{t_l} L(g, \dot{g}, \dots, \dot{g}^{(k)}) dt + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(g(t_i)Q_0, Q_{t_i}), \quad (4.19)$$

where we defined $L := \ell \circ \alpha_k$. If the curve $g(t)$ is a stationary point of (4.19), then $\xi = TR_{g^{-1}}\dot{g}$ is a stationary point of (4.2). Conversely, if ξ is a stationary point of (4.2), then the curve $g(t)$ defined by $g(0) = e$ and $\dot{g} = TR_g\xi$ is a stationary point of (4.19). But for $g(t)$ to be a stationary point of (4.19), $g|_{[t_i, t_{i+1}]}$, $0 \leq i < l$, must be a solution to Hamilton’s principle $\delta \int_{t_i}^{t_{i+1}} L(g, \dot{g}, \dots, \dot{g}^{(k)}) dt = 0$ with fixed end-points $g(t_i)$ and $g(t_{i+1})$. Correspondingly, since by construction L is right-invariant, $\xi|_{[t_i, t_{i+1}]}$ must satisfy the k -th order Euler–Poincaré equations. It is important to note that the second term of (4.2) does not enter our considerations here, since we kept $g(t)$ fixed at the boundaries of the open intervals. In other words, in order to calculate the equations governing curve evolutions on open intervals, one can ignore the mismatch term in (4.2).

4.3.3 Geometry of multipliers

We observe from (4.6)–(4.10) that on open intervals (t_i, t_{i+1}) , $i = 0, \dots, l-1$,

$$\mu^r = \sum_{j=0}^{k-r-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{r+j}} \quad (r = 0, \dots, k-1). \quad (4.20)$$

For $k = 2$, for example, we obtain

$$\mu^0 = \frac{\delta \ell}{\delta \xi^0} - \frac{d}{dt} \frac{\delta \ell}{\delta \xi^1}, \quad \mu^1 = \frac{\delta \ell}{\delta \xi^1}. \quad (4.21)$$

We will now discuss the geometric meaning of the identities (4.20). As argued in the previous paragraph, we can ignore the second term of (4.2). That is, equations (4.6)–(4.10) and therefore (4.20) are obtained by taking suitably constrained variations of

$$\mathcal{J}[g, \xi^0, \dots, \xi^{k-1}, \mu^0, \dots, \mu^{k-1}] := \int_a^b \left[\ell(\xi^0, \xi^1, \dots, \xi^{k-1}) + \langle \mu^0, TR_{g^{-1}}\dot{g} - \xi^0 \rangle \right]$$

$$+ \sum_{r=1}^{k-1} \langle \mu^r, \dot{\xi}^{r-1} - \xi^r \rangle \Big] dt \quad (4.22)$$

This variational principle is a higher-order generalisation of the *reduced Hamilton–Pontryagin principle* of first-order mechanics. In first-order mechanics, this principle provides a unified treatment of the Lagrangian and Hamiltonian descriptions of invariant mechanical systems on Lie groups (see [72] for a detailed discussion). In particular, the *Legendre transform* connecting the two descriptions is revealed by the variational calculus.

This remains true for higher-order mechanics. Indeed, (4.20) can be recognised to be the reduced Legendre transform that appears in [39, 50]. While we found (4.20) from a variational approach, these references take as starting point [38], where a coordinate free description of the higher-order Legendre transform on manifolds was given. We briefly review this approach here.

The Legendre transform of higher-order mechanics, given in [38], is a map $\text{Leg} : T^{(2k-1)}G \rightarrow T^*(T^{(k-1)}G)$. If Leg is a diffeomorphism (that is, L is *hyperregular*), it connects the Lagrangian and Hamiltonian descriptions just as in the first-order case. With respect to the right-trivialisations

$$T^{(2k-1)}G \cong G \times (2k-2)\mathfrak{g}, \quad T^*(T^{(k-1)}G) \cong G \times (k-2)\mathfrak{g} \times (k-1)\mathfrak{g}^* \quad (4.23)$$

it is given as [50]

$$\begin{aligned} \text{Leg}: (g, \xi^0, \dots, \xi^{2k-2}) &\mapsto (g, \xi^0, \dots, \xi^{k-2}, \mu^0, \dots, \mu^{k-1}), \\ \text{where } \mu^r &= \sum_{j=0}^{k-r-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{r+j}} \quad (r = 0, \dots, k-1). \end{aligned}$$

The same equations were seen in (4.20) to emerge from the Hamilton–Pontryagin principle (4.22). This means that, as for first order, the higher-order Hamilton–Pontryagin principle contains both Lagrangian and Hamiltonian descriptions of higher-order mechanics and provides a unified framework for both views.

More precisely, to obtain the Lagrangian description one may eliminate $\mu^0, \dots, \mu^{k-1}$ from (4.6)–(4.10) using (4.20). The resulting equations are the trivialised flow equations of the *Lagrangian vector field*, which is an element of $\mathfrak{X}(T^{(2k-1)}G)$ [38].

On the other hand, if (4.10) can be solved for ξ^{k-1} (this is the case, for example, when L is hyperregular) then (4.6)–(4.9) are the trivialised flow equations of the *Hamiltonian vector field* $X_H \in \mathfrak{X}(T^*(T^{(k-1)}G))$, which solves [38]

$$i_{X_H}\omega = dH, \quad (4.24)$$

where ω is the canonical symplectic form on $T^*(T^{(k-1)}G)$ and $H : T^*(T^{(k-1)}G) \rightarrow \mathbb{R}$ is given as

$$H(g, \xi^0, \dots, \xi^{k-2}, \mu^0, \dots, \mu^{k-1}) = \sum_{r=0}^{k-1} \langle \mu^r, \xi^r \rangle - \ell(\xi^0, \dots, \xi^{k-1}) \quad (4.25)$$

with respect to the trivialisation (4.23).

By consequence of (4.24) the flow map $F_t : T^*(T^{(k-1)}G) \rightarrow T^*(T^{(k-1)}G)$ of the Hamiltonian vector field preserves the symplectic form ω . For later reference we point out how this can be seen alternatively from the Hamilton–Pontryagin principle. This is a generalisation to higher order of a standard argument (see for example [71, Section 3] for the first-order case). If we omit end point constraints on the variations of $\mathcal{J}$ in (4.22), the integration by parts contributes boundary terms to $\delta\mathcal{J}$ (cf. (4.5)),

$$\delta\mathcal{J} = \int_a^b \dots dt + \langle \mu^0, TR_{g^{-1}}\delta g \rangle \Big|_a^b + \sum_{r=0}^{k-2} \langle \mu^{r+1}, \delta\xi^r \rangle \Big|_a^b = \int_a^b \dots dt + \theta(\delta x) \Big|_a^b, \quad (4.26)$$

where θ in the second equality is the canonical one-form on $T^*(T^{(k-1)}G)$ and we defined $\delta x(t)$ to be the curve in $TT^*(T^{(k-1)}G)$ whose trivialisation corresponds to the variations $(TR_{g^{-1}}\delta g, \delta\xi^0, \dots, \delta\xi^{(k-2)}, \delta\mu^0, \dots, \delta\mu^{(k-1)})$. If we restrict the variations to solution curves of (4.6)–(4.10), we may just as well express $\mathcal{J}$ as a function of initial conditions $\mathcal{J}_{\text{initial}} : T^*(T^{(k-1)}G) \rightarrow \mathbb{R}$. The integral part of (4.26) then vanishes and we obtain

$$\delta\mathcal{J} = d\mathcal{J}_{\text{initial}}(\delta x(a)) = (F_{b-a}^*\theta - \theta)(\delta x(a)).$$

Therefore, $d\mathcal{J}_{\text{initial}} = F_{b-a}^*\theta - \theta$. Taking into account that $d^2 = 0$, we obtain the desired identity $F_{b-a}^*\omega - \omega = d^2\mathcal{J}_{\text{initial}} = 0$.

4.3.4 Momentum conservation and Noether's theorem

We alluded to a conservation law in Remark (1) of Section 2.6 promising a more detailed discussion later in the thesis. Let us make good on that promise now. We notice from (4.6)–(4.10) that $\text{Ad}_g^* \mu^0$ is a conserved quantity on open time intervals. Indeed by (4.6)

$$\frac{d}{dt} \text{Ad}_g^* \mu^0 = \text{Ad}_g^* (\dot{\mu}^0 + \text{ad}_{\xi^0}^* \mu^0) = 0. \quad (4.27)$$

In the context of first-order Euler–Poincaré equations a similar momentum conservation is due to the invariance of the Lagrangian with respect to group multiplication operations. This is an instance of Noether's theorem, which roughly speaking guarantees that the momentum map associated with the action of a symmetry group is preserved (see for example [5, Chap. 11]). We now show that the situation is similar for (4.27).

The right action R of G on itself,

$$R : G \times G, \quad (h, g) \mapsto R_g(h) = hg,$$

can be lifted to an action on $T^{(k-1)}G$,

$$T^{(k-1)}R : T^{(k-1)}G \times G \rightarrow T^{(k-1)}G, \quad \left([h]_{h(0)}^{(k-1)}, g\right) \mapsto T^{(k-1)}R_g \left([h]_{h(0)}^{(k-1)}\right) = [hg]_{h(0)g}^{(k-1)}.$$

This action can subsequently be lifted to its cotangent lifted action ([5, Chap. 12.1])

$$T^*T^{(k-1)}R : T^*(T^{(k-1)}G) \times G \rightarrow T^*(T^{(k-1)}G),$$

given in trivialised form as

$$T^*T^{(k-1)}R_g(h, \xi^0, \dots, \xi^{k-2}, \mu^0, \dots, \mu^{k-1}) = (hg, \xi^0, \dots, \xi^{k-2}, \mu^0, \dots, \mu^{k-1}).$$

It is apparent that the Hamiltonian (4.25) is symmetric with respect to this group action. By Noether's theorem the associated momentum map is conserved.

What is this momentum map? By appealing to standard formulas¹⁵ ([5, Chap. 12.1]) we find that the momentum map $J : T^*(T^{(k-1)}G) \rightarrow \mathfrak{g}^*$ is

$$J(g, \xi^0, \dots, \xi^{k-2}, \mu^0, \dots, \mu^{k-1}) = \text{Ad}_g^* \mu^0,$$

¹⁵In essence one uses the standard formula (4.4) for cotangent lift momentum maps, however replacing Q by $T^{(k-1)}G$ and taking into account that we are dealing with a right action here.

with respect to the trivialisation (4.23). The conservation law observed in (4.27) can therefore be written as $\dot{J} = 0$. It thus arises from the right-invariance of the Lagrangian (respectively, the Hamiltonian) via Noether's theorem.

The conservation law can also be obtained from a variational perspective. This is well known in first-order mechanics, and it is also the case in higher-order mechanics. We take a solution of (4.6)–(4.10) on the time interval $[a, b]$ and vary it according to $\delta g = TL_g \nu$ for $\nu \in \mathfrak{g}$. For $\mathcal{J}$ as in (4.22) we have (cf. (4.26))

$$\delta \mathcal{J} = 0 = \langle \mu^0, TR_{g^{-1}} TL_g \nu \rangle \Big|_a^b = \langle \text{Ad}_g^* \mu^0, \nu \rangle \Big|_a^b. \quad (4.28)$$

The same argument holds after replacing the upper boundary b by any $b' \in [a, b]$. Since ν was arbitrary we conclude that $\text{Ad}_g^* \mu^0$ is conserved along a solution of (4.6)–(4.10).

4.3.5 Node equations

The remarks above concerned equations (4.6)–(4.10) on the open time intervals between nodes. We now come to the node equations (4.11)–(4.14). They specify the evolution across node times of the Lagrange multipliers μ^r , which we interpreted above as the reduced Legendre momenta of the system.

More specifically, the momenta μ^r , $r = 1, \dots, k-1$ are continuous on $[0, t_l]$, while the 0-th momentum μ^0 experiences jump discontinuities at the nodes. If the Lagrangian ℓ is hyperregular we can conclude that $g \in C^{2k-2}([0, t_l])$, that is, g is $(2k-2)$ times continuously differentiable on $[0, t_l]$. Furthermore, the node equations specify terminal values for the curves μ^r , $r = 1, \dots, k-1$.

For $a, b \in \mathbb{R}$ define $1_{a \leq b}$ to be equal to 1 if $a \leq b$ and 0 otherwise. We can now prove the following theorem, which generalises Theorem 2.11 (I).

Theorem 4.2. *For t in any of the open time intervals (t_s, t_{s+1}) as well as for $t \in \{0, t_l\}$,*

$$\mu^0(t) = -\frac{1}{\sigma^2} \text{Ad}_{g(t)^{-1}}^* \left(\sum_{s=1}^l 1_{t \leq t_s} d(t_s) \text{Ad}_{g_{t_s}}^* J^Q(d_1 d(t_s)) \right). \quad (4.29)$$

Proof. At final time $t = t_l$ (4.29) clearly holds because of (4.13). Since $\text{Ad}_g^* \mu^0$ is conserved on open intervals it follows that for $t \in (t_s, t_{s+1})$,

$$\mu^0(t) = \text{Ad}_{g(t)^{-1}}^* \text{Ad}_{g(t_{s+1})}^* \mu^0(t_{s+1}^-).$$

We can now obtain (4.29) by induction over the open time intervals, noting that at each node $t = t_s$ a term $-\frac{d(t_s)}{\sigma^2} J^Q(d_1 d(t_s))$ gets added on. $\square$

In order to formulate the following corollary we recall from Section 3.2.2 the notation $\mathfrak{g}_q$ for any point $q \in Q$, which denotes the Lie algebra of the isotropy subgroup G_q of that point, $G_q := \{g \in G \mid gq = q\}$.

Corollary 4.3. *For a solution of (4.6)–(4.14) we have, for t in any of the open time intervals (t_s, t_{s+1}) as well as for $t \in \{0, t_l\}$,*

$$\langle \mu^0(t), \rho \rangle = 0 \quad \text{for all } \rho \in \mathfrak{g}_{q(t)}. \quad (4.30)$$

In Section 4.5 we will develop a geometric algorithm that inherits an exact version of this corollary. This implies that the numerical search for the optimal initial value of μ^0 can be restricted to the subspace of $\mathfrak{g}^*$ that annihilates $\mathfrak{g}_{Q_0}$.

Proof. For t and ρ as in the statement of the corollary it follows from Theorem 4.2 that

$$\begin{aligned} \langle \mu^0(t), \rho \rangle &= -\frac{1}{\sigma^2} \sum_{s=1}^l 1_{t \leq t_s} \langle d(t_s) J^Q(d_1 d(t_s)), \text{Ad}_{g_{t_s}} \text{Ad}_{g(t)^{-1}} \rho \rangle \\ &= -\frac{1}{\sigma^2} \sum_{s=1}^l 1_{t \leq t_s} \left\langle d(t_s) d_1 d(t_s), (\text{Ad}_{g_{t_s}} \text{Ad}_{g(t)^{-1}} \rho)_Q(q(t_s)) \right\rangle = 0, \end{aligned}$$

where we used (4.4) for the second equality and noted that $\text{Ad}_{g_{t_s}} \text{Ad}_{g(t)^{-1}} \rho \in \mathfrak{g}_{q(t_s)}$ for the third. $\square$

4.3.6 Residual conservation law after partial symmetry breaking

A physically intuitive perspective on Corollary 4.3 is to understand it as a residual conservation law after partial symmetry breaking. We can see the sum in (4.3) as the integral over a time-dependent potential function $V : [0, t_l] \times G \rightarrow \mathbb{R}$ given by

$$V(t, g) = \frac{1}{2\sigma^2} \sum_{i=1}^l \delta(t - t_i) d^2(g(t) Q_0, Q_{t_i}), \quad (4.31)$$

where δ denotes the Dirac delta function. This potential produces instantaneous singular forces at node times t_i , which impart the jump discontinuities on the otherwise conserved momentum $J = \text{Ad}_g^* \mu^0$,

$$J(t_s^+) = J(t_s^-) + \frac{d(t_s)}{\sigma^2} \text{Ad}_{g(t_s)}^* J^Q(d_1 d(t_s)). \quad (4.32)$$

This is because the presence of this potential breaks the G -invariance of the variational problem, however a residual symmetry remains. Clearly, multiplication of g from the right by an element $h \in G_{Q_0}$ leaves V invariant. An adaptation of the argument surrounding equation (4.28), restricting ν to the subspace $\mathfrak{g}_{Q_0} \subset \mathfrak{g}$, then leads to

$$0 = \langle \text{Ad}_g^* \mu^0, \nu \rangle \Big|_0^t,$$

for any $t \in [0, t_l]$. But (4.13) guarantees that $\langle \text{Ad}_{g(t_l)}^* \mu^0(t_l), \nu \rangle = 0$. Therefore,

$$\langle \text{Ad}_{g(t)}^* \mu^0(t), \nu \rangle = 0$$

for any $t \in [0, t_l]$ and any $\nu \in \mathfrak{g}_{Q_0}$, which is equivalent to Corollary 4.3.

4.4 Applications

In this section and in the next chapter we discuss a number of examples that summon the inexact trajectory planning problem. We first briefly review why Riemannian cubics are particularly important types of curves in physical problems. We then treat the rigid body and molecular strands, leaving the discussion of finite-dimensional quantum systems for Chapter 5.

4.4.1 Riemannian cubics revisited

A curve $x(t)$ in Euclidean 3-space can be given physical content by understanding it as the trajectory of a classical particle. More specifically, Newton's second law demands that a particle of unit mass under the influence of an external force F moves along a trajectory that satisfies $\ddot{x} = F$. In particular, force-free motion ($F = 0$) is along straight lines at constant speed.

If instead configuration space is given by a Riemannian manifold Q , with kinetic energy represented by $K = \frac{1}{2} \|\dot{q}\|_q^2$, then one can understand $(D_t \dot{q})^\flat$ as a generalised force acting on a physical system whose time evolution is represented by the curve $q(t)$. Hence, force-free motion is along geodesics and the action functional $S = \int_a^b L dt$, with $L(q, \dot{q}, \ddot{q}) = \frac{1}{2} \|D_t \dot{q}\|_q^2$, measures the square of the L^2 -norm of the external force. As we saw in Section 2.2.3, Hamilton's principle leads to Riemannian cubics.

In the following we will discuss a number of physical systems whose configuration spaces are Lie groups and whose equation of motion in the absence of external forces is given by $D_t \dot{g} = 0$ for some right (or left) invariant Riemannian metric. We showed in Section 2.4 that in this context

$$D_t \dot{g} = \left(\dot{\xi} + \text{ad}_\xi^\dagger \xi \right) g, \quad \text{or} \quad D_t \dot{g} = g \left(\dot{\xi} - \text{ad}_\xi^\dagger \xi \right), \quad (4.33)$$

respectively, where $\xi(t) \in \mathfrak{g}$ is the right (or left) reduced velocity vector. We observed that, as a consequence, the Lagrangian $L = \frac{1}{2} \|D_t \dot{g}\|_g^2$ can be reduced to a function $\ell : 2\mathfrak{g} \rightarrow \mathbb{R}$

$$\ell(\xi^0, \xi^1) = \frac{1}{2} \left\| \xi^1 \pm \text{ad}_{\xi^0}^\dagger \xi^0 \right\|_{\mathfrak{g}}^2. \quad (4.34)$$

The Lagrangian (4.34) is a particularly natural choice for ℓ in the inexact trajectory planning problem since optimal curves minimise (in the L^2 sense) the amount of external forcing necessary to achieve them. It is clear from the equations of motion (4.6)–(4.12) that the solution $g(t)$ is a Riemannian cubic on open intervals, and twice continuously differentiable on the whole time interval $[0, t_l]$. As we mentioned in Remark 2.2, such curves are called *Riemannian cubic splines*.

Remark 4.4. *Let us point to a probabilistic interpretation of Riemannian cubics. More details can be found in Appendix A. See also [16], where a closely related idea is discussed in the context of stochastic modeling of biological growth. Let G be a Lie group with right-invariant metric γ and let e_i , $i = 1, \dots, d$, be an orthonormal basis of the d -dimensional Lie algebra $\mathfrak{g}$. Consider a curve $g(t) \in G$, whose right-reduced velocity $\xi = TR_{g^{-1}} \dot{g}$ satisfies the following Ito stochastic differential equation*

$$d\xi = -\text{ad}_\xi^\dagger \xi dt + \sigma_W \sum_{i=1}^d dW^i e_i, \quad (4.35)$$

where W^i , $i = 1, \dots, d$, are independent Brownian motions (see, for example, [73, Chap. 2.2] for a definition) and $\sigma_W \in \mathbb{R}$. Suppose the (noisy) data is given in a vector space V equipped with an inner product, whose norm we denote by $\|\cdot\|_V$. The noise distribution is assumed to be Gaussian, that is, the probability density function has the form $p(Q) \sim \exp(-\frac{1}{2\sigma_n^2}\|Q - \bar{Q}\|_V^2)$, where $\bar{Q}$ is the true state of the system and $\sigma_n \in \mathbb{R}$. Suppose experiments at times t_i , $i = 1, \dots, l$, measuring the trajectory $g(t)Q_0$ have given results q_{t_i} . Then the minimisation of

$$S = \int_0^{t_l} \ell(\xi, \dot{\xi}) dt + \frac{\sigma_W^2}{2\sigma_n^2} \sum_{i=1}^l \|g(t_i)Q_0 - q_{t_i}\|_V^2, \quad (4.36)$$

with ℓ as in (4.34), can formally be understood as the maximisation of the (logarithm of the) probability of the path $g(t)$, given the measurements. Alternative models of stochastic forcing will typically lead to minimisation problems of the same type, but with different choices of ℓ (see also Remark 4.5 below).

4.4.2 Rigid body splines

Let the Lie group G be the set of rigid rotations $SO(3)$, and let Q be the unit sphere $S^2 \subset \mathbb{R}^3$. We work with the conventions of Remark 3.6. Let γ be a left-invariant Riemannian metric on $SO(3)$. This defines an inner product on $\mathfrak{so}(3)$ which can be expressed as

$$\gamma_e(\Omega_1, \Omega_2) = \Omega_1 \cdot \mathbb{I}\Omega_2$$

for a symmetric, positive definite matrix $\mathbb{I}$. The geodesic equation $D_t \dot{g} = 0$ in Euler–Poincaré form is

$$\dot{\Omega} + \mathbb{I}^{-1}(\Omega \times \mathbb{I}\Omega) = 0, \quad \dot{g} = g\hat{\Omega}. \quad (4.37)$$

Consequently, the Lagrangian (4.34) takes the form

$$\begin{aligned} \ell(\Omega^0, \Omega^1) &= \frac{1}{2}(\Omega^1 + \mathbb{I}^{-1}(\Omega^0 \times \mathbb{I}\Omega^0)) \cdot \mathbb{I}(\Omega^1 + \mathbb{I}^{-1}(\Omega^0 \times \mathbb{I}\Omega^0)) \\ &= \frac{1}{2} \|\Omega^1 + \mathbb{I}^{-1}(\Omega^0 \times \mathbb{I}\Omega^0)\|_{\mathfrak{so}(3)}^2. \end{aligned} \quad (4.38)$$

Consider the inexact trajectory planning problem in the left-action, left-reduction form. That is, suppose an initial point $\mathbf{x}_0$ and targets $\mathbf{x}_{t_1}, \dots, \mathbf{x}_{t_l} \in S^2$ are given, as well as a tolerance parameter σ . We seek the minimiser of

$$\begin{aligned} S[g, \boldsymbol{\Omega}^0, \boldsymbol{\Omega}^1, \boldsymbol{\mu}^0, \boldsymbol{\mu}^1] = & \int_{t_0}^{t_l} \frac{1}{2} \left\| \boldsymbol{\Omega}^1 + \mathbb{I}^{-1}(\boldsymbol{\Omega}^0 \times \mathbb{I}\boldsymbol{\Omega}^0) \right\|_{\mathfrak{so}(3)}^2 + \langle \boldsymbol{\mu}^0, g^{-1}\dot{g} - \boldsymbol{\Omega}^0 \rangle \\ & + \langle \boldsymbol{\mu}^1, \dot{\boldsymbol{\Omega}}^0 - \boldsymbol{\Omega}^1 \rangle + \frac{1}{2\sigma^2} \sum_{i=1}^l \|g(t_i)\mathbf{x}_0 - \mathbf{x}_{t_i}\|^2. \end{aligned} \quad (4.39)$$

The physical interpretation is as follows. The group of rigid rotations, $SO(3)$, is the configuration manifold of a rigid body constrained to rotate around a fixed point. In the absence of external torques the motion is governed by the geodesic equation (4.37), where $\mathbb{I}$ is the moment of inertia tensor (see, for example, [74, Chap. 2.4]). The resulting curve $g(t)$ describes the orientation of the rigid body relative to a space-fixed reference frame.

Suppose the motion of a rigid body (with or without external torque) is partially observed in an experiment. At discrete times t_i the direction of a particular body fixed axis is measured in the space-fixed frame, generating a sequence of outcomes $\mathbf{x}_{t_i} \in S^2$. Therefore, if $\mathbf{x}_0$ is the initial direction of the axis and $g(t)$ describes the rigid body motion, then $g(t_i)\mathbf{x}_0 - \mathbf{x}_{t_i} = 0$, up to measurement error. One would like to model the trajectory $g(t)$, taking into account this information. The action functional (4.39) encodes one such model, yielding the curve $g(t)$ of minimal external torque (in the L^2 sense) that is consistent with the experiment. A natural choice for the parameter σ^2 is to set it equal to the variance of the measurement.

An example simulation can be seen in Figure 4.1, which was generated using the numerical algorithm discussed later (in Section 4.5). What is shown is the curve $g(t)$ that minimises (4.39) for a set of measurements $\mathbf{x}_{t_i}$; these measurements are taken to correspond to the points I_{t_i} defined earlier, in (2.64). In the figure we represent the rotation group in the following way: A given element of $SO(3)$ corresponds to a point on the radial line along the rotation axis, at a distance from the origin equal to the rotation angle. The radius of the inner sphere is π and the radius of the outer sphere is 2π . The centre, marked by a cross, and the boundary of the outer sphere thus both represent the identity matrix.

If the observer is instead moving with a body fixed frame and measuring a space

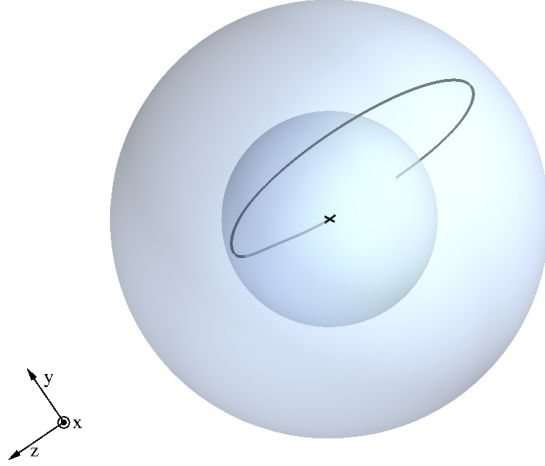


Fig. 4.1: *Riemannian cubic in the group of rigid rotations.* Optimal curve $g(t)$ in $SO(3)$ corresponding to the data points $\mathbf{x}_{t_i} := I_{t_i}$ of (2.64) and generated using the algorithm discussed later, in Section 4.5. The moment of inertia tensor was taken to be the identity matrix and the tolerance parameter was set to $\sigma = 0.025$. For more information on how $SO(3)$ is represented here, see the main text.

fixed direction, then $g(t_i)^{-1}\mathbf{x}_0 - \mathbf{x}_{t_i} = 0$, up to measurement error. In this case the inexact trajectory planning problem presents itself in the right-action, left-reduction form. Evidently, the formalism presented in this chapter applies to any (sufficiently smooth) choice of Lagrangian. For example,

$$\ell(\boldsymbol{\Omega}^0, \boldsymbol{\Omega}^1) = \frac{1}{2} \boldsymbol{\Omega}^0 \cdot (\mathbb{I} \boldsymbol{\Omega}^1 + \boldsymbol{\Omega}^0 \times \mathbb{I} \boldsymbol{\Omega}^0)$$

leads to optimal curves $g(t)$ with minimal work done by external torques.

Remark 4.5. *We mentioned in Remark 4.4 that the minimisation of (4.39) is related to a certain inverse problem given the stochastic evolution (4.35), see also Appendix A. Alternative stochastic models lead to forms of ℓ different from (4.38). For example,*

$$d\boldsymbol{\Omega} = -\mathbb{I}^{-1}(\boldsymbol{\Omega} \times \mathbb{I} \boldsymbol{\Omega})dt + \sigma_W d\mathbf{W},$$

where $d\mathbf{W} = (dW^1, dW^2, dW^3)^T$ is a vector of independent Brownian motions, leads to

$$\ell(\boldsymbol{\Omega}^0, \boldsymbol{\Omega}^1) = \frac{1}{2} \left\| \boldsymbol{\Omega}^1 + \mathbb{I}^{-1}(\boldsymbol{\Omega}^0 \times \mathbb{I} \boldsymbol{\Omega}^0) \right\|^2$$

with $\|\cdot\|$ being the Euclidean norm.

4.4.3 Quantum splines

Quantum splines will be the topic of Chapter 5. Suffice it to say at this stage that the object manifold will be n -dimensional complex projective space, $Q = \mathbb{CP}^n$, acted on by the group $G = SU(n+1)$ of special unitary matrices.

4.4.4 Macromolecular configurations

Equilibrium configurations of macromolecular structures and of DNA in particular can be modelled using the classical theory of elastic rods. See [75, 76] for examples of this approach. In this section we formulate an inexact trajectory planning problem in this context.

We start by describing how the configuration of an elastic rod can be described by a position curve $\mathbf{r}(s)$ in $\mathbb{R}^3$ and a curve $R(s)$ in the group of rigid rotations, $SO(3)$. Here, $s \in [0, 1]$ parametrises the cross sections of the rod along its length, whereby for a given value of s the vector $\mathbf{r}(s)$ points to the centre of mass of the respective cross section, as seen in the lab frame. Let $\mathbf{e}_i(s)$, $i = 1, 2, 3$ be an orthonormal frame such that $\mathbf{e}_1(s)$ and $\mathbf{e}_2(s)$ point along the principal axes of the moment of inertia tensor of the cross section. We will refer to this frame as the *body-fixed* frame. $R(s)$ is the rotation that transforms the initial frame at $s = 0$ (which we assume to coincide with the lab frame) to the body-fixed frame at s . Therefore the configuration of a macromolecule can be described by a curve $g(s) = (R(s), \mathbf{r}(s))$ in the special Euclidean group, $SE(3)$, originating at the identity.

The group multiplication rule of $SE(3)$ is

$$(R_1, \mathbf{r}_1)(R_2, \mathbf{r}_2) = (R_1 R_2, R_1 \mathbf{r}_2 + \mathbf{r}_1).$$

The Lie algebra $\mathfrak{se}(3)$ consists of elements $(\Omega, \mathbf{v})$ with $\Omega \in \mathfrak{so}(3)$ and $\mathbf{v} \in \mathbb{R}^3$. Applying the inverse of the hat map (2.45) to Ω we can represent $\mathfrak{se}(3)$ as $\mathbb{R}^6$. The ad operation becomes

$$\text{ad}_{\xi_1} \xi_2 = \text{ad}_{(\Omega_1, \mathbf{v}_1)}(\Omega_2, \mathbf{v}_2) = (\Omega_1 \times \Omega_2, \Omega_1 \times \mathbf{v}_2 - \Omega_2 \times \mathbf{v}_1).$$

If we identify the dual $\mathfrak{se}(3)^*$ with $\mathbb{R}^6$ using the standard dot product as duality pairing,

then

$$\text{ad}_{(\boldsymbol{\Omega}, \mathbf{v})}^*(\boldsymbol{\mu}, \mathbf{a}) = (-\boldsymbol{\Omega} \times \boldsymbol{\mu} - \mathbf{v} \times \mathbf{a}, -\boldsymbol{\Omega} \times \mathbf{a}). \quad (4.40)$$

The *body-fixed velocity* vector pertaining to a configuration $(R(s), \mathbf{r}(s))$ is defined as

$$\boldsymbol{\xi} = g^{-1}\dot{g} = (R^{-1}\dot{R}, R^{-1}\dot{\mathbf{r}}) = (\boldsymbol{\Omega}, R^{-1}\dot{\mathbf{r}})$$

where a superscript $\cdot$ means derivation with respect to s . In vector notation we can set $\mathbf{v} = R^{-1}\dot{\mathbf{r}}$ so that $\boldsymbol{\xi} = (\boldsymbol{\Omega}, \mathbf{v}) \in \mathbb{R}^6$. In physical terms this encodes the velocity (with respect to “time” s) of a fixed point in space that a *body-fixed observer* would measure, that is, an observer whose coordinate system is the body-fixed frame based at $\mathbf{r}(s)$. Indeed, a space-fixed point $\mathbf{x}_0 = \mathbf{x}(0)$ as seen by a body-fixed observer describes the curve $\mathbf{x}(s) = R^{-1}(s)\mathbf{x}_0 - R^{-1}(s)\mathbf{r}(s)$, whose velocity is $\dot{\mathbf{x}} = -R^{-1}\dot{R}\mathbf{x} + R^{-1}\dot{\mathbf{r}} = -\boldsymbol{\Omega} \times \mathbf{x} + \mathbf{v}$.

A macromolecule that is experimentally constrained to assume a configuration with given final rotation and displacement $g(1) = (R(1), \mathbf{r}(1))$ will relax into an equilibrium state that minimises potential energy with respect to all possible configurations respecting the constraint. For the case of DNA the authors of [76] propose to model this effect using the Lagrangian $L : TSE(3) \rightarrow \mathbb{R}$ whose left-reduced form $l : \mathfrak{se}(3) \rightarrow \mathbb{R}$ is given by

$$l(\boldsymbol{\xi}) = \frac{1}{2}(\boldsymbol{\xi} - \mathbf{z}) \cdot K(\boldsymbol{\xi} - \mathbf{z}).$$

The 6×6 matrix K is symmetric and positive definite and encodes the various stiffness properties. The double helix structure of DNA means that the equilibrium configuration for unconstrained end points retains a number n of rotations along its length. This is expressed by the vector

$$\mathbf{z} = \begin{pmatrix} 2\pi n \mathbf{e}_z \\ \mathbf{e}_z \end{pmatrix}, \quad (4.41)$$

where we assume without loss of generality that the equilibrium configuration for unconstrained end points is oriented along the spatial z -axis and has unit length.

In the case of constrained final rotation and displacement $g(1) = (R(1), \mathbf{r}(1))$ the equilibrium configuration minimises the action functional $S = \int_0^1 l(\boldsymbol{\xi}) ds$. Hamilton’s principle, $\delta S = 0$, leads to Euler–Poincaré equations

$$\dot{\boldsymbol{\xi}} = K^{-1} \text{ad}_{\boldsymbol{\xi}}^* K(\boldsymbol{\xi} - \mathbf{z}) = \text{ad}_{\boldsymbol{\xi}}^\dagger(\boldsymbol{\xi} - \mathbf{z}), \quad (4.42)$$

with ad^* operation given by (4.40) and the operation $\text{ad}^\dagger$ defined by $\text{ad}_{\xi_1}^\dagger \xi_2 := K^{-1} \text{ad}_{\xi_1}^* K \xi_2$.

This equation of motion can be used to design second-order Lagrangians to model non-equilibrium states of the DNA. For example, let us set

$$\begin{aligned} \ell(\xi^0, \xi^1) &= \frac{1}{2} (\xi^1 - \text{ad}_{\xi^0}^\dagger(\xi^0 - \mathbf{z})) \cdot K (\xi^1 - \text{ad}_{\xi^0}^\dagger(\xi^0 - \mathbf{z})) \\ &= \frac{1}{2} \left\| \xi^1 - \text{ad}_{\xi^0}^\dagger(\xi^0 - \mathbf{z}) \right\|_K^2. \end{aligned}$$

Suppose an experiment measures the position of the centre of mass $\mathbf{r}(s_i)$ at a number of parameter values s_i ($i = 1, \dots, l$) as well as a body fixed direction, say $\mathbf{e}_3(s_i)$. The space of measurement outcomes is $Q = S^2 \times \mathbb{R}^3 \subset \mathbb{R}^6$ with $SE(3)$ action given by

$$(R, \mathbf{r})(\mathbf{x}, \mathbf{y}) = (R\mathbf{x}, R\mathbf{y} + \mathbf{r}).$$

If the measurements yield the sequence $(\mathbf{x}_{s_i}, \mathbf{y}_{s_i})$ ($i = 1, \dots, l$) this suggests that up to measurement error the configuration $g(s) \in SE(3)$ satisfies

$$g(s_i)(\mathbf{e}_3(0), \mathbf{0}) = (\mathbf{x}_{s_i}, \mathbf{y}_{s_i}).$$

The task of modelling the configuration $g(s)$ can then be cast in the form of an inexact trajectory planning problem in the left-action, left-reduction form with cost functional

$$\begin{aligned} S[g, \xi^0, \xi^1, \mu^0, \mu^1] &= \int_0^{t_l} \frac{1}{2} \left\| \xi^1 - \text{ad}_{\xi^0}^\dagger(\xi^0 - \mathbf{z}) \right\|_K^2 + \langle \mu^0, g^{-1}\dot{g} - \xi^0 \rangle + \langle \mu^1, \dot{\xi}^0 - \xi^1 \rangle ds \\ &\quad + \frac{1}{2\sigma^2} \sum_{i=1}^l \|g(s_i)(\mathbf{e}_3(0), \mathbf{0}) - (\mathbf{x}_{s_i}, \mathbf{y}_{s_i})\|^2. \end{aligned}$$

Remark 4.6. *Due to the anisotropy in velocity space, expressed by the vector $\mathbf{z}$, the Euler–Poincaré equation (4.42) is not the reduced geodesic equation for the curve $g(t)$ with respect to the metric defined by K . Consequently, the solution to the inexact trajectory matching problem is not a Riemannian cubic spline.*

4.5 Geometric discretisation

The purpose of this section is to illustrate how the higher-order Hamilton–Pontryagin principle offers a direct route towards geometric numerical integrators. All one needs to do, in essence, is to provide a geometric discretisation of the constraint $TR_{g^{-1}}\dot{g} - \xi^0 = 0$ and

define a discrete Hamilton–Pontryagin principle accordingly. For first-order variational problems this idea was introduced in [71], building on the general theory of variational integrators (see [77] for an extensive review). We follow in the footsteps of [71] to treat second-order problems. Third and higher orders can be dealt with in a similar fashion.

Our main motivation for the development of a geometric integrator lies with the exact momentum behaviour exhibited by the discrete solution curves. Not only does the momentum conservation on open intervals (4.27) translate exactly into the discrete picture, but the behaviour at the nodes is given by discrete versions of the continuous time node equations (4.11)–(4.14). As a consequence, one can obtain discrete analogues of Theorem 4.2 and Corollary 4.3. This means that the numerical search for the optimal initial value of the momentum μ^0 can be restricted to a linear subspace of $\mathfrak{g}^*$ of the same dimension as the data manifold Q . As we shall see below, the variational nature of the integrator also means that the discrete flow map preserves the canonical symplectic form on $T^*(TG)$. The implementation of quantum splines in the next chapter will be based on the algorithm presented here.

4.5.1 A geometric integrator

In discrete mechanics the time axis $[t_0, t_l]$ is replaced by a set of discrete time points $t_k = t_0 + kh$, $k = 0, \dots, N$, where h is the step size and $t_l = t_0 + Nh$. We use integers N_i , $i = 1, \dots, l$, as node indices, that is, $t_i = t_0 + N_i h$. For convenience we also define $N_0 := 0$.

We will need a map $\tau : \mathfrak{g} \rightarrow G$ that approximates the Lie exponential and is an analytic diffeomorphism in a neighbourhood of 0 with $\tau(0) = e$ as well as $\tau(\xi)\tau(-\xi) = e$ for all $\xi \in \mathfrak{g}$. An example is the *Cayley transform*, which is a second-order approximation of the Lie exponential in quadratic matrix Lie groups. These include the applications discussed in Section 4.4. The Cayley transform is defined as

$$\tau(\xi) = (e - \xi/2)^{-1}(e + \xi/2).$$

More details on this and other examples can be found in [71].

Since τ is an approximate of the Lie exponential, a simple way of discretising the constraint $TR_{g^{-1}}\dot{g} - \xi^0 = 0$ is to require that $g_{k+1} = \tau(h\xi_k^0)g_k$, where h is the size

of a time step. Similarly one may translate $\dot{\xi}^0 = \xi^1$ to $\xi_{k+1}^0 = \xi_k^0 + h\xi_k^1$. With these considerations in mind we define a discretised version of the action functional (4.3) on discrete path space $\mathcal{C}_d$. Let

$$\mathcal{C}_d = \{g_0, \xi_0^0, \xi_0^1, (g_k, \xi_k^0, \xi_k^1, \check{\mu}_k^0, \mu_k^1)_{k=1}^N\} = (G \times 2\mathfrak{g}) \times (G \times 2\mathfrak{g} \times 2\mathfrak{g}^*)^N,$$

then we define $S_d : \mathcal{C}_d \rightarrow \mathbb{R}$ as

$$\begin{aligned} S_d = h & \left[\sum_{k=0}^{N-1} \ell(\xi_k^0, \xi_k^1) + \left\langle \check{\mu}_{k+1}^0, \frac{1}{h} \tau^{-1}(g_{k+1} g_k^{-1}) - \xi_k^0 \right\rangle + \left\langle \mu_{k+1}^1, \frac{1}{h} (\xi_{k+1}^0 - \xi_k^0) - \xi_k^1 \right\rangle \right] \\ & + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(g_{N_i} Q_0, Q_{t_i}). \end{aligned} \quad (4.43)$$

In analogy to the continuous time problem we assume that $g_0 = e$ and ξ_0^0 are given.

The discrete Euler–Lagrange equations follow from Hamilton’s principle, $\delta S_d = 0$. That is, they characterise paths $\gamma \in \mathcal{C}_d$, for which $\delta S_d := \delta \gamma(S) = 0$ for all variations $\delta \gamma \in T_\gamma \mathcal{C}_d$ with $\delta g_0 = 0$ and $\delta \xi_0^0 = 0$.

In the process of computing δS_d we need to calculate $\delta \tau^{-1}(g_{k+1} g_k^{-1})$. For that purpose it is convenient to introduce the left-trivialised differential of τ at $\xi \in \mathfrak{g}$,

$$D\tau_\xi : \mathfrak{g} \rightarrow \mathfrak{g}, \quad \eta \mapsto \tau(\xi)^{-1} (T_\xi \tau(\eta)),$$

whose inverse we denote by $D\tau_\xi^{-1}$. By taking a derivative of $\tau(\xi)\tau(-\xi) = e$ one can show that [71]

$$D\tau_\xi^{-1} = D\tau_{-\xi}^{-1} \circ \text{Ad}_{\tau(\xi)}. \quad (4.44)$$

Denoting $\eta_k := (\delta g_k) g_k^{-1}$ we find

$$\begin{aligned} \delta \tau^{-1}(g_{k+1} g_k^{-1}) &= T_{\tau(h\xi_k^0)} \tau^{-1}(\eta_{k+1} \tau(h\xi_k^0)) - T_{\tau(h\xi_k^0)} \tau^{-1}(\tau(h\xi_k^0) \eta_k) \\ &= T_{\tau(h\xi_k^0)} \tau^{-1}(\tau(h\xi_k^0) \text{Ad}_{\tau(-h\xi_k^0)} \eta_{k+1}) - T_{\tau(h\xi_k^0)} \tau^{-1}(\tau(h\xi_k^0) \eta_k) \\ &= D\tau_{h\xi_k^0}^{-1}(\text{Ad}_{\tau(-h\xi_k^0)} \eta_{k+1}) - D\tau_{h\xi_k^0}^{-1}(\eta_k) \\ &= D\tau_{-h\xi_k^0}^{-1}(\eta_{k+1}) - D\tau_{h\xi_k^0}^{-1}(\eta_k), \end{aligned}$$

where in the last equality we used (4.44). Introducing the quantities

$$\mu_0^1 := \mu_1^1 + h\check{\mu}_1^0 - h \frac{\delta \ell}{\delta \xi_0^0}, \quad \mu_0^0 := (D\tau_{h\xi_0^0}^{-1})^* \check{\mu}_1^0 \quad (4.45)$$

and

$$\mu_k^0 := (D\tau_{-h\xi_{k-1}^0}^{-1})^* \check{\mu}_k^0 \quad (k = 1, \dots, N), \quad (4.46)$$

we obtain, after rearranging terms,

$$\begin{aligned} \delta S_d = & h \left[\sum_{k=1}^{N-1} \left\langle \frac{\delta \ell}{\delta \xi_k^0} - \check{\mu}_{k+1}^0 + \frac{1}{h} \mu_k^1 - \frac{1}{h} \mu_{k+1}^1, \delta \xi_k^0 \right\rangle + \left\langle \frac{\delta \ell}{\delta \xi_k^1} - \mu_{k+1}^1, \delta \xi_k^1 \right\rangle \right. \\ & + \left\langle \frac{1}{h} \mu_k^0 - \frac{1}{h} (D\tau_{h\xi_k^0}^{-1})^* (D\tau_{-h\xi_k^0})^* \mu_{k+1}^0, \delta \xi_k^0 \right\rangle + \langle \delta \check{\mu}_{k+1}^0, \dots \rangle + \langle \delta \mu_{k+1}^1, \dots \rangle \Big] \\ & + h \left\langle \frac{\delta \ell}{\delta \xi_0^1} - \mu_1^1, \delta \xi_0^1 \right\rangle + h \langle \delta \check{\mu}_1^0, \dots \rangle + h \langle \delta \mu_1^1, \dots \rangle + \langle \mu_N^1, \delta \xi_N^0 \rangle - \langle \mu_0^1, \delta \xi_0^0 \rangle \\ & + \langle \mu_N^0, \eta_N \rangle - \langle \mu_0^0, \eta_0 \rangle + \frac{1}{\sigma^2} \sum_{i=1}^l \langle d_{N_i} J^Q(d_1 d_{N_i}), \eta_{N_i} \rangle, \end{aligned}$$

where we used abbreviations $d_{N_i} = d(g_{N_i} Q_0, Q_{t_i})$ and $d_1 d_{N_i} := d_1 d(g_{N_i} Q_0, Q_{t_i})$. The Euler–Lagrange equations are therefore composed of the following equations. The constraints

$$g_{k+1} = \tau(h\xi_k^0)g_k, \quad \xi_{k+1}^0 = \xi_k^0 + h\xi_k^1 \quad (k = 0, \dots, N-1), \quad (4.47)$$

the discrete equations for the Legendre–Ostrogradsky momenta

$$\mu_{k+1}^1 = \mu_k^1 - h(D\tau_{-h\xi_k^0})^* \mu_{k+1}^0 + h \frac{\delta \ell}{\delta \xi_k^0} \quad (k = 1, \dots, N-1) \quad (4.48)$$

$$\frac{\delta \ell}{\delta \xi_k^1} - \mu_{k+1}^1 = 0 \quad (k = 0, \dots, N-1), \quad (4.49)$$

the discrete version of the Euler–Poincaré equation for interior indices $k \neq N_i$ ($i = 1, \dots, l$)

$$\mu_{k+1}^0 = (D\tau_{-h\xi_k^0}^{-1})^* (D\tau_{h\xi_k^0})^* \mu_k^0 \quad (4.50)$$

and for node indices $k = N_i$ ($i = 1, \dots, l-1$)

$$\mu_{k+1}^0 = (D\tau_{-h\xi_k^0}^{-1})^* (D\tau_{h\xi_k^0})^* \left(\mu_k^0 + \frac{d_k}{\sigma^2} J^Q(d_1 d_k) \right). \quad (4.51)$$

Finally,

$$\mu_N^0 + \frac{d_N}{\sigma^2} J^Q(d_1 d_N) = 0, \quad (4.52)$$

$$\mu_N^1 = 0. \quad (4.53)$$

A solution $\gamma \in \mathcal{C}_d$ to (4.47)–(4.51) is said to have *initial conditions* $(g_0, \xi_0^0, \mu_0^0, \mu_0^1) \in G \times \mathfrak{g} \times 2\mathfrak{g}^* \cong T^*(TG)$, using the definitions (4.45). If the Lagrangian ℓ is hyperregular, then (4.49) can be solved for ξ_k^1 . This means that ξ_k^1 can be eliminated from equations (4.47)–(4.51), which can subsequently be integrated for given initial conditions. If in addition equations (4.52) and (4.53) are satisfied, then γ is a critical point of the action functional S_d .

Remark 4.7. *In Remark 4.1 we mentioned that besides the above left-action, right-reduction case, three other cases were available to be considered. In a similar manner to what we observed in that remark the right-action, right-reduction case introduces changes to equations (4.51) and (4.52). These become*

$$\begin{aligned} \mu_{k+1}^0 &= (D\tau_{-h\xi_k^0}^{-1})^* (D\tau_{h\xi_k^0})^* \left(\mu_k^0 - \frac{d_k}{\sigma^2} \text{Ad}_{g_k^{-1}}^* J^Q(d_1 d_k) \right) \quad \text{and} \\ \mu_N^0 - \frac{d_N}{\sigma^2} \text{Ad}_{g_N^{-1}}^* J^Q(d_1 d_N) &= 0, \end{aligned}$$

respectively. The remaining two cases (left-action, left-reduction; and right-action, left-reduction) are equivalent to the first two in just the same way as explained in Remark 4.1.

4.5.2 Geometric properties

As in Section 4.3.3 the interior equations are conveniently analysed by omitting the mismatch penalty term (4.43) from the action functional, so that

$$\mathcal{J}_d = h \left[\sum_{k=0}^{N-1} \ell(\xi_k^0, \xi_k^1) + \left\langle \check{\mu}_{k+1}^0, \frac{1}{h} \tau^{-1}(g_{k+1} g_k^{-1}) - \xi_k^0 \right\rangle + \left\langle \mu_{k+1}^1, \frac{1}{h} (\xi_{k+1}^0 - \xi_k^0) - \xi_k^1 \right\rangle \right]$$

The arguments that surround equation (4.26) and show symplecticity of the continuous time flow can then be applied in a straightforward manner to the discrete case. Indeed, interior equations (4.47)–(4.50) define a flow map $F_d : G \times \mathfrak{g} \times 2\mathfrak{g}^* \rightarrow G \times \mathfrak{g} \times 2\mathfrak{g}^*$, which integrates a solution γ for given initial conditions. That is, $(F_d)^k(g_0, \xi_0^0, \mu_0^0, \mu_0^1) = (g_k, \xi_k^0, \mu_k^0, \mu_k^1)$ or more succinctly $(F_d)^k(\gamma_0) = \gamma_k$. We restrict $\mathcal{J}_d$ to solutions of (4.47)–(4.50) and express it as a function $\mathcal{J}_{d,\text{initial}} : T^*(TG) \rightarrow \mathbb{R}$ of initial conditions $\gamma_0 \in T^*(TG)$. This means that if $\gamma \in \mathcal{C}_d$ is the solution obtained by integrating γ_0 then

$J_{d,\text{initial}}(\gamma_0) = \mathcal{J}_d(\gamma)$. It follows that

$$d\mathcal{J}_{d,\text{initial}}(\delta\gamma_0) = \langle \mu_N^1, \delta\xi_N^0 \rangle - \langle \mu_0^1, \delta\xi_0^0 \rangle + \langle \mu_N^0, \eta_N \rangle - \langle \mu_0^0, \eta_0 \rangle = (((F_d)^N)^*\theta - \theta)(\delta\gamma_0),$$

where θ is the canonical one-form on $T^*(TG)$. Hence, $d\mathcal{J}_{d,\text{initial}} = ((F_d)^N)^*\theta - \theta$. Taking an exterior derivative shows that the canonical symplectic form $\omega = d\theta$ is preserved by F_d . Hence the discrete flow map F_d is symplectic.

Similarly, the observations given in Section 4.3.4 can be translated to the discrete picture. We pointed out in the paragraph of equation (4.28) how to obtain the conservation of the momentum map $J = \text{Ad}_g^* \mu^0$ from a variational perspective. These arguments can be applied to the discrete variational principle to show that $\text{Ad}_{g_{k+1}}^* \mu_{k+1}^0 = \text{Ad}_{g_k}^* \mu_k^0$ for interior indices $k \neq N_i$. Equivalently, a manipulation of equation (4.50) using (4.44) shows that

$$\mu_{k+1}^0 = (D\tau_{-h\xi_k^0}^{-1})^*(D\tau_{h\xi_k^0})^* \mu_k^0 = \text{Ad}_{\tau(-h\xi_k^0)}^* \mu_k^0 = \text{Ad}_{g_k g_{k+1}^{-1}}^* \mu_k^0 = \text{Ad}_{g_{k+1}^{-1}}^* \text{Ad}_{g_k}^* \mu_k^0,$$

and therefore $\text{Ad}_{g_{k+1}}^* \mu_{k+1}^0 = \text{Ad}_{g_k}^* \mu_k^0$.

Remark 4.8. *One can show that the discrete flow map F_d is in general only accurate to first order in step size h . Indeed, consider the Lagrangian of Riemannian cubics for a bi-invariant metric, $\ell(\xi^0, \xi^1) = \frac{1}{2} \|\xi^1\|_{\mathfrak{g}}^2$. By a Taylor expansion of the continuous-time flow map F_t we obtain, for a solution originating at the identity $e \in G$,*

$$g(h) - g_1 = \frac{h^2}{2} (\mu_0^1)^\sharp + \mathcal{O}(h^3).$$

Here g_1 is the group element reached after the first step and μ_0^1 is the initial value of μ^1 . This is an important limitation of the integrator. In the present context this is acceptable, since we do not need to compute the solution curves for long times; if this were required, the development of more accurate methods would be necessary.

Remark 4.9. *The equations of motion on open intervals can be geometrically discretised in the purely Lagrangian picture by following [78]. One chooses a suitable discrete Lagrangian $L_d : G \times G \times G \rightarrow \mathbb{R}$ and applies variational calculus to the discrete action sum*

$$S_d = \sum_{k=0}^{N-2} L_d(g_k, g_{k+1}, g_{k+2}).$$

Let us define, for example, $\xi : G \times G \rightarrow \mathfrak{g}$ by $\xi(g_1, g_2) = h^{-1}\tau^{-1}(g_2g_1^{-1})$ and set

$$L_d(g_k, g_{k+1}, g_{k+2}) := h\ell(\xi(g_k, g_{k+1}), h^{-1}(\xi(g_{k+1}, g_{k+2}) - \xi(g_k, g_{k+1}))).$$

Boundary conditions being equal, the resulting optimal curve $(g_0, \dots, g_N)$ for this Lagrangian is the same as the one we obtained from the discrete Hamilton–Pontryagin principle. The Hamilton–Pontryagin principle has an advantage in situations where more sophisticated discretisations of the constraint $TR_{g^{-1}}\dot{g} = \xi^0$ are chosen. For example, in the preliminary study [79] Runge–Kutta–Munthe–Kaas methods were used to introduce a class of such integrators. Those integrators can still be understood in the purely Lagrangian framework, however the definition of the corresponding function $L_d(g_k, g_{k+1}, g_{k+2})$ is implicit in that evaluating it requires solving a variational problem. The Hamilton–Pontryagin approach circumvents this difficulty by building the discretisation of the constraint into the variational principle from the outset.

The node equation (4.51) reflects in a geometrically consistent way the jump discontinuities of $\text{Ad}_g^* \mu^0$ given in (4.32). Indeed, (4.51) says that

$$J_{k+1} = J_k + \frac{d_k}{\sigma^2} \text{Ad}_{g_k}^* J^Q(d_1 d_k)$$

when $k = N_i$. Moreover, the final time conditions (4.52) and (4.53) are exact analogues of (4.13) and (4.14). See Figure 4.2 for an example. Putting everything together leads to discrete versions of Theorem 4.2 and Corollary 4.3. The proofs are analogous to the continuous time case.

Theorem 4.10. *For $k = 0, \dots, N$*

$$\mu_k^0 = -\frac{1}{\sigma^2} \text{Ad}_{g_k}^* \left(\sum_{i=1}^l 1_{k \leq N_i} d_{N_i} \text{Ad}_{g_{N_i}}^* J^Q(d_1 d_{N_i}) \right).$$

Corollary 4.11. *For $k = 0, \dots, N$*

$$\langle \mu_k^0, \rho \rangle = 0 \quad \text{for all } \rho \in \mathfrak{g}_{g_k Q_0}.$$

4.6 Practicalities

The integrator derived above provides a way of finding a numerical solution to the inexact trajectory planning problem.

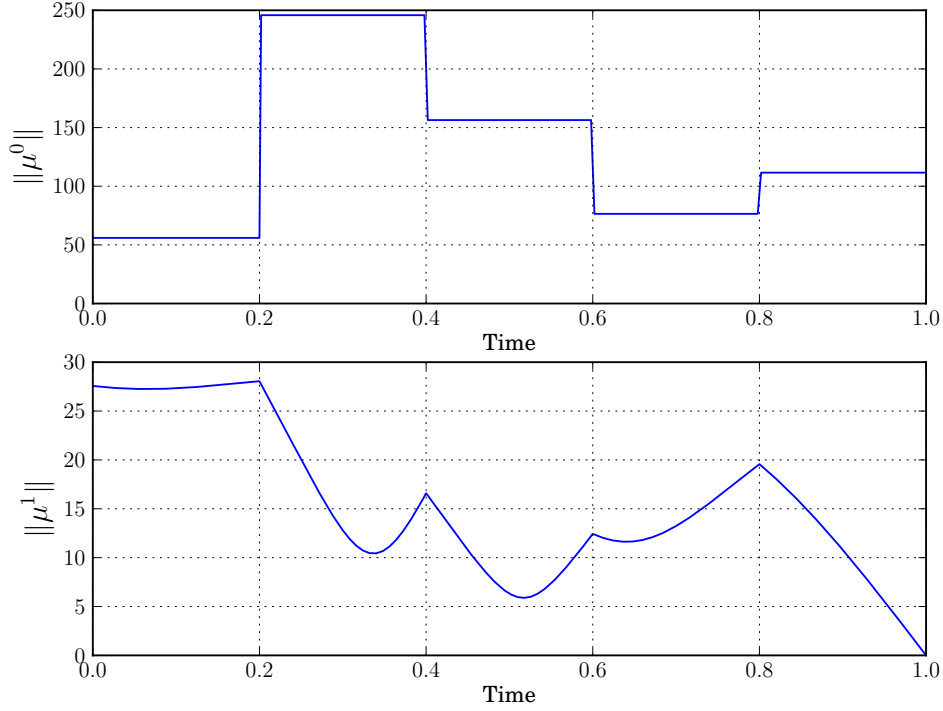


Fig. 4.2: *Momentum norms*. For the interpolating discrete cubic of Figure 4.1, the plot shows the norms of the momenta μ_k^0 and μ_k^1 . The norm of μ_k displays the momentum discontinuities at node indices as well as exact conservation at interior indices, in accordance with (4.50) and (4.51). The norm of μ_k^1 demonstrates continuity, as found in (4.48). Both curves respect terminal conditions (4.52) and (4.53).

The discrete equations of motion (4.47)–(4.51) can be employed to express the action functional $\mathcal{J}_d$ as a function $\mathcal{J}_{d,\text{initial}} : T^*(TG) \rightarrow \mathbb{R}$ of initial conditions $\gamma_0 = (g_0, \xi_0^0, \mu_0^0, \mu_0^1) \in T^*(TG) \cong G \times \mathfrak{g} \times 2\mathfrak{g}^*$. The minimiser in the space of initial conditions can then be determined by a gradient descent method. Since $g_0 = e$ and ξ_0^0 are given, the minimisation is in effect only over the variables (μ_0^0, μ_0^1) . By Corollary 4.11 the optimal μ_0^0 lies in the subspace of $\mathfrak{g}^*$ that annihilates $\mathfrak{g}_{Q_0}$, to which the search can therefore be restricted. On the other hand one still needs to consider all of $\mathfrak{g}^*$ for the optimisation of μ_0^1 .

The gradient of $\mathcal{J}_{d,\text{initial}}$ can be estimated via finite-difference methods. However, this requires the repeated forward integration of (4.47)–(4.51). The number of such integrations increases with the number of dimensions of the Lie group G , and for higher

dimensional systems this quickly becomes unfeasible. Such difficulties can be circumvented using the standard method of *adjoint equations*, which can be easily implemented for the geometric discretisation presented here (a detailed derivation is provided in the appendix). Then the *exact* gradient is obtained at the cost of integrating twice (once forward and once backward) a system of equations of the same complexity as the forward equations.

The significance of the exact preservation of final time constraints (4.13) and (4.14) in the form of (4.52), (4.53) is to provide verification that a (local) minimum has been found. When the tolerance is tight (σ is small) and in the absence of a good initial guess it may occur that the algorithm tends to a local minimum rather than the global one. Suitable initial guesses can be computed using a homotopy strategy. This means a step-by-step reduction of σ , where the optimum at one value of σ is taken as the initial guess at the next smaller value.

4.7 Final remarks

In this chapter we discussed a type of inexact trajectory planning problem whose optimal curves are required to pass near a sequence of fixed target positions at designated times. In Section 4.3 a new derivation of the Euler–Lagrange equations for this type of problem was obtained using a higher–order Hamilton–Pontryagin principle. This approach provided a geometric interpretation of the node equations in terms of Legendre–Ostrogradsky momenta. The highest-order momentum was seen to undergo discontinuous jumps at the node times as a consequence of a partially broken Lie group symmetry. This was the content of Theorem 4.2 and Corollary 4.3. In Section 4.4 some applications of the theory were discussed, which summoned the inexact trajectory planning problem both from a control theoretic viewpoint (quantum splines, see also the next chapter) as well as in the context of a type of inverse problem (rigid body splines, macromolecular configurations). Finally, Section 4.5 was concerned with the numerical approach to solving the problem at hand. The reduced Hamilton–Pontryagin principle was taken as the starting point to obtain a geometric discretisation of the Euler–Lagrange equations, which led to exact momentum behaviour with discrete versions of both Theorem 4.2 and Corollary 4.3. This meant in

particular that the search for the optimal initial value of the highest-order momentum could be restricted to a subspace of the Lie algebra dual of the Lie group whose action describes the motion.

In discussing the applications of Section 4.4, we were mainly interested in a certain type of inverse problem giving rise to the inexact trajectory planning problem. Without giving much detail we also mentioned an application in quantum control. This application will be the topic of the next chapter.

5 Quantum splines

A *quantum spline* is a smooth curve parametrised by time in the space of unitary transformations, whose associated orbit on the space of pure states traverses a designated set of quantum states at designated times, such that the trace norm of the time rate of change of the associated Hamiltonian is minimised. The solution to the quantum spline problem is obtained following the methods presented in the previous chapter. We then apply it in an example that illustrates quantum control of coherent states. We include some numerical simulations that are based on the algorithm presented in Section 4.5. The treatment is self-contained in the sense that the chapter can be read before or after Chapter 4. The content of this chapter has been published in condensed form in [41].

5.1 Some elements of quantum mechanics

Statics. The pure states of quantum systems are represented as points in *quantum state space*, which is a quotient of a complex Hilbert space. For a mathematically minded introduction to quantum theory, see [80]. In this chapter we consider quantum systems with finite dimensional Hilbert space $\mathcal{H} = \mathbb{C}^{n+1}$, for some $n \in \mathbb{N}$, with the standard Hermitian inner product. These are physically realised as systems of quantum mechanical angular momentum. The corresponding state space is n -dimensional complex projective space $\mathbb{CP}^n$, which derives from $\mathbb{C}_*^{n+1}$ upon considering the set of equivalence classes modulo the identification $|\psi\rangle \sim \lambda|\psi\rangle$, for $\lambda \in \mathbb{C}_*$. Here we introduced the notation $|\psi\rangle$ for elements of Hilbert space. We define $\langle\psi|$ to be the Hermitian transpose of $|\psi\rangle$ and write

$$\langle\psi|\phi\rangle := (|\psi\rangle^\dagger)|\phi\rangle = \sum_{i=1}^{n+1} \overline{\psi_i} \phi_i \quad (5.1)$$

for the inner product between vectors $|\psi\rangle$ and $|\phi\rangle$.

Dynamics. Quantum dynamics are encoded by the so-called *Hamiltonian operator* H of the system, which is a Hermitian matrix. The time evolution, in units $\hbar = 1$, is given by the *Schrödinger equation*

$$\frac{d}{dt}|\psi\rangle = -iH|\psi\rangle,$$

where $|\psi(0)\rangle$ is a homogeneous coordinate of the initial state. Equivalently,

$$\dot{U} = -iHU, \quad |\psi(t)\rangle = U(t)|\psi(0)\rangle, \quad (5.2)$$

where $U(t)$ is a curve in the group of unitary matrices whose corresponding orbit is the evolution on quantum state space. One may assume without loss of generality that the anti-Hermitian matrix iH is of zero trace, such that $U(t)$ is special unitary. This is due to the fact that the trace contributes a complex phase factor to the state evolution, which can be neglected in projective terms. If the Hamiltonian operator is constant, then the solution of the Schrödinger equation is given by the Lie exponential,

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle = U(t)|\psi(0)\rangle.$$

5.2 Motivation and problem statement

Controlling the evolution of the unitary transformations that generate quantum dynamics, as described above, is vital in quantum information processing. There is a substantial literature devoted to the investigation of the many aspects of quantum control.¹⁶ The objective of quantum control is the unitary transformation of one quantum state into another one, subject to certain criteria.

For example, one may wish to transform a given quantum state $|\psi\rangle$ into another state $|\phi\rangle$ unitarily in the shortest possible time, with finite energy resource [81, 82, 83]. When only the initial and final states are involved, many time-independent Hamiltonians are available that achieve the unitary evolution $|\psi\rangle \rightarrow |\phi\rangle$, and we simply need to find one that is optimal.

However, transforming a given quantum state $|\psi\rangle$ along a path that traverses through a *sequence* of designated quantum states $|\psi\rangle \rightarrow |\phi_1\rangle \rightarrow |\phi_2\rangle \rightarrow \cdots \rightarrow |\phi_n\rangle$ cannot in general be achieved by a time-independent Hamiltonian. To realise this chain of transformations in the shortest possible time, one chooses the optimal Hamiltonian H_j for each interval $|\phi_j\rangle \rightarrow |\phi_{j+1}\rangle$ [82, 83], and switches the Hamiltonian from H_j to H_{j+1} when the state has

¹⁶At the time of drafting this thesis there are 1000 papers posted on the arXiv that contain the terms “quantum” and “control” in the title.

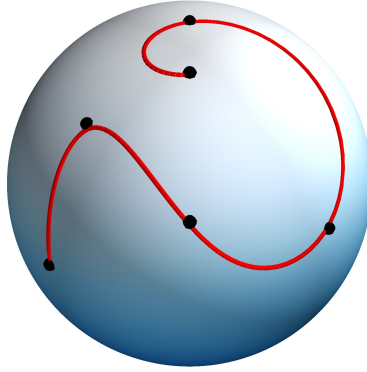


Fig. 5.1: *The orbit of a quantum spline for a two-level system.* The lower-left initial state and the targets are represented by black dots. The variational formulation of the problem requires to minimise a functional that measures both the cost related to the change of the Hamiltonian, and the amount of mismatch between the trajectory and the target points. *This figure appears in [41] – reproduced with kind permission from the American Physical Society.*

reached $|\phi_{j+1}\rangle$. However, instantaneous switching of the Hamiltonian is in general not experimentally feasible.

In this chapter, we consider the following quantum control problem: Let a set of quantum states $|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_m\rangle$ and a set of times $t_1, t_2, \dots, t_m$ be given. Starting from an initial state $|\psi_0\rangle$ at time $t_0 = 0$, find a time-dependent Hamiltonian $H(t)$ such that the evolution path $|\psi_t\rangle$ passes arbitrarily close to $|\phi_j\rangle$ at time $t = t_j$ for all $j = 1, \dots, m$, and such that the change in the Hamiltonian, in a sense defined below, is minimised. The solution to this problem will generate a continuous curve in the space of quantum states that interpolates through the designated states, just as a spline curve interpolates through a given set of data points. However, there is a difference between a classical spline curve and a quantum spline. In the classical context the solution curve passes through a given set of points, whereas in the quantum context, a curve on the space of pure states in itself has no operational meaning. After all, it is the Hamiltonian operator that can be designed in an experiment. Thus, instead of finding a curve in the space of pure states where the designated states lie, we must find a time-dependent curve in the space of Hamiltonians that in turn will generate the optimal curve in the unitary transformation group, the quantum spline. In other words, we shall seek a curve in the associated Lie algebra

$\mathfrak{su}(n+1)$, which of course is equivalent to the space of Hamiltonians, up to multiplication by $i = \sqrt{-1}$.

Since our optimality condition for quantum splines involves the time-derivative of $iH(t)$, we shall make use of the techniques developed in the earlier chapters for the higher-order calculus of variations on Lie groups and their algebras. First, we will derive the Euler–Lagrange equations (5.7) and (5.11) below that solve the quantum spline problem. An example of such a solution for a two-level quantum system is sketched in Fig. 5.1. As an application, we illustrate how the results transform a quantum state along a path that lies entirely on the coherent-state subspace. The discretisation used in the production of the figures in this chapter is described in Section 4.5.

5.3 Quantum spline equations

The optimal curve $H(t)$ that solves the quantum spline problem is the minimiser of a cost functional (action) consisting of two terms: The first term measures the overall change in the Hamiltonian during the evolution. For this purpose we shall consider the trace norm; i.e., for a pair of trace-free skew-Hermitian matrices A and B we define their inner product by

$$\langle A, B \rangle := -2 \operatorname{tr}(AB), \quad (5.3)$$

where the factor 2 is purely conventional. Thus, if H is a time-dependent Hamiltonian and $\dot{H}$ its time derivative, the instantaneous penalty arising from changing the Hamiltonian is given by $\frac{1}{2} \langle i\dot{H}, i\dot{H} \rangle = \operatorname{tr}(\dot{H}^2)$. The second term penalises the mismatch between the state $|\psi_{t_j}\rangle$ at time t_j and the target state $|\phi_j\rangle$. For this purpose we shall use the standard geodesic distance on complex projective space¹⁷

$$D(\psi, \phi) = 2 \arccos \sqrt{\frac{\langle \psi | \phi \rangle \langle \phi | \psi \rangle}{\langle \psi | \psi \rangle \langle \phi | \phi \rangle}} \quad (5.4)$$

¹⁷At this stage we can reveal our motivation for the factor of 2 in (5.3). The inner product (5.3) can be extended to a bi-invariant Riemannian metric on the $SU(n+1)$. The normal metric induced on $\mathbb{CP}^1$ by the action of $SU(n+1)$ then coincides with the so-called *Fubini–Study metric*, whose distance function is (5.4). This geometric relation between the respective metrics is useful when analysing the relationship between the Riemannian cubics on $\mathbb{CP}^1$ and those on $SU(n+1)$, along the lines of Chapter 3. In the present chapter however, such considerations will play no role.

for a pair of states $|\psi\rangle$ and $|\phi\rangle$. Writing $U(t)$ for the parametric family of unitary operators generated by $H(t)$ so that $|\psi_{t_j}\rangle = U(t_j)|\psi_0\rangle$, the mismatch penalty is chosen to be $D^2(U(t_j)\psi_0, \phi_j)/2\sigma^2$, where the tolerance $\sigma > 0$ is a tunable parameter so that the penalty is high when σ is small, and the factor of a half is purely conventional.

The action, of course, must be minimised subject to the constraint that the dynamical evolution of the state is unitary. That is, U must satisfy the Schrödinger equation $\dot{U} = -iHU$, given above in (5.2). Therefore, given an initial state $|\psi_0\rangle$ at time $t_0 = 0$, a set of target states $|\phi_1\rangle, \dots, |\phi_m\rangle$ at times $t_1, \dots, t_m$, and an initial Hamiltonian $H(0) = H_0$, we wish to find the minimiser of

$$\mathcal{J} = \int_{t_0}^{t_m} \left(\frac{1}{2} \langle i\dot{H}, i\dot{H} \rangle + \langle M, \dot{U}U^{-1} + iH \rangle \right) dt + \frac{1}{2\sigma^2} \sum_{j=1}^m D^2(U(t_j)\psi_0, \phi_j), \quad (5.5)$$

where the minimisation is over curves $U(t) \in SU(n+1)$ and $iH(t), M(t) \in \mathfrak{su}(n+1)$. Additionally, we require smoothness of these curves on open intervals (t_j, t_{j+1}) for $j = 0, \dots, m-1$; $U(0) = e$; and the continuity of $U(t)$ and $H(t)$ is assumed everywhere as well as the existence of certain limits (see below). The curve $M(t)$ acts as a Lagrange multiplier enforcing the kinematic constraint.

Before we proceed to vary the action $\mathcal{J}$ let us comment on the choice of the initial Hamiltonian H_0 . We let H_0 be such that the trajectory $e^{-iH_0 t}|\psi_0\rangle$ corresponds to the geodesic curve on the space of pure states joining $|\psi_0\rangle$ and $|\phi_1\rangle$; the construction of such a Hamiltonian can be found in [83]. Intuitively, since the first target time t_1 is fixed, this choice generates the most direct traverse $|\psi_0\rangle \rightarrow |\phi_1\rangle$, hence requiring least change in the Hamiltonian at initial times $t \ll t_1$.

The Euler–Lagrange equations governing stationary points of (5.5) are obtained by taking the variation of $\mathcal{J}$ and requiring $\delta\mathcal{J} = 0$. Writing $A = (\delta U)U^{-1}$ we have

$$\begin{aligned} \delta\mathcal{J} &= \int_{t_0}^{t_m} \left(\langle i\dot{H}, i\delta\dot{H} \rangle + \langle M, \dot{A} - [\dot{U}U^{-1}, A] + i\delta H \rangle + \langle \delta M, \dot{U}U^{-1} + iH \rangle \right) dt \\ &\quad + \frac{1}{2\sigma^2} \sum_{j=1}^m \delta D^2(\psi_{t_j}, \phi_j) \\ &= \int_{t_0}^{t_m} \left(\langle M - i\ddot{H}, i\delta H \rangle + \langle -\dot{M} + [\dot{U}U^{-1}, M], A \rangle + \langle \delta M, \dot{U}U^{-1} + iH \rangle \right) dt \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2\sigma^2} \sum_{j=1}^m \delta D^2(\psi_{t_j}, \phi_j) + \sum_{j=1}^{m-1} \left[\langle \Delta M(t_j), A(t_j) \rangle + \langle i\Delta \dot{H}(t_j), i\delta H(t_j) \rangle \right] \\
& + \langle M(t_m), A(t_m) \rangle + \langle i\dot{H}(t_m), i\delta H(t_m) \rangle,
\end{aligned} \tag{5.6}$$

where in the second step we have integrated by parts, and used the notations $\Delta M(t_j) = M(t_j^-) - M(t_j^+)$ and $\Delta \dot{H}(t_j) = \dot{H}(t_j^-) - \dot{H}(t_j^+)$, with $M(t_i^+) = \lim_{t \downarrow t_i} M(t)$ and $M(t_i^-) = \lim_{t \uparrow t_i} M(t)$; and similarly for $\dot{H}(t_j^\pm)$. It follows from (5.6) that on the open intervals (t_j, t_{j+1}) , $j = 0, \dots, m-1$, the following equations hold:

$$i\ddot{H} - M = 0, \quad \dot{M} + [M, \dot{U}U^{-1}] = 0, \quad \dot{U}U^{-1} + iH = 0. \tag{5.7}$$

Additionally, at the nodes $t = t_j$, we require matching conditions. To work them out, let us calculate the variation $\delta D^2 = 2D\delta D$ appearing in (5.6). From the definition (5.4) and the relation

$$\begin{aligned}
\frac{\langle \psi | e^{-\varepsilon A} | \phi \rangle \langle \phi | e^{\varepsilon A} | \psi \rangle}{\langle \psi | \psi \rangle \langle \phi | \phi \rangle} & \approx \frac{\langle \psi | (1 - \varepsilon A) | \phi \rangle \langle \phi | (1 + \varepsilon A) | \psi \rangle}{\langle \psi | \psi \rangle \langle \phi | \phi \rangle} \\
& = \frac{\langle \psi | \phi \rangle \langle \phi | \psi \rangle}{\langle \psi | \psi \rangle \langle \phi | \phi \rangle} + \frac{2\Re[\langle \psi | \phi \rangle \langle \phi | A | \psi \rangle]}{\langle \phi | \phi \rangle \langle \psi | \psi \rangle} \varepsilon + \mathcal{O}(\varepsilon^2),
\end{aligned} \tag{5.8}$$

which holds for any $A = -A^\dagger$, we find, bearing in mind that if $D(x) = 2 \arccos(\sqrt{x})$ then $dD/dx = -2/\sin(D)$,¹⁸

$$\delta D = \left. \frac{d}{d\varepsilon} D(e^{\varepsilon A} \psi, \phi) \right|_{\varepsilon=0} = \frac{-4\Re[\langle \psi | \phi \rangle \langle \phi | A | \psi \rangle]}{\sin(D) \langle \phi | \phi \rangle \langle \psi | \psi \rangle}. \tag{5.9}$$

From (5.9), and writing $D_j = D(\psi_{t_j}, \phi_j)$, we deduce that $\delta D_j^2 = 2D_j \langle F_j, A(t_j) \rangle$, where

$$F_j = \frac{\langle \psi_{t_j} | \phi_j \rangle | \psi_{t_j} \rangle \langle \phi_j | - \langle \phi_j | \psi_{t_j} \rangle | \phi_j \rangle \langle \psi_{t_j} |}{\sin(D_j) \langle \phi_j | \phi_j \rangle \langle \psi_{t_j} | \psi_{t_j} \rangle}. \tag{5.10}$$

The relevant matching conditions at the nodes are therefore given by:

$$\dot{H}(t_j^+) - \dot{H}(t_j^-) = 0, \quad M(t_j^+) - M(t_j^-) = \frac{D_j F_j}{\sigma^2}, \tag{5.11}$$

whereas we require $\dot{H}(t_m) = 0$ and $M(t_m) + D_m F_m / \sigma^2 = 0$ at the terminal point. Quantum spline problems are therefore solved by finding a solution to equations (5.7) and (5.11) that satisfies, in addition, the terminal conditions at t_m .

¹⁸We have

$$D'(x) = \arccos'(\sqrt{x}) \frac{1}{\sqrt{x}} = -\frac{1}{\sin(\arccos(\sqrt{x}))} \frac{1}{\sqrt{x}} = -\frac{1}{\sin(D/2)} \frac{1}{\sqrt{x}}.$$

Then use the standard formula $\sin(D) = 2\sin(D/2)\cos(D/2)$ to obtain $\sin(D/2) = \sin(D)/2\sqrt{x}$ and therefore $D'(x) = -2/\sin(D)$.

5.4 Geometry

The results of this section are presented in general terms in Section 4.3. Here we focus on the concrete example of quantum splines. On open time intervals (t_i, t_{i+1}) equation (5.7) yields

$$\ddot{H} + i[H, \ddot{H}] = 0. \quad (5.12)$$

Comparing with (2.43), we recognise this as the right-reduced equation for Riemannian cubics on $SU(n+1)$ with respect to the bi-invariant Riemannian metric induced by the inner product (5.3). That is, $U(t)$ is a Riemannian cubic on the open time intervals (t_i, t_{i+1}) . The node conditions (5.11) imply that $U(t)$ is a *Riemannian cubic spline*, a twice continuously differentiable curve that is composed of a series of cubics.

We remark on the important structure of the Lagrange multiplier $M(t)$ implied by the equations of motion that makes it sufficient to consider a subspace of $\mathfrak{su}(n+1)$ when searching for the optimal initial value $M(0)$. Let us denote by $\mathfrak{su}(n+1)_\psi$ the totality of trace-free skew-Hermitian generators of unitary motions that leave the projective image of the state $|\psi\rangle$ invariant, and $\mathfrak{su}(n+1)_\psi^\perp$ its complement with respect to the inner product (5.3). Then, we have the following

Lemma 5.1. $M(t) \in \mathfrak{su}(n+1)_{\psi_t}^\perp$

Proof. For t in any of the open intervals (t_i, t_{i+1}) it follows from the second equation of (5.7) that

$$M(t) = \text{Ad}_{U(t)U(t_{i+1})^{-1}} M(t_{i+1}^-).$$

Moreover, it is easy to check that

$$\mathfrak{su}(n+1)_{\psi_t}^\perp = \text{Ad}_{U(t)U(t_{i+1})^{-1}} \mathfrak{su}(n+1)_{\psi_{t_{i+1}}}^\perp.$$

Hence, if $M(t_{i+1}^-)$ lies in $\mathfrak{su}(n+1)_{\psi_{t_{i+1}}}^\perp$, then $M(t_i^+)$ lies in $\mathfrak{su}(n+1)_{\psi_{t_i}}^\perp$. $M(t)$ at final time t_m is defined by the terminal condition $M(t_m) = -D_m F_m / \sigma^2$, and it has discontinuities at nodal times t_j given by $D_j F_j / \sigma^2$, see (5.11). The proof is completed by noting that for any $j = 1, \dots, m$, the quantity F_j defined in (5.10) lies in $\mathfrak{su}(n+1)_{\psi_{t_j}}^\perp$. $\square$

This result is significant, because the search for the optimal $M(0)$ can be restricted to the $2n$ -dimensional subspace $\mathfrak{su}(n+1)_{\psi_0}^\perp$ of the $n(n+2)$ -dimensional Lie algebra $\mathfrak{su}(n+1)$.

Remark 5.2 (Momentum map). *It is interesting to compare (5.11) with the corresponding equation (4.11) in the more general setting of Chapter 4. We conclude in particular that*

$$[J^{\mathbb{CP}^n}(\mathrm{d}_1 D(\psi, \phi))]^\# = \frac{\langle \psi | \phi \rangle | \psi \rangle \langle \phi | - \langle \phi | \psi \rangle | \phi \rangle \langle \psi |}{\sin(D(\psi, \phi)) \langle \phi | \phi \rangle \langle \psi | \psi \rangle},$$

where $J^{\mathbb{CP}^n}$ is the cotangent lift momentum map for the left action of the special unitary group on complex projective space.

5.5 Two-level systems

Consider a two-level system ($n = 1$). We can think of this system as a spin- $\frac{1}{2}$ particle immersed in a magnetic field. If $\mathbf{n}(t)$ is the unit direction of the field at time t , the Hamiltonian of the system can be written in the form $H(t) = \omega(t)\boldsymbol{\sigma} \cdot \mathbf{n}(t)$, where $\omega(t)$ is the field strength and we recall the Pauli spin matrices $\sigma_x, \sigma_y, \sigma_z$,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In order to represent the evolution of quantum states we identify $\mathbb{CP}^1$ with the Bloch sphere $S^2 \subset \mathbb{R}^3$ through the diffeomorphism

$$\kappa : \mathbb{CP}^1 \rightarrow S^2, \quad \left[\begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix} \right] \mapsto \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}.$$

In words, for any given element of $\mathbb{CP}^1$ one chooses a unit length representative whose first entry is on the real axis, positive or negative. Then one chooses any pair (ϕ, θ) corresponding to the representative and computes the element in S^2 according to the formula.

Following the methods described in Section 4.5, we have implemented the optimisation for a set of target states on S^2 , an initial state $|\psi_0\rangle$, and a set of times. Using the resulting Hamiltonian we have generated the dynamics of the state, as illustrated in Fig. 5.1. In

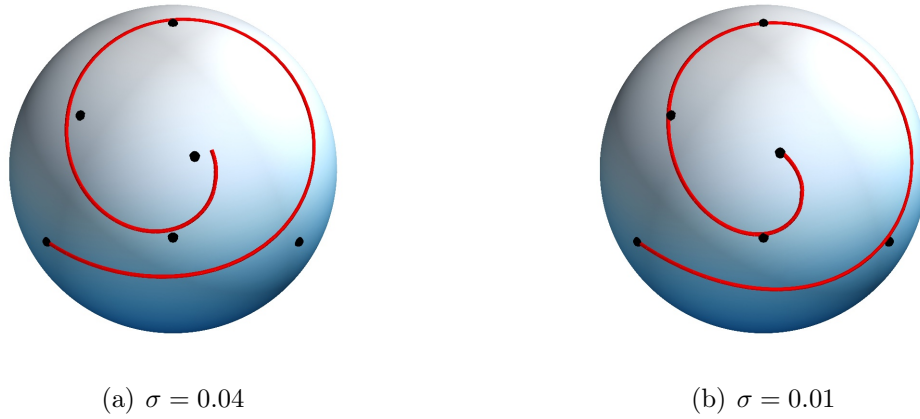


Fig. 5.2: *Orbits on state space generated by the solution to the quantum spline problem.* The black dots indicate the initial (lower left) and the target points. The optimal trajectories are shown for two different values of the tolerance parameter: $\sigma = 0.04$ and $\sigma = 0.01$. Lower values of the tolerance parameter translate, through the cost functional $\mathcal{J}$, into a stronger penalty on the mismatch. *This figure appears in [41] – reproduced with kind permission from the American Physical Society.*

Fig. 5.2 we have sketched the effect of choosing different tolerance levels. When the value of σ is reduced, the resulting orbit $|\psi_t\rangle$ traverses closer to the vicinities of the target states $\{|\phi_j\rangle\}$. From the action $\mathcal{J}$ in (5.5), one sees that this may be realised at the expense of varying the Hamiltonian $H(t)$ more rapidly. This effect can be visualised in the case of a two-level system, since the Hamiltonian is characterised by the unit vector $\mathbf{n}(t) \in \mathbb{R}^3$ and the scalar field strength $\omega(t)$. Thus in Fig. 5.3 we have plotted the end-point of the unit vector $\mathbf{n}(t)$ on a sphere, and the values of $\omega(t)$, for different choices of σ . These plots show that both $\mathbf{n}(t)$ and $\omega(t)$ vary more rapidly at smaller tolerance level (i.e. smaller σ).

5.6 Coherent states

Another example we consider here is a controlled motion of a quantum state on the coherent-state subspace of the state space. Consider $SU(n+1)$ coherent states [84, 85] in arbitrary dimensions. In the context of quantum information theory, these states correspond to totally disentangled states inside the symmetric subspace of the Hilbert space of the combined system. They can be generated by taking symmetric tensor products of

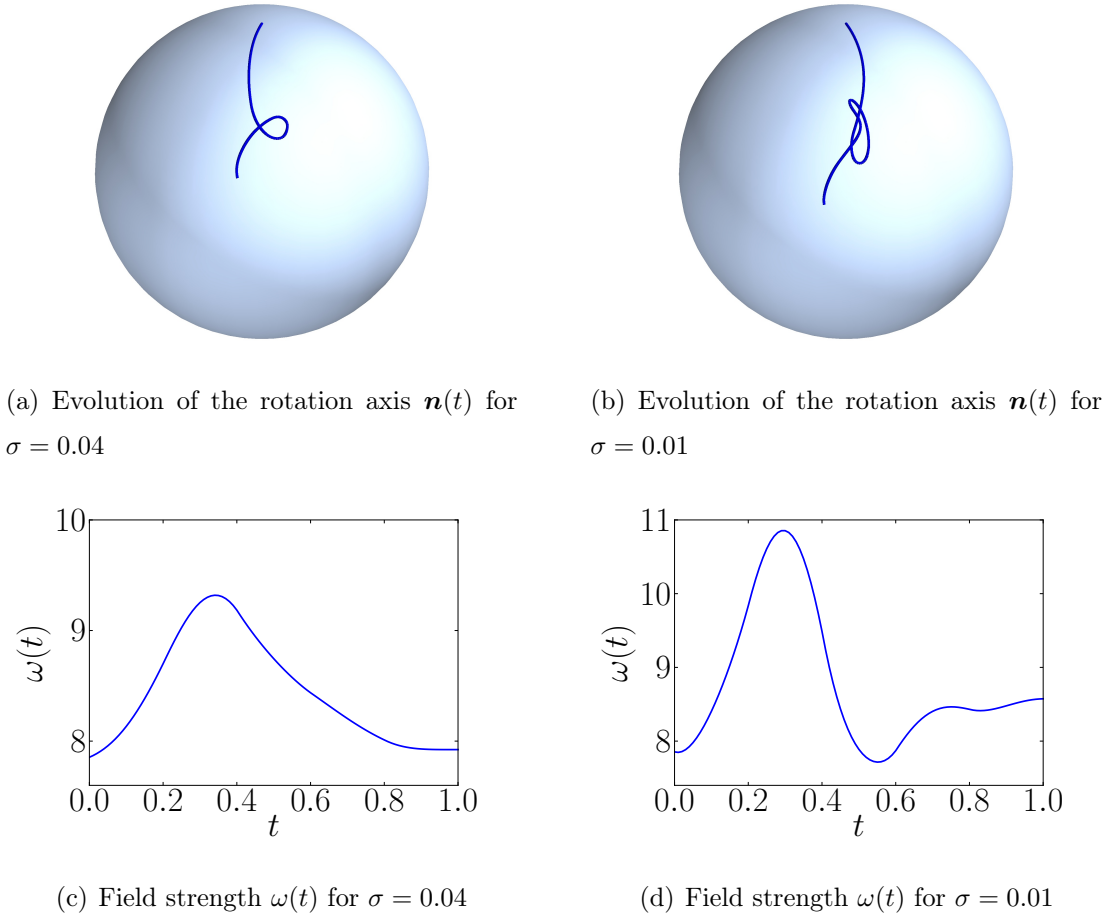


Fig. 5.3: *The quantum spline Hamiltonian $H(t)$.* Hamiltonians that generate the dynamical trajectories in Fig. 5.2. The top row shows the orbits of the endpoint of the rotation axis $\mathbf{n}(t)$. The bottom row shows the field strength $\omega(t)$. These images illustrate the fact that as the value of σ is decreased, the amount of change in the optimal Hamiltonian $H(t)$ increases. *This figure appears in [41] – reproduced with kind permission from the American Physical Society.*

single-particle states. Let us briefly review their construction.

One starts from a $(n + 1)$ -level single-particle quantum system, whose Hilbert space is $\mathcal{H} = \mathbb{C}^{n+1}$ and whose quantum state space is $\mathbb{CP}^n$. The bosonic multi-particle system composed of N such particles is defined on the Hilbert space $\mathcal{H}_N = \bigotimes_{\text{sym}}^N \mathbb{C}^{n+1}$ (symmetrised tensors), whose dimension is

$$d := \binom{n + N}{N}.$$

Consequently, the quantum state space is $\mathbb{CP}^{d-1}$. The special unitary group $SU(n+1)$ has an irreducible unitary representation on $\mathcal{H}_N$ given by the factor-wise linear action on single-particle states (the *product action*)

$$U \text{Sym}(|\psi_1\rangle \otimes \dots \otimes |\psi_N\rangle) = \text{Sym}(U|\psi_1\rangle \otimes \dots \otimes U|\psi_N\rangle), \quad (5.13)$$

where we wrote Sym for the symmetrisation operator. Notice that (5.13) is enough to define the representation on Hilbert space, which subsequently descends to an action on projective space.

The submanifold of $SU(n+1)$ *coherent states* is defined as the group orbit, in $\mathbb{CP}^{d-1}$, of some reference state. A common choice of reference state is the projective image of a highest weight vector $|\gamma\rangle$ of the irreducible representation [86]. For concreteness, let us set $|\gamma\rangle = \otimes^N e_1$, where e_1 is the first element of the standard basis of $\mathbb{C}^{n+1}$. Hence, the coherent state submanifold $\mathcal{C}_{\text{c.s.}}$ is given by

$$\mathcal{C}_{\text{c.s.}} = \{ [\otimes^N U e_1] ; U \in SU(n+1) \} \subset \mathbb{CP}^{d-1},$$

where we wrote square brackets to denote the point in projective space. A dispassionate look at what we have just done reveals that coherent states are made up of N identical copies of single-particle states and that the unitary group acts on the coherent states through its action on the single-particle states. It is therefore quite clear that the formulation of the quantum spline problem in this context is no different from the single-particle case. We now make this statement precise.

There is a natural distance function $D_{\text{c.s.}}$ on the coherent state submanifold that is induced by the standard geodesic distance of the ambient $\mathbb{CP}^{d-1}$. We use this distance function to formulate the quantum spline problem on the coherent state submanifold¹⁹, as follows. Given an initial coherent state $|\psi_0\rangle$ at time $t_0 = 0$, a set of target coherent states $|\phi_1\rangle, \dots, |\phi_m\rangle$ at times $t_1, \dots, t_m$, and an initial Hamiltonian $H(0) = H_0$, we wish to find the minimiser of

$$\mathcal{J}_{\text{c.s.}} = \int_{t_0}^{t_m} \left(\frac{1}{2} \langle \dot{H}, \dot{H} \rangle + \langle M, \dot{U}U^{-1} + iH \rangle \right) dt + \frac{1}{2\sigma^2} \sum_{j=1}^m D_{\text{c.s.}}^2(U(t_j)\psi_0, \phi_j). \quad (5.14)$$

¹⁹Alternatively, one could choose as distance function the standard geodesic distance between the (identical) single-particle states that make up the coherent states. This would remove the scaling factor in (5.15) below.

Let us verify that this is equivalent to the single-particle quantum spline problem on $\mathbb{CP}^n$.

First, we notice that the coherent state submanifold is equal to the image set of the *generalised Veronese embedding*

$$V : \mathbb{CP}^n \hookrightarrow \mathbb{CP}^{d-1}, \quad [|\psi\rangle] \mapsto [\otimes^N |\psi\rangle],$$

which embeds the single-particle state space into the symmetric multi-particle state space. This fact was observed in [87]. Evidently, the Veronese embedding commutes with the action of $SU(n+1)$ on the respective spaces. That is, for any $U \in SU(n+1)$ we have $U \circ V = V \circ U$. Moreover, $D_{\text{c.s.}}$ is related to the standard geodesic distance (5.4) on $\mathbb{CP}^n$ by

$$D_{\text{c.s.}}(\psi, \phi) = \sqrt{N} D(V^{-1}(\psi), V^{-1}(\phi)), \quad (5.15)$$

where we used the fact that the Veronese embedding can be inverted on the coherent state submanifold. Therefore, instead of (5.14) we can equivalently minimise the single-particle cost functional

$$\mathcal{J} = \int_{t_0}^{t_m} \left(\frac{1}{2} \langle i\dot{H}, i\dot{H} \rangle + \langle M, \dot{U}U^{-1} + iH \rangle \right) dt + \frac{N}{2\sigma^2} \sum_{j=1}^m D^2(U(t_j)V^{-1}(\psi_0), V^{-1}(\phi_j)),$$

where we scaled the distance function according to (5.15) and mapped initial and target states to $\mathbb{CP}^n$ using V^{-1} . We can then use the single-particle methods described above to solve the quantum spline problem on the coherent state submanifold.

5.7 Final remarks

In this chapter we discussed an application to quantum control of the methods of Chapter 4. We derived Euler–Lagrange equations satisfied by the optimal time-dependent Hamiltonian operator and treated in detail two applications; control of two-level quantum systems and of $SU(n+1)$ -coherent states. The simulations were carried out by following the methods of Section 4.5. For the numerical treatment of quantum systems with three and more levels ($n > 1$) it is crucial to have an efficient way of computing the gradient of the cost functional. This can be achieved using the method of adjoint equations described in Appendix B.

6 Conclusions and Outlook

This thesis has been concerned with group invariant higher-order variational principles with a focus on their reduction theory, geometric properties and applications. With the dust settled, let us review in some detail what we have done (and also, afterwards, what we have *not* done).

Chapter 2. To begin with we developed the theory of Euler–Poincaré reduction for invariant higher-order variational principles on Lie groups. The idea was the following: If a higher-order Lagrangian, defined on the k^{th} -order tangent bundle of a Lie group G , is invariant with respect to group multiplication (from the right or from the left), one can eliminate the redundant degree of freedom and pass to the corresponding reduced Lagrangian defined on k copies of the Lie algebra. Accordingly, the higher-order Hamilton’s principle (2.5) can be transformed into an equivalent reduced form, which is stated in terms of the reduced Lagrangian and (right or left) trivialised velocities. When taking variations of the reduced action functional one needs to be careful since only a certain type of variations is relevant, namely, those that stem from variations on the Lie group.²⁰ With this in mind we proceeded to take variations and arrived at (2.27). We summarised these results in Theorem 2.6 of Section 2.3.1.

A case of particular interest, Riemannian cubics and their higher-order generalisations, was treated in detail in the new framework. This was the content of Section 2.4. As we had described earlier, in Section 2.2.3, Riemannian cubics are solutions to a second-order variational principle on Riemannian manifolds, whose Lagrangian (2.8) measures the norm squared of the covariant acceleration. A more informal way of saying this is that the Lagrangian measures the extent to which a curve on the Lie group is not a geodesic. It should come as no surprise then that the reduced Lagrangian, found in Proposition 2.8, measures the extent to which the trivialised velocity curve does not satisfy the reduced geodesic equation, which, incidentally, is an example of a first-order Euler–Poincaré equation. Once this reduced Lagrangian was found, we followed the standard machinery

²⁰This issue is also encountered in first-order Euler–Poincaré reduction theory and had been well-understood in this context (see, for example, [5, Chapter 13]).

introduced earlier to arrive at the k^{th} -order Euler–Poincaré equations for Riemannian cubics (2.40).

Reduced equations for Riemannian cubics had appeared previously in the literature. Indeed, the original paper [18] on Riemannian cubics already contained the reduced equations for cubics on $SO(3)$ with a bi-invariant metric (the so-called NHP equation). Another example appeared in the thesis [52, pp. 118–121], which generalised the treatment to metrics with one-sided invariance. The strategy in both cases was to take variations first, before switching to reduced variables. That is, one starts from the general equation (2.10) and achieves a reformulation in terms of invariant quantities. In the higher-order Euler–Poincaré formalism, contrarily, one immediately performs the reduction and only then takes variations. Once the general formalism has been set up, the equations for cubics can be derived in a straightforward manner from the Euler–Poincaré equations (2.27), as we saw in Section 2.4.1.

As we said in the introduction to this thesis, our main motivation to study higher-order variational problems arose from their potential applications to longitudinal studies in computational anatomy. Thus, in Section 2.5, we considered a higher-order generalisation of first-order template matching (see [10] for an overview), whose task was to find an optimal curve on a diffeomorphism group²¹ whose orbit on some data vector space traversed near given target data points at prescribed times. In Theorem 2.11 and Remark 2.12 we derived the corresponding Euler–Lagrange equations and concluded that optimal curves satisfy higher-order Euler–Poincaré equations on the open time intervals between target (or node) times. The Euler–Lagrange equations also contained so-called node equations, which specify how the optimal curve segments are glued together across target times.

In our treatment we followed paper [37] in formulating everything on the level of general Lie groups acting on vector spaces. The cotangent lift momentum map (2.51) associated with the group action enters the fray in a natural manner, encoding important aspects of the group action. In order to descend to the level of a specific application, it remains to compute this momentum map in the context of interest. Momentum maps for various situations have been compiled in [37]. Another important stepping stone in the

²¹Recall that our computations were done at the formal level. See also the footnote on page 16.

proof of Theorem 2.11 was Lemma 2.4 of the same paper, which in turn is an adaptation from [36] and [57].

We concluded Chapter 2 with a finite-dimensional example of a second-order method for template matching on the sphere acted on by the rotation group. Using numerical simulations we visually contrasted this second-order method, shown in Figure 2.3, with the first-order methods shown in Figure 2.2.

Chapter 3. In template matching applications a Lie group of transformations acts on the vector space of data points. We called such spaces ‘object manifolds’. In variational interpolation problems involving such object manifolds one must make a choice: Either to formulate the optimality condition entirely in terms of the curve in the object manifold, or to also include dependencies on the corresponding curve on the Lie group. Such questions of model selection motivated our investigation, in Chapter 3, of the relationship between Riemannian cubics on object manifolds (equipped with normal metrics) and those on the corresponding Lie groups. For example, under what circumstances does one obtain a cubic on the group when horizontally lifting a cubic on the object manifold?

In order to get a grasp on these types of questions, we derived equation (3.40), which governs cubics for normal metrics. The quantity $\bar{J}$ appearing in this equation is the horizontal generator of the curve in the object manifold, which we had purposefully included from the outset in the variational derivation (see Section 3.4.3). It is closely related to cotangent lift momentum maps, as explained in part (iv) of Proposition 3.1.

To illustrate (3.40), in Section 3.4.4 we revisited cubics on Lie groups from the viewpoint of normal metrics. This approach was seen to be clearly more involved than the Euler–Poincaré procedure developed earlier. It is also, of course, more general. In particular, we showed in Section 3.4.5 that Type I symmetric spaces are object manifolds with normal metrics. By specialising (3.40) to symmetric spaces we arrived at the equation of motion for cubics in (3.56).

The equation for Riemannian cubics in symmetric spaces had been known previously. It was first derived in [19] in a different form (see Remark 3.5, where we commented on its relation to (3.56)). The approach of [19] was to start from the general equation of cubics

(2.10) and particularise it to symmetric spaces using some standard results on covariant derivatives and curvature in symmetric spaces, as can be found, for example, in [67]. Our strategy on the other hand, motivated by the question of horizontal lifting, was to include the horizontal generator of curves from the very beginning. In the case of the 2-sphere and the real projective plane our form of the equation, (3.56), had previously appeared in [26, Lemmas 3 & 4] and the similarity with the NHP equation for $SO(3)$, (2.44), was pointed out there. Remarkably, it was not the horizontality aspect that motivated the treatment in [26]. Rather, the author was interested in the envelope of the family of planar lines defined by a curve in the projective plane. It turns out that in homogeneous coordinates this envelope is exactly the horizontal generator curve $\bar{J}$. In particular, if the projective curve is a Riemannian cubic, the corresponding envelope satisfies (3.56).

Due to the presence of the horizontal generator, (3.40) is well-suited to addressing the question of whether cubics can be lifted horizontally. This was the topic of Section 3.5, which we began by studying Type I symmetric spaces. Theorem 3.7 provided a characterisation of all cubics that can be lifted horizontally. Its proof exploited the aforementioned similarity between (3.56) and (2.44). Namely, if a cubic can be lifted, its horizontal generator curve satisfies both of these equations, hence the commutator term has to vanish and the theorem follows. As a consequence, and with the help of some general theory from [67], we concluded in Theorem 3.8 that precisely those cubics can be lifted that lie in flat, totally geodesic submanifolds. A more informal way of saying this is that the presence of curvature is prohibitive to the horizontal lifting of cubics. Section 3.5.2 was a brief excursion out of the context of normal metrics into the more general territory of Riemannian submersions. This lead, in Theorem 3.11, to a condition for horizontal lifting in terms of O'Neill's A -tensor along the lifted curve applied to velocity and covariant acceleration vectors. In specialising again to normal metrics we obtained Theorem 3.15 and explained, in Remark 3.16, how to recover our previous result on symmetric spaces.

Until this point in our study of the relationship between cubics on object manifolds and those on groups we had ignored *non*-horizontal curves on the group. This was remedied in Section 3.6. The question about horizontal lifts of cubics turned into one about *projections* of cubics. More to the point, what cubics on the group project to cubics on the object manifold? Before summarising the intermediate steps, let us fast-forward to Section 3.6.5

at the end of the Chapter. The idea was the following: Instead of reducing the variational principle for cubics on the group by the *full* group symmetry we only reduced by the symmetry associated with the isotropy subgroup of a given point in the object manifold. We achieved this by the method of second-order Lagrange–Poincaré reduction developed in [50]. In this way we obtained the set of coupled equations (3.119) and (3.120), one describing the horizontal degree of freedom (the projected part) of the cubic on the group, the other describing the vertical degree of freedom (the isotropy part). We recognised in particular that if the right hand side of the horizontal equation (3.119) vanished, then the projected curve was a cubic. This, therefore, was the obstruction for a cubic on the group to project to a cubic on the object manifold.

Now let us back up and describe what came before. In Section 3.6.2 we introduced the first-order theory of Lagrange–Poincaré reduction following [7]. More precisely, we considered *geodesic* curves on the group, reducing the corresponding variational principle by the isotropy subgroup rather than the whole group. The resulting Lagrange–Poincaré equations (3.102) described, respectively, the horizontal and vertical degrees of freedom of the geodesic. In particular, the projected curve was seen to be a geodesic on the object manifold if the first equation of (3.102) had a vanishing right hand side. This equation therefore identified the obstruction for a geodesic on the group to project to a geodesic on the object manifold, akin to what we would later observe for cubics.

After having reduced geodesics in this manner, an obvious question offered itself, which we investigated in Section 3.6.4: Which *geodesics*²² on the group project to *cubics* on the object manifold? For symmetric spaces a necessary and sufficient condition on initial velocities was given in Theorem 3.22. Figure 3.2 visualised the cubics on the sphere that arose in this way as projections of geodesics in $SO(3)$. The rest of the chapter was devoted to the Lagrange–Poincaré reduction of cubics, as summarised above, leading to the identification of the obstruction as the right hand side of (3.119).

Chapter 4. In this chapter we returned, in more generality, to the higher-order template matching application first discussed in Section 2.5. We now allowed for general object manifolds, having previously only considered vector spaces. First, we presented a new

²²Notice that geodesics are also cubics.

derivation of the Euler–Lagrange equations using a set of Lagrange multipliers $\mu^0, \dots, \mu^{k-1}$ in the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$. This had the effect of breaking the explicit link between the velocity $\dot{g}(t)$ of the curve in the group and its trivialisation, $\xi(t)$. In particular, the variational calculus was carried out without taking recourse to formula (2.54), which meant the derivation became self-contained. More importantly, the node equations now took the form (4.11)–(4.14), encoding the continuity properties of optimal curves in terms of the hierarchy of Lagrange multipliers. All but one of the Lagrange multipliers were seen to evolve continuously across node times. μ^0 was recognised to jump discontinuously by a quantity lying in the image set of the momentum map J^Q associated with the group action.

The Lagrange multipliers turned out to be intimately connected to the higher-order version of Legendre’s transform, the so-called Legendre–Ostrogradsky map. Specifically, they were seen to coincide with the trivialised higher-order Legendre–Ostrogradsky momenta of the curve in the Lie group. These had previously been derived in [50] based on the treatment in [38] of higher-order Hamiltonian mechanics on manifolds. We explained these aspects in Section 4.3.3.

A thesis about group invariant variational principles must surely mention Noether’s theorem. Indeed, in an earlier remark in Section 2.6 we had noticed a conservation law implied by the higher-order Euler–Poincaré equations. In Section 4.3.4 we connected this conservation law to group invariance through Noether’s theorem. Due to the discontinuity in μ^0 the conservation law did not extend across node times. We gave a physical interpretation of this effect in Section 4.3.6 in terms of an instantaneous potential (4.31) that kicks the momentum μ^0 at node times. Importantly, the potential depended on the curve in the object manifold rather than on the curve in the group directly. Consequently, a residual symmetry associated with the isotropy group was retained across node times leading to the conservation law (4.30) of Corollary 4.3. This implied in particular that μ^0 at initial time was restricted to a subspace of $\mathfrak{g}^*$ of the same dimension as the object manifold. It should be noted that all of these matters were closely related to the right hand side of (4.11) being a momentum map quantity (the proof of Theorem 4.2 was based on this fact, for instance).

In Section 4.4 we illustrated the theory with a number of examples (rigid body splines,

macromolecular configurations). In both instances the higher-order template matching problem arose as a type of inverse problem, where the state of a physical system was modeled based on a set of partial observations. In Remarks 4.4 and 4.5 we allowed ourselves a sideways glance into probability theory to make the link with inverse problems more explicit. Further details were provided in Appendix A. We only briefly mentioned another application in the context of quantum control theory, devoting Chapter 5 to an in-depth treatment.

With a view to concrete numerical implementation we developed, in the final part of the chapter, a geometric discretisation of the problem at hand. This was done by generalising the methods of [71] to higher order. The resulting integrator was seen to represent faithfully the momentum behaviour of the continuous-time solutions. More precisely, we obtained Theorem 4.10 and Corollary 4.11 as the discrete analogs of Theorem 4.2 and Corollary 4.3, respectively. The practical significance of this was that the numerical search for the optimal starting value of μ^0 could be carried out on subspace of $\mathfrak{g}^*$. As for the other momenta we observed that equations (4.52) and (4.53) at final time were exact parallels of (4.13) and (4.14). In practical terms equations (4.52) and (4.53) could be used upon termination of the algorithm described in Section 4.6 in order to verify that a local minimum of the cost functional had been obtained.

Chapter 5. In the last core chapter of the thesis we considered an application in quantum control. The evolution of quantum systems is encoded in the Schrödinger equation (5.2), the Hamiltonian operator H representing the specifics of the experimental setup. If the Hamiltonian is constant in time, then the curve $U(t)$ in the special unitary matrices is a Lie exponential. In particular, it is a geodesic with respect to the bi-invariant metric induced by the inner product (5.3) on the Lie algebra.

Our goal was to find a *time-dependent* curve $H(t)$ in the space of Hamiltonians to steer an initial (finite-dimensional) quantum state through a series of given target states at prescribed times in a way that required least change to the experimental apparatus in the process. This motivated the introduction of the first term of the cost functional (5.5). By adding the second term we allowed for a mismatch, tunable through the parameter σ , between the achieved quantum trajectory and the target states. The resulting cost

functional brought us into the general framework of the previous chapter, with complex projective space playing the role of object manifold, acted on by the special unitary group. In deriving the Euler–Lagrange equations the only difficulty was to explicitly compute the discontinuity in the evolution of the momentum $M(t)$ in (5.11). Of course, as we recalled in Remark 5.2, this was related to computing the momentum map associated with the action of the special unitary group on complex projective space. The optimal curve $U(t)$ in the special unitary group, generated from $H(t)$ by the Schrödinger equation, was seen to be a Riemannian cubic spline with respect to the bi-invariant metric mentioned above. By following the geometric methods of Section 4.5, we presented numerical simulations of quantum splines for two-level systems in Figures 5.2 and 5.3. The figures showed in particular that, by decreasing σ , a passage closer to the target states could be realised at the expense of a more rapidly varying Hamiltonian. Finally, in Section 5.6, we considered the control of coherent states in bosonic multi-particle systems. We observed that the methods previously developed were directly applicable, since coherent states consist of a number of identical copies of single-particle states.

Further directions. Let us reflect on what we have *not* done in this thesis; where, with further work, one may be able to weaken assumptions and strengthen results.

One of our main motivations for the investigations in this thesis was their relevance in the design of smooth interpolation methods for longitudinal studies in computational anatomy. With this in mind we studied a class of higher-order template matching methods in Section 2.5. It is important to note that our calculations were done at the *formal* level. That is, we formally worked in the framework of finite-dimensional Lie groups acting on vector spaces, without entering into a detailed analysis of infinite-dimensional Lie groups. A mathematically rigorous treatment of the infinite-dimensional case was beyond the scope of this thesis and has been left for future work. Special attention is required, for example, in the definition of Riemannian cubics on diffeomorphism groups. For more details we refer to the last paragraph of Section 2.5.3.

Let us also make a brief remark concerning numerical implementations of the diffeomorphism case. In the introductory section 1.1 we mentioned that in geodesic LDM (large deformation matching) the optimal curves in the diffeomorphism group are horizon-

tal above the corresponding curves in the object manifold. In practical terms this means that the numerical implementation need only be concerned with the subspace of horizontal vector fields, which is of the same dimension as the object manifold. Indeed, in the notation of Chapter 4 we have $\ell = \frac{1}{2}\|\xi\|_{\mathfrak{g}}^2$ and, by Corollary 4.3, $\mu^0(t) = \frac{\delta\ell}{\delta\xi}(t) = \xi(t)^\flat$ lies in the subspace of $\mathfrak{g}^*$ that annihilates $\mathfrak{g}_{q(t)}$, where $q(t)$ is the curve in the object manifold. Hence, $\xi(t)$ is in $\mathfrak{g}_{q(t)}^\perp$.²³ This horizontality property of geodesic LDM does not in general transfer to the higher-order models of Section 2.5. In the case of the second-order cost functional $\ell = \frac{1}{2}\|\dot{\xi}\|_{\mathfrak{g}}^2$, for example²⁴, we have from (4.21) that $\mu^0(t) = -\ddot{\xi}(t)^\flat$. In particular, Corollary 4.3 makes a statement about the horizontality of $\ddot{\xi}(t)$, rather than $\xi(t)$ as in the geodesic case. A numerical implementation will therefore require a parametrisation of the full space of vector fields, rather than just the ones in the horizontal subspace. Hence, one will need to suitably discretise the space of vector fields, a topic which we have not touched upon in this thesis. Nevertheless, let us point to the recent paper [88], which may serve as a useful guide in developing geometric discretisation methods respecting conservation laws such as Corollary 4.3.

The results of Chapter 3 were presented at various levels of generality. Let us recall that bi-invariance of the metric on the group, together with the third relation of (3.53) were characteristic of Type I symmetric spaces. Dropping the latter condition, but retaining bi-invariance, we arrive at homogeneous spaces (see the footnote on page 85). After further relaxation, to right-invariant metrics on the group, we come to the normal metrics as defined in Section 3.2.2. Stepping back once more, one can view normal metrics in the more general framework of Riemannian submersions, as treated in Section 3.5.2. One can now see quite easily that Chapter 3 contained certain omissions, which deserve to be addressed in future work. Horizontal lifts of cubics, for example, were considered in the context of symmetric spaces (Section 3.5.1) as well as homogeneous spaces and Riemannian submersion (both in Section 3.5.2), but not explicitly in the case of general normal metrics. The treatment of Section 3.6.4 on ballistic curves and cubics was restricted entirely to symmetric spaces, whereas the second-order Lagrange–Poincaré reduction of

²³In particular, by the second equation of (3.13), the curve $g(t)$ generated by $\xi(t)$ lies horizontally above $q(t)$.

²⁴This is the reduced Lagrangian for cubics on a Lie group with bi-invariant metric, see Section 2.4.2.

Section 3.6.5 was done at the level of homogeneous spaces. It would be interesting to extend (or specialise, respectively) these consideration to the remaining cases. For example, it should not be too difficult to carry out the second-order Lagrange–Poincaré reduction for the reduced Lagrangian in the special case of symmetric spaces, which, in the notation of Section 3.6.5, takes the form

$$\ell(q, \dot{q}, \ddot{q}, \sigma, \dot{\sigma}) = \frac{1}{2} \left\| \dot{J} + [\bar{J}, \bar{\sigma}] \right\|_{\mathfrak{g}}^2 + \frac{1}{2} \left\| \dot{\sigma} - [\bar{J}, \bar{\sigma}] \right\|_{\mathfrak{g}}^2.$$

A greater degree of difficulty should be expected when attempting to drop the assumption of bi-invariance, as can be intuited by comparing the equation for cubics in the bi-invariant case, (2.43), to the significantly more complicated equation for one-sided invariance, (2.40).

Further, in Section 3.6.4 we used Theorem 3.22 to list the geodesics in $SO(3)$ that project to cubics on the sphere. In future research an effort should be made to try and obtain an exhaustive list of initial conditions consistent with Theorem 3.22 for general Type I symmetric spaces.

At the end of Chapter 3 we identified the obstruction for cubics on the group to project to cubics on the object manifold. Of course, earlier results of the chapter had already contained some implicit information about this obstruction term. For example, if a cubic can be horizontally lifted to a cubic, the obstruction term must vanish. Similarly, it vanishes when a geodesic on the group projects to a cubic. One of the main tasks ahead is to deepen the understanding of the obstruction term, and thereby determine additional situations in which it vanishes.

Chapter 4 also invites further development in several directions. An important limitation of the discrete flow map derived in Section 4.5 is that it is in general only accurate to first order in step size h ; see Remark 4.8. The development of geometric methods with a higher degree of accuracy would be a desirable addition. The class of integrators presented in the preliminary study [79] may prove useful in this regard. An alternative possibility is the Lagrangian approach of [78] together with sufficiently exact approximations of the Lagrangian function. Moreover, the numerical optimisation of the cost functional was based on a shooting method with gradient descent in the space of initial conditions. A comparison with more sophisticated methods of nonlinear programming

(like, for example, the SNOPT algorithm [89]) would be a useful guide for further development. Third, recall that the example of the macromolecular strand (Section 4.4.4) was solved as a problem of statics. Adding the consideration of dynamics of the strand brings one into the realm of so-called $SE(3)$ -Strands [90, 91], which are special $SE(3)$ -valued functions of 2 real variables (physical time as well as the parameter s of Section 4.4.4). It would be interesting to generalise the inexact trajectory planning problems of Chapter 4 to this field theoretic context.

Finally, at the end of Chapter 5 we discussed the control of coherent states associated with the special unitary group. Coherent states can be constructed more generally in the sense of [84, 85]. One starts from a Hilbert space that carries an irreducible unitary representation of some Lie group G . Then the coherent state submanifold is defined as the G -orbit of some reference state.²⁵ The question of controlling such states, subject to optimality conditions similar to the ones of Chapter 5, should be closely related to our work and is likely to prove physically interesting.

²⁵The reference state is in principle arbitrary, but is usually chosen to be a highest weight state of the irreducible representation, see [86, Section IIIB].

References

- [1] I. Newton. *Philosophiae Naturalis Principia Mathematica*. Royal Society, London, 1687. *First published in Latin. Translations into English are available online. A scan of the 1846 translation by A. Motte can be found at <http://archive.org/details/newtonspmathema00newtrich> (accessed 15 May 2013).*
- [2] V. I. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Annales de l'institut Fourier*, **16**(1):319–361, 1966.
- [3] B. A. Khesin and R. Wendt. The Geometry of Infinite-Dimensional Groups. In *Ergebnisse der Mathematik und ihrer Grenzgebiete*, volume 51. Springer, 2009.
- [4] T. Ratiu. The motion of the free n -dimensional rigid body. *Indiana University Mathematics Journal*, **29**(4):609–629, 1980.
- [5] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer, New York, 2nd edition, 2003.
- [6] D. D. Holm, J. E. Marsden, and T. S. Ratiu. The Euler–Poincaré Equations and Semidirect Products with Applications to Continuum Theories. *Advances in Mathematics*, **137**(1):1–81, 1998.
- [7] H. Cendra, J. E. Marsden, and T. S. Ratiu. Lagrangian Reduction by Stages. *Memoirs of the American Mathematical Society*, **152**(722):1–117, 2001.
- [8] F. Gay-Balmaz and C. Tronci. Reduction theory for symmetry breaking with applications to nematic systems. *Physica D*, **239**:1929–1947, 2010.
- [9] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*, **56**:617–694, 1998.
- [10] L. Younes, F. Arrate, and M. I. Miller. Evolutions equations in computational anatomy. *NeuroImage*, **45**(1, Supplement 1):40–50, 2009.

- [11] P. Dupuis, U. Grenander, and M. I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, **56**:587–600, 1998.
- [12] A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, **28**:213–221, 1998.
- [13] A. M. Bloch, J. Baillieul, P. E. Crouch, and J. E. Marsden. *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer, 2003.
- [14] M. F. Beg and A. Khan. Representation of time-varying shapes in the Large Deformation Diffeomorphic Metric Mapping Framework. In *International Symposium of Biomedical Imaging*, 2008.
- [15] S. Durrleman, X. Pennec, A. Trouvé, G. Gerig, and N. Ayache. *Spatiotemporal Atlas Estimation for Developmental Delay Detection in Longitudinal Datasets*, volume 5761 of *Lecture Notes in Computer Science*, pages 297–304. Springer Berlin / Heidelberg, 2009.
- [16] F.-X. Vialard and A. Trouvé. Shape Splines and Stochastic Shape Evolutions: A Second Order Point of View. *Quarterly of Applied Mathematics*, **70**:219–251, 2012.
- [17] S. Gabriel and J. Kajiya. *Spline interpolation in curved space. State of the art in image synthesis*. SIGGRAPH 1985 course notes. ACM Press, New York, 1985.
- [18] L. Noakes, G. Heinzinger, and B. Paden. Cubic Splines on Curved Spaces. *IMA Journal of Mathematical Control & Information*, **6**:465–473, 1989.
- [19] P. E. Crouch and F. Silva Leite. The dynamic interpolation problem: On Riemannian manifolds, Lie groups, and symmetric spaces. *Journal of Dynamical and Control Systems*, **1**(2):177–202, 1995.
- [20] M. Camarinha, F. Silva Leite, and P. E. Crouch. Splines of class C^k on non-Euclidean spaces. *IMA Journal of Mathematical Control & Information*, **12**:399–410, 1995.

- [21] P. E. Crouch, M. Camarinha, and F. Silva Leite. A second order Riemannian variational problem from a Hamiltonian perspective. *Pré-publicações do Departamento de Matemática da Universidade de Coimbra*, 1998. Preprint.
- [22] M. Camarinha, F. Silva Leite, and P. E. Crouch. On the geometry of Riemannian cubic polynomials. *Differential Geometry and its Applications*, **15**(2):107–135, 2001.
- [23] L. Noakes. Null cubics and Lie quadratics. *Journal of Mathematical Physics*, **44**:1436–1448, 2003.
- [24] L. Noakes. Non-null Lie quadratics in E3. *Journal of Mathematical Physics*, **45**:4334–4351, 2004.
- [25] L. Noakes. Duality and Riemannian cubics. *Advances in Computational Mathematics*, **25**:195–209, 2006.
- [26] K. A. Krakowski. Envelopes of splines in the projective plane. *IMA Journal of Mathematical Control and Information*, **22**:171–180, 2005.
- [27] F.C. Park and B. Ravani. Smooth Invariant Interpolation of Rotations. *ACM Transactions on Graphics*, **16**:277–295, 1997.
- [28] M. Zefran, V. Kumar, and C.B. Croke. On the generation of smooth three-dimensional rigid body motions. *IEEE Transactions on Robotics and Automation*, **14**(4):576–589, 1998.
- [29] L. Noakes. *Interpolating Camera Configurations*, volume 2756 of *Lecture Notes in Computer Science*, pages 714–721. Springer, 2003.
- [30] I. H. Hussein and A. M. Bloch. Dynamic interpolation on Riemannian manifolds: an application to interferometric imaging. In *Proceedings of the 2004 American Control Conference*, volume 1, pages 685–690 vol.1, 2004.
- [31] I. H. Hussein and A. M. Bloch. Optimal control on Riemannian manifolds with potential fields, 2004. 43rd IEEE Conference on Decision and Control, Paradise Island, Bahamas.

- [32] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Physical Review Letters*, **71**:1661–1664, 1993.
- [33] D. D. Holm and J. E. Marsden. Momentum maps and measure-valued solutions (peakons, filaments and sheets) for the EPDiff equation. *Progr. Math.*, **232**:203–235, 2004. In *The Breadth of Symplectic and Poisson Geometry, A Festschrift for Alan Weinstein*.
- [34] D. D. Holm, J. T. Ratnanather, A. Trouné, and L. Younes. Soliton dynamics in computational anatomy. *NeuroImage*, **23**(Supplement 1):170–178, 2004.
- [35] M. I. Miller, A. Trouné, and L. Younes. Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, **24**(2):209–228, 2006.
- [36] F.-X. Vialard. *Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework: From the Discontinuous Image Matching Problem to a Stochastic Growth Model*. PhD thesis, Ecole Normale Supérieure de Cachan, 2009. Available at <http://tel.archives-ouvertes.fr/tel-00400379/fr/> (accessed 13 July 2013).
- [37] M. Bruveris, F. Gay-Balmaz, D. D. Holm, and T. S. Ratiu. The momentum map representation of images. *Journal of Nonlinear Science*, **21**(1):115–150, 2011.
- [38] M. de León and P. R. Rodrigues. Formalisme Hamiltonien symplectique sur les fibrés tangents d’ordre supérieur. *Les Comptes Rendus de l’Académie des sciences, Paris*, **301**:455–458, 1985.
- [39] F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu, and F.-X. Vialard. Invariant higher-order variational problems. *Communications in Mathematical Physics*, **309**(2):413–458, 2012.
- [40] F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu, and F.-X. Vialard. Invariant higher-order variational problems II. *Journal of Nonlinear Science*, **22**(4):553–597, 2012.
- [41] D. C. Brody, D. D. Holm, and D. M. Meier. Quantum splines. *Physical Review Letters*, **109**:100501, 2012.

- [42] C. L. Burnett, D. D. Holm, and D. M. Meier. Inexact trajectory planning and inverse problems in the Hamilton–Pontryagin framework. *Proceedings of the Royal Society A*, **469**(2160):20130249, 2013.
- [43] M. de León and P. R. Rodrigues. *Generalized classical mechanics and field theory*. Number 112 in North-Holland Mathematics Studies. Elsevier Science Publishers B.V., 1985.
- [44] M. Crampin, W. Sarlet, and F. Cantrijn. Higher-order differential equations and higher-order Lagrangian mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, **99**(3):565–587, 1986.
- [45] J. W. Milnor. *Morse theory*. Number 51 in Annals of Mathematics Studies. Princeton University Press, 1963.
- [46] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Graduate Texts in Mathematics. Springer, 1997.
- [47] K. Hüper and F. Silva Leite. On the geometry of rolling and interpolation curves on S^n , SO_n and Grassmann manifolds. *Journal of Dynamical and Control Systems*, **13**(4):467–502, 2007.
- [48] F. Silva Leite and K. A. Krakowski. Covariant differentiation under rolling maps. *Pré-publicações do Departamento de Matemática da Universidade de Coimbra*, 2008. Preprint.
- [49] J. Hinkle, P. Muralidharan, P. T. Fletcher, and S. Joshi. Polynomial regression in Riemannian manifolds. In *Computer Vision – ECCV 2012*, volume 7574 of *Lecture Notes in Computer Science*. Springer, 2012.
- [50] F. Gay-Balmaz, D. D. Holm, and T. S. Ratiu. Higher order Lagrange–Poincaré and Hamilton–Poincaré reductions. *Bulletin of the Brazilian Mathematical Society*, **42**(4):579 – 606, 2011.
- [51] A. Kriegl and P. W. Michor. *The Convenient Setting of Global Analysis*, volume 53 of *Surveys and Monographs*. American Mathematical Society, 1997.

- [52] K. A. Krakowski. *Geometrical Methods of Inference*. PhD thesis, University of Western Australia, 2002. Available online at <http://turing.une.edu.au/~kris/publications.html> (accessed 28 July 2013).
- [53] L. Noakes and T. Popiel. Null Riemannian cubics in tension in $SO(3)$. *IMA Journal of Mathematical Control and Information*, **22**:477–488, 2005.
- [54] M. I. Miller, A. Trouvé, and L. Younes. On the metrics and Euler–Lagrange Equations of Computational Anatomy. *Annual Review of Biomedical Engineering*, **4**:375–405, 2002.
- [55] M. I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, **41**(1-2):61–84, 2001.
- [56] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM Journal on Mathematical Analysis*, **37**(1):17–59, 2005.
- [57] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, **61**(2):139–157, 2005.
- [58] L. Younes. *Shapes and Diffeomorphisms*. Applied Mathematical Sciences. Springer, 2010.
- [59] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, **56**(2):411–421, 2006.
- [60] J. Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, **38**:95–113, 2007.
- [61] F.-X. Vialard. A second-order model for diffeomorphic evolution of images, 2012. *Private communication*.
- [62] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.

- [63] L. Colombo and D. Martín de Diego. On the geometry of higher-order variational problems on Lie groups, April 2011. Preprint available at <http://arxiv.org/abs/1104.3221>.
- [64] F. Gay-Balmaz and T. S Ratiu. Clebsch optimal control formulation in mechanics. *The Journal of Geometric Mechanics*, **3**(1):41–79, 2011.
- [65] P. W. Michor. *Topics in Differential Geometry*, volume 93 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2008.
- [66] B. O’Neill. The fundamental equations of a submersion. *The Michigan Mathematical Journal*, **13**:459–469, 1966.
- [67] S. Helgason. *Differential Geometry and Symmetric Spaces*. Pure and Applied Mathematics. Academic Press, 1962.
- [68] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry, Volume II*. Interscience Publishers, Wiley & Sons, 1969.
- [69] D. Alekseevsky, A. Kriegl, M. Losik, and P. W. Michor. The Riemannian geometry of orbit spaces – the metric, geodesics, and integrable systems. *Publicationes Mathematicae Debrecen*, **62**:247–276, 2003.
- [70] F. Gay-Balmaz, D. D. Holm, and T. S Ratiu. Geometric dynamics of optimization. *Communications in Mathematical Sciences*, **11**(1):163–231, 2013.
- [71] N. Bou-Rabee and J. E. Marsden. Hamilton Pontryagin integrators on Lie groups part I: Introduction and structure-preserving properties. *Foundations of Computational Mathematics*, **9**:197–219, 2009.
- [72] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics part II: Variational structures. *Journal of Geometry and Physics*, **57**(1):209–250, 2006.
- [73] B. Øksendahl. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer, 6th edition, 2003.
- [74] D. D. Holm. *Geometric Mechanics, Part II: Rotating, Translating and Rolling*. Imperial College Press, 2nd edition, 2011.

- [75] Y. Shi and J. E. Hearst. The Kirchhoff elastic rod, the nonlinear Schrödinger equation, and DNA supercoiling. *The Journal of Chemical Physics*, **101**(6):5186–5200, 1994.
- [76] J. S. Kim and G. S. Chirikjian. Conformational analysis of stiff chiral polymers with end-constraints. *Molecular Simulation*, **32**(14):1139–1154, 2006.
- [77] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, **10**:357–514, 2001.
- [78] L. Colombo, F. Jimenez, and D. Martín de Diego. Discrete second-order Euler–Poincaré equations: Applications to optimal control. *International Journal of Geometric Methods in Modern Physics*, **09**(04):1250037–1250057, 2012.
- [79] C. L. Burnett, D. D. Holm, and D. M. Meier. Geometric integrators for higher-order mechanics on Lie groups, December 2011. Available at <http://arxiv.org/abs/1112.6037>.
- [80] A. Sudbery. *Quantum mechanics and the particles of nature – an outline for mathematicians*. Cambridge University Press, 1986.
- [81] D. C. Brody. Elementary derivation for passage times. *Journal of Physics A*, **36**(20):5587–5593, 2003.
- [82] A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira. Time-optimal quantum evolution. *Physical Review Letters*, **96**:060503, 2006.
- [83] D. C. Brody and D. W. Hook. On optimum Hamiltonians for state transformations. *Journal of Physics A*, **39**:L167–L170, 2006. Corrigendum *ibid.* **40**:10949, 2007.
- [84] A. M. Perelomov. Coherent states for arbitrary Lie groups. *Communications in Mathematical Physics*, **26**(3):222–236, 1972.
- [85] R. Gilmore. Geometry of symmetrized states. *Annals of Physics*, **74**(2):391–463, 1972.
- [86] W. M. Zhang, D. H. Feng, and R. Gilmore. Coherent states: Theory and some applications. *Reviews of Modern Physics*, **62**(4):867–927, 1990.

- [87] D. C. Brody and E. M. Graefe. Coherent states and rational surfaces. *Journal of Physics A*, **43**:255205–255219, 2010.
- [88] D. Pavlov, P. Mullen, Y. Tong, E. Kanso, J. E. Marsden, and M. Desbrun. Structure-preserving discretization of incompressible fluids. *Physica D*, **240**(6):443–458, 2011.
- [89] P. E. Gill, W. Murray, and M. A. Sauder. SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization. *SIAM Review*, **47**:99–131, 2005.
- [90] D. D. Holm, R. I. Ivanov, and J. R. Percival. G -strands. *Journal of Nonlinear Science*, **22**(4):517–551, 2012.
- [91] D. D. Holm and R. I. Ivanov. Matrix G -Strands, March 2013. Preprint available at <http://arxiv.org/abs/1305.4010>.
- [92] J. Jacod and P. Protter. *Probability essentials*. Universitext. Springer, 2nd edition, 2003.
- [93] E. Todorov. General duality between optimal control and estimation. In *Proceedings of the 47th IEEE Conference on Decision and Control*. IEEE, 2008.
- [94] D. J. Higham. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, **43**(3):525–546, 2001.

A Probabilistic interpretation of Riemannian cubics

In this appendix we shall present a formal argument supporting Remark 4.4. In the interest of brevity we shall not attempt to provide an introduction to probability theory here. Instead we will supply the reader with pointers to standard textbooks such as [92], where the probabilistic concepts and tools used here are explained. Our strategy is the following: We consider a discrete approximation to the stochastic differential equation (4.35) and explain how its most likely trajectory is linked to a discrete-time optimal control problem. As we subsequently let the time step go to zero we formally recover the cost functional (4.36). A very similar setting to ours is considered in [93, Section V], whose methods we adapt for our purposes.

We shall use the notation of Chapter 4. To begin with, we discretise the time axis $[t_0, t_l]$ by replacing it with discrete time points $t_k = t_0 + kh$, $k = 0, \dots, N$, where h is the step size and $t_l = t_0 + Nh$. We use integers N_i , $i = 1, \dots, l$, as node indices, that is, $t_i = t_0 + N_i h$. For convenience let us also define $N_0 := 0$, and note that $N_l = N$. We approximate (4.35) with a stochastic Euler scheme,

$$\xi_{k+1} - \xi_k = -h \operatorname{ad}_{\xi_k}^\dagger \xi_k + \sigma_W \sum_{i=1}^d (W_{k+1}^i - W_k^i) e_i, \quad (\text{A.1})$$

where the random variable W_k^i is equal to the Brownian motion W^i at time t_k (we refer to [73, Chap. 2.2] for a definition). Moreover, let us approximate $g(t)$ by the sequence

$$g_{k+1} = \exp(h\xi_k) g_k, \quad (\text{A.2})$$

where $\exp$ denotes the Lie group exponential. We collect the random variables W_k^i in d -dimensional vectors W^i and then again in a global vector $\mathbf{W} = (W_1, \dots, W_{N_l})$. Note that we omitted W_0 , since a Brownian motion always starts at zero. The measurements of the path $g_k Q_0$ at times $t_i = t_0 + N_i h$ are represented by random variables Q_{N_i} , which we summarise in a vector $\mathbf{Q}$.

The discrete version of the question we addressed in Remark 4.4 is the following: Suppose we are given a set of observations $q_{N_i} \in V$ (recall that V is the data vector space) at times t_i , $i = 1, \dots, l$. What is the most likely path g_k on the Lie group that would have led to these observations? More to the point, what are the most likely values

of the random variables W_k^i that would have led to these observations through evolution equations (A.1) and (A.2)? The answer to this question is found by maximising the conditional probability density for $\mathbf{W}$ given observations $\mathbf{q}$, which we denote by

$$f_{\mathbf{W}|\mathbf{Q}=\mathbf{q}}(\mathbf{w}). \quad (\text{A.3})$$

We refer to [92, p. 89] for the definition of conditional densities. Due to the Markov property of Brownian motion, the vector W_{k+1} is independent of W_j for $j = 1, \dots, k-1$. Together with Theorem 3.3 of [92] this leads to

$$f_{\mathbf{W}|\mathbf{Q}=\mathbf{q}}(\mathbf{w}) \sim f_{W_1}(w_1) f_{W_2|W_1=w_1}(w_2) \cdots f_{W_N|W_{N-1}=w_{N-1}}(w_N) \prod_{i=1}^l f_{Q_{N_i}|\mathbf{W}=\mathbf{w}}(q_{N_i}),$$

up to a factor that is independent of $\mathbf{w}$. The measurement process at a given time t_i only depends on the Brownian motion through the momentary true state $g_{N_i}Q_0$ of the observed system as defined by (A.1) and (A.2). Hence,

$$= f_{W_1}(w_1) f_{W_2|W_1=w_1}(w_2) \cdots f_{W_N|W_{N-1}=w_{N-1}}(w_N) \prod_{i=1}^l f_{Q_{N_i}|\text{state}=g_{N_i}Q_0}(q_{N_i}).$$

Brownian increments across a time step h are normally distributed with mean zero and variance h . By the assumptions of Remark 4.4, the measurement outcome given a certain true state is also normally distributed, with variance σ_n^2 and expected value equal to the true state. Hence,

$$\begin{aligned} &\sim \exp\left(-\frac{1}{2h} (\|w_1\|^2 + \|w_2 - w_1\|^2 + \cdots + \|w_N - w_{N+1}\|^2)\right) \\ &\quad \cdot \exp\left(-\frac{1}{2\sigma_n^2} \sum_{i=1}^l \|g_{N_i}Q_0 - q_{N_i}\|_V^2\right), \end{aligned}$$

up to some normalisation factors that are independent of $\mathbf{w}$. Here $\|\cdot\|$ denotes the standard Euclidean norm. Taking the logarithm of this expression and multiplying by -1 we obtain

$$\frac{1}{2h} (\|w_1\|^2 + \|w_2 - w_1\|^2 + \cdots + \|w_N - w_{N+1}\|^2) + \frac{1}{2\sigma_n^2} \sum_{i=1}^l \|g_{N_i}Q_0 - q_{N_i}\|_V^2.$$

The multiplication by -1 means we are now seeking a minimum of this expression rather than a maximum. Changing from the $\mathbf{w}$ variables to the ξ variables using (A.1) and

multiplying by σ_W^2 we get

$$\sum_{k=0}^{N-1} \frac{1}{2} \left\| \frac{\xi_{k+1} - \xi_k}{h} + \text{ad}_{\xi_k}^\dagger \xi_k \right\|_{\mathfrak{g}}^2 h + \frac{\sigma_W^2}{2\sigma_n^2} \sum_{i=1}^l \|g_{N_i} Q_0 - q_{N_i}\|_V^2. \quad (\text{A.4})$$

We can formally take the limit $h \rightarrow 0$ thereby obtaining the cost functional S in (4.36). Moreover, one can show that the discrete Euler–Lagrange equations for (A.4) produce a first-order accurate integrator for the continuous-time solution (see also Section 4.5 of this thesis). That is, the minimiser of (A.4) converges pointwise to the minimiser of (4.36) as $h \rightarrow 0$.

The missing link that would make Remark 4.4 into a mathematically rigorous statement (rather than a formal one) is to show that, as $h \rightarrow 0$, the discrete maximum likelihood curve $(W_1, \dots, W_N)$ converges to the maximum likelihood Brownian path $W(t)$ of the continuous-time stochastic differential equation (4.35). This would require a closer look at the convergence properties of the stochastic Euler scheme introduced above. This topic is far developed (see [94] for an introduction) and one might suspect that a thorough search of the literature would yield what is needed. However, such matters are beyond the scope of this appendix and shall not be pursued here.

B Gradient calculation via adjoint equations

The purpose of this appendix is to provide a detailed derivation of the adjoint equations for the inexact trajectory planning problem, as promised in Section 4.6.

In order to implement an efficient descent method for $\mathcal{J}_{d,\text{initial}}$ it is useful to have an expression for its gradient. One may obtain a gradient estimate using finite difference methods. One drawback lies with the inaccuracies inherent in the estimation. Moreover, if the dimension of the Lie algebra $\mathfrak{g}$ is large such estimations quickly become computationally costly. Both of these drawbacks can be circumvented in our case by the use of adjoint equations. In this way we obtain an exact expression for the gradient of $\mathcal{J}_{d,\text{initial}}$, in a computationally efficient way. We will derive the system of adjoint equations now. For simplicity we treat only the special case where $\ell = \frac{1}{2} \|\xi^1\|_\gamma^2$, where γ denotes an inner product on $\mathfrak{g}$ and $\|\cdot\|$ the corresponding norm. Moreover we will only treat the left-action, right-reduction case, however the others can be obtained in the same way. We recall from (4.47)–(4.51) the equations of motion

$$g_{k+1} = \tau(h\xi_k^0)g_k, \quad \xi_{k+1}^0 = \xi_k^0 + h(\mu_{k+1}^1)^\# \quad (\text{B.1})$$

$$\mu_{k+1}^1 = \mu_k^1 - h(D\tau_{-h\xi_k^0})^* \mu_{k+1}^0, \quad (\text{B.2})$$

$$\mu_{k+1}^0 = (D\tau_{-h\xi_k^0}^{-1})^* (D\tau_{h\xi_k^0})^* (\mu_k^0 + \Delta_k(g_k Q_0)), \quad (\text{B.3})$$

where we introduce functions $\Delta_k : Q \rightarrow \mathfrak{g}^*$ for $k = 0, \dots, N$ defined as

$$\Delta_{N_i}(q) := \frac{d_{N_i}}{\sigma^2} J^Q(d_1 d(g_{N_i} Q_0, Q_{t_i}))$$

when $k \in \{N_1, \dots, N_l\}$ and $\Delta_k = 0$ otherwise.

Let us define an augmented functional $\mathcal{G}$, in which these equations are paired with Lagrange multipliers. These Lagrange multipliers will be denoted $(P_k^0, P_k^1, V_k^0, V_k^1) \in 2\mathfrak{g}^* \times 2\mathfrak{g}$ for $k = 1, \dots, N$. Let us introduce the shorthand notation x representing the discrete path $(g_k, \xi_k^0, \mu_k^0, \mu_k^1)_{k=0}^N$ and λ representing the ensemble of Lagrange multiplier $(P_k^0, P_k^1, V_k^0, V_k^1)_{k=1}^N$. The augmented functional $\mathcal{G}$ is given by

$$\begin{aligned} \mathcal{G}(x, \lambda) = & h \left[\sum_{k=0}^{N-1} \frac{1}{2} \|(\mu_{k+1}^1)^\#\|_\gamma^2 + \langle P_{k+1}^0, \tau^{-1}(g_{k+1} g_k^{-1}) - h\xi_k^0 \rangle \right. \\ & \left. + \langle P_{k+1}^1, \xi_{k+1}^0 - \xi_k^0 - h(\mu_{k+1}^1)^\# \rangle + \langle \mu_{k+1}^1 - \mu_k^1 + h(D\tau_{-h\xi_k^0})^* \mu_{k+1}^0, V_{k+1}^0 \rangle \right] \end{aligned}$$

$$+ \left\langle (D\tau_{-h\xi_k^0})^* \mu_{k+1}^0 - (D\tau_{h\xi_k^0})^* (\mu_k^0 + \Delta_k(g_k Q_0)), V_{k+1}^1 \right\rangle \Big] + \frac{1}{2\sigma^2} \sum_{i=1}^l d^2(g_{N_i} Q_0, Q_{t_i}).$$

No constraints are assumed here, apart from the prescribed initial velocity ξ_0^0 and $g_0 = e$. It is clear that for any choice of Lagrange multipliers λ we have $\mathcal{G}(x, \lambda) = \mathcal{J}_{d,\text{initial}}(\mu_0^0, \mu_0^1)$, as long as x satisfies (B.1)–(B.3) for given initial values μ_0^0, μ_0^1 . A tedious, but straightforward calculation shows that

$$\delta\mathcal{G}(x, \lambda) = -h \langle \delta\mu_0^1, V_1^0 \rangle - h \langle \delta\mu_0^0, D\tau_{h\xi_0^0} V_1^1 \rangle, \quad (\text{B.4})$$

if x satisfies (B.1)–(B.3) and λ is a solution of the *adjoint equations*. We describe these now. We introduce functions $K^\pm : 2\mathfrak{g} \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ by the defining relation

$$\langle K_{\xi, \mu}^\pm V, \rho \rangle = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \langle (D\tau_{\pm h(\xi + \varepsilon\rho)})^* \mu, V \rangle, \quad \text{for all } \xi, V, \rho \in \mathfrak{g}, \mu \in \mathfrak{g}^*.$$

Moreover, for $k = 0, \dots, N$ we define functions $\mathcal{A}_k : Q \times \mathfrak{g} \rightarrow \mathfrak{g}^*$ by

$$\langle \mathcal{A}_k(q, \rho), \eta \rangle = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \langle \Delta_k(\exp(\varepsilon\eta)q), \rho \rangle,$$

for all $q \in Q$ and $\rho, \eta \in \mathfrak{g}$. The adjoint equations consist of conditions at the final time point,

$$P_N^0 = -h^{-1}(D\tau_{-h\xi_{N-1}^0})^* \Delta_N(g_N Q_0), \quad P_N^1 = 0, \quad (\text{B.5})$$

$$V_N^0 = -(\mu_N^1)^\sharp, \quad V_N^1 = -hV_N^0, \quad (\text{B.6})$$

and the following equations for $k = 1, \dots, N-1$,

$$P_k^0 = (D\tau_{-h\xi_{k-1}^0})^* \left[(D\tau_{h\xi_k^0}^{-1})^* P_{k+1}^0 + \mathcal{A}_k(g_k Q_0, D\tau_{h\xi_k^0} V_{k+1}^1) - h^{-1} \Delta_k(g_k Q_0) \right] \quad (\text{B.7})$$

$$P_k^1 = P_{k+1}^1 + hP_{k+1}^0 - hK_{\xi_k^0, \mu_{k+1}^0}^- V_{k+1}^0 - K_{\xi_k^0, \mu_{k+1}^0}^- V_{k+1}^1 + K_{\xi_k^0, \mu_k^0 + \Delta_k(g_k Q_0)}^+ V_{k+1}^1 \quad (\text{B.8})$$

$$V_k^0 = V_{k+1}^0 - (\mu_k^1)^\sharp + h(P_k^1)^\sharp, \quad (\text{B.9})$$

$$V_k^1 = -hV_k^0 + D\tau_{-h\xi_{k-1}^0}^{-1} D\tau_{h\xi_k^0} V_{k+1}^1. \quad (\text{B.10})$$

These equations are posed backwards. That is, solving the adjoint equations entails initialising the Lagrange multipliers at time point N according to (B.5)–(B.6) and then iterating backwards from $k = N$ to $k = 1$ using (B.7)–(B.10).

We now obtain an expression for the gradient of $\mathcal{J}_{d,\text{initial}}$ from (B.4). Indeed, let $(\mu_0^0(\varepsilon), \mu_0^1(\varepsilon))$ be a variation of initial conditions (μ_0^0, μ_0^1) , and let $x(\varepsilon)$ be the corresponding set of solutions to (B.1)–(B.3). Let λ be a solution to the adjoint equations (B.7)–(B.10) for $x = x(0)$, then

$$\begin{aligned} \delta \mathcal{J}_{d,\text{initial}} &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{J}_{d,\text{initial}}(\mu_0^0(\varepsilon), \mu_0^1(\varepsilon)) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{G}(x(\varepsilon), \lambda) \\ &= -h \langle \delta \mu_0^1, V_1^0 \rangle - h \langle \delta \mu_0^0, D\tau_{h\xi_0^0} V_1^1 \rangle. \end{aligned}$$

From this we can read off the gradient,

$$\frac{\delta \mathcal{J}_{d,\text{initial}}}{\delta \mu_0^0} = -h D\tau_{h\xi_0^0} V_1^1, \quad \frac{\delta \mathcal{J}_{d,\text{initial}}}{\delta \mu_0^1} = -h V_1^0.$$

C Permission documents

As required by Imperial College London, copies of the licences to reproduce figures in [39, 40] are appended below. As for [41], published by The American Physical Society, we refer to <http://publish.aps.org/copyrightFAQ.html#figures> (*accessed 14 August 2013*), which states that authors have the “right to use figures, tables, graphs, etc. in subsequent publications using files prepared and formatted by you or the APS-prepared versions. The appropriate bibliographic citation must be included.”

Rightslink Printable License

<https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...>

SPRINGER LICENSE TERMS AND CONDITIONS

Apr 22, 2013

This is a License Agreement between David M. Meier ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3131861179451
License date	Apr 18, 2013
Licensed content publisher	Springer
Licensed content publication	Communications in Mathematical Physics
Licensed content title	Invariant Higher-Order Variational Problems
Licensed content author	François Gay-Balmaz
Licensed content date	Jan 1, 2011
Volume number	309
Issue number	2
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Title of your thesis / dissertation	Invariant higher-order variational problems: Reduction, geometry and applications
Expected completion date	Sep 2013
Estimated size(pages)	150
Total	0.00 USD
Terms and Conditions	

Introduction

The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

1 of 4

22/04/13 16:09

Rightslink Printable License

<https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...>

Limited License

With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided its password protected or on the university's intranet or repository, including UMI (according to the definition at the Sherpa website: <http://www.sherpa.ac.uk/romeo/>). For any other electronic use, please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com).

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, subject to a courtesy information to the author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well).

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted

You may not alter or modify the material in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)).

Reservation of Rights

Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice/Disclaimer

You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer and the original publisher /journal title, volume, year of publication,

2 of 4

22/04/13 16:09

Rightslink Printable License

<https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...>

page, chapter/article title, name(s) of author(s), figure number(s), original copyright notice) is given to the publication in which the material was originally published, by adding, with kind permission from Springer Science and Business Media"

Warranties: None

Example 1: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Example 2: Springer Science + Business Media makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity

You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer Science + Business Media's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms

Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held

3 of 4

22/04/13 16:09

Rightslink Printable License

<https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...>

in The Netherlands, in accordance with Dutch law, and to be conducted under the Rules of the 'Netherlands Arbitrage Instituut' (Netherlands Institute of Arbitration).OR:

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other terms and conditions:

v1.3

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK501002633. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact Rightslink Customer Support: customer@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

4 of 4

22/04/13 16:09

Rightslink Printable License

https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...

SPRINGER LICENSE TERMS AND CONDITIONS

Apr 22, 2013

This is a License Agreement between David M Meier ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3131861337104
License date	Apr 18, 2013
Licensed content publisher	Springer
Licensed content publication	Journal of Nonlinear Science
Licensed content title	Invariant Higher-Order Variational Problems II
Licensed content author	François Gay-Balmaz
Licensed content date	Jan 1, 2012
Volume number	22
Issue number	4
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Title of your thesis / dissertation	Invariant higher-order variational problems: Reduction, geometry and applications
Expected completion date	Sep 2013
Estimated size(pages)	150
Total	0.00 USD
Terms and Conditions	

Introduction

The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

1 of 4

22/04/13 16:09

Rightslink Printable License

https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...

Limited License

With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided its password protected or on the university's intranet or repository, including UMI (according to the definition at the Sherpa website: <http://www.sherpa.ac.uk/romeo/>). For any other electronic use, please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com).

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, subject to a courtesy information to the author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well).

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted

You may not alter or modify the material in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)).

Reservation of Rights

Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice-Disclaimer

You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer and the original publisher /journal title, volume, year of publication,

2 of 4

22/04/13 16:09

Rightslink Printable License

https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...

page, chapter/article title, name(s) of author(s), figure number(s), original copyright notice) is given to the publication in which the material was originally published, by adding, with kind permission from Springer Science and Business Media"

Warranties: None

Example 1: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Example 2: Springer Science + Business Media makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity

You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer Science + Business Media's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms

Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held

3 of 4

22/04/13 16:09

Rightslink Printable License

https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref...

in The Netherlands, in accordance with Dutch law, and to be conducted under the Rules of the 'Netherlands Arbitrage Instituut' (Netherlands Institute of Arbitration). **OR:**

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other terms and conditions:

v1.3

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLKNS01002637. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customer@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

4 of 4

22/04/13 16:09