

Geometric mechanics and Lagrangian reduction

A thesis presented for the degree of
Doctor of Philosophy of the Imperial College London
and the
Diploma of Imperial College London (DIC)
by

David Ellis

Department of Mathematics
Huxley Building, 180 Queen's Gate
Imperial College London
London. SW7 2AZ

MAY 17, 2011

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed:

David Ellis

COPYRIGHT

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the doctorate thesis archive of the college central library. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in Imperial College London, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement. Further information on the conditions under which disclosures and exploitation may take place is available from the Imperial College London registry.

ABSTRACT

The purpose of this thesis is two-fold: Firstly, to contribute to the tools available to geometric mechanics; secondly, to apply the geometric perspective to two particular problems.

The thesis falls into three parts. The first part deals with the dynamics of charged molecular strands (CMS). The second part contributes general tools for use in geometric mechanics. The third part develops a new geometric modelling technique and applies it to image dynamics.

Part I develops equations of motion for the dynamical folding of CMS (such as DNA). The CMS are modelled as flexible continuous filamentary distributions of interacting rigid charge conformations, and their dynamics are derived via a modified Hamilton-Pontryagin variational formulation. The new feature is the inclusion of *nonlocal* screened Coulomb interactions, or Lennard-Jones potentials between pairs of charges. The CMS equations are shown to arise from a form of Lagrangian reduction initially developed for complex fluids. Subsequently, the equations are also shown to arise from Lagrange-Poincaré reduction of a field theory. This dual interpretation of the CMS equations motivates the undertakings of Part II.

In Part II, a general treatment of Lagrange-Poincaré (LP) reduction theory is undertaken. The LP equations are cast into a field theoretic context together with their associated constrained variational principle. An integrability/reconstruction condition is established that relates solutions of the original problem with those of the reduced problem. The new contribution of the LP framework is to unify the Lagrange-Poincaré field reduction with the canonical theory, which involves a single independent variable, and to extend LP field reduction to the general fibre bundle setting.

The Kelvin-Noether theorem is generalised in two new ways; from the Euler-Poincaré to the LP setting, and from the canonical to the field setting. The importance of the extended Kelvin-Noether theorem is elucidated by an application to

the CMS problem, yielding new qualitative insight into molecular strand dynamics.

Finally, Part III gives a full geometric development of a new technique called *un-reduction*, that uses the canonical LP reduction back-to-front. Application of un-reduction leads to new developments in image dynamics.

ACKNOWLEDGEMENTS

I would like to thank my collaborators, Dorje Brody, Martins Bruveris, Francois Gay-Balmaz, Darryl Holm, Tudor Ratiu and Vakhtang Putkaradze, for their guidance and many delightful endeavours. Thank you to Cesare Tronci for showing me the ropes. I would also like to thank Cardiomatics, Shape Focused Research Group and Geometric Mechanics and Control Network for numerous stimulating interactions. Thank you to Engineering and Physical Sciences Research Council (EPSRC) for funding my PhD studies. Finally, I would like to thank my supervisor, Darryl Holm, for his unwavering support and guidance, and for the fun we have had in developing this work.

TABLE OF CONTENTS

ABSTRACT	4
ACKNOWLEDGEMENTS	6
I CHARGED MOLECULAR STRAND DYNAMICS	13
1 INTRODUCTION	14
1.1 PHYSICAL SETUP	14
1.1.1 OUTLINE OF PART I	17
1.2 MATHEMATICAL SETUP	20
1.2.1 DESCRIPTION OF THE VARIABLES INVOLVED	20
1.2.2 THE CLASSICAL PERSPECTIVE	22
1.2.3 THE COVARIANT PERSPECTIVE	25
1.3 CONNECTION TO PREVIOUS STUDIES	27
1.3.1 KIRCHHOFF EQUATIONS FOR PURE ELASTIC ROD	27
1.3.2 REDUCTIONS FOR A RIGID FILAMENT	34
1.4 SUMMARY	37
2 MOTION OF EXACT SELF-INTERACTING GEOMETRIC RODS	39
2.1 PROBLEM SET-UP	39
2.1.1 NONLOCAL POTENTIAL ENERGY	40
2.1.2 KINEMATICS	44
2.1.3 GENERALISATIONS OF THE STRAND	46
2.2 FILAMENT DYNAMICS	48
2.2.1 LOCAL DYNAMICS	48
2.2.2 NONLOCAL DYNAMICS	52
2.3 CONSERVATION LAWS	56
2.4 SUMMARY	60
3 CLASSICAL FORMULATION OF MOLECULAR STRANDS	61
3.1 INTRODUCTION TO AFFINE LIE GROUP ACTIONS	61
3.1.1 CLASSICAL LAGRANGIAN APPROACH	61
3.1.2 MAIN GOALS OF THE LAGRANGIAN APPROACH	62
3.1.3 PEDAGOGICAL EXAMPLES	62
3.1.4 THE STRAND WITH NONLOCAL INTERACTIONS	64

3.2	BACKGROUND MATERIAL	65
3.2.1	NOTATIONS FOR SEMIDIRECT PRODUCTS	65
3.2.2	AFFINE EULER-POINCARÉ REDUCTION	66
3.3	APPLICATION TO THE CHARGED STRAND	68
3.3.1	RECOVERING THE HAMILTON-PONTRYAGIN APPROACH	71
3.3.2	CLASSICAL STRAND EQUATIONS	73
3.4	RELATION TO OTHER FORMULATIONS	76
3.4.1	ELASTIC FILAMENT DYNAMICS AND KIRCHHOFF'S THEORY	76
3.4.2	THE CHARGED MOLECULAR STRAND: GENERAL CASE	77
3.4.3	CONSERVATION LAWS AND SPATIAL FORMULATION	80
3.4.4	THE FIXED FILAMENT AND ITS CONSERVATION LAW	81
3.5	SUMMARY	83
4	COVARIANT FORMULATION OF MOLECULAR STRANDS	84
4.1	COORDINATE TRANSFORMATION	85
4.1.1	MOTIVATION	85
4.1.2	COORDINATE TRANSFORMATION	86
4.1.3	THE GENERAL CASE	92
4.1.4	SUMMARY	94
4.2	COVARIANT EULER-POINCARÉ FORMULATION	95
4.2.1	BACKGROUND MATERIAL	96
4.2.2	APPLICATION TO THE MOLECULAR STRAND	98
4.3	COVARIANT LAGRANGE-POINCARÉ FORMULATION	100
4.3.1	BACKGROUND MATERIAL	101
4.3.2	APPLICATION TO THE MOLECULAR STRAND	102
4.3.3	THE GENERAL CASE	103
4.4	SUMMARY	105
5	CONCLUSION OF PART I	106
II	LAGRANGE-POINCARÉ FIELD REDUCTION	109
6	INTRODUCTION	110
6.1	AIMS OF PART II	111
6.2	GEOMETRIC CONSTRUCTIONS	113
6.2.1	GEOMETRIC SETTING	113
6.2.2	REDUCED COVARIANT CONFIGURATION SPACE	120
6.3	SUMMARY	122

7	LAGRANGE-POINCARÉ FIELD REDUCTION	124
7.1	LAGRANGE-POINCARÉ FIELD EQUATIONS	124
7.1.1	REDUCED VARIATIONS	126
7.1.2	THE LAGRANGE-POINCARÉ FIELD EQUATIONS	134
7.1.3	RECONSTRUCTION	142
7.2	SUMMARY	148
8	CONSERVATION LAWS AND REPRESENTATIONS	150
8.1	REPRESENTATIONS AND NOETHER'S THEOREM	150
8.1.1	CONVECTIVE REPRESENTATION	150
8.1.2	SPATIAL REPRESENTATION	152
8.2	THE KELVIN-NOETHER THEOREM	153
8.3	KELVIN-NOETHER THEOREM FOR THE MOLECULAR STRAND	156
8.4	SUMMARY	159
9	CONCLUSION OF PART II	160
III	UN-REDUCTION	165
10	UNREDUCTION	166
10.1	INTRODUCTION	166
10.1.1	INTRODUCING UN-REDUCTION	167
10.1.2	AIMS OF PART III	170
10.2	LAGRANGIAN REDUCTION WITH A FORCE FIELD	171
10.3	DISTORTIONS OF PROJECTED DYNAMICS	174
10.3.1	DISTORTION IN THE LAGRANGE-POINCARÉ EQUATIONS.	177
10.4	UN-REDUCTION	178
10.4.1	THE UN-REDUCTION THEOREM	178
10.4.2	HORIZONTALITY CONDITIONS FOR GEODESIC PROBLEMS	180
10.5	CLOSED, SIMPLE PLANE CURVE MATCHING	183
10.5.1	GEOMETRIC SETUP	184
10.5.2	DERIVATION OF UN-REDUCTION EQUATIONS	187
10.5.3	EXAMPLES OF UN-REDUCED IMAGE DYNAMICS	191
10.6	SUMMARY	197
IV	OUTLOOK	199
11	OUTLOOK	200
A	APPENDIX OF NOTATION	202

LIST OF FIGURES

1.1	RIGID CONFORMATIONS OF CHARGES ARE DISTRIBUTED ALONG A CURVE. NOTE THAT THIS IS A SPATIAL REPRESENTATION OF THE ORIENTATIONS OF THESE CONFORMATIONS OF CHARGES.	18
1.2	A DIAGRAM THAT SHOWS THE RELATIONSHIP BETWEEN VARIABLES AND SPACES FOR THE CMS PROBLEM. IN EACH SEQUENCE, THE MIDDLE SPACE IS A FIBRE BUNDLE WHOSE FIBRE IS THE FIRST SPACE AND WHOSE BASE THE THIRD. THE PROJECTIONS ARE SHOWN BESIDE THE ARROW FROM THE FIBRE BUNDLE TO ITS BASE. THE VARIABLES Λ , r AND ρ APPEAR IN THE LEFT MOST SEQUENCE. THE REMAINING VARIABLES, BOTH REDUCED AND UNREDUCED, APPEAR AS DERIVATIVES OF THESE THREE.	27
5.1	A TABLE IS SHOWN THAT CLASSIFIES THE GEOMETRIC INTERPRETATION OF THE MOLECULAR STRAND EQUATIONS, ACCORDING TO LOCAL COORDINATES.	107
10.1	A TRAJECTORY IN Q FROM q_0 TO q_1 IS SHOWN THAT PROJECTS VIA π TO A TRAJECTORY IN Σ FROM ρ_0 TO ρ_1 . UN-REDUCTION POSES A BOUNDARY VALUE PROBLEM IN Q SUCH THAT THE PROJECTED TRAJECTORIES ARE SOLUTIONS TO THE EULER-LAGRANGE EQUATIONS IN Σ	168

*There will be time, there will be time
To prepare a face to meet the faces that you meet;
There will be time to murder and create,
And time for all the works and days of hands
That lift and drop a question on your plate;
Time for you and time for me,
And time yet for a hundred indecisions,
And for a hundred visions and revisions,
Before the taking of a toast and tea.*

from 'The Love Song of J. Alfred Prufrock' by T.S. Eliot

Part I

Charged molecular strand dynamics

CHAPTER 1 : INTRODUCTION

Part I reviews the undertakings of ELLIS ET AL. [2010], which provides a new continuum model and geometric framework for the dynamics of long molecules of charged units. The presentation also demonstrates how modern geometric mechanics discovers new abstract formulations through concrete calculations aimed at understanding important phenomena. Later, the strand problem motivates and elucidates the theory developed in Part II, and deeper understanding of the strand dynamics is gained through application of the new tools that become available.

1.1 Physical Setup

Long molecules are often modelled as strands of many individual charged units. Generally, the dynamics of such *charged molecular strands* (CMS) depends on both their local elastic deformations and the nonlocal (screened electrostatic) interactions of charged units across any loops in the molecule. These electrostatic interactions depend on the spatial distances and relative orientations between the individual charged units at different locations along the CMS.

One important approach to such a complex problem is a full molecular dynamics simulation, taking into account all (or most of) the forces between the atoms of the biological molecule as well as surrounding water molecules; so the nonlocal interactions appear naturally MORITA & KANEKO [2004]. This approach, whilst important for determining molecular properties, provides little insight for analytical understanding of the dynamics.

Many previous studies have also addressed the elastic dynamics of the charged strands using Kirchhoff's approach KIRCHHOFF [1859]. Historical reviews and citations of this approach may be found, for example, in DILL [1992]; DICHMANN ET AL. [1992]. A comprehensive survey of both the history and present state of the field can be found in ANTMAN [2004]. Recent advances using this approach, especially in the context of helical structures, appear in, e.g., GORIELY & TABOR

[1996]; GOLDSTEIN ET AL. [1998]; BALAEFF ET AL. [1999]; GOLDSTEIN ET AL. [2000]; HAUSRATH & GORIELY [2006]; NEUKIRCH ET AL. [2008].

Although many important results have been obtained by the traditional continuum theory approach, it has a limitation. Namely, the generalisation of the classical Kirchhoff theory to account for the torque caused by the long-range electrostatic interaction of molecules in different spatial locations along a flexible strand remains elusive, although the force due to electrostatic interaction has been captured by the traditional theory. See, for example, the article DICHMANN ET AL. [1992] which reviews progress in dynamical investigations of charged units distributed along a strand. In general, the lack of a consistent continuous model incorporating both torques and forces from electrostatic interactions has hampered analytical considerations; see for example BALAEFF ET AL. [1999] for additional discussion.

Using a geometric approach, Part I provides a continuum treatment of the CMS problem that includes both torques and forces arising from electrostatic interactions. This contribution to the analytical understanding of molecular strand dynamics, also provides a practical grounding from which to appreciate the value of the advances in geometric mechanics made in Part II.

It is important to note that even in the absence of a continuous model for nonlocal interactions, it is still possible to obtain *static* solutions using energy minimisation techniques. For example, interesting helical static solutions of pressed elastic tubes using interactions that prevent self-intersection were obtained in BANAVAR ET AL. [2007]. However, here we treat the *dynamical* problem, which does require a proper treatment of torques in long-range interactions.

The difficulty in computing the dynamical effects of torque due to long-range interactions among the molecular sub-units arises because the classical Kirchhoff theory is formulated in a frame moving with the strand, but deals with a mixture of variables, some measured in the fixed spatial frame and some in the body frame. Since the torque due to long-range interactions is applied at base points of a curve that is moving in space, it presents a particular difficulty for the mixed representations in the Kirchhoff theory. That is, the spatial Euclidean distances and relative orientations of the molecules must be reconstructed at each time step during the sinuous motion and twisting of the strand before any self-consistent computation can be made of the forces and torques due to long-range electrostatic interactions.

Introduction

In fact, even when electrostatic forces are not involved, the motion of realistic curves in space is inherently nonlocal, because of the requirement that the curve not cross itself during the dynamics. In the purely elastic Kirchhoff approach, such nonlocal considerations are neglected. Physically, however, self-intersections are prevented by the existence of a short-range potential (e.g., Lennard-Jones potential) that produces highly repulsive forces when two points along the curve approach each other. Thus, forces between segments of the strand that could be quite distant along its arc length are essential for the physical description of its dynamics.

Part I casts the CMS problem for an arbitrary inter-molecular potential into the *convective representation* of continuum dynamics introduced in HOLM ET AL. [1986] and applied in exact geometric rod theory in SIMÓ ET AL. [1988]. The spatial and convective representations of continuum dynamics are the analogues, respectively, of the spatial and body representations of rigid body dynamics on $SO(3)$, MARSDEN & RATIU [2002]; HOLM [2008A]. This analogy arises because the configuration spaces for continuum dynamics with micro-structure and for rigid bodies are both Lie groups. In both cases, spatial velocities are right-invariant vector fields, while the convective, or body, velocities are the corresponding left-invariant vector fields.

If the curve were rigidly fixed in space, and the attached molecules on this fixed curve were simply allowed to rotate freely at each position, the theory of motion based on nonlocal interaction between different molecules would be more straightforward. Of particular interest here is the work MEZIC [2006] where a single charge was attached at each point along a fixed filament by a rigid rod of constant length that was allowed to rotate in a transverse plane. These charges were allowed to interact locally with other nearby charges that were similarly attached to planar rotors of constant length mounted transversely to the fixed filament.

The model in MEZIC [2006] comprised a fixed base strand and rigid charge configurations described by $SO(2)$ (i.e., one rotor in each plane). This model is generalised here to allow flexible motion of the base strand (time-dependent bend, twist, writhe, and extension) while also including all the degrees of freedom of molecular orientation in $SO(3)$ excited during the process of, say, folding. According to this more general class of models, a long molecule is represented as a flexible filament or strand, along which are attached various different types of rigid conformations

of sub-molecules that may *swivel* relative to each other in three dimensions under their mutual interactions. The flexibility of the filament arises physically because the electrostatic interaction between any pair of these rigid conformations, either along the filament or across from one loop to another of its folds, is much weaker than the internal interactions that maintain the shape of an individual charged conformation. The application of the present model to DNA requires further research, however, because the framework presented here fails to model the primary feature of DNA – its unzipping MOAKHER & MADDOCKS [2005].

Aim of Part I

The primary aim of Part I is to formulate the dynamics of nonlocal interactions on a continuum strand carrying charged micro-structure by using Lie symmetry reduction in the convective representation. The new formulation of symmetry-reduced, nonlocal, convective strand dynamics raises many interesting and non-trivial issues for future research. Among these issues are the classification and stability analysis of equilibrium solutions, dynamics of conformational changes (folding/unfolding), modelling unzipping of a double strand, and formulation of computational approaches in the convective representation, all of which provide challenges for future research.

1.1.1 Outline of Part I

Part I considers rigid charge conformations (RCCs) that are mounted along a flexible moving filament. These RCCs are more complex than the planar pendula considered for a fixed base strand in MEZIC [2006]. They are mounted in orthonormal frames defined at each point along the strand. They are allowed to interact with each other via the standard method in the field of a nonlocal (e.g., screened electrostatic, or Lennard-Jones) potential.

Geometrically exact rod theory

The model for the motion of the filament derives from the *geometrically exact rod theory* of SIMÓ ET AL. [1988], which is expressed in the convective representation

Introduction

of continuum mechanics. The rotations of rigid charge conformations along the flexible filament are illustrated in Figure 1.1.

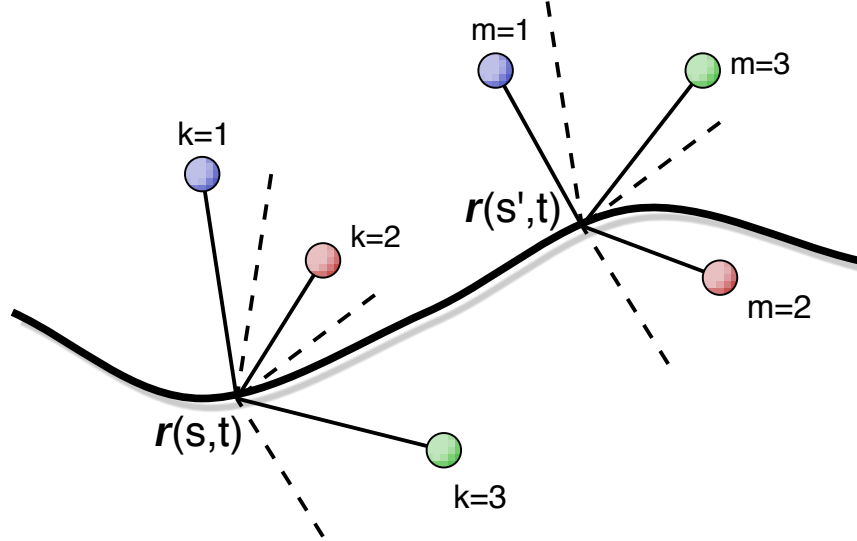


Figure 1.1: Rigid conformations of charges are distributed along a curve. Note that this is a spatial representation of the orientations of these conformations of charges.

These rigid conformations of multiple charges interact via a nonlocal effective many-body potential representing their screened electrostatic interactions. The nonlocal interactions among these RCCs depend on their separations in the ambient space and relative orientations, which are both allowed to evolve with the filament motion. Thus, the inertial motion of a pair of RCCs mounted at any two spatial points $\mathbf{r}(s, t)$ and $\mathbf{r}(s', t)$ along the filament is governed by an effective potential interaction energy that depends on their spatial separation and relative orientation. The filament is taken to be one-dimensional, although the orientations of the rigid charged conformations mounted along it are three-dimensional. A practical example to which the present filament approach would potentially apply is the vinylidene fluoride (VDF) oligomer NODA ET AL. [2003], which may be approximated by a strand carrying a dipole moment whose orientation is perpendicular to the axis of the strand. The VDF oligomer strand is approximately straight for small lengths, but it forms complex shapes due to electrostatic interactions for longer lengths. In the present framework, the undisturbed reference configuration of the VDF polymer will correspond to a straight elastic filament, along which a

rigid conformation of two opposite charges is positioned so that the dipole moment vector formed by those charges is perpendicular to the axis of the filament. The present work is limited to formulating the geometrically exact model of the charged molecular strand and studying its mathematical structure. Many challenges remain to be investigated in future research concerning the properties and solution behaviour of this model.

The dynamical influence of nonlocal electrostatic forces on rod mechanics is studied here using three approaches, including the Hamilton-Pontryagin variational method. Applying the $\text{Ad}_{(\Lambda, r)}^*$ transformation from convective to spatial variables in these equations streamlines their form and exposes the meaning of the interplay among their various local and nonlocal terms, relative to the Kirchhoff theory.

The convective formulation presented here applies equally well when the underlying substrate manifold (the filament here) becomes multi-dimensional; so this formulation would also be applicable to such problems as the motion of charged sheets, or charged elastically deformable media. Although we present part of the relevant geometry here, its applications in higher dimensions have been left for a later investigation.

Plan

Part I is written in two relatively independent, complementary phases that are meant to act as a ‘Rosetta stone’ for expressing applications of *Lagrangian reduction by symmetry* for CMS dynamics from the following two perspectives. The first phase consists of Chapter 2, in which the reduced dynamics of the charged strands are derived by means of ‘bare hands’ methods that use only variational principles and vector calculus. In contrast, the second phase, consisting of Chapter 3 and Chapter 4, contains a differential geometric perspective meant to elucidate the mathematical structure of the equations of motion derived in the first phase. With this two-pronged organisation, more abstract discussions embarked upon later in the thesis gain practical validation, since the abstract concepts are demonstrated to be inherent to a sophisticated understanding of applications rather than being imposed upon it. That is, the geometric tools naturally emerge from applications rather than being introduced to them by the mathematician.

§1.2 outlines the content of Part I in mathematical terms by giving an overview of the various spatial representations of filament dynamics discussed here from two different perspectives, known as the *classical* and *covariant* points of view.

§1.3 connects our results to the earlier literature. §1.3.1 will relate the theory presented here to the classical elastic rod approach pioneered by Kirchhoff. The need to keep track of spatial separations in long-range electrostatic interactions requires that we write the dynamics in either the spatial or convective representations, as opposed to using the Kirchhoff mixed representation. §1.3.2 will consider the simplified case when the orientations of the RCCs along the curve may depend on time, but the position of any point along the curve is fixed, thereby connecting to earlier work in MEZIC [2006].

Chapter 2 incorporates the flexible motion of the filament into the dynamics by using the geometrically exact rod theory given in SIMÓ ET AL. [1988]. The equations of motion are derived in convective form by using a Hamilton-Pontryagin approach, suitably modified to allow for nonlocal interactions. The strand equations in the convective representation will be formulated as conservation laws along the filament in §2.3.

In Chapter 3, §3.1 provides an introduction to affine Lie group actions in preparation for using them in §3.2, which explains the background for the affine Euler-Poincaré approach and applies it to the dynamics of charged strands.

Chapter 4 introduces a coordinate transformation that decouples the equations into their horizontal and vertical parts in a principal bundle framework. The geometric structure of this coordinate transformation is given by the covariant reduction theories found in Part II. A brief background is given and the covariant formulation of the CMS is obtained.

1.2 Mathematical Setup

1.2.1 Description of the variables involved

In the Lagrangian representation, the motion of a CMS is described by the variables $\Lambda(s, t) \in SO(3)$ and $\mathbf{r}(s, t) \in \mathbb{R}^3$. The vector $\mathbf{r}(s, t)$ is the spatial position of the filament and the variable $\Lambda(s, t)$ denotes the rotation of the RCC at the point s along the filament at time t . Here $s \in [0, L]$ is a parameter spanning a fixed in-

terval. The time and space derivatives yield the material velocity $(\dot{\Lambda}(s, t), \dot{\mathbf{r}}(s, t))$ and the angular and linear deformation gradients $(\Lambda'(s, t), \mathbf{r}'(s, t))$, where dot represents the partial derivative in the time variable, t , and prime denotes the partial derivative in, s , the space variable along the filament. Given Λ and $\mathbf{r}$, the *reduced variables* are denoted

$$\begin{aligned}\Omega &= \Lambda^{-1} \Lambda' && \in \mathfrak{so}(3) \\ \omega &= \Lambda^{-1} \dot{\Lambda} && \in \mathfrak{so}(3) \\ \Gamma &= \Lambda^{-1} \mathbf{r}' && \in \mathbb{R}^3 \\ \gamma &= \Lambda^{-1} \dot{\mathbf{r}} && \in \mathbb{R}^3 \\ \rho &= \Lambda^{-1} \mathbf{r} && \in \mathbb{R}^3.\end{aligned}\tag{1.1}$$

The physical interpretation of the variables (1.1) is as follows: The variable $\rho(s, t)$ represents the position of the filament in space as viewed by an observer who rotates with the RCC at (s, t) . The variables $(\Omega(s, t), \Gamma(s, t))$ describe the deformation gradients, and $(\omega(s, t), \gamma(s, t))$ describe the angular and linear velocities, as viewed by an observer who rotates with the RCC.

Definition 1.2.1 *The term ‘convective representation’ is the coordinate representation obtained by acting on the spatial quantities by the inverse of the spatial symmetry group component. For example, the reduced quantities (1.1) are in convective representation, since they arise when the spatial quantities, $(\Lambda, \mathbf{r})$ and their derivatives, are acted on by the inverse of the spatial $SO(3)$ component, Λ , i.e.*

$$\Lambda^{-1} \left(\Lambda, \mathbf{r}, \Lambda', \mathbf{r}', \dot{\Lambda}, \dot{\mathbf{r}} \right) = (e, \rho, \Omega, \Gamma, \omega, \gamma).$$

Remark 1.2.2 (Convective representation) The symmetry group component is always equal to the identity in the convective representation, and is usually omitted. The word ‘convective’ is derived from Latin, and means ‘tendency to go along with’. This terminology derives from the applications, where convective quantities describe the configuration of a system from the perspective of an observer that

Introduction

‘moves with’ part of it. For example, the convective quantities (1.1) describe the configuration of a CMS from the perspective of an observer who rotates with the RCC.

Remark 1.2.3 (Notation) Quantities defined using derivatives in s are denoted using capital Greek letters, whereas lower-case Greek letters (except for ρ) denote quantities whose definitions involve derivatives with respect to time. Bold letters, for example Γ , denote vectors in $\mathbb{R}^3$ whereas Ω is a 3×3 skew-symmetric matrix in the Lie algebra $\mathfrak{so}(3)$.

Definition 1.2.4 The *hat map* $\hat{\cdot} : (\mathbb{R}^3, \times) \rightarrow (\mathfrak{so}(3), [\cdot, \cdot])$ is the Lie algebra isomorphism given by $[\hat{\mathbf{u}}, \hat{\mathbf{v}}] = \mathbf{u} \times \mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$. Note that the infinitesimal action of $\hat{\mathbf{u}}$ on $\mathbf{b} \in \mathbb{R}^3$ is similarly given by $\hat{\mathbf{u}}\mathbf{b} = \mathbf{u} \times \mathbf{b}$. The hat map is just the adjoint representation of $\mathfrak{so}(3)$.

Thus, in an orthonormal basis of $\mathbb{R}^3$ and $\mathbf{u} \in \mathbb{R}^3$, the 3×3 skew-symmetric matrix $u := \hat{\mathbf{u}} \in \mathfrak{so}(3)$ has entries

$$u_{jk} = (\hat{\mathbf{u}})_{jk} = -\epsilon_{jkl} u^l. \quad (1.2)$$

Here the symbol ϵ_{jkl} with $j, k, l \in \{1, 2, 3\}$ denotes the totally skew-symmetric tensor density with $\epsilon_{123} = +1$ that defines the cross product of vectors in $\mathbb{R}^3$. In what follows, this notation is abbreviated by writing $\Omega := \hat{\Omega}$ and $\omega := \hat{\omega}$.

1.2.2 The classical perspective

The *classical perspective* of continuum dynamics derives the equations of motion by applying a process of reduction by symmetry to a system defined on a tangent bundle TQ of a configuration manifold Q , and governed by the *Euler-Lagrange equations* for a Lagrangian $L(q, \dot{q})$,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0.$$

Here the Lagrangian function $L : TQ \rightarrow \mathbb{R}$ is invariant under the action of the symmetry group.

Often an equivalent Hamiltonian formulation may be described. At the unreduced level, a Hamiltonian $H : T^*Q \rightarrow \mathbb{R}$ is defined on the cotangent bundle T^*Q , which is equipped with the canonical symplectic form. The equations of motion are then governed by the *canonical Hamilton's equations* for Hamiltonian $H(q, p)$,

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}.$$

Analogous reduction by symmetry techniques may then be applied to the Hamiltonian side. Our present investigation remains on the Lagrangian side. For more information about Hamiltonian mechanics and reduction techniques, see HOLM [2008A,B]; MARSDEN & RATIU [2002]; MARSDEN ET AL. [2007].

The classical approach has been extensively studied for fluids, see for example MARSDEN ET AL. [1984] for the Hamiltonian description and HOLM ET AL. [1998] for the Lagrangian side. In hydrodynamics, the configuration manifold $Q = G \times V^*$ is the product of a Lie group G and the dual of a representation space V on which the group acts linearly as $G \times V \rightarrow V$. The dual space, V^* , is the space of linearly *advected quantities* such as the mass density or the magnetic field in the reference configuration.

Definition 1.2.5 *An **advected quantity** is an element of the dual of a representation space whose motion arises solely from the action of the symmetry group.*

The associated process of reduction by symmetry under the action of G is called *Lagrangian reduction for semidirect products* HOLM ET AL. [1998]. Such systems (in the left version), employ the relations

$$\begin{aligned} \xi(t) &= g(t)^{-1} \dot{g}(t) \\ a(t) &= g(t)^{-1} a_{ref} \end{aligned} \tag{1.3}$$

where $g(t) \in G$ is the Lagrangian motion, $\xi(t)$ is the convective velocity, and $a(t) \in V^*$ is the evolution of the advected quantity for a given initial condition a_{ref} . Note that $a(t)$ is also a convective quantity, since it is acted upon by $g(t)^{-1}$ where $g(t)$ is the spatial symmetry group component.

For the molecular strand, $g = (\Lambda, \mathbf{r})$ and $a = (\Omega, \Gamma, \rho)$. However, the relations (1.1) cannot be recovered from (1.3) because the variables (Ω, Γ, ρ) are not linearly

Introduction

advected. Consequently, a generalisation of (1.3) is needed in which $g \in G$ acts on $a \in V^*$ by an *affine action*. Such a generalisation is given by the process of *affine Euler-Poincaré reduction* developed in the context of complex fluids in GAY-BALMAZ & RATIU [2008]. This theory, which will be reviewed in §3.2, produces the relations

$$\begin{aligned}\xi(t) &= g(t)^{-1} \dot{g}(t), \\ a(t) &= g(t)^{-1} a_{ref} + c(g(t)^{-1}),\end{aligned}\tag{1.4}$$

where the additional term c is a *group one-cocycle*, which is defined as follows.

Definition 1.2.6 (Group one-cocycle) *A map $c : G \rightarrow V^*$ is a group one-cocycle if and only if it satisfies the property*

$$c(fg) = c(f) + fc(g),$$

where f acts on $c(g)$ by a left representation.

Setting $a_{ref} = 0$, the advected quantity evolves in time as

$$a(t) = c(g(t)^{-1}).$$

Remarkably, the evolution of (Ω, Γ, ρ) in (1.1) is precisely of this form for a well chosen cocycle. The variables $(\Lambda, \mathbf{r})(s, t)$ are interpreted as time-parameterised curves in the infinite dimensional Lie group

$$G = \mathcal{F}([0, L], SE(3))$$

of all smooth functions on $[0, L]$ taking values in $SE(3)$. Here, $SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$ is the special Euclidean group, comprising the semidirect product action of three-dimensional rotations $SO(3)$ and translations $\mathbb{R}^3$. Further details may be found in Chapter 2.

Remark 1.2.7 The variables

$$(\omega, \gamma) = (\Lambda, \mathbf{r})^{-1}(\dot{\Lambda}, \dot{\mathbf{r}}),$$

their associated momenta

$$(\pi, \mu) := \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right)$$

for a reduced Lagrangian $l = l(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho})$, and the affine advected variables $(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho})$ are all convective quantities; see MARS DEN & RATIU [2002]. In this context, convective quantities may also be called *body quantities*, since they are defined in a frame following the motion of the molecular strand.

In contrast, the variables

$$(\boldsymbol{\pi}^{(S)}, \mathbf{p}^{(S)}) := \text{Ad}_{(\boldsymbol{\Lambda}, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) \quad (1.5)$$

are spatial quantities, i.e., they are defined at fixed points in Euclidean space.

1.2.3 The covariant perspective

The *covariant perspective* considers all independent variables on an equal footing by interpreting the Lagrangian quantities $(\boldsymbol{\Lambda}(s, t), \mathbf{r}(s, t))$ as a map

$$(s, t) \mapsto (\boldsymbol{\Lambda}(s, t), \mathbf{r}(s, t)), \quad (1.6)$$

from space-time into the finite dimensional configuration space $SE(3)$. This contrasts with the special interpretation given to the time variable t by the classical perspective, which considers a curve

$$t \mapsto (\boldsymbol{\Lambda}(\cdot, t), \mathbf{r}(\cdot, t))$$

in the infinite dimensional configuration space $\mathcal{F}([0, L], SE(3))$.

More precisely, the covariant perspective treats (s, t) as coordinates on a space-time manifold $X := [0, L] \times \mathbb{R}$. Note that this is exactly the point of view taken in field theories. The map (1.6) is regarded as a section of the trivial fibre bundle

$$\pi_{X,P} : P = X \times SE(3) \rightarrow X, \quad \pi_{X,P}(x, \boldsymbol{\Lambda}, \mathbf{r}) = x,$$

over space-time $X \ni (s, t) =: x$. Indeed, by definition, a section σ of a bundle $\pi_{X,P}$ is a smooth map $\sigma : X \rightarrow P$ verifying the property $\pi_{X,P} \circ \sigma = \text{id}_X$. Thus, since the bundle is trivial, a section reads

$$\sigma(x) = (x, \boldsymbol{\Lambda}(x), \mathbf{r}(x)).$$

Introduction

From the general covariant perspective, the spatial quantities are formulated as components of the *first jet extension* $j^1\sigma(x) := T_x\sigma$ of the section σ , where $T\sigma : TX \rightarrow TP$ denotes the tangent map of σ . Specifically, for the molecular strand,

$$j^1\sigma(x) \cong (\Lambda(x), \mathbf{r}(x), \Lambda'(x)ds + \dot{\Lambda}(x)dt, \mathbf{r}'(x)ds + \dot{\mathbf{r}}(x)dt). \quad (1.7)$$

The first jet extension $j^1\sigma$ is a section of a bundle over X called the *first jet bundle* $J^1P \rightarrow X$.

The Lagrangian map, $L : TQ \rightarrow \mathbb{R}$, of the classical perspective is replaced with a Lagrangian density

$$\mathcal{L} : J^1P \rightarrow \text{Den}(X).$$

In the case of the molecular strand, the Lagrangian density is defined according to $\mathcal{L}(j^1\sigma) = L(\Sigma, \mathbf{r}, \Lambda', \mathbf{r}', \dot{\Lambda}, \dot{\mathbf{r}}) ds dt$. The dynamics are given by the *covariant Euler-Lagrange equations* (7.1), which describe solutions to the covariant formulation of the variational principle,

$$\delta \int_X \mathcal{L} = 0.$$

For the molecular strand, $\mathcal{L}$ is $SO(3)$ -invariant, thus the first jet extension (1.7) induces the section

$$(\Lambda^{-1}\mathbf{r}, \Lambda^{-1}\Lambda' ds + \Lambda^{-1}\dot{\Lambda} dt, \Lambda^{-1}\mathbf{r}' ds + \Lambda^{-1}\dot{\mathbf{r}} dt)$$

by $SO(3)$ -reduction, recovering the reduced variables

$$(\boldsymbol{\rho}, \boldsymbol{\Omega} ds + \boldsymbol{\omega} dt, \boldsymbol{\Gamma} ds + \boldsymbol{\gamma} dt) \quad (1.8)$$

defined in (1.1). Reduction by the group $SO(3)$ yields a principal bundle structure on P given by

$$\pi_{\Sigma, P} : P \rightarrow \Sigma := P/SO(3) = X \times \mathbb{R}^3, \quad \pi_{\Sigma, P}(x, \Lambda, \mathbf{r}) = (x, \Lambda^{-1}\mathbf{r}).$$

This bundle over Σ should not be confused with the configuration bundle $\pi_{X, P} : P \rightarrow X$ that has the same total space, but a different base. The precise geometric setting underlying this covariant reduction process is explained in detail in

Chapters 4 and 6.

Figure 1.2 illustrates the relationship between the different variables and spaces of the covariant perspective of the CMS problem.

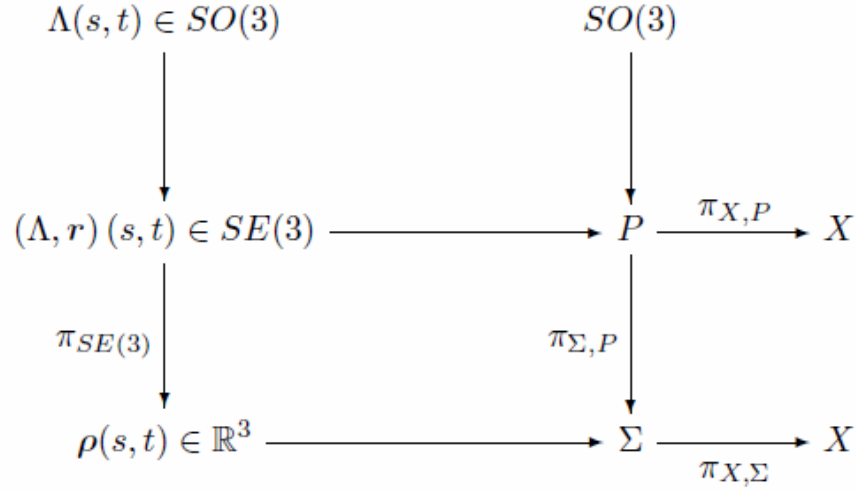


Figure 1.2: A diagram that shows the relationship between variables and spaces for the CMS problem. In each sequence, the middle space is a fibre bundle whose fibre is the first space and whose base is the third. The projections are shown beside the arrow from the fibre bundle to its base. The variables Λ , r and ρ appear in the left most sequence. The remaining variables, both reduced and unreduced, appear as derivatives of these three.

1.3 Connection to previous studies

1.3.1 Kirchhoff equations for pure elastic rod

The results of Part I in the convective representation may be compared with the classical Kirchhoff theory of the purely elastic rod, particularly in terms of the available conservation laws ANTMAN [2004]. This comparison was presented in SIMÓ ET AL. [1988] for the purely elastic case, when the reduced Lagrangian l is an explicit (local) function of the reduced variables $l = l(\omega, \gamma, \Omega, \Gamma, \rho)$ defined in equation (1.1). The work of SIMÓ ET AL. [1988] is extended in Part I to take account of *nonlocal* interactions.

The balance laws for angular and linear momenta are of particular interest. The notation of DICHMANN ET AL. [1992] has been used for this comparison. For sim-

Introduction

plicity, the position $\mathbf{r}(s)$ along the filament is assumed to be given by the arc length s . This assumption conveniently avoids extra factors of $|\Gamma(s)|$ in the expressions.

In order to connect to the Kirchhoff theory, fix $\Lambda(s) \in SO(3)$ as the transformation matrix from the fixed orthonormal basis $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$ of Euclidean space to the orthonormal basis of *directors* $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$ describing the orientation of the filament (see Figure 1.1). That is,

$$\mathbf{d}_i(s) = \Lambda_i^k(s) \mathbf{E}_k, \quad i = 1, 2, 3. \quad (1.9)$$

There is some ambiguity in the choice of the basis $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$ at each given point. The most popular selection of the basis is governed by the so-called ‘natural frame’. Further details of the natural frame are found in DICHMANN ET AL. [1992]. In principle, this particular choice of Λ need not have been taken, since for rigid charge conformations (RCC) the *relative* configuration of charges is conserved by the dynamics, and the configuration of an RCC state at each point s is completely described by a pair $(\Lambda, \mathbf{r}) \in SE(3)$. Taking Λ to be a different representation of an RCC would lead to a transformation $\Lambda(s, t) \mapsto A\Lambda(s, t)$ where $A \in SO(3)$ is a fixed factor. While the present description would be equivalent with such a factor, the explicit relation to Kirchhoff formulae would become less obvious.

Remark 1.3.1 Note that if the charge conformations were allowed to deform, then Λ would no longer be an element of $SO(3)$. Instead, the charge conformation would be described by a general invertible matrix $\Lambda \in GL(3)$ and a vector $\mathbf{r} \in \mathbb{R}^3$. No explicit relation to Kirchhoff’s formulae would then be possible.

As mentioned in §1.1, Kirchhoff’s approach precludes any simple computation of Euclidean distances between the charges, unless the spatial length-scale of the rigid charge conformations (RCCs) holding the charges at given point $\boldsymbol{\eta}_k(s)$ is negligible. It is interesting that in the more complex case considered here, the equations become formally equivalent to Kirchhoff’s equations, provided the effects of nonlocality are computed appropriately. In particular, there is an appropriate mapping from the convective representation to the Kirchhoff representation, as well as some identities connecting nonlocal contributions with the *total* derivatives of the Lagrangian. This mapping will be discussed in more detail in §2.3.

After these preliminaries, we are ready for a detailed comparison with Kirchhoff’s theory. If $\rho_d(s)$ is the local mass density of the rod, then the kinetic energy

due to linear motion K_{lin} is given by

$$K_{lin} = \frac{1}{2} \int \rho_d(s) \|\dot{\mathbf{r}}(s)\|^2 ds = \frac{1}{2} \int \rho_d(s) \|\Lambda^{-1} \dot{\mathbf{r}}(s)\|^2 ds = \frac{1}{2} \int \rho_d(s) \|\boldsymbol{\gamma}(s)\|^2 ds.$$

Consequently, the linear momentum densities $\mathbf{p} := \delta K_{lin} / \delta \dot{\mathbf{r}}$ and $\delta K_{lin} / \delta \boldsymbol{\gamma}$ are related by

$$\mathbf{p} = \rho_d \dot{\mathbf{r}} = \Lambda \rho_d \boldsymbol{\gamma} = \Lambda \frac{\delta K_{lin}}{\delta \boldsymbol{\gamma}}. \quad (1.10)$$

A point on a rod in Kirchhoff's theory is parameterised by the distance $\mathbf{r}(s, t)$ measured from a fixed point in space. The i -th component of the local angular momentum in the body frame $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$ is defined by $\pi^i(s) := \mathbb{I}_j^i(s) \omega^j(s)$, where $\omega^j(s)$ is the j -th component of body angular velocity given by $\hat{\boldsymbol{\omega}}(s) := \boldsymbol{\omega}(s) = \Lambda(s)^{-1} \dot{\Lambda}(s)$, and $\mathbb{I}_j^i(s)$ is the local value of the inertia tensor. Note that the inertia tensor $\mathbb{I}(s)$ expressed in body coordinates is time-independent. Thus, the local kinetic energy due to rotation is given by

$$K_{rot} = \frac{1}{2} \int \boldsymbol{\omega}(s) \cdot \mathbb{I}(s) \boldsymbol{\omega}(s) ds,$$

and the local body angular momentum density is given by

$$\boldsymbol{\pi} := \frac{\delta K_{rot}}{\delta \boldsymbol{\omega}} = \mathbb{I} \boldsymbol{\omega}.$$

To write the conservation laws, the angular momentum needs to be expressed in the fixed spatial frame $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$. The spatial angular momentum shall be denoted $\boldsymbol{\pi}^{(\mathbf{E})}$ to distinguish it from $\boldsymbol{\pi}$, which was expressed in the body frame $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$. This convention is also used for all other vectors. Equation (1.9) yields

$$\boldsymbol{\pi}(s) = \pi^i(s) \mathbf{d}_i(s) = \mathbb{I}_j^i(s) \omega^j(s) \mathbf{d}_i(s) = \mathbb{I}_j^i(s) \omega^j(s) \Lambda_i^k(s) \mathbf{E}_k = \boldsymbol{\pi}^{(\mathbf{E}),k}(s) \mathbf{E}_k,$$

so the k -th component of the spatial angular momentum is expressed in terms of the local body quantities $\mathbb{I}_j^i(s)$ and $\omega^k(s)$ as

$$\boldsymbol{\pi}^{(\mathbf{E}),k} = \Lambda_i^k \mathbb{I}_j^i \omega^j = [\Lambda \mathbb{I} \boldsymbol{\omega}]^k = \left[\Lambda \frac{\delta K_{rot}}{\delta \boldsymbol{\omega}} \right]^k. \quad (1.11)$$

Introduction

Thus, the vector $\pi^{(\mathbf{E})}(s)$ of *body* angular momentum expressed in the *spatial* frame is related to the local body quantities as

$$\pi^{(\mathbf{E})} = \Lambda \mathbb{I} \omega = \Lambda \frac{\delta K_{rot}}{\delta \omega}. \quad (1.12)$$

Similarly, the expression for $\mathbf{p}^{(\mathbf{E})}$ is

$$\mathbf{p}^{(\mathbf{E})} = \Lambda \frac{\delta l}{\delta \gamma} = \mathbf{p}.$$

Remark 1.3.2 The vector $\pi^{(\mathbf{E})}$ and all other vectors with the superscript $(\mathbf{E})$ do not have the physical meaning of the angular momentum in the fixed frame. The true angular and linear momenta in the spatial frame will be denoted (see immediately below) with the superscript (S) . The quantities with the superscript $(\mathbf{E})$ are just the transformations of vectors with respect to rotation of the base frame. No confusion should arise over this distinction.

In general, it may be assumed for physical reasons that the Lagrangian in Kirchhoff's formulation has the form

$$l(\omega, \gamma, \Omega, \Gamma) = K_{lin}(\gamma) + K_{rot}(\omega) - E(\Omega, \Gamma), \quad (1.13)$$

where $E(\Omega, \Gamma)$ is a certain explicit function of Ω and Γ (not necessarily quadratic). For Lagrangian (1.13), the body forces $\mathbf{n} = -\delta l / \delta \Gamma$ and torques $\mathbf{m} = -\delta l / \delta \Omega$ are similarly related to the transformed quantities $\mathbf{n}^{(\mathbf{E})}, \mathbf{m}^{(\mathbf{E})}$ in Kirchhoff's theory as

$$\mathbf{n}^{(\mathbf{E})} = -\Lambda \frac{\delta l}{\delta \Gamma}, \quad \mathbf{m}^{(\mathbf{E})} = -\Lambda \frac{\delta l}{\delta \Omega}. \quad (1.14)$$

Next, formula (1.5) may be used to transfer to spatial frame. Identifying elements of $\mathfrak{se}^*(3)$ with pairs of vectors $(\mu, \eta) \in \mathbb{R}^3$, produces a useful formula MARS DEN & RATIU [2002]; HOLM [2008B] for the coadjoint action $SE(3) \times \mathfrak{se}^*(3) \rightarrow \mathfrak{se}^*(3)$, expressed in terms of vector cross products in $\mathbb{R}^3$,

$$\text{Ad}_{(\Lambda, r)}^* (\mu, \eta) = (\Lambda \mu + r \times \Lambda \eta, \Lambda \eta). \quad (1.15)$$

Accordingly, the spatial momenta – denoted by a superscript (S) – become

$$\begin{aligned} (\boldsymbol{\pi}^{(S)}, \mathbf{p}^{(S)}) &:= \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) = \left(\Lambda \frac{\delta l}{\delta \boldsymbol{\omega}} + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \boldsymbol{\gamma}}, \Lambda \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) \\ &= (\boldsymbol{\pi}^{(\mathbf{E})} + \mathbf{r} \times \mathbf{p}^{(\mathbf{E})}, \mathbf{p}^{(\mathbf{E})}), \end{aligned} \quad (1.16)$$

upon using (1.10) and (1.12). Analogously, using (1.14), the spatial torques $\mathbf{m}^{(S)}$ and forces $\mathbf{n}^{(S)}$ are expressed as

$$\begin{aligned} (\mathbf{m}^{(S)}, \mathbf{n}^{(S)}) &:= \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(-\frac{\delta l}{\delta \boldsymbol{\Omega}}, -\frac{\delta l}{\delta \boldsymbol{\Gamma}} \right) = \left(-\Lambda \frac{\delta l}{\delta \boldsymbol{\Omega}} - \mathbf{r} \times \Lambda \frac{\delta l}{\delta \boldsymbol{\Gamma}}, -\Lambda \frac{\delta l}{\delta \boldsymbol{\Gamma}} \right) \\ &= (\mathbf{m}^{(\mathbf{E})} + \mathbf{r} \times \mathbf{n}^{(\mathbf{E})}, \mathbf{n}^{(\mathbf{E})}). \end{aligned} \quad (1.17)$$

The conservation laws in the Kirchhoff theory may now be written as

$$\frac{\partial}{\partial t} (\boldsymbol{\pi}^{(S)}, \mathbf{p}^{(S)}) = \frac{\partial}{\partial s} (\mathbf{m}^{(S)}, \mathbf{n}^{(S)}) + (\mathbf{T}, \mathbf{f}), \quad (1.18)$$

where $\mathbf{T}$ and $\mathbf{f}$ are external torques and forces, respectively. Equations (1.18) give, component-wise, the following linear and angular momentum conservation laws (cf. equations (2.5.5) and (2.5.7) of DICHMANN ET AL. [1992])

$$\frac{\partial}{\partial t} \mathbf{p}^{(\mathbf{E})} = \frac{\partial}{\partial s} (\mathbf{n}^{(\mathbf{E})} + \mathbf{F}) \quad (1.19)$$

$$\frac{\partial}{\partial t} (\boldsymbol{\pi}^{(\mathbf{E})} + \mathbf{r} \times \mathbf{p}^{(\mathbf{E})}) = \frac{\partial}{\partial s} (\mathbf{m}^{(\mathbf{E})} + \mathbf{r} \times \mathbf{n}^{(\mathbf{E})} + \mathbf{L}) \quad (1.20)$$

where $\mathbf{F}$ and $\mathbf{L}$ are defined as the indefinite integrals,

$$\mathbf{F} = \int^s \mathbf{f}(q) \, dq \quad \text{and} \quad \mathbf{L} = \int^s [\mathbf{r}(q) \times \mathbf{f}(q) + \mathbf{T}(q)] \, dq.$$

Opening the brackets in (1.19) and (1.20) gives the balances of linear and angular momenta in Kirchhoff's approach (cf. eqs. (2.3.5) and (2.3.6) of DICHMANN ET AL. [1992])

$$\frac{\partial \mathbf{p}^{(\mathbf{E})}}{\partial t} = \frac{\partial \mathbf{n}^{(\mathbf{E})}}{\partial s} + \mathbf{f} \quad (1.21)$$

$$\frac{\partial \boldsymbol{\pi}^{(\mathbf{E})}}{\partial t} = \frac{\partial \mathbf{m}^{(\mathbf{E})}}{\partial s} + \frac{\partial \mathbf{r}}{\partial s} \times \mathbf{n}^{(\mathbf{E})} + \mathbf{T}. \quad (1.22)$$

Introduction

To see how these Kirchhoff balance laws look in our convective representation, relations (1.16) and (1.17) may be substituted into (1.18) to obtain:

$$\frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) \right] + \frac{\partial}{\partial s} \left[\text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\Omega}}, \frac{\delta l}{\delta \boldsymbol{\Gamma}} \right) \right] = (\mathbf{T}, \mathbf{f}). \quad (1.23)$$

Equation (1.23) formulates the Kirchhoff model as a conservation law, which relates with the conservation law form of the CMS equations derived in §2.3. Expanding (1.23) gives the convective equations of motion,

$$(\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\omega}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta l}{\delta \boldsymbol{\Omega}} + \boldsymbol{\Gamma} \times \frac{\delta l}{\delta \boldsymbol{\Gamma}} = \Lambda^{-1} \mathbf{T} \quad (1.24)$$

$$(\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\gamma}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta l}{\delta \boldsymbol{\Gamma}} = \Lambda^{-1} \mathbf{f}. \quad (1.25)$$

Remark 1.3.3 In the case where $\mathbf{f} = \mathbf{T} = 0$, equations (1.24)–(1.25) are $SE(3)$ invariant. This may be verified since the Lagrangian (1.13) only depends on the variables $(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma})$ as defined by (1.1), and not on $\boldsymbol{\rho}$. Setting $g(s, t) = (\Lambda, \mathbf{r})(s, t) \in SE(3)$ leads to the quantities $g^{-1}\dot{g} = (\hat{\boldsymbol{\omega}}, \boldsymbol{\gamma})$ and $g^{-1}g' = (\hat{\boldsymbol{\Omega}}, \boldsymbol{\Gamma})$, both of which are $SE(3)$ invariant. In this context, dependence of the Lagrangian (1.13) on the variable $\boldsymbol{\rho}$ may be interpreted as symmetry breaking to an $SO(3)$ invariant system. It is more natural to then consider the $SE(3)$ as a manifold that is acted upon by the symmetry group $SO(3)$. This perspective leads to the invariant quantities $\Lambda^{-1}g = (e, \boldsymbol{\rho})$, $\Lambda^{-1}\dot{g} = (\hat{\boldsymbol{\omega}}, \boldsymbol{\gamma})$, and $\Lambda^{-1}g' = (\hat{\boldsymbol{\Omega}}, \boldsymbol{\Gamma})$.

The case of radially symmetric, potential external torques and forces is achieved by modifying the Lagrangian l to depend explicitly on the additional variable $\boldsymbol{\rho} = \Lambda^{-1}\mathbf{r}$. As will be shown in §2.3, the external torques $\mathbf{T}$ and forces $\mathbf{f}$ are given by

$$(\mathbf{T}, \mathbf{f}) = \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho}, \frac{\delta l}{\delta \boldsymbol{\rho}} \right). \quad (1.26)$$

Using formula (1.15), relationship (1.26) expands, then simplifies to

$$\begin{aligned} \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho}, \frac{\delta l}{\delta \boldsymbol{\rho}} \right) &= \left(\Lambda \left(\frac{\delta l}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho} \right) + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \boldsymbol{\rho}}, \Lambda \frac{\delta l}{\delta \boldsymbol{\rho}} \right) \\ &= \left(\left(\Lambda \frac{\delta l}{\delta \boldsymbol{\rho}} \right) \times (\Lambda \boldsymbol{\rho}) + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \boldsymbol{\rho}}, \Lambda \frac{\delta l}{\delta \boldsymbol{\rho}} \right) = \left(0, \Lambda \frac{\delta l}{\delta \boldsymbol{\rho}} \right), \end{aligned} \quad (1.27)$$

upon recalling that $\Lambda \rho = r$.

Remark 1.3.4 (Potential external forces produce no net torque) As (1.27) shows, radially symmetric potential external forces produce no net torque on the strand. This occurs because the external torque's effect on the directors is precisely offset by its effect on the centreline of the strand. Hence, the nonzero torques T in (1.23) must arise from non-potential or radially asymmetric forces.

Remark 1.3.5 (Reduction of static equations to the heavy top) An analogy exists between the stationary shapes of an elastic filament and the equations of motion of a heavy top, KEHRBAUM & MADDOCKS [1997] and NIZETTE & GORIELY [1999]. This shows that the geometric approach also applies to the problem of determining the steady equilibrium solutions of filament dynamics. The present discussion focuses, however, on the derivations and geometric structures underlying the dynamical equations, rather than on the solutions of the equations.

The conservation law (1.23) is formally equivalent to the classical expressions in (1.19) and (1.20), even if nonlocal interaction is present. This equivalence shows how the classical results (1.19) and (1.20) generalise for the case of nonlocal interactions. Clearly, the conservation laws are simpler in the Kirchhoff representation. However, if nonlocal interactions are present (called *self-interaction forces* in DICHMANN ET AL. [1992]), the computation of the required time-dependent Euclidean distances in the interaction energy becomes problematic in the classical Kirchhoff approach. As we shall see in §2.3, these conservation laws may be obtained, even when nonlocal interactions are present. Also in §2.3, the nonlocal forces will be shown to be included in the conservation law (1.23) and are expressed in the same form as a purely elastic conservation law.

The balance laws (1.19) and (1.20) are much simpler in appearance than the expressions in (1.23), as they do not involve computing (Λ, r) at each instant in time and point in space. Thus, for elastic rods in the absence of nonlocal interactions, the Kirchhoff mixed (convective-spatial) representation appears simpler than in either the convective or spatial representations. However, the presence of nonlocal terms summons the more general convective approach introduced for this problem in HOLM & PUTKARADZE [2009].

The Kirchhoff equations discussed here are governed purely by local elastic and external forces. In particular, no consideration of nonlocal self-interactions of

the filament takes place. In the sequel, §1.3.2, the filament remains fixed in time, however nonlocal interactions between RCCs that remain free to rotate in $SO(3)$ are considered. Therefore, the sequel demonstrates the relationship between the CMS dynamics as will be derived in Chapter 2 and MEZIC [2006].

1.3.2 Reductions for a rigid filament

We begin by applying the CMS model developed here to the particular case of a rigid filament, in order to compare the motion equations with those arising in MEZIC [2006]. The rigid filament case corresponds to the situation that arises when the filament is both inelastic and stationary. Such a restriction eliminates the variables Γ and γ in (1.1), since derivatives in $\mathbf{r}$ are assumed to vanish. The RCC, however, are still allowed to rotate by a general $SO(3)$ group element, and elastic energy may still be present, despite the filament being inelastic, due to twisting of the RCC. Such a system is a generalisation of that presented in MEZIC [2006] from $SO(2)$ to $SO(3)$ rotations.

The analysis of filament dynamics driven by nonlocal interactions simplifies in the case when the position of the filament is fixed as $\mathbf{r}(s)$ and does not depend on time. For simplicity, assume that the filament is straight and s is the arc length, so that $\mathbf{r}(s) = (s, 0, 0)^T$. The following reduced Lagrangian is invariant under the left action of the Lie group $SO(3)$:

$$l = \underbrace{\frac{1}{2} \int \boldsymbol{\omega}(s) \cdot \mathbb{I}(s) \boldsymbol{\omega}(s) ds}_{\text{Kinetic energy}} - \underbrace{\frac{1}{2} \int f(\boldsymbol{\Omega}(s)) ds}_{\text{Elastic energy}} - \underbrace{\iint U(\xi(s, s'), \boldsymbol{\kappa}(s, s')) ds ds'}_{\text{Potential energy}}. \quad (1.28)$$

A nonlocal interaction term appears in the potential energy in this Lagrangian. This nonlocal term involves the variables

$$\xi(s, s') = \Lambda^{-1}(s) \Lambda(s') \in SO(3) \quad \text{and} \quad \boldsymbol{\kappa}(s, s') = \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s))$$

which respectively define the relative orientation and relative separation of rigid charge conformations at two *different points* in space. The variables $(\xi, \boldsymbol{\kappa})(s, s') \in SE(3)$ are invariant with respect to simultaneous rotations of the coordinate frames for s and s' , but are not an elements of a Lie algebra. In particular, the two-

point relative orientation $\xi(s, s')$ is not a vector, but a rotation in $SO(3)$. The presence of nonlocal interactions introduces dependence on these relative orientations. The new dependence produces new types of nonlocal terms in the corresponding Euler-Poincaré dynamics, which are obtained by applying reduction by $SO(3)$ symmetry to Hamilton's principle. It is interesting to compare the role of ξ for the molecular strand with a similar quantity defined in CENDRA & MARSDEN [2005], which was used to model the relative orientations of two asteroids.

Euler-Poincaré dynamics

Euler-Poincaré dynamics for the angular dynamics on a fixed filament follows from stationarity of the left invariant action

$$S = \int l(\boldsymbol{\omega}, \boldsymbol{\Omega}, \xi, \boldsymbol{\kappa}) dt.$$

Note that this case does not require computation of the evolution equation for γ , since the filament is assumed to be fixed in space, i.e., $\gamma = \Lambda^{-1}\dot{\mathbf{r}} = 0$.

The variational derivative δS for such a Lagrangian is computed as

$$\delta S = \int \left\langle \frac{\delta l}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle + \left\langle \frac{\delta l}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle + \left\langle \xi^{-1} \frac{\delta l}{\delta \xi}, \xi^{-1} \delta \xi \right\rangle + \left\langle \frac{\delta l}{\delta \boldsymbol{\kappa}}, \delta \boldsymbol{\kappa} \right\rangle dt. \quad (1.29)$$

As Lemmas 2.2.2 and 2.2.3 will show, the variations of ω , Ω , ξ and κ are related to Σ by

$$\begin{aligned} \delta \boldsymbol{\omega} &= \dot{\Sigma} + [\boldsymbol{\omega}, \Sigma] = \dot{\Sigma} + \text{ad}_{\boldsymbol{\omega}} \Sigma \\ \delta \boldsymbol{\Omega} &= \Sigma' + [\boldsymbol{\Omega}, \Sigma] = \Sigma' + \text{ad}_{\boldsymbol{\Omega}} \Sigma \\ \xi^{-1}(s, s') \delta \xi(s, s') &= \Sigma(s') - \text{Ad}_{\xi^{-1}(s, s')} \Sigma(s) \\ \delta \boldsymbol{\kappa} &= -\boldsymbol{\Sigma} \times \boldsymbol{\kappa}, \end{aligned}$$

where the notation $\Sigma := \Lambda^{-1} \delta \Lambda$ has been introduced, and the identity $\delta \mathbf{r} = 0$ has been used to simplify the form of $\delta \boldsymbol{\kappa}$ from Lemma 2.2.3 to the fixed filament case. Substituting the variation formulae into (1.29), then integrating by parts in the

Introduction

time t and one-dimensional coordinate s along the fibre, yields

$$\begin{aligned} \delta S = \int \left\langle -(\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} - (\partial_s - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} - \int (\text{d}_2 U((\xi, \kappa)(s, s')) \times \kappa(s, s'))^\wedge \text{d}s' \right. \\ \left. - \int (\xi(s, s') \text{d}_1 U((\xi, \kappa)^{-1}(s, s')) - \text{d}_1 U((\xi, \kappa)(s, s')) \xi^{-1}(s, s')) \text{d}s', \Sigma \right\rangle \text{d}t, \end{aligned} \quad (1.30)$$

where d_1 denotes differentiation with respect to the $SO(3)$ component, and d_2 with respect to the $\mathbb{R}^3$ component. Thus, Hamilton's principle $\delta S = 0$ implies the *Euler-Poincaré equations*,

$$\begin{aligned} (\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\partial_s - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} = - \int (\text{d}_2 U((\xi, \kappa)(s, s')) \times \kappa(s, s'))^\wedge \text{d}s' \\ - \int (\xi(s, s') \text{d}_1 U((\xi, \kappa)^{-1}(s, s')) - \text{d}_1 U((\xi, \kappa)(s, s')) \xi^{-1}(s, s')) \text{d}s'. \end{aligned} \quad (1.31)$$

Note that these Euler-Poincaré equations are nonlocal. That is, they are integral-partial-differential equations.

Remark 1.3.6 Note that $\text{d}_1 U(\xi, \kappa) \in T_\xi^* SO(3)$ is a cotangent vector above $\xi \in SO(3)$. Therefore, the terms $\xi \text{d}_1 U((\xi, \kappa)^{-1})$ and $\text{d}_1 U((\xi, \kappa)) \xi^{-1}$ both live in $\mathfrak{so}^*(3)$, the dual of the Lie algebra $\mathfrak{so}(3)$. Similarly, $\text{d}_2 U(\xi, \kappa) \in T_\kappa^* \mathbb{R}^3$ is a cotangent vector over $\kappa \in \mathbb{R}^3$, so that $\mathbf{J}(\kappa, \text{d}_2 U(\xi, \kappa)) = \text{d}_2 U(\xi, \kappa) \times \kappa \in \mathfrak{so}^*(3)$, where $\mathbf{J}$ denotes the cotangent lift momentum map induced by the action of $SO(3)$ on $\mathbb{R}^3$.

Reformulating (1.31) in terms of vectors yields the following generalisation of equations considered by MEZIC [2006], written in a familiar vector form:

$$\begin{aligned} \left((\partial_t + \omega \times) \frac{\delta l}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta l}{\delta \Omega} \right)^\wedge = - \int (\text{d}_2 U((\xi, \kappa)(s, s')) \times \kappa(s, s'))^\wedge \text{d}s' \\ - \int (\xi(s, s') \text{d}_1 U((\xi, \kappa)^{-1}(s, s')) - \text{d}_1 U((\xi, \kappa)(s, s')) \xi^{-1}(s, s')) \text{d}s'. \end{aligned} \quad (1.32)$$

The set of integral-partial-differential equations on $\mathcal{F}([0, L], SO(3))$ for the fixed

filament is completed by computing the time derivative of $\xi(s, s')$:

$$\begin{aligned}\dot{\xi}(s, s') &= -\Lambda^{-1}(s')\dot{\Lambda}(s')\Lambda^{-1}(s')\Lambda(s) + \Lambda^{-1}(s')\dot{\Lambda}(s) \\ &= -\omega(s')\xi(s, s') + \xi(s, s')\omega(s).\end{aligned}\tag{1.33}$$

This expression is not quite a commutator because different positions s and s' appear in ω . However, operating with ξ^{-1} from the left in equation (1.33) gives a proper Lie-algebraic expression for the reconstruction of the relative orientation,

$$\xi^{-1}\dot{\xi}(s, s') = \omega(s) - \text{Ad}_{\xi^{-1}(s, s')}\omega(s').\tag{1.34}$$

Formulae (1.31) — (1.33) generalise the results in MEZIC [2006] for a fixed filament from $SO(2)$ to $SO(3)$ rotations.

1.4 Summary

At the outset, this chapter described the problem of charged molecular strand (CMS) dynamics. We saw that progress had been made in describing the elastic dynamics of CMS using Kirchhoff's approach KIRCHHOFF [1859]; DILL [1992], but that the generalisation of such descriptions to take account of torques arising from nonlocal electrostatic interactions of the strand with itself had remained elusive. We also saw that dynamical nonlocal interactions had been described in MEZIC [2006] in the case for a fixed filament with RCC dynamics occurring in $SO(2)$.

Next, we set out the goal of achieving a full dynamical description of both local elastic and nonlocal electrostatic interactions for the CMS by generalising the approach of SIMÓ ET AL. [1988] to take account of on-local self-interaction. This objective was set up in this chapter, but will be completed in Chapter 2.

The mathematical background to the CMS problem was touched upon in §1.2, where the reduced variables were introduced, and two alternative mathematical perspectives outlined. The *classical perspective* regarded the object of interest as a trajectory in the infinite dimensional configuration space $\mathcal{F}([0, L], SE(3))$. This contrasted with the *covariant perspective*, which con-

sidered space-time $X := [0, L] \times \mathbb{R}$ and sections $\sigma : X \rightarrow P := X \times SE(3)$ of a fibre bundle structure $\pi_{X,P} : P \rightarrow X$.

Finally, §1.3.1 describes the purely elastic Kirchhoff filament, and §1.3.2 the fixed filament with nonlocal interactions. Both of these simplifications will provide a useful reference point for the development of the CMS model in Chapter 2. The results of §1.3.2 represent the first new contribution, since they generalise the model of MEZIC [2006] to full $SO(3)$ motion of the RCC.

The next objective, achieved in Chapter 2, is to derive the CMS equations that take account of both local elastic deformations and nonlocal potential self-interactions.

CHAPTER 2 : MOTION OF EXACT SELF-INTERACTING GEOMETRIC RODS

In this chapter, the CMS equations for strand dynamics are derived in the convective representation. This derivation is based the theory of nonlinear elasticity set out in SIMÓ ET AL. [1988] suitably modified to take account of nonlocal self-interaction by potential forces. From a mathematical perspective, a Hamilton-Pontryagin (HP) approach is taken from control theory (see, for example, BLOCH [2003]), which has been modified to include additional terms describing the nonlocal contributions. A detailed background justification of the HP principle is found in YOSHIMURA & MARSDEN [2007].

The elegance and directness of the HP approach developed here relative to the alternative Euler-Poincaré (EP) approach in ELLIS ET AL. [2010] is accomplished by simplifying the interplay between the group action and the variational principle at the expense of introducing extra variables. The equivalent EP derivation is more elaborate than the HP derivation, because the EP approach invokes the Lie group action on the configuration space and thereby provides additional information. This alternative ‘bare hands’ derivation has been omitted from the thesis because the original EP derivation is not original work, being outlined in HOLM & PUTKARADZE [2009].

Chapter 3 shows that the derivation of the Euler-Poincaré equations and of the associated variational principle are corollaries of general theorems for systems whose configuration space is a Lie group. The complementary, but less transparent, HP route in that case reveals other perspectives and results whose abstract general formulation poses a challenge for future work.

2.1 Problem set-up

Suppose each rigid charge conformation (RCC) is identical and the k -th electrical charge is positioned near a given spatial point $\mathbf{r}$ through which the curve of base

points of the RCC passes. This curve is parameterised by a variable s which need not be the arc length. Rather, $s \in [0, L]$ is chosen here to be a parameter spanning a fixed interval.¹

The spatial reference (undisturbed) state for the k -th charge in a given RCC is the sum $\mathbf{r}(s) + \boldsymbol{\eta}_k(s)$. That is, $\boldsymbol{\eta}_k(s)$ is a constant vector that determines the position of the k -th electrical charge relative to the point $\mathbf{r}(s)$ along the curve in its reference configuration. The $\boldsymbol{\eta}_k(s)$ specify the shape of the RCC. At time t the position $\mathbf{c}_k$ of the k -th charge in the rigid conformation anchored at spatial position $\mathbf{r}(s, t)$ along the curve parameterised by s may rotate to a new position corresponding to the orientation $\Lambda(s, t)$ in the expression

$$\mathbf{c}_k(s, t) = \mathbf{r}(s, t) + \Lambda(s, t)\boldsymbol{\eta}_k(s), \quad \text{where} \quad \Lambda(s, 0) = \text{Id}. \quad (2.1)$$

This rigid conformational rotation is illustrated in Figure 1.1. In MEZIC [2006], the rotation is in the plane, so that $\Lambda \in SO(2)$, and there is only one charge, so $k = 1$. From now on notation for time dependence is suppressed without danger of confusion.

2.1.1 Nonlocal potential energy

One part of the potential energy of interaction between rigid conformations of charges at spatial coordinates $\mathbf{r}(s)$ and $\mathbf{r}(s')$ along the filament depends only on the magnitude $|\mathbf{c}_m(s') - \mathbf{c}_k(s)|$ of the vector from charge k at spatial position $\mathbf{c}_k(s)$ to charge m at spatial position $\mathbf{c}_m(s')$. This is the Euclidean spatial distance

$$d_{k,m}(s, s') = |\mathbf{c}_m(s') - \mathbf{c}_k(s)| \quad (2.2)$$

between the k -th and m -th charges in the two conformations whose base points are at $\mathbf{r}(s)$ and $\mathbf{r}(s')$, respectively. In this notation, the potential energy is given by

$$E = E_{loc}(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) - \sum_{k,m} \frac{1}{2} \iint U(d_{k,m}(s, s')) \left| \frac{d\mathbf{r}}{ds}(s) \right| \left| \frac{d\mathbf{r}}{ds}(s') \right| ds ds' \quad (2.3)$$

¹Note: limiting its parameterisation to a fixed interval does not mean that the filament is inextensible.

for an appropriate choice of the *interparticle interaction potential* $U(d_{k,m})$, and the quantities Ω, Γ (and ω, γ and ρ below) are defined in (1.1). The function $E_{loc}(\Omega, \Gamma)$ represents the elastic potential energy, and is usually taken to be a quadratic function of the deformations (Ω, Γ) , but more complex expressions are possible as well; we do not restrict the functional form of that dependence. The total Lagrangian

$$l = l_{loc}(\omega, \gamma, \Omega, \Gamma, \rho) + l_{np} \quad (2.4)$$

is then written as a sum of local l_{loc} and nonlocal l_{np} , where

$$l_{loc} = K(\omega, \gamma) - E_{loc}(\Omega, \Gamma, \rho), \quad l_{np} = -E_{np} := E_{loc} - E, \quad (2.5)$$

and K is the kinetic energy that depends only on the local velocities ω, γ .

The scalar distance $d_{k,m}$ in (2.2) and (2.3) may also be expressed in terms of vectors seen from the frame of orientation of the rigid body at a spatial point $\mathbf{r}(s)$ along the filament, as

$$\begin{aligned} d_{k,m}(s, s') &= |\mathbf{c}_m(s') - \mathbf{c}_k(s)| \\ &= |\Lambda^{-1}(s) (\mathbf{c}_m(s') - \mathbf{c}_k(s))| \\ &= |\Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) + \Lambda^{-1}(s)\Lambda(s')\boldsymbol{\eta}_m(s') - \boldsymbol{\eta}_k(s)| \\ &=: |\boldsymbol{\kappa}(s, s') + \xi(s, s')\boldsymbol{\eta}_m(s') - \boldsymbol{\eta}_k(s)|, \end{aligned} \quad (2.6)$$

where the following notation has been used

$$\boldsymbol{\kappa}(s, s') := \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) \in \mathbb{R}^3 \quad \text{and} \quad \xi(s, s') := \Lambda^{-1}(s)\Lambda(s') \in SO(3). \quad (2.7)$$

The first of these quantities is the spatial vector from $\mathbf{r}(s)$ to $\mathbf{r}(s')$, as seen by an observer who rotates with the RCC located at coordinate label s along the filament. The second is the relative orientation of the rigid charge conformations located at coordinate labels s and s' .

Remark 2.1.1 For later use, note the *transposition identities*,

$$\xi(s', s) = \xi(s, s')^T = \xi(s, s')^{-1}, \quad (2.8)$$

which follow from the definition of $\xi(s, s')$ in (2.7).

Proposition 2.1.2 (Left $SE(3)$ invariance) *The quantities $(\xi, \kappa) \in SO(3) \times \mathbb{R}^3$ defined in (2.7) are invariant under all transformations of the special Euclidean group $SE(3)$ acting on the left.*

Proof. As a set, the special Euclidean group is $SE(3) = SO(3) \times \mathbb{R}^3$, whose elements are denoted as $(\Lambda, \mathbf{r})$. Its group multiplication is given, e.g., in HOLM [2008B] by the *semidirect product action*,

$$(\Lambda_1, \mathbf{r}_1)(\Lambda_2, \mathbf{r}_2) = (\Lambda_1\Lambda_2, \mathbf{r}_1 + \Lambda_1\mathbf{r}_2), \quad (2.9)$$

where the concatenation $\Lambda\mathbf{r}$ denotes the action of $\Lambda \in SO(3)$ on $\mathbf{r} \in \mathbb{R}^3$ and the other notation is standard. For the choice

$$(\Lambda_1, \mathbf{r}_1) = (\Lambda, \mathbf{r})^{-1}(s) \quad \text{and} \quad (\Lambda_2, \mathbf{r}_2) = (\Lambda, \mathbf{r})(s'),$$

the $SE(3)$ multiplication rule (2.9) yields the quantities $(\xi, \kappa) \in SO(3) \times \mathbb{R}^3$ as

$$(\Lambda, \mathbf{r})^{-1}(s)(\Lambda, \mathbf{r})(s') = (\xi(s, s'), \kappa(s, s')). \quad (2.10)$$

This expression is invariant under the left action $(\Lambda, \mathbf{r}) \rightarrow (O, \mathbf{v})(\Lambda, \mathbf{r})$ of any element $(O, \mathbf{v})$ of the special Euclidean group $SE(3)$. ■

Remark 2.1.3 (Left $SO(3)$ invariance) Both the body separation vector $\kappa(s, s')$ and the relative orientation $\xi(s, s')$ defined in (2.7) are invariant under rotations of the spatial coordinate system obtained by the left action

$$(\mathbf{r}(s') - \mathbf{r}(s)) \mapsto O(\mathbf{r}(s') - \mathbf{r}(s)) \quad \text{and} \quad \Lambda \mapsto O\Lambda,$$

by any element O of the rotation group $SO(3)$. This is a corollary of Proposition 2.1.2 since $SO(3) \subset SE(3)$.

Remark 2.1.4 (Classical vs. covariant perspective) The variables $\Lambda, \mathbf{r}, \Omega, \omega, \Gamma, \gamma, \rho$ are interpreted as functions of the two variables s and t . For the classical perspective, it is important to see these variables as time-dependent curves with values in

function spaces. On the other hand, the covariant point of view requires the interpretation of these variables as fields over space-time. That is, one may interpret Λ as a curve

$$t \in \mathbb{R} \mapsto \Lambda(\cdot, t) \in \mathcal{F}([0, L], SO(3))$$

in the group $\mathcal{F}([0, L], SO(3))$ of all smooth functions defined on $[0, L]$ with values in $SO(3)$. Alternatively, one may interpret $\Lambda(s, t)$ as a function of space and time

$$(s, t) \in [0, L] \times \mathbb{R} \mapsto \Lambda(s, t) \in SO(3).$$

This observation leads to two different geometric approaches to the same equations: the *affine Euler-Poincaré* and the *covariant Lagrange-Poincaré* approaches.

The group operation for $\mathcal{F}([0, L], SO(3))$ is given by point-wise multiplication, and that $\mathcal{F}([0, L], SO(3))$ may be endowed with the structure of an infinite dimensional Fréchet Lie group. KRIEGL & MICHOR [1997] contains an account of Fréchet Lie groups in the framework of manifolds of maps from the point of view of convenient calculus.

Remark 2.1.5 Since $\Lambda \in SO(3)$, one finds that

$$\left| \frac{d\mathbf{r}}{ds}(s) \right| = \left| \Lambda^{-1} \frac{d\mathbf{r}}{ds}(s) \right| = |\mathbf{\Gamma}(s)|, \quad (2.11)$$

and the nonlocal potential energy in (2.3) reduces to

$$E_{np} = - \sum_{k,m} \frac{1}{2} \iint U(d_{k,m}(s, s')) |\mathbf{\Gamma}(s)| |\mathbf{\Gamma}(s')| ds ds'. \quad (2.12)$$

Remark 2.1.6 Everywhere in this discussion, the nonlocal Lagrangian, l_{np} , is assumed to be a function or functional of $\mathbf{\Gamma}$, ξ , and κ . It could, for example, be expressed in the integral form

$$l_{np}(\xi, \kappa, \mathbf{\Gamma}) = \iint U(\xi(s, s'), \kappa(s, s'), \mathbf{\Gamma}(s), \mathbf{\Gamma}(s')) ds ds' \quad (2.13)$$

or, indeed, a more general functional. In this work, formula (2.13) has been used to make calculations more explicit, although of course the methods would apply to more general functionals. Clearly, expression (2.3) is a reduction of (2.13) obtained

when the energy of the system of charges is a (half-)sum of interactions between all charges. This happens, for example, when investigating electrostatic or screened electrostatic charges in a linear medium.

Even though the expression $l_{np} = l_{np}(\xi, \kappa, \Gamma)$ is rather general, it is interesting to note that physical systems exist whose nonlocal interactions are more complex. For example, the electrostatic potential around a DNA molecule immersed in a fluid satisfies the nonlinear Poisson-Boltzmann equation and finding the potential in that case is a well-known problem for supercomputers BAKER ET AL. [2001]. It is usually impossible to find direct analytical solutions of this equation, except for some idealised cases. However, one may still consider the nonlocal Lagrangian to be a functional of the variables (ξ, κ, Γ) , and compute (at least in principle) the variational derivatives required by this theory. Therefore the present theory is applicable to the DNA molecule as well.

Having setup the nonlocal potential energy, the relative orientation $\xi(s, s')$ and relative body separation $\kappa(s, s')$, and the symmetry properties of the CMS problem, the next task is to derive *kinematic relations* that support the integral-partial-differential equations that arise from the variational principle.

From the covariant perspective, some of these kinematic relations obtain the geometric interpretation in part II as additional integrability/curvature relations required for reconstruction in the Lagrange-Poincaré field framework. A full treatment of this perspective is found in §7.1.3.

Meanwhile from the classical perspective, the kinematic relations are regarded as advection relations that arise from an affine action. This perspective will be fully explained in §3.3.

In the meantime, from the ‘bare hands’ perspective, the kinematic relations may be considered necessary auxiliary equations to the laws of motion that are required to close the dynamics among the reduced variables.

2.1.2 Kinematics

Auxiliary kinematic equations that hold without any reference to dynamics are required to close the dynamics. These go by the name *kinematic relations*, in order to distinguish them from the *dynamical equations* (derived later) that balance

the forces determined by the physics of the problem. By contrast, the kinematic relations hold for all strands, irrespective of their dynamic properties.

The first set of kinematic relations are derived by differentiating the definition of ρ in (1.1). First, the s -derivative along the filament is given by

$$\rho' = -\Lambda^{-1}\Lambda'\Lambda^{-1}\mathbf{r} + \Lambda^{-1}\mathbf{r}',$$

and hence (1.1) implies

$$\rho' = -\Omega\rho + \Gamma = -\Omega \times \rho + \Gamma. \quad (2.14)$$

Similarly,

$$\dot{\rho} = -\omega\rho + \gamma = -\omega \times \rho + \gamma. \quad (2.15)$$

The next set of kinematic relations is derived via equality of cross-derivatives with respect to t and s for any sufficiently smooth quantity. Symbolically, $\partial_s\partial_t\mathbf{r} = \partial_t\partial_s\mathbf{r}$. Differentiating γ and Γ gives

$$\gamma' = -\Omega \times \gamma + \Lambda^{-1}\partial_s\partial_t\dot{\mathbf{r}},$$

and

$$\dot{\Gamma} = -\omega \times \Gamma + \Lambda^{-1}\partial_t\partial_s\mathbf{r}.$$

Together, these two equations and the equality of cross-derivatives imply the following relation

$$\dot{\Gamma} + \omega \times \Gamma = \gamma' + \Omega \times \gamma. \quad (2.16)$$

Similarly, the equality of cross-derivatives $\partial_s\partial_t\Lambda = \partial_t\partial_s\Lambda$ yields the other kinematic relation,

$$\dot{\Omega} - \omega' = \omega \times \Omega. \quad (2.17)$$

Viewed from the covariant perspective, the latter of these equations is a type of *reconstruction relation*, specifically, (2.17) is an application of (7.29). The differential geometric meaning of these reconstruction relations will be discussed further in Chapter 7.

2.1.3 Generalisations of the strand

The molecular strand may be abstracted in two ways; from one-dimensional filaments to n -dimensional media, and from $SO(3)$ to arbitrary Lie groups. These generalisations provide useful for the extension of the theory to incorporate self-interacting membranes and deformable media, and also give a more transparent view of the geometric structure underlying the phenomena.

Consider the semidirect product $\mathcal{O} \ltimes E$ of a Lie group $\mathcal{O}$ with a *left* representation space E . The variables $\mathbf{r}$ and Λ defined above are now functions defined on a space-time $\mathcal{D} \times \mathbb{R}$, where $\mathcal{D}$ is a n -dimensional manifold:

$$\Lambda : (s, t) \in \mathcal{D} \times \mathbb{R} \rightarrow \Lambda(s, t) \in \mathcal{O}, \quad \text{and} \quad r : (s, t) \in \mathcal{D} \times \mathbb{R} \rightarrow r(s, t) \in E.$$

Here, the use of boldface notation has been avoided since the functions considered may be more general geometric quantities than vectors. As before, ‘dot’ ($\dot{}$) over a quantity denotes its time derivative. The derivative with respect to a variable in $\mathcal{D}$ is denoted by $\mathbf{d}$; for $\mathcal{D} = [0, L]$ this was previously denoted by ‘prime’ ($'$).

In this generalised situation, the definitions (1.1) become

$$\begin{aligned} \Omega &= \Lambda^{-1} \mathbf{d}\Lambda : T\mathcal{D} \rightarrow \mathfrak{o} \\ \omega &= \Lambda^{-1} \dot{\Lambda} : \mathcal{D} \rightarrow \mathfrak{o} \\ \Gamma &= \Lambda^{-1} \mathbf{d}r : T\mathcal{D} \rightarrow E \\ \gamma &= \Lambda^{-1} \dot{r} : \mathcal{D} \rightarrow E \\ \rho &= \Lambda^{-1} r : \mathcal{D} \rightarrow E. \end{aligned} \tag{2.18}$$

Thus, from the classical perspective, interpreting (Λ, r) as a curve in the group $\mathcal{F}(\mathcal{D}, \mathcal{O} \ltimes E)$, the previous definition may be rewritten as

$$(\omega, \gamma) = (\Lambda, r)^{-1} (\dot{\Lambda}, \dot{r}) \tag{2.19}$$

$$(\Omega, \Gamma, \rho) = c((\Lambda, r)^{-1}) \tag{2.20}$$

where c is defined by

$$c(\Lambda, r) = ((\Lambda, r) \mathbf{d}(\Lambda, r)^{-1}, -r). \tag{2.21}$$

Remarkably, c is a group one-cocycle, as defined in Definition 1.2.6. Therefore, recalling (1.4), equation (2.20) says that (Ω, Γ, ρ) are affine advected quantities with initial values equal to zero. This observation strongly suggests a relation with the affine Euler-Poincaré theory developed in the context of complex fluids in GAY-BALMAZ & RATIU [2008].

On the other hand, from the covariant point of view, interpreting (Λ, r) as a section of the trivial principal bundle

$$(\mathcal{D} \times \mathbb{R}) \times \mathcal{O} \otimes E \rightarrow \mathcal{D} \times \mathbb{R}$$

over space-time, definition (2.18) says that the variables (Ω, ω, ρ) are obtained by reduction by the subgroup $\mathcal{O}$ of the first jet extension of (Λ, r) . This, in turn, leads to a relation with the covariant Lagrange-Poincaré reduction for field theories developed in CASTRILLÓN-LÓPEZ & RATIU [2003].

Note that by choosing the one-dimensional interval $\mathcal{D} = [0, L]$, the Lie group $\mathcal{O} = SO(3)$ and left representation space $E = \mathbb{R}^3$, one recovers the situation of charged molecular strands.

Remark 2.1.7 Generalising to higher dimensions reveals certain distinct aspects of the underlying geometry of the problem that are not distinguished by considering the particular case of the charged strands. For example, in the case of *charged sheets* or *charged elastic deformed media*, $\mathcal{D}$ is a domain in $\mathbb{R}^n$ with $n = 2$ or 3 respectively, so the coordinate s has several dimensions. Therefore, Γ should be understood as a set of vectors $\Gamma_1, \dots, \Gamma_n$. Likewise, for the problem of flexible strands of rigid charge conformations the distinct objects E and $\mathfrak{o}$ both coincide with $\mathbb{R}^3$. This coincidence is removed in higher dimensions and thereby clarifies the underlying geometric structure of the theory.

In this section, we began by establishing the concrete modelling of the RCC. We moved on to the modelling of the nonlocal interactions, which naturally brought out the dependence on the relative body displacement, $\kappa(s, s')$ and the relative RCC orientation, $\xi(s, s')$. We also remarked on the symmetry properties of these quantities. We then moved on to establish the auxiliary kinematic relations in §2.1.2, and in §2.1.3 we remarked on the possible abstractions of the CMS model to higher dimensions and to an arbitrary Lie group. Throughout §2.1 we con-

trasted the differing conceptions brought by the classical and covariant perspectives.

In the sequel, §2.2, we shall move onto the derivation of the dynamical equations for the CMS model, which are obtained via a Hamilton-Pontryagin variational principle. These dynamical equations, when augmented with the auxiliary kinematic equations, complete the CMS equations.

2.2 Filament dynamics

It is convenient to split the variational principle into two parts, one concerning the local Lagrangian l_{loc} , the other concerning the nonlocal Lagrangian, l_{np} . The local part of the variational principle is performed in §2.2.1, meanwhile the nonlocal variational principle is embarked upon in §2.2.2. Each derivation is split into two parts, the necessary variations are calculated in a preparatory lemma, leaving the variational principle proper to be stated as a theorem. This organisation prevents cluttering in what is, necessarily, a long calculation.

Both §2.2.1 and §2.2.2 make use of a Hamilton-Pontryagin (HP) variational principle. The modification to the standard HP argument may be found in §2.2.2, where the variations of the relative variables require careful consideration. For more information on HP variational arguments, see YOSHIMURA & MARSDEN [2007]; BLOCH [2003].

2.2.1 Local dynamics

The Hamilton-Pontryagin approach is first applied to the local Lagrangian, so that $l = l_{loc}(\omega, \gamma, \Omega, \Gamma, \rho)$. Inspired by the classical HP approach, Lagrange multipliers are introduced for the holonomic constraints that impose the defining relations (1.1) for the five reduced variables $(\omega, \gamma, \Omega, \Gamma, \rho)$.

Remark 2.2.1 Since the local Lagrangian depends only on the reduced variables defined in (1.1), there is no need to study the relative variables defined in (2.7) until §2.2.2.

Theorem 2.2.1 (Hamilton-Pontryagin theorem for local filament dynamics)

The filament dynamics that arise from the variational principle $\delta S = 0$ with action S given by

$$S = \int l(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) dt + \iint \left(\boldsymbol{\pi} \cdot (\Lambda^{-1} \dot{\Lambda} - \boldsymbol{\omega}) + \boldsymbol{\Pi} \cdot (\Lambda^{-1} \Lambda' - \boldsymbol{\Omega}) + \mathbf{R} \cdot (\Lambda^{-1} \mathbf{r} - \boldsymbol{\rho}) + \boldsymbol{\mu} \cdot (\Lambda^{-1} \dot{\mathbf{r}} - \boldsymbol{\gamma}) + \mathbf{M} \cdot (\Lambda^{-1} \mathbf{r}' - \boldsymbol{\Gamma}) \right) ds dt,$$

is governed by the following equations

$$\frac{\delta l}{\delta \boldsymbol{\rho}} = \mathbf{R}, \quad \frac{\delta l}{\delta \boldsymbol{\omega}} = \boldsymbol{\pi}, \quad \frac{\delta l}{\delta \boldsymbol{\Omega}} = \boldsymbol{\Pi}, \quad \frac{\delta l}{\delta \boldsymbol{\gamma}} = \boldsymbol{\mu}, \quad \frac{\delta l}{\delta \boldsymbol{\Gamma}} = \mathbf{M},$$

$$\dot{\boldsymbol{\pi}} + \boldsymbol{\omega} \times \boldsymbol{\pi} + \boldsymbol{\Pi}' + \boldsymbol{\Omega} \times \boldsymbol{\Pi} + \boldsymbol{\gamma} \times \boldsymbol{\mu} + \boldsymbol{\Gamma} \times \mathbf{M} + \boldsymbol{\rho} \times \mathbf{R} = 0,$$

and

$$\dot{\boldsymbol{\mu}} + \boldsymbol{\omega} \times \boldsymbol{\mu} + \mathbf{M}' + \boldsymbol{\Omega} \times \mathbf{M} - \mathbf{R} = 0,$$

together with the constraints,

$$\Lambda^{-1} \dot{\Lambda} = \boldsymbol{\omega}, \quad \Lambda^{-1} \Lambda' = \boldsymbol{\Omega}, \quad \Lambda^{-1} \mathbf{r} = \boldsymbol{\rho}, \quad \Lambda^{-1} \dot{\mathbf{r}} = \boldsymbol{\gamma}, \quad \Lambda^{-1} \mathbf{r}' = \boldsymbol{\Gamma},$$

which correspond to (1.1).

To facilitate the proof of Theorem 2.2.1, it is convenient to first prove a lemma that will be helpful in computing the variations of the quantities appearing in the action S . Lemma 2.2.2 relates the variations of the reduced variables (1.1) with the unreduced variations. Note that Lemma 2.2.2 was also used in §1.3.2 to derive the equations for a fixed filament that generalised the model of MEZIC [2006].

Lemma 2.2.2 *The variations of the quantities in Λ and $\mathbf{r}$ of the formulas in (1.1) are*

$$\begin{aligned}
 \delta \left(\Lambda^{-1} \dot{\Lambda} \right) &= \frac{\partial \widehat{\Sigma}}{\partial t} + \left[\Lambda^{-1} \dot{\Lambda}, \widehat{\Sigma} \right], \\
 \delta \left(\Lambda^{-1} \Lambda' \right) &= \widehat{\Sigma}' + \left[\Lambda^{-1} \Lambda', \widehat{\Sigma} \right], \\
 \delta \left(\Lambda^{-1} \mathbf{r} \right) &= \Psi - \widehat{\Sigma} \left(\Lambda^{-1} \mathbf{r} \right), \\
 \delta \left(\Lambda^{-1} \dot{\mathbf{r}} \right) &= \dot{\Psi} - \widehat{\Sigma} \left(\Lambda^{-1} \dot{\mathbf{r}} \right) + \left(\Lambda^{-1} \dot{\Lambda} \right) \Psi, \\
 \delta \left(\Lambda^{-1} \mathbf{r}' \right) &= \Psi' - \widehat{\Sigma} \left(\Lambda^{-1} \mathbf{r}' \right) + \left(\Lambda^{-1} \Lambda' \right) \Psi,
 \end{aligned} \tag{2.22}$$

where the independent variations are defined by

$$\Psi(s) = \Lambda^{-1}(s) \delta \mathbf{r}(s) \quad \text{and} \quad \widehat{\Sigma}(s) = \Lambda^{-1}(s) \delta \Lambda(s).$$

Proof. Calculating the variations directly, one by one, yields,

$$\begin{aligned}
 \delta \left(\Lambda^{-1} \dot{\Lambda} \right) &= -\Lambda^{-1} \delta \Lambda \left(\Lambda^{-1} \dot{\Lambda} \right) + \Lambda^{-1} \delta \dot{\Lambda} \\
 &= -\Lambda^{-1} \delta \Lambda \left(\Lambda^{-1} \dot{\Lambda} \right) + \frac{\partial}{\partial t} \left(\Lambda^{-1} \delta \Lambda \right) + \left(\Lambda^{-1} \dot{\Lambda} \right) \left(\Lambda^{-1} \delta \Lambda \right) \\
 &= \frac{\partial \widehat{\Sigma}}{\partial t} + \left[\Lambda^{-1} \dot{\Lambda}, \widehat{\Sigma} \right].
 \end{aligned}$$

The variation of $\Lambda^{-1} \Lambda'$ takes on a similar form as

$$\delta \left(\Lambda^{-1} \Lambda' \right) = \widehat{\Sigma}' + \left[\Lambda^{-1} \Lambda', \widehat{\Sigma} \right].$$

Now, the variation of $\Lambda^{-1} \dot{\mathbf{r}}$ is given by

$$\begin{aligned}
 \delta \left(\Lambda^{-1} \dot{\mathbf{r}} \right) &= -\Lambda^{-1} \delta \Lambda \Lambda^{-1} \dot{\mathbf{r}} + \Lambda^{-1} \delta \dot{\mathbf{r}} \\
 &= -\left(\Lambda^{-1} \delta \Lambda \right) \left(\Lambda^{-1} \dot{\mathbf{r}} \right) + \frac{\partial}{\partial t} \left(\Lambda^{-1} \delta \mathbf{r} \right) + \left(\Lambda^{-1} \dot{\Lambda} \right) \left(\Lambda^{-1} \delta \mathbf{r} \right) \\
 &= \dot{\Psi} - \widehat{\Sigma} \left(\Lambda^{-1} \dot{\mathbf{r}} \right) + \left(\Lambda^{-1} \dot{\Lambda} \right) \Psi.
 \end{aligned}$$

Again, a similar argument yields the variation of $\Lambda^{-1}\mathbf{r}'$,

$$\delta(\Lambda^{-1}\mathbf{r}') = \Psi' - \widehat{\Sigma}(\Lambda^{-1}\mathbf{r}') + (\Lambda^{-1}\Lambda')\Psi.$$

Finally, the variation of $\Lambda^{-1}\mathbf{r}$ follows from

$$\delta(\Lambda^{-1}\mathbf{r}) = -\Lambda^{-1}\delta\Lambda\Lambda^{-1}\mathbf{r} + \Lambda^{-1}\delta\mathbf{r} = \Psi - \widehat{\Sigma}(\Lambda^{-1}\mathbf{r}).$$

This completes the list of formulas in the statement of the lemma. ■

Remark 2.2.2 Note that the derivation of the variation formulae in Lemma 2.2.2 are similar to that of the auxiliary kinematic equations in §2.1.2. This similarity may be understood from the covariant perspective where the two calculations arise in precisely the same fashion. More details on this calculation may be found in §7.1.1. Such similarities are more obscure from the classical perspective, where quantities such as, e.g., $\Lambda^{-1}\dot{\Lambda}$ and $\Lambda^{-1}\Lambda'$, are interpreted as totally distinct geometric objects, the former as a component of a tangent vector on $\mathcal{F}([0, L], SO(3))$, the latter as a component of a group one-cocycle, more details on this point and details on group on-cocycles will be discussed further in §3.3.

The variational formulae (2.22) are now used to prove Theorem 2.2.1, which derives the equations governing local filament dynamics.

Proof.

Variations with respect to the Lagrange multipliers impose the defining relations for the five quantities $\{\rho, \omega, \Omega, \gamma, \Gamma\}$. The conjugate variations give

$$\frac{\delta l}{\delta \rho} - \mathbf{R} = 0, \quad \frac{\delta l}{\delta \omega} - \pi = 0, \quad \frac{\delta l}{\delta \Omega} - \Pi = 0, \quad \frac{\delta l}{\delta \gamma} - \mu = 0, \quad \frac{\delta l}{\delta \Gamma} - \mathbf{M} = 0.$$

Collecting the variations proportional to Σ and Ψ yields the filament equations

$$\dot{\pi} + \omega \times \pi + \Pi' + \Omega \times \Pi + \gamma \times \mu + \Gamma \times \mathbf{M} + \rho \times \mathbf{R} = 0,$$

and

$$\dot{\mu} + \omega \times \mu + \mathbf{M}' + \Omega \times \mathbf{M} - \mathbf{R} = 0,$$

respectively. ■

Remark 2.2.3 The Hamilton-Pontryagin approach used here also allows nonholonomic constraints to be imposed on the motion of the strand, although that direction is too far afield for the current work. See HOLM [2008B] for a discussion of nonholonomic constraints using the Hamilton-Pontryagin approach.

2.2.2 Nonlocal dynamics

The HP derivation of the CMS equations is completed by calculating the variational principle for l_{np} , the nonlocal Lagrangian. In Lemma 2.2.3, the variations of the relative variables (ξ, κ) , defined in (2.7), are pre-calculated. These variations are subsequently used to derive the nonlocal part of the CMS equations in Theorem 2.2.4.

The action S_{np} for the nonlocal Lagrangian, l_{np} , is given by

$$\begin{aligned} S_{np} = & \int l_{np}(\xi, \kappa, \Gamma) dt + \iint \mathbf{m} \cdot (\Lambda^{-1}(s) \mathbf{r}'(s) - \Gamma) \, ds dt \\ & + \iiint \left(X \cdot (\Lambda^{-1}(s) \Lambda(s') - \xi) + \mathbf{K} \cdot (\Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) - \kappa) \right) \, ds \, ds' \, dt. \end{aligned}$$

Since the strand is assumed to be locally extensible, the stretch of the strand's base requires extra factors of $|\Gamma|$ multiplying the differential of parameter $|\Gamma(s)|$ along the strand ds in the expression for potential energy, as in (2.12). However, in order to simplify the formulas and avoid extra factors in the integrals, that factor of $|\Gamma(s)|$ has been incorporated into the nonlocal potential for the present derivation. No confusion with the discussion of §2.1.1 should arise here.

Lemma 2.2.3 *The additional variational formulas needed for calculating the equations of motion are given by*

$$\begin{aligned} \Lambda^{-1}(s') \Lambda(s) (\delta (\Lambda^{-1}(s) \Lambda(s'))) &= -\text{Ad}_{\Lambda^{-1}(s') \Lambda(s)} \widehat{\Sigma}(s) + \widehat{\Sigma}(s'), \\ \delta (\Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s))) &= -\widehat{\Sigma}(s) \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) \\ &\quad + \Lambda^{-1}(s) \Lambda(s') \Psi(s') + \Psi(s), \end{aligned}$$

where the independent variations are defined by

$$\Psi(s) = \Lambda^{-1}(s)\delta\mathbf{r}(s) \quad \text{and} \quad \widehat{\Sigma}(s) = \Lambda^{-1}(s)\delta\Lambda(s). \quad (2.23)$$

Proof. The first variational formula is calculated directly, as

$$\begin{aligned} \Lambda^{-1}(s')\Lambda(s) \left(\delta \left(\Lambda^{-1}(s)\Lambda(s') \right) \right) &= \Lambda^{-1}(s')\Lambda(s) \left(\Lambda^{-1}(s)\delta\Lambda(s') \right) \\ &\quad - \Lambda^{-1}(s')\Lambda(s) \left(\Lambda^{-1}(s)\delta\Lambda(s)\Lambda^{-1}(s)\Lambda(s') \right) \\ &= -\text{Ad}_{\Lambda^{-1}(s')\Lambda(s)}\widehat{\Sigma}(s) + \widehat{\Sigma}(s'). \end{aligned}$$

The second variational formula follows similarly from a direct calculation,

$$\begin{aligned} \delta \left(\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) \right) &= -\Lambda^{-1}(s)\delta\Lambda(s)\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) \\ &\quad + \Lambda^{-1}(s) \left(\delta\mathbf{r}(s') - \delta\mathbf{r}(s) \right) \\ &= -\widehat{\Sigma}(s)\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) \\ &\quad + \Lambda^{-1}(s)\Lambda(s')\Psi(s') + \Psi(s), \end{aligned}$$

which proves the lemma. ■

Theorem 2.2.4 *The equations that arise from the variational principle with the non-local action*

$$\begin{aligned} S_{np} &= \iiint U(\xi, \boldsymbol{\kappa}, \boldsymbol{\Gamma}) ds ds' dt + \iint \mathbf{m} \cdot (\Lambda^{-1}(s)\mathbf{r}'(s) - \boldsymbol{\Gamma}) ds dt \\ &\quad + \iiint (X \cdot (\Lambda^{-1}(s)\Lambda(s') - \xi)) ds ds' dt \\ &\quad + \iiint (\mathbf{K} \cdot (\Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) - \boldsymbol{\kappa})) ds ds' dt \end{aligned}$$

are given by:

$$X = \frac{\partial U}{\partial \xi}, \quad \mathbf{K} = \frac{\partial U}{\partial \boldsymbol{\kappa}}, \quad \mathbf{m} = \frac{\partial U}{\partial \boldsymbol{\Gamma}},$$

and

$$\mathbf{\Gamma} \times \mathbf{m} = \int (\xi(s, s')X(s', s) - X(s, s')\xi^{-1}(s, s') + \mathbf{K}(s, s') \times \boldsymbol{\kappa}(s, s')) ds',$$

$$\mathbf{m}' + \boldsymbol{\Omega} \times \mathbf{m} = \int (\xi(s, s')\mathbf{K}(s', s) - \mathbf{K}(s, s')) ds',$$

together with the constraints,

$$\xi = \Lambda^{-1}(s)\Lambda(s'), \quad \boldsymbol{\kappa} = \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) , \quad \mathbf{\Gamma} = \Lambda^{-1}(s)\mathbf{r}'(s) .$$

Proof. The proof is obtained by substituting the variations given in Lemma 2.2.3 into Hamilton's principle for the action in the statement of the theorem. Variations in X , $\mathbf{K}$ and $\mathbf{m}$ yield the constraints,

$$\xi = \Lambda^{-1}(s)\Lambda(s'), \quad \boldsymbol{\kappa} = \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) , \quad \mathbf{\Gamma} = \Lambda^{-1}(s)\mathbf{r}'(s) .$$

Variations in ξ , $\boldsymbol{\kappa}$, and $\mathbf{\Gamma}$ yield the relationships

$$X = \frac{\partial U}{\partial \xi}, \quad \mathbf{K} = \frac{\partial U}{\partial \boldsymbol{\kappa}}, \quad \mathbf{m} = \frac{\partial U}{\partial \mathbf{\Gamma}} .$$

Finally, the variations proportional to $\hat{\Sigma}(s)$ and $\Psi(s)$ yield

$$\mathbf{\Gamma} \times \mathbf{m} = \int (\xi(s, s')X(s', s) - X(s, s')\xi^{-1}(s, s') + \mathbf{K}(s, s') \times \boldsymbol{\kappa}(s, s')) ds'$$

and

$$\mathbf{m}' + \boldsymbol{\Omega} \times \mathbf{m} = \int (\xi(s, s')\mathbf{K}(s', s) - \mathbf{K}(s, s')) ds',$$

respectively. ■

Combining the nonlocal terms from Theorem 2.2.4 with the local strand equations from Theorem 2.2.1 results in the full set of nonlocal CMS equations, which

read

$$\begin{aligned} \dot{\pi} + \omega \times \pi + \Pi' + \Omega \times \Pi + \gamma \times \mu + \Gamma \times (\mathbf{M} + \mathbf{m}) + \rho \times \mathbf{R} \\ = \int (\mathbf{K}(s, s') \times \kappa(s, s') + \mathbf{Z}(s, s')) ds' \end{aligned} \quad (2.24)$$

and

$$\begin{aligned} \dot{\mu} + \omega \times \mu + (\mathbf{M} + \mathbf{m})' + \Omega \times (\mathbf{M} + \mathbf{m}) - \mathbf{R} \\ = \int (\xi(s, s') \mathbf{K}(s', s) - \mathbf{K}(s, s')) ds', \end{aligned} \quad (2.25)$$

where $\mathbf{Z}$ has been defined according to

$$\widehat{\mathbf{Z}}(s, s') := \xi(s, s') X(s', s) - X(s, s') \xi^{-1}(s, s'). \quad (2.26)$$

The functional derivative relations obtained from Theorems 2.2.1 and 2.2.4 may now be used to express the equations of motion in terms of the total Lagrangian, $l = l_{loc} + l_{np}$. These relations are

$$\begin{aligned} \mathbf{R} &= \frac{\delta l_{loc}}{\delta \rho}, & \pi &= \frac{\delta l_{loc}}{\delta \omega}, \\ \Pi &= \frac{\delta l_{loc}}{\delta \Omega}, & \mu &= \frac{\delta l_{loc}}{\delta \gamma}, \\ \mathbf{M} &= \frac{\delta l_{loc}}{\delta \Gamma}, & X &= \frac{\delta l_{np}}{\delta \xi}, \\ \mathbf{K} &= \frac{\delta l_{np}}{\delta \kappa}, & \mathbf{M} + \mathbf{m} &= \frac{\delta (l_{loc} + l_{np})}{\delta \Gamma}. \end{aligned}$$

Substituting these relations into the equations of motion (2.24) — (2.25) gives the full nonlocal equations of motion for the charged molecular strand in Lagrangian form.

$$\begin{aligned} (\partial_t + \omega \times) \frac{\delta l_{loc}}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta l_{loc}}{\delta \Omega} &= \frac{\delta l_{loc}}{\delta \gamma} \times \gamma + \frac{\delta (l_{loc} + l_{np})}{\delta \Gamma} \times \Gamma + \frac{\delta l_{loc}}{\delta \rho} \times \rho \\ &+ \int \left(\frac{\partial U}{\partial \kappa}(s, s') \times \kappa(s, s') + \mathbf{Z}(s, s') \right) ds', \end{aligned} \quad (2.27)$$

Motion of exact self-interacting geometric rods

$$\begin{aligned}
(\partial_t + \boldsymbol{\omega} \times) \frac{\delta l_{loc}}{\delta \boldsymbol{\gamma}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta (l_{loc} + l_{np})}{\delta \boldsymbol{\Gamma}} \\
= \frac{\delta l_{loc}}{\delta \boldsymbol{\rho}} + \int \left(\xi(s, s') \frac{\partial U}{\partial \boldsymbol{\kappa}}(s', s) - \frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \right) ds'. \quad (2.28)
\end{aligned}$$

The term $\widehat{\mathbf{Z}}(s, s')$ is the sought after contribution from the nonlocal part of the Lagrangian.

Remark 2.2.4 The dynamical equations (2.27) and (2.28) must be augmented by the kinematic relations (2.15), (2.16) and (2.17) in order to close the system.

Remark 2.2.5 The resulting system of equations describes an elastic filament with two additional nonlocal components, or degrees of freedom, compared to the ordinary Kirchhoff filament, to which the system reduces when $\boldsymbol{\kappa}$ and ξ are absent.

Remark 2.2.6 The two additional (nonlocal) degrees of freedom in $\boldsymbol{\kappa}$ and ξ may produce an important effect that distinguishes the behaviour of this system from that of the ordinary Kirchhoff filament. Namely, the presence of the two additional equations for $\boldsymbol{\kappa}$ and ξ *raises the order* of the equations. In turn, the increase in differential order of the system may be expected to produce additional modes of excitation for the waves that propagate along the filament when the system is linearised around the static solutions. This possibility is not pursued in the present work.

2.3 Conservation laws

In order to elucidate the physical meaning of the somewhat complex-looking equations (2.27) and (2.28), they may be written explicitly as conservation laws. For this purpose, one invokes the following identities valid for any Lie group G . Given a smooth curve $g(t) \in G$, $\eta \in \mathfrak{g}$, and $\mu \in \mathfrak{g}^*$, one has

$$\text{Ad}_{g^{-1}(t)} \frac{\partial}{\partial t} \text{Ad}_{g(t)} \eta = \text{ad}_{\sigma(t)} \eta \quad (2.29)$$

$$\text{Ad}_{g(t)}^* \frac{\partial}{\partial t} \text{Ad}_{g^{-1}(t)}^* \mu = -\text{ad}_{\sigma(t)}^* \mu, \quad (2.30)$$

where $\sigma(t) = g^{-1}\dot{g}(t) \in \mathfrak{g}$ and Ad^* denotes the coadjoint action of G on $\mathfrak{g}^*$ defined by $\langle \text{Ad}_g^* \mu, \eta \rangle := \langle \mu, \text{Ad}_g \eta \rangle$. Formula (2.30) generalises to a curve $\mu(t)$ as

$$\text{Ad}_{g(t)}^* \frac{\partial}{\partial t} \text{Ad}_{g^{-1}(t)}^* \mu(t) = \dot{\mu}(t) - \text{ad}_{\sigma(t)}^* \mu(t). \quad (2.31)$$

The derivation of the conservation form of equations (2.27) and (2.28) requires the group $G = SE(3)$ whose elements are denoted by $g = (\Lambda, \mathbf{r})$. Consider the function $(\Lambda(s, t), \mathbf{r}(s, t))$ defined on space-time. Then,

$$\sigma = (\Lambda, \mathbf{r})^{-1}(\dot{\Lambda}, \dot{\mathbf{r}}) = (\Lambda^{-1}\dot{\Lambda}, \Lambda^{-1}\dot{\mathbf{r}}) = (\boldsymbol{\omega}, \boldsymbol{\gamma}). \quad (2.32)$$

Recall, e.g., MARS DEN & RATIU [2002]; HOLM [2008B], that the coadjoint action on $(\boldsymbol{\mu}, \boldsymbol{\beta}) \in \mathfrak{se}(3)^*$ is given by (1.15) and hence

$$\text{ad}_{(\boldsymbol{\omega}, \boldsymbol{\gamma})}^*(\boldsymbol{\mu}, \boldsymbol{\beta}) = -(\boldsymbol{\omega} \times \boldsymbol{\mu} + \boldsymbol{\gamma} \times \boldsymbol{\beta}, \boldsymbol{\omega} \times \boldsymbol{\beta}). \quad (2.33)$$

Then, using equations (2.31) and (2.33) for the temporal dual Lie algebra elements $(\boldsymbol{\mu}, \boldsymbol{\beta}) = (\delta l / \delta \boldsymbol{\omega}, \delta l / \delta \boldsymbol{\gamma})$ yields

$$\begin{aligned} & \text{Ad}_{(\Lambda, \mathbf{r})}^* \frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, \mathbf{r})^{-1}}^* \left(\frac{\delta l_{loc}}{\delta \boldsymbol{\omega}}, \frac{\delta l_{loc}}{\delta \boldsymbol{\gamma}} \right) \right] \\ &= \frac{\partial}{\partial t} \left(\frac{\delta l_{loc}}{\delta \boldsymbol{\omega}}, \frac{\delta l_{loc}}{\delta \boldsymbol{\gamma}} \right) + \left(\boldsymbol{\omega} \times \frac{\delta l_{loc}}{\delta \boldsymbol{\omega}} + \boldsymbol{\gamma} \times \frac{\delta l_{loc}}{\delta \boldsymbol{\gamma}}, \boldsymbol{\omega} \times \frac{\delta l_{loc}}{\delta \boldsymbol{\gamma}} \right). \end{aligned} \quad (2.34)$$

For the derivative with respect to curve parameterisation s , recall that the non-local part of the potential depends on $\boldsymbol{\Gamma}$ as well. Thus,

$$\begin{aligned} & \text{Ad}_{(\Lambda, \mathbf{r})}^* \frac{\partial}{\partial s} \left[\text{Ad}_{(\Lambda, \mathbf{r})^{-1}}^* \left(\frac{\delta l_{loc}}{\delta \boldsymbol{\Omega}}, \frac{\delta(l_{loc} + l_{np})}{\delta \boldsymbol{\Gamma}} \right) \right] \\ &= \frac{\partial}{\partial s} \left(\frac{\delta l_{loc}}{\delta \boldsymbol{\Omega}}, \frac{\delta l_{loc}}{\delta \boldsymbol{\Gamma}} \right) + \left(\boldsymbol{\Omega} \times \frac{\delta l_{loc}}{\delta \boldsymbol{\Omega}} + \boldsymbol{\Gamma} \times \frac{\delta(l_{loc} + l_{np})}{\delta \boldsymbol{\Gamma}}, \boldsymbol{\Omega} \times \frac{\delta l_{loc}}{\delta \boldsymbol{\Gamma}} \right). \end{aligned} \quad (2.35)$$

Some additional identities derived below are needed in treating the non-local part of the potential.

The nonlocal terms may be expressed as a formal derivatives of the non-local part of the potential with respect to Lie algebra elements $\boldsymbol{\Omega}$ and $\boldsymbol{\Gamma}$. These formal derivatives are known as *total* derivatives, and result from an implicit dependence

Motion of exact self-interacting geometric rods

of the nonlocal terms on Ω and Γ that has no explicit expression. Further details on the mathematical nature of the total derivative will be given in §3.3.1. For the moment, a direct implicit calculation can be made as follows. Note that there are only *two* free variations $\hat{\Sigma} = \Lambda^{-1}\delta\Lambda$ and $\Psi = \Lambda^{-1}\delta\mathbf{r}$. On the other hand, the non-local part of the Lagrangian depends on three variables κ, ξ , and Γ . Thus, there must be a relation between the partial derivatives of the non-local part of the Lagrangian and the total derivatives with respect to Γ and Ω . This relation is computed as follows.

Upon identifying coefficients of the free variations $\hat{\Sigma} = \Lambda^{-1}\delta\Lambda$ and $\Psi = \Lambda^{-1}\delta\mathbf{r}$, the following identity relates different variational derivatives of the non-local potential l_{np} :

$$\begin{aligned} \delta l_{np} &= \left\langle \xi^{-1} \frac{\delta l_{np}}{\delta \xi}, \xi^{-1} \delta \xi \right\rangle + \left\langle \frac{\delta l_{np}}{\delta \kappa}, \delta \kappa \right\rangle + \left\langle \frac{\delta l_{np}}{\delta \Gamma}, \delta \Gamma \right\rangle \\ &= \left\langle \frac{\delta l_{np}}{\delta \Gamma} \Big|_{\text{Tot}}, \delta \Gamma \right\rangle + \left\langle \frac{\delta l_{np}}{\delta \Omega} \Big|_{\text{Tot}}, \delta \Omega \right\rangle. \end{aligned} \quad (2.36)$$

Here, the subscript on $(\cdot)|_{\text{Tot}}$ denotes the total derivative. Using expressions from Lemmas 2.2.2 and 2.2.3 for $\xi^{-1}\delta\xi$, $\delta\kappa$, $\delta\Omega$ and $\delta\Gamma$, then collecting terms proportional to the free variation Σ yields the following identity, which implicitly defines $\delta l_{np}/\delta\Omega$ in terms of known quantities,

$$\begin{aligned} -\frac{\partial}{\partial s} \frac{\delta l_{np}}{\delta \Omega} \Big|_{\text{Tot}} - \Omega \times \frac{\delta l_{np}}{\delta \Omega} \Big|_{\text{Tot}} &= \\ \int \frac{\partial U}{\partial \kappa}(s, s') \times \kappa(s, s') \, ds' + \int \mathbf{Z}(s, s') \, ds', \end{aligned} \quad (2.37)$$

where $\mathbf{Z}(s, s')$ is defined according to (2.26). Likewise, identifying terms multiplying Ψ gives

$$\begin{aligned} -\frac{\partial}{\partial s} \frac{\delta l_{np}}{\delta \Gamma} \Big|_{\text{Tot}} - \Omega \times \frac{\delta l_{np}}{\delta \Gamma} \Big|_{\text{Tot}} &= -\frac{\partial}{\partial s} \frac{\delta l_{np}}{\delta \Gamma} - \Omega \times \frac{\delta l_{np}}{\delta \Gamma} \\ &+ \int \frac{\partial U}{\partial \kappa}(s, s') - \xi(s, s') \frac{\partial U}{\partial \kappa}(s', s) \, ds'. \end{aligned} \quad (2.38)$$

Therefore, equations (2.27) and (2.28) are equivalent to the following equations expressed on $\mathfrak{se}^*(3)$ in conservative form using variations of the total Lagrangian,

$l := l_{loc} + l_{np}$:

$$\begin{aligned} \frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) \right] + \frac{\partial}{\partial s} \left[\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \Omega} \Big|_{\text{Tot}}, \frac{\delta l}{\delta \Gamma} \Big|_{\text{Tot}} \right) \right] \\ = \text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \rho} \times \rho, \frac{\delta l}{\delta \rho} \right). \end{aligned} \quad (2.39)$$

Here, the components of

$$\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right)$$

represent, respectively, the spatial angular momentum density and the spatial linear momentum density of the strand, whose centre of mass lies along its centreline. The components of

$$\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \rho} \times \rho, \frac{\delta l}{\delta \rho} \right) = \left(0, \Lambda \frac{\delta l}{\delta \rho} \right) \quad (2.40)$$

are the external torques and forces. (See (1.27) for the last simplification.) The vanishing torque from conservative $SO(3)$ invariant external forces is a consequence of Noether's theorem. More details on this will be given in §8.3. As mentioned above, only external forces arising from potentials are considered in this work. In principle, more general non-conservative forces and torques can be considered as well, but this question is left for further studies.

Remark 2.3.1 For future reference, it is advantageous to write out the conservation law (2.39) in convective form as

$$\begin{cases} (\partial_t + \omega \times) \frac{\delta l}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta l}{\delta \Omega} \Big|_{\text{Tot}} + \rho \times \frac{\delta l}{\delta \rho} + \Gamma \times \frac{\delta l}{\delta \Gamma} \Big|_{\text{Tot}} + \gamma \times \frac{\delta l}{\delta \gamma} = 0, \\ (\partial_t + \omega \times) \frac{\delta l}{\delta \gamma} + (\partial_s + \Omega \times) \frac{\delta l}{\delta \Gamma} \Big|_{\text{Tot}} - \frac{\delta l}{\delta \rho} = 0. \end{cases} \quad (2.41)$$

Here the total Lagrangian $l := l_{loc} + l_{np}$ has been used. Note that these equations coincide precisely with the equations for the purely elastic filaments derived in SIMÓ ET AL. [1988].

Note that the variations with respect to Ω and Γ are computed implicitly in (2.37, 2.38). To actually use these equations to explicitly describe non-local interactions, one must expand the derivatives with respect to ξ and κ in (2.41). However,

it bears further emphasis that it is interesting that the expressions for the non-local interactions formally coincide with the equations for the purely elastic motion.

2.4 Summary

In this chapter we have derived the nonlocal equations of motion for the charged molecular strand. The problem was set up in §2.1, where the nonlocal potential and symmetry properties were discussed in §2.1.1, and the kinematic relations that close the dynamics and hold for any Lagrangian were derived in §2.1.2.

The equations of motion were derived in §2.2. The calculation of the variational principle was long, and was therefore split into two parts. The first part focused on the local Lagrangian, while the second part concerned the nonlocal terms. The full equations of motion were given by (2.27) and (2.28).

Finally, the conservation law form of the molecular strand equations was developed in §2.3. After introducing the notion of the total derivative, the conservation equations were given by (2.39), and a version of (2.27) — (2.28) that uses total derivatives was shown in (2.41).

The remaining objectives of Part I consist of giving two geometric interpretations of the molecular strand equations derived here. The classical formulation, developed in Chapter 3, reveals that the molecular strand equations result from affine Euler-Poincaré reduction. Meanwhile, the covariant formulation of Chapter 4 shows that the molecular strand equations arise from covariant Lagrange-Poincaré reduction.

CHAPTER 3 : CLASSICAL FORMULATION OF MOLECULAR STRANDS

3.1 Introduction to affine Lie group actions

3.1.1 Classical Lagrangian approach

The Hamilton-Pontryagin method used in Chapter 2 to derive the strand equations of motion (2.27) and (2.28) strongly suggest that the dynamics of the molecular strand with nonlocal interactions may be obtained by a *classical* Lagrangian reduction. In other words, the evolution for the Lagrangian variables Λ and $\mathbf{r}$ should be given by the *standard Euler-Lagrange equations*

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\Lambda}} - \frac{\partial L}{\partial \Lambda} = 0, \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{r}}} - \frac{\partial L}{\partial \mathbf{r}} = 0$$

associated to a Lagrangian L defined on the tangent bundle TQ of the configuration space Q and invariant under the action of the symmetry group of the theory. The evolution of the reduced quantities $\omega, \gamma, \Omega, \Gamma, \rho$ may then be obtained by applying the general tools of Lagrangian reduction. Such an approach, has been applied with success to a wide range of mechanical systems with symmetry, from fluid dynamics and imaging to rigid bodies and particles with broken symmetries. It is therefore of great interest to obtain such a description for the molecular strand. This objective will be the main goal of §3.2.

Perhaps not surprisingly, the description of such a complex system with nonlocal interactions needs a somewhat sophisticated version of the classical Lagrangian reduction. This will be explained here, by introducing the geometric setting via simpler examples. Besides the fact that the configuration space $\mathcal{F}([0, L], SE(3))$ is infinite dimensional, there are two major difficulties to overcome.

The first is related to the observation that the advected variables (Ω, Γ, ρ) are not acted on linearly by the Lagrangian variables $(\Lambda, \mathbf{r})$; see (1.1). In order to understand their evolution, an affine action needs to be introduced.

The second difficulty is related to the nonlocal dependence of the reduced Lagrangian on the unreduced variables $\Lambda, \mathbf{r}$. Remarkably, the geometrisation of this nonlocal dependence is also solved by the presence of the affine term (the cocycle) in the action, which allows the choice of the zero value for the reference condition $(\Omega_{ref}, \Gamma_{ref}, \rho_{ref})$ without leading to trivial dynamics.

3.1.2 Main goals of the Lagrangian approach

Besides the technical difficulties still to be overcome, a *classical* geometric description of the strand has many advantages. First, at a pure mathematical level, it rigorously justifies the ‘bare hands’ derivation of the equations derived in Chapter 2 by the Hamilton-Pontryagin method. The geometric description also explains the somewhat mysterious vanishing of the explicit dependence on the nonlocal terms in the final equation of motion; see (2.41).

At a more applied level, this symmetry reduced Lagrangian approach provides a guide towards a generalisation to higher dimensional versions or to other matrix Lie groups describing the charge conformation. Moreover, this classical approach is also suitable for coupling the molecular strand with fluid dynamics, since the two systems are now described by the same simple geometric framework; canonical Lagrangian description on the tangent bundle of the configuration space.

3.1.3 Pedagogical examples

Three preparatory examples, which illustrate the main idea behind the classical Lagrangian approach are now briefly described.

Example 3.1.1 (Euler-Poincaré) The simplest setting of Lagrangian reduction is that of a G -invariant Lagrangian $L : TG \rightarrow \mathbb{R}$, defined on the tangent bundle of its symmetry Lie group G . In this case, the Euler-Lagrange equations on TG can be reduced to provide equations on the Lie algebra $\mathfrak{g}$ known as the *Euler-Poincaré equations*

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \quad \xi = g^{-1} \dot{g} \in \mathfrak{g}, \quad (3.1)$$

for the reduced Lagrangian $l(g^{-1}\dot{g}) = L(g, \dot{g})$; see, e.g., MARSDEN & RATIU [2002]. Even such a simple setting has many applications. For example, Euler-Poincaré (EP) dynamics on the orthogonal group $G = SO(3)$ produces Euler's rigid body equations, while EP dynamics on the volume preserving diffeomorphisms yields the Euler equations for ideal fluid flows.

Example 3.1.2 (Euler-Poincaré with advected variables) When passing from the free rigid body to the heavy top, the direction of gravity breaks $SO(3)$ -invariance of the Lagrangian. In the framework of Euler-Poincaré theory, this is understood as follows. A G -invariant Lagrangian $L = L(g, \dot{g}, a) : TG \times V^* \rightarrow \mathbb{R}$ is given, where the (dual) vector space V^* contains the advected quantity a . Fixing a particular reference value $a_{ref} \in V^*$, breaks the symmetry and produces a physical Lagrangian $L_{a_{ref}} : TG \rightarrow \mathbb{R}$ that is only $G_{a_{ref}}$ -invariant. Here $G_{a_{ref}}$ denotes the isotropy group of the parameter a_{ref} .

For the heavy top, this corresponds to the choice of a fixed direction for gravity, whereas for compressible hydrodynamics, this choice corresponds to fixing the mass density ρ_{ref} of the fluid in the reference configuration for the Lagrangian representation. In the convective picture (for the heavy top) and Eulerian picture (for fluids), these quantities are *linearly* advected by the flow of the Euler-Lagrange equation. For example, in the case of compressible fluids, this is easily seen in the continuity equation

$$\dot{\rho} + \operatorname{div}(\rho u) = 0$$

for the mass density. As §1.2.2 pointed out, such a linear evolution does not appear in the strand and affine advection needs to be considered.

Returning to the abstract formulation, the G -invariant function $L = L(g, \dot{g}, a_{ref})$ determines the reduced Lagrangian $l(\xi, a) = L(g^{-1}\dot{g}, g^{-1}a_{ref})$ on the space $\mathfrak{g} \times V^*$, and the presence of the new variable a acted on by G modifies the right hand side of the Euler-Poincaré equation (3.1). The theory of Euler-Poincaré reduction with advected variables may be found in HOLM ET AL. [1998].

Example 3.1.3 (Affine Euler-Poincaré) The extension of classical Lagrangian reduction to the case of fluids with internal structure, such as superfluids or spin glasses, requires quantities that are *affinely* advected by the Lagrangian flow, as opposed to the *linearly* advected quantities of Example 3.1.2. This observation is

made in GAY-BALMAZ & RATIU [2008] to explain the presence of cocycles in the Hamiltonian structure of these equations (see also HOLM & KUPERSHMIT [1988]) from a canonical point of view (that is, by reduction of the canonical Hamilton equations on phase space).

Roughly speaking, the main idea is to replace the linear action in Example 3.1.2 by an affine action. Such an action is of the form $a \mapsto ga + c(g)$, where $c(g)$ is a group one-cocycle. As before, a G -invariant function $L = L(g, \dot{g}, a_{ref}) : TG \times V^* \rightarrow \mathbb{R}$ is given, where V^* is the space of affine advected quantities. Fixing a_{ref} , produces a Lagrangian $L_{a_{ref}}$ that is only invariant under the isotropy group $G_{a_{ref}}$ of the affine action at a_{ref} .

3.1.4 The strand with nonlocal interactions

§3.2 will show that the classical Lagrangian formulation for the molecular strand with nonlocal interactions is a special case of Example 3.1.3.

The main challenge is that the reduced Lagrangian, l given in (2.4), is defined only for the particular value $a_{ref} = 0$. Therefore it is necessary to construct a $G_{a_{ref}}$ -invariant Lagrangian $L_{a_{ref}} : TG \rightarrow \mathbb{R}$ for the unreduced system such that the equations of motion derived via Example 3.1.3 restrict to the CMS equations (2.41) upon evaluating the parameter $a_{ref} = 0$. Therefore, the reduced equations of motion, (2.41), live on the space $\mathfrak{g} \times \mathcal{O}$, where $\mathcal{O} \subset V^*$ is the orbit of $a_{ref} = 0$ under the affine action of G .

For application to the strand equations, the configuration Lie group is

$$G = \mathcal{F}([0, L], SE(3)) \ni g = (\Lambda, \mathbf{r}),$$

while the affine advected quantities are represented by the variables (Ω, Γ, ρ) .

When the nonlocal interactions are neglected, then the dynamics of the Kirchhoff rod in the convective representation are recovered. This simpler case and its link to Lagrange-Poincaré reduction and Clebsch-constrained variational principles are explored in GAY-BALMAZ ET AL. [2009].

§3.3 will explain how the equations of the charged strand may be obtained by classical Lagrangian reduction by the symmetry group. In particular, the material representation of the dynamics is governed by the standard Euler-Lagrange

equations on the tangent bundle of the configuration group $\mathcal{F}(I, SE(3))$. To see this, the process of affine Euler-Poincaré reduction is applied. This proves that the charged strand admits the same geometrical description as complex fluids and spin systems.

3.2 Background material

It is suitable to begin by recalling the theory of *affine Euler-Poincaré reduction* from GAY-BALMAZ & RATIU [2008] modified for *left* rather than right reduction.

3.2.1 Notations for semidirect products

Let V be a vector space and assume that the Lie group G acts on the *left* by linear maps (and hence G also acts on the left on the dual space V^*). As a set, the semidirect product $S = G \ltimes V$ is the Cartesian product $S = G \times V$ whose group multiplication is given by

$$(g_1, v_1)(g_2, v_2) = (g_1 g_2, v_1 + g_1 v_2),$$

where the action of $g \in G$ on $v \in V$ is denoted by concatenation, gv . The Lie algebra of S is the semidirect product Lie algebra, $\mathfrak{s} = \mathfrak{g} \ltimes V$, whose Lie bracket has the expression

$$\text{ad}_{(\xi_1, v_1)}(\xi_2, v_2) = [(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], \xi_1 v_2 - \xi_2 v_1), \quad (3.2)$$

where ξv denotes the induced action of $\mathfrak{g}$ on V , that is,

$$\xi v := \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)v \in V.$$

For $(\xi, v) \in \mathfrak{s}$ and $(\mu, a) \in \mathfrak{s}^*$,

$$\text{ad}_{(\xi, v)}^*(\mu, a) = (\text{ad}_\xi^* \mu - v \diamond a, -\xi a), \quad (3.3)$$

Classical formulation of molecular strands

where $\xi a \in V^*$ and $v \diamond a \in \mathfrak{g}^*$ are given by

$$\xi a := \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)a \quad \text{and} \quad \langle v \diamond a, \xi \rangle_{\mathfrak{g}} := -\langle \xi a, v \rangle_V,$$

and where $\langle \cdot, \cdot \rangle_{\mathfrak{g}} : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$ and $\langle \cdot, \cdot \rangle_V : V^* \times V \rightarrow \mathbb{R}$ are dual pairings. The coadjoint action of S on $\mathfrak{s}^*$ has the expression

$$\text{Ad}_{(g,v)^{-1}}^*(\mu, a) = (\text{Ad}_{g^{-1}}^* \mu + v \diamond ga, ga). \quad (3.4)$$

Given a left representation of G on the vector space V^* , an affine left representation $\theta_g(a) := ga + c(g)$ may be formed, where $c \in \mathcal{F}(G, V^*)$ is a **left group one-cocycle**, that is, it verifies the property

$$c(gh) = c(g) + gc(h), \quad (3.5)$$

for all $g, h \in G$. Note that

$$\left. \frac{d}{dt} \right|_{t=0} \theta_{\exp(t\xi)}(a) = \xi a + \mathbf{d}c(\xi)$$

and

$$\langle \xi a + \mathbf{d}c(\xi), v \rangle_V = \langle \mathbf{d}c^T(v) - v \diamond a, \xi \rangle_{\mathfrak{g}},$$

where $\mathbf{d}c : \mathfrak{g} \rightarrow V^*$ is defined by $\mathbf{d}c(\xi) := T_e c(\xi)$, and $\mathbf{d}c^T : V \rightarrow \mathfrak{g}^*$ is defined by

$$\langle \mathbf{d}c^T(v), \xi \rangle_{\mathfrak{g}} := \langle \mathbf{d}c(\xi), v \rangle_V.$$

Therefore, under the affine action associated with cocycle c , the formulae (3.3) and (3.4) become

$$\text{ad}_{(\xi,v)}^*(\mu, a) = (\text{ad}_{\xi}^* \mu - v \diamond a + \mathbf{d}c^T(v), -\xi a - \mathbf{d}c(\xi)) \quad (3.6)$$

$$\text{Ad}_{(g,v)^{-1}}^*(\mu, a) = (\text{Ad}_{g^{-1}}^* \mu + v \diamond ga - \mathbf{d}c^T(v), ga + c(g)). \quad (3.7)$$

3.2.2 Affine Euler-Poincaré reduction

The general theory of affine Euler-Poincaré reduction is described as follows:

Suppose a function $L : TG \times V^* \rightarrow \mathbb{R}$ is given that is left G -invariant under the

affine action $(v_h, a) \mapsto (gv_h, \theta_g(a)) = (gv_h, ga + c(g))$. In particular, if $a_{ref} \in V^*$, and a parametric family of Lagrangians $L_{a_{ref}} : TG \rightarrow \mathbb{R}$ may be defined according to $L_{a_{ref}}(v_g) := L(v_g, a_{ref})$. The Lagrangian $L_{a_{ref}}$ is left invariant under the lift to TG of the left action of $G_{a_{ref}}^c$ on G , where $G_{a_{ref}}^c$ is the isotropy group of a_{ref} with respect to the affine action θ .

The reduced Lagrangian $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$ is defined by $l := L|_{\mathfrak{g} \times V^*}$, and left G -invariance of L yields

$$l(g^{-1}v_g, \theta_{g^{-1}}(a)) = L(v_g, a)$$

for all $g \in G$, $v_g \in T_g G$, $a \in V^*$. For a curve $g(t) \in G$, let $\xi(t) := g(t)^{-1}\dot{g}(t)$ and $a(t)$ be the unique solution of the following affine differential equation with time dependent coefficients

$$\dot{a} = -\xi a - \mathbf{d}c(\xi), \quad (3.8)$$

with initial condition

$$a(0) = g^{-1}(0)a_{ref} + c(g^{-1}(0)) \quad \text{for } g(0) \in G. \quad (3.9)$$

The solution of (3.8) can then be written as

$$a(t) = \theta_{g^{-1}(t)}(a_{ref}) = g^{-1}(t)a_{ref} + c(g^{-1}(t)). \quad (3.10)$$

Theorem 3.2.1 *In the preceding notation, the following are equivalent:*

i *With a_{ref} held fixed, Hamilton's variational principle*

$$\delta \int_{t_0}^{t_1} L_{a_{ref}}(g, \dot{g}) dt = 0, \quad (3.11)$$

holds, for variations $\delta g(t)$ of $g(t)$ vanishing at the endpoints.

ii *$g(t)$ satisfies the Euler-Lagrange equations for $L_{a_{ref}}$ on G .*

iii *The constrained variational principle*

$$\delta \int_{t_0}^{t_1} l(\xi, a) dt = 0, \quad (3.12)$$

holds on $\mathfrak{g} \times V^*$, upon using variations of the form

$$\delta\xi = \frac{\partial\eta}{\partial t} + [\xi, \eta], \quad \delta a = -\eta a - \mathbf{d}c(\eta),$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints.

iv The affine Euler-Poincaré equations hold on $\mathfrak{g} \times V^*$:

$$\frac{\partial}{\partial t} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l}{\delta a} \right). \quad (3.13)$$

See GAY-BALMAZ & RATIU [2008] for the proof and applications to *spin systems* and *complex fluids*.

3.3 Application to the charged strand

Equipped with Theorem 3.2.1, this section will apply affine Euler-Poincaré reduction to the charged molecular strand. After some preliminary remarks concerning the peculiarities of the molecular strand problem, §3.3.1 will describe the compatibility between the treatment of non-local terms in the Hamilton-Pontryagin principle in Chapter 2 and the approach adopted by the affine Euler-Poincaré point of view. Then, the equations proper will be derived in §3.3.2. Relationships with previous work, generalisations and conservation laws will be given in §3.4.

As will be shown in §3.3.2, the molecular strand equations (2.41) arise from the choice $a_{ref} = 0$. In that case the isotropy group is

$$G_0^c = \{g \in G \mid c(g) = 0\}.$$

Given a G_0^c -invariant Lagrangian $L_0 : TG \rightarrow \mathbb{R}$, the reduced Lagrangian l is defined on $\mathfrak{g} \times \mathcal{O}_0^c$ by

$$l(\xi, c(g^{-1})) = L_0(g\xi).$$

It is sufficient to restrict to Lagrangians for simple mechanical systems with symmetry, that is, of the form $L_0(v_g) = K(v_g) - P(g)$, where K is the kinetic energy

associated to a G_0^c -invariant Riemannian metric on G and the potential P is G_0^c -invariant. In this case, the reduced Lagrangian is

$$l(\xi, c(g^{-1})) = K(g\xi) - P(g). \quad (3.14)$$

Note that the right hand side of (3.14) is well defined on $\mathfrak{g} \times \mathcal{O}_0^c$, that is, it depends on g only through $c(g^{-1})$. Indeed, $c(g^{-1}) = c(h^{-1})$ if and only if $\theta_{g^{-1}}(0) = \theta_{h^{-1}}(0)$, which means that $hg^{-1} \in G_0^c$. Therefore, $P(h) = P((hg^{-1})g) = P(g)$ by left G_0^c -invariance of P . For the kinetic energy the same argument works since the metric is G_0^c -invariant.

Thus it is possible to write $L_0(v_g) = K(v_g) - E(c(g^{-1}))$ for the function $E : V^* \rightarrow \mathbb{R}$ uniquely determined by the relation $P(g) = E(c(g^{-1}))$. In this case,

$$l(\xi, c(g^{-1})) = K(g\xi) - E(c(g^{-1})).$$

For the Lagrangian of the charged molecular strand the potential energy is the sum of two terms. One term, denoted by E_{loc} , may be explicitly expressed in terms of $c(g^{-1})$. The other, denoted by E_{np} , does not have a concrete expression in terms of $c(g^{-1})$, although it is G_0^c -invariant. In addition, for the charged molecular strand the kinetic energy metric is not just G_0^c -invariant but G -invariant, which implies it is only a function of $\xi \in \mathfrak{g}$. For the molecular strand the Lagrangian is of the form

$$L_0(v_g) = K(v_g) - E_{loc}(c(g^{-1})) - E_{np}(\zeta(g), c(g^{-1})),$$

where ζ is a G_0^c -invariant function defined on G and the reduced Lagrangian is

$$\begin{aligned} l(\xi, c(g^{-1})) &= \underbrace{K(\xi) - E_{loc}(c(g^{-1}))}_{=l_{loc}} - E_{np}(\zeta(g), c(g^{-1})) \\ &= l_{loc}(\xi, c(g^{-1})) + l_{np}(\zeta(g), c(g^{-1})). \end{aligned}$$

Note that l can be expressed in terms of $(\xi, a) \in \mathfrak{g} \times \mathcal{O}_0^c$ as

$$l(\xi, a) = K(\xi) - E_{loc}(a) - E_{np}(\zeta(g_a), a) = l_{loc}(\xi, a) + l_{np}(\zeta(g_a), a), \quad (3.15)$$

where $g_a \in G$ is such that $c(g_a^{-1}) = a$. This g_a is determined only up to left multipli-

Classical formulation of molecular strands

cation by G_0^c . Since E_{np} is G_0^c -invariant, the function $a \mapsto E_{np}(g_a)$ is well-defined. Despite $a \mapsto l_{np}(\zeta(g_a), a)$ being a well-defined function of $a \in \mathcal{O}_0^c$, there is no explicit expression for $l_{np} : \mathcal{O}_0^c \rightarrow \mathbb{R}$ for the molecular strand. Therefore, the terminology $l_{np} = l_{np}(\zeta(g_a), a)$ rather than $l_{np} = l_{np}(a)$ must be used for explicit calculations; see (2.13).

It is now timely to identify all relevant objects that appear in the dynamics of the molecular strand as an example of this abstract setup. The Lie group is $G = \mathcal{F}([0, L], SE(3))$, its Lie algebra is $\mathfrak{g} = \mathcal{F}([0, L], \mathfrak{se}(3))$, the representation space is $V = \mathfrak{X}([0, L], \mathfrak{se}^*(3)) \oplus \mathcal{F}([0, L], \mathbb{R}^3)$, and its dual is $V^* = \Omega^1([0, L], \mathfrak{se}(3)) \oplus \mathcal{F}([0, L], \mathbb{R}^3)$.

The spaces $\Omega^1([0, L], \mathfrak{se}(3))$ and $\mathfrak{X}([0, L], \mathbb{R}^3)$ are vector valued one-forms and vector fields (contravariant one-tensors). Of course, since $[0, L]$ is one-dimensional, both of these spaces are naturally identified with smooth functions taking values in the respective vector spaces. However, in view of the generalisation presented here, it is useful to think of these spaces as described. Also, the various $\mathbb{R}^3$ appearing above play different roles: they can be the Lie algebra of $\mathfrak{so}(3)$, its dual, the natural representation space of $SO(3)$, or the dual of the representation space. When discussing the generalisation here, all these spaces are different.

The variables associated with these spaces are the following. Elements of G are $(\Lambda, \mathbf{r})$, where $\Lambda \in \mathcal{F}([0, L], SO(3))$ and $\mathbf{r} \in \mathcal{F}([0, L], \mathbb{R}^3)$. Elements of $\mathfrak{g}$ are (ω, γ) , with $\omega, \gamma \in \mathcal{F}([0, L], \mathbb{R}^3)$. Finally, elements of V^* are (Ω, Γ, ρ) , where $\Omega \in \Omega^1([0, L], \mathbb{R}^3)$, $\Gamma \in \Omega^1([0, L], \mathbb{R}^3)$, and $\rho \in C^\infty([0, L], \mathbb{R}^3)$.

The V^* -valued one-cocycle on G is given by

$$c((\Lambda, \mathbf{r})^{-1}) = ((\Lambda, \mathbf{r})^{-1}(\Lambda, \mathbf{r})', \Lambda^{-1}\mathbf{r}) = (\Lambda^{-1}\Lambda', \Lambda^{-1}\mathbf{r}', \Lambda^{-1}\mathbf{r}) = (\Omega, \Gamma, \rho). \quad (3.16)$$

The function ζ appearing in l_{np} is given in this case by

$$\zeta(s, s') = (\xi(s, s'), \kappa(s, s')) = (\Lambda, \mathbf{r})^{-1}(s)(\Lambda, \mathbf{r})(s') \in SE(3),$$

so $\zeta(_, s') \in G$. Note that the Lagrangian of the strand (see (2.5), (2.12), and (2.13)) is exactly of the form (3.15), with $\zeta(s, s') = (\xi(s, s'), \kappa(s, s')) \in SE(3)$ given above. In fact, (2.13) has an expression of the type $l_{np} = l_{np}(\zeta(g_a), a)$. These comments will be expanded upon and explained in detail in §3.3.2.

3.3.1 Recovering the Hamilton-Pontryagin approach

Theorem 3.2.1 demonstrates that the Euler-Lagrange equations of a G_0^c -invariant Lagrangian $L_0 : TG \rightarrow \mathbb{R}$ are equivalent to the affine Euler-Poincaré equations for a reduced Lagrangian $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$, that is,

$$\frac{\partial}{\partial t} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l}{\delta a} \right), \quad (3.17)$$

together with the advection and reconstruction relations

$$a(t) = c(g^{-1}(t)) \quad (3.18)$$

$$\dot{g} = g\xi. \quad (3.19)$$

In order to write these equations, the reduced Lagrangian l needs to be extended to $\mathfrak{g} \times V^*$. That is, a Lagrangian $L : TG \times V^* \rightarrow \mathbb{R}$ is required such that $L(v_g, 0) = L_0(v_g)$.

For the molecular strand, there is an additional complication coming from the fact that the Lagrangian

$$l(\xi, a) = l_{loc}(\xi, a) + l_{np}(\zeta(g_a), a), \quad (3.20)$$

being a well-defined function of $(\xi, a) \in \mathfrak{g} \times V^*$, is not explicitly written in terms of a . Therefore, when computing the affine Euler-Poincaré equations in concrete examples, there is still a dependence on g_a in the final equation, although this dependence may be replaced by a dependence in a uniquely, by the results above.

The next step is to apply the variational principle from Theorem 3.2.1 to the Lagrangian (3.20). Let $g(t)$ be a given curve in G . Take a family of curves $g_\varepsilon(t)$ satisfying $g_0(t) = g(t)$ and denote $\eta(t) := g^{-1}(t)\delta g(t)$, where $\delta g(t) = \partial_\varepsilon g_\varepsilon(t)|_{\varepsilon=0}$. Then $\delta \int_{t_0}^{t_1} l(\xi(t), c(g(t)^{-1})) dt = 0$ implies

$$\frac{\partial}{\partial t} \frac{\delta l_{loc}}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l_{loc}}{\delta \xi} + \frac{\delta(l_{loc} + l_{np})}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta(l_{loc} + l_{np})}{\delta a} \right) + g^{-1} \frac{\delta l_{np}}{\delta \zeta} T_g \zeta. \quad (3.21)$$

Note that (3.21) is the abstract generalisation of (2.27) and (2.28).

Recall from the abstract theory that l_{np} depends only on $a \in \mathcal{O}_0^c$. However, l_{np}

Classical formulation of molecular strands

is given as a function of $(\zeta(g), c(g^{-1}))$. Let

$$\left. \frac{\delta l_{np}}{\delta a} \right|_{Tot}$$

denote the functional derivative of l_{np} viewed as a function of $a \in \mathcal{O}_0^c$ only. Since every curve in $\mathcal{O}_0^c$ through $a = c(g^{-1}) \in \mathcal{O}_0^c$ is of the form $c(g_\varepsilon^{-1})$, where $g_0 = g$,

$$\begin{aligned} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} l_{np}(\zeta(g_\varepsilon), c(g_\varepsilon^{-1})) &= \left\langle \left. \frac{\delta l_{np}}{\delta a} \right|_{Tot}, \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} c(g_\varepsilon^{-1}) \right\rangle = - \left\langle \left. \frac{\delta l_{np}}{\delta a} \right|_{Tot}, \eta a + \mathbf{d}c(\eta) \right\rangle \\ &= \left\langle \left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} \diamond a - \mathbf{d}c^T \left(\left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} \right), \eta \right\rangle, \end{aligned} \quad (3.22)$$

where $\eta := g^{-1}\delta g$. On the other hand,

$$\begin{aligned} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} l_{np}(\zeta(g_\varepsilon), c(g_\varepsilon^{-1})) &= \left\langle \frac{\delta l_{np}}{\delta \zeta}, T_g \zeta(g\eta) \right\rangle - \left\langle \left. \frac{\delta l_{np}}{\delta a} \right|_{Tot}, \eta a + \mathbf{d}c(\eta) \right\rangle \\ &= \left\langle g^{-1} \frac{\delta l_{np}}{\delta \zeta} T_g \zeta + \left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} \diamond a - \mathbf{d}c^T \left(\left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} \right), \eta \right\rangle. \end{aligned} \quad (3.23)$$

Equations (3.22) and (3.23) prove the following identity:

$$\left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} \diamond a - \mathbf{d}c^T \left(\left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} \right) = \frac{\delta l_{np}}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l_{np}}{\delta a} \right) + g^{-1} \frac{\delta l_{np}}{\delta \zeta} T_g \zeta, \quad (3.24)$$

where $a = c(g^{-1})$. Substituting (3.24) into (3.21) results in the affine Euler-Poincaré equations (3.17), since

$$\frac{\delta l}{\delta a} = \left. \frac{\delta l_{np}}{\delta a} \right|_{Tot} + \frac{\delta l_{loc}}{\delta a}.$$

Thus, affine Euler-Poincaré reduction theory reproduces the equations of motion resulting from the Hamilton-Pontryagin approach described in Chapter 2. Note that the affine Euler-Poincaré form of the equations is given in (2.41), while the explicit equations (2.27) — (2.28), derived via the Hamilton-Pontryagin principle, must use $\delta l_{np}/\delta a$ rather than $\delta l_{np}/\delta a|_{Tot}$.

3.3.2 Classical strand equations

Recall that the variables needed for this problem consist of the Lagrangian quantities $g = (\Lambda, \mathbf{r}) : [0, L] \rightarrow SE(3)$, together with the convective variables defined by

$$\Omega := \Lambda^{-1}\Lambda', \quad \omega := \Lambda^{-1}\dot{\Lambda} : [0, L] \rightarrow \mathfrak{so}(3)$$

and

$$\Gamma := \Lambda^{-1}\mathbf{r}', \quad \gamma := \Lambda^{-1}\dot{\mathbf{r}}, \quad \rho := \Lambda^{-1}\mathbf{r} : [0, L] \rightarrow \mathbb{R}^3.$$

In order to give a more transparent vision of the underlying geometric structures, consider the n -dimensional generalisation described in §2.1.3, that is, replace the interval $[0, L]$ by an arbitrary manifold $\mathcal{D}$ and $SE(3)$ by the semidirect product $S = \mathcal{O} \ltimes E$ of a Lie group $\mathcal{O}$ with a *left* representation space E . Given a manifold $\mathcal{D}$, define the group $G := \mathcal{F}(\mathcal{D}, S)$ and the dual vector space $V^* := \Omega^1(\mathcal{D}, \mathfrak{s}) \oplus \mathcal{F}(\mathcal{D}, E)$. The elements of the group G are denoted by (Λ, r) , where $\Lambda : \mathcal{D} \rightarrow \mathcal{O}$ and $r : \mathcal{D} \rightarrow E$. The elements of V^* are denoted by (Ω, Γ, ρ) , where $\Omega \in \Omega^1(\mathcal{D}, \mathfrak{o})$, $\Gamma \in \Omega^1(\mathcal{D}, E)$, and $\rho : \mathcal{D} \rightarrow E$. The space V^* may be seen as the dual of $V = \mathfrak{X}(\mathcal{D}, \mathfrak{s}^*) \oplus \mathcal{F}(\mathcal{D}, E^*)$, where $\mathfrak{X}(\mathcal{D}, \mathfrak{s})$ is the space of $\mathfrak{s}$ -valued vector fields on $\mathcal{D}$.

Consider the representation of G on V^* defined by

$$(\Lambda, r)(\Omega, \Gamma, \rho) = (\text{Ad}_{(\Lambda, r)}(\Omega, \Gamma), \Lambda\rho) \quad (3.25)$$

where the adjoint action is that of S , acting here on functions defined on $\mathcal{D}$, and $\Lambda\rho$ denotes the left representation of $\mathcal{O}$ on E , acting on functions. The main object for this approach is the group one-cocycle c already appearing implicitly in the definition of the variables Ω, Γ, ρ in (1.1), (2.18), and explicitly in (2.21). Recall that it is given by (3.16), now rewritten as

$$c(\Lambda, r) := ((\Lambda, r)\mathbf{d}(\Lambda, r)^{-1}, -r). \quad (3.26)$$

Next, the cocycle identity is verified for the first component $(\Lambda, r)\mathbf{d}(\Lambda, r)^{-1}$. To

Classical formulation of molecular strands

simplify notation, denote $g_i := (\Lambda_i, r_i) \in \mathcal{F}(\mathcal{D}, S)$, $i \in \{1; 2\}$. This yields

$$\begin{aligned} g_1 g_2 \mathbf{d}(g_1 g_2)^{-1} &= g_1 g_2 \mathbf{d}(g_2^{-1} g_1^{-1}) = g_1 g_2 \mathbf{d}(g_2^{-1}) g_1^{-1} + g_1 g_2 g_2^{-1} \mathbf{d}(g_1^{-1}) \\ &= \text{Ad}_{g_1} (g_2 \mathbf{d} g_2^{-1}) + g_1 \mathbf{d}(g_1^{-1}). \end{aligned}$$

Since the second coordinate of $((\Lambda_1, r_1)(\Lambda_2, r_2))$ is equal to $r = r_1 + \Lambda_1 r_2$, one finds

$$\begin{aligned} c((\Lambda_1, r_1)(\Lambda_2, r_2)) &= \\ &= (\text{Ad}_{(\Lambda_1, r_1)} ((\Lambda_2, r_2) \mathbf{d}(\Lambda_2, r_2)^{-1}) + (\Lambda_1, r_1) \mathbf{d}((\Lambda_1, r_1)^{-1}), -r_1 - \Lambda_1 r_2) \\ &= (\text{Ad}_{(\Lambda_1, r_1)} ((\Lambda_2, r_2) \mathbf{d}(\Lambda_2, r_2)^{-1}), -\Lambda_1 r_2) + ((\Lambda_1, r_1) \mathbf{d}((\Lambda_1, r_1)^{-1}), -r_1) \\ &= (\Lambda_1, r_1) c(\Lambda_2, r_2) + c(\Lambda_1, r_1). \end{aligned}$$

This shows that c verifies the cocycle property (3.5) relative to the representation (3.25). Note that the first component of c is the left version of the cocycle appearing in the theory of complex fluids; see GAY-BALMAZ & RATIU [2008]. From the covariant perspective, which will be discussed fully in Chapter 4, this quantity is considered to be a connection form on a principal bundle rather than a cocycle.

Let $g_{ref} = (\Lambda_{ref}, \mathbf{r}_{ref}) : \mathcal{D} \rightarrow S$ be a reference configuration for the strand. That is $\mathbf{r}_{ref} : \mathcal{D} \rightarrow E$ describes a reference embedded submanifold, and $\Lambda_{ref} : \mathcal{D} \rightarrow \mathcal{O}$ describes the reference group element at points on the submanifold. Consider the relative variable $\tilde{g} \in S$ that describes the configuration g with respect to the reference configuration g_{ref} . The relative variable $\tilde{g} = (\tilde{\Lambda}, \tilde{\mathbf{r}}) : \mathcal{D} \rightarrow S$ is therefore given by $g_{ref} \tilde{g} = g$.

Using the tilde ($\tilde{\cdot}$) notation for relative quantities, the subscript $_{ref}$ for reference quantities, and unadorned notation for the reduced variables as in (1.1); the transformation from the reduced to relative variables is described by

$$\begin{aligned} (\Omega, \Gamma) &= \text{Ad}_{\tilde{g}^{-1}} (\Omega_{ref}, \Gamma_{ref}) + (\tilde{\Omega}, \tilde{\Gamma}) \\ (\omega, \gamma) &= (\tilde{\omega}, \tilde{\gamma}) \\ \rho &= \tilde{\Lambda}^{-1} \mathbf{r}_{ref} + \tilde{\rho}. \end{aligned} \tag{3.27}$$

Therefore the reduced variables Ω, Γ and ρ may be interpreted as affinely advected quantities in the relative representation. That is, combining (3.27) with (3.25) and

(3.26) results in the equation

$$(\Omega, \Gamma, -\rho) = \tilde{g}^{-1} a_{ref} + c(\tilde{g}^{-1}) = \theta_{\tilde{g}^{-1}}(a_{ref}) = \tilde{a}, \quad (3.28)$$

where $a_{ref} := (\Omega_{ref}, \Gamma_{ref}, -r_{ref}) = c(g_{ref})$ has been defined. Note that when $a_{ref} = 0$ then $c(g^{-1}) = (\Omega, \Gamma, -\rho) = c(\tilde{g}^{-1})$. The remaining variables, (ω, γ) , remain unchanged under the transformation. Consequently, the Lagrangian $L_0(g, \dot{g}) = l(g^{-1}\dot{g}, c(g^{-1}))$ may be extended to a function $\tilde{L}(\tilde{g}, \dot{\tilde{g}}, a_{ref}) = l(g^{-1}\dot{g}, c(g^{-1}))$, which satisfies the relation $\tilde{L}(g, \dot{g}, 0) = L_0(g, \dot{g})$. That is, $\tilde{L} : TG \times V^* \rightarrow \mathbb{R}$ is an extension of L_0 .

Applying Theorem 3.2.1 to the Lagrangian $\tilde{L}(g, \dot{g}, 0)$, and using the expressions

$$(u, w, f) \diamond (\Omega, \Gamma, \rho) = (\text{ad}_{\Omega_i}^* u^i + w^i \diamond \Gamma_i + f \diamond \rho, -\Omega_i w^i),$$

$$\mathbf{d}c(\omega, \gamma) = (-\mathbf{d}\omega, -\mathbf{d}\gamma, -\gamma), \quad \text{and} \quad \mathbf{d}c^T(u, w, f) = (\text{div}(u), \text{div}(w) - f),$$

the affine Euler-Poincaré equations (3.13) become

$$\begin{cases} (\partial_t - \text{ad}_{\omega}^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_{\Omega}^*) \frac{\delta l}{\delta \Omega} = \frac{\delta l}{\delta \gamma} \diamond \gamma + \frac{\delta l}{\delta \Gamma} \diamond \Gamma + \frac{\delta l}{\delta \rho} \diamond \rho, \\ (\partial_t + \omega) \frac{\delta l}{\delta \gamma} + (\text{div} + \Omega) \frac{\delta l}{\delta \Gamma} = \frac{\delta l}{\delta \rho}. \end{cases} \quad (3.29)$$

The *advection relations* are

$$\begin{cases} \partial_t \Omega + \text{ad}_{\omega} \Omega = \mathbf{d}\omega, \\ (\partial_t + \omega) \Gamma = (\mathbf{d} + \Omega) \gamma, \\ \partial_t \rho + \omega \rho = \gamma. \end{cases} \quad (3.30)$$

Together (3.29) and (3.30) are therefore equivalent to the Euler-Lagrange equations for Lagrangian $L_0(g, \dot{g})$.

Remark 3.3.1 To write these equations, one must suppose that the dynamics is described by a Lagrangian l given explicitly in terms of the variables $(\omega, \gamma, \Omega, \Gamma, \rho)$. Equivalently, one must assume that l is explicitly expressible by an affine left-invariant Lagrangian L defined on $TG \times V^*$. As §2.1.1 pointed out, such a hypothesis is not verified when nonlocal terms are taken into account. For the molec-

ular strand, the local Kirchhoff situation was dealt with in §3.4.1, while the more general nonlocal strand is left to §3.4.2.

Remark 3.3.2 (Affine reduction at a fixed parameter) The approach taken here is different from that of ELLIS ET AL. [2010] where a new form of reduction called *affine reduction at a fixed parameter* is developed specifically for the molecular strand problem. The reduction at fixed parameter approach considers $l : \mathfrak{g} \times \mathcal{O}_0^c \rightarrow \mathbb{R}$ that is only defined on the orbit $\mathcal{O}_0^c \subset V^*$, bypassing the calculation of an extended reduced Lagrangian $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$. The approach adopted here requires the calculation of an extended Lagrangian, but does not refer to new abstract results beyond those contained in GAY-BALMAZ & RATIU [2008].

3.4 Relation to other formulations

3.4.1 Elastic filament dynamics and Kirchhoff's theory

Suppose that the dynamics of the strand is described by a Lagrangian

$$l = l(\omega, \gamma, \Omega, \Gamma, \rho)$$

defined on $\mathfrak{g} \times V^*$, where $\mathfrak{g} = \mathcal{F}(\mathcal{D}, \mathfrak{s})$ and $V^* = \Omega^1(\mathcal{D}, \mathfrak{s}) \oplus \mathcal{F}(\mathcal{D}, E)$. The Lagrangian l is induced by a left invariant Lagrangian L defined on $TG \times V^*$, where $G = \mathcal{F}(\mathcal{D}, S)$.

Note that there is no restriction in the way l depends on the variables. In particular the dependence can be nonlocal. However, it is assumed here that l depends explicitly on the variables $(\omega, \gamma, \Omega, \Gamma, \rho)$. Recall that such an hypothesis is verified for the Lagrangian of Kirchhoff's theory (1.13) but not for the Lagrangian of the charged strand (2.5).

The explicit Kirchhoff equations may be recovered as follows. Fix the reference values $(\Omega_{ref}, \Gamma_{ref}, \rho_{ref})$ and define the Lagrangian

$$L_{(\Omega_{ref}, \Gamma_{ref}, \rho_{ref})}(\Lambda, r) := L(\Lambda, r, \Omega_{ref}, \Gamma_{ref}, \rho_{ref}).$$

Consider a curve $(\Lambda, r) \in G$ and define the quantities

$$\begin{aligned} (\Omega, \Gamma, \rho) &= (\Lambda, r)^{-1}(\Omega_{ref}, \gamma_{ref}, \rho_{ref}) + c((\Lambda, r)^{-1}) \\ &= (\text{Ad}_{\Lambda^{-1}} \Omega_{ref}, \Lambda^{-1}(\Gamma_{ref} + \Omega_{ref}r), \Lambda^{-1}\rho_{ref}) + (\Lambda^{-1}\mathbf{d}\Lambda, \Lambda^{-1}\mathbf{d}r, \Lambda^{-1}r). \end{aligned}$$

and

$$\omega = \Lambda^{-1}\dot{\Lambda}, \quad \gamma = \Lambda^{-1}\dot{r}.$$

Note that when the initial values $\Omega_{ref}, \Gamma_{ref}, \rho_{ref}$ are zero, the definitions of the variables $\omega, \gamma, \Omega, \Gamma, \rho$ coincide with those given in (1.1) and (2.18).

Then the curve (Λ, r) is a solution of the Euler-Lagrange equations associated with $L_{(\Omega_{ref}, \Gamma_{ref}, r_0)}$ on TG if and only if $(\omega, \gamma, \Omega, \Gamma, \rho)$ is a solution of the Euler-Poincaré equations (3.29).

When $\mathcal{D}$ is the interval $[0, L]$ and S is the semidirect product of $\mathcal{O} = SO(3)$ with $E = \mathbb{R}^3$, then the dynamical equations of the charged strand (2.41) are recovered from (3.29), since

$$\text{ad}^* \mapsto - \times \quad \text{and} \quad \diamond \mapsto \times.$$

These equations are the convective representation of Kirchhoff's equations. The advection relations, which were derived §2.1.2, are recovered from (3.30).

3.4.2 The charged molecular strand: general case

Recall from §2.1.1 that the Lagrangian of the charged strand has the expression

$$l = l_{loc}(\omega, \gamma, \Omega, \Gamma, \rho) + l_{np}(\xi, \kappa, \Gamma),$$

where l_{loc} is a local function of the form

$$l_{loc}(\omega, \gamma, \Omega, \Gamma, \rho) = K(\omega, \gamma) - E_{loc}(\Omega, \Gamma, \rho) \tag{3.31}$$

and l_{np} is of the form

$$l_{np}(\xi, \kappa, \Gamma) = \iint U(\xi(s, s'), \kappa(s, s'), \Gamma(s), \Gamma(s')) \, ds \, ds',$$

where

$$U : SE(3) \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} \quad \text{and} \quad (\xi(s, s'), \kappa(s, s')) := (\Lambda, \mathbf{r})^{-1}(s)(\Lambda, \mathbf{r})(s').$$

Remark 3.4.1 (Two crucial observations) Note the following:

1. The nonlocal Lagrangian l_{np} is induced by an $SO(3)$ -invariant potential $E_{np} = E_{np}(\Lambda, \mathbf{r})$. Thus the total Lagrangian l may be seen to be induced by the $SO(3)$ -invariant Lagrangian $L_0 = L_0(\Lambda, \dot{\Lambda}, \mathbf{r}, \dot{\mathbf{r}})$ given by

$$L_0(\Lambda, \dot{\Lambda}, \mathbf{r}, \dot{\mathbf{r}}) = K(\Lambda, \dot{\Lambda}, \mathbf{r}, \dot{\mathbf{r}}) - E_{loc}(c((\Lambda, \mathbf{r})^{-1})) - E_{np}(\Lambda, \mathbf{r}),$$

where K is the $\mathcal{F}(\mathcal{D}, SE(3))$ -left invariant extension of the kinetic energy K in (3.31). Note that the dependence of E_{loc} on (Ω, Γ, ρ) has been replaced by a dependence on $(\Lambda, \mathbf{r})$ through the cocycle c . The affine Euler-Poincaré dynamics yield the relation $(\Omega, \Gamma, \rho) = c((\Lambda, \mathbf{r})^{-1})$ which allows us to recover the dependence of the potential on (Ω, Γ, ρ) .

2. The group $SO(3)$ is precisely the isotropy group

$$G_0^c = \mathcal{F}(\mathcal{D}, SE(3))_0^c = \{(\Lambda, \mathbf{r}) \in G \mid c(\Lambda, \mathbf{r}) = 0\}$$

of the affine action at zero.

These two remarks enable the dynamics of the molecular strand by the affine reduction processes described in Theorem 3.2.1. As before, it is convenient to work with the general framework involving $\mathcal{D}$ and $\mathcal{O} \otimes E$. The present approach is applicable to any $\mathcal{O}$ -invariant Lagrangian

$$L_0 = L_0(\Lambda, \dot{\Lambda}, r, \dot{r}) : T[\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)] \rightarrow \mathbb{R}.$$

Note there are no conditions on the dependence of L_0 on the variables (Λ, r) . In particular, L_0 may be nonlocal, and may depend on the derivatives of Λ and r . An important class of such Lagrangians is given by

$$L_0(\Lambda, \dot{\Lambda}, r, \dot{r}) = K(\Lambda, \dot{\Lambda}, r, \dot{r}) - P(\Lambda, r),$$

where K is the kinetic energy associated to an $\mathcal{O}$ -invariant metric on $\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)$ and the potential P is an $\mathcal{O}$ -invariant function on $\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)$. In particular, P may be nonlocal, or depend on derivatives of Λ and r ; see (1.13) for an example. In the case of the molecular strand, K is assumed to be left-invariant and P is given by

$$P(\Lambda, r) = E_{loc} (c((\Lambda, r)^{-1})) + E_{np}(\Lambda, r),$$

where

$$E_{np}(\Lambda, r) := \iint_{\mathcal{D}} U(\xi(s, s'), \kappa(s, s'), \Lambda^{-1} \mathbf{d}r(s), \Lambda^{-1} \mathbf{d}r(s')) \, \mathbf{d}s \, \mathbf{d}s'$$

$$(\xi(s, s'), \kappa(s, s')) := (\Lambda, r)^{-1}(s)(\Lambda, r)(s') \in \mathcal{O} \otimes E$$

and one readily sees that E_{np} is $\mathcal{O}$ -invariant. Recall that the cocycle is

$$c((\Lambda, r)^{-1}) = (\Lambda^{-1} \mathbf{d}\Lambda, \Lambda^{-1} \mathbf{d}r, \Lambda^{-1} r).$$

Thus, a straightforward and potentially useful generalisation of E_{np} is

$$E_{np}(\Lambda, r) := \iint_{\mathcal{D}} U\left(\xi(s, s'), \kappa(s, s'), c((\Lambda, r)^{-1})(s), c((\Lambda, r)^{-1})(s')\right) \, \mathbf{d}s \, \mathbf{d}s'.$$

Theorem 3.2.1 with L_0 yields the same affine Euler-Poincaré equations (3.29), where all derivatives are total derivatives. Equivalently the Hamilton-Pontryagin approach may be used to obtain the equations

$$\left\{ \begin{array}{l} (\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} = \frac{\delta l}{\delta \gamma} \diamond \gamma + \frac{\delta l}{\delta \Gamma} \diamond \Gamma + \frac{\delta l}{\delta \rho} \diamond \rho, \\ \quad + \int \left[\xi(s, s') \frac{\partial U}{\partial \xi}(s', s) - \frac{\partial U}{\partial \xi}(s, s') \xi(s', s) - \kappa(s, s') \diamond \frac{\partial U}{\partial \kappa}(s, s') \right] \mathbf{d}s' \\ (\partial_t + \omega) \frac{\delta l}{\delta \gamma} + (\text{div} + \Omega) \frac{\delta l}{\delta \Gamma} = \frac{\delta l}{\delta \rho} + \int \left[\xi(s, s') \frac{\partial U}{\partial \kappa}(s', s) - \frac{\partial U}{\partial \kappa}(s, s') \right] \mathbf{d}s'. \end{array} \right. \quad (3.32)$$

Note that here the derivatives are not total derivatives, see the discussion in §3.3.1.

3.4.3 Conservation laws and spatial formulation

This paragraph generalises the approach taken in §2.3 and reformulates the equations (3.29) for the generalised charged strand as a conservation law. The first step is to provide an n -dimensional generalisation of formula (2.31). Given a Lie group G , a map $g : \mathcal{D} \rightarrow G$ defined on a n -dimensional manifold $\mathcal{D}$, $s \in \mathcal{D}$, and a $\mathfrak{g}^*$ -valued vector field w on $\mathcal{D}$, one has

$$\text{Ad}_g^* [\text{div} (\text{Ad}_{g^{-1}}^* w)] = \text{div } w - \text{ad}_{\sigma_i}^* w^i =: \text{div}^\sigma w, \quad \sigma := g^{-1} \mathbf{d}g \in \Omega^1(\mathcal{D}, \mathfrak{g}). \quad (3.33)$$

This formula, (2.31), the expression of ad^* associated to the semidirect product $\mathcal{O} \ltimes E$, and the equalities

$$(\omega, \gamma) = (\Lambda, r)^{-1}(\dot{\Lambda}, \dot{r}), \quad (\Omega, \Gamma) = (\Lambda, r)^{-1} \mathbf{d}(\Lambda, r),$$

combine to give

$$\begin{aligned} \text{Ad}_{(\Lambda, r)}^* \frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) \right] \\ = \frac{\partial}{\partial t} \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) + \left(-\text{ad}_\omega^* \frac{\delta l}{\delta \omega} + \gamma \diamond \frac{\delta l}{\delta \gamma}, \omega \frac{\delta l}{\delta \gamma} \right), \end{aligned}$$

and

$$\begin{aligned} \text{Ad}_{(\Lambda, r)}^* \text{div} \left[\text{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \Omega}, \frac{\delta l}{\delta \Gamma} \right) \right] \\ = \text{div} \left(\frac{\delta l}{\delta \Omega}, \frac{\delta l}{\delta \Gamma} \right) + \left(-\text{ad}_\Omega^* \frac{\delta l}{\delta \Omega} + \Gamma \diamond \frac{\delta l}{\delta \Gamma}, \Omega \frac{\delta l}{\delta \Gamma} \right). \end{aligned}$$

Thus, equations (3.32) can be rewritten in the form of a conservation law, namely

$$\begin{aligned} \frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) \right] + \text{div} \left[\text{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \Omega}, \frac{\delta l}{\delta \Gamma} \right) \right] \\ = \text{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho, \frac{\delta l}{\delta \rho} \right). \end{aligned} \quad (3.34)$$

Using (3.4), the right hand side expands to

$$\begin{aligned} \text{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho, \frac{\delta l}{\delta \rho} \right) &= \left(\text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho \right) + r \diamond \left(\Lambda \frac{\delta l}{\delta \rho} \right), \Lambda \frac{\delta l}{\delta \rho} \right) \\ &= \left(\left(\Lambda \frac{\delta l}{\delta \rho} \diamond \Lambda \rho \right) + r \diamond \left(\Lambda \frac{\delta l}{\delta \rho} \right), \Lambda \frac{\delta l}{\delta \rho} \right) = \left(0, \Lambda \frac{\delta l}{\delta \rho} \right), \end{aligned}$$

since $\rho = \Lambda^{-1}r$. Note that this is the exact analogue of the simplification (1.27).

Such a conservation law is valid for each solution of the affine Euler-Poincaré equation (3.2.1) associated to a G_0^c -invariant Lagrangian $L_0 : TG \rightarrow \mathbb{R}$. In particular, it is valid for the Kirchhoff's theory as is shown at end of §1.3.1.

A short computation shows that, in general, the conservation law reads

$$\frac{\partial}{\partial t} \left[\text{Ad}_{g^{-1}}^* \frac{\delta l}{\delta \xi} \right] + \mathbf{d}c^T \left(g \frac{\delta l}{\delta a} \right) = 0. \quad (3.35)$$

When a_{ref} is not necessarily zero, the formula becomes

$$\frac{\partial}{\partial t} \left[\text{Ad}_{g^{-1}}^* \frac{\delta l}{\delta \xi} \right] + \mathbf{d}c^T \left(g \frac{\delta l}{\delta a} \right) = \text{Ad}_{g^{-1}}^* \left(\frac{\delta l}{\delta a} \diamond g^{-1} a_{ref} \right). \quad (3.36)$$

3.4.4 The fixed filament and its conservation law

The equations (1.32) for a fixed filament may also be obtained via affine Euler-Poincaré reduction. Theorem 3.2.1 suffices with the group $G = \mathcal{F}(\mathcal{D}, \mathcal{O})$ acting on the vector space $\Omega^1(\mathcal{D}, \mathfrak{o}) \times \mathcal{F}(\mathcal{D}, E) \ni (\Omega, \rho)$ by the affine action

$$(\Omega, \rho) \mapsto \theta_\Lambda(\Omega, \rho) := (\text{Ad}_\Lambda \Omega + \Lambda \mathbf{d}\Lambda^{-1}, \Lambda \rho).$$

Note that the cocycle is $c(\Lambda) = (\Lambda \mathbf{d}\Lambda^{-1}, 0)$. Using the expressions

$$(u, f) \diamond (\Omega, \rho) = \text{ad}_{\Omega_i}^* u^i + f \diamond \rho$$

$$\mathbf{d}c(\omega) = (-\mathbf{d}\omega, 0) \quad \text{and} \quad \mathbf{d}c^T(u, f) = \text{div}(u),$$

the affine Euler-Poincaré equations (3.13) become

$$(\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} = \frac{\delta l}{\delta \rho} \diamond \rho \quad (3.37)$$

Classical formulation of molecular strands

and the advection equations are

$$\begin{cases} \partial_t \Omega + \text{ad}_\omega \Omega = \mathbf{d}\omega \\ \partial_t \rho + \omega \rho = 0. \end{cases} \quad (3.38)$$

Recall from §1.3.2 that the Lagrangian for a fixed filament is of the form

$$l = l_{loc}(\omega, \Omega) + l_{np}(\xi, \rho),$$

$$l_{loc}(\omega, \Omega) = K(\omega) - \frac{1}{2} \int f(\Omega(s)) ds, \quad l_{np}(\xi, \rho) = - \iint U(\rho(s), \xi(s, s')) ds ds'$$

where

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad U : \mathbb{R}^3 \times SO(3) \rightarrow \mathbb{R}, \quad \xi(s, s') := \Lambda^{-1}(s) \Lambda(s').$$

The relations $\omega = \Lambda^{-1} \dot{\Lambda}$, $\Omega = \Lambda^{-1} \Lambda'$, and $\rho = \Lambda^{-1} \rho_{ref}$, where $\rho_{ref}(s) := (s, 0, 0)^T$, show that l is induced by an $SO(2)$ -invariant Lagrangian $L_{(0, \mathbf{r})} = L_{(0, \mathbf{r})}(\Lambda, \dot{\Lambda})$. Note that $SO(2)$ is precisely the isotropy group of $(0, \mathbf{r})$ relative to the affine action.

These observations enable the derivation of the equations for the fixed filament by the affine reduction processes described in Theorem 3.2.1. Using the general framework involving $\mathcal{D}$ and $\mathcal{O} \ltimes E$, one obtains the equations

$$(\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} = \frac{\delta l}{\delta \rho} \diamond \rho + \int \left[\xi(s, s') \frac{\partial U}{\partial \xi}(s', s) - \frac{\partial U}{\partial \xi}(s, s') \xi(s', s) \right] ds'$$

which coincides with (1.31) in the case of the fixed filament. Using total derivatives, these equations can be rewritten as (3.37).

The general formula (3.36) yields the conservation law

$$\frac{\partial}{\partial t} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \omega} \right] + \text{div} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \Omega} \right] = \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho \right).$$

From the general theory it follows that the solution of the advection equations (3.38) in terms of Λ are given by $\Omega = \Lambda^{-1} \mathbf{d}\Lambda$ and $\rho = \Lambda^{-1} \rho_{ref}$.

For the fixed filament, let $\mathcal{D} = [0, L]$, $E = \mathbb{R}^3$, the symmetry group $\mathcal{O} = SO(3)$, and the reference strand $\rho_{ref}(s) = \mathbf{r}(s) = (s, 0, 0)^T$. Then,

$$\frac{\partial}{\partial t} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \omega} \right] + \frac{\partial}{\partial s} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \Omega} \right] = \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \times \rho \right). \quad (3.39)$$

In this case, the torque does not vanish. The explanation is that the reference value ρ_{ref} of ρ is not zero, so (3.36) rather than (3.35) is required.

Also,

$$\text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \times \rho \right) = \Lambda \frac{\delta l}{\delta \rho} \times \Lambda \rho = \Lambda \frac{\delta l}{\delta \rho} \times \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix}.$$

More generally, the right hand side is

$$\left(\Lambda \frac{\delta l}{\delta \rho} \right) \times r,$$

where r describes the fixed filament.

The conservation law (3.39) is a particular case of the general formula (3.36), and did not appear in §1.3.2.

3.5 Summary

In this chapter we have provided a geometric interpretation of the molecular strand equations (2.27) and (2.28) as resulting from affine Euler-Poincaré reduction of a canonical Lagrangian system. A brief introduction to Lagrangian reduction from the classical perspective was given in §3.1, which was followed by some background material in §3.2. In particular, the general affine EP reduction theory was reviewed in §3.2.2.

The application to the molecular strand posed two issues. First, the treatment of nonlocal terms. Second, the extension of an unreduced Lagrangian that is required for the application of affine EP reduction. The former was dealt with in §3.3.1, the latter in §3.3.2. This culminated in the affine EP equations with advection relations (3.29) and (3.30).

Finally, §3.4 related the affine EP molecular strand equations with the Kirchhoff theory, the generalised strand, the fixed filament, and the conservation law formulation.

The final objective for Part I is to promote an alternative geometric interpretation of the molecular strand equations from the covariant perspective.

CHAPTER 4 : COVARIANT FORMULATION OF MOLECULAR STRANDS

As a counterpoint to the classical perspective, this chapter advances an alternative geometric description for the Hamilton-Pontryagin approach of Chapter 2. This alternative characterisation formulates the problem as a covariant field theory as was described in §1.2.3. The covariant approach contrasts with classical setup, which was introduced in §1.2.2 and developed in Chapter 3.

The discovery of the covariant formulation has its roots in a particular change of coordinates that drastically simplifies the form of the molecular strand equations. This coordinate transformation, and its generalisations, will be the subject of §4.1.

One particular feature of the covariant formulation is that the geometric description of the molecular strand problem changes radically depending on whether the Lagrangian (2.4) depends on the variable ρ . The presence of ρ in the Lagrangian breaks the $SE(3)$ symmetry to a $SO(3)$ symmetry, as is mentioned in Remark 1.3.3. This symmetry breaking may be handled by the Hamilton-Pontryagin approach of Chapter 2 by adding an extra variable, and by the affine Euler-Poincaré approach of Chapter 3 by extending the cocycle. The situation on the covariant side, however, is more involved.

The molecular strand problem with $SE(3)$ symmetry is an example of covariant Euler-Poincaré field equations, as developed in CASTRILLÓN-LÓPEZ ET AL. [2000]. On the other hand, the symmetry broken molecular strand is an example of covariant Lagrange-Poincaré field equations of CASTRILLÓN-LÓPEZ & RATIU [2003]. The geometric setup of the covariant Euler-Poincaré and Lagrange-Poincaré field equations differ considerably, leading to differences in the molecular strand problem viewed from the covariant perspective depending on the presence of ρ . The geometric description when ρ is absent will be given in §4.2, while the situation with ρ present will be described in §4.3.

4.1 Coordinate transformation

The coordinate transformation will be motivated in §4.1.1 by the identification of certain covariant differential operators. These operators naturally lead to a coordinate transformation that relates them with the corresponding partial differential operators. Subsequently, the coordinate transformation will be effected for the molecular strand in §4.1.2 leading to a dramatic simplification in the equations of motion (2.41). The transformation works in general, and the abstract formulation is the subject of §4.1.3.

4.1.1 Motivation

The molecular strand problem naturally defines its own differential operators. These operators arise by rearranging (1.1) to show that ρ , Γ , and γ satisfy the following relations

$$(\partial_s + \Omega \times) \rho = \Gamma, \quad (\partial_t + \omega \times) \rho = \gamma. \quad (4.1)$$

Therefore, the reduced variables (1.1) lead naturally to two covariant derivatives,

$$\frac{D}{Ds} = (\partial_s + \Omega \times), \quad \frac{D}{Dt} = (\partial_t + \omega \times), \quad (4.2)$$

one with respect to space and the other with respect to time.

With this interpretation Γ and γ are regarded as the tangent vectors above ρ that result from covariant differentiation,

$$\frac{D\rho}{Ds} = \Gamma, \quad \frac{D\rho}{Dt} = \gamma. \quad (4.3)$$

The operators from (4.2) also appear in the equations of motion for the molecular strand since the second equation of (2.41) may be written in the form

$$\frac{D}{Dt} \frac{\delta l}{\delta \gamma} + \frac{D}{Ds} \frac{\delta l}{\delta \Gamma} - \frac{\delta l}{\delta \rho} = 0. \quad (4.4)$$

When (4.3) and (4.4) are taken together, (4.4) is seen to be in the form of the Euler-Lagrange equations where the partial derivatives have been replaced by the

covariant derivatives (4.2). With this interpretation in mind it is natural to ask whether, by a coordinate transformation, it is possible to transform (4.4) into the canonical Euler-Lagrange form. The objective of §4.1 will be to provide such a change of variables. This line of enquiry leads to the covariant Euler-Poincaré and Lagrange-Poincaré points of view, which will be given in §4.2 and §4.3 respectively.

Rearranging equations (4.1) yields the time and space derivatives of ρ ,

$$\partial_s \rho = \Gamma - \Omega \times \rho, \quad \partial_t \rho = \gamma - \omega \times \rho.$$

Therefore, the natural candidate for the coordinate transformation is

$$\begin{aligned} \mathcal{F}(I, \mathfrak{so}(3)) \times \mathcal{F}(I, \mathbb{R}^3) \times \Omega^1(I, \mathfrak{so}(3)) \times \Omega^1(I, \mathbb{R}^3) \times \mathcal{F}(I, \mathbb{R}^3) \\ \ni (\omega, \gamma, \Omega, \Gamma, \rho) \mapsto (\rho, \rho_s, \rho_t, \omega, \Omega) \in \\ \mathcal{F}(I, \mathbb{R}^3) \times \Omega^1(I, \mathbb{R}^3) \times \mathcal{F}(I, \mathbb{R}^3) \times \mathcal{F}(I, \mathfrak{so}(3)) \times \Omega^1(I, \mathfrak{so}(3)), \end{aligned} \quad (4.5)$$

where two new variables have been introduced

$$\rho_s := \Gamma - \Omega \times \rho, \quad \rho_t := \gamma - \omega \times \rho. \quad (4.6)$$

This candidate coordinate transformation relates the differential operators (4.2) with their corresponding partial derivatives. The objective is to show that the equations of motion (2.41) have simple expressions if one uses this change of variables. Before ELLIS ET AL. [2010] this transformation had not been noticed before, in either the local or nonlocal elastic filament dynamics setting.

Notation 4.1.1 Below, the notation $\bar{l}$ is adopted to denote the Lagrangian l in terms of the new variables given by (4.5), that is,

$$l_{loc}(\omega, \gamma, \Omega, \Gamma, \rho) + l_{np}(\xi, \kappa, \Gamma) = \bar{l}_{loc}(\rho, \rho_s, \rho_t, \omega, \Omega) + \bar{l}_{np}(\xi, \kappa, \rho_s, \rho, \Omega).$$

4.1.2 Coordinate transformation

As in the Hamilton-Pontryagin approach, it is convenient to split the calculation of the change of variables into two parts, one for the local terms, another for the nonlocal terms. In other words, first the equations of motion for Lagrangian $\bar{l}_{loc}$

are calculated, which corresponds to fixing $\bar{l}_{np} = 0$. Subsequently, the equations for Lagrangian $\bar{l}_{np}$ are calculated, which is equivalent to setting $\bar{l}_{loc} = 0$. Finally, the full equations of motion are derived for the total Lagrangian $\bar{l} = \bar{l}_{loc} + \bar{l}_{np}$ by adding the local and nonlocal terms together. Here, linearity of the equations of motion with respect to the Lagrangian function is used. For the local terms, the coordinate transformation is described by the following theorem.

Theorem 4.1.2 *The local terms in (2.27) — (2.28) transform under the coordinate transformation (4.5) into the following equations involving $\bar{l}_{loc}$,*

$$(\partial_s + \boldsymbol{\Omega} \times) \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\Omega}} + (\partial_t + \boldsymbol{\omega} \times) \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\omega}} = 0, \quad (4.7)$$

$$\frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}} - \partial_t \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}_t} - \partial_s \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}_s} = 0. \quad (4.8)$$

Remark 4.1.3 Equations (4.7) and (4.8) have now formally decoupled, although the equations themselves must be solved simultaneously because the Lagrangian l depends on all the variables. Also note that equation (4.8) is equivalent, for local Lagrangians, to (2.28) with the covariant derivatives replaced by partial derivatives (but relative to the new variables).

One approach to proving Theorem 4.1.2 is to via the variational principle. From Notation 4.1.1 it is clear that

$$\begin{aligned} 0 &= \delta \int l_{loc}(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) dt \\ &= \int \left[\left\langle \frac{\delta l}{\delta \boldsymbol{\rho}}, \delta \boldsymbol{\rho} \right\rangle + \left\langle \frac{\delta l_{loc}}{\delta \boldsymbol{\gamma}}, \delta \boldsymbol{\gamma} \right\rangle + \left\langle \frac{\delta l_{loc}}{\delta \boldsymbol{\Gamma}}, \delta \boldsymbol{\Gamma} \right\rangle + \left\langle \frac{\delta l_{loc}}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle + \left\langle \frac{\delta l_{loc}}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle \right] dt \\ &= \delta \int \bar{l}_{loc}(\boldsymbol{\rho}, \boldsymbol{\rho}_s, \boldsymbol{\rho}_t, \boldsymbol{\omega}, \boldsymbol{\Omega}) dt \\ &= \int \left[\left\langle \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}}, \delta \boldsymbol{\rho} \right\rangle + \left\langle \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}_s}, \delta \boldsymbol{\rho}_s \right\rangle + \left\langle \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}_t}, \delta \boldsymbol{\rho}_t \right\rangle \right. \\ &\quad \left. + \left\langle \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle + \left\langle \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle \right] dt. \end{aligned} \quad (4.9)$$

Covariant formulation of molecular strands

Therefore, defining free variations $\Psi(s) = \Lambda(s)^{-1}\delta\mathbf{r}(s)$ and $\hat{\Sigma}(s) = \Lambda(s)^{-1}\delta\Lambda(s)$ and calculating results in the required equations. This calculation is shown in the proof below.

Proof. In order to expand the variational principle, the variations $\delta\boldsymbol{\rho}_t$ and $\delta\boldsymbol{\rho}_s$ must first be computed. Combining the variations of (4.6) with those from Lemma 2.2.2 results in the following term in (4.9) proportional to $\delta\bar{l}_{loc}/\delta\boldsymbol{\rho}_t$,

$$\begin{aligned} \left\langle \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \delta\boldsymbol{\rho}_t \right\rangle &= \left\langle \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \dot{\boldsymbol{\Psi}} + \boldsymbol{\omega} \times \boldsymbol{\Psi} - \boldsymbol{\Sigma} \times \boldsymbol{\gamma} \right. \\ &\quad \left. - \left(\dot{\boldsymbol{\Sigma}} + \boldsymbol{\omega} \times \boldsymbol{\Sigma} \right) \times \boldsymbol{\rho} - \boldsymbol{\omega} \times \left(-\boldsymbol{\Sigma} \times \boldsymbol{\rho} + \boldsymbol{\Psi} \right) \right\rangle \\ &= \left\langle -\frac{\partial}{\partial t} \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \boldsymbol{\Psi} \right\rangle \\ &\quad + \left\langle \frac{\partial}{\partial t} \left(\boldsymbol{\rho} \times \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t} \right) - \boldsymbol{\gamma} \times \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t} - \left(\boldsymbol{\rho} \times \boldsymbol{\omega} \right) \times \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \boldsymbol{\Sigma} \right\rangle, \quad (4.10) \end{aligned}$$

where the Jacobi identity has been used to simplify two triple cross products. Employing the kinematic condition for the derivative of $\boldsymbol{\rho}$, $\partial_t\boldsymbol{\rho} = \boldsymbol{\gamma} - \boldsymbol{\omega} \times \boldsymbol{\rho}$, to simplify the $\boldsymbol{\Sigma}$ term in (4.10) yields

$$\left\langle \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \delta\boldsymbol{\rho}_t \right\rangle = \left\langle -\frac{\partial}{\partial t} \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \boldsymbol{\Psi} \right\rangle + \left\langle -\boldsymbol{\rho} \times \frac{\partial}{\partial t} \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_t}, \boldsymbol{\Sigma} \right\rangle. \quad (4.11)$$

Analogously,

$$\left\langle \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_s}, \delta\boldsymbol{\rho}_s \right\rangle = \left\langle -\frac{\partial}{\partial s} \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_s}, \boldsymbol{\Psi} \right\rangle + \left\langle -\boldsymbol{\rho} \times \frac{\partial}{\partial s} \frac{\delta\bar{l}_{loc}}{\delta\boldsymbol{\rho}_s}, \boldsymbol{\Sigma} \right\rangle. \quad (4.12)$$

Completing the variational principle (4.9), the only terms containing $\boldsymbol{\Psi}$ are the derivatives with respect to $\boldsymbol{\rho}$, $\boldsymbol{\rho}_s$ and $\boldsymbol{\rho}_t$ since, from Lemma 2.2.2, the variations in $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}$ only involve $\boldsymbol{\Sigma}$. This results in (4.8).

Gathering the terms proportional to $\boldsymbol{\Sigma}$, another cancellation appears. All the terms involving cross products with respect to $\boldsymbol{\rho}$ cancel, as they are each be multiplied by the left hand side of (4.8) which vanishes. Thus, derivatives with respect to $\boldsymbol{\rho}$, $\boldsymbol{\rho}_s$, and $\boldsymbol{\rho}_t$ do not contribute to the terms proportional to $\boldsymbol{\Sigma}$. Collecting those terms yields exactly (4.7). ■

Another approach to performing the change of variables that highlights the decoupling arises from the expression for $\delta\rho$ in terms of the free variations Ψ and Σ ,

$$\delta\rho = \Psi - \Sigma \times \rho. \quad (4.13)$$

Equation (4.13) may be interpreted as providing a choice of any two of Σ , Ψ , and $\delta\rho$ as a free variation, with the third variation being determined by the relationship. In practice, $\delta\Omega$ and $\delta\omega$ only depend on Σ , which must therefore be selected. This leaves a choice of Ψ or $\delta\rho$ for the second choice of free variation. The relations $\rho_s = \partial_s \rho$ and $\rho_t = \partial_t \rho$ prompt the choice of $\delta\rho$, and following expression for $\delta\rho_s$ in terms of $\delta\rho$ results

$$\delta\rho_s = \delta\partial_s \rho = \partial_s \delta\rho.$$

Similarly, $\delta\rho_t = \partial_t \delta\rho$. Together, referring to Lemma 2.2.2, the variations are

$$\delta\omega = \dot{\Sigma} + \omega \times \Sigma, \quad \delta\Omega = \Sigma' + \Omega \times \Sigma \quad (4.14)$$

$$\delta\rho_s = \partial_s \delta\rho, \quad \delta\rho_t = \partial_t \delta\rho, \quad (4.15)$$

which is obviously augmented by the trivial relation $\delta\rho = \delta\rho$. An alternative proof of Theorem 4.1.2 goes as follows.

Proof.

The terms arising from $\delta\Omega$ and $\delta\omega$ are identical to before and only depend on Σ . Therefore, the Σ variation results in,

$$(\partial_t + \omega \times) \frac{\delta\bar{l}_{loc}}{\delta\omega} + (\partial_s + \Omega \times) \frac{\delta\bar{l}_{loc}}{\delta\Omega} = 0.$$

The variations (4.14), (4.15) yield, for example,

$$\left\langle \frac{\delta\bar{l}_{loc}}{\delta\rho_s}, \delta\rho_s \right\rangle = \left\langle \frac{\delta\bar{l}_{loc}}{\delta\rho_s}, \partial_s \delta\rho \right\rangle = - \left\langle \partial_s \frac{\delta\bar{l}_{loc}}{\delta\rho_s}, \delta\rho \right\rangle.$$

The second equation comes from terms proportional to $\delta\rho$ which is

$$\partial_t \frac{\delta\bar{l}_{loc}}{\delta\rho_t} + \partial_s \frac{\delta\bar{l}_{loc}}{\delta\rho_s} - \frac{\delta\bar{l}_{loc}}{\delta\rho} = 0.$$

These are the required equations in Theorem 4.1.2. ■

Remark 4.1.4 It is tempting to refer to Euler-Poincaré reduction with advected quantities, which is described in Example 3.1.2, at this point and ask whether a similar change of variables might simplify the its equations of motion. The answer, alas, is negative but is nevertheless interesting. The crucial property was to regard $\delta\rho$ as a free variation. Given an advected quantity $\mathbf{a} = \Lambda^{-1}\mathbf{a}_{ref}$, could one consider $\delta\mathbf{a}$ as a free variation? Unfortunately, the variation $\delta\mathbf{a}$ is given by

$$\delta\mathbf{a} = -\Sigma\mathbf{a}.$$

Therefore $\delta\mathbf{a}$ is determined by Σ and no interpretation of $\delta\mathbf{a}$ as a free variation can be made.

The change of variables is available for problems posed on semidirect products when the Euler-Lagrange equations hold on the whole of $\mathcal{O} \otimes V^*$, and reduction is performed by an $\mathcal{O}$ symmetry. This contrasts with the situation where the Lagrangian is defined on $\mathcal{O}$ with $a_{ref} \in V^*$ held fixed, and the symmetry is the isotropy group $\mathcal{O}_{a_{ref}}$. This alternative situation, which includes the molecular strand, is called *covariant Lagrange-Poincaré reduction by a subgroup* and, having been developed in CASTRILLÓN-LÓPEZ & RATIU [2003], provides a particular example throughout Part II. Discussion on the underlying geometric structure to the change of variables may be found in §4.3.

Having addressed the local part of the dynamics, the next task is to deal with the nonlocal terms. The change of variables for the nonlocal dynamics is outlined by the following theorem.

Theorem 4.1.5 *The nonlocal terms in (2.27) — (2.28) transform under (4.5) into the following equations involving nonlocal Lagrangian $\bar{l}_{np}$,*

$$(\partial_s + \Omega \times) \frac{\delta \bar{l}_{np}}{\delta \Omega} = \int \left(\mathbf{Z}(s, s') + \frac{\delta \bar{U}}{\delta \boldsymbol{\kappa}}(s, s') \times \boldsymbol{\kappa}(s, s') \right) ds', \quad (4.16)$$

$$\frac{\delta \bar{l}_{np}}{\delta \boldsymbol{\rho}} - \partial_s \frac{\delta \bar{l}_{np}}{\delta \boldsymbol{\rho}_s} = \int \left(\xi(s, s') \frac{\delta \bar{U}}{\delta \boldsymbol{\kappa}}(s', s) - \frac{\delta \bar{U}}{\delta \boldsymbol{\kappa}}(s, s') \right) ds' \quad (4.17)$$

where, as in (2.26),

$$\mathbf{Z}(s, s') = \xi^{-1}(s, s') \left(\frac{\delta \bar{U}}{\delta \xi} \right)^T (s, s') - \frac{\delta \bar{U}}{\delta \xi}(s, s') \xi(s, s').$$

Proof. The free variations are $\widehat{\Sigma}(s) = \Lambda^{-1}(s)\delta\Lambda(s)$ and $\delta\boldsymbol{\rho}(s) = \delta(\Lambda^{-1}(s)\mathbf{r}(s)) = \Psi(s) - \Sigma(s) \times \boldsymbol{\rho}$, which may be used on the variational principle for the nonlocal Lagrangian

$$\bar{l}_{np}(\boldsymbol{\rho}, \boldsymbol{\rho}_s, \boldsymbol{\Omega}, \xi, \boldsymbol{\kappa}) = \iint \bar{U}(\xi, \boldsymbol{\kappa}, \boldsymbol{\rho}, \boldsymbol{\rho}_s, \boldsymbol{\Omega}) ds ds'.$$

The variations of ξ and $\boldsymbol{\kappa}$ from (2.7) are

$$\begin{cases} \xi^{-1}(s, s')\delta\xi(s, s') &= \widehat{\Sigma}(s') - \text{Ad}_{\xi^{-1}(s, s')} \widehat{\Sigma}(s) \\ \delta\boldsymbol{\kappa}(s, s') &= (\Sigma(s') - \xi^{-1}(s, s')\Sigma(s)) \times \boldsymbol{\rho}(s') \\ &\quad + \xi(s, s')\delta\boldsymbol{\rho}(s') - \delta\boldsymbol{\rho}(s). \end{cases} \quad (4.18)$$

Using the variations (4.18) in combination with (4.14) — (4.15), the terms in the variational principle proportional to $\Sigma(s)$ yield (4.16). Meanwhile, the terms in the variational principle proportional to $\delta\boldsymbol{\rho}(s)$ give (4.17). ■

Combining Theorems 4.1.2 and 4.1.5 yields the complete equations of motion in $\boldsymbol{\rho}_s$ and $\boldsymbol{\rho}_t$ variables, which read

$$\begin{aligned} (\partial_t + \boldsymbol{\omega} \times) \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\omega}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta (\bar{l}_{loc} + \bar{l}_{np})}{\delta \boldsymbol{\Omega}} = \\ \int \left(\mathbf{Z}(s, s') + \frac{\delta \bar{U}}{\delta \boldsymbol{\kappa}}(s, s') \times \boldsymbol{\kappa}(s, s') \right) ds', \end{aligned} \quad (4.19)$$

$$\begin{aligned} \partial_t \frac{\delta \bar{l}_{loc}}{\delta \boldsymbol{\rho}_t} + \partial_s \frac{\delta (\bar{l}_{loc} + \bar{l}_{np})}{\delta \boldsymbol{\rho}_s} - \frac{\delta (\bar{l}_{loc} + \bar{l}_{np})}{\delta \boldsymbol{\rho}} = \\ \int \left(\xi(s, s') \frac{\delta \bar{U}}{\delta \boldsymbol{\kappa}}(s', s) - \frac{\delta \bar{U}}{\delta \boldsymbol{\kappa}}(s, s') \right) ds'. \end{aligned} \quad (4.20)$$

Equations (4.19) — (4.20) are the transformed versions of (2.27) — (2.28) respectively. Note that (4.19) are Euler-Poincaré equations HOLM [2008A], while (4.20) are Euler-Lagrange equations, both modified with nonlocal interactions. The ability to identify the familiar Euler-Lagrange and Euler-Poincaré operators in (4.19) — (4.20) highlights the geometric structure of the equations of motion (2.27) — (2.28).

Using total derivatives in ρ , ρ_s and Ω , analogously to equation 3.24 derived in §??, enables (4.19) — (4.20) to be written in the form

$$\begin{cases} (\partial_t + \omega \times) \frac{\delta l}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta l}{\delta \Omega} \Big|_{\text{Tot}} = 0 \\ \partial_t \frac{\delta l}{\delta \rho_t} + \partial_s \frac{\delta l}{\delta \rho_s} \Big|_{\text{Tot}} - \frac{\delta l}{\delta \rho} \Big|_{\text{Tot}} = 0 \end{cases} \quad (4.21)$$

Equations (4.21) are the transformed versions of (2.41), and shall be identified in §4.3 as a form of covariant Lagrange-Poncaré equations CASTRILLÓN-LÓPEZ & RATIU [2003].

It is also interesting to note that the coordinate transformation is not available in the classical Kirchhoff theory, described in §1.3.1, unless ρ is present. §4.2 will show that this situation may be understood as covariant Euler-Poincaré equations CASTRILLÓN-LÓPEZ ET AL. [2000]. Breaking the $SE(3)$ symmetry of the Kirchhoff theory, for example by adding a ρ dependent external potential, enables the coordinate change and the theory becomes a Lagrange-Poincaré reduction.

4.1.3 The general case

Having effected the coordinate transformation (4.5) for the molecular strand, this paragraph considers the new form of the equations in the generalised setting of §2.1.3.

Recall from §2.1.3 that the generalisation of the molecular strand considers $(\Lambda, r) \in \mathcal{F}(\mathcal{D}, S)$, $(\Omega, \Gamma) \in \Omega^1(\mathcal{D}, \mathfrak{s})$, and $\rho \in \mathcal{F}(\mathcal{D}, E)$, where $S = \mathcal{O} \ltimes E$ is the semidirect product of a Lie group $\mathcal{O}$ with a vector space E .

Differentiating the definition of ρ in (2.18) gives

$$\mathbf{d}\rho = \mathbf{d}(\Lambda^{-1}r) = -\Lambda^{-1}\mathbf{d}\Lambda\Lambda^{-1}r + \Lambda^{-1}\mathbf{d}r = \Gamma - \Omega\rho.$$

This motivates the introduction of the new variables $\rho_s \in \Omega^1(\mathcal{D}, E)$ and $\rho_t \in \mathcal{F}(\mathcal{D}, E)$, playing the respective roles of space and time derivatives of ρ ,

$$\rho_s := \Gamma - \Omega\rho \quad \rho_t := \gamma - \omega\rho. \quad (4.22)$$

The relations (4.22) define a diffeomorphism from $(\omega, \gamma, \Omega, \Gamma, \rho)$ to the new variables $(\omega, \Omega, \rho_s, \rho_t, \rho)$, which generalises (4.6). In terms of the new variables, the local Lagrangian is denoted by $\bar{l}$ and is characterised by the relation

$$\bar{l}(\rho, \rho_s, \rho_t, \omega, \Omega) = l(\omega, \gamma, \Omega, \Gamma, \rho).$$

For simplicity, only the case of a local Lagrangian has been treated. There are two equivalent methods to obtain the equations of motion in terms of $\bar{l}$.

The first method is via a variational principle, as in the particular case of the charged strand in §4.1.2. This method involves calculating the constrained variations of $(\omega, \gamma, \Omega, \Gamma, \rho)$, defined in (2.18), and applying them to the variational principle for the general Lagrangian.

The second method involves computing the functional derivatives of l in terms of those of $\bar{l}$, which results in

$$\begin{aligned} \frac{\delta l}{\delta \omega} &= \frac{\delta \bar{l}}{\delta \omega} - \rho \diamond \frac{\delta \bar{l}}{\delta \rho_t}, & \frac{\delta l}{\delta \Omega} &= \frac{\delta \bar{l}}{\delta \Omega} - \rho \diamond \frac{\delta \bar{l}}{\delta \rho_s}, \\ \frac{\delta l}{\delta \gamma} &= \frac{\delta \bar{l}}{\delta \rho_t}, & \frac{\delta l}{\delta \Gamma} &= \frac{\delta \bar{l}}{\delta \rho_s}, \end{aligned}$$

and

$$\frac{\delta l}{\delta \rho} = \frac{\delta \bar{l}}{\delta \rho} + \Omega_i \frac{\delta \bar{l}}{\delta \rho_{s_i}} + \omega \frac{\delta \bar{l}}{\delta \rho_t}.$$

These relations may be substituted into the general equations (3.32), with the non-local terms removed for simplicity.

Both these methods lead to the same transformed generalised equations,

$$\begin{cases} \left(\frac{d}{dt} - \text{ad}_\omega^* \right) \frac{\delta \bar{l}}{\delta \omega} + \text{div}^\Omega \frac{\delta \bar{l}}{\delta \Omega} = 0, \\ \frac{d}{dt} \frac{\delta \bar{l}}{\delta \rho_t} + \text{div} \frac{\delta \bar{l}}{\delta \rho_s} - \frac{\delta \bar{l}}{\delta \rho} = 0, \end{cases} \quad (4.23)$$

Covariant formulation of molecular strands

where $\text{div}^\Omega : \mathfrak{X}(\mathcal{D}, \mathfrak{o}^*) \rightarrow \mathcal{F}(\mathcal{D}, \mathfrak{o}^*)$ is defined by $\text{div}^\Omega w := \text{div } w - \text{ad}_{\Omega_i}^* w^i \in \mathcal{F}(\mathcal{D}, \mathfrak{o}^*)$.

Equations (4.23) coincide with (4.7) and (4.8) in the particular case $\mathcal{D} = [0, L]$ and $S = SE(3)$. The auxiliary equations are computed to be

$$\begin{cases} \dot{\rho}_s + \omega \rho_s = \mathbf{d}\rho_t + \Omega \rho_t, \\ \dot{\Omega} + \text{ad}_\omega \Omega = \mathbf{d}\omega, \\ \dot{\rho} = \rho_t. \end{cases} \quad (4.24)$$

Here, the second equation of 4.24 will be interpreted as a zero-curvature (or integrability) relation in §7.1.3. Such relations appear in the covariant perspective as extra conditions that ensure the solution of the unreduced equations may be reconstructed from that of the reduced equations. The first equation is easily verified by an equality of cross derivatives argument using the relation $\mathbf{d}\rho = \rho_s$ and the third equation. Therefore, (4.24) may be replaced with

$$\begin{cases} \mathbf{d}\rho = \rho_s, \\ \dot{\Omega} + \text{ad}_\omega \Omega = \mathbf{d}\omega, \\ \dot{\rho} = \rho_t. \end{cases} \quad (4.25)$$

4.1.4 Summary

In this section we have seen that a coordinate transformation of the dynamical variables for the molecular strand problem results in equations of motion (4.21) that exhibits a number of interesting properties. Firstly, a great deal of cancelation occurs, and the form of the equations reduce to one Euler-Lagrange equation and one Euler-Poincaré equation, both with nonlocal forcing. Secondly, the space variable s and time variable t encounter an exchange symmetry, that is, time and space become formally interchangeable. Finally, the equations of motion formally decouple among the filament variables (ρ, ρ_s, ρ_t) and the RCC variables (Ω, ω) .

Here, the change of variables was seen to apply both local and nonlocal terms via the direct Hamilton-Pontryagin, or ‘bare hands’ method used in Chapter 2. In §4.1.3, the coordinate transformation was cast in the general

setting for local Lagrangians.

Having identified the coordinate transformation using direct calculations, the challenge that remains for this chapter is to provide an underlying theoretical understanding that accounts for the remarkable formal decomposition of the equations and exchange symmetry. Indeed, the exchange symmetry introduces a formal indistinguishability between time and space variables, which is indicative of a field theory. On the other hand, the formal decomposition of the equations of motion is reminiscent of a splitting of the configuration space into two bundles. Both of these insights will be justified mathematically in §4.3 by applying the theory developed in CASTRILLÓN-LÓPEZ & RATIU [2003]. In the meantime, §4.2 simplifies the situation slightly by removing dependence on ρ , in order to develop the full theory and introduce the relevant concepts in a more manageable fashion.

4.2 Covariant Euler-Poincaré formulation

In order to develop the covariant theory of molecular strands, it is convenient to begin by simplifying the application to the particular case where the Lagrangian does not depend on the ρ variable. This particular case includes the Kirchhoff model and the fixed filament, described in §1.3.1 and §1.3.2 respectively.

The relevant theoretical machinery for an abstract formulation of this particular case is provided by CASTRILLÓN-LÓPEZ ET AL. [2000], which is the first in a series of papers on covariant Lagrangian reduction, and deals with the extension of classical Euler-Poincaré reduction of variational principles to the field theoretic context.

The objective here is to understand certain cases of molecular strand dynamics as applications of CASTRILLÓN-LÓPEZ ET AL. [2000]. In the process, various geometric concepts are introduced that provide an applied background for the more technical developments encountered in Part II.

In §4.2.1 the background material for the covariant perspective will be introduced. This material main reduction theorem is called *covariant Euler-Poincaré reduction*, and was first developed in CASTRILLÓN-LÓPEZ ET AL. [2000]. Other material and notation has come from ELLIS ET AL. [2009]. Following the presen-

tation of the background material, the application to the strand itself will come in §4.2.2, where the equations of motion are derived and related with the molecular strand problem. This achievement provides a theoretical understanding of particular cases of the molecular strand equations as arising from covariant Euler-Poincaré reduction of a field theory.

4.2.1 Background material

Various geometric concepts are required for a full appreciation of the covariant perspective. These are introduced briefly, and sometimes in a simplified form. Detailed definitions in a more technical setting may be found in §6.2.

As §1.2.3 pointed out, the covariant perspective considers a space-time $X := [0, L] \times \mathbb{R}$. Next a principal G -bundle over X is introduced, $P := X \times SE(3)$ with projection $\pi_{X,P} : P \rightarrow X$. A **section** of π is defined as a smooth map $\sigma : X \rightarrow P$, such that

$$\pi \circ \sigma = \text{id}_X, \quad (4.26)$$

where id_X is the identity map on X . The space of sections of π is denoted $\Gamma\pi$.

The role of the unreduced configuration space for field theories is played by the **first jet bundle** J^1P , which constitutes an affine bundle over P that may be defined fibre-wise by

$$J_p^1P = \{ \gamma_p \in L(T_xX, T_pP) \mid T_p\pi \circ \gamma_p = \text{id}_{T_xX} \},$$

with projection $\pi_{P,J^1P} : J^1P \rightarrow P$ given by $\pi_{P,J^1P}(\gamma_p) = p$, as in ELLIS ET AL. [2009]. The first jet bundle may be thought of as the field theoretic version of the tangent bundle in the classical perspective.

Despite the definition of J^1P as an affine bundle over P , the more important conception for field theory is J^1P as a fibre bundle over X with projection

$$\pi_{X,J^1P} := \pi \circ \pi_{P,J^1P} : J^1P \rightarrow X. \quad (4.27)$$

Indeed, the tangent lift of a section $\sigma \in \Gamma\pi$, interpreted as a map $x \mapsto T_x\sigma$, constitutes a section of π_{X,J^1P} since for $T_x\sigma \in L(T_xX, T_{\sigma(x)}P)$ equation (4.26) yields

$$T_{\sigma(x)}\pi \circ T_x\sigma = T_x(\pi \circ \sigma) = T_x\text{id}_X = \text{id}_{T_xX}.$$

The geometry introduced here so far is succinctly visualised and organised by commutative diagrams. The following commutative diagram exhibits the geometry of the jet bundle:

$$\begin{array}{ccc}
 J^1P & \xrightarrow{\pi_{X,J^1P}} & X \\
 & \searrow \pi_{P,J^1P} & \nearrow \iota \\
 & P &
 \end{array}$$

Arrows between spaces indicate maps from the space at the tail of the arrow to the space at its head. Sometimes arrows are adorned with the name of the maps they represent. Different paths through the diagram are equivalent in terms of composition of the associated maps; therefore, this diagram also communicates (4.27).

The geometry of the reduced jet bundle, J^1P/G , may be understood via the use of a connection form $\mathcal{A} : TP \rightarrow \mathfrak{g}$ ¹. Any connection form $\mathcal{A}$ introduces a vector bundle isomorphism

$$J^1P/G \rightarrow \Lambda^1(X; \text{Ad}P), \quad (4.28)$$

where, for any vector bundle E over X , $\Lambda^1(X; E)$ denotes the bundle over X defined fibre-wise by

$$\Lambda_x^i(X; E) = L_a^i(T_x X, E_x),$$

where L_a^i denotes the i -linear, skew-symmetric maps. Also, $\text{Ad}P$ denotes the *ad-joint bundle* associated to the principal bundle P defined as the quotient space

$$\text{Ad}P := (P \times \mathfrak{g}) / G,$$

relative to the diagonal action of $g \in G$

$$(p, \xi) \in P \times \mathfrak{g} \mapsto (gp, \text{Ad}_g \xi) \in P \times \mathfrak{g}.$$

¹A $\mathfrak{g}$ -valued one form on P that satisfies $\mathcal{A}(gv_p) = \text{Ad}_g \mathcal{A}(v_p)$ and $\mathcal{A}(\xi_P(p)) = \xi$ for all $v_p \in TP$, $g \in G$, and $\xi \in \mathfrak{g}$.

Denoting the equivalence class of $(p, \xi) \in P \times \mathfrak{g}$ by

$$[p, \xi]_{\mathfrak{g}} \in \text{Ad}P,$$

the bundle isomorphism $J^1P/G \rightarrow \Lambda^1(X; \text{Ad}P)$ reads

$$[T\sigma] \mapsto \bar{\sigma} := [\sigma, \sigma^* \mathcal{A}]_{\mathfrak{g}}. \quad (4.29)$$

Therefore, $J^1P/G \cong \Lambda^1(X; \text{Ad}P)$. At the level of the sections, $\Gamma\pi_{X, J^1P/G} = \Omega^1(X; \text{Ad}P)$, that is the G -invariant sections of π_{X, J^1P} are $\text{Ad}P$ -valued one forms on X .

The covariant Euler-Poincaré reduction theorem was first given as Theorem 3.1 in CASTRILLÓN-LÓPEZ ET AL. [2000]. In Part II a broader theory of Lagrange-Poincaré reduction is developed. Consequently, the same result will be given as particular case **ii** following Theorem 7.1.8 in §7.1.2. In order to avoid repetition, details of the reduction theorem are left to Part II where a full treatment is undertaken.

4.2.2 Application to the molecular strand

For the molecular strand the space-time is $x := (s, t) \in X := [0, L] \times \mathbb{R}$. When the ρ variable is absent, that is, when the local Lagrangian is independent of ρ indicating the absence of an external central potential, the symmetry group is $SE(3)$. The principal bundle is the trivial $SE(3)$ -bundle, $P := X \times SE(3)$. Therefore the structure group of P is also the symmetry group, as required.

Recall that a principal bundle admits a global section if and only if it is trivial. Therefore, a section $\sigma \in \Gamma\pi_{X, P}$ may be identified with a map $g : X \rightarrow SE(3)$ such that $\sigma(x) = (x, g(x)) = ((s, t), (\Lambda, \mathbf{r})(s, t))$. That is, a section $\sigma \in \Gamma\pi$ is equivalent to an integral curve of unreduced Euler-Lagrange equations.

The first jet extension of σ may be expressed as

$$j^1\sigma := (\text{id}_{TX}, Tg) = \left(\text{id}_{TX}, \Lambda, \mathbf{r}, \Lambda' ds + \dot{\Lambda} dt, \mathbf{r}' ds + \dot{\mathbf{r}} dt \right) \in \Gamma\pi_{X, J^1P}.$$

The connection may be fixed as $\mathcal{A}(v_x, v_g) = v_g g^{-1}$. Therefore, the bundle iso-

morphism (4.29) may be used to express the reduced section as

$$[j^1\sigma] \cong [\sigma, dgg^{-1}]_{\mathfrak{se}(3)} =: \bar{\sigma}.$$

Since the bundle $P \rightarrow X$ is trivial, there is a bundle isomorphism $\text{Ad } P \rightarrow X \times \mathfrak{g}$ given by

$$[(x, h), \xi]_{\mathfrak{se}(3)} \mapsto (x, \text{Ad}_{h^{-1}}\xi), \text{ for all } h \in G, \xi \in \mathfrak{g}, x \in X. \quad (4.30)$$

Therefore the reduced section $\bar{\sigma} \in \Omega^1(X; \text{Ad } P) \cong \Omega^1(X; \mathfrak{g})$ may be identified with a $\mathfrak{g}$ -valued one form on $\mathfrak{g}$, which is given by

$$[j^1\sigma] \cong g^{-1}dg = (\Omega ds + \omega dt, \Gamma ds + \gamma dt). \quad (4.31)$$

Note that (4.31) recovers all of the reduced coordinates from (1.1) except the ρ variable. Therefore a reduced Lagrangian is given by $l([j^1\sigma]) = l(\omega, \Omega, \gamma, \Gamma)$.

Under the bundle isomorphism (4.30), the differential operator ∇^A transforms as

$$\nabla^A [\sigma, \xi]_{\mathfrak{se}(3)} = [\sigma, d\xi - [\sigma^* \mathcal{A}, \xi]]_{\mathfrak{se}(3)} \mapsto (x, \text{Ad}_{g^{-1}}d\xi - [g^{-1}dg, \text{Ad}_{g^{-1}}\xi]) = (x, d(\text{Ad}_{g^{-1}}\xi)).$$

Let $\mu, v : X \rightarrow \mathfrak{se}^*(3)$, so that $(x, \mu\partial_t + v\partial_s) \in \mathfrak{X}(X) \otimes \mathfrak{se}^*(3) = \Omega^1(X; \mathfrak{se}^*(3))^*$ is an element in the dual space to $J^1P/G \cong \Omega^1(X; \mathfrak{se}(3))$. The operator div^A , which is the adjoint of ∇^A , acts on $\mathfrak{X}(X) \otimes \mathfrak{se}^*(3)$ according to

$$\text{div}^A(x, \mu\partial_s + v\partial_t) = (x, \partial_s\mu + \partial_tv). \quad (4.32)$$

Combining (4.32) with the general expression for ad^* operator on a semidirect product (3.3), the covariant Euler-Poincaré equations from Theorem 7.1.8 may be written as

$$\begin{cases} (\partial_s - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} + (\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + \Gamma \diamond \frac{\delta l}{\delta \Gamma} + \gamma \diamond \frac{\delta l}{\delta \gamma} = 0 \\ (\partial_t + \omega) \frac{\delta l}{\delta \gamma} + (\partial_s + \Omega) \frac{\delta l}{\delta \Gamma} = 0. \end{cases} \quad (4.33)$$

Employing the hat map from Definition 1.2.4, equations (4.33) may be written in the form

$$\begin{cases} (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta l}{\delta \boldsymbol{\Omega}} + (\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\omega}} + \boldsymbol{\Gamma} \times \frac{\delta l}{\delta \boldsymbol{\Gamma}} + \boldsymbol{\gamma} \times \frac{\delta l}{\delta \boldsymbol{\gamma}} = 0 \\ (\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\gamma}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta l}{\delta \boldsymbol{\Gamma}} = 0. \end{cases} \quad (4.34)$$

Consequently, particular case **ii** of Theorem 7.1.8 in Part II says that (4.33) come from Euler-Poincaré reduction of a covariant field theory on J^1P governed by the Euler-Lagrange equations. Note that (4.33) exactly correspond with (2.41) with the terms involving $\boldsymbol{\rho}$ dropped. Therefore, in the absence of $\boldsymbol{\rho}$ dependence in the local Lagrangian, (2.41) are a form of Euler-Poincaré reduced field equations. This case includes the fixed filament dynamics of §1.3.2 and the Kirchhoff model of §1.3.1.

4.3 Covariant Lagrange-Poincaré formulation

As discussed in §4.2, the molecular strand without $\boldsymbol{\rho}$ dependency is an example of Euler-Poincaré reduction, as laid out in CASTRILLÓN-LÓPEZ ET AL. [2000]. When the $\boldsymbol{\rho}$ dependency is present in the local Lagrangian, then an alternative theory must be employed to understand the dynamics at the abstract geometric level. This alternative theory is called *subgroup Lagrange-Poincaré reduction*, and was first derived in CASTRILLÓN-LÓPEZ & RATIU [2003].

In contrast with the classical perspective, where adding a term to the cocycle provides the necessary augmentation, the presence of $\boldsymbol{\rho}$ in the covariant perspective calls for an total review of the geometric framework of the theory. The extended geometric framework, in turn, provides a geometric understanding of the coordinate transformation set out in §4.1. The discussion in this section describes the full molecular strand dynamics, including the $\boldsymbol{\rho}$ variable, from the covariant perspective. The presentation given here is adapted from ELLIS ET AL. [2009] in the interests of brevity, and to avoid repetition in Part II. A more in depth treatment of the material can be found in ELLIS ET AL. [2010].

The theory is laid out similarly to that in §4.2, beginning with more background material in §4.3.1, and continuing with the application to the molecular strand in

§4.3.2

4.3.1 Background material

In order to avoid repetition, the background material presented here follows on from §4.2.1. Again, various geometric concepts are introduced in form adapted to the application, for a more technical definition of the concepts involved, see §6.2.

It is convenient for the present purposes to make certain assumptions in order to simplify the theory. These assumptions are the following:

1. The bundle $\pi_{X,P} : P \rightarrow X$ is trivial, i.e. $P = X \times G$.
2. The connection form $\mathcal{A}$ is flat. That is the curvature of $\mathcal{A}$, as defined in Proposition 7.1.3, vanishes.

Departing from §4.2, P is also regarded as a principal H -bundle where $H \subset G$ is a subgroup of the structure group. The quotient space $\Sigma := P/H = X \times G/H$ is referred to as *shape space*, and there is a projection $\pi_{\Sigma,P} : P \rightarrow \Sigma$.

Now there are two principal bundle structures on P given by $\pi_{X,P}$ and $\pi_{\Sigma,P}$. These induce a fibre bundle structure on Σ given by

$$\pi_{X,\Sigma} : \Sigma \rightarrow X, \quad \pi_{X,\Sigma}(x, \rho) = x.$$

The geometry is described by the following commutative diagram.

$$\begin{array}{ccc} P & \xrightarrow{\pi_{X,P}} & X \\ & \searrow \pi_{\Sigma,P} & \nearrow \pi_{X,\Sigma} \\ & \Sigma & \end{array}$$

The bundle isomorphism (4.29) gains an additional component that accounts for the shape space. Given the connection $\mathcal{A} : TP \rightarrow \mathfrak{h}$, the bundle isomorphism is now

$$\alpha_{\mathcal{A}} : J^1 P/H \rightarrow J^1 \Sigma \oplus_{\Sigma} L(\pi_{X,\Sigma}^* TX, \text{Ad} P)$$

given by

$$\alpha_{\mathcal{A}}([T\sigma]_H) \mapsto \left(T(\pi_{\Sigma,P} \circ \sigma), [\sigma, \sigma^* \mathcal{A}]_{\mathfrak{h}} \right). \quad (4.35)$$

More details and proof of (4.35) will be provided in §6.2.2.

The subgroup Lagrange-Poincaré reduction theorem for field theories was first given in CASTRILLÓN-LÓPEZ & RATIU [2003]. Similarly to the Euler-Poincaré case, the theorem is a particular case of the more general Lagrange-Poincaré framework developed in Part II, where the subgroup reduction is particular case **iii** after Theorem 7.1.8.

4.3.2 Application to the molecular strand

Recall from §4.2 that in the absence of ρ dependency, the molecular strand may be understood as reduction of the trivial bundle $P := X \times SE(3)$ by the structure group, $G := SE(3)$. Since $(\Lambda, \mathbf{r}) \mapsto \Lambda^{-1}\mathbf{r}$ is only $SO(3)$ invariant, the presence of $\rho := \Lambda^{-1}\mathbf{r}$ in the local Lagrangian indicates that only $H := SO(3) \subset G$ reduction may be carried out, rather than the full $SE(3)$ reduction. Therefore, the symmetry breaking introduced by ρ dependence in the Lagrangian requires the *subgroup Lagrange-Poincaré reduction* discussed in §4.3.1, rather than the Euler-Poincaré reduction of §4.2.

With the notation of §4.2.2, P may be regarded as a principal $SO(3)$ -bundle over $\Sigma := P/H = X \times \mathbb{R}^3$ with the following projection

$$\pi_{\Sigma, P} : P \rightarrow \Sigma, \quad \pi_{\Sigma, P}(x, \mathbf{r}, \Lambda) = (x, \Lambda^{-1}\mathbf{r}) = (x, \rho).$$

Due to the assumption that $\pi_{\Sigma, P}$ is a trivial bundle, the adjoint bundle $\text{Ad } P$ may be trivialised using the isomorphism that corresponds to (4.30). That is,

$$[(x, \Lambda, \mathbf{r}), \xi]_{\mathfrak{so}(3)} \mapsto ((x, \rho), \text{Ad}_{\Lambda^{-1}}\xi).$$

Using the Maurer-Cartan connection, $\mathcal{A}(v_x, v_\Lambda, v_r) = v_\Lambda \Lambda^{-1}$, the isomorphism (4.35) reads

$$[(x, \mathbf{r}, \Lambda, \text{id}, d\mathbf{r}, d\Lambda)]_{SO(3)} \mapsto (x, \rho, d\rho, \Lambda^{-1}d\Lambda) = (x, \rho, \dot{\rho}dt + \rho'ds, \omega dt + \Omega ds), \quad (4.36)$$

where $\rho = \Lambda^{-1}\mathbf{r}$, $\omega = \Lambda^{-1}\dot{\Lambda}$, $\Omega = \Lambda^{-1}\Lambda'$, $(\dot{\cdot})$ denotes differentiation with respect to t , and $(\cdot')$ denotes differentiation with respect to s .

Applying Theorem ?? to an arbitrary reduced Lagrangian, $l(\rho, \rho_s, \rho_t, \Omega, \omega)$, re-

sults in the following Lagrange-Poincaré equations written in local coordinates.

$$\begin{cases} (\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\omega}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta l}{\delta \boldsymbol{\Omega}} = 0, \\ \partial_t \frac{\delta l}{\delta \dot{\boldsymbol{\rho}}} + \partial_s \frac{\delta l}{\delta \dot{\boldsymbol{\rho}'}} - \frac{\delta l}{\delta \boldsymbol{\rho}} = 0. \end{cases} \quad (4.37)$$

Equations (4.37) are identical to (4.21), which describe the molecular strand with nonlocal interactions after the coordinate transformation (4.5); this shows that the molecular strand equations (4.21) result from subgroup Lagrange-Poincaré reduction of a system governed by the Euler-Lagrange equations.

Equations (4.37) need to be augmented with an integrability condition to allow reconstruction. This integrability/reconstruction condition is related to the curvature of the connection form $\mathcal{A}$. For the molecular strand the required reconstruction condition is

$$\partial_t \boldsymbol{\Omega} - \partial_s \boldsymbol{\omega} - \boldsymbol{\omega} \times \boldsymbol{\Omega} = 0. \quad (4.38)$$

Relation (4.38) appears as a kinematic condition (2.17), and also in the transformed and generalised form in (4.25). The remaining kinematic conditions in (4.25) show that $(\boldsymbol{\rho}, \dot{\boldsymbol{\rho}} dt + \dot{\boldsymbol{\rho}'} ds) = j^1 \boldsymbol{\rho}$ is the first jet extension of $\boldsymbol{\rho}$. More details about the reconstruction condition can be found in §7.1.3, and in particular in Theorem 7.1.9 and subsequent remarks.

4.3.3 The general case

The subgroup Lagrange-Poincaré approach generalises to the higher dimensional strand with an arbitrary Lie group structure $\mathcal{O}$, in the setting of §2.1.3, as follows. Consider the $(n+1)$ -dimensional space-time $X := \mathcal{D} \times \mathbb{R}$ and the trivial fibre bundle

$$\pi_{X,P} : P := X \times S \rightarrow X,$$

where $S = \mathcal{O} \otimes E$ and E is a representation space of $\mathcal{O}$. A section σ of P reads

$$\sigma(x) = (x, \Lambda(x), r(x)), \quad x = (s, t) \in X,$$

Covariant formulation of molecular strands

and its first jet extension is

$$j^1\sigma(x) = (T_x\Lambda, T_xr) = (\mathbf{d}\Lambda(x) + \dot{\Lambda}(x)dt, \mathbf{d}r(x) + \dot{r}(x)dt),$$

where $\mathbf{d}$ is the partial derivative with respect to space (that is, the derivative on $\mathcal{D}$), and the dot is the partial derivative with respect to time.

There is a natural $\mathcal{O}$ -principal bundle structure on S given by

$$\pi_{E,S} : S \rightarrow E, \quad \pi_{ES}(\Lambda, r) = \Lambda^{-1}r = \rho.$$

This principal bundle structure on the fibre S induces a principal $\mathcal{O}$ -bundle structure on P given by

$$\pi_{\Sigma,P} : P \rightarrow X \times E, \quad \pi_{\Sigma P}(x, \Lambda, r) = (x, \Lambda^{-1}r).$$

There is also a natural connection $\mathcal{A}$ on $\pi_{\Sigma P} : P \rightarrow \Sigma := X \times E$ given by

$$\mathcal{A}(v_x, v_\Lambda, (r, u)) = v_\Lambda \Lambda^{-1},$$

which allows identification of the reduced jet bundle $J^1P/\mathcal{O}$ with the fibre bundle $J^1\Sigma \times_\Sigma L(TX, \text{Ad}P)$. Using the same notations as before,

$$\alpha_{\mathcal{A}} \circ [j^1\sigma(x)]_{\mathcal{O}} = T_x\rho \oplus \bar{\mathcal{A}} \circ T_x\sigma \cong (x, \rho(x), \mathbf{d}\rho(x) + \dot{\rho}(x), \Omega(x) + \omega(x)dt),$$

by (2.18). The vertical and horizontal variations are given by

$$\begin{aligned} \delta^v [j^1\sigma(x)]_{\mathcal{O}} &= (x, \Lambda^{-1}r(x); 0, \mathbf{d}\Sigma + [\Omega, \Sigma] + (\dot{\Sigma} + [\omega, \Sigma])dt), \\ \delta^h [j^1\sigma(x)]_{\mathcal{O}} &= (x, \rho(x); \mathbf{d}(\delta\rho) + (\delta\rho)_t dt, 0). \end{aligned}$$

Correspondingly, the vertical and horizontal Lagrange-Poincaré equations are

$$(\partial_t - \text{ad}_\omega^*) \frac{\delta \bar{l}}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta \bar{l}}{\delta \Omega} = 0, \quad (4.39)$$

$$\frac{\delta \bar{l}}{\delta \rho} - \partial_t \frac{\delta \bar{l}}{\delta \rho_t} - \text{div} \frac{\delta \bar{l}}{\delta \rho_s} = 0. \quad (4.40)$$

Of course, as expected, these equations coincide with equations (4.23) obtained

from the affine Euler-Poincaré equations (3.29) by the change of variables (4.22).

4.4 Summary

In this chapter we have established an alternative geometric interpretation of the molecular strand equations from the classical perspective. The alternative interpretation comes from the covariant perspective, and describes the equations of motion as coming from subgroup Lagrange-Poincaré reduction.

The derivation began by introducing a change of coordinates in §4.1. Similarly to the Hamilton-Pontryagin approach, the length of the calculation meant that it was presented in two parts. The local transformation was dealt with in Theorem 4.1.2, while the nonlocal terms were calculated in Theorem 4.1.5. The coordinate transformation was then generalised in §4.1.3.

The geometric interpretation of the coordinate transformed equations of motion fell into two cases, depending on whether or not ρ was present in the local Lagrangian.

The case when ρ is absent was identified as an application of covariant Euler-Poincaré reduction in §4.2. Some background material was given, after which an application of Theorem 7.1.8 to the molecular strand was made in §4.2.2.

In §4.3, subgroup Lagrange-Poncaré reduction yielded a geometric interpretation of the molecular strand equations from the covariant perspective when ρ is present. Further background material was given that built upon §4.2.1 in §4.3.1, which culminated in Theorem ???. Subsequently, the application to the strand was made in §4.3.2. The generalisation of the molecular strand was treated from the covariant perspective in §4.3.3.

CHAPTER 5 : CONCLUSION OF PART I

Part I formulates the problem of strand dynamics for an arbitrary long-range inter-molecular potential in the convective representation HOLM ET AL. [1986] of exact geometric rod theory SIMÓ ET AL. [1988]. Its methods would also apply in the consideration of Lennard-Jones potentials and the constrained motion of non-self-interacting curves.

Three different formulations of the same continuum equations of motion for an elastic strand experiencing nonlocal (for example, electrostatic or Lennard-Jones) interactions are derived and compared. These are: the Hamilton-Pontryagin approach, the affine Euler-Poincaré approach and the covariant Lagrange-Poincaré approach.

Part I concentrates primarily on the one-dimensional strand, which is a major object of interest for biological applications. However, these approaches for deriving continuum motion equations possess more significance and applicability than might be suggested by the one-dimensional developments illustrated here. For example, the geometrical considerations and nonlinear context of the present investigation would also apply in formulating the dynamics of the higher dimensional case. That is, when s has more than one component, the approaches discussed here still apply.

Besides passing to higher dimensions, future studies will consider both linear and nonlinear wave propagation on electrostatically charged strands, as well as the description of nontrivial stationary states that arise from nonlocal interactions, such as for the VDF oligomers mentioned in the Chapter 1.

Yet another interesting question for future studies concerns the possibility of enhancing the internal structure of the rigid charge conformations. This will allow even richer dynamics. While the resulting equations may be different (and more complex), the methods developed here will still be applicable when the dynamics takes place in spaces that possess richer conformational structure than rigid rotations.

Many interesting and non-trivial issues for future research are raised by the symmetry reduced formulation of convective dynamics introduced here for non-local interactions of charged strands. As mentioned earlier, these issues include classification and stability analysis of equilibrium solutions, dynamics of conformational changes (folding/unfolding) and adaptation of these methods to computational approaches, all of which we must place beyond the scope of the present work.

The geometric content of the choice of coordinates provides an interesting comparison between the classical and covariant descriptions of the molecular strand. The coordinates (Γ, γ) arise when the symmetry group is $SE(3)$. This occurs for the classical perspective of affine Euler-Poincaré reduction described in Chapter 3 and for covariant Euler-Poincaré reduction when ρ is absent in §4.2. The coordinates (ρ_s, ρ_t) arise for subgroup covariant Lagrange-Poincaré reduction, where the symmetry group is taken to be $SO(3)$. Therefore the following classification may be given according to the local coordinates that arise in the theory:

Local Coordinates	Geometric Interpretation	Perspective
(Γ, γ) without ρ	covariant and affine Euler-Poincaré	Both
(Γ, γ) with ρ	affine Euler-Poincaré	Classical
(ρ_s, ρ_t) with or without ρ	subgroup covariant Lagrange-Poincaré	Covariant

Figure 5.1: A table is shown that classifies the geometric interpretation of the molecular strand equations, according to local coordinates.

*There is a tide in the affairs of men
Which taken at the flood, leads on to fortune;
Omitted, all the voyage of their life
Is bound in shallows and in miseries.*

from 'Julius Caesar' by William Shakespeare

Part II

Lagrange-Poincaré field reduction

CHAPTER 6 : INTRODUCTION

Lagrangian field theories, and their Hamiltonian counterparts, have aided the development of many diverse mathematical models in geometric mechanics, including string theory BECKER ET AL. [2007]; NAKAHARA [2003], particle physics SKYRME [1961]; GISIGER & PARANJAPPE [1998], fluid dynamics WHITAM [1966]; MARSDEN & SHKOLLER [1999], and molecular strand dynamics as in Part I.

As Part I highlights, reduction by symmetry techniques assist the development of these models by lowering the order of the partial differential equations, and providing fundamental geometric understanding of the models developed.

Two main approaches to reduction by symmetry of Lagrangian field theories have been developed. One approach, investigated, for example, in GOTAY ET AL. [2004], employs multisymplectic geometry to extend symplectic reduction using a momentum map of to the field theoretic setting.

The second approach, familiar from §4, was developed by CASTRILLÓN-LÓPEZ ET AL. [2000] and CASTRILLÓN-LÓPEZ & RATIU [2003], and reduces the variational principle itself. In §4.2 the molecular strand equations without the ρ variable are understood as an example of covariant Euler-Poincaré reduction as developed in CASTRILLÓN-LÓPEZ ET AL. [2000]. The reduction consisted of a principal bundle being reduced by its structure group. Subsequently, §4.3 extends the results of §4.2 by describing the molecular strand equations with the ρ variable as an example of subgroup Lagrange-Poincaré reduction, as in CASTRILLÓN-LÓPEZ & RATIU [2003]. In §4.3, a principal $SE(3)$ -bundle is reduced by the subgroup $SO(3) \subset SE(3)$ of the structure group.

These ideas go a long way to progress the covariant reduction of variational principles, however comparing the reduction of classical variational principles (that is, those with a single independent variable), shows that these results are not sufficient to recover the classical reduction from the covariant reduction. Despite this, the molecular strand application in Part I provides an example of reduction that may be transformed between the classical and covariant perspectives via a coordinate transformation. In order to better understand the relationship between

classical and covariant perspectives, it is desirable to develop a paradigm for reduction that captures both the classical and covariant perspectives. This objective is a subject of Part II, and the form of reduction that encompasses both classical and covariant perspectives called *fibre bundle reduction*, as in ELLIS ET AL. [2009].

Although a reduction theory that contains both the existing covariant reduction and the classical reduction is achieved, there are a number of points that are left to future work. The relationship between affine Euler-Poincaré reduction and covariant Euler-Poincaré reduction was established in GAY-BALMAZ & RATIU [2009]. One challenge that remains, however, is to relate the affine Euler-Poincaré reduction with subgroup Lagrange-Poincaré reduction in general. It is clear that some relationship must hold in certain situations, since the molecular strand of Part I provides an example that bridges both theories. The extent to which these two types of reduction overlap, and the precise relationship between them remains to be understood.

Part II presents the results of ELLIS ET AL. [2009], which achieves a full generalisation of the classical Lagrangian reduction method, HOLM ET AL. [1998]; CENDRA ET AL. [1997, 2001], to the field setting by extending the covariant variational principle theory, CASTRILLÓN-LÓPEZ ET AL. [2000]; CASTRILLÓN-LÓPEZ & RATIU [2003].

In §6.1 a brief review of the relevant reduction on the classical side CENDRA ET AL. [2001] is presented, and the natural extension to the covariant perspective is discussed. This discourse leads to a criteria for the new type of reduction. The remainder of §6, namely §6.2, introduces the geometric tools required to develop fibre bundle reduction.

6.1 Aims of Part II

The classical Lagrange-Poincaré reduction of CENDRA ET AL. [2001] is formulated as follows. A variational principle is formulated on a principal bundle $\pi : Q \rightarrow Q/G$ and a principal connection $\mathcal{A}$ is introduced on Q . The connection yields a bundle isomorphism

$$(TQ)/G \rightarrow T(Q/G) \oplus_{Q/G} \text{Ad}Q$$

Introduction

given by

$$[v_q] \mapsto \left(T\pi(v_q), \llbracket q, \mathcal{A}(v_q) \rrbracket_{\mathfrak{g}} \right).$$

Thus, a curve $q(t) \in Q$ induces the two curves

$$\rho(t) := \pi(q(t)) \in Q/G \quad \text{and} \quad \bar{\sigma}(t) = \llbracket q(t), \mathcal{A}(\dot{q}(t)) \rrbracket_{\mathfrak{g}} \in \text{Ad}Q.$$

Classical Lagrangian reduction states that the Euler-Lagrange equations on Q with a G invariant Lagrangian L are equivalent to the Lagrange-Poincaré equations on $TQ/G \cong T(Q/G) \oplus_{Q/G} \text{Ad}Q$ with reduced Lagrangian l . The Lagrange-Poincaré equations read

$$\begin{cases} \frac{D}{Dt} \frac{\delta l}{\delta \bar{\sigma}} - \text{ad}_{\bar{\sigma}}^* \frac{\delta l}{\delta \bar{\sigma}} = 0, \\ \frac{\delta l}{\delta \rho} - \frac{D}{Dt} \frac{\delta l}{\delta \dot{\rho}} = \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \mathbf{i}_{\dot{\rho}} \tilde{\mathcal{B}} \right\rangle, \end{cases}$$

where $\tilde{\mathcal{B}}$ is the reduced curvature form associated to the connection $\mathcal{A}$ and D/Dt denotes a suitable covariant derivative.

When the classical Lagrange-Poincaré reduction is generalised to field theories, the natural structure to consider is the trivial fibre bundle

$$\pi_{X,P} : P := X \times Q \rightarrow X, \quad \pi_{X,P}(x, q) := x.$$

Now $\Gamma\pi_{X,P}$, the space of sections of $\pi_{X,P}$, generalises the space of curves in Q , and the principal bundle structure π on Q gives a natural principal bundle structure

$$\pi_{\Sigma,P} : P := X \times Q \rightarrow \Sigma := X \times (Q/G), \quad \pi_{\Sigma,P}(x, q) := (x, \pi(q)).$$

More generally, the structure that arises is described by the following commutative diagram

$$\begin{array}{ccc} P & \xrightarrow{\pi_{X,P}} & X \\ & \searrow \pi_{\Sigma,P} & \nearrow \pi_{X,\Sigma} \\ & \Sigma & \end{array}$$

where $\pi_{X,P} : P \rightarrow X$ is *any* fibre bundle and $\pi_{\Sigma,P} : P \rightarrow \Sigma$ is a principal bundle, whose group action preserves the fibres of $\pi_{X,P}$.

Following these observations, a specification for covariant Lagrangian reduction theory emerges: The reduction theory should

1. Capture the natural generalisation of classical Lagrange-Poincaré reduction of CENDRA ET AL. [2001] to field theories.
2. Reduce to the work of CASTRILLÓN-LÓPEZ ET AL. [2000] and CASTRILLÓN-LÓPEZ & RATIU [2003] as particular cases.

These objectives are constantly kept in mind, and provide a guide for the development of the theory. To ensure that the objectives are met, four particular examples are referred to. These particular examples are: the *unreduced case*, *principal bundle reduction*, *subgroup reduction*, and *classical reduction*. The method to check that the theory being developed achieves the desired goals is to routinely specialise to each of the four particular cases, to ensure that consistency is achieved with the established theory.

In §6.2 some geometric tools that are necessary for performing reduction are introduced. The relationship between two bundle structures on the same manifold P is also studied. In §7, the Lagrange-Poincaré field equations are developed and a method for reconstruction is given. The Kelvin-Noether theorem associated to the Lagrange-Poincaré field equations is presented in §8.

6.2 Geometric constructions

There are two main geometric constructions of interest. The first is the interaction between two bundle structures, $\pi_{X,P}$ and $\pi_{\Sigma,P}$ on P . The second is the reduction of the jet bundle J^1P by the structure group G . The first of these constructions is the subject of §6.2.1, the second construction is addressed in §6.2.2.

6.2.1 Geometric setting

Consider a locally trivial fibre bundle $\pi_{X,P} : P \rightarrow X$.

Introduction

Definition 6.2.1 A *section* of $\pi_{X,P}$ is a smooth map $\sigma : X \rightarrow P$ such that

$$\pi_{X,P} \circ \sigma = \text{id}_X.$$

The space of smooth sections of $\pi_{X,P}$ is denoted $\Gamma\pi_{X,P}$.

Remark 6.2.2 It is necessary to introduce many fibre bundle projections during the development of the theory. The notation indicates the source and target space, e.g. $\pi_{X,P} : P \rightarrow X$, where the first subscript denotes the base space and the second the total space. The order of the subscripts allows one to write, for example,

$$\pi_{X,Y} \circ \pi_{Y,Z} = \pi_{X,Z}.$$

Definition 6.2.3 The *first jet bundle* of $\pi_{X,P}$ is the affine bundle $\pi_{P,J^1P} : J^1P \rightarrow P$ whose fibre at p is the affine space

$$J_p^1P = \{\gamma_p \in L(T_xX, T_pP) \mid T_p\pi_{X,P} \circ \gamma_p = \text{id}_{T_xX}, x = \pi_{X,P}(p)\},$$

where $L(T_xX, T_pP)$ denotes the space of linear maps $\gamma_p : T_xX \rightarrow T_pP$.

Remark 6.2.4 The first jet bundle is the natural generalisation of the tangent bundle to the field theoretic context. Therefore J^1P plays the role of the *unreduced state space* in applications. The manifold J^1P may also be regarded as a locally trivial fibre bundle over X , that is, $\pi_{X,J^1P} : J^1P \rightarrow X$ with $\pi_{X,J^1P}(\gamma_p) := \pi_{X,P}(p)$.

Definition 6.2.5 Given $\sigma \in \Gamma\pi_{X,P}$, the *first jet extension* of σ is defined by

$$j^1\sigma(x) := T_x\sigma.$$

$j^1\sigma$ is a section of π_{X,J^1P} .

Suppose there is a free and proper left action Φ of a Lie group G on P such that

$$\pi_{X,P} \circ \Phi_g = \pi_{X,P}, \quad \text{for all } g \in G. \quad (6.1)$$

Equation (6.1) is equivalent to the assumption that the action of G preserves the fibres of $\pi_{X,P}$. Since the action is free and proper, there exists a principal bundle

$\pi_{\Sigma,P} : P \rightarrow \Sigma$, where $\Sigma := P/G$. Here Σ is the equivalent of *shape space* in applications. Since, by (6.1), the projection $\pi_{X,P}$ is G -invariant, it induces a surjective submersion $\pi_{X,\Sigma} : \Sigma \rightarrow X$ via the relation

$$\pi_{X,\Sigma} \circ \pi_{\Sigma,P} = \pi_{X,P}. \quad (6.2)$$

It is easily verified that if $\pi_{X,P}$ is proper then $\pi_{X,\Sigma}$ is also proper.

More generally, if $\pi_{X,P}$ is a locally trivial fibre bundle then $\pi_{X,\Sigma}$ is also a locally trivial fibre bundle. To see this, take a fibre bundle chart $\psi : \pi_{X,P}^{-1}(U) \rightarrow U \times F$, where U is an open subset of X and the manifold F is the model of the fibre. By definition, $p_1 \circ \psi = \pi_{X,P}$, where $p_1 : U \times F \rightarrow U$ is the projection onto the first factor. Property (6.1) implies that $\pi_{X,P}^{-1}(U)$ is a G -invariant subset. Thus, the diffeomorphism ψ bestows a well-defined G -action on $U \times F$ which turns out to be free, proper, and acts only on the component F . The model fibre F thereby attains a principal bundle structure $F \rightarrow F/G$ induced by (and depending upon) the chart ψ . Since ψ is an equivariant diffeomorphism, it drops to a diffeomorphism $\bar{\psi} : \pi_{X,P}^{-1}(U)/G \rightarrow U \times F/G$. Also, since $\pi_{X,P}^{-1}(U)/G = \pi_{X,\Sigma}^{-1}(U)$ and $p_1 \circ \bar{\psi} = \pi_{X,\Sigma}$, the map $\bar{\psi}$ is a fibre bundle chart of $\pi_{X,\Sigma} : \Sigma \rightarrow X$. For principal bundles, it is necessary to work with local sections, since a principal bundle does not have global sections unless it is trivial.

Remark 6.2.6 There are now two different bundles $\pi_{X,P} : P \rightarrow X$ and $\pi_{\Sigma,P} : P \rightarrow \Sigma$ with the same total space P . In general, the associated vertical distributions do not coincide, although (6.1) provides the inclusion

$$\ker(T_p \pi_{\Sigma,P}) \subset \ker(T_p \pi_{X,P}).$$

Thus, it is possible to associate two different jet bundles to P . Here, the only one of interest is $\pi_{X,P} : P \rightarrow X$ and hence there is no ambiguity in the notation $J^1 P$.

Lagrangian field theories are described by a Lagrangian density $\mathcal{L} : J^1 P \rightarrow \Lambda^{n+1} X$ defined on the first jet bundle. Here $\Lambda^{n+1} X$ denotes the bundle of $(n+1)$ -forms on X , where $n+1 = \dim X$. In this context the G -action on P , lifted to $J^1 P$, should be interpreted as a symmetry of the Lagrangian density. The associated reduction process, described in the next section, is called the *covariant Lagrange-Poincaré reduction*.

Particular cases

Various other theories may be identified as particular cases of the geometric setting developed in Part II. These examples are referred to throughout the discourse and serve to illustrate the ideas introduced in a more familiar context while demonstrating how the objective of capturing previous theories in this new paradigm is fulfilled.

- i If $\Sigma = P$, that is, $G = \{1\}$, there are no symmetries. The principal bundle structure disappears and the geometric setting for *covariant Lagrangian field theory*, referred to as the *unreduced case*, emerges. The commutative diagram that describes this case is

$$\begin{array}{ccc}
 P & \xrightarrow{\pi_{X,P}} & X \\
 \pi_{\Sigma,P} = \text{id}_P \searrow & & \nearrow \pi_{X,\Sigma} = \pi_{X,P} \\
 & \Sigma = P &
 \end{array}$$

where $\pi_{X,P}$ is a fibre bundle.

- ii Assume that the configuration space $\pi_{X,P} : P \rightarrow X$ is itself a principal G -bundle and G is also the symmetry group. Then $\Sigma = X$ and $\pi_{X,\Sigma}$ is the identity map. This recovers the geometric setting for *covariant Euler-Poincaré reduction*, or *principal bundle reduction* in CASTRILLÓN-LÓPEZ ET AL. [2000] which is used to study, for example, the molecular strand in §4.2. The commutative diagram that describes this case is

$$\begin{array}{ccc}
 P & \xrightarrow{\pi_{X,P}} & X \\
 \pi_{\Sigma,P} = \pi_{X,P} \searrow & & \nearrow \pi_{X,\Sigma} = \text{id}_X \\
 & \Sigma = X &
 \end{array}$$

where $\pi_{X,P}$ is a principal bundle.

- iii The formulation of CASTRILLÓN-LÓPEZ & RATIU [2003] is recovered when $\pi_{X,P} : P \rightarrow X$ is a principal bundle such that the symmetry group G is a *subgroup* of the structure group H . This is the geometric setting for the formulation of the molecular strand from §4.3, and is referred to as *subgroup covariant Lagrange-Poincaré reduction* or simply *subgroup reduction*. The commutative diagram that describes this case is

$$\begin{array}{ccc} P & \xrightarrow{\pi_{X,P}} & X \\ & \searrow \pi_{\Sigma,P} & \nearrow \pi_{X,\Sigma} \\ & \Sigma & \end{array}$$

where $\pi_{X,P}$ and $\pi_{\Sigma,P}$ are respectively G and H -principal bundles.

- iv If $X = \mathbb{R}$, $\Sigma = \mathbb{R} \times M$, and $P = \mathbb{R} \times Q$, where $\pi_{M,Q} : Q \rightarrow M$ is a G -principal bundle, the geometric setting for Lagrangian reduction in classical mechanics, known as *classical reduction*, becomes apparent. Here Q plays the role of the configuration space. There are two well-known particular cases: If $Q = G$ (thus $M = \{m\}$) the geometric setting for *Euler-Poincaré reduction* surfaces (this is also a particular case of ii), where the configuration space coincides with the group of symmetries; if $G = \{1\}$ (thus $Q = M$) there are no symmetries and the geometric setting for unreduced classical Lagrangian mechanics emerges (this is also a particular case of iii). The commutative diagram that describes this case is

$$\begin{array}{ccc} P = \mathbb{R} \times Q & \xrightarrow{\pi_{X,P} = p_1} & X = \mathbb{R} \\ & \searrow \pi_{\Sigma,P} & \nearrow \pi_{X,\Sigma} = p_1 \\ & \Sigma = \mathbb{R} \times M & \end{array}$$

where $\pi_{M,Q}$ is a principal bundle and $\pi_{\Sigma,P} := \text{id}_{\mathbb{R}} \times \pi_{M,Q}$.

Adjoint bundle

Definition 6.2.7 The *adjoint bundle* associated with the principal bundle $\pi_{\Sigma, P} : P \rightarrow \Sigma$ is a vector bundle $\pi_{\Sigma, \text{Ad } P} : \text{Ad } P \rightarrow \Sigma$. The total space $\text{Ad } P$ is the quotient space $P \times_G \mathfrak{g}$ relative to the following left action of G on $P \times \mathfrak{g}$: $(p, \xi) \mapsto (\Phi_g(p), \text{Ad}_g \xi)$.

Notation 6.2.8 Elements in the adjoint bundle are equivalence classes $\llbracket p, \xi \rrbracket_{\mathfrak{g}}$ and the projection is described by $\pi_{\Sigma, \text{Ad } P} \left(\llbracket p, \xi \rrbracket_{\mathfrak{g}} \right) = \pi_{\Sigma, P}(p)$.

The adjoint bundle is, in fact, a Lie algebra bundle. That is, each fibre $\text{Ad}_s P$, $s \in \Sigma$, has a natural Lie bracket

$$\left[\llbracket p, \xi \rrbracket_{\mathfrak{g}}, \llbracket p, \eta \rrbracket_{\mathfrak{g}} \right]_s := \llbracket p, [\xi, \eta] \rrbracket_{\mathfrak{g}},$$

where $\pi_{\Sigma, P}(p) = s$, $\xi, \eta \in \mathfrak{g}$ and these Lie brackets depend smoothly on the base variable $s \in \Sigma$.

This Lie algebra bundle structure enables the introduction of a wedge product. For 1-forms this wedge product $\wedge : \Omega^1(\Sigma; \text{Ad } P) \times \Omega^1(\Sigma; \text{Ad } P) \rightarrow \Omega^2(\Sigma; \text{Ad } P)$ is defined by

$$(\alpha \wedge \beta)(u_s, v_s) := [\alpha(u_s), \beta(v_s)]_s - [\alpha(v_s), \beta(u_s)]_s, \quad (6.3)$$

where $s \in \Sigma$ and $u_s, v_s \in T_s \Sigma$.

The different equivalence classes are interpreted as different representations of the dynamics. For given $s \in \Sigma$ and $p \in \pi_{\Sigma, P}^{-1}(s)$, the map $\chi_p : \text{Ad}_s P \rightarrow \mathfrak{g}$ given by

$$\chi_p \left(\llbracket q, \eta \rrbracket_{\mathfrak{g}} \right) = \xi, \quad (6.4)$$

where $\xi \in \mathfrak{g}$ is such that $\llbracket p, \xi \rrbracket_{\mathfrak{g}} = \llbracket q, \eta \rrbracket_{\mathfrak{g}}$, defines a p -dependent Lie algebra isomorphism.

The choice of $p \in \pi_{\Sigma, P}^{-1}(s)$ determines the representation of the dynamics. Thus the *convective* or the *spatial* representation of the dynamics may be found by altering the element p for χ_p . Investigation of the local representations of the dynamics is engaged in further in §8.

Definition 6.2.9 A *connection form* $\mathcal{A}$ on the principal bundle $\pi_{\Sigma, P} : P \rightarrow \Sigma$ is a

one-form $\mathcal{A} \in \Omega^1(P; \mathfrak{g})$ such that

$$\Phi_g^* \mathcal{A} = \text{Ad}_g \circ \mathcal{A} \quad \text{and} \quad \mathcal{A}(\xi_P(p)) = \xi, \quad \text{where} \quad \xi_P(p) := \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp(t\xi)}(p)$$

is the infinitesimal generator associated to the Lie algebra element $\xi \in \mathfrak{g}$.

Definition 6.2.10 The *horizontal distribution* associated to $\mathcal{A}$ is the subbundle $HP \subset TP$ defined by

$$H_p P := \ker (\mathcal{A}(p)).$$

The horizontal distribution is complementary to the vertical distribution

$$V_p P := \ker (T_p \pi_{\Sigma P}),$$

and consequently TP decomposes as $T_p P = H_p P \oplus V_p P$

Definition 6.2.11 The *horizontal lift operator* $\text{Hor}_p^{\mathcal{A}} : T_s \Sigma \rightarrow H_p P$ associated with the connection $\mathcal{A}$ is defined according to

$$\text{Hor}_p^{\mathcal{A}}(v_s) \in H_p P \quad \text{and} \quad T_s \pi_{\Sigma P} \circ \text{Hor}_p^{\mathcal{A}} = \text{id}_{T_s \Sigma}$$

where $s \in \Sigma$, $p \in \pi_{\Sigma, P}^{-1}(s)$ and $v_s \in T_s \Sigma$.

Definition 6.2.12 The *curvature form* $\mathcal{B} \in \Omega^2(P; \mathfrak{g})$ associated with the connection form $\mathcal{A}$ on $\pi_{\Sigma, P}$ is defined by the formula

$$\mathcal{B} = d\mathcal{A} - \frac{1}{2} \mathcal{A} \wedge \mathcal{A}.$$

Remark 6.2.13 (Curvature form and integrability) The horizontal distribution associated with the connection form $\mathcal{A}$ is integrable if and only if the curvature form $\mathcal{B}$ vanishes. A connection with zero curvature is referred to as *flat*. This assertion is well-known, a proof follows from Frobenius' Theorem and can be found in, for example, CENDRA ET AL. [2001]. The curvature form may therefore be interpreted as a geometric object that measures nonintegrability of the connection.

Introduction

Definition 6.2.14 *The covariant derivative induced on the adjoint bundle by the connection $\mathcal{A}$,*

$$\nabla^{\mathcal{A}} : \Gamma\pi_{\Sigma, \text{Ad } P} \rightarrow \Gamma\pi_{\Sigma, L(T\Sigma, \text{Ad } P)}, \quad (6.5)$$

is defined by

$$\nabla_{v_s}^{\mathcal{A}} \sigma(s) = \left. \frac{D^{\mathcal{A}}}{Dt} \right|_{t=0} \sigma(c(t)) = \llbracket p(s), \mathbf{d}\xi(v_s) - [\mathcal{A}(Tp(v_s)), \xi(s)] \rrbracket_{\mathfrak{g}}, \quad (6.6)$$

where $\xi : \Sigma \rightarrow \mathfrak{g}$ and $p : \Sigma \rightarrow P$ are such that $\sigma(s) = \llbracket p(s), \xi(s) \rrbracket_{\mathfrak{g}}$ and $c(t)$ is a curve in Σ such that $\dot{c}(0) = v_s$ (see CENDRA ET AL. [2001], Lemma 2.3.4).

Remark 6.2.15 The covariant derivative $\nabla^{\mathcal{A}}$ also has an interpretation as a bilinear map

$$\nabla^{\mathcal{A}} : \mathfrak{X}(\Sigma) \times \Gamma\pi_{\Sigma, \text{Ad } P} \rightarrow \Gamma\pi_{\Sigma, \text{Ad } P}, \quad (X, \sigma) \mapsto \nabla_X^{\mathcal{A}} \sigma.$$

6.2.2 Reduced covariant configuration space

The free and proper action $\Phi : G \times P \rightarrow P$ induces a free and proper action $\Phi^1 : G \times J^1 P \rightarrow J^1 P$ defined by

$$\Phi_g^1(\gamma_p) := T\Phi_g \circ \gamma_p, \quad \gamma_p \in J_p^1 P. \quad (6.7)$$

Note that this action preserves $J^1 P$ since, by (6.1),

$$T\pi_{X, P} \circ \Phi_g^1(\gamma_p) = T\pi_{X, P} \circ T\Phi_g \circ \gamma_p = T\pi_{X, P} \circ \gamma_p = \text{id}_{T_x X}.$$

Thus it is valid to consider the quotient manifold $J^1 P/G \ni [\gamma_p]_G$.

Remark 6.2.16 Recall that $J^1 P$ denotes the first jet bundle of P as a fibre bundle over X and not as a principal bundle over Σ .

The connection $\mathcal{A}$ on the principal bundle $\pi_{\Sigma P} : P \rightarrow \Sigma$ introduces the smooth map $\beta_{\mathcal{A}}$, which is defined by

$$\beta_{\mathcal{A}} : J^1 P/G \rightarrow J^1 \Sigma \oplus_{\Sigma} \mathcal{P}, \quad \beta_{\mathcal{A}}([\gamma_p]_G) := \left(T\pi_{\Sigma, P} \circ \gamma_p, \llbracket p, \mathcal{A}(\gamma_p(-)) \rrbracket_{\mathfrak{g}} \right), \quad (6.8)$$

where

$$\pi_{\Sigma, \mathcal{P}} : \mathcal{P} \rightarrow \Sigma$$

is the vector bundle whose fibre at $s \in \Sigma$ is $L(T_{\pi_X, \Sigma(s)} X, \text{Ad}_s P)$.

The map $\beta_{\mathcal{A}}$ is an diffeomorphism, the inverse being given by

$$\beta_{\mathcal{A}}^{-1} : J^1 \Sigma \oplus_{\Sigma} \mathcal{P} \rightarrow J^1 P/G, \quad \beta_{\mathcal{A}}^{-1}(\delta_s, l_s) = [\text{Hor}_p^{\mathcal{A}} \circ \delta_s + \zeta_p \circ l_s]_G,$$

where $p \in P$ is such that $\pi_{\Sigma, P}(p) = s$ and

$$\zeta_p : \text{Ad}_s P \rightarrow V_p P$$

is defined by $\zeta_p \left(\llbracket q, \eta \rrbracket_{\mathfrak{g}} \right) = \xi_P(p)$ where $\xi \in \mathfrak{g}$ is such that $\llbracket q, \eta \rrbracket_{\mathfrak{g}} = \llbracket p, \xi \rrbracket_{\mathfrak{g}}$. Note that $\text{Hor}_p^{\mathcal{A}} \circ \delta_s + \zeta_p \circ l_s \in J_p^1 P$ and that its equivalence class does not depend on which p is chosen. Also, note that the diffeomorphism $\beta_{\mathcal{A}}$ endows the manifold $J^1 P/G$ with the structure of an affine bundle over Σ .

The isomorphism $\beta_{\mathcal{A}}$ is interpreted as follows in the four particular cases:

- i** Here $G = \{1\}$ thus the principal bundle structure disappears. The bundle isomorphism is the identity on $J^1 P$.
- ii** Here $\Sigma = X$, thus $\beta_{\mathcal{A}}$ is a bundle map over X and

$$\beta_{\mathcal{A}} : J^1 P/G \rightarrow \Lambda^1(X; \text{Ad } P) \cong \mathcal{P}, \quad \beta_{\mathcal{A}}([\gamma_p]) = \llbracket p, \mathcal{A}(\gamma_p(-)) \rrbracket_{\mathfrak{g}}.$$

Therefore, the isomorphism used in the Euler-Poincaré reduction is recovered, see formula (2.5) in CASTRILLÓN-LÓPEZ ET AL. [2000].

- iii** The isomorphism used in the subgroup Lagrange-Poincaré reduction re-emerges, see Proposition 3 in CASTRILLÓN-LÓPEZ & RATIU [2003].
- iv** Since $X = \mathbb{R}$ and $P = \mathbb{R} \times Q$, the jet bundle $J^1 P$ may be identified with $\mathbb{R} \times TQ$. Thus, $J^1 P/G \simeq \mathbb{R} \times (TQ/G)$. Similarly, $J^1 \Sigma$ may be identified with $\mathbb{R} \times TM$ and $\text{Ad } P$ with $\mathbb{R} \times \text{Ad } Q$. The bundle $J^1 \Sigma \oplus_{\Sigma} \mathcal{P}$ can therefore be identified with $\mathbb{R} \times (TM \oplus_M \text{Ad } Q)$. A connection γ on Q naturally induces a connection $\mathcal{A}$ on P and the bundle map $\beta_{\mathcal{A}}$ reads $\beta_{\mathcal{A}} : \mathbb{R} \times (TQ/G) \rightarrow \mathbb{R} \times (TM \oplus_M \text{Ad } Q)$,

$$\beta_{\mathcal{A}}(t, [v_q]_G) = \left(t, T\pi_{M, Q}(v_q), \llbracket q, \gamma(v_q) \rrbracket_{\mathfrak{g}} \right). \quad (6.9)$$

Therefore the usual connection dependent isomorphism $TQ/G \simeq TM \oplus_M \text{Ad } Q$ used in classical Lagrangian reduction is recovered, as in CENDRA ET AL. [2001].

6.3 Summary

In §6 we have looked at the need for a covariant Lagrange-Poincaré reduction, or fibre bundle reduction, that is compatible with both the classical reduction, described in CENDRA ET AL. [2001], and the covariant reduction in CASTRILLÓN-LÓPEZ ET AL. [2000]; CASTRILLÓN-LÓPEZ & RATIU [2003]. These objectives were codified in §6.1.

The geometric setting for fibre bundle reduction consisted of the interaction of two bundle structures on a manifold P , this was treated in §6.2 along with the introduction of various spaces and geometric tools.

Subsequently, §6.2.2 tackled the issue of characterising the reduced state space, J^1P/G . This was accomplished by the introduction of a vector bundle isomorphism (6.8).

Throughout, reference was made to four particular cases: unreduced, principle, subgroup and classical reduction. These particular cases demonstrate at each stage through the development of the theory that the objectives set out in §6.1 are satisfied.

The development proceeds in §7 with the derivation of the reduction and reconstruction theorems, Theorems 7.1.8 and 7.1.9 respectively. These follow some preliminary calculations concerning required the variations and derivatives.

Following on, §8 focuses on the local representations of the dynamics in vector bundle trivialisations. The spatial and convective representations are studied, and the Kelvin-Noether Theorem 8.2.4 extends the result from CENDRA ET AL. [1997] two ways, from Euler-Poincaré to Lagrange-Poincaré, and from the classical to the covariant setting. The Kelvin-Noether Theorem is then

applied to the molecular strand.

CHAPTER 7: LAGRANGE-POINCARÉ FIELD REDUCTION

7.1 Lagrange-Poincaré field equations

Consider a G -invariant Lagrangian density $\mathcal{L} : \Gamma\pi_{X,J^1P} \rightarrow \Omega^{n+1}(X)$. For simplicity, suppose that X is orientable and fix a volume form μ on X . The Lagrangian density may thereby be expressed as $\mathcal{L} = L\mu$, where $L : \Gamma\pi_{X,J^1P} \rightarrow C^\infty(X)$.

Let $U \subset X$ be an open subset whose closure $\bar{U}$ is compact. Recall that a section $\sigma : \bar{U} \rightarrow P$ of $\pi_{X,P}$ is, by definition, smooth if for every point $x \in \bar{U}$ there is an open neighbourhood U_x of x and a smooth section $\sigma_x : U_x \rightarrow P$ extending σ .

Definition 7.1.1 *A critical section of the variational problem associated to $\mathcal{L}$ is defined as a smooth local section $\sigma : \bar{U} \rightarrow P$ of $\pi_{X,P}$ that satisfies*

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U \mathcal{L}(j^1\sigma_\varepsilon) = 0,$$

for all smooth variations $\sigma_\varepsilon : \bar{U} \rightarrow P$ such that $\sigma_0 = \sigma$ and $\sigma_\varepsilon|_{\partial U} = \sigma|_{\partial U}$.

Since

$$\delta\sigma(x) := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \sigma_\varepsilon(x) \in V_{\sigma(x)}P := \ker(T_{\sigma(x)}\pi_{X,P}) \quad \text{and} \quad \delta\sigma|_{\partial U} = 0,$$

it is possible to assume, without loss of generality, that $\sigma_\varepsilon = \phi_\varepsilon \circ \sigma$, where ϕ_ε is the flow of a vertical (with respect to the bundle structure $\pi_{X,P}$) vector field $V \in \mathfrak{X}^V(P)$ such that $V \circ \sigma(x) = 0$ for all $x \in \partial U$. The smooth Tietze extension theorem facilitates V 's definition over the whole manifold P , but values of V outside $\sigma(\bar{U})$ will not play any role in any subsequent consideration. Note that $\delta\sigma(x) = V \circ \sigma(x)$ for all $x \in \bar{U}$. Consequently,

$$\delta j^1\sigma(x) := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} j^1\sigma_\varepsilon(x) = j^1V \circ j^1\sigma(x) \in V_{j^1\sigma(x)}J^1P = \ker(T_{j^1\sigma(x)}\pi_{X,J^1P}),$$

where $V \in \mathfrak{X}^V(P) \mapsto j^1V \in \mathfrak{X}^V(J^1P)$ is the 1-jet lift of vector fields. Thus, σ is a critical section of the variational problem defined by $\mathcal{L}$ if and only if

$$0 = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U L(j^1\sigma_\varepsilon) \mu = \int_U \left\langle \frac{\delta L}{\delta \sigma}, j^1V \circ j^1\sigma \right\rangle \mu,$$

where $(\delta L/\delta \sigma)(x) \in V_{j^1\sigma(x)}^* J^1P$ is the differential along $j^1\sigma$. That is,

$$\left\langle \frac{\delta L}{\delta \sigma}, Z \circ j^1\sigma \right\rangle = \mathbf{d}L(j^1\sigma)(Z \circ j^1\sigma)$$

for arbitrary vector fields $Z \in \mathfrak{X}^V(J^1P)$ that are vertical with respect to π_{X,J^1P} . Denoting by $\mathcal{EL}(L)$ the bundle morphism $\mathcal{EL}(L) : J^1P \rightarrow V^*P$ defined by the condition

$$\int_U \langle \mathcal{EL}(L)(j^1\sigma), V \circ \sigma \rangle \mu = \int_U \left\langle \frac{\delta L}{\delta \sigma}, j^1V \circ j^1\sigma \right\rangle \mu, \text{ for all } V \in \mathfrak{X}^V(P),$$

the covariant Euler-Lagrange equations can be written intrinsically as

$$\mathcal{EL}(L) = 0.$$

Here $\mathcal{EL}$ is represented locally by

$$\mathcal{EL}(L) = \left[\frac{\partial L}{\partial y^\alpha}(j^1\sigma) - \frac{\partial}{\partial x^i} \left(\frac{\partial L}{\partial v_i^\alpha}(j^1\sigma) \right) \right] dy^\alpha,$$

where $\mathcal{L} = L(x^i, y^\alpha, v_i^\alpha) d^{n+1}x$. Thus, in coordinates, the covariant Euler-Lagrange equations take the standard form

$$\frac{\partial L}{\partial y^\alpha}(j^1\sigma) - \frac{\partial}{\partial x^i} \left(\frac{\partial L}{\partial v_i^\alpha}(j^1\sigma) \right) = 0. \quad (7.1)$$

These equations may be written globally by using a connection on the affine bundle $\pi_{P,J^1P} : J^1P \rightarrow P$; this point of view is used at the reduced level.

By G -invariance, L induces a reduced Lagrangian $l : \Gamma\pi_{X,J^1P/G} \rightarrow C^\infty(X)$. Fixing a connection $\mathcal{A}$ on the principal bundle $\pi_{\Sigma,P} : P \rightarrow \Sigma$ brings in the bundle isomorphism $\beta_{\mathcal{A}}$ defined in (6.8), thereby permitting the definition of the reduced Lagrangian l on $\Gamma\pi_{X,J^1\Sigma \oplus_\Sigma P} \cong \Gamma\pi_{X,J^1P/G}$.

Lagrange-Poincaré field reduction

A section $\sigma \in \Gamma\pi_{X,P}$ of the configuration bundle induces a section

$$\rho := \pi_{\Sigma,P} \circ \sigma \in \Gamma\pi_{X,\Sigma}$$

by (6.2). Similarly, the *reduced section* is defined as

$$\bar{\sigma} := \llbracket \sigma, \sigma^* \mathcal{A} \rrbracket_{\mathfrak{g}} \in \Gamma\pi_{X,\mathcal{P}}. \quad (7.2)$$

Thus,

$$(j^1\rho, \bar{\sigma}) = \beta_{\mathcal{A}}([j^1\sigma]_G) : X \rightarrow J^1\Sigma \oplus_{\Sigma} \mathcal{P}.$$

The two components are not independent since ρ can be obtained from $\bar{\sigma}$; explicitly,

$$\rho = \pi_{\Sigma,\mathcal{P}} \circ \bar{\sigma}.$$

Note that $(j^1\rho, \bar{\sigma})$ is a section of the bundle $J^1\Sigma \oplus_{\Sigma} \mathcal{P}$ viewed as a fibre bundle over X , and not as an affine bundle over Σ . These definitions and the G -invariance of $\mathcal{L}$ (and hence of L) yield

$$\mathcal{L}(j^1\sigma) = L(j^1\sigma)\mu = l(j^1\rho, \bar{\sigma})\mu \quad (7.3)$$

for any $\sigma \in \Gamma\pi_{X,P}$.

The previous considerations hold without changes when σ is a local section $\sigma : \bar{U} \rightarrow P$.

The fact that $J^1\Sigma \oplus_{\Sigma} \mathcal{P}$ is a locally trivial fibre bundle over X follows from the following observation: G acts on the locally trivial fibre bundle $\pi_{X,J^1P} : J^1P \rightarrow X$ by a free and proper action Φ^1 , such that $\pi_{X,J^1P} \circ \Phi_g^1 = \pi_{X,J^1P}$. Therefore, by the argument used in §6.2.1, $J^1P/G \rightarrow X$ is a locally fibre bundle. Thus, the isomorphism $\beta_{\mathcal{A}}$ ensures that $J^1\Sigma \oplus_{\Sigma} \mathcal{P}$ is a locally trivial fibre bundle over X .

7.1.1 Reduced variations

Using the bundle isomorphism $\beta_{\mathcal{A}}$, the variation of the action defined by $\mathcal{L}$ gives

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U \mathcal{L}(j^1\sigma_{\varepsilon}) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U l(j^1\rho_{\varepsilon}, \bar{\sigma}_{\varepsilon})\mu = 0,$$

where $U \subset X$ is an open subset and $\sigma_\varepsilon : \bar{U} \rightarrow P$ is a smooth variation of the smooth section $\sigma : \bar{U} \rightarrow P$.

A covariant derivative on the locally trivial fibre bundle

$$\pi_{X, \text{Ad } P} := \pi_{X, \Sigma} \circ \pi_{\Sigma, \text{Ad } P} : \text{Ad } P \rightarrow X$$

is required to compute the reduced variations. Recall the following general construction.

General constructions

Let $\tau : E \rightarrow \Sigma$ a vector bundle endowed with a covariant derivative ∇ . Recall that ∇ induces a covariant exterior derivative $d^\nabla : \Omega^k(\Sigma; E) \rightarrow \Omega^{k+1}(\Sigma; E)$ whose formula is a direct adaptation of the standard Cartan formula for k -forms on a manifold, by replacing all directional derivatives by covariant derivatives relative to ∇ . In particular, for one-forms

$$d^\nabla \alpha(U, V) = \nabla_U(\alpha(V)) - \nabla_V(\alpha(U)) - \alpha([U, V]), \quad (7.4)$$

where $\alpha \in \Omega^1(\Sigma; E)$ and $U, V \in \mathfrak{X}(\Sigma)$.

Definition 7.1.2 *Let X be an arbitrary manifold and $f : X \rightarrow E$ a smooth function. The ∇ -derivative of f is defined by*

$$\tilde{\nabla}_{v_x} f(x) := \left. \frac{D^\nabla}{Dt} \right|_{t=0} f(c(t)) \in E_{\tau(f(x))}, \quad (7.5)$$

where $c(t) \in X$ is a curve such that $\dot{c}(0) = v_x$ and D^∇/Dt is the usual covariant time derivative associated to ∇ of the curve $t \mapsto f(c(t))$ in E .

Note that $\tilde{\nabla} f(x) \in L(T_x X, E_{\tau(f(x))})$, and when $U \in \mathfrak{X}(X)$, the derivative $\tilde{\nabla}_U f$ is a function on X taking values in E .

The variational principle calculations require an exterior covariant derivative of forms on X with values in E . To make sense of this, assume that there is a smooth map $\pi : \Sigma \rightarrow X$. Recall that $\tau^k : \mathcal{P}_{X,E}^k \rightarrow \Sigma$ is the vector bundle whose fibre at $s \in \Sigma$ is $(\mathcal{P}_{X,E}^k)_s = L_a^k(T_{\pi(s)} X, E_s)$, the k -linear anti-symmetric maps from $T_{\pi(s)} X$

Lagrange-Poincaré field reduction

to E_s . Define the E -valued k -forms on X by

$$\Omega_\pi^k(X; E) := \Gamma\pi_{X, \mathcal{P}_{X,E}^k},$$

where $\pi_{X, \mathcal{P}_{X,E}^k} := \pi \circ \tau^k : \mathcal{P}_{X,E}^k \rightarrow X$. Note that this is not a vector bundle and thus $\Omega_\pi^k(X; E)$ are not the usual vector bundle valued k -forms on X . In fact, $\Omega_\pi^k(X; E)$ is not even a vector space, since addition is not defined between elements of E in different fibres over Σ . In spite of this, there is a derivation, analogous to the usual exterior covariant differentiation (7.4) on forms. While the definition of this operator holds for general elements in $\Omega_\pi^k(X; E)$ and is again based on Cartan's classical formula, only the definition for one-forms is needed:

$$\mathbf{d}^{\tilde{\nabla}}\alpha(U, V) = \tilde{\nabla}_U(\alpha(V)) - \tilde{\nabla}_V(\alpha(U)) - \alpha([U, V]), \quad (7.6)$$

where $\alpha \in \Omega_\pi^1(X; E)$ and $U, V \in \mathfrak{X}(X)$. Note that $\alpha(U), \alpha(V) : X \rightarrow E$, hence (7.5) is valid, also note that $\mathbf{d}^{\tilde{\nabla}}\alpha \in \Omega_\pi^2(X; E)$.

Since $\mathcal{P}_{X,E}^0 = E$ by definition, $\tau^0 = \tau : E \rightarrow \Sigma$ and thus $\Omega_\pi^0(X; E) = \Gamma\pi_{X,E} \subset C^\infty(X, E)$, $\pi_{X,E} := \pi \circ \tau^0$. Therefore the operator $\mathbf{d}^{\tilde{\nabla}}$ on $\Omega_\pi^0(X; E)$ coincides with $\tilde{\nabla}$ as defined in (7.5).

Covariant derivatives

Returning to the case at hand, the general construction specialises to the covariant derivative $\nabla^{\mathcal{A}}$ on the vector bundle $\text{Ad } P \rightarrow \Sigma$. Thus if $\xi \in \Gamma\pi_{X, \text{Ad } P}$ and $u_x \in T_x X$, Definition 7.1.2, gives the $\nabla^{\mathcal{A}}$ -derivative of ξ by

$$\tilde{\nabla}_{u_x}^{\mathcal{A}}\xi(x) := \left. \frac{D^{\mathcal{A}}}{Dt} \right|_{t=0} \xi(c(t)),$$

where $c(t)$ is a curve in X such that $\dot{c}(0) = u_x$. Writing $\xi(x) = \llbracket p(x), \zeta(x) \rrbracket_{\mathfrak{g}}$ yields

$$\tilde{\nabla}_{u_x}^{\mathcal{A}}\xi(x) = \llbracket p(x), \mathbf{d}\zeta(u_x) - [\mathcal{A}(T_x p(u_x)), \zeta(x)] \rrbracket_{\mathfrak{g}} \quad (7.7)$$

(see CENDRA ET AL. [2001], Lemma 2.3.4). Note that the $\nabla^{\mathcal{A}}$ -derivative is a map

$$\tilde{\nabla}^{\mathcal{A}} : \Gamma\pi_{X, \text{Ad } P} \rightarrow \Gamma\pi_{X, \mathcal{P}},$$

not to be confused with (6.5), and it can be interpreted as a map

$$\tilde{\nabla}^A : \mathfrak{X}(X) \times \Gamma\pi_{X, \text{Ad } P} \rightarrow \Gamma\pi_{X, \text{Ad } P}.$$

Also, $\tilde{\nabla}^A \xi$, $\tilde{\nabla}_X^A \xi$, and ξ project to the same section $\rho \in \Gamma\pi_{X, \Sigma}$. That is,

$$\pi_{\Sigma, P} \circ \tilde{\nabla}^A \xi = \pi_{\Sigma, \text{Ad } P} \circ \tilde{\nabla}_X^A \xi = \pi_{\Sigma, \text{Ad } P} \circ \xi = \rho.$$

Next, in the present situation, the covariant exterior derivative

$$d^A := \mathbf{d}^{\tilde{\nabla}^A} : \Omega_{\pi_{X, \Sigma}}^1(X; \text{Ad } P) \rightarrow \Omega_{\pi_{X, \Sigma}}^2(X; \text{Ad } P) \quad (7.8)$$

is attained from (7.6). That is,

$$d^A \xi(u_x, v_x) = \tilde{\nabla}_U^A(\xi(V)) - \tilde{\nabla}_V^A(\xi(U)) - \xi([U, V]), \quad (7.9)$$

where $\xi \in \Omega_{\pi_{X, \Sigma}}^1(X; \text{Ad } P)$, $u_x, v_x \in T_x X$, $U, V \in \mathfrak{X}(X)$ satisfying $U(x) = u_x$, $V(x) = v_x$.

Variations

The wedge product (6.3) in $\Omega^1(\Sigma, \text{Ad } P)$ extends to elements $\alpha, \beta \in \Omega_{\pi_{X, \Sigma}}^1(X, \text{Ad } P)$ that project to the same element $\rho \in \Gamma\pi_{X, \Sigma}$ by

$$(\alpha \wedge \beta)(u_x, v_x) := [\alpha(u_x), \beta(v_x)]_{\rho(x)} - [\alpha(v_x), \beta(u_x)]_{\rho(x)}, \quad (7.10)$$

where $u_x, v_x \in T_x X$.

Proposition 7.1.3 *Let $\sigma : X \rightarrow P$ be a smooth section of $\pi_{X, P} : P \rightarrow X$. Let $\mathcal{A}$ be a connection on the principal bundle $\pi_{\Sigma, P} : P \rightarrow \Sigma$ and $\bar{\sigma} = \llbracket \sigma, \sigma^* \mathcal{A} \rrbracket_{\mathfrak{g}} \in \Gamma\pi_{X, P}$ the reduced section. Then*

$$d^A \bar{\sigma} + \frac{1}{2} \bar{\sigma} \wedge \bar{\sigma} = \rho^* \tilde{\mathcal{B}},$$

where ρ is the section of $\pi_{X, \Sigma} : \Sigma \rightarrow X$ defined by $\rho := \pi_{\Sigma, P} \circ \sigma$ and $\tilde{\mathcal{B}}$ is the $\text{Ad } P$ -valued two-form induced on Σ by the curvature $\mathcal{B}$.

Proof. For $U, V \in \mathfrak{X}(X)$, (7.7) gives

$$\begin{aligned}
 d^A \bar{\sigma}(U, V) &= \tilde{\nabla}_U^A(\bar{\sigma}(V)) - \tilde{\nabla}_V^A(\bar{\sigma}(U)) - \bar{\sigma}([U, V]), \\
 &= \llbracket \sigma, \mathbf{d}(\sigma^* \mathcal{A}(V))U - [\sigma^* \mathcal{A}(U), \sigma^* \mathcal{A}(V)] \rrbracket_{\mathfrak{g}} \\
 &\quad - \llbracket \sigma, \mathbf{d}(\sigma^* \mathcal{A}(U))V - [\sigma^* \mathcal{A}(V), \sigma^* \mathcal{A}(U)] \rrbracket_{\mathfrak{g}} \\
 &\quad - \llbracket \sigma, \sigma^* \mathcal{A}([U, V]) \rrbracket_{\mathfrak{g}} \\
 &= \llbracket \sigma, \mathbf{d}(\sigma^* \mathcal{A})(U, V) - 2[\sigma^* \mathcal{A}(U), \sigma^* \mathcal{A}(V)] \rrbracket_{\mathfrak{g}}, \\
 &= \left[\left[\sigma, \sigma^* \mathbf{d} \mathcal{A} - \frac{1}{2} \sigma^* (\mathcal{A} \wedge \mathcal{A})(U, V) \right] \right]_{\mathfrak{g}} - \left[\left[\sigma, \frac{1}{2} \sigma^* (\mathcal{A} \wedge \mathcal{A})(U, V) \right] \right]_{\mathfrak{g}}, \\
 &= \llbracket \sigma, \sigma^* \mathcal{B}(U, V) \rrbracket_{\mathfrak{g}} - \frac{1}{2}(\bar{\sigma} \wedge \bar{\sigma})(U, V), \\
 &= \rho^* \tilde{\mathcal{B}}(U, V) - \frac{1}{2}(\bar{\sigma} \wedge \bar{\sigma})(U, V).
 \end{aligned}$$

Thus giving the formula

$$d^A \bar{\sigma} + \frac{1}{2} \bar{\sigma} \wedge \bar{\sigma} = \rho^* \tilde{\mathcal{B}}$$

as required. ■

Since $\bar{\sigma} \in \Omega^1(X; \rho^* \text{Ad } P)$, where $\rho := \pi_{\Sigma, P} \circ \sigma$ and $\rho^* \text{Ad } P$ is the pull-back vector bundle over X , $d^A \bar{\sigma}, \bar{\sigma} \wedge \bar{\sigma} \in \Omega^2(X; \rho^* \text{Ad } P)$. Also, the reduced curvature $\tilde{\mathcal{B}}$ belongs to the space $\Omega^1(\Sigma; \text{Ad } P)$, the pullback $\rho^* \tilde{\mathcal{B}} \in \Omega^2(X; \rho^* \text{Ad } P)$. Therefore, the formula $d^A \bar{\sigma} + \bar{\sigma} \wedge \bar{\sigma} = \rho^* \tilde{\mathcal{B}}$ is well-defined as an equality in $\Omega^2(X; \rho^* \text{Ad } P)$.

Corollary 7.1.4 *Let σ_ε be a smooth variation of the smooth section $\sigma : \bar{U} \rightarrow P$, $\mathcal{A}$ a connection on the principal bundle $\pi_{\Sigma, P} : P \rightarrow \Sigma$ and $u_x \in T_x X$. Then*

$$\begin{aligned}
 \left. \frac{D^{\mathcal{A}}}{D\varepsilon} \right|_{\varepsilon=0} \left(\llbracket \sigma_\varepsilon, \sigma_\varepsilon^* \mathcal{A} \rrbracket_{\mathfrak{g}}(u_x) \right) &= \tilde{\nabla}_{u_x}^{\mathcal{A}} \llbracket \sigma, \mathcal{A}(\delta\sigma) \rrbracket_{\mathfrak{g}} + \llbracket \sigma, \sigma^* (\mathbf{i}_{\delta\sigma} \mathcal{B}) \rrbracket_{\mathfrak{g}}(u_x) \\
 &\quad - \llbracket \sigma, [\mathcal{A}(\delta\sigma), \sigma^* \mathcal{A}] \rrbracket_{\mathfrak{g}}(u_x),
 \end{aligned}$$

where $\mathcal{B}$ is the curvature of the connection.

Thus, the infinitesimal variations of $\bar{\sigma}$ are of the form

$$\delta^A \bar{\sigma} = \tilde{\nabla}^A \bar{\eta} - [\bar{\eta}, \bar{\sigma}] + \rho^* (\mathbf{i}_{\delta\rho} \tilde{\mathcal{B}}),$$

where $\rho := \pi_{\Sigma, P} \circ \bar{\sigma}$, $\delta\rho$ is an arbitrary variation of ρ vanishing on ∂U , $\bar{\eta}$ is an arbitrary section in $\Gamma\pi_{X, \text{Ad } P}$ that projects to ρ and vanishes on ∂U , and $\tilde{\mathcal{B}}$ denotes the $\text{Ad } P$ -valued two-form induced on Σ by the curvature $\mathcal{B}$.

Proof. The second formula is a direct consequence of the first. To prove the first, it is sufficient to verify the identity in local bundle charts. However, a global proof based on the previous lemma is preferable.

Extending the bundle geometry in order to explicitly take account of variations achieves the objective. Consider $\hat{P} = \mathbb{R} \times P$ and $\hat{X} = \mathbb{R} \times X$ with the projection $\hat{\pi}_{\hat{X}, \hat{P}}(\varepsilon, p) = (\varepsilon, \pi_{X, P}(p))$. Smooth sections of $\hat{\pi}_{\hat{X}, \hat{P}}$, $\hat{\sigma} : \hat{X} \rightarrow \hat{P}$ are in bijective correspondence with smooth variations of smooth sections of $\pi_{X, P}$, as follows:

$$\hat{\sigma}(\varepsilon, x) = (\varepsilon, \sigma_\varepsilon(x)).$$

Let G act on $\hat{P}$ by extending the action trivially to the $\mathbb{R}$ -factor. Thus, $\hat{\Sigma} : \hat{P}/G = \mathbb{R} \times \Sigma$ and $\text{Ad } \hat{P} = \mathbb{R} \times \text{Ad } P$. Since $\hat{\rho} := \pi_{\hat{X}, \hat{\Sigma}} \circ \hat{\sigma}$, it is clear that $\hat{\rho}(\varepsilon, x) = (\varepsilon, \rho_\varepsilon(x))$. Similarly, the connection $\mathcal{A}$ extends to a connection $\hat{\mathcal{A}} \in \Omega^1(\hat{P}, \mathfrak{g})$ by setting $\hat{\mathcal{A}}(\partial_\varepsilon, u_p) := \mathcal{A}(u_p)$, for any $u_p \in T_p P$.

The section $\hat{\sigma} : \hat{X} \rightarrow \hat{P}$ induces the reduced section $\bar{\sigma}$ (see (7.2) for the general definition) whose explicit expression may be computed as follows: For $(\varepsilon, x) \in \hat{X}$, $u_x \in T_x X$, letting $\delta\sigma_\varepsilon := \frac{d}{d\varepsilon}\sigma_\varepsilon$, and using

$$T\hat{\sigma}(\partial_\varepsilon, u_x) = (\partial_\varepsilon, T\sigma_\varepsilon(u_x) + \delta\sigma_\varepsilon(x)),$$

generates

$$\begin{aligned} \bar{\sigma}(\partial_\varepsilon, u_x) &:= \left[\hat{\sigma}(\varepsilon, x), \hat{\mathcal{A}}(T\hat{\sigma}(\partial_\varepsilon, u_x)) \right]_{\mathfrak{g}} = \left[(\varepsilon, \sigma_\varepsilon(x)), \mathcal{A}(T\sigma_\varepsilon(u_x) + \delta\sigma_\varepsilon(x)) \right]_{\mathfrak{g}} \\ &= \left(\varepsilon, \bar{\sigma}_\varepsilon(u_x) + \left[\sigma_\varepsilon(x), \mathcal{A}(\delta\sigma_\varepsilon(x)) \right]_{\mathfrak{g}} \right). \end{aligned} \quad (7.11)$$

The required formula is attained by evaluating the identity in Proposition 7.1.3,

$$d^{\hat{\mathcal{A}}}\bar{\sigma} + \frac{1}{2}\bar{\sigma} \wedge \bar{\sigma} = \hat{\rho}^*\tilde{\mathcal{B}},$$

Lagrange-Poincaré field reduction

on the pair of vectors $(\partial_\varepsilon, 0), (\partial_\varepsilon, v_x)$ for $v_x \in T_x X$. A direct computation shows that

$$\begin{aligned}\hat{\rho}^* \tilde{\mathcal{B}}((\partial_\varepsilon, 0), (\partial_\varepsilon, v_x)) &= \left(\varepsilon, \llbracket \sigma_\varepsilon(x), \mathcal{B}(\delta\sigma_\varepsilon(x), T\sigma_\varepsilon(v_x)) \rrbracket_{\mathfrak{g}} \right), \\ \frac{1}{2} (\bar{\sigma} \wedge \bar{\sigma})((\partial_\varepsilon, 0), (\partial_\varepsilon, v_x)) &= \left(\varepsilon, \llbracket \sigma_\varepsilon(x), [\mathcal{A}(\delta\sigma_\varepsilon(x)), \mathcal{A}(T\sigma_\varepsilon(v_x))] \rrbracket_{\mathfrak{g}} \right).\end{aligned}$$

To calculate $d^{\hat{\mathcal{A}}\bar{\sigma}}((\partial_\varepsilon, 0), (\partial_\varepsilon, v_x))$, let $V \in \mathfrak{X}(X)$ be such that $V(x) = v_x$ and use (7.9), (7.11), and $[((\partial_\varepsilon, 0), (\partial_\varepsilon, V))] = 0$ to get

$$\begin{aligned}d^{\hat{\mathcal{A}}\bar{\sigma}}((\partial_\varepsilon, 0), (\partial_\varepsilon, v_x)) &= \left(\tilde{\nabla}_{(\partial_\varepsilon, 0)}^{\hat{\mathcal{A}}}(\bar{\sigma}(\partial_\varepsilon, V)) \right)(\varepsilon, x) - \left(\tilde{\nabla}_{(\partial_\varepsilon, v_x)}^{\hat{\mathcal{A}}}(\bar{\sigma}(\partial_\varepsilon, 0)) \right)(\varepsilon, x) \\ &\quad - \bar{\sigma}([((\partial_\varepsilon, 0), (\partial_\varepsilon, V))])(\varepsilon, x) \\ &= \tilde{\nabla}_{(\partial_\varepsilon, 0)}^{\hat{\mathcal{A}}}(\varepsilon, \bar{\sigma}_\varepsilon(V))(\varepsilon, x) - \tilde{\nabla}_{(0, v_x)}^{\hat{\mathcal{A}}} \left(\varepsilon, \llbracket \sigma_\varepsilon(x), \mathcal{A}(\delta\sigma_\varepsilon(x)) \rrbracket_{\mathfrak{g}} \right) \\ &= \frac{D^{\hat{\mathcal{A}}}}{D\varepsilon}(\varepsilon, \bar{\sigma}_\varepsilon(v_x)) - \left(\varepsilon, \tilde{\nabla}_{v_x}^{\mathcal{A}} \llbracket \sigma_\varepsilon(x), \mathcal{A}(\delta\sigma_\varepsilon(x)) \rrbracket_{\mathfrak{g}} \right) \\ &= \left(\varepsilon, \frac{D^{\mathcal{A}}}{D\varepsilon} \bar{\sigma}_\varepsilon(v_x) - \tilde{\nabla}_{v_x}^{\mathcal{A}} \llbracket \sigma_\varepsilon(x), \mathcal{A}(\delta\sigma_\varepsilon(x)) \rrbracket_{\mathfrak{g}} \right).\end{aligned}$$

The last three identities prove the first stated formula. ■

A covariant derivative ∇^Σ on the tangent bundle $\tau_\Sigma : T\Sigma \rightarrow \Sigma$ is needed in order to compute the variation of $T_x \rho$. Given ∇^Σ , (7.5) defines the ∇^Σ -derivative $\tilde{\nabla}^\Sigma$ which acts on maps $X \rightarrow T\Sigma$ and thus (7.6) provides the operator

$$d^\Sigma := \mathbf{d}^{\tilde{\nabla}^\Sigma} : \Omega_{\pi_{X,\Sigma}}^1(X; T\Sigma) \rightarrow \Omega_{\pi_{X,\Sigma}}^2(X; T\Sigma)$$

defined by

$$d^\Sigma \lambda(U, V) = \tilde{\nabla}_U^\Sigma(\lambda(V)) - \tilde{\nabla}_V^\Sigma(\lambda(U)) - \lambda([U, V]), \quad (7.12)$$

where $\lambda \in \Omega_{\pi_{X,\Sigma}}^1(X; T\Sigma)$ and $U, V \in \mathfrak{X}(X)$. In particular, since

$$\Gamma \pi_{X, J^1 P} \subset \Omega_{\pi_{X,\Sigma}}^1(X, T\Sigma),$$

d^Σ operates on sections of the bundle $J^1 \Sigma \rightarrow X$. This differential operator satisfies the following property.

Proposition 7.1.5 *Let $\rho : X \rightarrow \Sigma$ be a smooth section of $\pi_{X,\Sigma}$. Then*

$$d^\Sigma(j^1\rho) = \rho^*T^\Sigma.$$

where $T^\Sigma(U, V) = \nabla_U^\Sigma V - \nabla_V^\Sigma U - [U, V]$ is the torsion tensor of the connection ∇^Σ .

Proof. Recall that the section $j^1\rho : X \rightarrow J^1\Sigma$ is interpreted in this formula in the following way. Given $s \in \Sigma$, let $x := \pi_{X,\Sigma}(s) \in X$ and so $j^1\rho(x) = T_x\rho : T_xX \rightarrow T_{\rho(x)}\Sigma$, that is, one thinks of $j^1\rho$ as an element of $\Omega^1_{\pi_{X,\Sigma}}(X; T\Sigma)$. Given $u_x, v_x \in T_xX$ and having chosen two vector fields $U, V \in \mathfrak{X}(X)$ such that $U(x) = u_x$ and $V(x) = v_x$, (7.12) and (7.5) confer

$$\begin{aligned} d^\Sigma(j^1\rho)(u_x, v_x) &= \tilde{\nabla}_U^\Sigma(T\rho(V)) - \tilde{\nabla}_V^\Sigma(T\rho(U)) - T\rho([U, V]) \\ &= \left. \frac{D^{\nabla^\Sigma}}{Dt} \right|_{t=0} T\rho(V(c_1(t))) - \left. \frac{D^{\nabla^\Sigma}}{Dt} \right|_{t=0} T\rho(U(c_2(t))) - T\rho([U, V]), \end{aligned}$$

where $c_1(t)$ and $c_2(t)$ are curves in X such that $c_1(0) = c_2(0) = x$ and $\dot{c}_1(0) = u_x$, $\dot{c}_2(0) = v_x$. Since ρ is a section of $\pi_{X,\Sigma}$, it is an embedding and the image $\rho(X)$ is a submanifold of Σ . Thus, there exists vector fields $\bar{U}, \bar{V} \in \mathfrak{X}(\Sigma)$ such that $\bar{U}(\rho(x)) = T\rho(U(x))$ and $\bar{V}(\rho(x)) = T\rho(V(x))$. Accordingly,

$$\begin{aligned} \rho^*T^\Sigma(u_x, v_x) &= T^\Sigma(T\rho(U(x)), T\rho(V(x))) = T^\Sigma(\bar{U}, \bar{V})(\rho(x)) \\ &= (\nabla_{\bar{U}}^\Sigma \bar{V} - \nabla_{\bar{V}}^\Sigma \bar{U} - [\bar{U}, \bar{V}])(\rho(x)) \\ &= \left. \frac{D^{\nabla^\Sigma}}{Dt} \right|_{t=0} \bar{V}(\rho(c_1(t))) - \left. \frac{D^{\nabla^\Sigma}}{Dt} \right|_{t=0} \bar{U}(\rho(c_2(t))) - [\bar{U}, \bar{V}](\rho(x)) \\ &= \left. \frac{D^{\nabla^\Sigma}}{Dt} \right|_{t=0} T\rho(V(c_1(t))) - \left. \frac{D^{\nabla^\Sigma}}{Dt} \right|_{t=0} T\rho(U(c_2(t))) - T\rho([U, V]), \end{aligned}$$

which proves the statement. ■

The next result may be obtained using the previous formula by extending the bundle geometry to $\hat{P}$ as is achieved in the proof of Corollary 7.1.4 using Proposition 7.1.3. This time, however, a different proof based on a standard formula for the torsion is provided.

Corollary 7.1.6 *Let ρ_ε be a smooth variation of the section $\rho : \bar{U} \rightarrow \Sigma$, $u_x \in T_x X$, and let ∇^Σ be a covariant derivative on $T\Sigma$. Then*

$$\left. \frac{D^{\nabla^\Sigma}}{D\varepsilon} \right|_{\varepsilon=0} T_x \rho_\varepsilon(u_x) = \tilde{\nabla}_{u_x}^\Sigma \delta \rho + T^\Sigma(\delta \rho(x), T_x \rho(u_x)).$$

where $\tilde{\nabla}^\Sigma$ is the ∇^Σ -derivative and $T^\Sigma(U, V) = \nabla_U^\Sigma V - \nabla_V^\Sigma U - [U, V]$ is the torsion tensor of the connection ∇^Σ .

Proof. This is a direct consequence of the formula

$$\frac{D^{\nabla^\Sigma}}{D\varepsilon} \frac{d}{dt} \alpha(\varepsilon, t) - \frac{D^{\nabla^\Sigma}}{Dt} \frac{d}{d\varepsilon} \alpha(\varepsilon, t) = T^\Sigma \left(\frac{d}{d\varepsilon} \alpha(\varepsilon, t), \frac{d}{dt} \alpha(\varepsilon, t) \right),$$

where $(\varepsilon, t) \in I \times J \mapsto \alpha(\varepsilon, t) \in \Sigma$ is a smooth function. Here it suffices to choose $\alpha(\varepsilon, t) = \rho_\varepsilon(c(t))$, where c is a smooth curve in X such that $\dot{c}(0) = u_x$. ■

For simplicity, a torsion free connection ∇^Σ has been assumed in the subsequent calculations of this discussion. In this case, the formulae from Proposition 7.1.5 and Corollary 7.1.6 simplify to

$$d^\Sigma(j^1 \rho) = 0 \quad \text{and} \quad \left. \frac{D^{\nabla^\Sigma}}{D\varepsilon} \right|_{\varepsilon=0} T_x \rho_\varepsilon(u_x) = \tilde{\nabla}_{u_x}^\Sigma \delta \rho.$$

7.1.2 The Lagrange-Poincaré field equations

Let $l : \Gamma \pi_{X, J^1 \Sigma} \oplus \Omega_{\pi_{X, \Sigma}}^1(X, \text{Ad} P) \rightarrow \mathbb{R}$ be the reduced Lagrangian, as in (7.3). This section computes the Lagrange-Poincaré equations given by the variational principle

$$0 = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U \mathcal{L}(j^1 \sigma_\varepsilon) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U l(j^1 \rho_\varepsilon, \bar{\sigma}_\varepsilon) \mu.$$

An affine connection on the affine bundle $J^1 \Sigma \oplus_\Sigma \mathcal{P} \rightarrow \Sigma$ is required in order to obtain explicit formulae. Since the principal connection $\mathcal{A}$ brings the covariant derivative on $\text{Ad} P \rightarrow \Sigma$ from Definition 6.2.14, it suffices to choose a covariant derivative ∇^Σ on the vector bundle $T\Sigma \rightarrow \Sigma$. This induces a connection on $J^1 \Sigma$

given by

$$\nabla_Z^{J^1\Sigma}\gamma := \text{ver}^\gamma \circ \nabla_Z^L\gamma, \quad (7.13)$$

where ver^γ is the vertical projection associated with the section $\gamma \in \Gamma\pi_{\Sigma, J^1\Sigma}$, interpreted as a connection on $\pi_{X, \Sigma}$, and Z is a vector field on Σ . Here ∇^L denotes the covariant derivative induced from ∇^Σ and from a covariant derivative ∇^X on TX according to

$$(\nabla_\xi^L v)(U)(s) := \nabla_\xi^\Sigma(v(U))(s) - v\left(\nabla_{T\pi_{X, \Sigma}(\xi(s))}^X U\right). \quad (7.14)$$

Here v is a section of the vector bundle $L(\pi_{X, \Sigma}^*TX, T\Sigma) \rightarrow \Sigma$, $\xi \in \mathfrak{X}(\Sigma)$, $U \in \mathfrak{X}(X)$, $s \in \Sigma$, and recall that $v(U)(s) := v(s)(U(\pi_{X, \Sigma}(s)))$. However, the final result only depends on ∇^Σ and not on ∇^X , see JANYŠKA & MODUGNO [1996]. In the current discussion it is also shown that if ∇^Σ is projectable onto a covariant derivative on X , then $\nabla_Z^{J^1\Sigma}\gamma$ is an affine connection.

Thus assuming a projectable covariant derivative ∇^Σ is given, an affine connection on $J^1\Sigma \oplus_\Sigma \mathcal{P}$ is obtained. Given a smooth Lagrangian l on sections of this affine bundle, the fibre derivatives may be defined as

$$\begin{aligned} \left\langle \frac{\delta l}{\delta j^1\rho}, v \right\rangle &:= \left(\frac{d}{d\varepsilon} \Big|_{\varepsilon=0} l(j^1\rho + \varepsilon v, \bar{\sigma}) \right) \\ \left\langle \frac{\delta l}{\delta \bar{\sigma}}, w \right\rangle &:= \left(\frac{d}{d\varepsilon} \Big|_{\varepsilon=0} l(j^1\rho, \bar{\sigma} + \varepsilon w) \right), \end{aligned}$$

where $v \in \Omega_{\pi_{X, \Sigma}}^1(X, V\Sigma)$ and $w \in \Omega_{\pi_{X, \Sigma}}^1(X, \text{Ad}P)$ are arbitrary sections. Note that $\delta l / \delta j^1\rho$ and $\delta l / \delta \bar{\sigma}$ are sections of the bundles $L(\pi_{X, \Sigma}^*T^*X, V^*\Sigma) \rightarrow X$ and $L(\pi_{X, \Sigma}^*T^*X, \text{Ad}^*P) \rightarrow X$; both project to ρ . The derivative with respect to ρ is the horizontal derivative defined at $(j^1\rho, \bar{\sigma})$ by

$$\left\langle \frac{\delta l}{\delta \rho}, u \right\rangle := \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} l\left(\sigma(\varepsilon)_{(j^1\rho, \bar{\sigma})}^h\right), \quad (7.15)$$

where $u \in \mathfrak{X}(\Sigma)$ is a vector field on Σ , $\sigma(\varepsilon)$ is a curve in $\Gamma\pi_{X, \Sigma}$ such that $\dot{\sigma}(0) = u$, and for each x , $\sigma(\varepsilon)_{(j^1\rho, \bar{\sigma})}^h(x)$ is the unique horizontal curve in $J^1\Sigma \oplus \text{Ad}P$ starting at $(j^1\rho(x), \bar{\sigma}(x))$ and projecting to $\sigma(\varepsilon)(x)$. Note that here no attempt is made to make these calculations analytically rigorous, these arguments focus on the for-

Lagrange-Poincaré field reduction

malism arising from the geometry involved. A further analytical investigation of the equations is left to future work.

Consider a variation σ_ε of a given local section $\sigma : \bar{U} \rightarrow P$ and the reduced section $\bar{\sigma}_\varepsilon$. Employing the decomposition of the ε -derivative into its vertical and horizontal parts yields

$$\begin{aligned} \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \int_U l(j^1 \rho_\varepsilon, \bar{\sigma}_\varepsilon) \mu &= \int_U \mathbf{dl} \left(\frac{d}{d\varepsilon} \Big|_{\varepsilon=0} (j^1 \rho_\varepsilon, \bar{\sigma}_\varepsilon) \right) \mu \\ &= \left\langle \frac{\delta l}{\delta \rho}, \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \rho_\varepsilon \right\rangle + \left\langle \frac{\delta l}{\delta j^1 \rho}, \frac{D^{J^1 \Sigma}}{D\varepsilon} \Big|_{\varepsilon=0} j^1 \rho_\varepsilon \right\rangle \\ &\quad + \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \frac{D^L}{D\varepsilon} \Big|_{\varepsilon=0} \bar{\sigma}_\varepsilon \right\rangle, \end{aligned} \quad (7.16)$$

where $D^{J^1 \Sigma}/D\varepsilon$ and $D^L/D\varepsilon$ denote the covariant derivatives associated to the connection $\nabla^{J^1 \Sigma}$ on $J^1 \Sigma \rightarrow \Sigma$ and to the induced covariant derivative on $\mathcal{P} \rightarrow \Sigma$, respectively.

The second term may be computed using the following relation:

$$\left(\frac{D^{J^1 \Sigma}}{D\varepsilon} \Big|_{\varepsilon=0} j^1 \rho_\varepsilon(x) \right) (v_x) = \frac{D^{T\Sigma}}{D\varepsilon} \Big|_{\varepsilon=0} (j^1 \rho_\varepsilon(x)(v_x)). \quad (7.17)$$

This relation is obtained from the definition of the induced covariant derivative $D^L/D\varepsilon$ on $L(\pi_{X,\Sigma}^* TX, T\Sigma)$. Given a curve $\gamma_\varepsilon \in J^1 \Sigma$, (7.14) shows that

$$\left(\frac{D^L}{D\varepsilon} \Big|_{\varepsilon=0} \gamma_\varepsilon \right) (v_x) = \left(\frac{D^{T\Sigma}}{D\varepsilon} \Big|_{\varepsilon=0} (\gamma_\varepsilon \cdot v_\varepsilon) \right) - \gamma_0 \cdot \left(\frac{D^{TX}}{D\varepsilon} \Big|_{\varepsilon=0} v_\varepsilon \right),$$

where $v_\varepsilon \in TX$ is a curve such that $v_\varepsilon \in T_{x_\varepsilon} X$ and $x_\varepsilon \in X$ is such that $\gamma_\varepsilon \in L(T_{x_\varepsilon} X, T_{s_\varepsilon} \Sigma)$. In the present case $\gamma_\varepsilon = j^1 \rho_\varepsilon(x)$ and variations in TX are not considered, so $x_\varepsilon = x$ and $v_\varepsilon = v_x$. Thus

$$\left(\frac{D^L}{D\varepsilon} \Big|_{\varepsilon=0} (j^1 \rho_\varepsilon(x)) \right) (v_x) = \frac{D^{T\Sigma}}{D\varepsilon} \Big|_{\varepsilon=0} (j^1 \rho_\varepsilon(x) \cdot v_x). \quad (7.18)$$

Denoting the connector map of ∇^Σ by $K^{T\Sigma}$ and recalling that ∇^Σ is projectable

allows the following calculation:

$$\begin{aligned}
 T\pi_{X,\Sigma} \left(\left. \frac{D^{T\Sigma}}{D\varepsilon} \right|_{\varepsilon=0} (j^1 \rho_\varepsilon \cdot v_x) \right) &= T\pi_{X,\Sigma} \left(K^{T\Sigma} \left(\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} (j^1 \rho_\varepsilon \cdot v_x) \right) \right) \\
 &= K^{TX} \left(TT\pi_{X,\Sigma} \left(\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} (j^1 \rho_\varepsilon \cdot v_x) \right) \right) = K^{TX} \left(\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} T\pi_{X,\Sigma} (j^1 \rho_\varepsilon \cdot v_x) \right) \\
 &= K^{TX} \left(\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} v_x \right) = 0.
 \end{aligned}$$

This proves that the expression (7.18) is vertical. Thus, by the definition (7.13) of $\nabla^{J^1\Sigma}$, the identity (7.17) is proved.

The third term in equation (7.16) may be evaluated using the equality

$$\left(\left. \frac{D^L}{D\varepsilon} \right|_{\varepsilon=0} \bar{\sigma}_\varepsilon(x) \right) (v_x) = \left. \frac{D^A}{D\varepsilon} \right|_{\varepsilon=0} \bar{\sigma}_\varepsilon(x)(v_x), \quad (7.19)$$

whose proof is similar to that of (7.17).

Using (7.17), (7.19), and Lemmas 7.1.4, 7.1.6, the expression (7.16) may be rewritten

$$\begin{aligned}
 &\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_U l(j^1 \rho_\varepsilon, \bar{\sigma}_\varepsilon) \mu \\
 &= \left\langle \frac{\delta l}{\delta \rho}, \delta \rho \right\rangle + \left\langle \frac{\delta l}{\delta j^1 \rho}, \tilde{\nabla}^\Sigma \delta \rho + T^\Sigma (\delta \rho, T\rho) \right\rangle \\
 &\quad + \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \tilde{\nabla}^A \bar{\eta} - [\bar{\eta}, \bar{\sigma}] + \rho^* (\mathbf{i}_{\delta \rho} \tilde{\mathcal{B}}) \right\rangle \\
 &= \left\langle -\operatorname{div}^A \frac{\delta l}{\delta \bar{\sigma}} + \operatorname{ad}_\sigma^* \frac{\delta l}{\delta \bar{\sigma}}, \bar{\eta} \right\rangle \\
 &\quad + \left\langle \frac{\delta l}{\delta \rho} - \operatorname{div}^\Sigma \frac{\delta l}{\delta j^1 \rho} - \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \mathbf{i}_{T\rho} \tilde{\mathcal{B}} \right\rangle + \left\langle \frac{\delta l}{\delta j^1 \rho}, \mathbf{i}_{T\rho} T^\Sigma \right\rangle, \delta \rho \right\rangle
 \end{aligned}$$

Since $\delta \rho$ and $\bar{\eta}$ are arbitrary, this results in the *vertical* and *horizontal Lagrange-Poincaré equations* given by

$$\operatorname{div}^A \frac{\delta l}{\delta \bar{\sigma}} - \operatorname{ad}_\sigma^* \frac{\delta l}{\delta \bar{\sigma}} = 0 \quad \text{and} \quad \frac{\delta l}{\delta \rho} - \operatorname{div}^\Sigma \frac{\delta l}{\delta j^1 \rho} + \left\langle \frac{\delta l}{\delta j^1 \rho}, \mathbf{i}_{T\rho} T^\Sigma \right\rangle = \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \mathbf{i}_{T\rho} \tilde{\mathcal{B}} \right\rangle, \quad (7.20)$$

respectively, where the second equation has to be considered as an equation in $V_{\rho(x)}^* \Sigma$.

Lagrange-Poincaré field reduction

Remark 7.1.7 The second Lagrange-Poincaré equation is a modified form of the Euler-Lagrange equations on Σ with two additional terms. One term arises from the torsion of ∇^Σ . From now on it is supposed that ∇^Σ is torsion free, as is the case, for example, whenever it is a Levi-Civita connection.

The second modification involves the curvature, $\mathcal{B}$, of the the principal connection used to describe the reduced state space. This term is famous in theoretical physics since it was used by T. Kaluza and O. Klein to attempt unification of gravitational and electro-magnetic forces, KLEIN [1921].

Here, $\text{div}^{\mathcal{A}}$ denotes the divergence associated with $\tilde{\nabla}^{\mathcal{A}}$,

$$\text{div}^{\mathcal{A}} : \Gamma\pi_{X,L}(\pi_{X,\Sigma}^* T^* X, \text{Ad}^* P) \rightarrow \Gamma\pi_{X,\text{Ad}^* P},$$

which is defined as minus the adjoint differential operator to $\tilde{\nabla}^{\mathcal{A}}$:

$$\langle \text{div}^{\mathcal{A}} z, \xi \rangle = - \langle z, \tilde{\nabla}^{\mathcal{A}} \xi \rangle,$$

for all $z \in \Gamma\pi_{X,L}(\pi_{X,\Sigma}^* T^* X, \text{Ad}^* P)$ and $\xi \in \Gamma\pi_{X,\text{Ad}^* P}$ such that $\pi_{\Sigma,L}(\pi_{X,\Sigma}^* T^* X, \text{Ad}^* P) \circ z = \pi_{\Sigma,\text{Ad}^* P} \circ \xi$. In the vertical equation, ad^* denotes the map

$$\text{ad}^* : \Gamma\pi_{X,L}(\pi_{X,\Sigma}^* TX, \text{Ad}^* P) \times \Gamma\pi_{X,L}(\pi_{X,\Sigma}^* T^* X, \text{Ad}^* P) \rightarrow \Gamma\pi_{X,\text{Ad}^* P}, \quad (\bar{\sigma}, \bar{\mu}) \mapsto \text{ad}_{\bar{\sigma}}^* \bar{\mu},$$

well-defined when $\pi_{\Sigma,L}(\pi_{X,\Sigma}^* T^* X, \text{Ad}^* P) \circ \bar{\mu} = \pi_{\Sigma,L}(\pi_{X,\Sigma}^* TX, \text{Ad}^* P) \circ \bar{\sigma}$. Similarly, the operator

$$\text{div}^\Sigma : \Gamma\pi_{X,L}(\pi_{X,\Sigma}^* T^* X, V^* \Sigma) \rightarrow \Gamma\pi_{X,V^* \Sigma} \quad (7.21)$$

is the divergence associated to the ∇^Σ -derivative $\tilde{\nabla}^\Sigma$ restricted to vertical valued sections:

$$\tilde{\nabla}^\Sigma : \Gamma\pi_{X,V\Sigma} \rightarrow \Gamma\pi_{X,L}(\pi_{X,\Sigma}^* TX, V\Sigma).$$

Note that such a restriction is possible since ∇^Σ is projectable. The results obtained above are summarised by the following theorem.

Theorem 7.1.8 (Lagrange-Poincaré reduction theorem) *Let $\pi_{X,P} : P \rightarrow X$ be a locally trivial fibre bundle over an oriented manifold with volume form μ . Let $L :$*

$\Gamma\pi_{X,J^1P} \rightarrow C^\infty(X)$ be a Lagrangian which is invariant under a free and proper left action $\Phi : G \times P \rightarrow P$ such that

$$\pi_{X,P} \circ \Phi_g = \pi_{X,P}, \quad \text{for all } g \in G.$$

Let $\pi_{\Sigma,P} : P \rightarrow \Sigma := P/G$ be the associated principal bundle.

Fix a connection $\mathcal{A}$ on $\pi_{\Sigma,P}$ and let $l : \Gamma\pi_{X,J^1\Sigma} \oplus \Omega_{\pi_{X,\Sigma}}^1(X, \text{Ad } P) \rightarrow C^\infty(X)$ be the reduced Lagrangian induced on the quotient by means of the identification (6.8). Let $\sigma : \bar{U} \rightarrow P$ be a smooth local section of $\pi_{X,P}$, define the reduced local section $\bar{\sigma} \in \Omega_{\pi_{X,\Sigma}}^1(X, \text{Ad } P)$ by

$$\bar{\sigma}(x) = \llbracket \sigma(x), \mathcal{A}(T_x\sigma(-)) \rrbracket_{\mathfrak{g}},$$

and the local section $\rho := \pi_{\Sigma,P} \circ \sigma$ of $\pi_{X,\Sigma}$. Fix a projectable covariant derivative ∇^Σ on $T\Sigma$ and suppose, for simplicity, that ∇^Σ is torsion free. Then the following are equivalent:

i The variational principle

$$\delta \int_U L(j^1\sigma)\mu = 0,$$

holds for arbitrary vertical variations $\delta\sigma$ vanishing on ∂U .

ii The section σ satisfies the covariant Euler-Lagrange equations for $L\mu$.

iii The variational principle

$$\delta \int_X l(j^1\rho, \bar{\sigma})\mu = 0,$$

holds, for variations of the form $\delta^{\mathcal{A}}\bar{\sigma} = \nabla^{\mathcal{A}}\bar{\eta} - [\bar{\eta}, \bar{\sigma}] + \tilde{\mathcal{B}}(\delta\rho, T\rho)$, where $\delta\rho$ is an arbitrary variation of ρ vanishing on ∂U and $\bar{\eta}$ is an arbitrary section of $\pi_{X,\text{Ad } P}$ vanishing on ∂U and such that $\pi_{\Sigma,\text{Ad } P} \circ \bar{\eta} = \rho$.

iv The section $\bar{\sigma}$ satisfies the Lagrange-Poincaré field equations

$$\begin{cases} \frac{\delta l}{\delta \rho} - \text{div}^\Sigma \frac{\delta l}{\delta j^1\rho} = \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \mathbf{i}_{T\rho} \tilde{\mathcal{B}} \right\rangle, \\ \text{div}^{\mathcal{A}} \frac{\delta l}{\delta \bar{\sigma}} - \text{ad}_\sigma^* \frac{\delta l}{\delta \bar{\sigma}} = 0. \end{cases} \quad (7.22)$$

In the case of a connection ∇^Σ with torsion, a term involving the torsion tensor has to be added in the horizontal Lagrange-Poincaré field equations, see (7.20).

Each particular case of the Lagrange-Poincaré field equations is now examined.

- i** If $G = \{1\}$ then $\Sigma = P$, $\rho = \sigma$, $l = L$, and there is no reduction. In this case (7.22) becomes

$$\frac{\delta L}{\delta \sigma} - \operatorname{div}^P \frac{\delta L}{\delta j^1 \sigma} = 0, \quad (7.23)$$

which is just a restatement of the covariant Euler-Lagrange equations, using a projectable and torsion free covariant derivative ∇^P on $TP \rightarrow P$.

- ii** If $\pi_{X,P} : P \rightarrow X$ is a principal bundle and the symmetry group is the structure group, then $\Sigma = X$ and the section ρ is absent since it is the identity on X . Therefore, the reduced variation reads $\delta^{\mathcal{A}} \bar{\sigma} = \nabla^{\mathcal{A}} \bar{\eta} - [\bar{\eta}, \bar{\sigma}]$, where $\bar{\eta}$ is an arbitrary section of $\pi_{X, \operatorname{Ad} P}$, and the Lagrange-Poincaré field equations (7.22) read

$$\operatorname{div}^{\mathcal{A}} \frac{\delta l}{\delta \bar{\sigma}} - \operatorname{ad}_{\bar{\sigma}}^* \frac{\delta l}{\delta \bar{\sigma}} = 0.$$

Thus the covariant Euler-Poincaré equations are recovered; see Theorem 3.1 of CASTRILLÓN-LÓPEZ ET AL. [2000].

- iii** If $\pi_{X,P} : P \rightarrow X$ is a principal bundle whose structure group contains the symmetry group as a subgroup, the equations (7.22) coincide with the equations (4.11) obtained in CASTRILLÓN-LÓPEZ & RATIU [2003].
- iv** If $P = \mathbb{R} \times Q$ where $\pi_{M,Q} : Q \rightarrow M$ is a G -principal bundle then $\Sigma = \mathbb{R} \times M$. The sections $\sigma \in \Gamma \pi_{\mathbb{R},P}$ and $\rho \in \Gamma \pi_{\mathbb{R},\Sigma}$ read $\sigma(t) = (t, q(t))$ and $\rho(t) = (t, m(t))$, where $m(t) = \pi_{M,Q}(q(t)) \in M$. The first jet extensions $j^1 \sigma$ and $j^1 \rho$ are identified with $(t, \dot{q}(t))$ and $(t, \dot{m}(t))$.

In practice the connection $\mathcal{A}$ on P is chosen to be induced by a connection γ on Q . Therefore, the reduced section $\bar{\sigma}$ is identified with $(t, \bar{v}(t))$,

where $\bar{v}(t) = [q(t), \gamma(\dot{q}(t))]_G$. Similarly, a section $\xi \in \Gamma\pi_{\mathbb{R}, \text{Ad } P}$ covering $\rho(t) = (t, m(t))$ reads $\xi(t) = (t, \bar{\xi}(t))$, where $\bar{\xi}(t) \in \text{Ad}_{m(t)} Q$. The ∇^A -derivative $\tilde{\nabla}^A$ of ξ can thus be identified with the covariant time derivative $(D^\gamma/Dt) \bar{\xi}(t)$. Using all these observations, the second equation of (7.22) reads

$$\frac{D^\gamma}{Dt} \frac{\delta l}{\delta \bar{v}} - \text{ad}_{\bar{v}}^* \frac{\delta l}{\delta \bar{v}} = 0$$

and the variation of $\bar{v}$ is $\delta^\gamma \bar{v} = \frac{D^\gamma}{dt} \bar{\eta} - [\bar{\eta}, \bar{v}] + \tilde{B}(\delta m, \dot{m})$, where $\tilde{B}$ is the reduced curvature of γ . Recall that writing the horizontal equation requires a projectable covariant derivative ∇^Σ on $T\Sigma$, which is also assumed to be torsion free for simplicity. In the classical case, the covariant derivative is constructed from a torsion free covariant derivative ∇ on TM and the natural covariant derivative on $T\mathbb{R}$. In this case, ∇^Σ is obviously projectable and torsion free. The first equation of (7.22) reads

$$\frac{\delta l}{\delta m} - \frac{D^\nabla}{Dt} \frac{\delta l}{\delta \dot{m}} = \left\langle \frac{\delta l}{\delta \bar{v}}, \mathbf{i}_{\dot{m}} \tilde{B} \right\rangle.$$

Thus the classical Lagrange-Poincaré equations are recovered; see Theorem 3.4.1 in CENDRA ET AL. [2001]. Note that here the Lagrangian is allowed to be time-dependent.

If ∇ has a torsion T^∇ , the horizontal equation reads

$$\frac{\delta l}{\delta m} - \frac{D^\nabla}{Dt} \frac{\delta l}{\delta \dot{m}} + \left\langle \frac{\delta l}{\delta \dot{m}}, \mathbf{i}_{\dot{m}} T^\nabla \right\rangle = \left\langle \frac{\delta l}{\delta \bar{v}}, \mathbf{i}_{\dot{m}} \tilde{B} \right\rangle,$$

see (7.20).

In the particular case, $G = \{1\}$, there is no reduction and the vertical equation is absent. In this case the horizontal equation reads

$$\frac{D^\nabla}{Dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = 0. \quad (7.24)$$

Of course, (7.24) recovers the standard Euler-Lagrange equation written with the help of a connection. In the case when the connection has torsion,

(7.24) reads

$$\frac{D^\nabla}{Dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = \left\langle \frac{\delta l}{\delta \dot{q}}, \mathbf{i}_{\dot{q}} T^\nabla \right\rangle;$$

see (3) in GAMBOA & SOLOMIN [2003]. Recall that the usual way to write the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

makes sense only locally; see (7.1).

Another particular case arises when $Q = G$. In this case, there is no horizontal equation and the vertical equation gives the Euler-Poincaré equation. Indeed, all the connections are equivalent (the bundle $Q \rightarrow Q/G$ is over a point) and the covariant time derivative on the adjoint bundle becomes the ordinary time derivative on the Lie algebra $\mathfrak{g}$. These observations and (7.22) recover the Euler-Poincaré equation together with the associated constrained variations

$$\frac{d}{dt} \frac{\delta l}{\delta v} - \text{ad}_v^* \frac{\delta l}{\delta v} = 0 \quad \text{and} \quad \delta v = \frac{d}{dt} \eta - [\eta, v].$$

Theorem 7.1.8 and the subsequent particular cases demonstrate that the objective of deriving a reduction theory that satisfies the criteria set out in §6.1 has now been accomplished. For completeness, a corresponding reconstruction theory must be developed. This objective is accomplished in §7.1.3.

7.1.3 Reconstruction

Having derived the Lagrange-Poincaré field equations the next stage it is reconstruct a solution section $\sigma : \bar{U} \rightarrow P$ of the original Euler-Lagrange equations from a solution section $\bar{\sigma} : \bar{U} \rightarrow L(\pi_{X,\Sigma}^* TX, \text{Ad } P)$ of the Lagrange-Poincaré equations. Note that Theorem 7.1.8 does not consider this problem, since the section σ is given a priori. §7.1.3 deals with the reconstruction problem and demonstrates that reconstruction of field theories requires an extra integrability condition.

Induced connection

A section $\bar{\sigma} \in \Omega_{\pi_{X,\Sigma}}^1(X, \text{Ad } P)$ induces a G -principal bundle $P^\rho \rightarrow X$ and a connection $\mathcal{A}^{\bar{\sigma}}$ on it as follows: The subset $P^\rho \subset P$ is defined by

$$P^\rho := \pi_{\Sigma,P}^{-1}(\rho(X)) = \{p \in P \mid \pi_{\Sigma,P}(p) \in \rho(X)\},$$

where $\rho := \pi_{\Sigma,L(\pi_{X,\Sigma}^* TX, \text{Ad } P)} \circ \bar{\sigma} \in \Gamma \pi_{X,\Sigma}$. Since ρ is a section, it is an injective immersion and a homeomorphism onto its image. Thus, the image $\rho(X)$ is a submanifold of Σ . Now, since $\pi_{\Sigma,P}$ is a submersion, it is transversal to the submanifold $\rho(X)$. This proves that P^ρ is a submanifold of P , whose tangent space at p is

$$T_p P^\rho = (T_p \pi_{\Sigma,P})^{-1}(T_x \rho(T_x X)), \quad x = \pi_{X,P}(p).$$

The manifold P^ρ may be endowed with the structure of a principal G -bundle over X by restriction of the G -action on P . Note that P^ρ can be identified with the pull-back bundle $\rho^* P = \{(x, p) \mid \pi_{\Sigma,P}(p) = \rho(x)\}$, the identification being given by

$$p \in P^\rho \mapsto (\pi_{X,P}(p), p) \in \rho^* P.$$

The section $\bar{\sigma}$ may be regarded as a section of the vector bundle $L(TX, \rho^* \text{Ad } P) \simeq L(TX, \text{Ad } P^\rho)$, and thus induces an equivariant, vertical one-form $\omega^{\bar{\sigma}} \in \overline{\Omega^1}(P^\rho, \mathfrak{g})$. This vertical one-form constitutes the affine term that relates the principal connection $\mathcal{A}^\rho$, chosen arbitrarily to perform reduction, with a new principal connection, $\mathcal{A}^{\bar{\sigma}}$, which is the correct choice of connection for reconstruction. The isomorphism $\overline{\Omega^1}(P^\rho, \mathfrak{g}) \leftrightarrow L(TX, \text{Ad } P^\rho)$ is written explicitly as follows.

$$\omega^{\bar{\sigma}} \in \overline{\Omega^1}(P^\rho, \mathfrak{g}) \mapsto \bar{\sigma} \in L(TX, \text{Ad } P^\rho), \quad \bar{\sigma}(u_x) := \llbracket p, \omega^{\bar{\sigma}}(u_p) \rrbracket_{\mathfrak{g}}, \quad (7.25)$$

where $u_x = T\pi_{X,P}(u_p) \in T_x X$, $u_p \in T_p P$. The connection $\mathcal{A}$ on $\pi_{\Sigma,P} : P \rightarrow \Sigma$ naturally induces a connection $\mathcal{A}^\rho$ on $P^\rho \rightarrow X$. A new connection $\mathcal{A}^{\bar{\sigma}}$ is thereby obtained on $P^\rho \rightarrow X$. Concretely,

$$\mathcal{A}^{\bar{\sigma}} := \mathcal{A}^\rho - \omega^{\bar{\sigma}}.$$

Lagrange-Poincaré field reduction

Thus the vertical solution of the Lagrange-Poincaré field equations may be interpreted as describing an affine modification to the a priori connection $\mathcal{A}^\rho$. The modified connection $\mathcal{A}^{\bar{\sigma}}$ is the correct choice of connection for reconstruction, as is explained below.

Reconstruction condition

This paragraph proves that if $\bar{\sigma}$ is the reduced section associated to a section $\sigma \in \Gamma\pi_{X,P}$ then $\mathcal{A}^{\bar{\sigma}}$ is flat. Indeed, in this case $P^\rho = \{\Phi_g(\sigma(x)) \mid g \in G, x \in X\}$ and for $p = \sigma(x) \in P^\rho$ and $v_p \in T_p P^\rho$ formula (7.25) gives

$$\mathcal{A}^{\bar{\sigma}}(v_p) = \mathcal{A}^\rho(v_p) - \omega^{\bar{\sigma}}(v_p) = \mathcal{A}(v_p) - \mathcal{A}(T_x \sigma(T_p \pi_{X,P}(v_p))) \quad (7.26)$$

since $\bar{\sigma}(x) = \llbracket \sigma(x), \sigma^* \mathcal{A}(x) \rrbracket_{\mathfrak{g}}$ for all $x \in X$. Recall that $u_p \in T_p P^\rho$ if and only if $T_p \pi_{\Sigma,P}(u_p) \in T_x \rho(T_x X)$. That is, in terms of σ ,

$$T_p \pi_{\Sigma,P}(u_p) \in T_{\sigma(x)} \pi_{\Sigma,P}(T_x \sigma(T_x X)).$$

This proves that

$$T_{\sigma(x)} P^\rho = T_x \sigma(T_x X) + V_{\sigma(x)} P$$

at $p = \sigma(x)$, where $V_{\sigma(x)} P = \ker(T_{\sigma(x)} \pi_{\Sigma,P})$ is the vertical space relative to $\pi_{\Sigma,P}$. Thus, for $p = \sigma(x)$, any $v_p \in T_p P^\rho$ reads $v_p = T_x \sigma(v_x) + \xi_P(p)$. Inserting this expression for v_p into (7.26), reveals the condition $\mathcal{A}^{\bar{\sigma}}(T_x \sigma(v_x)) = 0$, which proves that the $\mathcal{A}^{\bar{\sigma}}$ -horizontal subspace at $\sigma(x)$ is given by

$$H_{\sigma(x)}^{\mathcal{A}^{\bar{\sigma}}} P = T_x \sigma(T_x X).$$

This horizontal distribution is integrable, the integral leaves being $\{\Phi_g(\sigma(x)) \mid x \in X\} = \Phi_g(\text{Im}(\sigma))$, for each $g \in G$. Thus, the connection $\mathcal{A}^{\bar{\sigma}}$ on P^ρ is flat and the **horizontality condition**

$$\sigma^* \mathcal{A}^{\bar{\sigma}} = 0 \quad (7.27)$$

is a necessary condition for reconstruction.

Conversely, consider a section $\bar{\sigma}$ of $\pi_{X,L(TX, \text{Ad } P)}$ such that the connection $\mathcal{A}^{\bar{\sigma}}$ on P^ρ is flat and has trivial holonomy. Since the connection $\mathcal{A}^{\bar{\sigma}}$ is flat, the horizontal

distribution is integrable and the leaves cover the base, that is, given a leaf L , each fibre intersects the leaf L at least once. Since the holonomy is trivial, each fibre intersects the leaf exactly once. This construction shows that each integral leaf of the horizontal distribution defines a section of the bundle $P^\rho \rightarrow X$. Thus a family of sections $\Phi_g \circ \sigma$ of $\pi_{X,P}$ that project via $\pi_{\Sigma,P}$ to ρ is attained. Since

$$\llbracket \sigma, \sigma^* \mathcal{A} \rrbracket_{\mathfrak{g}} = \llbracket \sigma, \sigma^* \mathcal{A}^{\bar{\sigma}} + \omega^{\bar{\sigma}} \rrbracket_{\mathfrak{g}} = \llbracket \sigma, \omega^{\bar{\sigma}} \rrbracket_{\mathfrak{g}} = \bar{\sigma},$$

the section $\bar{\sigma}$ is the reduced section associated to the family of sections $\Phi_g \circ \sigma$ for each $g \in G$. The horizontality condition (7.27) is, of course, satisfied.

Recall that the flatness of the connection does not imply that the holonomy is trivial unless the base is simply connected or the holonomy group is connected. Note that this fact implies that the holonomy of a flat connection is locally trivial, that is, for every $x \in X$, there exists an open neighbourhood U such that the holonomy of $P|_U$ is trivial.

The situation is summarised in the following reconstruction theorem.

Theorem 7.1.9 (Lagrange-Poincaré reconstruction theorem) *Choose a fixed connection $\mathcal{A}$ on the principal bundle $\pi_{\Sigma,P} : P \rightarrow \Sigma$ and consider a G -invariant Lagrangian L and the reduced Lagrangian l .*

If $\sigma : \bar{U} \rightarrow P$ is a solution of the Euler-Lagrange field equations, then the reduced section $\bar{\sigma}$ is a solution of the Lagrange-Poincaré field equations. Moreover the connection $\mathcal{A}^{\bar{\sigma}}$ on P^ρ is flat and the horizontality condition (7.27) holds.

Conversely, given a solution $\bar{\sigma}$ of the Lagrange-Poincaré equations on $\bar{U}$ such that $\mathcal{A}^{\bar{\sigma}}$ is flat and has trivial holonomy over an open set containing $\bar{U}$, the family $\Phi_g \circ \sigma$, $g \in G$, of solutions of the Euler-Lagrange field equations are given by the integral leaves of the horizontal distribution associated to $\mathcal{A}^{\bar{\sigma}}$. In addition, the horizontality condition (7.27) holds. If the connection $\mathcal{A}^{\bar{\sigma}}$ is flat it is always possible to restrict it to an open simply connected set contained in U so that its holonomy on U is automatically zero.

Note that the curvature of $\mathcal{A}^{\bar{\sigma}}$ is $\mathcal{B} - \mathbf{d}^{\mathcal{A}} \omega^{\bar{\sigma}} - (1/2) \omega^{\bar{\sigma}} \wedge \omega^{\bar{\sigma}}$. Therefore, the *recon-*

Lagrange-Poincaré field reduction

struction condition is

$$\mathcal{B} - \mathbf{d}^A \omega^{\bar{\sigma}} - \frac{1}{2} \omega^{\bar{\sigma}} \wedge \omega^{\bar{\sigma}} = 0 \quad \text{on} \quad P^\rho. \quad (7.28)$$

Equation (7.28) has to be seen as an equality in the space $\overline{\Omega^2}(P^\rho, \mathfrak{g})$ of equivariant vertical two-forms. The isomorphism (7.25) shows it is equivalent to assert that the corresponding two-form in $\Omega^2(X, \text{Ad } P^\rho) = \Omega^2(X, \rho^* \text{Ad } P)$ vanishes. Applying (7.25) to equation (7.28) recovers the formula

$$d^A \bar{\sigma} + \frac{1}{2} \bar{\sigma} \wedge \bar{\sigma} = \rho^* \tilde{\mathcal{B}}, \quad (7.29)$$

which is familiar from Proposition 7.1.3.

Reconstruction equation

When reconstructing solutions of the Euler-Lagrange field equations it is necessary to add (7.29) to the reduced field equations (7.22) since there could be solutions to the Lagrange-Poincaré field equations (7.22) that do not correspond to the original Euler-Lagrange system. An example of this is (2.17), the second equation of (4.25), or (4.38) all of which refer to the same kinematic condition for the molecular strand. This equation is referred to in the Hamilton-Pontryagin derivation in §2.1.2 as a kinematic relation, since it holds without reference to the Lagrangian. In §3.3.2 the same relation appears as the first of (3.30), where it is interpreted as an advection relation from the classical perspective. It is now clear that the co-variant perspective interprets (2.17) as an additional reconstruction/integrability condition to the Lagrange-Poincaré equations.

Given a solution $(\rho, \bar{\sigma})$ that satisfies both the Lagrange-Poincaré equations and the reconstruction equation (7.29), (7.25) uniquely determines $\omega^{\bar{\sigma}}$ by the formula

$$\bar{\sigma}(u_x) = \llbracket p, \omega^{\bar{\sigma}}(\text{Hor}_p^A(T_x \rho(u_x))) \rrbracket_{\mathfrak{g}}, \quad p \in \pi_{\Sigma, P}^{-1}(\rho(x)), \quad u_x \in T_x X, \quad (7.30)$$

since $T_p \pi_{X, P}(\text{Hor}_p^A(T_x \rho(u_x))) = u_x$. Thus, $\omega^{\bar{\sigma}}$ is completely determined in terms of $(\rho, \bar{\sigma})$.

For a section $\sigma \in \Gamma\pi_{X,P^\rho}$, the horizontality condition (7.27) for $\mathcal{A}^{\bar{\sigma}}$ is

$$0 = \sigma^* \mathcal{A}^{\bar{\sigma}} = \sigma^* \mathcal{A}^\rho - \sigma^* \omega^{\bar{\sigma}} = \sigma^* \mathcal{A}^\rho - \rho^* (\text{Hor}_\sigma^{\mathcal{A}})^* \omega^{\bar{\sigma}} \quad (7.31)$$

because $\omega^{\bar{\sigma}}(T_x \sigma(u_x)) = \omega^{\bar{\sigma}}(\text{Hor}_{\sigma(x)}^{\mathcal{A}}(T_x \rho(u_x)))$ since $\omega^{\bar{\sigma}} \in \overline{\Omega^1}(P^\rho, \mathfrak{g})$. Note that following the determination of $\omega^{\bar{\sigma}}$ by (7.30) the only unknown quantity in (7.31) is σ .

Now, (7.31) gives a first order PDE that determines σ as follows: If $u_x \in T_x X$, then by (7.31) and the horizontal-vertical decomposition relative to the connection $\mathcal{A}^\rho$,

$$\begin{aligned} T_x \sigma(u_x) &= \text{Hor}_{\sigma(x)}^{\mathcal{A}^\rho}(T_{\sigma(x)} \pi_{\Sigma,P}(T_x \sigma(u_x))) + (\mathcal{A}^\rho(T_x \sigma(u_x)))_P(\sigma(x)) \\ &= \text{Hor}_{\sigma(x)}^{\mathcal{A}^\rho}(T_x \rho(u_x)) + (\sigma^* \mathcal{A}^\rho(u_x))_P(\sigma(x)) \\ &= \text{Hor}_{\sigma(x)}^{\mathcal{A}^\rho}(T_x \rho(u_x)) + \left(\rho^* (\text{Hor}_\sigma^{\mathcal{A}^\rho})^* \omega^{\bar{\sigma}} \right)_P(\sigma(x)). \end{aligned}$$

This gives the following first order reconstruction PDE for σ :

$$T_x \sigma = \text{Hor}_{\sigma(x)}^{\mathcal{A}} \circ T_x \rho + (\omega^{\bar{\sigma}} \circ \text{Hor}_{\sigma(x)}^{\mathcal{A}^\rho} \circ T_x \rho)_P(\sigma(x)). \quad (7.32)$$

Theorem 7.1.9 may now be interpreted as asserting that given a solution $(\rho, \bar{\sigma})$ of equations (7.22) and (7.29), *there exists a unique solution σ to the reconstruction equation (7.32) in a neighbourhood where $\mathcal{A}^{\bar{\sigma}}$ has trivial holonomy.* This section σ solves the corresponding Euler-Lagrange equations for the unreduced problem.

As a final comment, note that (7.32) is the field theoretic analogue of the classical reconstruction equation $\xi = g^{-1} \dot{g}$ associated to the Euler-Poincaré equations.

Particular cases

The reconstruction condition specialises to the particular cases as follows:

- i If $G = \{1\}$ there is no reduction and, therefore, no reconstruction condition.
- ii In this case the variable ρ is absent, so $P^\rho = P$. Moreover, the reduced

section $\bar{\sigma}$ turns out to be associated, via the map $\beta_{\mathcal{A}}$, to a section ζ of $J^1P/G \rightarrow X$, that can be interpreted as a connection on P . This connection ζ does not depend on the chosen $\mathcal{A}$ and $\mathcal{A}^{\bar{\sigma}}$ turns out to be the connection one-form associated to ζ . The reconstruction condition is simply that the curvature of this connection (or of ζ) is zero. This recovers the reconstruction condition that in the case of covariant Euler-Poincaré reduction; see §3.2 of CASTRILLÓN-LÓPEZ ET AL. [2000].

- iii The reconstruction condition is the same as in CASTRILLÓN-LÓPEZ & RATIU [2003].
- iv Since $X = \mathbb{R}$, the base is one-dimensional and every connection is flat. Since $\mathbb{R}$ is simply connected the holonomy is trivial. The reconstruction condition is always satisfied, agreeing with the fact that in classical Lagrangian reduction, the solution of the Euler-Lagrange equations can always be constructed from that of the reduced equations.

7.2 Summary

In §7 we have developed the covariant Lagrange-Poincaré theory. This was achieved by carefully constructing appropriate derivatives in the geometric setting inherited from §6. The reduced variations were calculated, and then used to derive the reduced variational principle. This work culminated in Theorem 7.1.8.

Next, §7.1.3 established the reconstruction formula and integrability condition for the covariant Lagrange-Poincaré setting. The Reconstruction Theorem 7.1.9 gave the main result, while (7.32) gave an explicit formula that generalised the classical reconstruction equation.

Throughout the development of the theory, the particular cases have demonstrated the compatibility of covariant Lagrange-Poincaré setting with both the classical and covariant perspectives. This constant comparison makes

it clear that Theorems 7.1.8 and 7.1.9 satisfy the criteria set out in §6.1.

With the main objective of Part II already completed, §8 rounds off the discussion by considering the local representations and conservation laws of covariant Lagrange-Poincaré dynamics.

CHAPTER 8 : CONSERVATION LAWS AND REPRESENTATIONS

In applications there is often a natural choice of gauge that is used to formulate the Lagrange-Poincaré field equations in a convenient local form.

§8 describes the two predominant choices of representations that occur, the *spatial* and *convective representations*. The Lagrange-Poincaré equations (7.22) are given locally using these choices of gauge. The spatial representation yields Noether's Theorem as the vertical equation whilst the convective representation has the Euler-Poincaré equation as its vertical equation. This observation shows that the Lagrange-Poincaré equations are equivalent to Noether's Theorem, a statement often found in the literature when dealing with concrete applications.

Finally, a global version of the Kelvin-Noether Theorem is formulated that generalises the result for classical systems given, for example, in CENDRA ET AL. [1997]; HOLM ET AL. [1998]. The Kelvin Noether theorem is then applied to the molecular strand problem from Part I.

8.1 Representations and Noether's Theorem

A section $\sigma \in \Gamma\pi_{X,P}$ introduces a representation of $\text{Ad}^* P$ which, in turn, yields local equations for the vertical part of (7.22). The two natural choices of section and their associated representations are described below.

8.1.1 Convective representation

Suppose a local solution $\sigma : U \subset X \rightarrow P$ of (7.22) and (7.32) is sought in a trivialisation of P over $U \subset X$. Let $\rho := \pi_{\Sigma,P} \circ \sigma$ and $V := \rho(U) \subset \Sigma$. Suppose further that a flat connection $\mathcal{A}$ exists on $P \rightarrow V$. Then, there exists a unique section $\gamma : V \rightarrow P$ such that $T_s\gamma(v_s) \in H_sP$, for all $s \in \Sigma$ and $v_s \in T_s\Sigma$. Therefore, the section $\sigma^h := \gamma \circ \rho \in \Gamma\pi_{U,P}$ has the property that $T_x\sigma^h(v_x) \in H_{\sigma^h(x)}P$ for all $v_x \in TX|_U$. Such a section is called a *horizontal section*.

Remark 8.1.1 It may not be possible to find a flat connection $\mathcal{A}$ on an arbitrary open set $V \subset \Sigma$. The convective representation is not defined in such cases. In applications one may find that shrinking the set $U \subset X$ yields a suitable $V \subset \Sigma$ such that the convective representation makes sense. In classical Lagrangian reduction when $X = \mathbb{R}$ it is always possible to construct a local horizontal section. Therefore the convective representation is always well-defined for classical systems. Note, however, that when the reconstruction equation is satisfied then a flat connection can be constructed. Therefore the convective representation is always well-defined when there exists a solution to the Lagrange-Poincaré equations (7.22) together with the reconstruction condition (7.29).

Under these circumstances, the problem of finding a local solution $\sigma : U \rightarrow P$ reduces to finding a map $g : U \subset X \rightarrow G$ defined according to $\sigma(x) = \Phi_{g(x)}\sigma^h(x)$ for all $x \in U$. Recalling (6.4), g must satisfy

$$\chi_{\sigma^h}(\bar{\sigma}) = \chi_{\sigma^h} \left(\llbracket \sigma, \sigma^* \mathcal{A} \rrbracket_{\mathfrak{g}} \right) = g^{-1} dg =: \xi \in \Omega^1(X, \mathfrak{g}). \quad (8.1)$$

Consequently, the vertical part of equations (7.22) composed with χ_{σ^h} yields

$$\chi_{\sigma^h} \left(\operatorname{div}^{\mathcal{A}} \frac{\delta l}{\delta \bar{\sigma}} - \operatorname{ad}_{\bar{\sigma}}^* \frac{\delta l}{\delta \bar{\sigma}} \right) = \chi_{\sigma^h} \left(\left\llbracket \sigma^h, \operatorname{div} \frac{\delta l}{\delta \xi} - \operatorname{ad}_{\xi}^* \frac{\delta l}{\delta \xi} \right\rrbracket_{\mathfrak{g}^*} \right) = \operatorname{div} \frac{\delta l}{\delta \xi} - \operatorname{ad}_{\xi}^* \frac{\delta l}{\delta \xi}.$$

Thus, the local representation of the vertical Lagrange-Poincaré equation in this gauge is

$$\operatorname{div} \frac{\delta l}{\delta \xi} - \operatorname{ad}_{\xi}^* \frac{\delta l}{\delta \xi} = 0 \quad (8.2)$$

which recovers the Euler-Poincaré equation. This choice of gauge is called the *convective representation*, see CENDRA ET AL. [1997].

Remark 8.1.2 In §4.2 and §4.3 the molecular strand is described from the covariant perspective. For the molecular strand, the bundles involved are trivial and the Maurer-Cartan form provides a flat connection. As a result it is possible to seek a global solution section $\sigma : X \rightarrow P$ in the convective representation. The relation (8.1) appears as (4.31) and (4.36). The vertical covariant Lagrange-Poincaré equation in convective representation (8.2) appears as (4.33), (4.34) and as the first equation in (4.37).

8.1.2 Spatial representation

With the same notation as for the convective representation, $\sigma = g\sigma^h$, the map χ_σ applied to $\bar{\sigma}$ yields

$$\chi_\sigma(\bar{\sigma}) = \chi_\sigma\left(\llbracket \sigma, \sigma^* \mathcal{A} \rrbracket_{\mathfrak{g}}\right) = dg g^{-1} =: \Xi \in \Omega^1(X, \mathfrak{g}).$$

Accordingly, the vertical part of equations (7.22) composed with χ_σ reads

$$\chi_\sigma\left(\operatorname{div}^{\mathcal{A}} \frac{\delta l}{\delta \bar{\sigma}} - \operatorname{ad}_{\bar{\sigma}}^* \frac{\delta l}{\delta \bar{\sigma}}\right) = \chi_\sigma\left(\llbracket \sigma, \operatorname{div} \frac{\delta l}{\delta \Xi} \rrbracket_{\mathfrak{g}^*}\right) = \operatorname{div} \frac{\delta l}{\delta \Xi}.$$

Thus, the local representation of the vertical Lagrange-Poincaré equation (7.22) in this gauge is

$$\operatorname{div} \frac{\delta l}{\delta \Xi} = 0, \tag{8.3}$$

which is Noether's Theorem. This choice of gauge is called the *spatial representation*.

Note that since $\operatorname{Ad}_g \xi = \Xi$, each of the Euler-Poincaré equations and Noether's Theorem are local representations of equations (7.22) corresponding to a particular choice of gauge. In particular, the Euler-Poincaré equation is equivalent to Noether's Theorem.

Remark 8.1.3 When the convective representation cannot be defined it is still possible to fix $\Xi = \sigma^* \mathcal{A}$ and proceed with the construction of the spatial representation without the use of σ^h . Thus the spatial representation of the Lagrange-Poincaré equations is well-defined even when there is no solution to the reconstruction equation, while the convective representation is not. In practice this observation is just an academic aside, since the reconstruction equation is always asserted in applications.

Remark 8.1.4 The name 'convective', which derives from Latin (literally 'inclined to carry with'), alludes to the conception of the convective representation, in physical application, as moving with the dynamics. This contrasts with the name 'spatial' since the spatial representation often arises in physical problems described with respect to an inertial frame of reference that is 'fixed in space'.

Remark 8.1.5 The vertical covariant Lagrange-Poincaré equation in spatial representation (8.3) appears for the molecular strand as the first equation in (2.39), after (2.40) has been taken into account. Indeed, if l is independent of ρ , then $\delta l / \delta \rho = 0$ and (2.40) shows that the right hand side of (2.39) vanishes. In that case, the whole of (2.39) may be regarded as an example of (8.3). This observation is related to the symmetry breaking properties that cause the differences between §4.2 and §4.3. The whole content of §2.3 maybe therefore be summarised as changing from the convective representation to the spatial representation of the vertical part of (7.22).

8.2 The Kelvin-Noether theorem

This paragraph extends the Kelvin-Noether Theorem, as in CENDRA ET AL. [1997]; HOLM ET AL. [1998], in two ways, from the classical to the covariant setting, and from the Euler-Poincaré to the Lagrange-Poincaré setting.

Definition 8.2.1 *Given any manifold $\mathcal{C}$ on which G acts, the **associated bundle** is a fibre bundle over Σ defined by*

$$\mathcal{H}_{\mathcal{C}} := P \times_G \mathcal{C} = (P \times \mathcal{C}) / G,$$

where the action of G on $P \times \mathcal{C}$ is the diagonal action.

Note that the adjoint and coadjoint bundles, $\text{Ad } P$ and $\text{Ad}^* P$ are associated bundles with $\mathcal{C} = \mathfrak{g}$ and $\mathcal{C} = \mathfrak{g}^*$ respectively. The action of G on $\mathfrak{g}$ for $\text{Ad } P$ is the adjoint action whilst the action on G on $\mathfrak{g}^*$ is the coadjoint action. The equivalence class of $(p, c) \in P \times \mathcal{C}$ is denoted

$$[p, c]_{\mathcal{C}} \in P \times_G \mathcal{C}.$$

The lifted action of G on $T\mathcal{C}$ enables the definition of $\mathcal{H}_{T\mathcal{C}} = P \times_G T\mathcal{C}$.

Definition 8.2.2 *The **infinitesimal action** $\text{Ad } P \times \mathcal{H}_{\mathcal{C}} \rightarrow \mathcal{H}_{T\mathcal{C}}$ is defined on $\mathcal{H}_{\mathcal{C}}$ as follows:*

$$[p, \xi]_{\mathfrak{g}} \cdot [p, c]_{\mathcal{C}} = [p, \xi_{\mathcal{C}}(c)]_{T\mathcal{C}}, \quad (8.4)$$

where the vector field $\xi_{\mathcal{C}} \in \mathfrak{X}(\mathcal{C})$ denotes the infinitesimal generator of $\xi \in \mathfrak{g}$ on $\mathcal{C}$.

Conservation laws and representations

Definition 8.2.3 Given a connection form $\mathcal{A}$ on $\pi_{\Sigma, P}$, the *covariant tangent functor* $T^{\mathcal{A}}$ defined on sections of $\pi_{X, \mathcal{H}_C}$ by

$$T^{\mathcal{A}} : \Gamma \pi_{X, \mathcal{H}_C} \rightarrow \Gamma \pi_{X, L(\pi_{X, \Sigma}^* TX, \mathcal{H}_{TC})}, \quad T^{\mathcal{A}} \llbracket \sigma, c \rrbracket_C = \llbracket \sigma, Tc - (\sigma^* \mathcal{A})_C \circ c \rrbracket_{TC}. \quad (8.5)$$

Therefore, if $\bar{c}$ is a section of $\pi_{X, \mathcal{H}_C}$ covering $\rho \in \Gamma \pi_{X, \Sigma}$, then $T^{\mathcal{A}} \bar{c}(v_x) \in (\mathcal{H}_{TC})_{\rho(x)}$.

A map $K : \mathcal{H}_C \rightarrow \text{Ad}^{**} P$ may be represented by a G -equivariant map $\mathcal{K} : \mathcal{C} \rightarrow \mathfrak{g}^{**}$ defined by the relation

$$K(\llbracket p, c \rrbracket_C) = \llbracket p, \mathcal{K}(c) \rrbracket_{\mathfrak{g}^{**}}.$$

For a detailed discussion on the relationship linking maps between associated bundles with equivariant maps, see KOLÁŘ ET AL. [1993]. Here $\mathfrak{g}^{**}$ denotes the double dual of the Lie algebra. For an example where the distinction between $\mathfrak{g}$ and $\mathfrak{g}^{**}$ arises, see HOLM ET AL. [1998].

The derivative of $K : \mathcal{H}_C \rightarrow \text{Ad}^{**} P$ may be defined as follows:

$$dK(\llbracket p, v_c \rrbracket_{TC}) = \llbracket p, d\mathcal{K}(v_c) \rrbracket_{\mathfrak{g}^{**}}.$$

Note that $dK : (\mathcal{H}_{TC})_s \rightarrow \text{Ad}_s^{**} P$, that is, dK is fibre preserving; and $d^{\mathcal{A}}(K \circ \bar{c}) = dK \circ T^{\mathcal{A}} \bar{c}$, where $\bar{c} = \llbracket \sigma, c \rrbracket_C$ denotes a section of $\pi_{X, \mathcal{H}_C}$. Indeed,

$$\begin{aligned} d^{\mathcal{A}}(K \circ \bar{c}) &= \llbracket \sigma, d(\mathcal{K} \circ c) - (\sigma^* \mathcal{A}) \cdot (\mathcal{K} \circ c) \rrbracket_{\mathfrak{g}^{**}} \\ &= \llbracket \sigma, d\mathcal{K} \circ Tc - d\mathcal{K} \circ (\sigma^* \mathcal{A})_C \circ c \rrbracket_{\mathfrak{g}^{**}} \\ &= \llbracket \sigma, d\mathcal{K} \circ (Tc - (\sigma^* \mathcal{A})_C \circ c) \rrbracket_{\mathfrak{g}^{**}} \\ &= dK \circ T^{\mathcal{A}} \bar{c}. \end{aligned}$$

This relation is described by the following commutative diagram:

$$\begin{array}{ccc} \Gamma \pi_{X, L(\pi_{X, \Sigma}^* TX, \mathcal{H}_{TC})} & \xrightarrow{dK} & \Gamma \pi_{X, L(\pi_{X, \Sigma}^* TX, \text{Ad}^{**} P)} \\ \uparrow T^{\mathcal{A}} & & \uparrow d^{\mathcal{A}} \\ \Gamma \pi_{X, \mathcal{H}_C} & \xrightarrow{K} & \Gamma \pi_{X, \text{Ad}^{**} P} \end{array}$$

The infinitesimal actions described in (8.4) on $\text{Ad}^{**}P$ and $\mathcal{H}_C$ are related via

$$\llbracket p, \xi \rrbracket_{\mathfrak{g}} \cdot K(\bar{c}) = dK \left(\llbracket p, \xi \rrbracket_{\mathfrak{g}} \cdot \bar{c} \right). \quad (8.6)$$

Additionally, observe the following relationship:

$$\text{div} \langle \bar{\nu}, \bar{\mu} \rangle = \langle d^A \bar{\nu}, \bar{\mu} \rangle + \langle \bar{\nu}, \text{div}^A \bar{\mu} \rangle, \quad (8.7)$$

where $\bar{\nu}$ and $\bar{\mu}$ are sections of $\pi_{X, \text{Ad}^{**}P}$ and $\pi_{X, L(\pi_{X, \Sigma}^* T^* X, \text{Ad} P^*)}$ respectively, and both cover the same section ρ of $\pi_{X, \Sigma}$.

With this background, the generalisation of the Kelvin-Noether theorem may be stated as follows.

Theorem 8.2.4 (Kelvin-Noether theorem) *Let $\bar{\sigma} \in \Omega_{\pi_{X, \Sigma}}^1(X, \text{Ad} P)$ be a solution to the Lagrange-Poincaré equations (7.22), and $\bar{c} \in \Gamma \pi_{X, \mathcal{H}_C}$ cover $\rho := \pi_{\Sigma, L(\pi_{X, \Sigma}^* T^* X, \text{Ad} P^*)} \circ \bar{\sigma} \in \Gamma \pi_{X, \Sigma}$ while satisfying*

$$T^A \bar{c} + \bar{\sigma}_C \circ \bar{c} = 0. \quad (8.8)$$

*If $K : \mathcal{H}_C \rightarrow \text{Ad}^{**}P$ fibre-preserving map that covers the identity on Σ then the associated circulation*

$$I := \left\langle K \circ \bar{c}, \frac{\delta l}{\delta \bar{\sigma}} \right\rangle \in \mathfrak{X}(X)$$

satisfies

$$\text{div} I = 0. \quad (8.9)$$

Conservation laws and representations

Proof. The result is obtained via a direct calculation that uses (7.22), (8.7), (8.6), and (8.8) as follows:

$$\begin{aligned}
 \operatorname{div} I &= \operatorname{div} \left\langle K \circ \bar{c}, \frac{\delta l}{\delta \bar{\sigma}} \right\rangle \\
 &= \left\langle d^{\mathcal{A}}(K \circ \bar{c}), \frac{\delta l}{\delta \bar{\sigma}} \right\rangle + \left\langle K \circ \bar{c}, \operatorname{div}^{\mathcal{A}} \frac{\delta l}{\delta \bar{\sigma}} \right\rangle \\
 &= \left\langle (d^{\mathcal{A}} + \operatorname{ad}_{\bar{\sigma}}^{**})(K \circ \bar{c}), \frac{\delta l}{\delta \bar{\sigma}} \right\rangle + \left\langle K \circ \bar{c}, (\operatorname{div}^{\mathcal{A}} - \operatorname{ad}_{\bar{\sigma}}^*) \frac{\delta l}{\delta \bar{\sigma}} \right\rangle \\
 &= \left\langle (d^{\mathcal{A}} + \operatorname{ad}_{\bar{\sigma}}^{**})(K \circ \bar{c}), \frac{\delta l}{\delta \bar{\sigma}} \right\rangle \\
 &= \left\langle dK \circ (T^{\mathcal{A}}\bar{c} + \bar{\sigma}_C \circ \bar{c}), \frac{\delta l}{\delta \bar{\sigma}} \right\rangle = 0,
 \end{aligned}$$

as required. ■

Recall that classical Lagrangian reduction (particular case **iv**), used the formulation $\sigma(t) = (t, q(t))$ and $\bar{\sigma}(t) = \llbracket q(t), \mathcal{A}(\dot{q}(t)) \rrbracket_{\mathfrak{g}}$. In this case (8.8) becomes

$$T^{\mathcal{A}}\bar{c} + \bar{\sigma}_C \circ \bar{c} = \llbracket q, \dot{c} - (\mathcal{A}(\dot{q}))_C \circ c \rrbracket_{TC} + \llbracket q, (\mathcal{A}(\dot{q}))_C \circ c \rrbracket_{TC} = \llbracket q, \dot{c} \rrbracket_{TC} = 0.$$

Therefore (8.8) diminishes to the assumptions for the classical Kelvin-Noether theorem; see HOLM ET AL. [1998]. Furthermore the conclusion to Theorem 8.2.4 in this context becomes

$$\frac{d}{dt} \left\langle \mathcal{K}(c), \frac{\delta l}{\delta \xi} \right\rangle = 0.$$

These results extend those of CENDRA ET AL. [1997] to the Lagrange-Poincaré context.

8.3 Kelvin-Noether theorem for the molecular strand

The Kelvin-Noether theorem (8.2.4) gives qualitative information about the behaviour of solutions to the covariant Lagrange-Poincaré equations. Recall the molecular strand problem of Part I, as described from the covariant point of view in §4. Denoting divergence relative to the variable $x = (s, t) \in [0, L] \times \mathbb{R}$ by div_x ,

Theorem (4.1.2) may be rewritten as

$$\begin{aligned}
 0 &= \text{Ad}_{\Lambda^{-1}}^* \left(\partial_s \frac{\delta \bar{l}_{loc}}{\delta \Omega} + \Omega \times \frac{\delta \bar{l}_{loc}}{\delta \Omega} + \partial_t \frac{\delta \bar{l}_{loc}}{\delta \omega} + \omega \times \frac{\delta \bar{l}_{loc}}{\delta \omega} \right) \\
 &= \partial_s \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \Omega} \right) + \partial_t \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \right) \\
 &= \text{div}_x \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \Omega} \partial_s + \text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \partial_t \right).
 \end{aligned} \tag{8.10}$$

Equation (8.10) yields the following qualitative information: when the strand is a closed loop, the change of variables (1.16) implies that the circulation of its spatial angular momentum integrated around the moving loop in the convective frame is conserved. That is,

$$\frac{d}{dt} \oint \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \right) ds = - \oint \partial_s \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \Omega} \right) ds = 0. \tag{8.11}$$

By using the divergence theorem, this circulation theorem may be re-expressed covariantly as

$$\begin{aligned}
 0 &= \int_S \text{div}_x \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} \partial_s + \text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \partial_t \right) ds dt \\
 &= \int_{\partial S} \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \Omega} \partial_s + \text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \partial_t \right) \cdot \mathbf{n} d\ell \\
 &= \int_{\partial S} \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta \bar{l}_{loc}}{\delta \Omega} dt - \frac{\delta \bar{l}_{loc}}{\delta \omega} ds \right),
 \end{aligned}$$

where $\mathbf{n}$ is the outward pointing unit normal to the boundary ∂S and the following identity has been used

$$\left(\frac{\delta \bar{l}_{loc}}{\delta \Omega} \partial_s + \frac{\delta \bar{l}_{loc}}{\delta \omega} \partial_t \right) \cdot \mathbf{n} d\ell = \frac{\delta \bar{l}_{loc}}{\delta \Omega} dt - \frac{\delta \bar{l}_{loc}}{\delta \omega} ds. \tag{8.12}$$

Thus, the covariant expression of the circulation theorem (8.11) becomes

$$\int_{\partial S} \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta \bar{l}_{loc}}{\delta \Omega} dt - \frac{\delta \bar{l}_{loc}}{\delta \omega} ds \right) = 0, \tag{8.13}$$

which may be interpreted as a zero-flux theorem in space-time X .

Conservation laws and representations

Combining (8.10) with (4.19) shows that,

$$\begin{aligned} \operatorname{div}_x \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta (\bar{l}_{loc} + \bar{l}_{np})}{\delta \Omega} \partial_s + \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \partial_t \right) = \\ \int \left(\frac{\partial U}{\partial \kappa}(s, s') \times \kappa(s, s') + \mathbf{Z}(s, s') \right) \mathrm{d}s'. \end{aligned} \quad (8.14)$$

Therefore, the nonlocal terms destroy the conservation of average spatial angular momentum around the strand when it forms a loop. Let $s_1, s_2 \in [0, L]$, the relation

$$\begin{aligned} \frac{d}{dt} \int_{s_1}^{s_2} \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}_{loc}}{\delta \omega} \mathrm{d}s = \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta (\bar{l}_{loc} + \bar{l}_{np})}{\delta \Omega} \Big|_{s_1}^{s_2} \\ + \int_{s_1}^{s_2} \int \left(\frac{\partial U}{\partial \kappa}(s, s') \times \kappa(s, s') + \mathbf{Z}(s, s') \right) \cdot \mathbf{n} \mathrm{d}s' \mathrm{d}s \end{aligned} \quad (8.15)$$

shows that the average angular momentum along a part of the moving strand is not only related to the amount the stand twists over the interval, but is also *driven* by the nonlocal interactions. This observation will provide the topic for further study and, perhaps, the beginnings of an explicit continuum dynamical description of protein coiling and folding. Also note that the accuracy of these conclusions would provide a solid basis from which to test this theory of molecular strand dynamics with experimental data.

For the general setting,

$$\begin{aligned} \operatorname{div}_x \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} + \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \omega} \partial_t \right) &= \operatorname{div} \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} \right) + \partial_t \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \omega} \right) \\ &= \operatorname{Ad}_{\Lambda^{-1}}^* \left(\operatorname{div} \frac{\delta \bar{l}}{\delta \Omega} - \operatorname{ad}_{\Omega}^* \frac{\delta \bar{l}}{\delta \Omega} + \partial_t \frac{\delta \bar{l}}{\delta \omega} - \operatorname{ad}_{\omega}^* \frac{\delta \bar{l}}{\delta \omega} \right) = 0. \end{aligned}$$

The divergence theorem gives a generalisation of the zero flux theorem (8.13), now applied to an n -dimensional strand (or surface),

$$\int_{\partial V} \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} + \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \omega} \partial_t \right) \cdot \mathbf{n} \, d\sigma = 0,$$

where $\mathbf{n}$ is the outward pointing unit normal to the boundary ∂V of a given domain

$V \subset \mathcal{D} \times \mathbb{R}$ and $d\sigma$ is the induced boundary volume element of ∂V .

8.4 Summary

In §8 we have looked at two local representations covariant Lagrange-Poincaré dynamics, extended the Kelvin-Noether Theorem to the covariant Lagrange-Poincaré setting, and applied these results to the molecular strand problem of Part I.

In §8.1 the convective and spatial representations were developed by applying suitable representations to the Lagrange-Poincaré equations (7.22). These discussions helped explain the conservation and Hamilton-Pontryagin forms of the molecular strand equations derived in §2.

The Kelvin-Noether Theorem was addressed in §8.2. Some preliminary definitions were made, after which Theorem 8.2.4 extended the Kelvin-Noether Theorem in two ways; from the classical to the covariant setting, and from the Euler-Poincaré to the Lagrange-Poincaré setting.

§8.3 applied the newly derived Kelvin-Noether Theorem to the molecular strand dynamics derived in §2 and §4. The analysis culminated in (8.15), which shows that average angular momentum along the molecular strand is driven by the nonlocal interaction forces. This potentially provides the basis for future research moving towards dynamical modelling of protein coiling and folding.

CHAPTER 9 : CONCLUSION OF PART II

Part II presented a theory for covariant Lagrange-Poincaré reduction that includes each of the particular cases **i** - **iv**. On one hand, the work of CASTRILLÓN-LÓPEZ ET AL. [2000] and CASTRILLÓN-LÓPEZ & RATIU [2003] has been extended to apply to the general fibre bundle case. On the other hand, the classical Lagrange-Poincaré theory developed in CENDRA ET AL. [1997] and CENDRA ET AL. [2001] has been extended to the field theoretic setting. The key landmarks include Theorem 7.1.8 and the subsequent particular cases, taken together these achieve the objective of a unifying reduction theory called for in §6.1. The covariant Lagrange-Poincaré theory relies on the interaction of two bundle structures on a manifold P , as indicated by the following commutative diagram.

$$\begin{array}{ccc}
 P & \xrightarrow{\pi_{X,P}} & X \\
 & \searrow \pi_{\Sigma,P} & \nearrow \pi_{X,\Sigma} \\
 & \Sigma &
 \end{array}$$

In its most general form the diagram describes the geometric setting for fibre bundle reduction. By making various assumptions and identifications, the diagram recovers the geometric setting for each of the particular cases **i** - **iv**.

In addition to the reduction theorem, a cohesive reconstruction theory has been developed for the fibre bundle reduction. This culminated in Theorem 7.1.9 and the explicit reconstruction equation (7.32). The reconstruction theorem may be regarded as an existence and uniqueness theorem for the explicit reconstruction PDE, which requires an additional relation that was first observed in CASTRILLÓN-LÓPEZ & RATIU [2003]. The additional condition, and other axioms of the reconstruction theorem are always satisfied in the classical case, which coincides with the fact that the reconstruction equation can always be solved in classical mechanics.

In §8 two particular gauges were chosen to represent the vertical covariant Lagrange-Poincaré equation locally. These choices of gauge resulted in the Euler-Poincaré equation for the convective representation, and Noether's Theorem for the spatial representation. It was noted that the convective representation may not always exist in the full fibre bundle reduction, although always does for the classical case. Even though the convective representation may not exist, the general $\text{Ad}P$ -valued objects do exist and they can be used either to study the dynamics or to find an alternative representation. The convective representation was related with the covariant perspective of the molecular strand set out in §4. Meanwhile, the spatial representation was related with the conservation laws derived in §2.3, which explained the vanishing torque phenomenon and related the conservation law form of the molecular strand equations (2.3) with the covariant perspective described in §4.

Finally, the Kelvin-Noether Theorem 8.2.4 extended the result from CENDRA ET AL. [1997] to the Lagrange-Poincaré field setting. This extension achieved two generalisations. Firstly, the theory has been extended from Euler-Poincaré systems where there is no shape space to the Lagrange-Poincaré setting, where a shape space is present. Secondly, the Kelvin-Noether Theorem has been generalised from the classical to the covariant context. The extended Kelvin-Noether Theorem constitutes a major tool for gaining qualitative information about any problem formulated within the scope of the Lagrange-Poincaré field theory. An application was made to the molecular strand dynamics in §8.3, where the average twist of the molecular strand was seen to be driven by the nonlocal interactions. This could provide the basis for future work aimed at providing a dynamical model for protein coiling and folding.

The covariant Lagrange-Poincaré theory of Part II provides a solid foundation to explicitly link the techniques used in §3 with those used in §4. In §3 the configuration space is TG where $G = \mathcal{F}([0, L]; SE(3))$ is the space of smooth maps from $[0, L]$ into $SE(3)$. The theory proceeded as a classical reduction, considering the variation of curves in the infinite dimensional Lie group G . In contrast, the configuration space for the covariant side is the first jet bundle J^1P , which is finite dimensional. The covariant theory proceeds by varying sections of π_{X, J^1P} . Indeed, GAY-BALMAZ & RATIU [2009] establishes such a connection for the covariant Euler-Poincaré reduction in CASTRILLÓN-LÓPEZ ET AL. [2000] and the affine

Conclusion of Part II

Euler-Poincaré theory in GAY-BALMAZ & RATIU [2008]. The results there rely on the fact that the connection form $\mathcal{A}$ may also be interpreted as a cocycle in a similar way to the trick used to extend the reduced Lagrangian in §3.3.2. This result shows how such a relation could be established for the full covariant Lagrange-Poincaré framework.

When developing a link between the classical and covariant perspectives, a particular difficulty arises in dealing with the shape space components. For the molecular strand the shape space variable was $\rho(s, t) \in \mathbb{R}^3$, as was demonstrated in §4.3. The vector space structure of the shape space for the molecular strand allowed the interpretation of ρ as a cocycle in §3. In general, however, the shape space need neither be a vector space, nor even a group. In that case, no interpretation of the shape space variable as a cocycle can be made. We may therefore tentatively expect that affine Euler-Poincaré theory in the classical perspective to only apply to a subset of the covariant theories that may be written in classical form.

The alternation between the classical and covariant perspectives may be referred to as *classical-covariant duality*. That is, a mechanical system possesses classical-covariant duality when it may be interpreted as both a covariant or classical theory with a means to pass from one perspective to the other. An example of a classical-covariant duality is provided by the molecular strand, where §3 provides the classical interpretation, while §4.3 provides the corresponding covariant interpretation. §4.1 demonstrates the procedure for passing between the two perspectives, which is manifested at the local level by a coordinate transformation.

The extension of classical-covariant duality to covariant Lagrange-Poincaré theory provides an interesting project for the future that has the potential to yield new insights for classical reduction.

*Seize the moment of excited curiosity on any subject to solve your doubts;
For if you let it pass, the desire may never return,
And you may remain in ignorance.*

William Wirt

Part III

Un-reduction

CHAPTER 10 : UNREDUCTION

§3 and §4 demonstrate how the reduced molecular strand equations (2.41) resulted from affine Euler-Poincaré reduction and subgroup covariant Lagrange-Poincaré reduction respectively. The reduced dynamics, however, require a reconstruction equation such as (7.32) in order to relate them to the dynamics of the Euler-Lagrange equations. Assuming the reconstruction equation is not satisfied, there are potentially many sections and curves in the unreduced state space that project onto the solution of the reduced equations.

§10 describes the freedom that may be obtained by relaxing the reconstruction equation. The motivation for such an enquiry does not come from the molecular strand, as has been the case in Parts I and II, but rather from an application to image dynamics. Part III relates the results of BRUVERIS ET AL. [2011].

10.1 Introduction

Recently there has been a certain degree of interest in geodesic shape matching, as in MICHOR & MUMFORD [2007]; COTTER & HOLM [2009]; BAUER ET AL. [2010] and references therein. This interest has grown from the desire to assist applications in medical imaging by providing a coherent, quantitative method for comparing shapes.

Here, a *shape* is taken to be an embedded or immersed submanifold of an ambient space. For example, a closed, simple, planar curve is a shape since it is the image of an embedding $S^1 \hookrightarrow \mathbb{R}^2$. The space of shapes (or shape space) may be realised as a quotient. Given two smooth manifolds $\mathcal{M}$ and $\mathcal{N}$, the shape space $\mathcal{M}$ -type submanifolds of $\mathcal{N}$ may be identified with the quotient space $\Sigma := \text{Emb}(\mathcal{M}, \mathcal{N}) / \text{Diff}(\mathcal{M})$. That is, the space of embeddings of $\mathcal{M}$ into $\mathcal{N}$ up to reparameterisation of $\mathcal{M}$.

In order to compare two shapes, a natural way to proceed is to construct a path between them. This is known as the *matching problem*. For example, the matching problem for closed, simple planar curves may be stated as follows: Let ρ_0 and ρ_1

be two smooth submanifolds in $\mathbb{R}^2$ of S^1 -type. Find a path of submanifolds of S^1 -type, $\rho(t)$, such that $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$. In practice there are many such paths, therefore a variational principle is used to select a preferred one.

Problem 10.1.1 (The matching problem) *Minimise the functional*

$$\int_0^1 \ell^\Sigma(\rho, \dot{\rho}) dt$$

subject to the boundary conditions $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$, where Σ denotes the space of all submanifolds of S^1 -type in the plane and $\ell^\Sigma : T\Sigma \rightarrow \mathbb{R}$ is a Lagrangian defined on the tangent bundle of Σ .

As is well known, the solution to Problem 10.1.1 satisfies the Euler-Lagrange equations

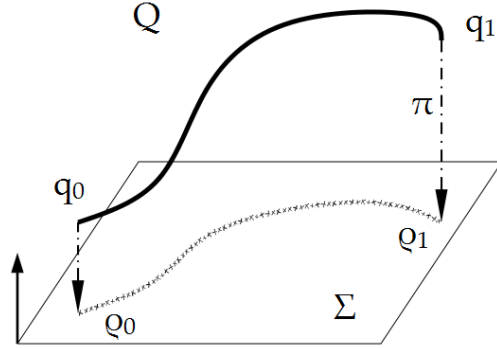
$$\frac{d}{dt} \frac{\delta \ell^\Sigma}{\delta \dot{\rho}} - \frac{\delta \ell^\Sigma}{\delta \rho} = 0. \quad (10.1)$$

In order for these equations to make sense rigorously, Σ must be endowed with a smooth manifold structure. This is achieved in MICHOR [1980]; KRIEGL & MICHOR [1997]. The key concern with Problem 10.1.1, however, is that numerical implementation of (10.1) is rendered unfeasible by a lack of coordinates for the shape space Σ . Sadly, the solution derived via the approach of direct application of the variational principle lacks practical utility, despite the analytical understanding of the problem. The task for Part III is to investigate alternative formulations of Problem 10.1.1 that takes advantage of the analytical results in such a way that their solutions are tractable to practical implementation.

10.1.1 Introducing un-reduction

Recent progress has been made for the matching problem with simple, planar, closed curves in the geodesic case, with $\ell^\Sigma(\rho, \dot{\rho}) = \mathcal{G}_\rho^\Sigma(\dot{\rho}, \dot{\rho})$ where $\mathcal{G}^\Sigma$ is a Riemannian metric on the shape space $\Sigma = \text{Emb}(S^1, \mathbb{R}^2) / \text{Diff}(S^1)$. This approach involved formulating a related geodesic problem on $\text{Emb}(S^1, \mathbb{R}^2)$ rather than Σ , MICHOR & MUMFORD [2007]; COTTER & HOLM [2009]. These papers study the

Figure 10.1: A trajectory in Q from q_0 to q_1 is shown that projects via π to a trajectory in Σ from ρ_0 to ρ_1 . Un-reduction poses a boundary value problem in Q such that the projected trajectories are solutions to the Euler-Lagrange equations in Σ .



geodesic problem by adapting techniques commonly used for reduction by symmetry of variational principles, as in MONTGOMERY [1993], in a novel way. Recalling particular case **iv** of Part II, given a principal G -bundle $\pi : Q \rightarrow \Sigma$ and a G -invariant Lagrangian $L : TQ \rightarrow \mathbb{R}$, it is possible to write down a reduced variational principle on TQ/G . This procedure is called *classical Lagrange-Poincaré reduction*. The novel insight in MICHOR & MUMFORD [2007]; COTTER & HOLM [2009] is to apply this procedure in reverse. That is, when interested in geodesics on the base space Σ , a geodesic problem is formulated on Q that reduces to the desired geodesic problem on Σ , rather than applying a variational principle directly on Σ . That is, a form of *un-reduction* is employed, and the resulting equations on Q are more convenient to deal with numerically. This approach has also been applied recently in BAUER ET AL. [2010] to achieve promising results for higher dimensional geodesic problems.

Some numerical difficulties persist, in that, although the geodesic in shape space may be successfully represented, its parameterisation may still evolve in an undesirable way. For example, when the geodesic equations are implemented on $\text{Emb}(S^1, \mathbb{R}^2)$, phenomena such as clustering of data points and bad parameterisations present themselves, as shown in Figure 1 of COTTER & HOLM [2009]. One approach developed in COTTER & HOLM [2009] deals with these issues by adding a step in the numerical procedure that reparameterises the initial conditions to evolve onto a prescribed parameterisation of the target shape. This initial reparameterisation procedure commutes with evolution along the matching trajectory, and may also be iterated a number of times throughout the evolution without significantly changing the problem. This iterative approach effectively breaks the matching problem up into a number of sub-problems. While the periodic reparam-

eterisation method works well in practice, it is *extrinsic* in nature, and sidesteps the problem with the matching equations rather than dealing with it directly. It would be preferable to seek matching equations for which the parameterisation problem never arises. That is, for the parameterisation issues to be dealt with *intrinsically*. Indeed, an intrinsic method is also described in (loc. cit.), which involves relaxing constraints on the initial conditions. The extra initial degrees of freedom thereby obtained allows matching of any boundary data irrespective of parameterisation, however nothing is said about the parameterisation dynamics along the trajectory. One of the aims in developing un-reduction is to provide the necessary tools to derive matching equations that intrinsically control parameterisation dynamics along the entire trajectory. Such a derivation is given in §10.5, which appears in BRUVERIS ET AL. [2011], represents the first geodesic matching equations with intrinsic, dynamic parameterisation control.

The papers reporting progress so far, such as MICHOR & MUMFORD [2007]; COTTER & HOLM [2009]; BAUER ET AL. [2010], have constrained themselves to geodesic problems. More precisely, geodesics on Σ are obtained by projection of horizontal geodesics on Q , relative to a G -invariant Riemannian metric on Q that projects to the given metric on Σ . Whilst the geodesic problem is important, it may be useful to match curves in a variety of other ways depending on the application. It turns out that in order to achieve the desired goals in parameterisation dynamics, the geodesic properties must be partially sacrificed. Further, in certain situations it may be useful to assert that the curve dynamics respect different properties such as that of being described as a graph in the ambient space.

Future work may include, for example, of the geometric splines approach for image analysis as recently discussed in VIALARD [2009], which may fit naturally into the un-reduction framework that is develop here. However, an application to image analysis using geometric splines is beyond our present scope and will be discussed elsewhere.

The framework used in developing geodesic matching needs to be extended in order to incorporate non-geodesic properties. The motivation in Part III is to provide one possible extension by developing extra geometric tools, and to demonstrate their use on an established problem from curve matching.

10.1.2 Aims of Part III

The objective of Part III is to investigate the extent to which the un-reduction approach may be used in the design of numerical methods, by investigating its geometrical context. The investigation, inspired by the novel approach taken in COTTER & HOLM [2009] that introduced reduction-by-symmetry techniques in reverse for the particular case of geodesic simple, closed, planar curve matching, reveals a rich geometric framework for un-reduction. This framework turns out to be broader in scope than the pioneering curve matching examples. The essential features of an un-reduction procedure are highlighted in purely geometric terms, giving clarity and rigorously capturing the generic notion of ‘doing reduction-by symmetry in reverse’. When cast in the language of geometry, the un-reduction technique is sufficiently general to tentatively it to find many other productive uses, even in different fields such as data assimilation. See also COTTER & HOLM [2009] for additional outlook toward further potential applications of un-reduction.

The general theory is applied to the curve matching problem in §10.5 goes some way towards answering the call for greater control over parameterisation dynamics in COTTER & HOLM [2009]. For example, the investigation of the geometry shows that the problem of quality control in parameterisations, and the extension to include potential forces may be addressed *simultaneously*. Further, the general un-reduction algorithms admit the introduction of a family of exogenous design factors that may be used either to formulate optimal control problems based on parameterisation dynamics, or for modelling purposes. This extra modelling capability may then lead to greater functionality of the solution, such as adaptive matching algorithms in which data points may evolve to preserve properties such as uniform parameterisation over time.

After explaining the basic ideas in the method of reduction by symmetry in §10.2, the approach for applying the symmetry-reduction method *in reverse* as un-reduction is considered in §10.3. Having achieved a good understanding of the fundamental issues, a general procedure for applying the un-reduction method is then developed in §10.4. Parallels are drawn both with particular cases of un-reduction and with reduction by symmetry. Finally, in §10.5 a new un-reduction procedure is applied to the curve matching problem, and comparisons are made with previous treatments.

10.2 Lagrangian reduction with a force field

Un-reduction is closely related to classical Lagrange-Poincaré reduction, which served as particular case **iv** in Part II. §10.2 briefly reviews the main results of Lagrangian reduction in the slightly generalised case when equivariant external forces are present. Such a force is encoded as a fibre-preserving map $F : TQ \rightarrow T^*Q$ over the identity, see §7.8 in MARS DEN & RATIU [2002]. The G -equivariance property reads

$$F(v_q g) = F(v_q)g, \quad \text{for all } g \in G,$$

where G acts by tangent and cotangent lifted actions, respectively.

Let $L : TQ \rightarrow \mathbb{R}$ be a G -invariant Lagrangian under the tangent lifted action of G on TQ and consider the associated reduced Lagrangian $l : TQ/G \rightarrow \mathbb{R}$. Fix a connection $\mathcal{A}$ on Q and consider the vector bundle isomorphism $\alpha_{\mathcal{A}} : TQ/G \rightarrow T\Sigma \oplus \text{Ad } Q$ over Σ given by

$$\alpha_{\mathcal{A}}([v_q]_G) = T\pi(v_q) \oplus \bar{\mathcal{A}}(v_q), \quad (10.2)$$

as in (6.9), where the notation $\bar{\mathcal{A}}(v_q) = \llbracket q, \mathcal{A}(v_q) \rrbracket_{\mathfrak{g}}$ has been introduced. The inverse is

$$\alpha_{\mathcal{A}}^{-1}(u_{\rho} \oplus \bar{\xi}) = [\text{Hor}_q u_{\rho} + \bar{\xi}_Q(q)]_G,$$

where for $\bar{\xi} = \llbracket q, \xi \rrbracket_{\mathfrak{g}}$, $\bar{\xi}_Q(q) = \xi_Q(q)$, which defines $\bar{\xi}_Q$ as a G -invariant vector field on Q . Thus, the reduced Lagrangian l may be regarded as a map $l : T\Sigma \oplus \text{Ad } Q \rightarrow \mathbb{R}$.

Definition 10.2.1 *The cotangent lift momentum map is defined as the map $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$ such that*

$$\langle \mathbf{J}(p_q), \xi \rangle = \langle p_q, \xi_Q(q) \rangle$$

for all $p_q \in T_q^*Q$ and $\xi \in \mathfrak{g}$.

Departing from Part II, consider a G -equivariant force field $F : TQ \rightarrow T^*Q$. The reduced force fields F^{Σ} and F^{Ad} are defined by

$$F^{\text{Ad}} : T\Sigma \oplus \text{Ad } Q \rightarrow \text{Ad}^* Q, \quad F^{\Sigma}(\rho, \dot{\rho}, \bar{\sigma}) := \bar{\mathbf{J}}(F(q, \dot{q})) \quad (10.3)$$

$$F^{\Sigma} : T\Sigma \oplus \text{Ad } Q \rightarrow T^*\Sigma, \quad \langle F^{\Sigma}(\rho, \dot{\rho}, \bar{\sigma}), v_{\rho} \rangle := \langle F(q, \dot{q}), \text{Hor}_q(v_{\rho}) \rangle, \quad (10.4)$$

Unreduction

where $(q, \dot{q}) \in TQ$ are such that $\alpha_{\mathcal{A}}([q, \dot{q}]_G) = (\rho, \dot{\rho}, \bar{\sigma})$, and $\bar{\mathbf{J}}(p_q) := \llbracket q, \mathbf{J}(p_q) \rrbracket_{\mathfrak{g}^*}$. Note that by G -equivariance of F , the right hand side does not depend on the choice of $(q, \dot{q})$ in the equivalence class. The force field F is related to F^Σ and F^{Ad} as follows:

$$\begin{aligned} \langle F(q, \dot{q}), w_q \rangle &= \langle F(q, \dot{q}), \text{Hor}_q(T\pi(w_q)) \rangle + \left\langle F(q, \dot{q}), (\mathcal{A}(w_q))_Q(q) \right\rangle \\ &= \langle \text{Hor}_q^*(F(q, \dot{q})), T\pi(w_q) \rangle + \langle \mathbf{J}(F(q, \dot{q})), \mathcal{A}(w_q) \rangle_{\mathfrak{g}^* \times \mathfrak{g}} \\ &= \langle F^\Sigma(\rho, \dot{\rho}, \bar{\sigma}), T\pi(w_q) \rangle + \langle \bar{\mathbf{J}}(F(q, \dot{q})), \bar{\mathcal{A}}(w_q) \rangle \\ &= \langle F^\Sigma(\rho, \dot{\rho}, \bar{\sigma}), T\pi(w_q) \rangle + \langle F^{\text{Ad}}(\rho, \dot{\rho}, \bar{\sigma}), \bar{\mathcal{A}}(w_q) \rangle. \end{aligned}$$

It is convenient to also define the vertical and horizontal part of F by considering the decomposition $T^*Q = (VQ)^\circ \oplus (HQ)^\circ$, where

$$VQ^\circ := \{\alpha_q \in T^*Q \mid \langle \alpha_q, v_q \rangle = 0, \text{ for all } v_q \in V_q Q\} \quad \text{and} \quad (10.5)$$

$$HQ^\circ := \{\alpha_q \in T^*Q \mid \langle \alpha_q, v_q \rangle = 0, \text{ for all } v_q \in H_q Q\}, \quad (10.6)$$

are the annihilators of the vertical and horizontal distributions, respectively. These allow the following definitions

$$F^h : TQ \rightarrow VQ^\circ \quad \text{and} \quad F^v : TQ \rightarrow HQ^\circ$$

as $F^h(v_q) := P_h^*(F(v_q)) \in VQ^\circ$ and $F^v(v_q) := P_v^*(F(v_q)) \in HQ^\circ$ and obtain the relations

$$F^h(v_q) = \pi^*(F^\Sigma(\alpha_{\mathcal{A}}(v_q))) \quad \text{and} \quad F^v(v_q) = \bar{\mathcal{A}}^*(F^{\text{Ad}}(\alpha_{\mathcal{A}}(v_q))). \quad (10.7)$$

The classical Lagrange-Poincaré reduction theorem (cf. particular case **iv** following Theorem 7.1.8) in the case that external forces are allowed reads as follows.

Theorem 10.2.1 (Classical Lagrangian reduction with forces) *Consider a curve $q(t) \in Q$ and define the two curves*

$$\rho(t) := \pi(q(t)) \in \Sigma \quad \text{and} \quad \bar{\sigma}(t) := \bar{\mathcal{A}}(\dot{q}(t)) \in \text{Ad } Q,$$

where $\mathcal{A}$ is a fixed connection on $\pi : Q \rightarrow \Sigma$. Then, the following statements are equivalent:

1. Hamilton's variational principle,

$$\delta \int_{t_1}^{t_2} L(q(t), \dot{q}(t)) dt + \int_{t_1}^{t_2} F(q(t), \dot{q}(t)) \delta q(t) dt = 0, \quad (10.8)$$

holds for variations $\delta q(t)$ of $q(t)$ vanishing at the endpoints.

2. The curve $q(t)$ satisfies the Euler-Lagrange equations with external forces:

$$\frac{D^{\tilde{\nabla}}}{Dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = F(q, \dot{q}), \quad (10.9)$$

where $\tilde{\nabla}$ is an arbitrary torsion free affine connection on Q .

3. The constrained variational principle (of Lagrange-d'Alembert type)

$$\begin{aligned} & \delta \int_{t_1}^{t_2} l(\rho(t), \dot{\rho}(t), \bar{\sigma}(t)) dt \\ & + \int_{t_1}^{t_2} (\langle F^{\text{Ad}Q}(\rho(t), \dot{\rho}(t), \bar{\sigma}(t)), \bar{\eta}(t) \rangle + \langle F^{\Sigma}(\rho(t), \dot{\rho}(t), \bar{\sigma}(t)), \delta \rho(t) \rangle) dt = 0 \end{aligned}$$

holds, using variations $\delta \rho(t)$ of $\rho(t)$ vanishing at the endpoints, and variations of $\bar{\sigma}(t)$ of the form

$$\delta^{\mathcal{A}} \bar{\sigma}(t) = \frac{D^{\mathcal{A}}}{Dt} \bar{\eta}(t) - [\bar{\sigma}(t), \bar{\eta}(t)] - \tilde{\mathcal{B}}(\dot{\rho}(t), \delta \rho(t))$$

vanishing at the endpoints.

4. The Lagrange-Poincaré equations for l with external forcing $F^{\text{Ad}Q}$ and F^{Σ} hold:

$$\begin{cases} \frac{D^{\mathcal{A}}}{Dt} \frac{\delta l}{\delta \bar{\sigma}} + \text{ad}_{\bar{\sigma}}^* \frac{\delta l}{\delta \bar{\sigma}} = F^{\text{Ad}Q}(\rho, \dot{\rho}, \bar{\sigma}) \\ \frac{D^{\nabla}}{Dt} \frac{\delta l}{\delta \dot{\rho}} - \frac{\delta l}{\delta \rho} = F^{\Sigma}(\rho, \dot{\rho}, \bar{\sigma}) - \left\langle \frac{\delta l}{\delta \bar{\sigma}}, \mathbf{i}_{\dot{\rho}} \tilde{\mathcal{B}} \right\rangle, \end{cases} \quad (10.10)$$

where ∇ denotes a torsion-free affine connection on Σ .

Unreduction

Remark 10.2.2 In Theorem 10.2.1 the Euler-Lagrange equations (10.9) are written with the help of a torsion free connection $\tilde{\nabla}$ on Q , in order to have a global (i.e. coordinate independent) formulation. These equations may also be written locally as

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = F(q, \dot{q}),$$

where $\delta L / \delta q$ denotes the partial derivative of L in a local chart.

Remark 10.2.3 (Energy and momentum map) In the presence of a force field, the time derivative of the energy $E(q, \dot{q}) = \left\langle \dot{q}, \frac{\delta L}{\delta \dot{q}} \right\rangle - L(q, \dot{q})$ along a solution of (10.9) is

$$\frac{d}{dt} E(q(t), \dot{q}(t)) = \langle F(q(t), \dot{q}(t)), \dot{q}(t) \rangle.$$

A direct computation, using the same arguments as in §2.7 MARSDEN [1992], shows that in the presence of forces, Noether's theorem is replaced by the relation

$$\frac{d}{dt} \mathbf{J} \left(\frac{\delta L}{\delta \dot{q}} \right) = \mathbf{J} (F(q, \dot{q})) = \mathbf{J} (F^v(q, \dot{q})). \quad (10.11)$$

Remark 10.2.4 (Euler-Poincaré reduction) In the particular case $Q = G$, the G -invariant force field is completely determined by a smooth map $f : \mathfrak{g} \rightarrow \mathfrak{g}^*$, and $\langle F(v_g), w_g \rangle = \langle f(v_g g^{-1}), w_g g^{-1} \rangle$. In this case, the Lagrange-Poincaré equations (10.10) recover the Euler-Poincaré equations with a force,

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi} = f(\xi).$$

These equations are equivalent to Noether's Theorem (10.11), since they can both be written as

$$\frac{d}{dt} \text{Ad}_g^* \frac{\delta l}{\delta \xi} = \text{Ad}_g^* (f(\xi)), \quad \xi(t) = \dot{g}(t)g(t)^{-1},$$

which corresponds to the spatial representation of the vertical Lagrange-Poincaré equation with a force and the reconstruction equation respectively.

10.3 Distortions of projected dynamics

In order to formulate an un-reduction procedure, the Lagrange-Poincaré equations (10.10) must be related with the Euler-Lagrange equations (10.1) on Σ Theorem

10.2.1. Here, differences between the Lagrange-Poincaré equations and the Euler-Lagrange equations are called *distortions*. This terminology is suitable since, in the context of un-reduction, such differences are unwanted and the main goal of the un-reduction procedure is to eliminate them.

The conception of certain terms in the reduced equations as unwanted barriers to the objective stands in contrast with applications of reduction by symmetry, where such differences instead attain physical meaning. For example, in ELLIS ET AL. [2010] in the context of molecular strand dynamics where micro-structure of a strand couples with the reduced filament dynamics.

There are two distortions to contend with:

1. **Coupling distortion:** Often a G -invariant Lagrangian $L : TQ \rightarrow \mathbb{R}$ is given as opposed to a Lagrangian on $T\Sigma$, since coordinates on Σ are unavailable. Since (10.2) shows that $TQ/G \cong T\Sigma \oplus \text{Ad } Q$, the Lagrangian L certainly depends on $\text{Ad } Q$ if it is non-degenerate. It is therefore necessary to remove any coupling between the $\text{Ad } Q$ dependencies and the $T\Sigma$ dependencies from the unreduced Euler-Lagrange equations on Q . In general, there is no preferred way to deal with these dependencies, and therefore they must be tailored to each particular un-reduction problem. This ambiguity of the Lagrangian on TQ is referred to hereafter as *coupling distortion*.
2. **Curvature distortion:** The Lagrange-Poincaré equations (10.10) differ from the Euler-Lagrange equations (10.1) in that the right hand side contains a driving term that arises from the curvature of the principal connection. This driving term is referred to as *curvature distortion*.

The next task is to describe a particular class of G -invariant Lagrangians $L : TQ \rightarrow \mathbb{R}$ well appropriate for the formulation of un-reduction. The G -invariant Lagrangian is said to *decouple relative to a connection* $\mathcal{A}$ if it can be written as the sum of two G -invariant Lagrangians, $L = L^h + L^v$, where $L^h : TQ \rightarrow \mathbb{R}$ and $L^v : TQ \rightarrow \mathbb{R}$ are such that

$$L^h(v_q) = L^h(P^h(v_q)) \quad \text{and} \quad L^v(v_q) = L^v(P^v(v_q)), \quad \text{for all } v_q \in TQ,$$

that is, $L^h(v_q)$ depends only on the horizontal part of v_q and $L^v(v_q)$ depends only on the vertical part of v_q . Now, the bundle isomorphism $\alpha_{\mathcal{A}} : TQ/G \rightarrow T\Sigma \oplus \text{Ad } Q$

Unreduction

induces two bundle isomorphisms

$$\alpha_{\mathcal{A}}|_{VQ/G} : VQ/G \rightarrow \text{Ad } Q \quad \text{and} \quad \alpha_{\mathcal{A}}|_{HQ/G} : HQ/G \rightarrow T\Sigma.$$

Therefore, L^h is completely determined by a Lagrangian $l^\Sigma : T\Sigma \rightarrow \mathbb{R}$ and L^v is completely determined by a function $l^{\text{Ad}} : \text{Ad } Q \rightarrow \mathbb{R}$ through the relations

$$L^h(\text{Hor}_q(v_\rho)) = l^\Sigma(v_\rho) \quad \text{and} \quad L^v(\xi_Q(q)) = l^{\text{Ad}}(\llbracket q, \xi \rrbracket_{\mathfrak{g}}).$$

Also note that

$$\frac{\delta L^h}{\delta \dot{q}} = \pi^* \frac{\delta l^\Sigma}{\delta \dot{\rho}} \in VQ^\circ \quad \text{and} \quad \frac{\delta L^v}{\delta \dot{q}} = \bar{\mathcal{A}}^* \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}} \in HQ^\circ, \quad (10.12)$$

where the last equality implies the relation $\frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}} = \bar{\mathbf{J}} \left(\frac{\delta L^v}{\delta \dot{q}} \right)$.

Example 10.3.1 (Geodesic problems) Let $\mathcal{G}$ be a G -invariant Riemannian metric on Q , and consider the Lagrangian $L(v_q) := \frac{1}{2} \|v_q\|_{\mathcal{G}}^2$ associated to the metric. The Lagrangian L decouples relative to a particular connection, which may be seen as follows:

Recall that a G -invariant Riemannian metric naturally induces a connection defined by $H_q Q = (V_q Q)^\perp$ and called the *mechanical connection*. The associated connection form $\mathcal{A}_{\text{mech}}$ is determined by the relation

$$\mathcal{G}_q(v_q, q \cdot \sigma) = \langle \mathbb{I}(q) \mathcal{A}_{\text{mech}}(v_q), \sigma \rangle_{\mathfrak{g}^* \times \mathfrak{g}}, \quad \text{for all } \sigma \in \mathfrak{g},$$

where $\mathbb{I}(q) : \mathfrak{g} \rightarrow \mathfrak{g}^*$ is the locked inertia tensor defined by

$$\mathcal{G}_q(q \cdot \nu, q \cdot \sigma) = \langle \mathbb{I}(q) \nu, \mu \rangle_{\mathfrak{g}^* \times \mathfrak{g}}, \quad \text{for all } \nu, \sigma \in \mathfrak{g}.$$

The Lagrangian $L(v_q) = \frac{1}{2} \|v_q\|_{\mathcal{G}}^2$ decouples relative to the mechanical connection, since

$$L(v_q) = \frac{1}{2} \|v_q\|_{\mathcal{G}}^2 = \frac{1}{2} \|P_v(v_q)\|_{\mathcal{G}}^2 + \frac{1}{2} \|P_h(v_q)\|_{\mathcal{G}}^2 = L^v(v_q) + L^h(v_q).$$

A vector bundle isomorphism $\mathcal{I} : \text{Ad } Q \rightarrow \text{Ad}^* Q$ over the identity, defined by

$$\mathcal{I} \left(\llbracket q, \sigma \rrbracket_{\mathfrak{g}} \right) = \llbracket q, \mathbb{I}(q)\sigma \rrbracket_{\mathfrak{g}^*},$$

is required in order to write l^{Ad} explicitly. The induced functions $l^{\text{Ad}} : \text{Ad } Q \rightarrow \mathbb{R}$ and $l^{\Sigma} : T\Sigma \rightarrow \mathbb{R}$ are therefore given by

$$l^{\text{Ad}}(\bar{\sigma}) = \frac{1}{2} \langle \mathcal{I}(\bar{\sigma}), \bar{\sigma} \rangle \quad \text{and} \quad l^{\Sigma}(\rho, \dot{\rho}) = \frac{1}{2} \|\dot{\rho}\|_{\pi_* \mathcal{G}}^2,$$

where $\pi_* \mathcal{G}$ is the Riemannian metric on Σ induced by $\mathcal{G}$. $\blacklozenge$

10.3.1 Distortion in the Lagrange-Poincaré equations.

Example 10.3.1 shows that the class of geodesic Lagrangians is contained within the restricted class of Lagrangians on Q that decouples relative to an appropriate choice of connection form. In this case, the reduced Lagrangian $l : TQ/G \rightarrow \mathbb{R}$ takes the form $l = l^{\Sigma} + l^{\text{Ad}}$. Consequently, using the relations

$$\frac{\delta l}{\delta \rho} = \frac{\delta l^{\Sigma}}{\delta \rho} + \frac{\delta l^{\text{Ad}}}{\delta \rho}, \quad \frac{\delta l}{\delta \dot{\rho}} = \frac{\delta l^{\Sigma}}{\delta \dot{\rho}}, \quad \frac{\delta l}{\delta \bar{\sigma}} = \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}},$$

the Lagrange-Poincaré equations (10.10) read:

$$\begin{cases} \frac{D^{\mathcal{A}}}{Dt} \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}} - \text{ad}_{\bar{\sigma}}^* \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}} = F^{\text{Ad}}(\rho, \dot{\rho}, \bar{\sigma}) \\ \frac{D^{\nabla}}{Dt} \frac{\delta l^{\Sigma}}{\delta \dot{\rho}} - \frac{\delta l^{\Sigma}}{\delta \rho} = F^{\Sigma}(\rho, \dot{\rho}, \bar{\sigma}) + \frac{\nabla l^{\text{Ad}}}{\delta \rho} - \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \mathbf{i}_{\dot{\rho}} \tilde{\mathcal{B}} \right\rangle. \end{cases} \quad (10.13)$$

Here the third term on the right hand side of the second equation in (10.13) is the curvature distortion, while the second term gives an explicit form of the coupling distortion. Note that the left hand side is the Euler-Lagrange operator on Σ for the Lagrangian $l^{\Sigma} : T\Sigma \rightarrow \mathbb{R}$. Thus, the goal of the un-reduction method is to selecting solutions to a Lagrangian system with forcing on Q such that the right hand side of the second equation vanishes. Such solutions in Q project onto solutions of the Euler-Lagrange equations on Σ with Lagrangian l^{Σ} .

10.4 Un-reduction

Having investigated the differences between the Lagrange-Poincaré equations and the Euler-Lagrange equations on $T\Sigma$, the next step, in light of (10.13) is to use the external force F^Σ to cancel the right hand side of the second equation in (10.13). Clearly, the reduced force $F^\Sigma : T\Sigma \oplus \text{Ad } Q \rightarrow T^*\Sigma$ needs to be defined by

$$F^\Sigma(\rho, \dot{\rho}, \bar{\sigma}) := \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \mathbf{i}_{\dot{\rho}} \bar{\mathcal{B}} \right\rangle - \frac{\delta l^{\text{Ad}}}{\delta \rho}. \quad (10.14)$$

For the moment the other component $F^{\text{Ad}} : T\Sigma \oplus \text{Ad } Q \rightarrow \text{Ad}^* Q$ of the reduced force is left unspecified. However, once F^{Ad} is fixed, then the force field $F : TQ \rightarrow T^*Q$ is completely determined by the equality

$$F(v_q) = F^h(v_q) + F^v(v_q) = \pi^* (F^\Sigma(\alpha_{\mathcal{A}}(v_q))) + \bar{\mathcal{A}}^* (F^{\text{Ad}}(\alpha_{\mathcal{A}}(v_q)))$$

as is shown in (10.7).

10.4.1 The un-reduction theorem

In §10.4.1, the variational principle (10.8) is used in order to compute the unreduced Euler-Lagrange equations associated to this forcing term F . Using the same notations as Theorem 10.2.1,

$$\begin{aligned} 0 &= \delta \int_0^1 L^h(q, \dot{q}) dt + \delta \int_0^1 L^v(q, \dot{q}) dt + \int_0^1 F^h(q, \dot{q}) \delta q dt + \int_0^1 F^v(q, \dot{q}) \delta q dt \\ &= \delta \int_0^1 L^h(q, \dot{q}) dt + \int_0^1 \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \delta \bar{\sigma} \right\rangle + \left\langle \frac{\delta l^{\text{Ad}}}{\delta \rho}, \delta \rho \right\rangle dt \\ &\quad + \int_0^1 \langle F^\Sigma(\rho, \dot{\rho}, \bar{\sigma}), \delta \rho \rangle dt + \int_0^1 \langle F^{\text{Ad}}(\rho, \dot{\rho}, \bar{\sigma}), \bar{\mathcal{A}}(\delta q) \rangle dt \\ &= \delta \int_0^1 L^h(q, \dot{q}) dt + \int_0^1 \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \delta \bar{\sigma} \right\rangle + \int_0^1 \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \bar{\mathcal{B}}(\dot{\rho}, \delta \rho) \right\rangle \\ &\quad + \int_0^1 \langle F^{\text{Ad}}(\rho, \dot{\rho}, \bar{\sigma}), \bar{\mathcal{A}}(\delta q) \rangle dt, \end{aligned}$$

where (10.14) has been used in the last equality.

The expression for the constrained variation $\delta\bar{\sigma}$ yields

$$\begin{aligned}
 \int_0^1 \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \delta \bar{\sigma} + \tilde{\mathcal{B}}(\dot{\rho}, \delta \rho) \right\rangle dt &= \int_0^1 \left\langle \frac{\delta l^{\text{Ad}}}{\delta \bar{\sigma}}, \frac{D^{\mathcal{A}}}{Dt} \bar{\eta} - [\bar{\sigma}, \bar{\eta}] \right\rangle dt \\
 &= \int_0^1 \left\langle \bar{\mathbf{J}} \left(\frac{\delta L^v}{\delta \dot{q}} \right), \llbracket q(t), \dot{\eta}(t) \rrbracket_{\mathfrak{g}^*} \right\rangle dt \\
 &= \int_0^1 \left\langle \mathbf{J} \left(\frac{\delta L^v}{\delta \dot{q}} \right), \dot{\eta}(t) \right\rangle dt \\
 &= - \int_0^1 \left\langle \frac{d}{dt} \mathbf{J} \left(\frac{\delta L^v}{\delta \dot{q}} \right), \mathcal{A}(\delta q) \right\rangle dt,
 \end{aligned}$$

since $\bar{\eta}(t) = \llbracket q(t), \mathcal{A}(\delta q(t)) \rrbracket_{\mathfrak{g}}$. Therefore, the variational principle yields the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\delta L^h}{\delta \dot{q}} - \frac{\delta L^h}{\delta q} = \bar{\mathcal{A}}^* F^{\text{Ad}}(\rho, \dot{\rho}, \bar{\sigma}) - \mathcal{A}^* \frac{d}{dt} \mathbf{J} \left(\frac{\delta L^v}{\delta \dot{q}} \right). \quad (10.15)$$

Equation (10.15) may be split into horizontal and vertical parts by observing that the terms on the right hand side belong to HQ° and the terms on the left hand side belong to VQ° , since

$$\frac{\delta L^h}{\delta \dot{q}} = \pi^* \frac{\delta l^\Sigma}{\delta \dot{\rho}} \quad \text{and} \quad \frac{\delta L^h}{\delta q} = \pi^* \frac{\delta l^\Sigma}{\delta \rho}. \quad (10.16)$$

Thus, the un-reduction equations (10.15) can be written equally well as

$$\frac{d}{dt} \frac{\delta L^h}{\delta \dot{q}} - \frac{\delta L^h}{\delta q} = 0 \quad (10.17)$$

$$\frac{d}{dt} \mathcal{A}^* \mathbf{J} \left(\frac{\delta L^v}{\delta \dot{q}} \right) = F^v(q, \dot{q}), \quad (10.18)$$

where $F^v : TQ \rightarrow HQ^\circ$ is completely determined by F^{Ad} . Note that the first equality in (10.16) yields

$$\mathbf{J} \left(\frac{\delta L^h}{\delta \dot{q}} \right) = 0,$$

and that equation (10.18) can be rewritten as

$$\frac{d}{dt} \mathbf{J} \left(\frac{\delta L^v}{\delta \dot{q}} \right) = \mathbf{J} (F^v(q, \dot{q})), \quad (10.19)$$

Unreduction

since $F^v(q, \dot{q}) \in HQ^\circ$. Equation (10.19) can also be deduced from the relation (10.11).

The results obtained so far are summarised in the following theorem:

Theorem 10.4.1 (Un-reduction) *Let $l^\Sigma : T\Sigma \rightarrow \mathbb{R}$ be an arbitrary Lagrangian defined on the base of a principal bundle $\pi : Q \rightarrow Q/G = \Sigma$. Let $L = L^h + L^v$ be a G -invariant Lagrangian such that $L^h \circ P_h = L^h$ and $L^v \circ P_v = L^v$, where L^h is uniquely determined by l^Σ , and choose an arbitrary force $F^v : TQ \rightarrow HQ^\circ$.*

Then, solutions $q(t)$ of the equations (10.17) — (10.18) project to solutions $\rho(t) := \pi(q(t))$ of the Euler-Lagrange equations for l^Σ .

Remark 10.4.1 Recall here that equations (10.17) — (10.18) are the unreduced Euler-Lagrange equations on TQ for a Lagrangian $L = L^v + L^h$ and a force field $F = F^v + F^h$, in the special case when F^h is constructed from the given Lagrangian l^Σ by the formula (10.14). Note that once a connection has been fixed, the choice of l^Σ determines L^h and vice versa. However, the choice of L^v and F^v is left open. The main content of un-reduction is that the projected dynamics on Σ is independent of the choice of L^v and F^v . Therefore, both these functions may be chosen arbitrarily, a freedom that constitutes a distinct modelling step for the particular application at hand. In practice, vertical dynamics, that is equation (10.18), may be chosen to give suitable numerical properties or to add additional functionality to the solution being developed. An example of such a modelling procedure is given in §10.5.

10.4.2 Horizontality conditions for geodesic problems

The horizontal methods for geodesic problems, developed and applied in MICHOR & MUMFORD [2007]; COTTER & HOLM [2009], may be recovered from the general un-reduction procedure developed here. The horizontal approach takes the Euler-Lagrange equations on Q with the geodesic Lagrangian and asserts, in addition, the condition

$$\mathbf{J} \left(\frac{\delta L}{\delta \dot{q}} \right) = 0. \quad (10.20)$$

To gain a clearer grasp of the horizontal approach for geodesic problems, consider first the broader problem of asserting (10.20) for a general Lagrangian $L = L^h + L^v$. Since $\delta L^h / \delta \dot{q} \in VQ^\circ$, the relation $\mathbf{J}(\delta L^h / \delta \dot{q}) = 0$ is always satisfied, therefore (10.20) implies $\mathbf{J}(\delta L^v / \delta \dot{q}) = 0$, and consequently $\delta l^{\text{Ad}} / \delta \bar{\rho} = \bar{\mathbf{J}}(\delta L^v / \delta \dot{q}) = 0$. The assertion of (10.20) therefore reduces the Lagrange-Poincaré equations (10.13) with zero external forces to the following single equation,

$$\frac{D^\nabla}{Dt} \frac{\delta l^\Sigma}{\delta \bar{\rho}} - \frac{\delta l^\Sigma}{\delta \rho} = \frac{\delta l^{\text{Ad}}}{\delta \rho}.$$

At the un-reduced level, since $\delta L^v / \delta \dot{q} \in HQ^\circ$ by (10.12), the condition

$$\mathbf{J}(\delta L^v / \delta \dot{q}) = 0$$

implies that $\delta L^v / \delta \dot{q} = 0$. Consequently, the Euler-Lagrange equations in the vanishing momentum case take the form

$$\frac{D^{\tilde{\nabla}}}{Dt} \frac{\delta L^h}{\delta \dot{q}} - \frac{\delta L^h}{\delta q} = \frac{\delta L^v}{\delta q}. \quad (10.21)$$

Therefore, the vanishing momentum map condition eliminates curvature distortion, although coupling distortion may persist.

Note that in general (10.21) does not coincide with the un-reduction equations with the vanishing momentum map condition imposed. Indeed, applying (10.20) to (10.17) — (10.18) results in

$$\tilde{\nabla}_t \frac{\delta L^h}{\delta \dot{q}} - \frac{\tilde{\nabla} L^h}{\delta q} = 0 \quad (10.22)$$

$$\mathbf{J}\left(\frac{\delta L}{\delta \dot{q}}\right) = 0, \quad (10.23)$$

where the coupling distortion term $\tilde{\nabla} L^v / \delta q$ has now been cancelled by the horizontal forcing.

Specialising to the particular case when L is the geodesic Lagrangian and $\mathcal{A}$ is the corresponding mechanical connection, the vanishing momentum map condi-

Unreduction

tion (10.20) becomes a horizontality condition,

$$\left(\frac{\delta L}{\delta \dot{q}}\right)^\sharp = \dot{q} \in HQ.$$

In addition, for the geodesic Lagrangian, $\tilde{\nabla} L^v / \delta q = \tilde{\nabla} L / \delta q = 0$ where $\tilde{\nabla}$ is the Levi-Civita connection. Thus, in the case of geodesic motion with vanishing momentum, the un-reduction equations are equivalent to the horizontal geodesic equations themselves

$$\tilde{\nabla}_t \dot{q} = 0 \tag{10.24}$$

$$\mathbf{J} \left(\frac{\delta L}{\delta \dot{q}} \right) = 0. \tag{10.25}$$

Whilst (10.24) - (10.25) may appear over-determined, the observation that (10.24) is an Euler-Lagrange equation with symmetry allows the interpretation of (10.25) as Noether's Theorem. Therefore (10.24) - (10.25) are consistent, and (10.25) asserts a constraint on the initial conditions. Namely, $\mathcal{A}(\dot{q}_0) = 0$, that is the initial velocity must be horizontal. Thus, un-reduction applied to the geodesic problem together with the vanishing momentum map condition recovers to the horizontal shooting method for geodesic problems.

This coincidence occurs because, in addition to the property of vanishing momentum to eliminate curvature distortion, the particular choice of the geodesic Lagrangian does not introduce any coupling distortion. Therefore, the forcing (both horizontal and vertical) is equal to zero, and the un-reduction procedure drops to reconstruction of a symmetry reduced Lagrangian system. That is, the un-reduction equations are the original Euler-Lagrange equations (i.e. the geodesic equation) with the vanishing momentum map condition asserted on the initial condition.

For a general Lagrangian, however, the vanishing momentum map condition is not admissible without modification since it does not result in the Euler-Lagrange equations for l^Σ on Σ . In this case, it is necessary to use the un-reduction approach developed here as opposed to asserting the vanishing momentum condition.

10.5 Closed, simple plane curve matching

§10.5 develops equations for matching closed, simple, planar curves using the unreduction techniques developed in Part III. The matching problem for such curves has been treated in, for example, MICHOR & MUMFORD [2006, 2007]; COTTER [2008]; COTTER & HOLM [2009]. The formulation of the problem differs between authors, however, the methods developed achieve consensus in agreeing that the path followed by the matching algorithms should be geodesics. Here, we present an alternative perspective is presented. Since any observer of the dynamics is only sensible of the geometric information, or ‘shape’, of the curve, it is natural to propose that matching requires only that the shape need follow a geodesic. Meanwhile, the dynamics of the whole curve, which includes information about both parameterisation and shape, may evolve in a non-geodesic fashion.

This broader notion of curve matching allows the introduction of an entire family of matching dynamics, one of whom is the original geodesic dynamics. This new family of matching dynamics, which is described efficiently by the unreduction technique outlined here, incorporates enough flexibility to implicitly overcome other difficulties. Here, for example, the extra flexibility brought by unreduction is used to address the unwieldy parameterisation dynamics displayed by geodesic matching, as in Figure 1 in COTTER & HOLM [2009], by giving dynamics that remain uniformly parameterised throughout the matching procedure. This achieves the objective set out in §10.1 to derive a set of matching dynamics that addressed the parameterisation problems faced by curve matching algorithms in an intrinsic, dynamic way along the entire trajectory.

Additionally, a member of the unreduction family is presented whose dynamics respect the property of curves being written as a graph in polar coordinates, wherever it is sensible to do so. In higher dimensional problems, such as that studied in BAUER ET AL. [2010], the graph preservation property may significantly simplify numerical implementation of the dynamics. These two examples of non-geodesic matching dynamics will, hopefully, stimulate the description of matching dynamics with other additional properties.

10.5.1 Geometric setup

Following MICHOR & MUMFORD [2006], consider $Q = \text{Emb}^+(S^1, \mathbb{R}^2)$, the smooth manifold of all positively oriented embeddings from S^1 to $\mathbb{R}^2$. Q may be thought of as the space of simple, closed, planar curves. An element $c \in Q$ contains information about a shape, namely a submanifold of $\mathbb{R}^2$ of S^1 -type, and a parameterisation of the shape. The space of shapes may therefore be identified with the quotient manifold $\Sigma = Q/G$, where $G = \text{Diff}^+(S^1)$ is the group of orientation preserving diffeomorphisms of S^1 which acts freely and properly on the right of Q by composition. See KRIEGL & MICHOR [1997] for a detailed discussion of the manifold structure of $\text{Emb}(S^1, \mathbb{R}^2)$ and Σ , and to MICHOR & MUMFORD [2006] for a discussion on the geometry of spaces of immersions and their quotients.

In particular, $\text{Emb}^+(S^1, \mathbb{R}^2)$ is an open subset of $C^\infty(S^1, \mathbb{R}^2)$. The quotient manifold, however, is not feasible for numerical implementation since there are no natural coordinates and analytical considerations can only be achieved via the use of equivalence classes. Therefore, it is natural to resort to an un-reduction procedure that takes advantage of the principal $\text{Diff}^+(S^1)$ -bundle structure of Q over Σ .

Recall that an embedding $c : S^1 \rightarrow \mathbb{R}^2$ is a smooth injective immersion that is a homeomorphism onto its image. Thus, Q may be identified with smooth, periodic maps $C_1^\infty(\mathbb{R}; \mathbb{R}^2)$, with, say, period 1, that are injective on each interval $[a, 1+a)$ for all $a \in \mathbb{R}$, and satisfy $|c_\theta|(\theta) \neq 0$ for all $\theta \in \mathbb{R}$. Here the subscript denotes differentiation.

Similarly, an element $f \in G := \text{Diff}^+(S^1)$ may be identified with a smooth, strictly monotonic map $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(\theta + 1) = f(\theta) + 1$ for all $\theta \in \mathbb{R}$.

Since Q is an open subset of $C^\infty(S^1, \mathbb{R}^2)$, a tangent vector $U_c \in T_c Q$ is represented by a pair of maps (c, U) , where $U \in C_1^\infty(\mathbb{R}, \mathbb{R}^2)$. The Lie algebra of G is given by the space $\mathfrak{g} := \mathfrak{X}(S^1)$ of vector fields on S^1 , which may also be represented by real valued, periodic functions $C_1^\infty(\mathbb{R}, \mathbb{R})$. The infinitesimal generator associated to $u \in \mathfrak{g}$ reads $u_Q(c) = (c, uc_\theta)$.

The following are some geometric quantities which are convenient to use due to their behavior under the action of G .

- Derivative along the curve

$$D_\theta = \frac{1}{|c_\theta|} \partial_\theta$$

- Length of the curve

$$l(c) = \int_0^1 |c_\theta| d\theta$$

- Unit tangent vector

$$\tau(c) = \frac{c_\theta}{|c_\theta|}$$

- Unit normal vector

$$n(c) = J\tau(c), \text{ where } J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

- Curvature

$$\kappa(c) = (D_\theta \tau(c)) \cdot n(c)$$

- Volume form

$$\text{vol}(c) = |c_\theta| d\theta$$

The cotangent space T_c^*Q at a point c comprises the one-form densities on c and $\mathfrak{g}^* := \mathfrak{X}(S^1)^*$, the dual of $\mathfrak{g} = \mathfrak{X}(S^1)$, comprises the one-form densities on S^1 . This is seen by introducing the following pairings

$$\langle (c, P \otimes \Omega), (c, U) \rangle := \int_{S^1} P(U) \Omega, \quad \text{and} \quad \langle \mu \otimes \omega, u \rangle = \int_{S^1} \mu(u) \omega,$$

respectively, where $P \otimes \Omega \in T_c^*Q$, is a one-form density on c ; $\mu \otimes \omega \in \mathfrak{g}^*$ is a one-form density on S^1 ; $(c, U) \in T_cQ$ is a tangent vector at c ; and $u \in \mathfrak{g}$ is a vector field on S^1 .

The cotangent lift momentum map, $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$, that arises with this setup is calculated to be

$$\mathbf{J}(c, P \otimes \Omega) = (P \cdot c_\theta) \otimes \Omega.$$

Fixing the density on S^1 to be the Lebesgue measure, $\omega = d\theta$, the momentum map becomes

$$\mathbf{J}(c, P \otimes \text{vol}(c)) = |c_\theta|(P \cdot \tau(c)) \otimes \text{vol}(c). \quad (10.26)$$

The simplest metric to consider would be the L^2 metric. However, as pointed out in MICHOR & MUMFORD [2006], this metric has arbitrarily small geodesic dis-

Unreduction

tance between any two curves. The simplest metric that is of practical importance is therefore the L^2 metric weighted by curvature. That is,

$$\mathcal{G}_c(U, V) = \int_0^1 (1 + \kappa^2(c)) (U \cdot V) \operatorname{vol}(c). \quad (10.27)$$

For a detailed discussion of the properties of this and other metrics such as Sobolev metrics, see MICHOR & MUMFORD [2006, 2007]. The mechanical connection corresponding to $\mathcal{G}$ is given by

$$\mathcal{A}_{\text{mech}}((c, U)) = \frac{U \cdot \tau(c)}{|c_\theta|}. \quad (10.28)$$

Indeed, since

$$\mathcal{G}_c(u_Q(c), v_Q(c)) = \int_0^1 (1 + \kappa^2(c)) uv |c_\theta|^2 \operatorname{vol}(c),$$

the locked inertia tensor is $\mathbb{I}(c)u = (1 + \kappa^2(c)) |c_\theta|^2 u \otimes \operatorname{vol}(c)$. Therefore, the identity

$$\langle \mathbb{I}(c) \mathcal{A}_{\text{mech}}(c, U), u \rangle = \mathcal{G}_c((c, U), (c, uc_\theta))$$

implies formula (10.28). The projections associated with $\mathcal{A}_{\text{mech}}$ are

$$P^v(c, U) = (c, (U \cdot \tau(c))\tau(c)) \quad \text{and} \quad P^h(c, U) = (c, (U \cdot n(c))n(c)).$$

Thus, the horizontal-vertical split of a vector field along c is just its decomposition into normal and tangent components.

Let L be the geodesic Lagrangian corresponding to (10.27),

$$L(c, c_t) = \frac{1}{2} \|c_t\|_Q^2 = \frac{1}{2} \int_0^1 (1 + \kappa^2(c)) |c_t|^2 \operatorname{vol}(c). \quad (10.29)$$

The horizontal part of L is

$$L^h(c, c_t) = \frac{1}{2} \int_0^1 (1 + \kappa^2(c)) h(c, c_t)^2 \operatorname{vol}(c) \quad (10.30)$$

where the notation $h(c, c_t) := c_t \cdot n(c)$, so that $P^h(c, c_t) = h(c, c_t)n(c)$.

10.5.2 Derivation of un-reduction equations

In order to derive the un-reduction equations (10.17) and (10.19) for the planar curve matching problem it is necessary to write down the Euler-Lagrange equations for the Lagrangian L^h . This is best accomplished by using a variational principle. For the variational approach, first the variations must be calculated. This calculation is achieved by the following lemma, after which un-reduction equations are derived.

Lemma 10.5.1 *The variations δh , $\delta \text{vol}(c)$ and $\delta \kappa$ are*

$$\begin{aligned}\delta h &= (\partial_t - sD_\theta)(\delta c \cdot n) + D_\theta h(\delta c \cdot \tau) \\ \delta \text{vol}(c) &= (D_\theta(\delta c \cdot \tau) - (\delta c \cdot n)\kappa) \text{vol}(c), \\ \delta \kappa &= (D_\theta^2 + \kappa^2)(\delta c \cdot n) + D_\theta \kappa(\delta c \cdot \tau).\end{aligned}$$

where for simplicity we write h, n, τ, κ instead of $h(c, c_t), n(c), \tau(c), \kappa(c)$.

Proof. It is convenient to calculate the commutator of differential operators, $[\delta, D_\theta]$, before continuing. That is,

$$\begin{aligned}[\delta, D_\theta] &= \delta \frac{\partial_\theta}{|c_\theta|} - \frac{\partial_\theta}{|c_\theta|} \delta \\ &= -\frac{c_\theta \cdot \delta c_\theta}{|c_\theta|^3} \partial_\theta \\ &= -(\tau \cdot D_\theta \delta c) D_\theta \\ &= ((n \cdot \delta c) \kappa - D_\theta(\tau \cdot \delta c)) D_\theta,\end{aligned}\tag{10.31}$$

where the identity $D_\theta \tau = \kappa n$ has been used. Employing (10.31), allows the variation of the tangent field τ to be pre-calculated,

$$\begin{aligned}\delta \tau &= \delta D_\theta c \\ &= D_\theta \delta c + [\delta, D_\theta] c \\ &= D_\theta((\tau \cdot \delta c) \tau + (n \cdot \delta c) n) + ((n \cdot \delta c) \kappa - D_\theta(\tau \cdot \delta c)) \tau \\ &= ((\tau \cdot \delta c) \kappa + D_\theta(n \cdot \delta c)) n.\end{aligned}\tag{10.32}$$

Unreduction

where $D_\theta n = -\kappa\tau$ has been used. Similarly, for the time derivative,

$$\partial_t \tau = ((\tau \cdot c_t) \kappa + D_\theta (n \cdot c_t)) n = ((\tau \cdot c_t) \kappa + D_\theta h) n.$$

Multiplying these variations by J results in the following equations for n ,

$$\delta n = J\delta\tau = -(D_\theta (\delta c \cdot n) + (\delta c \cdot \tau) \kappa) \tau \quad \text{and} \quad \partial_t n = J\partial_t \tau = -(D_\theta h + s\kappa) \tau,$$

where the notation $s := c_t \cdot \tau$ has been introduced.

Proceeding with the expression for δh ,

$$\begin{aligned} \delta h &= \delta (c_t \cdot n) = \delta c_t \cdot n + c_t \cdot \delta n \\ &= \partial_t (\delta c \cdot n) - \delta c \cdot \partial_t n - (D_\theta (\delta c \cdot n) + (\delta c \cdot \tau) \kappa) s \\ &= (\partial_t - sD_\theta) (\delta c \cdot n) + D_\theta h (\delta c \cdot \tau) \end{aligned}$$

as required.

The expression for $\delta \text{vol}(c)$ follows since

$$\begin{aligned} \delta \text{vol}(c) &= \frac{\delta c_\theta \cdot c_\theta}{|c_\theta|} d\theta \\ &= (D_\theta \delta c \cdot \tau) \text{vol}(c) \\ &= (D_\theta (\delta c \cdot \tau) - (\delta c \cdot n) \kappa) \text{vol}(c). \end{aligned}$$

Finally, upon noting that $\delta\tau \propto D_\theta\tau \propto n$, and similarly reversing τ and n , the derivation of the expression for $\delta\kappa$ reads

$$\begin{aligned} \delta\kappa &= \delta (D_\theta\tau \cdot n) \\ &= D_\theta \delta\tau \cdot n + [\delta, D_\theta]\tau \cdot n + D_\theta\tau \cdot \delta n \\ &= D_\theta (\delta\tau \cdot n) + [\delta, D_\theta]\tau \cdot n \\ &= D_\theta ((\tau \cdot \delta c) \kappa + D_\theta (n \cdot \delta c)) + \kappa ((n \cdot \delta c) \kappa - D_\theta (\tau \cdot \delta c)) \\ &= (D_\theta^2 + \kappa^2) (\delta c \cdot n) + D_\theta \kappa (\delta c \cdot \tau), \end{aligned}$$

as required. ■

Proposition 10.5.2 *The un-reduction equations for simple, planar, closed curves with the*

geodesic Lagrangian associated with $\mathcal{G}$ read

$$h_t = D_\theta(sh) - \frac{\kappa(1+3\kappa^2)}{2(1+\kappa^2)}h^2 + \frac{1}{(1+\kappa^2)}(D_\theta^2(\kappa h^2) - 2\kappa h D_\theta^2 h)$$

$$s_t = f(c, s, h)$$

$$c_t = hn + s\tau.$$

where f is an arbitrary smooth, G -invariant map $\text{Emb}(S^1, \mathbb{R}^2) \times C_1^\infty(\mathbb{R}) \times C_1^\infty(\mathbb{R}) \rightarrow C_1^\infty(\mathbb{R})$.

Proof.

The first step is to calculate the variational principle for Lagrangian (10.30) using the variations derived in Lemma 10.5.1. The variational principle may be calculated as

$$\begin{aligned} 0 &= \delta \int_0^1 \int_0^1 \frac{1}{2} (1 + \kappa^2) h^2 \text{vol}(c) dt \\ &= \int_0^1 \int_0^1 (1 + \kappa^2) h \delta h \text{vol}(c) dt + \int_0^1 \int_0^1 \frac{1}{2} (1 + \kappa^2) h^2 \delta \text{vol}(c) dt \\ &\quad + \int_0^1 \int_0^1 \kappa h^2 \delta \kappa \text{vol}(c) dt. \end{aligned}$$

After substituting the variations from Lemma 10.5.1 and integrating by parts, the terms proportional to $(\delta c \cdot \tau)$ vanish as expected since the Lagrangian L^h is known to be degenerate. The remaining terms, those proportional to $(\delta c \cdot n)$, yield the equation

$$h_t = D_\theta(sh) - \frac{\kappa h^2}{2} + \frac{1}{1+\kappa^2} ((D_\theta^2 + \kappa^2)(\kappa h^2) - h(\partial_t - sD_\theta)(1 + \kappa^2)). \quad (10.33)$$

Now, replacing δ with ∂_t in the expression for $\delta \kappa$ from Lemma 10.5.1 yields the following result

$$(\partial_t - sD_\theta)\kappa = (D_\theta^2 + \kappa^2)h.$$

Unreduction

Employing this relation on (10.33) gives

$$h_t = D_\theta(sh) - \frac{\kappa h^2}{2} + \frac{1}{1 + \kappa^2} ((D_\theta^2 + \kappa^2)(\kappa h^2) - 2\kappa h(D_\theta^2 + \kappa^2)h).$$

Rearranging terms yields the desired result.

The equation for c is simply a decomposition of c_t into horizontal and vertical parts. That is,

$$c_t = P^h(c_t) + P^v(c_t) = hn + s\tau. \quad (10.34)$$

Next, for the equation for s , note that (10.26) implies the relation

$$\mu = \mathbf{J} \left(\frac{\delta L}{\delta c_t} \right) = |c_\theta| (1 + \kappa^2) s \otimes \text{vol}(c) = \mathbb{I}(c)s \quad (10.35)$$

so equation (10.19) reads $\frac{d}{dt}(\mathbb{I}(c)s) = \mathbf{J}(F^v(c, c_t))$. Computing the time derivative on the left hand side and rearranging the terms yields the relation

$$s_t = \mathbb{I}(c)^{-1} \left(\mathbf{J}(F^v(c, c_t)) - \frac{d}{dt}(\mathbb{I}(c))s \right). \quad (10.36)$$

Substituting (10.34) on the right hand side of (10.36) yields an equation for s that may be expressed as

$$s_t = f(c, s, h),$$

for an arbitrary smooth, G -invariant map f . Of course, there is a relationship between f and $\mathbf{J}(F^v(c, c_t))$, however, this relationship is not necessary for the present purposes, so it has been omitted. ■

The un-reduction equations for curve matching derived in Proposition 10.5.2 represent a family of dynamical equations parameterised by an exogenous choice of smooth function $f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$. For any initial conditions $(c_0, \dot{c}_0) \in TQ$ and any choice of exogenous function f , the projected dynamics in Q/Σ , obtained by taking the image of a curve c , are guaranteed to follow geodesics under the metric induced on shape space by the curvature-weighted L^2 metric on Q . Consequently, the choice of parameterisation of the boundary data in matching algorithms, and the choice of forcing may be freely chosen for convenience of the user, to pose an optimal control problem, or to introduce extra modelling into the dynamics without needing to re-derive the equations. A couple of uses of this extra flexibility

are presented in §10.5.3. It should be noted, however, that the examples we give only scratch the surface of possible uses of the extra functionality brought by unreduction.

10.5.3 Examples of un-reduced image dynamics

Horizontal geodesics

The horizontal geodesics case has already been described in some detail in §10.4.2. Recall that c is horizontal if $s = 0$ (which is consistent with the vanishing momentum map condition, as shown by (10.35)). This case corresponds to the choice $f = 0$ together with the initial condition $s_0 = 0$, and recovers the horizontal shooting method of MICHOR & MUMFORD [2007]; COTTER & HOLM [2009]. The equations for this case are given by

$$h_t = -\frac{\kappa(1+3\kappa^2)}{2(1+\kappa^2)}h^2 + \frac{1}{(1+\kappa^2)}(D_\theta^2(\kappa h^2) - 2\kappa h D_\theta^2 h) \quad (10.37)$$

$$c_t = hn. \quad (10.38)$$

Note that by choosing $f = 0$, but allowing a free choice of the initial condition s_0 , one recovers the method introduced in COTTER & HOLM [2009] to address parameterisation concerns. These equations read

$$h_t = D_\theta(s_0 h) - \frac{\kappa(1+3\kappa^2)}{2(1+\kappa^2)}h^2 + \frac{1}{(1+\kappa^2)}(D_\theta^2(\kappa h^2) - 2\kappa h D_\theta^2 h) \quad (10.39)$$

$$c_t = hn + s_0 \tau. \quad (10.40)$$

Equations (10.39) — (10.40) incorporate an extra term s_0 compared with (10.37) — (10.38) which allow matching without the need to reparameterise the boundary data. Note that nothing is asserted about the parameterisation away from the boundary data. Differences between the equations derived here and those in COTTER & HOLM [2009] are attributed to different approaches being used. Their equations live on $\mathfrak{X}(\mathbb{R}^2)^*$, the one-form densities on $\mathbb{R}^2$, with the equations on T^*Q being given by a momentum map. Our equations and variational principle are formu-

lated directly on TQ . Whilst the precise relation between the two sets of equations remains unclear, the method of adding an initial vector field on S^1 in order to enable matching parameterised boundary data is in agreement. Indeed, comparing equations before and after adding in the extra parameter both here and in COTTER & HOLM [2009] shows that each set of equations are modified by adding precisely the same terms.

Uniformly parameterised morphing

The second case considers forcing f that ensures that the curve c is parameterised uniformly for all time. As BRUVERIS ET AL. [2011], this is referred to as *uniformly parameterised morphing*. That is, that the equation

$$|c_\theta| = l(c) := \int_0^1 \text{vol}(c) \quad (10.41)$$

holds for all time. Notice that the right hand side of (10.41) is independent of θ . MICHOR & MUMFORD [2006] shows that such curves, which are equally well described by the relation $|c_\theta| = \text{constant}$, form a submanifold $\mathcal{U} \subset \text{Emb}(S^1, \mathbb{R}^2)$. Differentiating (10.41) shows that a tangent vector $(c, U) \in T\mathcal{U}$ if and only if

$$|c_\theta| (D_\theta (U \cdot \tau) - (U \cdot n) \kappa) = - \int_0^1 (U \cdot n) \kappa \text{vol}(c). \quad (10.42)$$

Indeed, if $\delta c = U$, the variation of the left hand side of (10.41) is

$$\delta |c_\theta| = \frac{c_\theta \cdot \delta c_\theta}{|c_\theta|} = c_\theta \cdot D_\theta \delta c = c_\theta \cdot (\delta \tau - [\delta, D_\theta] c) = |c_\theta| (D_\theta (U \cdot \tau) - (U \cdot n) \kappa),$$

where $c_\theta \cdot \delta \tau = 0$ has been used. The variation of the right hand side is

$$\delta \int_0^1 \text{vol}(c) = \int_0^1 (D_\theta (U \cdot \tau) - (U \cdot n) \kappa) \text{vol}(c) = - \int_0^1 (U \cdot n) \kappa \text{vol}(c)$$

Applying (10.42) to $U = c_t$ reveals the following relation

$$|c_\theta| (D_\theta s - h\kappa) = - \int_0^1 h\kappa \text{vol}(c). \quad (10.43)$$

Solving the relation (10.43) for s in Proposition 10.5.2 would yield the required vertical term to keep the evolution of c uniformly parameterised. There is still a problem to deal with, namely that (10.42) does not respect the action of G . This corresponds to breaking the G -invariance of f in the un-reduction equations, which is disallowed since G -invariance is required by the un-reduction theory.

A solution to this issue may be found by noting that one only requires (10.42) *on the submanifold of uniformly parameterised curves, $\mathcal{U}$* . The intersection of each fibre of $Q \rightarrow \Sigma$ with $\mathcal{U}$ are the orbits of rigid rotation of S^1 . Indeed, (10.42) is invariant under the action of S^1 on Q given by $(\alpha, c(\theta)) \mapsto c(\theta + \alpha)$. Therefore, it is possible to define f on $\mathcal{U}$ such that (10.42) is satisfied, and then extend to the whole of Q by enforcing G -invariance.

Having described the way in which a suitable f may be found in principle, it now appears that there is a short cut that allows the explicit calculation of f to be bypassed. Note that on the submanifold $\mathcal{U}$, the relation $l/|c_\theta| = 1$ holds. Thus, multiplying this factor on the *left hand side only* of (10.42) results in a relation that is precisely equivalent to the original on $\mathcal{U}$, but is also G -invariant. This relation, applied to $U = c_t$, is

$$l(D_\theta s - h\kappa) = - \int_0^1 h\kappa \, \text{vol}(c). \quad (10.44)$$

Solving (10.44) for s yields the same result as integrating the second un-reduction equation from Proposition 10.5.2 with f constructed as described above. Rearranging (10.44) yields

$$D_\theta s = h\kappa - \frac{1}{l} \int_0^1 h\kappa \, \text{vol}(c). \quad (10.45)$$

Equation (10.45) states that the rate of change of s as one moves along the curve is equal to the deviation of the quantity $h\kappa$ from its average around the curve. Since uniform parameterisation is a nonlocal property of the curve, it is natural to expect that nonlocal terms appear in the equations, these terms take the form of the average of $h\kappa$.

Combining (10.45) with the un-reduction equations from Proposition 10.5.2

Unreduction

yields the following equations

$$h_t = D_\theta(sh) - \frac{\kappa(1+3\kappa^2)}{2(1+\kappa^2)}h^2 + \frac{1}{(1+\kappa^2)}(D_\theta^2(\kappa h^2) - 2\kappa h D_\theta^2 h) \quad (10.46)$$

$$s(\theta) = s(0) + \int_0^\theta h\kappa \operatorname{vol}(c) - \frac{1}{l} \int_0^\theta \operatorname{vol}(c) \int_0^1 h\kappa \operatorname{vol}(c), \quad (10.47)$$

$$c_t = hn + s\tau. \quad (10.48)$$

Equations (10.46)—(10.48) may be described as the lift of geodesics on Σ under the metric induced by $\mathcal{G}$ to the submanifold, $\mathcal{U} \subset \operatorname{Emb}(S^1, \mathbb{R}^2)$, of uniformly parameterised embeddings. (10.46)—(10.48) are a new set of dynamical matching equations that appeared first in BRUVERIS ET AL. [2011]. The main interest in these un-reduction equations is that they solve the major issue concerning the horizontal shooting method, (10.37)—(10.38). That is, the parameterisation of solutions of (10.46)—(10.48) behaves regularly for all time. This property solves the problems concerning clustering of data points and bad parameterisations implicitly along the entire trajectory, thereby solving the parameterisation issues described in §10.1.

Section morphing

During a recent visit to London, M. Bauer and P. Harms asked whether or not unreduction could be used to achieve geodesic morphing of shapes whilst respecting the property of being described as a graph. Working with Bauer and Harms it was possible to show that this objective is indeed attainable. The argument is presented here for the case of simple, closed planar curves.

To capture the setup geometrically, it is helpful to begin with slight generalisations. Consider $Q = \operatorname{Emb}(\mathcal{M}, \mathcal{N})$, which is a principal bundle over $\Sigma = Q/G$ where $G = \operatorname{Diff}(\mathcal{M})$, as in KRIEGL & MICHOR [1997]. Suppose that $\mathcal{N} = \mathcal{M} \times \mathcal{S}$ is a trivial fibre bundle over $\mathcal{M}$ with fibre $\mathcal{S}$ and projection $\pi_{\mathcal{N}}$ given by projection on the first factor. A **graph** is then a section $\eta \in \Gamma(\pi_{\mathcal{N}})$. Also, a curve $c(t) \subset Q$ is said to respect the graph property if and only if $c(t) \subset \Gamma(\pi_{\mathcal{N}}) \cap Q$. That is, $\pi_{\mathcal{N}} \circ c(t) = \operatorname{id}_{\mathcal{M}}$.

Interesting choices could involve $(\mathcal{M}, \mathcal{N}) = (S^2, \mathbb{R}^3/\{0\} \equiv S^2 \times \mathbb{R}^+)$, which

could be used to match spheres embedded in $\mathbb{R}^3$ under small deformations. The example $(\mathcal{M}, \mathcal{N}) = (\mathbb{R}^2, \{(x, y, z) \in \mathbb{R}^3 | z > 0\} \equiv \mathbb{R}^2 \times \mathbb{R}^+)$ could be used to compare different topographies.

These examples are beyond the scope of the present work, and are left to future investigation. The case of closed, simple planar curves corresponds to the particular choice $(\mathcal{M}, \mathcal{N}) = (S^1, \mathbb{R}^2/\{0\} \equiv S^1 \times \mathbb{R}^+)$. It is expedient to identify S^1 with the unit circle in $\mathbb{R}^2$, and fix

$$\pi_{\mathcal{N}} : \mathcal{N} = \mathbb{R}^2/\{0\} \rightarrow \mathcal{M} = S^1, \quad \pi_{\mathcal{N}}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|.$$

Section morphing is not available for all curves in the plane, but rather is constrained to curves $c \in \Gamma(\pi_{\mathcal{N}}) \cap Q$ that contain the origin. Considering curves up to translations in the plane, the choice of origin becomes arbitrary, and therefore the origin condition is ignored here.

The approach to deriving the un-reduction equations is similar to that taken for uniform parameterisation morphing in §10.5.3. That is, a relation is derived on the submanifold $\mathcal{U} := \Gamma(\pi_{\mathcal{N}}) \cap Q \subset Q$. The submanifold $\mathcal{U}$ is not G -invariant, in fact, it is only preserved by the identity. Despite this, in contrast with §10.5.3, the relation turns out to be G -invariant without the need for modification. Therefore, the implied forcing may be extended to the G -invariant submanifold $\mathcal{U} \cdot G \subset Q$. This observation again allows the equations to be identified as un-reduction equations without explicitly deriving the vertical forcing.

For the graph condition to remain invariant over time, the map $\pi_{\mathcal{N}} \circ c(t)$ must be time-invariant. This leads to the following calculation

$$0 = \partial_t \left(\frac{c}{|c|} \right) = \frac{1}{|c|} \left(c_t - c \frac{(c \cdot c_t)}{|c|^2} \right). \quad (10.49)$$

Equation (10.49) implies

$$s = c_t \cdot \tau = (c \cdot \tau) \frac{c \cdot c_t}{|c|^2} \quad \text{and} \quad h = c_t \cdot n = (c \cdot n) \frac{c \cdot c_t}{|c|^2}$$

which results in

$$s = h \frac{(c \cdot \tau)}{(c \cdot n)}. \quad (10.50)$$

Note that equation (10.50) is G -invariant, therefore it may be extended along

Unreduction

the whole of $\mathcal{U} \cdot G$. Differentiating (10.50) and substituting in equations from Proposition 10.5.2 would again result in a complicated expression for the exogenous parameter f , which would be the graph preservation forcing term. Given that (10.50), this calculation is omitted. Collecting the un-reduction equations from Proposition 10.5.2 together with (10.50) yields the section morphing un-reduction equations

$$\begin{aligned} h_t = & D_\theta \left(h^2 \frac{(c \cdot \tau)}{(c \cdot n)} \right) - \frac{\kappa(1 + 3\kappa^2)}{2(1 + \kappa^2)} h^2 \\ & + \frac{1}{(1 + \kappa^2)} (D_\theta^2 (\kappa h^2) - 2\kappa h D_\theta^2 h) \end{aligned} \quad (10.51)$$

$$c_t = \frac{h}{(c \cdot n)} c. \quad (10.52)$$

Introducing polar coordinates in the plane, $c(t, \theta) = r(t, \theta)(\cos \theta, \sin \theta)$, equations (10.51) — (10.52) become

$$h_t = D_\theta \left(h^2 \frac{r_\theta}{r} \right) - \frac{\kappa(1 + 3\kappa^2)}{2(1 + \kappa^2)} h^2 + \frac{1}{(1 + \kappa^2)} (D_\theta^2 (\kappa h^2) - 2\kappa h D_\theta^2 h) \quad (10.53)$$

$$r_t = h \sqrt{\left(\frac{r_\theta}{r} \right)^2 + 1}. \quad (10.54)$$

Note that equations (10.53) — (10.54) describe a set of differential equations on functions taking values in $\mathbb{R}$, whereas, the full un-reduction equations from Proposition 10.5.2 constitute a set of differential equations on functions taking values in $\mathbb{R}^2$. Therefore the particular choice of morphing that respects (10.49) has allowed the un-reduction equations to be integrated twice. In higher dimensional problems, this method of reduction which we call *section morphing* could prove invaluable for simplifying numerical implementation. Note that, in this example, clustering of data points is prevented by the graph preservation property. The only parameterisation effect that may cause a problem is the distance between data points becomes large as the radius $\max_\theta r(\theta, t)$ increases. This issue may be averted by first scaling the shapes to be matched, and then using section morphing. A detailed discussion on such developments is beyond the scope of the present work, and is left to future enquiries.

10.6 Summary

In §10 we have developed a general theory called un-reduction that applies Lagrangian reduction 'in reverse', and applied it to the matching problem in image dynamics.

The background material was described in §10.2 and culminated in the classical Lagrange-Poincaré reduction with forces theorem in Theorem 10.2.1.

Two types of distortion were identified in §10.3. These were coupling and curvature distortion. Also, a class of Lagrangians was identified that is amenable to the development of un-reduction. The geodesic Lagrangian was shown to belong to this class in Example 10.3.1.

The un-reduction procedure itself was developed in §10.4, and the main result came in Theorem 10.4.1. Following the un-reduction theorem, §10.4.2 related the new technique to the vanishing momentum and horizontal methods that are used to deal with geodesic problems in the literature.

Finally, §10.5 applied the un-reduction technique to the problem of matching closed, simple, planar curves. The un-reduction equations for planar curve matching were derived in Proposition 10.5.2. Subsequently, §10.5.3 specialised three specific sets of curve matching equations. The first of these corresponded with the horizontal method. The second and third were new dynamics called *uniformly parameterised morphing* and *section morphing*. These two new sets of matching equations are the first that exhibit intrinsic, dynamical control of the parameterisation.

*Every day you may make progress. Every step may be fruitful.
Yet there will stretch out before you an ever-lengthening, ever-ascending, ever-improving path.
You know you will never get to the end of the journey.
But this, so far from discouraging, only adds to the joy and glory of the climb.*

Sir Winston Churchill

Part IV

Outlook

CHAPTER 11 : OUTLOOK

This thesis has developed a dynamical model charged molecular strand dynamics, contributed to Lagrangian reduction by symmetry, and developed a new geometric framework called 'un-reduction' for performing reduction by symmetry in reverse, which we have been able to apply to derive new image matching equations.

While these achievements solve a range of problems, they also present new challenges for the future. Some of these problems are described here:

1. **Dynamical theory of protein coiling and folding:**

To build upon the qualitative information gained in §8 to provide a dynamical theory for protein folding and coiling. Such a study could find static equilibria and normal modes of the charged molecular strand, and compare the results with and without the nonlocal interactions. The static equilibria and normal modes could may then be related with empirical observations of, for example, the VDF oligomer NODA ET AL. [2003].

2. **Affine Euler-Poincaré vs. covariant Lagrange-Poincaré reduction:**

To relate the affine Euler-Poincaré reduction theorem (Theorem 3.2.1) with the covariant Lagrange-Poincaré reduction theorem (Theorem 7.1.8), as is accomplished for particular case ii in Part II by GAY-BALMAZ & RATIU [2009]. Particular difficulties are presented by reconciling the general shape space in covariant LP reduction with the necessity for a vector space structure required by the cocycle in the affine EP approach, as is pointed out in §9.

3. **Covariant Lagrange-Poincaré reduction vs. the multisymplectic approach:**

To relate the covariant Lagrange-Poincaré reduction theory developed in Part II with a multisymplectic reduction theory in the style of GOTAY ET AL. [2004]. At the unreduced level the relationship is well established, therefore the multisymplectic and variational method approach should also be related at the reduced level. The challenge is to write down a concrete transformation that relates the geometric structures and equations of each approach.

This achievement would further consolidate the various approaches to reduction of field theories, and could potentially lead to new insight in the field.

4. Application of un-reduction other types of image:

To apply the un-reduction theory developed in Part III to solve parameterisation issues with other types of image. Specifically, the notion of an image as an embedding or immersion extends to higher dimensions than the case studied in §10, as in BAUER ET AL. [2010]. It may be possible to alleviate parameterisation issues in higher dimensions using un-reduction. An entirely different interpretation of an image is as a density on a compact manifold. Since un-reduction is developed in abstract geometric terms in §10, one may tentatively expect a family of un-reduction equations for matching densities using the geometric setting outlined in, for example, KHESIN & WENDT [2009].

These challenges promise to yield deeper geometric insight, and continue to provide state of the art tools for a molecular strand dynamics, image dynamics, and for a range of future applications that have yet to be addressed.

APPENDIX A : APPENDIX OF NOTATION

Part I notation

ρ	Reduced convective strand position
Ω	Reduced convective RCC angular strain
Γ	Reduced convective filament stress
Ω	Reduced convective RCC angular velocity
γ	Reduced convective filament velocity
ξ	Relative RCC orientation
κ	Relative filament position
$\pi_{X,P}$	Projection from P onto X
σ	Section of $\pi_{X,P}$
ρ	Projected section; $\rho = \pi_{\Sigma,P} \circ \sigma$
$J^1 P$	First jet bundle of $\pi_{X,P}$
$j^1 \sigma$	First jet extension of sigma
$\mathcal{A}$	Principal connection
$\mathcal{B}$	Curvature of principal connection
∇^Σ	Affine connection on Σ
T^Σ	Torsion of ∇^Σ
$\text{Ad} P$	Adjoint bundle associated with P
$\mathcal{H}_\mathcal{C}$	Associated bundle $P \times_G \mathcal{C}$
$\llbracket p, c \rrbracket_\mathcal{C}$	Element of $\mathcal{H}_\mathcal{C}$ corresponding to $(p, c) \in P \times \mathcal{C}$
$\Omega^k(\Sigma; \text{Ad } P)$	Ad P -valued k -forms on Σ
$\bar{\sigma}$	Reduced section; element of $\Omega^1(X; \rho^* \text{Ad } P)$
P^ρ	Pull-back bundle $\rho^* P$
$\mathcal{A}^\rho$	Connection induced on P^ρ by $\mathcal{A}$
$\mathcal{A}^{\bar{\sigma}}$	Connection used for reconstruction
$\omega^{\bar{\sigma}}$	Vertical one-form on P^ρ induced by $\bar{\sigma}$
$T^\mathcal{A}$	Covariant tangent functor associated with connection $\mathcal{A}$

REFERENCES

- Antman, S. S. (2004). *Nonlinear Problems of Elasticity*. New York: Springer-Verlag.
- Baker, N., Sept, D., Joseph, S., Holst, M., & McCammon, J. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci*, 98, 10037–10041.
- Balaeff, A., Mahadevan, L., & Schulten, K. (1999). Elastic rod model of a DNA loop in the lac operon. *Phys. Rev. Lett*, 83, 4900–4903.
- Banavar, J., Hoang, T., Maddocks, H., Maritan, A., Poletto, C., Stasiak, A., & Trovato, A. (2007). Structural motifs of biomolecules. *Proc. Natl. Acad. Sci*, 104, 17283–17286.
- Bauer, M., Harms, P., & Michor, P. (2010). Almost local metrics on shape space of hypersurfaces in n-space. arXiv:1001.0717v2.
- Becker, K., Becker, M., & Schwarz, J. (2007). *String theory and M-theory: a modern introduction*. Cambridge University Press.
- Bloch, A. (2003). *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Applied Mathematics*. New York: Springer-Verlag. With the collaboration of J. Baillieul, P. Crouch and J. Marsden, With scientific input from P. S. Krishnaprasad, R. M. Murray and D. Zenkov, Systems and Control.
- Bruveris, M., Ellis, D., Gay-Balmaz, F., & Holm, D. (2011). Un-reduction. arXiv:1012.0076, Submitted to Jour. Geom. Mech.
- Castrillón-López, M. & Ratiu, T. (2003). Reduction in principal bundles: Covariant Lagrange-Poincaré equations. *Communications in Mathematical Physics*, 236, 223–250.
- Castrillón-López, M., Ratiu, T., & Shkoller, S. (2000). Reduction in principal fiber bundles: Covariant Euler-Poicaré equations. *Proc. Amer. Math. Soc.*, 128, 2155–2164.

REFERENCES

- Cendra, H., Holm, D., Marsden, J., & Ratiu, T. (1997). Lagrange reduction, the Euler-Poincaré equations and semidirect products. *Amer. Math. Soc.*, 186, 1—25.
- Cendra, H. & Marsden, J. (2005). Geometric mechanics and the dynamics of asteroid pairs. *Dynamical Systems*, 20, 3–21.
- Cendra, H., Marsden, J., & Ratiu, T. (2001). *Lagrangian Reduction by Stages*, volume 152. Memoirs American Mathematical Society.
- Cotter, C. (2008). The variational particle-mesh method for matching curves. *J. Phys. A: Math. Theor.*, 41, 18.
- Cotter, C. & Holm, D. (2009). Geodesic boundary value problems with symmetry. arXiv:0911.2205.
- Dichmann, D., Li, Y., & Maddocks, J. (1992). *Mathematical approaches to biomolecular structure and dynamics*, volume Mathematical Approaches to Biomolecular Structure and Dynamics, chapter Hamiltonian formulation and symmetries in rod mechanics, (pp. 71—115). New York: Springer IMA.
- Dill, E. (1992). Kirchhoff's theory of rods. *Arch. Hist. Exact Sci.*, 44, 1–23.
- Ellis, D., Gay-Balmaz, F., Holm, D., Putkaradze, V., & Ratiu, T. (2010). Symmetry reduced dynamics of charged molecular strands. *Arch. Rational Mech. Anal.*, 197, 811—902.
- Ellis, D., Gay-Balmaz, F., Holm, D., & Ratiu, T. (2009). Lagrange-Poincaré field equations. arXiv:0910.0874. Submitted to Journal of Nonlinear Science.
- Gamboa, R. & Solomin, J. (2003). On the global version of Euler-Lagrange equations. *J. Phys. A: Math. Gen.*, 36, 7301–7305.
- Gay-Balmaz, F., Holm, D., & Ratiu, T. (2009). Variational principles for spin systems and the Kirchhoff rod. *Journal of Geometric Mechanics*, 1, 417—444.
- Gay-Balmaz, F. & Ratiu, T. (2008). The geometric structure of complex fluids. *Advances in Applied Mathematics*, 42, 176—275.
- Gay-Balmaz, F. & Ratiu, T. (2009). A new lagrangian dynamic reduction in field theory. *Annales de l'institut Fourier*, 60, 1125—1160.

- Gisiger, T. & Paranjape, M. (1998). Recent mathematical developments in the skyrme model. *Physics Reports*, 306, 109.
- Goldstein, R., Goriely, A., Huber, G., & Wolgemuth, C. (2000). Bistable helices. *Phys. Rev. Lett.*, 84, 1631–1634.
- Goldstein, R., Powers, T., & Wiggins, C. (1998). Viscous nonlinear dynamics of twist and writhe. *Phys. Rev. Lett.*, 80, 5232–5235.
- Goriely, A. & Tabor, M. (1996). New amplitude equations for thin elastic rods. *Phys. Rev. Lett.*, 77, 3537–3540.
- Gotay, M., Isenberg, J., Marsden, J., Montgomery, R., Śniatycki, J., & Yasskin, P. (2004). Momentum maps and classical fields, part i: Covariant field theory. arXiv:physics/9801019v2.
- Hausrath, A. & Goriely, A. (2006). Repeat protein architectures predicted by a continuum representation of fold space. *Protein Science*, 15, 1–8.
- Holm, D. (2008a). *Geometric Mechanics Part 1: Dynamics and Symmetry*. Imperial College Press.
- Holm, D. (2008b). *Geometric Mechanics Part 2: Rotating, Translating and Rolling*. Imperial College Press.
- Holm, D., Marsden, J., & Ratiu, T. (1986). *Hamiltonian Structure and Lyapunov Stability for Ideal Continuum Dynamics*. Montreal: University of Montreal Press.
- Holm, D., Marsden, J., & Ratiu, T. (1998). The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.*, 137, 1–81.
- Holm, D. D. & Kupershmidt, B. (1988). The analogy between spin glasses and Yang-Mills fluids. *J. Math Phys.*, 29, 21–30.
- Holm, D. D. & Putkaradze, V. (2009). Nonlocal orientation-dependent dynamics of molecular strands. *C. R. Acad. Sci. Paris*, 347, 1093–1098.
- Janyška, J. & Modugno, M. (1996). Relations between linear connections on the tangent bundle and connections on the jet bundle of a fibred manifold. *Arch. Math. (Brno)*, 32, 281–288.

REFERENCES

- Kehrbaum, S. & Maddocks, J. (1997). Elastic rods, rigid bodies, quaternions, and the last quadrature. *Phil. Trans*, 355, 2117–2136.
- Khesin, B. & Wendt, R. (2009). *The geometry of infinite dimensional groups*. Springer.
- Kirchhoff, G. (1859). Über das Gleichgewicht und die Bewegung eines unendlich dünnen elastischen Stabes. *J. f. reine angew. Math.*, 56, 285–313.
- Klein, O. (1921). Quantentheorie und fünfdimensionale relativitätstheorie. *Zeitschrift für Physik a Hadrons and Nuclei*, 37, 895–906.
- Kolář, I., Michor, P., & Slovák, J. (1993). *Natural operations in differential geometry*. Springer-Verlag.
- Kriegl, A. & Michor, P. W. (1997). *The convenient setting of global analysis*, volume 53 of *Mathematical Surveys and Monographs*. Providence, RI: Amer. Math. Soc.
- Marsden, J. (1992). *Lectures on Mechanics*. Cambridge University Press.
- Marsden, J., Misiolek, G., Ortega, J.-P., Perlmutter, M., & Ratiu, T. (2007). *Hamiltonian Reduction by Stages*, volume 1913 of *Lecture Notes in Mathematics*. Berlin: Springer.
- Marsden, J. & Ratiu, T. (2002). *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems (Texts in Applied Mathematics)* (2nd ed.). Springer.
- Marsden, J., Ratiu, T., & Weinstein, A. (1984). Semidirect products and reduction in mechanics. *Trans. Am. Math. Soc.*, 281, 147–177.
- Marsden, J. & Shkoller, S. (1999). Multisymplectic geometry, covariant hamiltonians, and water waves. *Math. Proc. Camb. Phil. Soc.*, 125, 553.
- Mezic, I. (2006). On the dynamics of molecular conformation. *Proc. Natl. Acad. Sci*, 103, 7542–7547.
- Michor, P. (1980). Manifolds of smooth maps III: The principal bundle of embeddings of non-compact smooth manifolds. *Cah. Top. Géom. Diff. Cat.*, 21, 325–337.

- Michor, P. & Mumford, D. (2006). Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.*, 8, 1–48.
- Michor, P. & Mumford, D. (2007). An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Appl. Comput. Harmon. Anal.*, 23, 74–113.
- Moakher, M. & Maddocks, J. (2005). A double-strand elastic rod theory. *Arch. Rational Mech. Anal.*, 177, 53–91.
- Montgomery, R. (1993). Gauge theory of the falling cat. *Fields Inst. Comm.*, 1, 193–218.
- Morita, H. & Kaneko, K. (2004). Self-organized bottleneck in energy relaxation. *Europhysics Letters*, 66, 198–204.
- Nakahara, M. (2003). *Geometry, Topology and Physics*. IOP Publishing.
- Neukirch, S., Goriely, A., & Hausrath, A. (2008). Chirality of coiled coils: Elasticity matters. *Phys. Rev. Lett.*, 100, 038105.
- Nizette, M. & Goriely, A. (1999). Towards a classification of Euler-Kirchhoff filaments. *J. Math. Phys.*, 40, 2830–2866.
- Noda, K., Ishida, K., Kubono, A., Horiuchi, T., & Yamada, H. (2003). Remanent polarization of evaporated films of vinylidene fluoride oligomers. *J. Appl. Phys.*, 93, 2866–2870.
- Simó, J., Marsden, J., & Krishnaprasad, P. (1988). The Hamiltonian structure of nonlinear elasticity: The material and convective representations of solids, rods, and plates. *Arch. Rat. Mech. Anal.*, 104, 125–183.
- Skyrme, T. (1961). A nonlinear field theory. *Proc. Roy. Soc. Series A, Math. Phy. Sci.*, 260, 127–138.
- Vialard, F.-X. (2009). *Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework: From the Discontinuous Image Matching Problem to a Stochastic Growth Model*. Ph.D. Thesis, École Normale Supérieure de Cachan.

REFERENCES

- Whitam, G. (1966). Nonlinear dispersion of water waves. *J. Fluid Mech.*, 27, 399–412.
- Yoshimura, H. & Marsden, J. (2007). Reduction of Dirac structures and the Hamilton-Pontryagin principle. *Rep. Math. Phys.*, 60(3), 381–426.

INDEX

- adjoint bundle, 97
- advected quantity, 23
- advected variables, 63
- advection relations, 75
- affine action, 23, 62, 64
- body
 - representation, 16
- charged molecular strand, 14
 - equations, 55
- classical
 - perspective, 20, 22, 42, 61
- classical-covariant duality, 163
- cocycle, 24, 47, 64, 66, 70, 73
- commutative diagram, 96
- conservation law, 31, 56, 79–81
- convective
 - representation, 16, 17, 21, 28, 29, 151, 152
- covariant
 - perspective, 20, 25, 84
- covariant derivative, 85, 130
- directors, 28
- distortion, 175
 - coupling, 175
 - curvature, 175
- dynamical equation, 44
- energy
 - kinetic, 28
 - potential, 30, 40
- Euler-Lagrange
 - equations, 22, 26, 61
- Euler-Poincaré
 - covariant, 95
- Euler-Poincaré, 35, 39, 62, 63
 - affine, 20, 23, 43, 47, 63, 65, 67
 - classical, 119
 - covariant, 118, 141, 142, 149
 - equations, 36
- Euler-Poincaré
 - affine, 66
- fibre bundle, 101
 - reduction, 113
- filament
 - elastic, 59, 86
 - fixed, 81
- first jet
 - bundle, 25, 96
 - extension, 25, 47
- fixed filament, 34
- geometrically exact rod, 17
- graph, 194
- Hamilton's equations, 22
- Hamilton-Pontryagin, 39, 71, 79
- hat map, 22
- heavy top, 33, 63
- ideal fluid, 63
- interparticle potential, 41

INDEX

- Kelvin-Noether, 154, 156
- kinematic relation, 44
- kinematic relations, 44
- Kirchhoff
 - representation, 28
- Kirchhoff theory, 30, 76
- Lagrange-Poincaré, 112, 117, 126
 - classical, 141, 142, 150, 168
 - covariant, 43, 47, 140–142, 149
 - equation, 106
 - subgroup, 100, 119
- matching problem, 166
- Maurer-Cartan, 104
- mechanical connection, 176
- molecular strand, 68
- multisymplectic, 112
- nonlocal
 - contribution, 55
 - dynamics, 52
 - interaction, 17, 34, 59, 64
 - potential, 40, 43, 90
- particular cases, 117
- perspective
 - classical, 20, 22, 42, 61
 - covariant, 20, 25, 84
- reconstruction, 144, 147
 - condition, 104, 145
 - relation, 45
- reduced variables, 21
- reduction
 - classical, 113
 - fibre bundle, 113
- representation
 - body, 16
 - convective, 16, 17, 21, 28, 29, 151, 152
 - Kirchhoff, 28
 - spatial, 16, 29, 151, 153
- rigid body, 63
- rigid charge conformation, 28, 39
- section, 96
 - critical, 126
 - horizontal, 151
 - morphing, 196
 - reduced, 128
- semidirect product, 24, 46, 65, 73
 - action, 42
 - reduction, 23
- shape space, 101
- spatial
 - representation, 16, 29, 151, 153
- strand generalisation, 46, 73, 79, 92, 105
- total derivative, 58
- transposition identities, 41
- un-reduction, 168
- uniformly parameterised morphing, 192
- variational principle, 48, 53, 86, 90