GEOMETRIC DYNAMICS OF OPTIMIZATION[*]

FRANÇOIS GAY-BALMAZ[†], DARRYL D. HOLM[‡], AND TUDOR S. RATIU[§]

Abstract. This paper investigates a family of dynamical systems arising from an evolutionary re-interpretation of certain optimal control and optimization problems. We focus particularly on the application in image registration of the theory of *metamorphosis*. Metamorphosis is a means of tracking the optimal changes of shape that are necessary for registration of images with various types of data structures, without requiring that the transformations of shape be diffeomorphisms, but penalizing them if they are not. The possibilities of this approach are just beginning to be developed. In particular, metamorphosis and its related variants in the geometric approach to control and optimization can be expected to produce many exciting opportunities for new applications and analysis in geometric dynamics.

Key words. Optimal control, optimization, geometric mechanics, symmetry.

AMS subject classifications. 34H05, 49J20, 49J27.

1. Introduction

With the advent of new devices capable of seeing objects and structures not previously imagined, the realm of science and medicine has been extended in a multitude of different ways. The impact of this technology has been to generate new challenges associated with the problems of formation, acquisition, compression, transmission, and analysis of images. These challenges cut across the disciplines of mathematics, physics, computational science, engineering, biology, medicine, and statistics.

For example, in computational anatomy (CA) biomedical images are compared quantitatively by calculating the "distance" between them, along a path that is *optimal* in transforming one such image to another. The optimal path is traversed along a curve of deformations in the group of smooth invertible maps with smooth inverses (i.e., the diffeomorphisms) and it is governed by a partial differential equation (PDE) called the EPDiff equation, which takes its simplest form as [42, 43, 72, 64]

$$\frac{d}{dt}\frac{\delta\ell}{\delta\xi} = \mp \operatorname{ad}_\xi^* \frac{\delta\ell}{\delta\xi}. \tag{1.1}$$

The term EPDiff is an abbreviation for 'the Euler-Poincaré equation on the group of diffeomorphisms'. EPDiff arises from Hamilton's principle $\delta S = 0$ for $S = \int \ell(\xi)\,dt$ for a Lagrangian $\ell(\xi) : \mathfrak{X} \to \mathbb{R}$ defined on the Lie algebra of vector fields $\mathfrak{X}$ with Lie bracket $-[\xi,\eta] = \operatorname{ad}_\xi\eta : \mathfrak{X} \times \mathfrak{X} \to \mathfrak{X}$. The dual operation is $\operatorname{ad}_\xi^*\mu : \mathfrak{X} \times \mathfrak{X}^* \to \mathfrak{X}^*$. The sign in equation (1.1) is $+$ (resp. $-$) for left (resp. right) invariant vector fields.

The EPDiff equation (1.1) governs geodesic flow on the group of diffeomorphisms, with respect to any prescribed metric. This flow from one shape to another also has an evolutionary interpretation that invites ideas from the analysis of evolutionary equations. In particular, the *momentum map* for EPDiff, identified first in [17] and

explained more completely in [40], yields the canonical Hamiltonian formulation of the dynamics of the singular evolutionary solutions of EPDiff. Moreover, in an optimization sense, this momentum map also provides a complete representation of the landmarks and contours (outlines) of images to be matched, in terms of the canonical positions and momenta associated with the evolutionary interpretation [44]. In addition, it provides a natural strategy for finding the optimal path between two configurations of either landmarks or contours [72]. Thus, the momentum map (a concept from Hamiltonian systems) is crucial in the construction of an isomorphism between the data structures used in the optimal matching of images and the evolutionary singular solutions of the EPDiff equation. This isomorphism has already suggested new dynamical paradigms for CA, as well as new strategies for assimilation of data in other image representations, for example, as gray-scale densities [47, 72]. The converse benefit may also develop, in which methods of optimal control and optimization of data assimilation used in image matching for CA may suggest new strategies for investigating dynamical systems of evolutionary PDE. In short, the variational formulations, Lie symmetries and associated momentum maps encountered in applications of EPDiff have led to a convergence in the analysis of both its evolutionary properties and its optimization equations.

This paper focuses on the evolutionary aspects of the PDE that are summoned by adopting a dynamical interpretation of the optimal control and optimization methods used in the registration of various types of images. The paper does not perform any applications of optimization methods to image registration, nor does it develop any numerical algorithms for making such applications. Instead, the paper re-interprets the endeavor of image registration from a dynamical systems viewpoint. In particular, as we shall explain, a recent development in Large Deformation Diffeomorphic Metric Mapping (LDM), in an approach for image registration called *metamorphosis*[1] [61, 68, 47], introduces a new type of evolutionary equation that may be called *optimization dynamics*. In following this line of reasoning, the geometric mechanics approach for evolutionary PDE provides a framework that we hope will inform both optimization and dynamics.The primary example in the line of reasoning leading to optimization dynamics is the EPDiff equation [42, 43, 72].

A brief history of the EPDiff equation. The EPDiff equation (1.1) stems from the recognition by Arnold in [1] that incompressible fluid dynamics could be characterized as geodesic flow in the group of volume preserving diffeomorphisms, with respect to the kinetic energy metric (L^2 norm of the fluid velocity). A few years later, the one-dimensional compressible version of EPDiff reappeared as the dispersionless limit of the Camassa-Holm (CH) equation [17]. The CH equation is a completely integrable evolution equation for shallow water waves, whose soliton solutions develop sharp peaks in the dispersionless limit. Its peaked soliton solutions (peakons) correspond to concentrations of momentum into delta-function singularities and are solutions of EPDiff in one dimension with the H^1 kinetic energy metric. Slightly later, the incompressible version of EPDiff with the H^1 kinetic energy metric was generalized to higher dimensions in [42, 43] by using its symmetry-reduced variational principle, and was interpreted as Euler's fluid equations, averaged following Lagrangian particle trajectories. This interpretation soon led to the introduction of viscosity and some interesting applications of the resulting viscous equations as a

---

[1] Although the term "metamorphosis" has a precise mathematical definition that will be given below, it also satisfies its proper dictionary definition, as "a change of physical form, structure, or substance". This paper interprets the change as a type of evolution.

turbulence model by Chen et al. [20, 21].

Around the same time, EPDiff arose independently in a completely different context. Namely, it arose as the governing equation in the optimization problem for Large Deformation Diffeomorphic Metric Mapping (LDM) in image registration [66, 67, 70]. The recognition that EPDiff was arising in these two different contexts provided a fruitful opportunity for dual interpretations of the solutions of the same equation. In particular, the "peakons" of the CH equation in the water wave context were soon recognized to be the "landmarks" in images in the LDM context [44]. Since then, the two types of problems have continued their optimization-dynamics interplay and have been found to inform each other, while also showing intriguing differences and similarities that arise in their dual formulations as initial value problems on one hand and boundary value problems on the other. In particular, the concept of symmetry reduction and momentum maps from geometric mechanics, that had previously been applied so effectively in fluid dynamics [1] and shallow water soliton theory [17], has recently been recognized also as a unifying approach for developing multi-mode LDM methods for images whose data structure may comprise arbitrary tensors, or tensor densities [16]. This is a rich and rapidly developing area of science, for which a complete literature review would be beyond our scope here.

The convergence of these two independent endeavors has led to dual interpretations of the same equation and the same key ideas in different but complementary contexts. This convergence is fascinating, and we continue our investigation of it here. In the present paper, we emphasize the dynamical interpretations of the equations and approaches that are applied in optimal image matching. This is not to say that we solve optimal matching problems for images at all in this paper. Rather, being cognizant of the ideas and variational formulations underlying the optimal matching approach, we shall apply these formulations to study certain classes of equations that arise in the problem of image registration, not from the viewpoint of optimization, but rather from the evolutionary viewpoint of *geometric mechanics* [45, 54].

The geometric mechanics approach emphasizes Lie group actions on manifolds, momentum maps, and reduction by symmetry. This approach leads in the present paper to an understanding of certain classes of control and optimization problems as systems of evolutionary equations. In particular, the Lie symmetry ideas underlying the process of optimal image assimilation known as *metamorphosis* [61, 68, 47] in combination with the evolutionary geometric mechanics viewpoint leads the family of EPDiff equations into the realm of *optimization dynamics*. Optimization dynamics extends the previous association of image matching ideas with soliton theory [44] to produce new results, such as the derivation and re-interpretation of the two-component CH system (CH2) as an equation for the dynamics of metamorphosis of gray-scale images [47]. The CH2 system is a completely integrable evolutionary system of equations that was recently discovered using isospectral methods for solitons [22]. Its inverse scattering transform is discussed in [38]. Recognizing that some systems of equations arising in optimization dynamics for image analysis may be associated with soliton theory raises many questions about the mathematical properties of these systems and their solutions, particularly when the equations are nonlocal. For example, the initial value problems for some of the nonlocal equations obtained in optimization dynamics investigated here allow *emergent* singular solutions, in which the evolution of a smooth, spatially confined, initial condition becomes singular by concentrating itself into delta function distributions. In particular, EPDiff has that property and so does the corresponding system of equations for the optimization dynamics of meta-

morphosis. See [42, 43, 45] and [71, 72], respectively, for further discussions of EPDiff from the different but complementary viewpoints of geometric mechanics and image matching.

1.1. LDM approach, EPDiff, and momentum maps. The LDM approach is based on minimizing the sum of a time-integrated *kinetic energy metric* whose value defines the length of an optimal deformation path, plus a *penalty norm* that ensures an acceptable tolerance in image mismatch. The main reason why matching cannot be exact is that the sets of level curves of two generic images are rarely topologically equivalent, thus the images cannot be matched exactly.

LDM approaches were introduced and systematically developed in Trouvé [66, 67], Dupuis et al. [25], Joshi and Miller [48], Miller et al. [61, 60], Beg [5], and Beg et al. [6]. The LDM approaches of those papers are based on Grenander's *deformable template* paradigm for image registration [32]. Grenander's paradigm, in turn, is a development of a biometric strategy introduced by D'Arcy Thompson [65] of comparing a template image I_0 to a target image I_1 by finding a smooth invertible transformation of coordinates that maps one image to the other. This transformation is assumed to belong to a Lie group G of diffeomorphisms that acts on the set of templates containing I_0 and I_1. The effect of the transformation on the data structure that is encoded in the set of templates is called the *action* of the Lie group G on the set of images. As discussed below, the optimal path in the transformation group is the one that costs the least in time-integrated 'kinetic energy' for a given tolerance. This concept of optimization summons a control theory approach into the analysis and registration of images. For a comprehensive presentation of the mathematical foundations of the LDM approach, see [71].

In applications of the LDM approach, the optimal transformation path is often sought by using a variational optimization method such as the one developed in [25, 66, 67]. Using this method, the optimal path for the matching transformation in this problem is obtained from a gradient-descent algorithm based on the Euler-Lagrange equation arising from stationary balance between kinetic energy and tolerance. This gradient-descent approach does indeed determine an optimal matching path. However, from the viewpoint of dynamical systems theory, it misses the following potentially interesting question:

> *What information and perspective may be obtained by interpreting the Euler-Lagrange equations associated to the LDM approach from a dynamical systems viewpoint?*

The answer to this question may be sought by interpreting the variational optimization method in the LDM approach as a form of Hamilton's principle. Hamilton's principle for the variational construction of optimal paths with minimal kinetic energy for a given tolerance in image mismatch yields an associated set of Euler-Lagrange equations that may then be given an evolutionary interpretation. The optimal solutions of these equations have been investigated as evolutionary motion on the Lie group of diffeomorphisms in the *absence* of additional penalty terms by Arnold [1, 2], Holm et al. [42, 43], Marsden and Ratiu [54], and for the particular application to template matching in Miller et al. [60]. As mentioned earlier, the optimal paths in these cases are geodesics with respect to the metric provided by the kinetic energy. The kinetic energy for LDM is invariant under right translations on the diffeomorphism group. Reducing Hamilton's principle with respect to this symmetry and then invoking the Euler-Poincaré theory applied to diffeomorphisms produces the EPDiff

equation (1.1) [42, 43].

The solution of the EPDiff equation yields the spatial representation of the geodesic velocity, i.e., the tangent vector to the optimal path of deformations along which the minimal distance from one image to another is measured. The geodesics themselves may be obtained from the solutions of EPDiff for the velocity by a reconstruction process that inverts the previous reduction by symmetry after the solution to the EPDiff equation for velocity has been obtained. This is analogous to the reconstruction process in classical mechanics that recovers the symmetry coordinate conjugate to a conserved momentum as the final step in the solution, after the other degrees of freedom have been determined in the reduced space.

Composing the evolutionary solutions of EPDiff with the reconstruction process provides an important representation of diffeomorphisms that relates the endpoint of a geodesic to the initial value for *momentum* in the EPDiff equation. This relation is the momentum representation of the deformation. The long-time existence of this representation is based on conservation by EPDiff of the kinetic energy norm, which may be chosen so that its boundedness affords enough smoothness on the velocities to ensure the long-time existence of solutions of EPDiff. In this case, EPDiff admits emergent weak momentum solutions; for example, delta-function distributions of momentum that emerge from smooth, spatially confined initial conditions [17, 40]. This singular behavior is well understood analytically only in certain one-dimensional cases. In particular, it is understood for the completely integrable case of the Camassa-Holm equation; see, e.g., [52, 62] and references therein.

The EPDiff equation is of central importance in computational anatomy [72]. This is because the optimal paths sought by LDM on the image template space defined on a manifold M are inherited from the geodesics on $\mathrm{Diff}(M)$, the Lie group of diffeomorphisms acting on the manifold M. These, in turn, are governed by EPDiff. Consequently, any solution of the LDM problem for optimal geodesics must involve EPDiff [72]. Conversely, solving the LDM problem directly produces the momentum representation of the optimal diffeomorphism. The momentum representation arising from this evolutionary interpretation is then available for analyzing anatomical data sets. In any case, despite the disparate forms that the geodesic equations may take for the various data structures in the various types of images, all of them are instances of EPDiff with the corresponding representation for momentum. The specific representation for momentum in terms of the image data structure in a given case is called the *momentum map*. The momentum map for images is another dynamical systems concept that emerges as a central feature in this paper. The EPDiff equation and its associated momentum map for various image data structures are discussed in Section 8.4.

An interesting example of the momentum map relating solutions of LDM to solutions of EPDiff arises for the case of *landmark* data structure, in which the momentum is singularly concentrated at points. The relation between these singular geodesic solutions and evolutionary soliton solutions, called *peakons* for a shallow water wave equation introduced in Camassa and Holm [17], has been examined in the context of computational anatomy in Holm et al. [44]. A numerical analysis of the stability of these equations is also given in McLachlan and Marsland [57]. See also Micheli [58] for other recent developments involving the curvature of the space of landmark shapes. Holm and Marsden [40] explain that two independent momentum maps for EPDiff are available in the case that the image data structure comprises the manifold $\mathrm{Emb}(S^1, \mathbb{R}^2)$ of embedded closed curves (embedded images of S^1) in the plane $\mathbb{R}^2$.

The left action of the group of diffeomorphisms $\mathrm{Diff}(\mathbb{R}^2)$ of the plane deforms the curve by a smooth invertible transformation of the coordinate system in which it is embedded, while leaving the parameterization of the curve invariant. The right action of the group of diffeomorphisms $\mathrm{Diff}(S^1)$ of the circle corresponds to smooth invertible reparameterizations of the domain S^1 of the coordinates of the curve. In this case, one momentum map corresponds to action from the left by the diffeomorphisms on $\mathbb{R}^2$, the other to their action from the right on the embedded curves. Optimal control and reparameterization methods for matching closed curves in the plane using these two momentum maps for the space of closed curves in the plane have recently been developed in Cotter and Holm [24].

In summary, LDM image analysis is based on optimization methods that are formulated as boundary value problems. However, the re-interpretation of their governing equations as evolutionary systems by using symmetry reduction of the corresponding Hamilton's principle allows various concepts from dynamical systems theory to be profitably applied in the solution and interpretation of image analysis problems. Thus, the transfer of concepts and ideas between these two fields in the context of image registration has the potential to enrich them both.

1.2. Distributed optimization dynamics, or evolutionary metamorphosis. As we have been discussing, the paper focuses on the geometric dynamics interpretation of the optimization problems designed for image registration. However, rather than concentrating on the development of *solutions* of optimization problems, the treatment here focuses on the *dynamics* that are produced in applying the method of reduction by Lie group symmetry to families of optimization problems posed in a geometric setting. This is a new arena for geometric dynamics and several new departures are being taken. Among these new departures is the investigation of the evolutionary dynamics that arises when distributed or nonlocal *penalties* are imposed in Hamilton's principle, rather than local constraints. For lack of a better name, we call this sort of problem *distributed optimization dynamics*. It is the evolutionary counterpart of the *metamorphosis* approach in imaging science [61, 68, 47], which, in turn, is a modification and development of LDM that allows the evolution $n(t)$ of the image template to deviate from pure deformation. That is, metamorphosis only *penalizes* the spatial average of the deviation away from the infinitesimal action of the vector fields on an image manifold, rather than *enforcing* it as a local pointwise constraint. This approach, in turn, modifies the EPDiff equation and thereby introduces a wealth of new structure and new examples that we shall investigate in this paper.

An explicit comparison for the case that the image templates are gray-scale density distributions may help to understand the difference between the LDM approach and the metamorphosis approach.

LDM approach: Given the source and target templates for the images characterized as scalar densities n_0 and n_T at the initial time $t=0$ and the final time $t=T$, respectively, minimize the quantity

$$\int_0^T \ell(u(t))dt + \frac{1}{2\sigma^2}\|n_0\circ\eta_T^{-1} - n_T\|_{L^2}^2 \tag{1.2}$$

over the time dependent vector field $u(t)$, where η_T is the flow of $u(t)$ evaluated at time $t=T$, and the formula

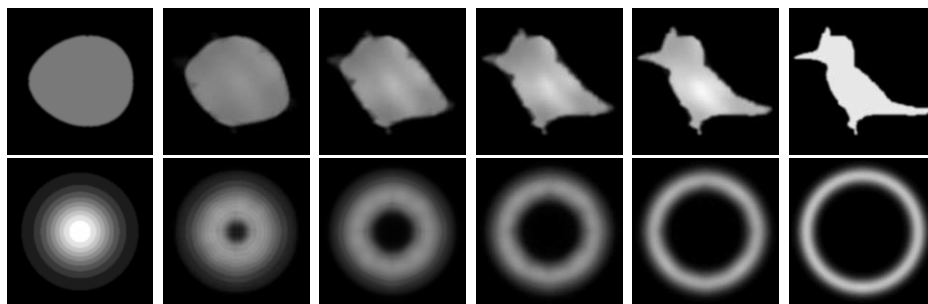$$\dot{n}(t) + \mathrm{div}\big(n(t)u(t)\big) = 0$$

FIG. 1.1. *These gray-scale images show optimal metamorphoses between two density distributions with equal total mass from [47]. The optimization approach would compute the distance along the optimal path between between the first and last density in each row. In the evolutionary approach, the optimal* trajectories *for n(t) are computed. The images between the endpoints show snapshots along the optimal path n(t) in each row at intermediate points in time. In particular, the second row shows that metamorphosis allows a change in topology along its optimal path. Our interest focuses on the evolutionary equations for the process of metamorphosis. The dynamical system of metamorphosis equations obtained in registering such gray-scale image densities is given in Section 8 as one of the examples of the general approach. In one dimension, the metamorphosis equations for this class of images comprises a completely integrable Hamiltonian system [47].*

is its infinitesimal action on a smooth density $n(t) = n_0 \circ \eta_t^{-1}$ defined over time $0 \le t \le T$ on the domain of flow.

Metamorphosis approach: Given n_0 and n_T, minimize

$$\int_0^T \left(\ell(u(t)) + \frac{1}{2\sigma^2} \| \dot{n}(t) + \mathrm{div}\big(n(t)u(t)\big) \|_{L^2}^2 \right) dt \qquad (1.3)$$

over time dependent vector field $u(t)$ and scalar densities $n(t)$. As one sees in figure 1.1 for the metamorphosis of shapes characterized as densities, the term "metamorphosis" introduced in [68] for this process can be understood in practice by its ordinary meaning, as "change of shape", such as the gradual and continuous metamorphosis of a tadpole into a frog.

The paper begins by contrasting optimal control problems with distributed optimization problems in a geometric setting. In particular, we discuss the geometric properties of Lie algebra controls acting on state space manifolds. The latter optimal control approach parallels the familiar Clebsch variational formulation of dynamical equations continuum mechanics (e.g., [39]). In fact, continuum mechanics was one of the early paradigms for image registration [70]. The Clebsch variational formulation of continuum mechanics has recently been developed and applied in the study of the dynamical aspects of optimal control problems in a geometric setting (see [29, 37]). Conversely, our concern here is to continue this parallel development by studying the implications for dynamics of the geometric approach to *distributed* optimization problems.

1.3. Main content of the paper.

Context. In [29] a general formulation for a large class of optimal control problems was given. These problems, called *Clebsch optimal control problems*, are associated to an action $\Phi \colon G \times Q \to Q$ of a Lie group G on a manifold Q and to a cost

function $\ell \colon \mathfrak{g} \times Q \to \mathbb{R}$, where $\mathfrak{g}$ denotes the Lie algebra of G. The Clebsch optimal control problem for the curves $\xi(t) \in \mathfrak{g}$ and $n(t) \in Q$ is, by definition,

$$\min_{\xi(t)} \int_0^T \ell(\xi(t), n(t)) dt, \qquad (1.4)$$

subject to the following conditions:

(A) Either $\quad \dot{n}(t) = \xi(t)_Q(n(t)), \quad$ or $\quad$ (A)$'$ $\quad \dot{n}(t) = -\xi(t)_Q(n(t));$

(B) Both $\quad n(0) = n_0 \quad$ and $\quad n(T) = n_T,$

where ξ_Q denotes the infinitesimal generator of the G-action, that is,

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

These optimal control problems comprise abstract formulations of many systems such as the symmetric representation of the rigid body and Euler fluid equations [9, 37], the double bracket equations on symmetric spaces [8], the singular solutions of the Camassa-Holm equation [17], control problems on Stiefel manifolds [13], and others [7, 12].

Optimal control problems on Lie groups have a long history; see [7], [49], and references therein. Some of the earliest papers dealing with such problems are [14] and [33].

Goals of the paper. The first goal of the present paper is to replace the constraints in the Clebsch optimal control problem (1.4) with a penalty function added to the cost function and to obtain in this way a classical (unconstrained) optimization problem. The fundamental idea is to use the constraints to form a quadratic penalty function in order to get the Lagrangian

$$\int_0^T \left(\ell(u, n) + \frac{1}{2\sigma^2} \| \dot{n} \mp u_Q(n) \|^2 \right) dt. \qquad (1.5)$$

We first determine necessary and sufficient conditions characterizing the critical points of this Lagrangian. Taking the time derivative of one of the conditions and using the others leads directly to certain equations of motion. We then show that these equations are naturally obtained by Lagrangian reduction and that they are the Lagrange-Poincaré equations of a Lagrangian function in the material representation that is the sum of the original Lagrangian plus the square of the norm on the velocity vector. This approach links directly to the approach used in [47] in the study of the metamorphosis of shapes. From a variational point of view, one replaces the Hamilton-Pontryagin variational principle in the Clebsch framework

$$\delta \int_0^T \left(\ell(u, n) + \langle \alpha, \dot{n} \mp u_Q(n) \rangle \right) dt = 0$$

by the principle

$$\delta \int_0^T \left(\ell(u, n) + \frac{1}{2\sigma^2} \| \dot{n} \mp u_Q(n) \|^2 \right) dt = 0,$$

which is the basis of *Clebsch distributed optimization*.

This paper traces how the dynamical equations change on moving from constraints (optimal control) to optimization via imposition of a cost, and then on to metamorphosis. Passing from optimal control to optimization preserves the momentum map, but this passage modifies the reconstruction relation. The evolution is no longer only for the momentum map of the reduced Lagrangian. Instead, the momentum canonically conjugate to the velocity on the configuration manifold becomes coupled to the momentum map equations (which are the Euler-Poincaré equations), with coupling constant σ^2.

Another feature of the paper, directly related to the dynamics of our optimization problem, is the description of the equations of motion by Lagrangian and Hamiltonian reduction. In particular, we carry out a certain type of Lagrangian reduction adapted to the problem, that we naturally call *metamorphosis reduction*, since it was directly inspired by the example of the metamorphosis approach to image dynamics [47]. This Lagrangian reduction leads to the expression of the associated variational principles and Hamiltonian structures. In metamorphosis, the optimization problem involves Riemannian structures induced by Lie group actions on themselves and on Lie subgroups by group homomorphisms. This is a rich field whose possibilities are still being developed. In particular, metamorphosis and related variants of the geometric approach to control and optimization can be expected to produce opportunities for new applications and analysis in geometric dynamics.

1.4. Plan of the paper.   In the remainder of the paper, we compare the dynamical equations that arise from optimal control problems with those arising from distributed optimization. This comparison provides several examples of how the two approaches differ and, in particular, how their dynamical equations differ when their variational problem is regarded as Hamilton's principle for the dynamics. Their comparison also identifies the aspects of these approaches that are fundamentally the same.

- Section 2 begins by explaining the dynamical set up for standard optimal control problems treated by the Pontryagin Maximum Principle. Section 2.2 provides several examples illustrating the consequences of applying Lie group controls acting on state manifolds by using the Clebsch framework for optimal control. These examples introduce the *momentum map* for the cotangent-lifted action of the Lie group controls on the state manifold. The cotangent-lift momentum map is a fundamental concept in the application of geometric mechanics methods in the Clebsch framework for optimal control. It turns out that the same momentum map is also the organizing principle for the distributed optimization dynamics introduced in Section 2.3. After establishing this background for our comparison of optimization and dynamical systems methods, Section 1.3 provides an overview of the rest of the paper.

- Section 3 begins by reviewing the Clebsch framework for optimal control problems introduced and studied in [29]. A new class of optimization problems is then introduced, which is the subject of study of this paper. The stationarity conditions are obtained and the associated equations of motion are determined.

- Inspired by the extremum problems presented earlier, Section 4 presents two Lagrangian reduction procedures for Lagrangian functions defined on $T(G \times Q)$, where G is a Lie group acting on the manifold Q.

- These reduction methods are used in Section 5 to rederive the equations of

motion that were found in Section 3.

- Hamiltonian reduction is carried out in Section 6. As before, there are two reduction methods and, in the case of a representation, one of them leads to Lie-Poisson equations with a symplectic cocycle on the dual of a larger semidirect product Lie algebra.

- In Section 7 we apply these Hamiltonian reduction methods to the optimization problems introduced earlier.

- Section 8, by far the longest of the paper, presents a panorama of examples for the purpose of illustrating the breadth of the applications of the general theory. We begin by studying examples where G is represented on a vector space. The concrete examples treated are the heavy top and a class of problems using the adjoint representation. For example, we find a modification of the pair of double bracket equations studied in [8], [9]. Next, we study optimization problems associated to affine actions. Actions by group multiplication is the next topic. The concrete examples include Euler's equations for an ideal incompressible homogeneous and for a barotropic fluid. The N-dimensional Camassa-Holm equation is presented from this optimization point of view, inspired by the construction of singular solutions. Finally, the optimization problem is used to obtain the equations of metamorphosis dynamics for use in computational anatomy.

- Section 9 briefly summarizes the paper and gives an outlook for future work.

2. Review of optimal control problems

2.1. Definitions.   We begin by recalling the definition of optimal control problems.

DEFINITION 2.1 (Optimal control problems). *A standard optimal control problem comprises:*

- *a differentiable manifold Q on which state variables $n \in Q$ evolve in time t during an interval $I = [0,T]$ along a curve $n : I \to Q$ from $n(0) = n_0$ to $n(T) = n_T$, with specified values $n_0, n_T \in Q$;*

- *a vector space U of control variables $u \in U$ whose time dependence $u : I \to U$ is at our disposal to affect the evolution $n(t)$ of the state variables;*

- *a smooth map $F : Q \times U \to TQ$ such that $F(\cdot, u) : Q \to TQ$ is a vector field on Q for any $u \in U$ whose associated evolution equation[2]*

$$\dot{n} = F(n, u) \tag{2.1}$$

relates the unknown state and control variables $(n(t), u(t)) : I \to Q \times U$;

- *a cost functional depending on the state and control variables*

$$S := \int_0^T \ell(u(t), n(t)) \, dt, \tag{2.2}$$

---

[2]The over-dot notation in $\dot{n}$ means time derivative. Several forms of time derivative appear in applications and the meaning should be clear from the usage. Besides the over-dot notation, we shall use the equivalent notation d/dt to mean either partial or ordinary time derivative in the abstract formulas, as needed in the context. For fluids, we shall also use ∂_t for the Eulerian time derivative at fixed spatial location. Finally, the covariant time derivation on a Riemannian manifold will be denoted as D/Dt.

subject to the prescribed initial and final conditions, at $n(0) = n_0$ and $n(T) = n_T$. The integrand $\ell : Q \times U \to \mathbb{R}$, called the Lagrangian, is assumed to be C^1 on $Q \times U$.

The goal of the optimal control problem is to find the evolution $(n(t), u(t))$ of the state and control variables such that S is minimal subject to the prescribed dynamics (2.1) and the prescribed initial and final conditions $n(0) = n_0$, $n(T) = n_T$.

The coupling between the control and state variables may be made explicit by using the pairing $\langle \cdot, \cdot \rangle_Q : T^*Q \times TQ \to \mathbb{R}$ and a Lagrange multiplier $\alpha \in T^*Q$ that imposes the state system as a constraint on the cost functional,

$$S_c := \int_0^T \left[\ell(u,n) + \langle \alpha, \dot{n} - F(n,u) \rangle_Q \right] dt. \tag{2.3}$$

This is a consequence of the well-known *Pontryagin maximum principle* [7, 49].

The variable $\alpha \in T^*Q$ is called a *costate variable*. We now compute the equations associated to the variational principle $\delta S_c = 0$. For simplicity, we suppose here that the state manifold Q is a vector space, say W. In this case the cotangent space is $T^*W = W \times W^*$ and the costate variable is of the form $\alpha = (n,p) \in W \times W^*$. The stationary variations of the constrained cost function S_c in (2.3) yield

$$0 = \delta S_c = \int_0^T \left[\left\langle \frac{\partial \ell}{\partial n} - \left(\frac{\partial F}{\partial n} \right)^T p - \dot{p}, \delta n \right\rangle_W + \left\langle \frac{\delta \ell}{\delta u} - \left(\frac{\delta F}{\delta u} \right)^T p, \delta u \right\rangle_U \right.$$
$$\left. + \langle \delta p, \dot{n} - F(n,u) \rangle_W \right] dt + \langle p, \delta n \rangle_W \Big|_0^T,$$

where $\langle \cdot, \cdot \rangle_U : U^* \times U \to \mathbb{R}$ denotes the duality pairing for the control vector space U and the symbols $\frac{\delta \ell}{\delta n}$, $\frac{\delta F}{\delta n}$, $\frac{\partial \ell}{\partial n}$, and $\frac{\partial F}{\partial u}$ denote the functional and partial derivatives of the functions ℓ and F, respectively.

Stationarity in the variations δu gives a relation that determines the controls u in terms of the state and costate variables, n and α, respectively, while stationarity in the variations $(\delta n, \delta \alpha)$ determines the evolution equations for the state and costate variables that minimize the cost function S. Since the values of n at the endpoints in time are fixed, δn vanishes at the endpoints. We thus get the stationarity conditions

$$\frac{\delta \ell}{\delta u} = \left(\frac{\delta F}{\delta u} \right)^T p, \qquad \dot{n} = F(n,u), \qquad \dot{p} = \frac{\delta \ell}{\delta n} - \left(\frac{\delta F}{\delta n} \right)^T p.$$

REMARK 2.1. Although we shall confine our considerations to the Lagrangian description, we point out that the relation to the *Pontryagin Maximum Principle* in the Hamiltonian description is obtained via the Legendre transformation of the integrand in the cost functional given by (2.3) which, for each point u in the control space U, defines the corresponding Hamiltonian $H_u : T^*Q \to \mathbb{R}$ by

$$H_u(\alpha_n) = \langle \alpha_n, F(n,u) \rangle_Q - \ell(n,u). \tag{2.4}$$

The notation α_n for a covector in T^*Q means that it belongs to the fiber T_n^*Q of the cotangent bundle. For more information about the Hamiltonian approach to geometric optimal control theory and the Pontryagin Maximum Principle, see [7, 49]. In particular, the Pontryagin principle admits abnormal solutions that are not

included in the Clebsch formulation. Very few results are known about the existence of strict abnormal minimizers for these problems [4]. One might conjecture that abnormal solutions do not arise in optimization dynamics, regarded as an initial value problem. However, we have no results in this matter.

2.2. Examples: Lie group controls acting on state manifolds.   As an example that illustrates the theory developed in this paper, we consider the case of continuum mechanical systems with advected quantities; see Section 6 in [42]. In this case, the state manifold M is a vector subspace V^* of $\mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, the tensor field densities on a manifold $\mathcal{D}$. For example, in nonlinear elasticity, the stress is a tensor field density. We will denote by $a \in V^*$ these tensor field densities. The group $\mathrm{Diff}(\mathcal{D})$ of all diffeomorphisms of the manifold $\mathcal{D}$ acts on V^* by pull back, that is,

$$a \mapsto \eta^* a, \quad \text{for all} \quad \eta \in \mathrm{Diff}(\mathcal{D}).$$

It is thus a *right* representation of $\mathrm{Diff}(\mathcal{D})$ on $\mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$. We consider here the group $\mathrm{Diff}(\mathcal{D})$ of diffeomorphism as an infinite dimensional Lie group (either formally or in some Fréchet sense) whose Lie algebra is given by vector fields $\mathbf{v} \in \mathfrak{X}(\mathcal{D})$. The right action of the Lie algebra $\mathfrak{X}(\mathcal{D})$ on V^* is given by the Lie derivative

$$\left. \frac{d}{dt} \right|_{t=0} \exp(t\mathbf{v})^* a := \boldsymbol{\mathcal{L}}_{\mathbf{v}} a,$$

where $t \mapsto \exp(t\mathbf{v})$ denotes the flow of $\mathbf{v}$.

Example 1. We present a simple example of an optimal control problem based on the geometric formulation of continuum mechanics described above. In this example, the control space U is the Lie algebra $\mathfrak{X}(\mathcal{D})$ and thus the control variable is a vector field $u := \mathbf{v} \in \mathfrak{X}(\mathcal{D})$. The state manifold Q is the vector space V^* of tensor field densities. The state variable $n := a \in V^*$ is constrained to evolve according to the ODE

$$\dot{a} = F(a, \mathbf{v}) := \boldsymbol{\mathcal{L}}_{\mathbf{v}} a,$$

and one wants to minimize

$$S := \frac{1}{2} \int_0^T \|\mathbf{v}\|_{\mathfrak{g}}^2 dt,$$

where $\| \cdot \|_{\mathfrak{g}}$ is an inner product norm on the Lie algebra $\mathfrak{g} = \mathfrak{X}(\mathcal{D})$. Note that we are in the setting of Definition 2.1 with $M = V^*$ and $U = \mathfrak{X}(\mathcal{D})$. This is an example of a *Clebsch optimal control problem*, as studied from a geometric point of view in [29]. For this class of problems, the vector field F is given by the infinitesimal generator associated to a group action on the state manifold. In the present example, this infinitesimal generator turns out to be the Lie derivative.

According to (2.3), the constrained cost function in this case is

$$S_c = \int_0^T \left(\frac{1}{2} \|\mathbf{v}\|_{\mathfrak{g}}^2 + \langle p, \dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a \rangle_V \right) dt,$$

where $p \in V$ is the costate variable. This is nothing else than the *Clebsch approach* to continuum mechanics; see, e.g., [39]. The variational principle $\delta S_c = 0$ gives the control

$$\mathbf{v} = -(p \diamond a)^{\sharp} \in \mathfrak{g},$$

where $\sharp : \mathfrak{g}^* \to \mathfrak{g}$ is the sharp operator associated to the inner product on $\mathfrak{g}$ and the bilinear operator $\diamond : V \times V^* \to \mathfrak{g}^*$ is defined by

$$\langle p \diamond a, \mathbf{v} \rangle := -\langle \boldsymbol{\mathcal{L}}_{\mathbf{v}} a, p \rangle, \quad \text{for all} \quad p \in V, \quad a \in V^*, \quad \mathbf{v} \in \mathfrak{g}. \tag{2.5}$$

The other stationarity conditions are

$$\begin{cases} \dot{a} + \boldsymbol{\mathcal{L}}_{(p \diamond a)^\sharp} a = 0, \\ \dot{p} - \boldsymbol{\mathcal{L}}^{\mathsf{T}}_{(p \diamond a)^\sharp} p = 0, \end{cases} \tag{2.6}$$

where $\boldsymbol{\mathcal{L}}^{\mathsf{T}}_{\mathbf{v}} p \in V$ is defined by

$$\left\langle a, \boldsymbol{\mathcal{L}}^{\mathsf{T}}_{\mathbf{v}} p \right\rangle = \langle \boldsymbol{\mathcal{L}}_{\mathbf{v}} a, p \rangle, \quad \text{for all} \quad p \in V, \quad a \in V^*, \quad \mathbf{v} \in \mathfrak{g}. \tag{2.7}$$

The Clebsch state-costate equations (2.6) are canonically Hamiltonian with

$$H(a,p) = \frac{1}{2} \|(p \diamond a)^\sharp\|^2_{\mathfrak{g}} = \frac{1}{2} \left\langle p \diamond a, (p \diamond a)^\sharp \right\rangle_{\mathfrak{g}}.$$

As is well known [39], using the cotangent-lift momentum map given by $\Pi = -p \diamond a$ to project the equations (2.6) on T^*M to $\mathfrak{g}^*$ yields the (left) Lie-Poisson bracket on the dual Lie algebra $\mathfrak{g}^*$. Explicitly, this Lie-Poisson bracket is given by

$$\dot{\Pi} = \operatorname{ad}^*_{\delta h / \delta \Pi} \Pi = \operatorname{ad}^*_{\Pi^\sharp} \Pi, \tag{2.8}$$

where the Hamiltonian has the expression

$$h(\Pi) = \frac{1}{2} \left\langle \Pi, \Pi^\sharp \right\rangle_{\mathfrak{g}}. \tag{2.9}$$

Example 2. This example will use the geometric setting of continuum mechanics as described before. However, the control vector space will now be given by $U := \mathfrak{g} \times V^* \ni (\mathbf{v}, \nu)$. We choose the quadratic Lagrangian

$$\ell(\mathbf{v}, \nu) := \frac{1}{2} \|\mathbf{v}\|^2_{\mathfrak{g}} + \frac{1}{2\sigma^2} \|\nu\|^2_{L^2},$$

where $\|\cdot\|_{L^2}$ denotes an L^2 norm on $V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \operatorname{Den}(\mathcal{D})$. As before, the state manifold Q is V^* and the state variable $a \in V^*$ is constrained to evolve as

$$\dot{a} = F(a, \mathbf{v}, \nu) := \boldsymbol{\mathcal{L}}_{\mathbf{v}} a + \nu.$$

Note that the advection law $\dot{a} = \boldsymbol{\mathcal{L}}_{\mathbf{v}} a$ is not imposed. Instead, the penalty term in the Lagrangian introduces the additional term ν into the advection law.

Thus, the constrained action (2.3) becomes in this case

$$S_c = \int_0^T \left(\frac{1}{2} \|\mathbf{v}\|^2_{\mathfrak{g}} + \frac{1}{2\sigma^2} \|\nu\|^2_{L^2} + \langle p, \dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a - \nu \rangle_V \right) dt, \tag{2.10}$$

whose stationary variation results in

$$0 = \delta S_c = \int_0^T \left[\left\langle -\boldsymbol{\mathcal{L}}^{\mathsf{T}}_{\mathbf{v}} p - \dot{p}, \delta a \right\rangle_V + \left\langle \mathbf{v}^\flat + p \diamond a, \delta \mathbf{v} \right\rangle_{\mathfrak{g}} \right.$$

$$+\left\langle \frac{1}{\sigma^2}\nu^\flat - p, \delta\nu\right\rangle_V + \langle \delta p, \dot a - \boldsymbol{\mathcal{L}}_{\mathbf v}a - \nu\rangle_V\bigg]dt + \langle p, \delta a\rangle_V\bigg|_0^T,$$

where the flat operators $\flat:\mathfrak{g}\to\mathfrak{g}^*$ and $\flat:V^*\to V$ are associated to the inner products on $\mathfrak{g}$ and V^*, respectively. Here the endpoint terms vanish because the values of a at the endpoints in time are fixed. According to the variational formula for δS_c, the cost functional in (2.10) is optimized when the controls satisfy

$$\mathbf v = -(p\diamond a)^\sharp \in \mathfrak{g} \qquad \text{and} \qquad \nu = \sigma^2 p^\sharp \in V^*, \tag{2.11}$$

in which the sharp maps are the inverses of the flat maps defined above. For the controls $(\mathbf v,\nu)\in\mathfrak{g}\times V^*$, the state and costate variables $(a,p)\in V^*\times V$ evolve according to the closed system

$$\begin{cases} \dot a + \boldsymbol{\mathcal{L}}_{(p\diamond a)^\sharp}a = \sigma^2 p^\sharp, \\ \dot p - \boldsymbol{\mathcal{L}}_{(p\diamond a)^\sharp}^\mathsf{T}p = 0. \end{cases} \tag{2.12}$$

These are Hamilton's canonical equations for the Hamiltonian

$$H(p,a) = \frac{1}{2}\big\langle (p\diamond a),(p\diamond a)^\sharp\big\rangle_\mathfrak{g} + \frac{\sigma^2}{2}\langle p, p^\sharp\rangle_V. \tag{2.13}$$

REMARK 2.2. Thus, the evolution of the state a and costate p variables occurs by the corresponding Lie derivative actions of the vector field $(p\diamond a)^\sharp\in\mathfrak{g}=\mathfrak{X}(\mathcal{D})$ calculated by applying the sharp map $\sharp$ to raise indices on the cotangent momentum map $(a,p)\in V^*\times V = T^*V^*\mapsto \mathbf{J}(a,p) = -p\diamond a\in\mathfrak{g}^*$ of the cotangent-lifted action.

The evolution of the momentum $V^*\times V\to\mathfrak{g}^*$ itself is the last formula to be found, just as in the Clebsch approach [39].

PROPOSITION 2.2. *Denote the momentum map of the cotangent-lifted action by*

$$\Pi := -p\diamond a,$$

and its dual vector field by

$$\mathbf v := -(p\diamond a)^\sharp = \Pi^\sharp.$$

Then the state and costate equations (2.12) *imply the following Euler-Poincaré equation for the evolution for the momentum map:*

$$\dot\Pi = -\boldsymbol{\mathcal{L}}_{\mathbf v}^*\Pi - \sigma^2 p\diamond p^\sharp, \tag{2.14}$$

where the operator $\boldsymbol{\mathcal{L}}_{\mathbf v}^*:\mathfrak{g}^*\to\mathfrak{g}^*$ is defined by $\langle\boldsymbol{\mathcal{L}}_{\mathbf v}^*\Pi,\mathbf u\rangle := \langle\Pi,[\mathbf v,\mathbf u]_{JL}\rangle$ for any $\mathbf u,\mathbf v\in\mathfrak{g} = \mathfrak{X}(\mathcal{D})$ and $\Pi\in\mathfrak{g}^* = \Omega^1(\mathcal{D})\otimes\mathrm{Den}(\mathcal{D})$, and where $[\mathbf v,\mathbf u]_{JL} = \boldsymbol{\mathcal{L}}_{\mathbf v}\mathbf u$ denotes the standard Lie bracket of vector fields.

Proof. The proof proceeds by a direct calculation. In the computation below we use the standard Jacobi-Lie bracket of vector fields $[X,Y]_{JL}(f) = X(Y(f)) - Y(X(f))$ for any $f\in C^\infty(\mathcal{D})$. For a fixed Lie algebra element $Z\in\mathfrak{g} = \mathfrak{X}(\mathcal{D})$, we compute

$$\begin{aligned} \left\langle\dot\Pi, Z\right\rangle &= -\langle\dot p\diamond a + p\diamond\dot a, Z\rangle \\ &= \langle\dot p, \boldsymbol{\mathcal{L}}_Z a\rangle + \langle p, \boldsymbol{\mathcal{L}}_Z\dot a\rangle \end{aligned}$$

$$= -\left\langle \boldsymbol{\mathcal{L}}_{\mathbf{v}}^{\mathsf{T}}p, \boldsymbol{\mathcal{L}}_Z a \right\rangle + \left\langle p, \boldsymbol{\mathcal{L}}_Z \boldsymbol{\mathcal{L}}_{\mathbf{v}} a \right\rangle + \sigma^2 \left\langle p, \boldsymbol{\mathcal{L}}_Z p^{\sharp} \right\rangle$$

$$= \left\langle p, \boldsymbol{\mathcal{L}}_{[Z,\mathbf{v}]} a \right\rangle + \sigma^2 \left\langle p, \boldsymbol{\mathcal{L}}_Z p^{\sharp} \right\rangle$$

$$= -\left\langle p \diamond a, [Z,\mathbf{v}] \right\rangle - \sigma^2 \left\langle p \diamond p^{\sharp}, Z \right\rangle$$

$$= -\left\langle \Pi, \boldsymbol{\mathcal{L}}_{\mathbf{v}} Z \right\rangle - \sigma^2 \left\langle p \diamond p^{\sharp}, Z \right\rangle$$

$$= -\left\langle \boldsymbol{\mathcal{L}}_{\mathbf{v}}^{*} \Pi, Z \right\rangle - \sigma^2 \left\langle p \diamond p^{\sharp}, Z \right\rangle,$$

which proves the proposition.   □

Note that we identified the dual space $\mathfrak{X}(\mathcal{D})^{*}$ with one-form densities $\Omega^1(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$. If $\Pi = \alpha \otimes m \in \Omega^1(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, then the pairing of $\alpha \otimes m$ with $\mathbf{v} \in \mathfrak{X}(\mathcal{D})$ is given by

$$\langle \alpha \otimes m, \mathbf{v} \rangle = \int_{\mathcal{D}} (\alpha \cdot \mathbf{v}) m,$$

where $\alpha \cdot \mathbf{v}$ is the standard contraction of a one-form with a vector field.

REMARK 2.3 (Lie algebra formulation of the equations). Recall the the Lie algebra bracket $[\mathbf{u}, \mathbf{v}] = \mathrm{ad}_{\mathbf{u}} \mathbf{v}$ on $\mathfrak{g}$ is minus the Lie bracket of vector fields, that is,

$$[\mathbf{u}, \mathbf{v}] = -[\mathbf{u}, \mathbf{v}]_{JL} := -(\mathbf{u} \cdot \nabla \mathbf{v} - \mathbf{v} \cdot \nabla \mathbf{u}).$$

We may thus identify $\boldsymbol{\mathcal{L}}_{\mathbf{v}}^{*} = -\mathrm{ad}_{\mathbf{v}}^{*}$ and the previous equations can be rewritten as

$$\begin{cases} \dot{\Pi} = \mathrm{ad}_{\mathbf{v}}^{*} \Pi - \sigma^2 p \diamond p^{\sharp}, \\ \dot{a} = -\boldsymbol{\mathcal{L}}_{\mathbf{v}} a + \sigma^2 p^{\sharp}, \\ \dot{p} = \boldsymbol{\mathcal{L}}_{\mathbf{v}}^{\mathsf{T}} p. \end{cases} \tag{2.15}$$

These are Lie-Poisson equations with a cocycle for the Hamiltonian

$$h(\Pi, a, p) = \frac{1}{2} \left\langle \Pi, \Pi^{\sharp} \right\rangle_{\mathfrak{g}} + \frac{\sigma^2}{2} \left\langle p, p^{\sharp} \right\rangle_V \tag{2.16}$$

with respect to the Lie-Poisson bracket given by

$$\begin{bmatrix} \dot{\Pi} \\ \dot{a} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} \mathrm{ad}_{\square}^{*} \Pi & a \diamond \square & -p \diamond \square \\ -\boldsymbol{\mathcal{L}}_{\square} a & 0 & 1 \\ \boldsymbol{\mathcal{L}}_{\square}^{\mathsf{T}} p & -1 & 0 \end{bmatrix} \begin{bmatrix} \delta h/\delta \Pi = \Pi^{\sharp} = \mathbf{v} \\ \delta h/\delta a = 0 \\ \delta h/\delta p = \sigma^2 p^{\sharp} \end{bmatrix}, \tag{2.17}$$

in which the variational derivatives of the Hamiltonian are to be substituted into the corresponding places indicated by a box ($\square$). This matrix is identified as the Hamiltonian operator for the Lie-Poisson bracket dual to the semidirect product Lie algebra $\mathfrak{g} \, \circledS \, (V^{*} \times V)$ *plus* a symplectic 2-cocycle on $(a, p) \in V \times V^{*}$.

REMARK 2.4. This Hamiltonian matrix will block-diagonalize in the Lagrange-Poincaré formulation discussed in Section 4. Roughly speaking, this amounts to transforming variables $\Pi \to \tilde{\Pi} := (\Pi + p \diamond a)$ and $(a, \nu) \to (a, \dot{a})$.

Example 3. We now consider an example analogous to the preceding one but in finite dimensions. We let the orthogonal group $G = SO(3)$ act on $\mathbb{R}^3$ by matrix multiplication on the *left* and we choose $U := \mathfrak{so}(3) \times \mathbb{R}^3 \ni (\mathbf{\Omega}, \nu)$ as the control space. As usual, we identify the Lie algebra $\mathfrak{so}(3)$ with $\mathbb{R}^3$. We choose the quadratic Lagrangian $\ell : \mathfrak{so}(3) \times \mathbb{R}^3 \to \mathbb{R}$ given by

$$\ell(\mathbf{\Omega}, \nu) := \frac{1}{2} \mathbb{I}\mathbf{\Omega} \cdot \mathbf{\Omega} + \frac{1}{2\sigma^2} \mathbb{K}\nu \cdot \nu,$$

for symmetric positive definite matrices $\mathbb{I}$ and $\mathbb{K}$. We impose the evolution equation

$$\dot{\mathbf{X}} = -\mathbf{\Omega} \times \mathbf{X} + \nu \tag{2.18}$$

for the state variable $\mathbf{X} \in \mathbb{R}^3 =: Q$. As before, the variational principle $\delta S_c = 0$ with

$$S_c = \int_0^T \left(\frac{1}{2} \mathbb{I}\mathbf{\Omega} \cdot \mathbf{\Omega} + \frac{1}{2\sigma^2} \mathbb{K}\nu \cdot \nu + \mathbf{P} \cdot \left(\dot{\mathbf{X}} + \mathbf{\Omega} \times \mathbf{X} - \nu \right) \right) dt$$

yields the controls

$$\mathbb{I}\mathbf{\Omega} = \mathbf{P} \times \mathbf{X} \quad \text{and} \quad \mathbb{K}\nu = \sigma^2 \mathbf{P},$$

as in (2.11). Note that $\mathbf{\Omega} = \mathbb{I}^{-1}(\mathbf{P} \times \mathbf{X}) = (\mathbf{P} \times \mathbf{X})^\sharp$ and $\mathbb{K}^{-1}\mathbf{P} = \mathbf{P}^\sharp$, by the definition of the sharp maps. Then the state and costate evolution equations (2.12) take canonical Hamiltonian form with Hamiltonian function

$$H(\mathbf{X}, \mathbf{P}) = \frac{1}{2}(\mathbf{P} \times \mathbf{X}) \cdot (\mathbf{P} \times \mathbf{X})^\sharp + \frac{\sigma^2}{2} \mathbf{P} \cdot \mathbf{P}^\sharp. \tag{2.19}$$

Intriguingly, the resulting canonical Hamiltonian equations,

$$\begin{cases} \dot{\mathbf{X}} = \dfrac{\partial H}{\partial \mathbf{P}} = -(\mathbf{P} \times \mathbf{X})^\sharp \times \mathbf{X} + \sigma^2 \mathbf{P}^\sharp, \\[2mm] \dot{\mathbf{P}} = -\dfrac{\partial H}{\partial \mathbf{X}} = -(\mathbf{P} \times \mathbf{X})^\sharp \times \mathbf{P}, \end{cases} \tag{2.20}$$

involve the double cross product of the state and costate vectors $(\mathbf{X}, \mathbf{P}) \in \mathbb{R}^3 \times \mathbb{R}^3$. The double cross products correspond to the Lie derivatives in equations (2.12) which for this case become cross products. For more information about the roots of the Hamiltonian approach in geometric control theory, see [3].

Upon defining the vector $\mathbf{\Pi} := \mathbb{I}\mathbf{\Omega} = \mathbf{P} \times \mathbf{X}$, equations (2.20) imply

$$\begin{cases} \dot{\mathbf{\Pi}} = -\mathbf{\Omega} \times \mathbf{\Pi} - \sigma^2 (\mathbb{K}^{-1}\mathbf{P}) \times \mathbf{P}, \\ \dot{\mathbf{X}} = -\mathbf{\Omega} \times \mathbf{X} + \sigma^2 \mathbf{P}^\sharp, \\ \dot{\mathbf{P}} = -\mathbf{\Omega} \times \mathbf{P}, \end{cases} \tag{2.21}$$

which recovers the momentum map system (2.15) for this case. Indeed, one may compute directly that

$$\dot{\mathbf{\Pi}} = \dot{\mathbf{P}} \times \mathbf{X} + \mathbf{P} \times \dot{\mathbf{X}}$$

$$= (-\mathbf{\Omega} \times \mathbf{P}) \times \mathbf{X} + \mathbf{P} \times \left(-\mathbf{\Omega} \times \mathbf{X} + \sigma^2 \mathbf{P}^\sharp \right)$$

$$= (\mathbf{P} \times \mathbf{\Omega}) \times \mathbf{X} + (\mathbf{\Omega} \times \mathbf{X}) \times \mathbf{P} + \sigma^2 \mathbf{P} \times \mathbf{P}^\sharp$$

$$= -(\mathbf{X} \times \mathbf{P}) \times \mathbf{\Omega} + \sigma^2 \mathbf{P} \times (\mathbb{K}^{-1}\mathbf{P})$$

$$= \mathbf{\Pi} \times \mathbf{\Omega} + \sigma^2 \mathbf{P} \times (\mathbb{K}^{-1}\mathbf{P}),$$

from which the result follows.

REMARK 2.5 (Lie algebra formulation). The Lie algebra bracket on $\mathfrak{se}(3) \simeq \mathfrak{so}(3) \circledS \mathbb{R}^3$ may be written on $\mathbb{R}^3 \times \mathbb{R}^3$ as

$$\mathrm{ad}_{(\mathbf{\Omega},\boldsymbol{\alpha})}(\tilde{\mathbf{\Omega}}, \tilde{\boldsymbol{\alpha}}) = \left[(\mathbf{\Pi}, \boldsymbol{\alpha}), (\tilde{\mathbf{\Omega}}, \tilde{\boldsymbol{\alpha}}) \right] = \left(\mathbf{\Omega} \times \tilde{\mathbf{\Omega}}, \mathbf{\Omega} \times \tilde{\boldsymbol{\alpha}} - \tilde{\mathbf{\Omega}} \times \boldsymbol{\alpha} \right).$$

Its dual operation is

$$\mathrm{ad}^*_{(\mathbf{\Omega},\boldsymbol{\alpha})}(\mathbf{\Pi}, \mathbf{P}) = \left(-\mathbf{\Omega} \times \mathbf{\Pi} - \boldsymbol{\alpha} \times \mathbf{P}, -\mathbf{\Omega} \times \mathbf{P} \right).$$

In terms of the ad* operation on $\mathfrak{se}(3)^*$, the motion equations for $(\mathbf{\Pi}, \mathbf{P})$ in (2.21) can be rewritten as

$$\left(\dot{\mathbf{\Pi}}, \dot{\mathbf{P}} \right) = \left(-\mathbf{\Omega} \times \mathbf{\Pi} - \sigma^2 \mathbf{P}^\sharp \times \mathbf{P}, -\mathbf{\Omega} \times \mathbf{P} \right)$$

$$= \left(\mathrm{ad}^*_{\mathbf{\Omega}} \mathbf{\Pi} + \sigma^2 \mathbf{P} \diamond \mathbf{P}^\sharp, -\mathbf{\Omega} \times \mathbf{P} \right)$$

$$= \mathrm{ad}^*_{(\mathbf{\Omega}, \sigma^2 \mathbf{P}^\sharp)} (\mathbf{\Pi}, \mathbf{P}).$$

The result of the last calculation may be rewritten in Lie-Poisson bracket form as

$$\left(\dot{\mathbf{\Pi}}, \dot{\mathbf{P}} \right) = \mathrm{ad}^*_{(\partial h/\partial \mathbf{\Pi}, \partial h/\partial \mathbf{P})} (\mathbf{\Pi}, \mathbf{P}), \tag{2.22}$$

with Hamiltonian (2.19) rewritten in these variables as

$$h(\mathbf{\Pi}, \mathbf{P}) = \frac{1}{2} \mathbf{\Pi} \cdot \mathbf{\Pi}^\sharp + \frac{\sigma^2}{2} \mathbf{P} \cdot \mathbf{P}^\sharp, \tag{2.23}$$

and using the (left) Lie-Poisson bracket defined on the dual Lie algebra $\mathfrak{se}(3)^*$. This is the Hamiltonian and Lie-Poisson bracket for the motion of an ellipsoidal underwater vehicle in the body representation. See, e.g., [36] for more discussion and references to the literature about the geometrical approach to the dynamics and control of underwater vehicles.

We have seen that equations (2.20) for the state-costate vectors $(\mathbf{X}, \mathbf{P})$ are canonically Hamiltonian and that the system (2.22) for $(\mathbf{\Pi}, \mathbf{P})$ is Lie-Poisson on the dual of a semidirect product Lie algebra. Now, it remains to include the dynamics of the coordinate $\mathbf{X}$ into a single structure for the entire system (2.21) for $(\mathbf{\Pi}, \mathbf{X}, \mathbf{P})$. We observe that equations (2.21) may be put into Lie-Poisson form, as

$$\begin{bmatrix} \dot{\mathbf{\Pi}} \\ \dot{\mathbf{X}} \\ \dot{\mathbf{P}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi} \times & \mathbf{X} \times & \mathbf{P} \times \\ \mathbf{X} \times & 0 & 1 \\ \mathbf{P} \times & -1 & 0 \end{bmatrix} \begin{bmatrix} \delta h/\delta \mathbf{\Pi} \\ \delta h/\delta \mathbf{X} \\ \delta h/\delta \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi} \times & \mathbf{X} \times & \mathbf{P} \times \\ \mathbf{X} \times & 0 & 1 \\ \mathbf{P} \times & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{\Omega} \\ 0 \\ \sigma^2 \mathbf{P}^\sharp \end{bmatrix}. \tag{2.24}$$

This is the Lie-Poisson bracket dual to the semidirect product Lie algebra $\mathfrak{so}(3) \circledS (\mathbb{R}^3 \times \mathbb{R}^3)$ plus a symplectic 2-cocycle on $(\mathbf{X}, \mathbf{P}) \in \mathbb{R}^3 \times \mathbb{R}^3$.

REMARK 2.6. As mentioned earlier, the Lagrange-Poincaré and Hamilton-Poincaré formulations in sections 4 and 6 will block-diagonalize this Hamiltonian matrix.

REMARK 2.7 (Comparison of the examples). The major difference between Example 1 and Examples 2 and 3 is the following. In Example 1, we impose the *advection equation* $\dot{a} = \boldsymbol{\mathcal{L}}_{\mathbf{v}} a$ as a *constraint* on the minimization problem. This is done, as usual, by introducing a new variable $\mathbf{p}$ and adding the term $\langle \mathbf{p}, \dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a \rangle$ in the action functional. In Examples 2 and 3, the advection law is not imposed *exactly*, but only up to an error term

$$\nu := \dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a,$$

whose norm is added to the Lagrangian as a penalty, and needs to be minimized. Of course, in this case, the relation $\nu = \dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a$ is a constraint as seen in the term $\langle \mathbf{p}, \dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a - \nu \rangle$. As we have seen in Proposition 2.2, this error term implies a modification of the equations of motion.

    One of the aims of the present paper is to transform the control problem corresponding to the cost function in (2.10) into an optimization problem in which the penalty term $\|\dot{a} - \boldsymbol{\mathcal{L}}_{\mathbf{v}} a\|^2$ appears. This objective motivates the introduction of the distributed optimization problem in the next section.

    2.3. Control problems versus optimization problems.   Let $\ell = \ell(u, n)$: $U \times Q \to \mathbb{R}$ be a cost function and F a vector field on Q. We associate to these objects the following problems.

 (1) In the **Clebsch Optimal Control Problem** [29] one minimizes the integral

$$S := \int_0^T \ell(u, n) dt \quad \text{subject to the conditions} \quad \dot{n} = F(n, u)$$

    and the usual endpoint conditions. The resolution of this problem uses the Pontryagin maximum principle which, under a sufficient smoothness condition, implies that a solution of this problem is necessarily a solution of the variational principle $\delta S_c = 0$, with

$$S_c = \delta \int_0^T \Big(\ell(u, n) + \langle \alpha, \dot{n} - F(n, u) \rangle \Big) dt = 0. \tag{2.25}$$

    Example 1 in Section 2.2, for which the cost function is a kinetic energy and the vector field F is given by a Lie derivative, illustrates this method.

 (2) In **Clebsch Distributed Optimization** one minimizes the integral

$$S_p := \int_0^T \Big(\ell(u, n) + \frac{1}{2\sigma^2} \|\dot{n} - F(n, u)\|^2 \Big) dt, \tag{2.26}$$

    where a norm $\| \cdot \|$ is chosen, associated to a Riemannian metric on Q. In this cost functional, the Clebsch constraint for the state system dynamics has been relaxed to the status of a penalty, with positive tolerance, $\sigma^2 > 0$, and the variations are subject to the usual endpoint conditions. Of course, the solutions of this problem are necessarily solutions of the variational principle $\delta S_p = 0$.

REMARK 2.8. Despite the analogy between the two variational principles $\delta S_c = 0$ and $\delta S_p = 0$, the origins of these principles are quite different.

(1) In Clebsch Optimal Control, the functional S_c in (2.25) is minimized subject to the constraint of the state system dynamics by introducing the costate variable α. The well-known Pontryagin approach tells us that the solutions of the optimal control problem are necessarily critical points of S_c.

(2) In Clebsch Distributed Optimization, the stationarity condition for the variational principle is implied by *optimization* of the functional S_p with penalty in (2.26), without other constraints, except the endpoint conditions.

We may initially regard this second approach as simply modifying the cost function in the optimal control problem (2.3) by introducing a penalty based on a norm of the state system. We will show later that the equations of motion for the two types of optimization problems coincide in the limit $\sigma^2 \to 0$.

In the case where Q is a vector space, denoted by W, and the norm is associated to an inner product, the variations of the distributed cost function S_p in (2.26) now yield

$$\delta S_p = \int_0^T \left[\left\langle \frac{\delta \ell}{\delta n} - \left(\frac{\delta F}{\delta n} \right)^T p - \dot{p}, \delta n \right\rangle_W + \left\langle \frac{\delta \ell}{\delta u} - \left(\frac{\delta F}{\delta u} \right)^T p, \delta u \right\rangle_V \right] dt + \langle p, \delta n \rangle_W \Big|_0^T,$$

$$(2.27)$$

where the momentum variable p obtained from the variation with respect to the vector field $\dot{n} \in W$ is defined by

$$\sigma^2 p := \left(\dot{n} - F(n, u) \right)^\flat \in W^*, \tag{2.28}$$

and in this case the $\flat$ map (index lowering) is applied with respect to the inner product on W.

Example. Let us return to Example 2 above and treat it as Clebsch distributed optimization problem. As in Section 2.2, we consider the geometric setting of continuum mechanics. Contrary to Example 1 above, we do not impose the advection equation $\dot{a} = \pounds_{\mathbf{v}} a$ as a constraint but as a penalty. The problem is now to minimize the expression

$$S_p := \int_0^T \left[\frac{1}{2} \|\mathbf{v}\|_{\mathfrak{g}}^2 + \frac{1}{2\sigma^2} \|\dot{a} - \pounds_{\mathbf{v}} a\|_{L^2}^2 \right] dt,$$

where $\| \cdot \|_{L^2}$ is a L^2 norm on the space of tensor field densities. This problem is clearly equivalent to that of Example 2 in Section 2.2. The variational principle $\delta S_p = 0$ yields the control

$$\mathbf{v} = -(p \diamond a)^\sharp \in \mathfrak{g},$$

and the same equations as before,

$$\begin{cases} \dot{a} + \pounds_{(p \diamond a)^\sharp} a = \sigma^2 p^\sharp, \\ \dot{p} - \pounds_{(p \diamond n)^\sharp}^T p = 0, \end{cases} \tag{2.29}$$

where we have defined the variable p by

$$p := \frac{1}{\sigma^2} (\dot{a} - \pounds_{\mathbf{v}} a)^\flat \in V. \tag{2.30}$$

It is important to observe that in this approach the variable p is not really needed, since it is *defined* in terms of the other variables. This is not the case for the Clebsch approach described in the Examples of Section 2.2 for which p is an independent variable. For the Clebsch approach, the relation (2.30) is recovered as a consequence of the variational principle $\delta S_c = 0$.

3. Clebsch optimal control versus distributed optimization

In this section we begin with a brief review in Section 3.1 of the Clebsch optimal control problem studied in [29]. Then in Section 3.2 we introduce the class of *Clebsch distributed optimization problems* investigated in this paper, obtained by adding to the cost function a penalty given by the norm of the constraints in the previous approach.

3.1. Review of Clebsch optimal control.

Clebsch optimal control formulation and main results. We remind the reader that we previously defined the Clebsch optimal control problem [29] as follows. Let $\Phi : G \times Q \to Q$ be a left (resp. right) action of a Lie group G on the manifold Q and let $\ell : \mathfrak{g} \times Q \to \mathbb{R}$ be a cost function. The *Clebsch optimal control problem* for the curves $\xi(t) \in \mathfrak{g}$ and $n(t) \in Q$ is

$$\min_{\xi(t)} \int_0^T \ell(\xi(t), n(t)) dt \qquad (3.1)$$

subject to the following conditions:

(A) Either $\quad \dot{n}(t) = \xi(t)_Q(n(t)), \quad$ or $\quad$ (A)$'$ $\quad \dot{n}(t) = -\xi(t)_Q(n(t))$;

(B) Both $\quad n(0) = n_0 \quad$ and $\quad n(T) = n_T$,

where ξ_Q denotes the infinitesimal generator of the G-action associated to $\xi \in \mathfrak{g}$, that is,

$$\xi_Q(q) := \frac{d}{ds}\bigg|_{s=0} \Phi_{\exp(s\xi)}(q), \qquad q \in Q.$$

If condition (A) is assumed, then by applying the Pontryagin maximum principle, we obtain that an extremal curve $n(t) \in Q$ is necessarily the projection of a curve $\alpha(t) \in T^*Q$ that is a solution of the equations [29]

$$\frac{\delta \ell}{\delta \xi} = \mathbf{J}(\alpha), \quad \dot{\alpha} = \xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n}. \qquad (3.2)$$

Here $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ denotes the momentum map associated to the cotangent-lifted action of G on T^*Q. Recall that $\mathbf{J}$ is given by [54]

$$\langle \mathbf{J}(\alpha_q), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle = -\langle \alpha_q \diamond q, \xi \rangle, \quad \text{where} \quad \xi_Q(q) = \pounds_\xi q.$$

The expression $\frac{\delta \ell}{\delta \xi} \in \mathfrak{g}^*$ denotes the usual functional derivative of $\ell(\cdot, n)$ for each fixed $n \in Q$ whereas $\frac{\partial \ell}{\partial n} := \mathbf{d}\ell(\xi, \cdot) \in T_n^*Q$ denotes the differential of the function $\ell(\xi, \cdot) : Q \to \mathbb{R}$ for each fixed $\xi \in \mathfrak{g}$. For $\alpha, \beta \in T_q^*Q$, the map $\mathrm{Ver}_\alpha \beta$ denotes the *vertical lift of $\beta \in T_q^*Q$ relative to $\alpha \in T_q^*Q$*, defined by

$$\mathrm{Ver}_\alpha \beta := \frac{d}{ds}\bigg|_{s=0} (\alpha + s\beta) \in T_\alpha(T^*Q).$$

In (3.2), ξ_{T^*Q} denotes the infinitesimal generator of the cotangent-lifted action of G on T^*Q. Note that the vector field $\xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n}$ on T^*Q is the Hamiltonian vector field associated to the Hamiltonian

$$\alpha_n \in T^*Q \mapsto \langle \alpha_n, \xi_Q(n) \rangle - \ell(\xi, n) \in \mathbb{R},$$

in which the Lie algebra element $\xi \in \mathfrak{g}$ is regarded as a parameter. Using these equations, we determine that the optimal control ξ is the solution of the equations

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = -\mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi} + \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right), \quad \text{resp.} \quad \frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi} + \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right). \tag{3.3}$$

If condition (A)$'$ is assumed, then (3.2) is replaced by

$$\frac{\delta \ell}{\delta \xi} = -\mathbf{J}(\alpha), \quad \dot{\alpha} = -\xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n}, \tag{3.4}$$

and the optimal control ξ is the solution of the equations

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi} - \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right), \quad \text{resp.} \quad \frac{d}{dt}\frac{\delta \ell}{\delta \xi} = -\mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi} - \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right). \tag{3.5}$$

We refer to [29] for proofs of these statements and further discussion.

PROPOSITION 3.1 (Variational principle). *The equations* (3.2) *or* (3.4), *together with the constraint* $\dot{n} = \pm \xi_Q(n)$, *follow from the variational principle*

$$\delta \int_0^T \left(\ell(\xi, n) + \langle \alpha, \dot{n} \mp \xi_Q(n) \rangle\right) dt = 0, \tag{3.6}$$

for curves $t \mapsto \xi(t) \in \mathfrak{g}$ *and* $t \mapsto \alpha(t) \in T^*_{n(t)}Q$. *The variations* $\delta \xi$ *are free, whereas the variations* $\delta \alpha$ *are such that the induced variations* δn *vanish at the endpoints, that is,* $\delta n(0) = \delta n(T) = 0$.

Proof. To see this, let $\xi_s \in \mathfrak{g}$ and $\alpha_s \in T^*_{n_s}Q$ be curves whose infinitesimal variations at $s=0$ are $\delta \xi \in \mathfrak{g}$ and $\delta \alpha \in T^*_n Q$. We have

$$\delta \int_0^T \left(\ell(\xi, n) + \langle \alpha, \dot{n} \mp \xi_Q(n) \rangle\right) dt$$

$$= \int_0^T \left\langle \frac{\delta \ell}{\delta \xi}, \delta \xi \right\rangle dt + \int_0^T \left\langle \frac{\partial \ell}{\partial n}, \delta n \right\rangle dt$$

$$+ \frac{d}{ds}\bigg|_{s=0} \int_0^T \langle \alpha_s, \dot{n}_s \rangle dt \mp \frac{d}{ds}\bigg|_{s=0} \int_0^T \langle \mathbf{J}(\alpha_s), \xi_s \rangle dt. \tag{3.7}$$

A direct computation in canonical coordinates, using $\delta n(0) = \delta n(T) = 0$ in an integration by parts, shows that

$$\frac{d}{ds}\bigg|_{s=0} \int_0^T \langle \alpha_s, \dot{n}_s \rangle dt = \int_0^T \Omega_{\mathrm{can}}(\dot{\alpha}, \delta \alpha) dt, \tag{3.8}$$

where Ω_{can} denotes the canonical symplectic form on T^*Q. In addition, using the definition of the momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ we have

$$\frac{d}{ds}\bigg|_{s=0} \langle \mathbf{J}(\alpha_s), \xi_s \rangle = \langle T_\alpha \mathbf{J}(\delta \alpha), \xi \rangle + \langle \mathbf{J}(\alpha), \delta \xi \rangle = \Omega_{\mathrm{can}}(\xi_{T^*Q}(\alpha), \delta \alpha) + \langle \mathbf{J}(\alpha), \delta \xi \rangle. \tag{3.9}$$

Using relations (3.8) and (3.9) in formula (3.7) yields (3.2) and (3.4). $\square$

Alternative form of the stationarity conditions. Note that the *equations*

$$\dot{\alpha} = \pm\xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n} \tag{3.10}$$

imply the constraint $\dot{n} = \pm\xi_Q(n)$. To see this, it suffices to apply the tangent map $T\pi$ to (3.10), where $\pi : T^*Q \to Q$ is the projection, and recall that ξ_{T^*Q} and ξ_Q are π-related. By introducing a Riemannian metric g on Q, it is possible to rewrite the stationarity condition in a more explicit way, as we show in the following lemma.

LEMMA 3.1. *Suppose that Q is endowed with a Riemannian metric g and denote by ∇ and D/Dt the associated Levi-Civita covariant derivatives.*
  Then the equation $\dot{\alpha} = \pm\xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n}$ in (3.10) is equivalent to the system

$$\begin{cases} \dot{n} = \pm\xi_Q(n), \\ \dfrac{D}{Dt}\alpha = \mp\langle\alpha, \nabla\xi_Q(n)\rangle + \dfrac{\partial \ell}{\partial n}. \end{cases} \tag{3.11}$$

Proof.   We begin by recalling the definition and main property of the *connector* $K : TTQ \to TQ$ associated to a Riemannian manifold (Q, g). A general detailed treatment for connectors associated to linear connections can be found in [59], Section 19.11. In infinite dimensions we need to assume that the given weak Riemannian metric has a smooth geodesic spray $S \in \mathfrak{X}(TQ)$. In natural local charts of TTQ, the intrinsic map K is defined by

$$K_{loc}(x, e, u, v) = (x, v + \Gamma(x)(e, u)), \tag{3.12}$$

where $\Gamma(x)$ is the Christoffel map defined by the quadratic form in the fourth component of the geodesic spray $S(x, u) = (x, u, u, -\Gamma(x)(u, u))$. In finite dimensions, the Christoffel map has the familiar expression $\Gamma(x)(e, u)^i = \Gamma^i_{jk}(x)e^j u^k$, where Γ^i_{jk} are the usual Christoffel symbols associated to the metric g. The relation between the connector and the Levi-Civita covariant derivative is given for all $X, Y \in \mathfrak{X}(Q)$ by

$$\nabla_Y X = K \circ TX \circ Y. \tag{3.13}$$

The connector K induces an intrinsic map, also denoted by $K : TT^*Q \to T^*Q$ defined in natural local charts by

$$K_{loc}(x, \beta, u, \gamma) = (x, \gamma - \beta(\Gamma(x)(u, \cdot))). \tag{3.14}$$

The associated covariant derivative

$$\nabla_X \alpha := K \circ T\alpha \circ X \tag{3.15}$$

on T^*Q recovers the Levi-Civita connection on one-forms $\alpha \in \Omega^1(Q)$. Although the same notation is used for the connector on TQ and on T^*Q, it will be clear from the context which one is meant.
  The proof of Lemma 3.1 begins by recalling the vector bundle isomorphism $TT^*Q \to T^*Q \oplus TQ \oplus T^*Q$ given by

$$X \mapsto (\sigma_{T^*Q}(X), T\pi(X), K(X)),$$

where $\sigma_{T^*Q} : TT^*Q \to T^*Q$ is the projection. Therefore, to prove the equivalence it suffices to apply the maps $T\pi$ and K to the equation $\dot{\alpha} = \pm\xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n}$. As

we have seen before, applying $T\pi$ yields the first equation in the system (3.11). The Definition (3.14) of K and (3.15) immediately imply the equalities

$$K(\dot{\alpha}) = \frac{D}{Dt}\alpha \quad \text{and} \quad K\left(\text{Ver}_\alpha \frac{\partial \ell}{\partial n}\right) = \frac{\partial \ell}{\partial n}.$$

Thus, to finish the proof, it suffices to compute $K(\xi_{T^*Q}(\alpha))$. Given $v_n \in T_n Q$, $\alpha_n \in T_n^* Q$, and $\xi \in \mathfrak{g}$, we have

$$\left\langle T^* \Phi^{-1}_{\exp(s\xi)}(\alpha_n), T\Phi_{\exp(s\xi)}(v_n) \right\rangle = \langle \alpha_n, v_n \rangle.$$

Taking the s-derivative at $s = 0$ yields

$$\langle K(\xi_{T^*Q}(\alpha_n)), v_n \rangle + \langle \alpha_n, K(\xi_{TQ}(v_n)) \rangle = 0. \tag{3.16}$$

Noting the equalities $K(\xi_{TQ}(v_n)) = K(T\xi_Q(v_n)) = \nabla_{v_n}\xi_Q(n)$, we obtain the formula

$$K(\xi_{T^*Q}(\alpha_n)) = -\langle \alpha_n, \nabla \xi_Q(n) \rangle,$$

which proves Lemma 3.1, that the stationarity conditions (3.10) and (3.11) are equivalent for a Riemannian manifold. □

System (3.11) may also be obtained directly from the variational principle $\delta S_c = 0$,

$$S_c = \int_0^T \left(\ell(\xi, n) + \langle \alpha, \dot{n} \mp \xi_Q(n) \rangle\right) dt,$$

by using a Riemannian metric on Q. However, we have chosen to derive the stationarity conditions (3.2) or (3.4) together with the constraint $\dot{n} = \pm \xi_Q(n)$ for the functional S_c without introducing a Riemannian metric; see (3.7)–(3.9) above.

Lagrangian and Hamiltonian approach. Equations (3.3) and (3.5) can be obtained via Euler-Poincaré reduction for the G-invariant function $\mathcal{L}: TG \times Q \to \mathbb{R}$ induced by ℓ. More precisely, upon fixing $q \in Q$ and defining the Lagrangian $\mathcal{L}_q(u_g) := \mathcal{L}(u_g, q)$ on TG, one finds that the equations (3.3) and (3.5) are equivalent to the Euler-Lagrange equations for $\mathcal{L}_q$ by invoking a generalization of the Euler-Poincaré reduction theorem. We refer to [30] for a proof of this assertion and for applications to systems with broken symmetry. If Q is a representation space of G, one recovers the Euler-Poincaré reduction theorem for semidirect products; see [42, 43] and Section 5 below.

If the Legendre transform $\xi \in \mathfrak{g} \mapsto \frac{\delta \ell}{\delta \xi} \in \mathfrak{g}^*$ is a diffeomorphism, we can form the associated Hamiltonian $h: \mathfrak{g}^* \times Q \to \mathbb{R}$ defined by

$$h(\mu, n) := \langle \mu, \xi \rangle - \ell(\xi, n), \quad \text{where} \quad \frac{\delta \ell}{\delta \xi} = \mu.$$

In this case, the Lagrangian $\mathcal{L}$ is hyperregular on TG, the variable $q \in Q$ being considered as a parameter, and we can form the Hamiltonian $\mathcal{H}: T^*G \times Q \to \mathbb{R}$. More precisely, fixing $q \in Q$, we define

$$\mathcal{H}_q := E_q \circ \mathbb{F}\mathcal{L}_q^{-1},$$

where E_q is the energy associated to the Lagrangian $\mathcal{L}_q: TG \to \mathbb{R}$ and $\mathbb{F}\mathcal{L}_q: TQ \to T^*Q$ is the classical Legendre transform of $\mathcal{L}_q$. The function $\mathcal{H}: T^*G \times Q \to \mathbb{R}$ is

then defined by $\mathcal{H}(\alpha_g, q) := \mathcal{H}_q(\alpha_g)$. Equations (3.3) and (3.5) can be written in Hamiltonian form as

$$
\begin{cases}
\dot{\mu} = \mp \, \mathrm{ad}^*_{\frac{\delta h}{\delta \mu}} \mu - \mathbf{J}\left(\dfrac{\partial h}{\partial n}\right), \\[2ex]
\dot{n} = \left(\dfrac{\delta h}{\delta \mu}\right)_Q (n),
\end{cases}
\tag{3.17}
$$

and

$$
\begin{cases}
\dot{\mu} = \pm \, \mathrm{ad}^*_{\frac{\delta h}{\delta \mu}} \mu + \mathbf{J}\left(\dfrac{\partial h}{\partial n}\right), \\[2ex]
\dot{n} = -\left(\dfrac{\delta h}{\delta \mu}\right)_Q (n),
\end{cases}
\tag{3.18}
$$

respectively. They are obtained by Poisson reduction of Hamilton's equations for $\mathcal{H}$ on $T^*G \times Q$, where Q is endowed with the zero Poisson structure.

In terms of h, the equations (3.2) or (3.4) read

$$
\mu = \mathbf{J}(\alpha), \quad \dot{\alpha} = \left(\frac{\delta h}{\delta \mu}\right)_{T^*Q}(\alpha) - \mathrm{Ver}_\alpha \frac{\partial h}{\partial n},
\tag{3.19}
$$

and

$$
\mu = -\mathbf{J}(\alpha), \quad \dot{\alpha} = -\left(\frac{\delta h}{\delta \mu}\right)_{T^*Q}(\alpha) - \mathrm{Ver}_\alpha \frac{\partial h}{\partial n}.
\tag{3.20}
$$

As in Lemma 3.1, by introducing a Riemannian metric g on Q, these equations can be rewritten as

$$
\mu = \pm \mathbf{J}(\alpha), \quad \dot{n} = \pm \left(\frac{\delta h}{\delta \mu}\right)_Q (n), \quad \frac{D}{Dt}\alpha = \mp \left\langle \alpha, \nabla \left(\frac{\delta h}{\delta \mu}\right)_Q (n) \right\rangle - \frac{\partial h}{\partial n}.
\tag{3.21}
$$

3.2. Clebsch distributed optimization problems: using penalties.

As before, we consider a left (resp. right) action $\Phi : G \times Q \to Q$ and a cost function $\ell : \mathfrak{g} \times Q \to \mathbb{R}$. We suppose that the manifold Q is endowed with a Riemannian metric g. The basic idea is to treat the condition (A) or (A)$'$ as a penalty rather than a constraint. Therefore, in the case of condition (A) above, we consider the minimization problem

$$
\min_{\xi, n} \int_0^T \left(\ell(\xi, n) + \frac{1}{2\sigma^2} \|\dot{n} - \xi_Q(n)\|^2 \right) dt,
\tag{3.22}
$$

and if condition (A)$'$ holds, we consider

$$
\min_{\xi, n} \int_0^T \left(\ell(\xi, n) + \frac{1}{2\sigma^2} \|\dot{n} + \xi_Q(n)\|^2 \right) dt.
\tag{3.23}
$$

These two problems are subject to the condition

$$
n(0) = n_0 \quad \text{and} \quad n(T) = n_T,
$$

for given $n_0, n_T \in Q$. Here the norm is taken with respect to the Riemannian metric g on Q and $\sigma \neq 0$.

Stationarity conditions. In order to find the critical curves, we consider the variational principle

$$\delta \int_0^T \left(\ell(\xi, n) + \frac{1}{2\sigma^2} \|\dot{n} \mp \xi_Q(n)\|^2 \right) dt = 0 \tag{3.24}$$

for the two curves $(\xi, n) : [0, T] \mapsto \mathfrak{g} \times Q$, where n has fixed endpoints. That is, the variation $\delta\xi$ is free and the variation δn vanishes at the endpoints.

We will treat condition (A) and (A)$'$ simultaneously. In all the expressions below, the upper sign refers to condition (A) and the lower sign refers to condition (A)$'$. The ξ-variation yields the condition

$$\frac{\delta\ell}{\delta\xi} = \pm\frac{1}{\sigma^2} \mathbf{J}(\nu_n^\flat), \quad \text{where} \quad \nu_n := \dot{n} \mp \xi_Q(n), \tag{3.25}$$

and $\nu_n^\flat := g(n)(\nu_n, \cdot) \in T_n^* Q$. We now compute the variations of n, where we denote by ∇ and D/Dt the covariant derivatives associated to the Levi-Civita connection of the metric g. For $\delta n = \frac{d}{ds}\big|_{s=0} n_s$, we have

$$\int_0^T \left(\left\langle \frac{\partial\ell}{\partial n}, \delta n \right\rangle + \frac{1}{\sigma^2} \left\langle \nu_n^\flat, \left. \frac{D}{Ds} \right|_{s=0} \dot{n} \mp \left. \frac{D}{Ds} \right|_{s=0} \xi_Q(n_s) \right\rangle \right) dt$$

$$= \int_0^T \left(\left\langle \frac{\partial\ell}{\partial n}, \delta n \right\rangle - \frac{1}{\sigma^2} \left\langle \frac{D}{Dt} \nu_n^\flat, \delta n \right\rangle \mp \frac{1}{\sigma^2} \left\langle \nu_n^\flat, \nabla_{\delta n} \xi_Q(n) \right\rangle \right) dt.$$

Upon exchanging the order of derivatives, $\frac{D}{Dt}\frac{d}{ds} = \frac{D}{Ds}\frac{d}{dt}$ (which is allowed because the Levi-Civita connection has no torsion) one finds the equation

$$\frac{D}{Dt}\nu_n^\flat = \mp g(\nu_n, \nabla\xi_Q) + \sigma^2 \frac{\partial\ell}{\partial n}. \tag{3.26}$$

Consequently, (ξ, n) is a solution of (3.24) if and only if (3.25) and (3.26) hold. In what follows, equations (3.25) and (3.26) will be called the *stationarity conditions*.

Note that here, in contrast to the argument in Section 3.1, specific use of the Riemannian metric is made in computing the stationarity equations from the condition $\delta S_p = 0$, where

$$S_p := \int_0^T \left(\ell(\xi, n) + \frac{1}{2\sigma^2} \|\dot{n} \mp \xi_Q(n)\|^2 \right) dt. \tag{3.27}$$

This is natural, because a Riemannian metric is provided by the penalty term in the problem statement. Using the notation

$$\pi := \frac{1}{\sigma^2}\nu_n^\flat = \frac{1}{\sigma^2}(\dot{n} \mp \xi_Q(n))^\flat \in T^*Q$$

enables the stationarity conditions (3.25) and (3.26) to be written as

$$\frac{\delta\ell}{\delta\xi} = \pm\mathbf{J}(\pi), \quad \dot{n} = \pm\xi_Q(n) + \sigma^2\pi^\sharp, \quad \frac{D}{Dt}\pi = \mp\langle\pi, \nabla\xi_Q\rangle + \frac{\partial\ell}{\partial n}. \tag{3.28}$$

These equations should be compared with the other stationarity conditions (3.2) and (3.11),

$$\frac{\delta\ell}{\delta\xi} = \pm\mathbf{J}(\alpha), \quad \dot{n} = \pm\xi_Q(n), \quad \frac{D}{Dt}\alpha = \mp\langle\alpha, \nabla\xi_Q\rangle + \frac{\partial\ell}{\partial n}, \tag{3.29}$$

associated to the Clebsch optimal control problem. These two sets of stationarity conditions are analogous. However, the corresponding variables α and π have different origins. Namely, the costate variable α was introduced as the Lagrange multiplier in formulating the constrained Clebsch variational principle (3.6), whereas the variable π arises as a canonical momentum, dual to the penalty variable ν_n in the unconstrained variational principle (3.24).

Recall from Lemma 3.1 that the last two stationarity conditions of the system (3.29) are equivalent to

$$\dot{\alpha} = \pm \xi_{T^*Q}(\alpha) + \mathrm{Ver}_\alpha \frac{\partial \ell}{\partial n}.$$

An analogous result concerning the stationarity conditions of the distributed optimal control problem is given by the following lemma. Let $\sharp := \flat^{-1} : T^*Q \to TQ$.

LEMMA 3.2. *The system of two equations*

$$\begin{cases} \dot{n} = \pm \xi_Q(n) + \sigma^2 \pi^\sharp, \\ \dfrac{D}{Dt}\pi = \mp \langle \pi, \nabla \xi_Q(n) \rangle + \dfrac{\partial \ell}{\partial n} \end{cases} \tag{3.30}$$

is equivalent to the single equation

$$\dot{\pi} = \pm \xi_{T^*Q}(\pi) + \mathrm{Ver}_\pi \frac{\partial \ell}{\partial n} + \sigma^2 \mathcal{S}(\pi),$$

where $\mathcal{S} \in \mathfrak{X}(T^*Q)$ is the Hamiltonian vector field associated to the kinetic energy of the Riemannian metric.

Proof. It suffices to observe that the vector field $\mathcal{S}$ satisfies the properties

$$K(\mathcal{S}(\alpha)) = 0 \quad \text{and} \quad T\pi(\mathcal{S}(\alpha)) = \alpha^\sharp,$$

for all $\alpha \in T^*Q$. Then the proof is similar to that of Lemma 3.1.   □

REMARK 3.2. In terms of the Hamiltonian h associated to ℓ, the stationarity conditions (3.25) and (3.26) read

$$\dot{n} = \pm \left(\frac{\delta h}{\delta \kappa} \right)_Q (n) + \sigma^2 \pi^\sharp, \quad \frac{D}{Dt}\pi = \mp \left\langle \pi, \nabla \left(\frac{\delta h}{\delta \kappa} \right)_Q (n) \right\rangle - \frac{\partial h}{\partial n},$$

or, equivalently,

$$\dot{\pi} = \pm \left(\frac{\delta h}{\delta \kappa} \right)_{T^*Q} (\pi) - \mathrm{Ver}_\pi \frac{\partial h}{\partial n} + \sigma^2 \mathcal{S}(\pi).$$

These equations should be compared to their analogues in (3.19) – (3.21).

Equations of motion associated to the stationarity conditions. We now compute the differential equation associated to condition (3.25), that is, the analogue of equations (3.3), (3.5). The formulation will involve the following $\mathfrak{g}^*$-valued $(1,1)$ tensor field.

DEFINITION 3.3. *Consider a Lie group G acting on a Riemannian manifold (Q, g). We define the $\mathfrak{g}^*$-valued $(1,1)$ tensor field $\mathcal{F}^\nabla : T^*Q \times TQ \to \mathfrak{g}^*$ associated to the Levi-Civita connection ∇ by*

$$\left\langle \mathcal{F}^\nabla(\alpha_q, u_q), \eta \right\rangle := \left\langle \alpha_q, \nabla_{u_q} \eta_Q(q) \right\rangle, \tag{3.31}$$

for all $u_q \in T_q Q$, $\alpha_q \in T_q^* Q$, and $\eta \in \mathfrak{g}$.

The main properties of the tensor field $\mathcal{F}^\nabla$ are given in the following lemmas.

LEMMA 3.4. *For all $\alpha_q \in T_q^* Q$, $u_q \in T_q Q$, and $\xi \in \mathfrak{g}$,*

$$\langle \mathcal{F}^\nabla(\alpha_q, u_q), \xi \rangle = \langle \alpha_q, K(\xi_{TQ}(u_q)) \rangle = -\langle K(\xi_{T^*Q}(\alpha_q)), u_q \rangle,$$

where K denotes the connectors of the covariant derivatives on TQ and T^*Q, respectively (see formulas (3.12)-(3.15)).

Proof. It suffices to use formula (3.16) in the proof of Lemma 3.1.   □

The following important property of $\mathcal{F}^\nabla$ is valid when G acts by isometries.

LEMMA 3.5. *If G acts by isometries, then $\mathcal{F}^\nabla$ is antisymmetric, that is*

$$\mathcal{F}^\nabla(\alpha_q, u_q) = -\mathcal{F}^\nabla(u_q^\flat, \alpha_q^\sharp),$$

for all $u_q \in T_q Q$, $\alpha_q \in T_q^* Q$.

Proof. Since G acts by isometries, $\pounds_{\xi_Q} g = 0$ which implies $(\nabla \xi_Q)^T = -\nabla \xi_Q$.   □

We also need the following preparatory lemma, valid for any action.

LEMMA 3.6. *Let $\mathbf{J} : T^* Q \to \mathfrak{g}$, $\langle \mathbf{J}(\alpha_q), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle$ be the momentum map of the cotangent-lifted action of G on T^*Q and let g be a Riemannian metric on Q. Then for a curve $\alpha(t) \in T_{q(t)}^* Q$ we have*

$$\frac{d}{dt} \mathbf{J}(\alpha(t)) = \mathbf{J}\left(\frac{D}{Dt}\alpha(t)\right) + \mathcal{F}^\nabla(\alpha(t), \dot{q}(t)),$$

where D/Dt and ∇ denote the Levi-Civita covariant derivatives associated to g.

Proof. For all $\eta \in \mathfrak{g}$, we have

$$\frac{d}{dt}\langle \mathbf{J}(\alpha(t)), \eta \rangle = \frac{d}{dt}\langle \alpha(t), \eta_Q(q(t)) \rangle = \left\langle \frac{D}{Dt}\alpha(t), \eta_Q(q(t)) \right\rangle + \left\langle \alpha(t), \frac{D}{Dt}\eta_Q(q(t)) \right\rangle$$

$$= \left\langle \mathbf{J}\left(\frac{D}{Dt}\alpha(t)\right), \eta \right\rangle + \left\langle \alpha(t), \nabla_{\dot{q}(t)}\eta_Q(q(t)) \right\rangle.$$

Using the definition of $\mathcal{F}^\nabla$ implies the required formula.   □

Note that this proof of Lemma 3.6 did not assume that the metric is G-invariant and that the formula is valid for left and right actions.

Lemma 3.6 and equations (3.25), (3.26) enable one to compute the motion equations associated to the minimization problems (3.22), (3.23) as follows:

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \pm \frac{d}{dt}\frac{1}{\sigma^2}\mathbf{J}(\nu_n^\flat) = \pm \frac{1}{\sigma^2}\mathbf{J}\left(\frac{D}{Dt}\nu_n^\flat\right) \pm \frac{1}{\sigma^2}\mathcal{F}^\nabla(\nu_n^\flat, \dot{n})$$

$$= \pm \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right) - \frac{1}{\sigma^2}\mathbf{J}\left(\langle \nu_n^\flat, \nabla \xi_Q \rangle\right) \pm \frac{1}{\sigma^2}\mathcal{F}^\nabla(\nu_n^\flat, \nu_n) + \frac{1}{\sigma^2}\mathcal{F}^\nabla(\nu_n^\flat, \xi_Q(n))$$

$$= \pm \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right) \pm \frac{1}{\sigma^2}\mathcal{F}^\nabla(\nu_n^\flat, \nu_n) + (\mp)\frac{1}{\sigma^2}\mathrm{ad}_\xi^* \mathbf{J}(\nu_n^\flat)$$

$$= \pm \mathbf{J}\left(\frac{\partial \ell}{\partial n}\right) \pm \frac{1}{\sigma^2}\mathcal{F}^{\nabla}(\nu_n^{\flat}, \nu_n) \pm (\mp)\,\mathrm{ad}_{\xi}^{*}\frac{\delta \ell}{\delta \xi},$$

where in $(\mp)$ one chooses $-$ (resp. $+$) when G acts on Q by a left (resp. right) action; so in the last term there are four choices of sign. Thus, when the penalty is given by $\|\dot{n} - \xi_Q(n)\|^2$ (condition (A)), the critical curves of the variational principle (3.24) are solutions of

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta \ell}{\delta \xi} = \mp\,\mathrm{ad}_{\xi}^{*}\dfrac{\delta \ell}{\delta \xi} + \mathbf{J}\left(\dfrac{\partial \ell}{\partial n}\right) + \dfrac{1}{\sigma^2}\mathcal{F}^{\nabla}(\nu_n^{\flat}, \nu_n), \\[2mm] \dfrac{D}{Dt}\nu_n^{\flat} = -\langle \nu_n^{\flat}, \nabla\xi_Q\rangle + \sigma^2\dfrac{\partial \ell}{\partial n}, \quad \nu_n := \dot{n} - \xi_Q(n). \end{cases} \tag{3.32}$$

When the penalty $\|\dot{n} + \xi_Q(n)\|^2$ (condition (A)$'$) is chosen instead, one finds,

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta \ell}{\delta \xi} = \pm\,\mathrm{ad}_{\xi}^{*}\dfrac{\delta \ell}{\delta \xi} - \mathbf{J}\left(\dfrac{\partial \ell}{\partial n}\right) - \dfrac{1}{\sigma^2}\mathcal{F}^{\nabla}(\nu_n^{\flat}, \nu_n), \\[2mm] \dfrac{D}{Dt}\nu_n^{\flat} = \langle \nu_n^{\flat}, \nabla\xi_Q\rangle + \sigma^2\dfrac{\partial \ell}{\partial n}, \quad \nu_n := \dot{n} + \xi_Q(n). \end{cases} \tag{3.33}$$

REMARK 3.3. The motion equations (3.32) and (3.33) should be compared to the analogous motion equation (3.3) and (3.5), respectively, obtained by the Clebsch optimal control approach. Note that the term $\mathcal{F}^{\nabla}(\nu_n^{\flat}, \nu_n)$ is an additional force term that is due to the presence of the quantity ν_n. The variable $\nu_n = \dot{n} \pm \xi_Q(n)$ measures the inexact matching and evolves according to the second equation $\frac{D}{Dt}\nu_n^{\flat} = \pm g(\nu_n, \nabla\xi_Q) + \sigma^2\frac{\partial \ell}{\partial n}$. We shall return to the discussion of inexact matching for images in Section 8.5.

Thanks to Lemma 3.5 we obtain the following important result, when G acts by isometries.

THEOREM 3.7. *Let G be a Lie group acting on a manifold Q and let $\ell : \mathfrak{g} \times Q \to \mathbb{R}$ be a cost function. We consider the two associated Clebsch optimal control and distributed optimization problems. Suppose that the Riemannian metric used in the penalty term is G-invariant. Then the two problems yield the same equations of motion.*

Proof. It suffices to use Lemma 3.5, and to compare equations (3.33), (3.32) with equations (3.3), (3.5). □

For completeness we rewrite below the equations (3.32) and (3.33) in the particular case where G acts by isometries. Using $\mathcal{F}^{\nabla}(\nu_n^{\flat}, \nu_n) = 0$ and $\nabla\xi_Q^{T} = -\nabla\xi_Q$ for this case yields

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta \ell}{\delta \xi} = \mp\,\mathrm{ad}_{\xi}^{*}\dfrac{\delta \ell}{\delta \xi} + \mathbf{J}\left(\dfrac{\partial \ell}{\partial n}\right), \\[2mm] \dfrac{D}{Dt}\nu_n = \nabla_{\nu_n}\xi_Q + \sigma^2\dfrac{\partial \ell}{\partial n}^{\sharp}, \quad \nu_n := \dot{n} - \xi_Q(n) \end{cases} \tag{3.34}$$

and

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta \ell}{\delta \xi} = \pm\,\mathrm{ad}_{\xi}^{*}\dfrac{\delta \ell}{\delta \xi} - \mathbf{J}\left(\dfrac{\partial \ell}{\partial n}\right), \\[2mm] \dfrac{D}{Dt}\nu_n = -\nabla_{\nu_n}\xi_Q + \sigma^2\dfrac{\partial \ell}{\partial n}^{\sharp}, \quad \nu_n := \dot{n} + \xi_Q(n). \end{cases} \tag{3.35}$$

REMARK 3.4. The remainder of the present paper will investigate these equations as dynamical systems, rather than as optimal control problems. See [44], in which a similar approach is taken.

4. Lagrange-Poincaré and metamorphosis reduction

In this section, we present two Lagrangian reduction approaches that will be useful in understanding the geometry of the equations (3.33), (3.32) associated to the minimization problem (3.23), (3.22).

Let G act on the left (resp. right) on Q. Let $L:T(G \times Q) \to \mathbb{R}$ be a left (resp. right)-invariant Lagrangian under the action of G given by

$$(u_g, u_q) \mapsto (hu_g, hu_q) \quad \text{resp.} \quad (u_g, u_q) \mapsto (u_g h, u_q h).$$

Two reduction processes are discussed. The first uses Lagrangian reduction (see [19]) and the second is a formulation of the reduction used for metamorphosis in [47].

THEOREM 4.1 (Lagrange-Poincaré reduction). *Let $g \in G$ and $q \in Q$ be two curves and define the curves $n := g^{-1}q \in Q$ and $\xi := g^{-1}\dot{g} \in \mathfrak{g}$ (resp. $n := qg^{-1} \in Q$ and $\xi := \dot{g}g^{-1} \in \mathfrak{g}$).*

Then (g,q) is a solution of the Euler-Lagrange equations for L if and only if (n,ξ) is a solution of the Lagrange-Poincaré equations

$$
\begin{cases}
\dfrac{d}{dt}\dfrac{\delta\ell_{LP}}{\delta\xi} = \mathrm{ad}^*_\xi \dfrac{\delta\ell_{LP}}{\delta\xi}, \\[2mm]
\dfrac{D}{Dt}\dfrac{\partial\ell_{LP}}{\partial\dot{n}} - \dfrac{\partial\ell_{LP}}{\partial n} = 0, \quad \dfrac{d}{dt}n = \dot{n},
\end{cases}
\quad resp. \quad
\begin{cases}
\dfrac{d}{dt}\dfrac{\delta\ell_{LP}}{\delta\xi} = -\mathrm{ad}^*_\xi \dfrac{\delta\ell_{LP}}{\delta\xi}, \\[2mm]
\dfrac{D}{Dt}\dfrac{\partial\ell_{LP}}{\partial\dot{n}} - \dfrac{\partial\ell_{LP}}{\partial n} = 0, \quad \dfrac{d}{dt}n = \dot{n},
\end{cases}
\tag{4.1}
$$

where the Lagrange-Poincaré Lagrangian $\ell_{LP} = \ell_{LP}(n,\dot{n},\xi):TQ \times \mathfrak{g} \to \mathbb{R}$ is induced from L by the quotient map

$$T(G \times Q) \to TQ \times \mathfrak{g}, \quad (u_g, u_q) \mapsto (n,\dot{n},\xi) := (\nu_n - \xi_Q(n), \xi) \tag{4.2}$$

for $n := g^{-1}q$, $\nu_n := g^{-1}u_q$, $\xi := g^{-1}u_g$ (resp. $n := qg^{-1}$, $\nu_n := u_qg^{-1}$, $\xi := u_gg^{-1}$).

These equations are equivalent to the variational principle

$$\delta \int_0^T \ell_{LP}(n,\dot{n},\xi)dt = 0$$

for arbitrary variations δn and constrained variations $\delta\xi = \dot{\eta} + [\xi,\eta]$ (resp. $\delta\xi = \dot{\eta} - [\xi,\eta]$).

In the Lagrange-Poincaré equations, D/Dt and $\partial\ell_{LP}/\partial n$ denote the covariant derivative and the partial derivative associated to a fixed torsion free connection ∇ on Q.

Proof. We treat the case of a left action and apply the results of [19]. The projection associated to the G-action reads $\pi: G \times Q \to Q$, $\pi(q,g) = g^{-1}q$. Thus, by taking the tangent map, we find $T\pi(u_g, u_q) = \left(g^{-1}u_g - (g^{-1}u_g)_Q(g^{-1}q)\right)$. The adjoint bundle $\mathrm{Ad}(G \times Q)$ can be identified with the trivial vector bundle $Q \times \mathfrak{g}$ via the identification $[(g,q),\xi] \simeq (g^{-1}q, \mathrm{Ad}_{g^{-1}}\xi)$. Using the principal connection $\mathcal{A}(u_g, u_q) := u_gg^{-1}$, the diffeomorphism $(T(G \times Q))/G \cong TQ \times \mathfrak{g}$ is given by $[u_g, u_q] \mapsto \left(g^{-1}u_q - (g^{-1}u_g)_Q(g^{-1}q), g^{-1}u_g\right)$. Thus, the Lagrange-Poincaré reduction map has the required expression (4.2). Since the chosen principal connection is flat, we obtain the Lagrange-Poincaré equations (4.1). $\square$

For the same G-invariant Lagrangian $L:T(G\times Q)\to\mathbb{R}$ as before, we define another reduced Lagrangian $\ell_M=\ell_M(\nu_n,\xi):TQ\times\mathfrak{g}\to\mathbb{R}$ associated to the quotient map

$$T(G\times Q)\to TQ\times\mathfrak{g},\quad (u_g,u_q)\mapsto(\nu_n,\xi):=(g^{-1}u_q,g^{-1}u_g),\quad\text{resp.}\quad(\nu_n,\xi):=(u_qg^{-1},u_gg^{-1}).$$

This reduced Lagrangian differs from the Lagrange-Poincaré Lagrangian ℓ_{LP} defined above, but one can pass from the one to the other by the vector bundle isomorphism

$$TQ\times\mathfrak{g}\to TQ\times\mathfrak{g},\quad(\nu_n,\xi)\mapsto(\nu_n-\xi_Q(n),\xi),\tag{4.3}$$

that is, we have

$$\ell_{LP}(n,\dot n,\xi)=\ell_M(n,\dot n+\xi_Q(n),\xi)$$

for both the left and right cases. The reduction associated to this quotient map will be called *metamorphosis reduction*, since it is the abstract framework underlying the metamorphosis dynamics described in [47].

THEOREM 4.2 (Metamorphosis reduction). *Let $g\in G$ and $q\in Q$ be two curves and define the curves $\nu_n:=g^{-1}\dot q\in TQ$ and $\xi:=g^{-1}\dot g\in\mathfrak{g}$ (resp. $\nu_n:=\dot qg^{-1}\in TQ$ and $\xi:=\dot gg^{-1}\in\mathfrak{g}$).*

Then (g,q) is a solution of the Euler-Lagrange equations associated to L if and only if (ν,ξ) is a solution of the equations

$$\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell_M}{\delta\xi}=\pm\operatorname{ad}_\xi^*\dfrac{\delta\ell_M}{\delta\xi}-\mathbf{J}\left(\dfrac{\partial\ell_M}{\partial n}\right)-\mathcal{F}^\nabla\left(\dfrac{\partial\ell_M}{\partial\nu_n},\nu_n\right),\\[3mm]\dfrac{D}{Dt}\dfrac{\partial\ell_M}{\partial\nu_n}=\left\langle\dfrac{\partial\ell_M}{\partial\nu_n},\nabla\xi_Q\right\rangle+\dfrac{\partial\ell_M}{\partial n},\qquad\dfrac{d}{dt}n=\nu_n-\xi_Q(n),\end{cases}\tag{4.4}$$

where $+$ (resp. $-$) occurs when G acts on Q by a left (resp. right) action, and $\mathcal{F}^\nabla$ is the $\mathfrak{g}^*$-valued $(1,1)$ tensor field defined in (3.31). In (4.4), $\partial\ell_M/\partial n$ and $\partial\ell_M/\partial\nu_n$ denote the horizontal and fiber derivatives, respectively.

These equations are equivalent to the variational principle

$$\delta\int_0^T\ell_M(\nu,\xi)dt=0,$$

with variations $\delta\xi=\dot\eta+[\xi,\eta]$ (resp. $\delta\xi=\dot\eta-[\xi,\eta]$) and $\delta\nu=\dfrac{D}{Dt}\omega+\nabla_\omega\xi_Q-\nabla_\nu\eta_Q$.

The proof will use the following lemma.

LEMMA 4.3. *Consider the two reduced Lagrangians ℓ_{LP} and ℓ_M. Then we have the relations*

$$\frac{\delta\ell_{LP}}{\delta\xi}=\frac{\delta\ell_M}{\delta\xi}+\mathbf{J}\left(\frac{\partial\ell_M}{\partial\nu_n}\right),\quad\frac{\partial\ell_{LP}}{\partial n}=\frac{\partial\ell_M}{\partial n}+\left\langle\frac{\partial\ell_M}{\partial\nu_n},\nabla\xi_Q\right\rangle,\quad\frac{\partial\ell_{LP}}{\partial\dot n}=\frac{\partial\ell_M}{\partial\nu_n}.\tag{4.5}$$

Proof. Using the relation $\ell_{LP}(n,\dot n,\xi)=\ell_M(n,\dot n+\xi_Q(n),\xi)$, we easily obtain the first and third expression. For the second we recall that partial derivatives $\frac{\partial\ell_{LP}}{\partial n},\frac{\partial\ell_M}{\partial n}$ are defined with the help of a connection ∇ on Q. Let $c(t)\in T_{m(t)}Q$ be a smooth horizontal curve covering a curve $m(t)\in Q$ and such that $c(0)=\dot n,\ \dot m(0)=u_n\in T_nQ$. By using the decomposition of TTQ into its vertical and horizontal part, we have

$$\left\langle\frac{\partial\ell_{LP}}{\partial n}(n,\dot n,\xi),u_n\right\rangle=\frac{d}{dt}\bigg|_{t=0}\ell_{LP}(c(t),\xi)=\frac{d}{dt}\bigg|_{t=0}\ell_M(c(t)+\xi_Q(m(t)),\xi)$$

$$
= \mathbf{d}_{TQ} \ell_M (n, \dot{n}, \xi) \left(\frac{d}{dt} \Big|_{t=0} c(t) + \frac{d}{dt} \Big|_{t=0} \xi_Q(m(t)) \right)
$$

$$
= \left\langle \frac{\partial \ell_M}{\partial n} (\dot{n} + \xi_Q(n), \xi), T\pi \left(\frac{d}{dt} \Big|_{t=0} c(t) + \frac{d}{dt} \Big|_{t=0} \xi_Q(m(t)) \right) \right\rangle
$$

$$
+ \left\langle \frac{\partial \ell_M}{\partial \nu_n} (\dot{n} + \xi_Q(n), \xi), K \left(\frac{d}{dt} \Big|_{t=0} c(t) + \frac{d}{dt} \Big|_{t=0} \xi_Q(m(t)) \right) \right\rangle
$$

$$
= \left\langle \frac{\partial \ell_M}{\partial n} (\dot{n} + \xi_Q(n), \xi), u_n \right\rangle + \left\langle \frac{\partial \ell_M}{\partial \nu_n} (\dot{n} + \xi_Q(n), \xi), \nabla_{u_n} \xi_Q \right\rangle,
$$

where $K : TTQ \to TQ$ denotes the connector map associated to ∇. Here $\mathbf{d}_{TQ}$ is the exterior derivative on TQ and the fourth equality is a general formula, valid for linear connections, that links the total derivative to the horizontal and vertical derivatives. $\square$

Proof. [Proof of Theorem 4.2.] We treat simultaneously the case of a left and right action. Using the second equation in (4.1) and Lemma 4.3, we directly obtain the equations

$$
\frac{D}{Dt} \frac{\partial \ell_M}{\partial \nu_n} - \frac{\partial \ell_M}{\partial n} = \left\langle \frac{\partial \ell_M}{\partial \nu_n}, \nabla \xi_Q \right\rangle, \quad \frac{d}{dt} n = \nu_n - \xi_Q(n).
$$

By Lemma 3.6, for any $\eta \in \mathfrak{g}$, we have

$$
\left\langle \frac{d}{dt} \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right), \eta \right\rangle = \left\langle \mathbf{J} \left(\frac{D}{Dt} \frac{\partial \ell_M}{\partial \nu_n} \right), \eta \right\rangle + \left\langle \frac{\partial \ell_M}{\partial \nu_n}, \nabla_{\dot{n}} \eta_Q(n) \right\rangle
$$

$$
= \left\langle \mathbf{J} \left(\frac{\partial \ell_M}{\partial n} \right), \eta \right\rangle + \left\langle \frac{\partial \ell_M}{\partial \nu_n}, \nabla_{\eta_Q} \xi_Q(n) \right\rangle
$$

$$
+ \left\langle \frac{\partial \ell_M}{\partial \nu_n}, \nabla_{\nu_n} \eta_Q(n) \right\rangle - \left\langle \frac{\partial \ell_M}{\partial \nu_n}, \nabla_{\xi_Q} \eta_Q(n) \right\rangle
$$

$$
= \left\langle \mathbf{J} \left(\frac{\partial \ell_M}{\partial n} \right), \eta \right\rangle + \left\langle \mathcal{F}^\nabla \left(\frac{\partial \ell_M}{\partial \nu_n}, \nu_n \right), \eta \right\rangle \mp \left\langle \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right), [\eta, \xi] \right\rangle,
$$

where we use the equalities $\nabla_{\eta_Q} \xi_Q - \nabla_{\xi_Q} \eta_Q = [\eta_Q, \xi_Q] = \mp [\eta, \xi]_Q$. We thus obtain

$$
\frac{d}{dt} \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right) = \mathbf{J} \left(\frac{\partial \ell_M}{\partial n} \right) + \mathcal{F}^\nabla \left(\frac{\partial \ell_M}{\partial \nu_n}, \nu_n \right) \pm \mathrm{ad}_\xi^* \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right).
$$

Inserting the formula $\frac{\delta \ell_{LP}}{\delta \xi} = \frac{\delta \ell_M}{\delta \xi} + \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right)$ in the first equation of (4.1) and using the previous expression for $\frac{d}{dt} \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right)$, we get the required equation

$$
\frac{d}{dt} \frac{\delta \ell_M}{\delta \xi} = \pm \mathrm{ad}_\xi^* \frac{\delta \ell_M}{\delta \xi} - \mathbf{J} \left(\frac{\partial \ell_M}{\partial n} \right) - \mathcal{F}^\nabla \left(\frac{\partial \ell_M}{\partial \nu_n}, \nu_n \right).
$$

$\square$

Left (right) reduction and right (left) action. In some applications, we need to consider left-invariant (resp. right-invariant) Lagrangians whereas G acts on Q by a right (resp. left) action. We quickly present here the situation, by giving the main formulas in this case. Let G act on the left (resp. right) on Q. We consider here the

case of a right (resp. left) invariant Lagrangian $L: T(G \times Q) \to \mathbb{R}$ under the action of G given by

$$(u_g, u_q) \mapsto (u_g h, h^{-1} u_q) \quad \text{resp.} \quad (u_g, u_q) \mapsto (h u_g, u_q h^{-1}).$$

The Lagrange-Poincaré Lagrangian $\ell_{LP}: TQ \times \mathfrak{g} \to \mathbb{R}$ is now induced by the quotient map

$$T(G \times Q) \to \mathbb{R}, \quad (u_g, u_q) \mapsto (n, \dot{n}, \xi) := (\nu_n + \xi_Q(n), \xi), \tag{4.6}$$

for $n := gq$, $\nu_n := gu_q$, $\xi := u_g g^{-1}$ (resp. $n := qg$, $\nu_n := u_q g$, $\xi := g^{-1} u_g$). The Lagrange-Poincaré equations are now given by

$$
\begin{cases}
\dfrac{d}{dt} \dfrac{\delta \ell_{LP}}{\delta \xi} = -\operatorname{ad}_\xi^* \dfrac{\delta \ell_{LP}}{\delta \xi}, \\[2mm]
\dfrac{D}{Dt} \dfrac{\partial \ell_{LP}}{\partial \dot{n}} - \dfrac{\partial \ell_{LP}}{\partial n} = 0, \quad \dfrac{d}{dt} n = \dot{n},
\end{cases}
\quad \text{resp.} \quad
\begin{cases}
\dfrac{d}{dt} \dfrac{\delta \ell_{LP}}{\delta \xi} = \operatorname{ad}_\xi^* \dfrac{\delta \ell_{LP}}{\delta \xi}, \\[2mm]
\dfrac{D}{Dt} \dfrac{\partial \ell_{LP}}{\partial \dot{n}} - \dfrac{\partial \ell_{LP}}{\partial n} = 0, \quad \dfrac{d}{dt} n = \dot{n}.
\end{cases}
\tag{4.7}
$$

Note the change in the sign when compared to (4.1), and that we now have the relation $\ell_{LP}(n, \dot{n}, \xi) = \ell_M(\dot{n} - \xi_Q(n), \xi)$. Therefore, the conclusions of Lemma 4.3 should be replaced by

$$\frac{\delta \ell_{LP}}{\delta \xi} = \frac{\delta \ell_M}{\delta \xi} - \mathbf{J} \left(\frac{\partial \ell_M}{\partial \nu_n} \right), \quad \frac{\partial \ell_{LP}}{\partial n} = \frac{\partial \ell_M}{\partial n} - \left\langle \frac{\partial \ell_M}{\partial \nu_n}, \nabla \xi_Q \right\rangle, \quad \frac{\partial \ell_{LP}}{\partial \dot{n}} = \frac{\partial \ell_M}{\partial \nu_n}.$$

Thus, equations (4.4) are replaced by

$$
\begin{cases}
\dfrac{d}{dt} \dfrac{\delta \ell_M}{\delta \xi} = \mp \operatorname{ad}_\xi^* \dfrac{\delta \ell_M}{\delta \xi} + \mathbf{J} \left(\dfrac{\partial \ell_M}{\partial n} \right) + \mathcal{F}^\nabla \left(\dfrac{\partial \ell_M}{\partial \nu_n}, \nu_n \right), \\[3mm]
\dfrac{D}{Dt} \dfrac{\partial \ell_M}{\partial \nu_n} = -\left\langle \dfrac{\partial \ell_M}{\partial \nu_n}, \nabla \xi_Q \right\rangle + \dfrac{\partial \ell_M}{\partial n}, \quad \dfrac{d}{dt} n = \nu_n + \xi_Q(n),
\end{cases}
\tag{4.8}
$$

where $-$ (resp. $+$) occurs when G act on Q by a left (resp. right) action.

Alternative form of the equations. For completeness, we give here an alternative form for the second and third equations in systems (4.4), (4.8). This alternative form is analogous to that given in lemmas 3.1, 3.2, and reads

$$\frac{d}{dt} \frac{\partial \ell_M}{\partial \nu_n} = \pm \xi_{T^* Q} \left(\frac{\partial \ell_M}{\partial \nu_n} \right) + \operatorname{Ver}_{\frac{\partial \ell_M}{\partial \nu_n}} \frac{\partial \ell_M}{\partial n} + \operatorname{Hor}_{\frac{\partial \ell_M}{\partial \nu_n}} \nu_n, \tag{4.9}$$

where, for $\alpha_n \in T_n^* Q$, $\operatorname{Hor}_{\alpha_n}: T_n Q \to T_{\alpha_n} T^* Q$ denotes the horizontal lift associated to the Levi-Civita connection on $T^* Q$. Note that we have the formula $\operatorname{Hor}_{\nu_n^\flat} \nu_n = \mathcal{S}(\nu_n)$, where as before, $\mathcal{S} \in \mathfrak{X}(T^* Q)$ is the Hamiltonian vector field associated to the kinetic energy of the Riemannian metric.

5. Optimization, the Lagrangian approach

In this section, we show how to obtain the motion equations associated to the distributed optimization problem by using Lagrangian reduction. More precisely, we will use the metamorphosis reduction, starting from the unreduced Lagrangian associated to ℓ, augmented by the square of the norm of the velocity vector.

Let G act on the left (resp. right) on Q and consider a cost function $\ell := \ell(\xi, n)$ on $\mathfrak{g} \times Q$. Let $\mathcal{L}: TG \times Q \to \mathbb{R}$ be the associated G-invariant Lagrangian on $TG \times Q$. The definition of $\mathcal{L}$ depends on the condition ((A) or (A)') we want to impose.

- If (A) holds, we suppose that $\mathcal{L}$ is invariant under the right (resp. left) action

$$(u_g, q) \mapsto (u_g h, h^{-1} q), \quad \text{resp.} \quad (u_g, q) \mapsto (h u_g, q h^{-1}), \qquad (5.1)$$

  i.e., we define $\mathcal{L}(u_g, q) := \ell(u_g g^{-1}, gq)$, resp. $\mathcal{L}(u_g, q) := \ell(g^{-1} u_g, qg)$.
- If (A)$'$ holds, we suppose that $\mathcal{L}$ is invariant under the left (resp. right) action

$$(u_g, q) \mapsto (h u_g, h q), \quad \text{resp.} \quad (u_g, q) \mapsto (u_g h, q h), \qquad (5.2)$$

  i.e., we define $\mathcal{L}(u_g, q) := \ell(g^{-1} u_g, g^{-1} q)$, resp. $\mathcal{L}(u_g, q) := \ell(u_g g^{-1}, q g^{-1})$.

Definition of the unreduced Lagrangian. The G-invariant Lagrangian $\mathcal{L} : TG \times Q \to \mathbb{R}$ produces the function ℓ by reduction. We now want to modify $\mathcal{L}$ in order to obtain, by reduction, the integrand

$$\ell(\xi, n) + \frac{1}{2\sigma^2} \| \dot{n} \pm \xi_Q(n) \|^2 \qquad (5.3)$$

of the distributed optimization problem. This will be done by constructing, from $\mathcal{L}$, a G-invariant Lagrangian L defined on the tangent bundle $T(G \times Q)$. Of course, the definition of L depends on the condition ((A) or (A)$'$) we want to impose.

- If (A) holds, we define $L : T(G \times Q) \to \mathbb{R}$ by

$$L(u_g, u_q) := \mathcal{L}(u_g, q) + \frac{1}{2\sigma^2} \| g u_q \|^2, \quad \text{resp.} \quad L(u_g, u_q) := \mathcal{L}(u_g, q) + \frac{1}{2\sigma^2} \| u_q g \|^2. \qquad (5.4)$$

- If (A)$'$ holds, we define $L : T(G \times Q) \to \mathbb{R}$ by

$$L(u_g, u_q) := \mathcal{L}(u_g, q) + \frac{1}{2\sigma^2} \| g^{-1} u_q \|^2, \quad \text{resp.} \quad L(u_g, u_q) := \mathcal{L}(u_g, q) + \frac{1}{2\sigma^2} \| u_q g^{-1} \|^2. \qquad (5.5)$$

Of course, the norm appearing in the second term of the Lagrangian is the same as the norm used in the integrand (5.3) of the distributed optimization problem. It is associated to a Riemannian metric on the manifold Q. The presence of the group action in the second term is needed in order to make the Lagrangian G-invariant.

In the particular case where the Riemannian metric is G-invariant, the Lagrangian L is simply given by

$$L(u_g, u_q) := \mathcal{L}(u_g, q) + \frac{1}{2\sigma^2} \| u_q \|^2,$$

and the associated Euler-Lagrange equations for L read

$$\frac{D}{Dt} \dot{q} = \sigma^2 \frac{\partial \mathcal{L}}{\partial q}, \qquad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{g}} - \frac{\partial \mathcal{L}}{\partial g} = 0,$$

where D/Dt denotes the covariant derivative associated to the Riemannian metric on Q.

Lagrangian reduction. Using the quotient maps (4.6) and (4.2) associated to Lagrange-Poincaré reduction, we can compute the reduced Lagrangian associated to L. When the G-invariance (5.1) (condition (A)) holds, we get

$$\ell_{LP}(n, \dot{n}, \xi) = \ell(\xi, n) + \frac{1}{2\sigma^2} \| \dot{n} - \xi_Q(n) \|^2,$$

and when the G-invariance (5.2) (condition (A)') holds, we get

$$\ell_{LP}(n,\dot{n},\xi)=\ell(\xi,n)+\frac{1}{2\sigma^2}\|\dot{n}+\xi_Q(n)\|^2.$$

We have thus obtained the integrand of the distributed optimization problem by Lagrange-Poincaré reduction. However, in order to compute the associated equations of motion, it will be more appropriate to use metamorphosis reduction. For this approach, the reduced Lagrangian is readily seen to be

$$\ell_M(\nu_n,\xi)=\ell(\xi,n)+\frac{1}{2\sigma^2}\|\nu_n\|^2,$$

in all cases.

We now compute the reduced equations of motions. Since the functional derivatives of ℓ_M are

$$\frac{\delta\ell_M}{\delta\xi}=\frac{\delta\ell}{\delta\xi},\qquad\frac{\partial\ell_M}{\partial\nu_n}=\frac{1}{\sigma^2}\nu_n^\flat,\quad\text{and}\qquad\frac{\partial\ell_M}{\partial n}=\frac{\partial\ell}{\partial n},$$

the reduced equations (4.8) (associated to condition (A)) and (4.4) (associated to condition (A)$'$) become, respectively,

$$\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=\mp\operatorname{ad}_\xi^*\dfrac{\delta\ell}{\delta\xi}+\mathbf{J}\left(\dfrac{\partial\ell}{\partial n}\right)+\dfrac{1}{\sigma^2}\mathcal{F}^\nabla\left(\nu_n^\flat,\nu_n\right),\\[2mm]\dfrac{D}{Dt}\nu_n^\flat=-\langle\nu_n^\flat,\nabla\xi_Q\rangle+\sigma^2\dfrac{\partial\ell}{\partial n},\quad\dot{n}=\nu_n+\xi_Q(n)\end{cases}\tag{5.6}$$

and

$$\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=\pm\operatorname{ad}_\xi^*\dfrac{\delta\ell}{\delta\xi}-\mathbf{J}\left(\dfrac{\partial\ell}{\partial n}\right)-\dfrac{1}{\sigma^2}\mathcal{F}^\nabla\left(\nu_n^\flat,\nu_n\right),\\[2mm]\dfrac{D}{Dt}\nu_n^\flat=\langle\nu_n^\flat,\nabla\xi_Q\rangle+\sigma^2\dfrac{\partial\ell}{\partial n},\quad\dot{n}=\nu_n-\xi_Q(n).\end{cases}\tag{5.7}$$

These are exactly the equations (3.32) and (3.33) that are verified by the extremals of the distributed optimization problem, obtained here by metamorphosis reduction.

REMARK 5.1. The fact that metamorphosis reduction recovers the motion equations verified by the extremals of the distributed optimization problem is natural in the following sense. The extremals are given by the *unconstrained variational principle*

$$0=\delta S_p=\delta\int_0^T\left(\ell(\xi,n)+\frac{1}{2\sigma^2}\|\dot{n}\pm\xi_Q(n)\|^2\right)dt;$$

this gives the stationarity conditions (3.25), (3.26). These imply (but are not equivalent to) the metamorphosis equations (3.32), (3.33) obtained form the same action S_p under *constrained variations*.

6. Hamilton-Poincaré and metamorphosis reduction

In this section, we present the Hamiltonian side of the two Lagrangian reduction approaches described in Section 4.

As before, we let G act on the left (resp. right) on Q. We consider a left (resp. right)-invariant Hamiltonian $H:T^*(G\times Q)\to\mathbb{R}$ under the action of G given by

$$(\alpha_g,\alpha_q)\mapsto(h\alpha_g,h\alpha_q)\quad\text{resp.}\quad(\alpha_g,\alpha_q)\mapsto(\alpha_gh,\alpha_qh).$$

As before, there are two reduction processes. The first uses Hamilton-Poincaré reduction (see [18]) and the second is the Hamiltonian version of the metamorphosis reduction described in Section 4.

THEOREM 6.1 (Hamilton-Poincaré reduction). *Let $\alpha_g \in T^*G$ and $\alpha_q \in T^*Q$ be two curves and define the curves $\pi_n := g^{-1}\alpha_q \in T^*Q$ and $\mu := g^{-1}\alpha_g + \mathbf{J}(g^{-1}\alpha_q) \in \mathfrak{g}^*$ (resp. $\pi_n := \alpha_q g^{-1} \in T^*Q$ and $\mu := \alpha_g g^{-1} + \mathbf{J}(\alpha_q g^{-1}) \in \mathfrak{g}^*$).*

Then (α_g, α_q) is a solution of the canonical Hamilton equations for H on $T^*G \times T^*Q$ if and only if (π_n, μ) is a solution of the Hamilton-Poincaré equations

$$
\begin{cases}
\dfrac{d}{dt}\mu = \mathrm{ad}^*_{\frac{\delta h_{HP}}{\delta \mu}}\mu, \\[2mm]
\dfrac{d}{dt}n = \dfrac{\partial h_{HP}}{\partial \pi}, \quad \dfrac{d}{dt}\pi = -\dfrac{\partial h_{HP}}{\partial n},
\end{cases}
\quad resp. \quad
\begin{cases}
\dfrac{d}{dt}\mu = -\mathrm{ad}^*_{\frac{\delta h_{HP}}{\delta \mu}}\mu, \\[2mm]
\dfrac{d}{dt}n = \dfrac{\partial h_{HP}}{\partial \pi}, \quad \dfrac{d}{dt}\pi = -\dfrac{\partial h_{HP}}{\partial n},
\end{cases}
$$

(6.1)

where the Hamilton-Poincaré Hamiltonian $h_{HP} = h_{HP}(\pi_n, \mu) : T^*Q \times \mathfrak{g}^* \to \mathbb{R}$ is induced from H by the quotient map

$$
T^*(G \times Q) \to T^*Q \times \mathfrak{g}^*, \quad (\alpha_g, \alpha_q) \mapsto (\pi_n, \mu) := (\pi_n, \kappa + \mathbf{J}(\pi_n)) \tag{6.2}
$$

for $n := g^{-1}q$, $\pi_n := g^{-1}\alpha_q$, $\kappa := g^{-1}\alpha_g$, (resp. $n := qg^{-1}$, $\pi_n := \alpha_q g^{-1}$, $\kappa := \alpha_g g^{-1}$).

In the Hamilton-Poincaré equations (6.1), the second equation is written in Darboux coordinates. One can write it intrinsically as

$$
\frac{d}{dt}\pi_n = X_{h_{HP}}(\pi_n),
$$

where $X_{h_{HP}}$ is the Hamiltonian vector field of h_{HP} viewed as a function on T^*Q, the variable $\mu \in \mathfrak{g}^*$ being considered as a parameter.

Proof. We treat the case of a left action and apply the results in [18]. The coadjoint bundle $\mathrm{Ad}^*(G \times Q)$ can be identified with the trivial vector bundle $Q \times \mathfrak{g}^*$ via the identification $[(g,q), \mu] \simeq (g^{-1}q, \mathrm{Ad}^*_g \mu)$. Using the principal connection $\mathcal{A}(u_g, u_q) := u_g g^{-1}$, the diffeomorphism $(T^*(G \times Q))/G \cong T^*Q \times \mathfrak{g}^*$ is given by $[\alpha_g, \alpha_q] \mapsto (g^{-1}\alpha_q, g^{-1}\alpha_g + \mathbf{J}(g^{-1}\alpha_q))$. Indeed, the horizontal-lift associated to $\mathcal{A}$ reads $\mathrm{Hor}_{(g,q)} : T_nQ \to T_gG \times T_qQ$, $\mathrm{Hor}_{(g,q)} v_n = (0_g, gv_n)$, its dual map is $\left[\mathrm{Hor}_{(g,q)}\right]^*(\alpha_g, \alpha_q) = g^{-1}\alpha_q$, and the momentum map $\mathbb{J} : T^*(G \times Q) \to \mathfrak{g}^*$ is $\mathbb{J}(\alpha_g, \alpha_q) = \alpha_g g^{-1} + \mathbf{J}(\alpha_q)$. Thus, the Hamilton-Poincaré reduction map has the required expression (6.2). Since the chosen principal connection is flat, we obtain the Hamilton-Poincaré equations (6.1). $\square$

It is convenient to write the equations of motion (6.1) in matrix form, namely

$$
\frac{d}{dt}\begin{bmatrix} \mu \\ \pi_n \end{bmatrix} = \begin{bmatrix} \pm \mathrm{ad}^*_\square \mu & 0 \\ 0 & \Omega^\sharp_{can}(\pi_n) \end{bmatrix} \begin{bmatrix} \dfrac{\delta h_{HP}}{\delta \mu} \\[2mm] \mathbf{d}_{T^*Q} h_{HP} \end{bmatrix}, \tag{6.3}
$$

where $\mathbf{d}_{T^*Q}$ is the exterior derivative on T^*Q.

For the same G-invariant Hamiltonian $H : T^*(G \times Q) \to \mathbb{R}$ as before, we define another reduced Hamiltonian $h_M = h_M(\pi_n, \kappa) : T^*Q \times \mathfrak{g}^* \to \mathbb{R}$ associated to the quotient map

$$
T^*(G \times Q) \to T^*Q \times \mathfrak{g}^*,
$$

$$(\alpha_g, \alpha_q) \mapsto (\pi_n, \kappa) := (g^{-1}\alpha_q, g^{-1}\alpha_g), \quad \text{resp.} \quad (\pi_n, \kappa) := (\alpha_q g^{-1}, \alpha_g g^{-1}).$$

This reduced Hamiltonian differs from the Hamilton-Poincaré Hamiltonian h_{HP} defined above, but one can pass from the one to the other by the vector bundle isomorphism

$$T^*Q \times \mathfrak{g}^* \to T^*Q \times \mathfrak{g}^*, \quad (\pi_n, \mu) \mapsto (\pi_n, \mu - \mathbf{J}(\pi_n)),$$

that is, we have

$$h_{HP}(\pi_n, \kappa + \mathbf{J}(\pi_n)) = h_M(\pi_n, \kappa),$$

for both the left and right cases. Of course, the previous isomorphism is the dual map to (4.3).

As on the Lagrangian side, we fix a Riemannian metric g on Q. This allows us to write the reduced Hamilton equation a bit more explicitly. Note, however, that it is possible to write the reduced Hamilton equations without the help of a metric; see (6.5) below.

THEOREM 6.2 (Metamorphosis reduction). *Let $\alpha_g \in T^*G$ and $\alpha_q \in T^*Q$ be two curves and define the curves $\pi_n := g^{-1}\alpha_q \in T^*Q$ and $\kappa := g^{-1}\alpha_g \in \mathfrak{g}^*$ (resp. $\pi_n := \alpha_q g^{-1} \in T^*Q$ and $\kappa := \alpha_g g^{-1} \in \mathfrak{g}^*$).*

Then (α_g, α_q) is a solution of the canonical Hamilton equations for H on $T^*G \times T^*Q$ if and only if (π_n, κ) is a solution of the equations

$$
\begin{cases}
\dfrac{d}{dt}\kappa = \pm \operatorname{ad}^*_{\frac{\delta h_M}{\delta \kappa}} \kappa + \mathbf{J}\left(\dfrac{\partial h_M}{\partial n}\right) - \mathcal{F}^\nabla\left(\pi_n, \dfrac{\partial h_M}{\partial \pi_n}\right), \\[4mm]
\dfrac{d}{dt}\pi_n = -\left(\dfrac{\delta h_M}{\delta \kappa}\right)_{T^*Q}(\pi_n) + X_{h_M}(\pi_n),
\end{cases}
\tag{6.4}
$$

where X_{h_M} is the Hamiltonian vector field associated to h_M viewed as a function of π_n.

Proof. We treat simultaneously the case of left and right actions and apply Poisson reduction. The reduced Poisson structure on $T^*Q \times \mathfrak{g}^*$ associated to the quotient map $(\alpha_g, \alpha_q) \mapsto (g^{-1}\alpha_g, g^{-1}\alpha_q)$, resp. $(\alpha_g, \alpha_q) \mapsto (\alpha_g g^{-1}, \alpha_q g^{-1})$ is given for any $f, g \in C^\infty(T^*Q \times \mathfrak{g}^*)$ by

$$\{f, g\}_{T^*Q \times \mathfrak{g}^*} = \mp \left\langle \mu, \left[\dfrac{\delta f}{\delta \mu}, \dfrac{\delta g}{\delta \mu}\right]\right\rangle - \left\langle \boldsymbol{\mathcal{J}}(\mathbf{d}f(\pi_n)), \dfrac{\delta g}{\delta \mu}\right\rangle + \left\langle \boldsymbol{\mathcal{J}}(\mathbf{d}g(\pi_n)), \dfrac{\delta f}{\delta \mu}\right\rangle + \{f, g\}_{T^*Q},$$

where $\boldsymbol{\mathcal{J}} : T^*(T^*Q) \to \mathfrak{g}^*$ is the cotangent bundle momentum map and the last term is the canonical Poisson bracket on T^*Q; see Proposition 10.3.1 in [53]. Consequently, the reduced Hamilton's equation are

$$
\begin{cases}
\dfrac{d}{dt}\kappa = \pm \operatorname{ad}^*_{\frac{\delta h_M}{\delta \kappa}} \kappa + \boldsymbol{\mathcal{J}}(\mathbf{d}h_M(\pi_n)), \\[4mm]
\dfrac{d}{dt}\pi_n = -\left(\dfrac{\delta h_M}{\delta \kappa}\right)_{T^*Q}(\pi_n) + X_{h_M}(\pi_n).
\end{cases}
\tag{6.5}
$$

Now it suffices to decompose the derivative $\mathbf{d}h_M$ into the vertical (fiber) and horizontal partial derivatives and use Lemma 3.4 to write

$$\boldsymbol{\mathcal{J}}(\mathbf{d}h_M(\pi_n)) = \mathbf{J}\left(\dfrac{\partial h_M}{\partial n}\right) - \mathcal{F}^\nabla\left(\pi_n, \dfrac{\partial h_M}{\partial \pi_n}\right).
\tag{6.6}$$

This proves the result.   □

Equations (6.4) can be conveniently written in matrix form as

$$
\frac{d}{dt}\begin{bmatrix} \kappa \\ \pi_n \end{bmatrix} = \begin{bmatrix} \pm \operatorname{ad}^*_\square \kappa & \mathcal{J} \\ -(\square)_{T^*Q}(\pi_n) & \Omega^\sharp_{can}(\pi_n) \end{bmatrix} \begin{bmatrix} \frac{\delta h_M}{\delta \kappa} \\ \mathbf{d}_{T^*Q} h_M \end{bmatrix},
\tag{6.7}
$$

where in the (1,2) entry one uses formula (6.6). We shall see in Section 8.1 that if Q is a representation space of G, this formula gives rise to a Lie-Poisson equation on a semidirect product with a cocycle.

Left (right) reduction and right (left) action. We quickly present here the equations arising when the Hamiltonian $H : T^*(G \times Q) \to \mathbb{R}$ is invariant under the action of G given by

$$
(\alpha_g, \alpha_q) \mapsto (\alpha_g h, h^{-1} \alpha_q) \qquad \text{resp.} \qquad (\alpha_g, \alpha_q) \mapsto (h\alpha_g, \alpha_q h^{-1}).
$$

The Hamilton-Poincaré Hamiltonian $h_{HP} : T^*Q \times \mathfrak{g}^* \to \mathbb{R}$ is now induced by the quotient map

$$
T^*(G \times Q) \to \mathbb{R}, \quad (\alpha_g, \alpha_q) \mapsto (\pi_n, \mu) := (\pi_n, \kappa - \mathbf{J}(\pi_n)),
\tag{6.8}
$$

for $n := gq$, $\pi_n := g\alpha_q$, $\kappa := \alpha_g g^{-1}$ (resp. $n := qg$, $\pi_n := \alpha_q g$, $\kappa := g^{-1} u_g$). The resulting Hamilton-Poincaré equations are given by

$$
\begin{cases} \dfrac{d}{dt}\mu = -\operatorname{ad}^*_{\frac{\delta h_{HP}}{\delta \mu}} \mu, \\ \dfrac{d}{dt}n = \dfrac{\partial h_{HP}}{\partial \pi}, \quad \dfrac{d}{dt}\pi = -\dfrac{\partial h_{HP}}{\partial n}, \end{cases} \quad \text{resp.} \quad \begin{cases} \dfrac{d}{dt}\mu = \operatorname{ad}^*_{\frac{\delta h_{HP}}{\delta \mu}} \mu, \\ \dfrac{d}{dt}n = \dfrac{\partial h_{HP}}{\partial \pi}, \quad \dfrac{d}{dt}\pi = -\dfrac{\partial h_{HP}}{\partial n}. \end{cases}
\tag{6.9}
$$

Note the change in the sign when compared to (6.1), and that we now have the relation $h_{HP}(\pi_n, \kappa - \mathbf{J}(\pi_n)) = h_M(\pi_n, \kappa)$.

Likewise, equations (6.4) are replaced by

$$
\begin{cases} \dfrac{d}{dt}\kappa = \mp \operatorname{ad}^*_{\frac{\delta h_M}{\delta \kappa}} \kappa - \mathbf{J}\left(\dfrac{\partial h_M}{\partial n}\right) + \mathcal{F}^\nabla\left(\pi_n, \dfrac{\partial h_M}{\partial \pi_n}\right), \\ \dfrac{d}{dt}\pi_n = \left(\dfrac{\delta h_M}{\delta \kappa}\right)_{T^*Q}(\pi_n) + X_{h_M}(\pi_n), \end{cases}
\tag{6.10}
$$

where $-$ (resp. $+$) occurs when G acts on Q by a left (resp. right) action. As in (6.3) and (6.7), equations (6.9) and (6.10) may be re-expressed in matrix form as

$$
\frac{d}{dt}\begin{bmatrix} \mu \\ \pi_n \end{bmatrix} = \begin{bmatrix} \mp \operatorname{ad}^*_\square \mu & 0 \\ 0 & \Omega^\sharp_{can}(\pi_n) \end{bmatrix} \begin{bmatrix} \frac{\delta h_{HP}}{\delta \mu} \\ \mathbf{d}_{T^*Q} h_{HP} \end{bmatrix},
\tag{6.11}
$$

$$
\frac{d}{dt}\begin{bmatrix} \kappa \\ \pi_n \end{bmatrix} = \begin{bmatrix} \mp \operatorname{ad}^*_\square \kappa & -\mathcal{J} \\ (\square)_{T^*Q}(\pi_n) & \Omega^\sharp_{can}(\pi_n) \end{bmatrix} \begin{bmatrix} \frac{\delta h_M}{\delta \kappa} \\ \mathbf{d}_{T^*Q} h_M \end{bmatrix}.
\tag{6.12}
$$

REMARK 6.1 (Link with the untangling map). Recall that the vector bundle isomorphism

$$(\xi, \nu_n) \in \mathfrak{g} \times TN \mapsto (\xi, \nu_n \pm \xi_Q(n)) = (\xi, n, \dot{n}) \in \mathfrak{g} \times TN$$

allows one to pass from the metamorphosis reduced equation to the Lagrange-Poincaré equations. Its dual map

$$(\mu, \pi_n) \in \mathfrak{g}^* \times T^*N \mapsto (\mu \pm \mathbf{J}(\pi_n), \pi_n) = (\kappa, \pi_n) \in \mathfrak{g}^* \times T^*N,$$

naturally passes from the Hamilton-Poincaré description to the metamorphosis approach. The inverse of this map is known as the *untangling map* in applications ([34]) since it transforms the Hamiltonian structure of the metamorphosis equation into the direct sum of the Lie-Poisson bracket on $\mathfrak{g}^*$ and the canonical Poisson bracket on T^*N; see (6.1)-(6.7) and (6.11)-(6.12). Recent theoretical developments and new applications of the untangling map appear in [30].

Legendre transformation and alternative formulation. When the Hamiltonian H comes from a Lagrangian L by Legendre transformation, then we have the following relations between the reduced objects:

$$h_M(\pi_n, \kappa) = \langle \pi_n, \nu_n \rangle + \langle \kappa, \xi \rangle - \ell_M(\nu_n, \xi), \quad \kappa = \frac{\delta \ell_M}{\delta \xi}, \quad \pi_n = \frac{\partial \ell_M}{\partial \nu_n}.$$

The partial derivatives with respect to n are related by the formula

$$\frac{\partial h_M}{\partial n} = -\frac{\partial \ell_M}{\partial n}.$$

In this case, the reduced equations on the Lagrangian and Hamiltonian side ((4.4), (4.8) and (6.4), (6.10)) are readily seen to be equivalent. To see this, it suffices to use the formula

$$X_h(\alpha_n) = \mathrm{Hor}_{\alpha_n} \frac{\partial h}{\partial \alpha_n} - \mathrm{Ver}_{\alpha_n} \frac{\partial h}{\partial n}$$

for the Hamiltonian vector field, together with the alternative formulation (4.9) for the reduced equations on the Lagrangian side. This also shows that the second equation of the systems (6.4) and (6.10) can be equivalently written as

$$\frac{D}{Dt} \pi_n = \mp \left\langle \pi_n, \nabla \left(\frac{\delta h_M}{\delta \kappa} \right)_Q (n) \right\rangle - \frac{\partial h_M}{\partial n}, \qquad \frac{d}{dt} n = \pm \left(\frac{\delta h_M}{\delta \kappa} \right)_Q (n) + \frac{\partial h_M}{\partial \pi_n}.$$

7. Optimization, the Hamiltonian approach

Suppose we are given a left (resp. right) action of G on Q and a cost function $\ell = \ell(\xi, q)$ on $\mathfrak{g} \times Q$. Let the map $\xi \mapsto \frac{\delta \ell}{\delta \xi}$ be a diffeomorphism and consider the associated Hamiltonian $h : \mathfrak{g}^* \times Q \to \mathbb{R}$ defined by

$$h(\mu, q) := \langle \mu, \xi \rangle - \ell(\xi, q), \quad \frac{\delta \ell}{\delta \xi} = \mu.$$

As in Section 5 for ℓ, the function h induces a G-invariant function $\mathcal{H} : T^*G \times Q \to \mathbb{R}$. Of course, $\mathcal{H}$ can be obtained from $\mathcal{L}$ by a Legendre transformation, the variable q being considered as a parameter. Recall that given a Riemannian metric g on Q, we

associated to $\mathcal{L}$ a G-invariant Lagrangian on $T(G \times Q)$ by adding to $\mathcal{L}$ a G-invariant expression involving the norm of the vector in TQ; see (5.4), (5.5). For example, in the case of condition (A)$'$ and if G acts on the left we have defined

$$L(u_g, u_q) := \mathcal{L}(u_g, q) + \frac{1}{2\sigma^2} \|g^{-1}u_q\|^2.$$

Taking the Legendre transformation of this hyperregular Lagrangian yields the G-invariant Hamiltonian H on $T^*(G \times Q)$ given by

$$H(\alpha_g, \alpha_q) = \mathcal{H}(\alpha_g, q) + \frac{\sigma^2}{2} \|g^{-1}\alpha_q\|^2.$$

The reduced Hamiltonian associated to metamorphosis reduction reads

$$h_M(\pi_n, \kappa) = h(\kappa, n) + \frac{\sigma^2}{2} \|\pi_n\|^2.$$

When condition (A)$'$ is assumed, the reduced Hamilton-Poincaré Hamiltonian reads

$$h_{HP}(\pi_n, \mu) = h_M(\pi_n, \mu - \mathbf{J}(\pi_n)) = h(\mu - \mathbf{J}(\pi_n), n) + \frac{\sigma^2}{2} \|\pi_n\|^2.$$

In the case of condition (A), we have

$$h_{HP}(\pi_n, \mu) = h_M(\pi_n, \mu + \mathbf{J}(\pi_n)) = h(\mu + \mathbf{J}(\pi_n), n) + \frac{\sigma^2}{2} \|\pi_n\|^2.$$

Using the relations

$$\frac{\partial h_M}{\partial \pi_n} = \sigma^2 \pi_n^\sharp \quad \text{and} \quad \frac{\partial h_M}{\partial n} = \frac{\partial h}{\partial n},$$

the reduced equations (6.4) and (6.10) become, respectively,

$$\begin{cases} \dfrac{d}{dt}\kappa = \pm \operatorname{ad}^*_{\frac{\delta h}{\delta \kappa}} \kappa + \mathbf{J}\left(\dfrac{\partial h}{\partial n}\right), \\[2mm] \dfrac{d}{dt}\pi_n = -\left(\dfrac{\delta h}{\delta \kappa}\right)_{T^*Q}(\pi_n) - \operatorname{Ver}_{\pi_n}\dfrac{\partial h}{\partial n} + \sigma^2 \mathcal{S}(\pi_n) \end{cases} \tag{7.1}$$

and

$$\begin{cases} \dfrac{d}{dt}\kappa = \mp \operatorname{ad}^*_{\frac{\delta h}{\delta \kappa}} \kappa - \mathbf{J}\left(\dfrac{\partial h}{\partial n}\right), \\[2mm] \dfrac{d}{dt}\pi_n = \left(\dfrac{\delta h}{\delta \kappa}\right)_{T^*Q}(\pi_n) - \operatorname{Ver}_{\pi_n}\dfrac{\partial h}{\partial n} + \sigma^2 \mathcal{S}(\pi_n), \end{cases} \tag{7.2}$$

where $\mathcal{S} \in \mathfrak{X}(T^*Q)$ is the Hamiltonian vector field associated to the kinetic energy

$$\frac{1}{2}\|\pi_n\|^2 = \frac{1}{2}g(\pi_n^\sharp, \pi_n^\sharp).$$

These equations recover the motion equations associated to the distributed optimization, in Hamiltonian form; cf. Remark 3.2.

8. Examples

In this section we apply the general theory to various group actions.

8.1. Action by representation and advected quantities. Let G be a Lie group acting by left (resp. right) representation on the dual vector space $Q = V^*$. Given a Lie algebra element ξ, we denote by $\xi_{V^*}(a) = \xi a$ (resp. $\xi_{V^*}(a) = a\xi$) the associated infinitesimal generator. Using the diamond operator $\diamond : V \times V^* \to \mathfrak{g}^*$ defined for $p \in V$ and $a \in V^*$ by $\langle p \diamond a, \xi \rangle := -\langle \xi a, p \rangle$ (resp. $\langle p \diamond a, \xi \rangle := -\langle a\xi, p \rangle$), for any $\xi \in \mathfrak{g}$, the cotangent bundle momentum map is $\mathbf{J}(a,p) = -p \diamond a$. Since Q is a vector space, we can choose ∇ to be the ordinary derivative; therefore we have $\nabla_b \xi_{V^*}(a) = \xi b$ (resp. $\nabla_b \xi_{V^*}(a) = b\xi$), for all $a, b \in V^*$. From (3.31), we thus obtain the expression $\mathcal{F}^\nabla((a,v),(a,b)) = -v \diamond b$.

Metamorphosis reduction and Lie-Poisson formulation with cocycles. Assume that V is a left representation space of G and that reduction has been performed on the left. The other cases have similar formulations. In view of the identities above, equations (4.4) become

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta \ell_M}{\delta \xi} = \mathrm{ad}^*_\xi \dfrac{\delta \ell_M}{\delta \xi} + \dfrac{\delta \ell_M}{\delta a} \diamond a + \dfrac{\delta \ell_M}{\delta \nu} \diamond \nu, \\[2mm] \dfrac{d}{dt}\dfrac{\delta \ell_M}{\delta \nu} = -\xi \dfrac{\delta \ell_M}{\delta \nu} + \dfrac{\delta \ell_M}{\delta a}, \quad \dfrac{d}{dt} a = -\xi a + \nu, \end{cases}$$

where $\ell_M = \ell_M(\xi, a, \nu) : \mathfrak{g} \times V^* \times V^* \to \mathbb{R}$ is the reduced Lagrangian. Performing the Legendre transformation $h_M(\kappa, a, \pi) := \langle \kappa, \xi \rangle + \langle \pi, \nu \rangle - \ell_M(\xi, a, \nu)$, where

$$\frac{\delta \ell_M}{\delta \xi} = \kappa, \quad \frac{\delta \ell_M}{\delta \nu} = \pi,$$

one finds the corresponding Hamilton equations for $h_M = h_M(\kappa, a, \pi) : \mathfrak{g}^* \times V^* \times V \to \mathbb{R}$ as

$$\begin{cases} \dfrac{d}{dt}\kappa = \mathrm{ad}^*_{\frac{\delta h_M}{\delta \kappa}}\kappa - \dfrac{\delta h_M}{\delta a} \diamond a + \pi \diamond \dfrac{\delta h_M}{\delta \pi}, \\[2mm] \dfrac{d}{dt}\pi = -\dfrac{\delta h_M}{\delta \kappa}\pi - \dfrac{\delta h_M}{\delta a}, \quad \dfrac{d}{dt}a = -\dfrac{\delta h_M}{\delta \kappa}a + \dfrac{\delta h_M}{\delta \pi}. \end{cases} \tag{8.1}$$

These equations recover (6.4) for the case of a left G-representation.

Note that the inverse Legendre transformation (assuming it is a diffeomorphism) is given by $\delta h_M/\delta \kappa = \xi$, $\delta h_M/\delta \pi = \nu$ and that $\delta h_M/\delta a = -\delta \ell_M/\delta a$.

The proof of the following theorem is a direct verification.

THEOREM 8.1. *The equations of motion* (8.1) *are Lie-Poisson on* $(\mathfrak{g} \circledS (V \times V^*))^*$ *with the cocycle* $C : (V \times V^*) \times (V \times V^*) \to \mathbb{R}$ *given by the canonical symplectic structure* Ω_{can} *on* $T^*V = V \times V^*$, *where the* $\mathfrak{g}$-*left representation on* $V \times V^*$ *is given by* $(\xi, v) \mapsto \xi v$, $(\xi, \nu) \mapsto \xi \nu$, *for* $\xi \in \mathfrak{g}$, $v \in V$, $\nu \in V^*$. *Thus these equations can be written in matrix form as*

$$\frac{\partial}{\partial t}\begin{bmatrix} \kappa \\ a \\ \pi \end{bmatrix} = \begin{bmatrix} \mathrm{ad}^*_\square \kappa & -\square \diamond a & \pi \diamond \square \\ -\square a & 0 & 1 \\ -\square \pi & -1 & 0 \end{bmatrix} \begin{bmatrix} \delta h_M/\delta \kappa \\ \delta h_M/\delta a \\ \delta h_M/\delta \pi \end{bmatrix}. \tag{8.2}$$

The Clebsch optimal control approach. Given a cost function $\ell:\mathfrak{g}\times V^*\to\mathbb{R}$, the Clebsch optimal control problem with condition (A)$'$ yields (for left representation) the stationarity conditions

$$\frac{\delta\ell}{\delta\xi}=-\mathbf{J}(a,p)=p\diamond a,\qquad \dot a=-\xi a,\qquad \dot p=-\xi p+\frac{\partial\ell}{\partial a}. \tag{8.3}$$

For a representation on the right, one replaces ξa, ξp by $a\xi$, $p\xi$. These equations imply the Euler-Poincaré equations of motion

$$\frac{d}{dt}\frac{\delta\ell}{\delta\xi}=\pm\operatorname{ad}^*_\xi\frac{\delta\ell}{\delta\xi}+\frac{\partial\ell}{\partial a}\diamond a.$$

When condition (A) is assumed, we get the stationarity conditions

$$\frac{\delta\ell}{\delta\xi}=\mathbf{J}(a,p)=-p\diamond a,\qquad \dot a=\xi a,\qquad \dot p=\xi p+\frac{\partial\ell}{\partial a}, \tag{8.4}$$

and the motion equations

$$\frac{d}{dt}\frac{\delta\ell}{\delta\xi}=\mp\operatorname{ad}^*_\xi\frac{\delta\ell}{\delta\xi}-\frac{\partial\ell}{\partial a}\diamond a.$$

These are the *Euler-Poincaré equations for semidirect products*, useful for the study of physical systems with advected quantities; see [42, 43].

Note that when the Lagrangian ℓ is given by the kinetic energy associated to an inner product on $\mathfrak{g}$, the control is given by $\xi=\pm(p\diamond a)^\sharp$, where $\sharp:\mathfrak{g}^*\to\mathfrak{g}$ is associated to the inner product on $\mathfrak{g}$. We get the equations

$$\dot a+(p\diamond a)^\sharp a=0,\quad \dot p+(p\diamond a)^\sharp p=0.$$

This is the abstract formulation of the *double bracket equations*; see Section 8.1.2 below.

The distributed optimization approach. In order to state the optimization problem with penalty, we endow V^* with a inner product. The corresponding functional is thus

$$\ell(\xi,a)+\frac{1}{2\sigma^2}\|\dot a\pm\xi a\|^2.$$

Recall that in the present case we have $\mathcal{F}^\nabla((a,v),(a,b))=-v\diamond b$. So, if condition (A)$'$ is assumed, the motion equations (3.33) read

$$\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=\operatorname{ad}^*_\xi\dfrac{\delta\ell}{\delta\xi}+\dfrac{\delta\ell}{\delta a}\diamond a+\dfrac{1}{\sigma^2}\nu^\flat\diamond\nu,\\[2mm]\dfrac{d}{dt}\nu^\flat-\sigma^2\dfrac{\delta\ell}{\delta a}=-\xi\nu^\flat,\quad \nu=\dot a+\xi a,\end{cases}\quad\text{resp.}\quad\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=-\operatorname{ad}^*_\xi\dfrac{\delta\ell}{\delta\xi}+\dfrac{\delta\ell}{\delta a}\diamond a+\dfrac{1}{\sigma^2}\nu^\flat\diamond\nu,\\[2mm]\dfrac{d}{dt}\nu^\flat-\sigma^2\dfrac{\delta\ell}{\delta a}=-\nu^\flat\xi,\quad \nu=\dot a+a\xi,\end{cases} \tag{8.5}$$

where $\flat:V^*\to V$ is the flat isomorphism associated to the inner product on V^*. When condition (A) is assumed, we have (see (3.32))

$$\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=-\operatorname{ad}^*_\xi\dfrac{\delta\ell}{\delta\xi}-\dfrac{\delta\ell}{\delta a}\diamond a-\dfrac{1}{\sigma^2}\nu^\flat\diamond\nu,\\[2mm]\dfrac{d}{dt}\nu^\flat-\sigma^2\dfrac{\delta\ell}{\delta a}=\xi\nu^\flat,\quad \nu=\dot a-\xi a,\end{cases}\quad\text{resp.}\quad\begin{cases}\dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=\operatorname{ad}^*_\xi\dfrac{\delta\ell}{\delta\xi}-\dfrac{\delta\ell}{\delta a}\diamond a-\dfrac{1}{\sigma^2}\nu^\flat\diamond\nu,\\[2mm]\dfrac{d}{dt}\nu^\flat-\sigma^2\dfrac{\delta\ell}{\delta a}=\nu^\flat\xi,\quad \nu=\dot a-a\xi.\end{cases} \tag{8.6}$$

As we have seen in the general theory, these motion equations arise by metamorphosis reduction associated to the Lagrangian $\mathcal{L}(u_g,a)+\frac{1}{2\sigma^2}\|\dot{a}\|^2$. They can be obtained by the stationarity conditions (3.25), (3.26). In our case, for a left representation they read

$$\frac{\delta\ell}{\delta\xi}=\pm\frac{1}{\sigma^2}\nu^\flat\diamond a,\quad \dot{a}=\mp\xi a+\nu,\quad \dot{\nu}^\flat=\mp\xi\nu^\flat+\sigma^2\frac{\partial\ell}{\partial a}.$$

As usual, to compare these conditions with the stationarity conditions (8.3), (8.4) given by the Clebsch optimal control approach, we define

$$p:=\frac{1}{\sigma^2}\nu^\flat\in V,$$

and we get

$$\frac{\delta\ell}{\delta\xi}=\pm p\diamond a,\qquad \dot{a}=\mp\xi a+\sigma^2 p^\sharp,\qquad \dot{p}=\mp\xi p+\frac{\partial\ell}{\partial a}. \qquad (8.7)$$

For a right representation one simply replaces ξa, ξp by $a\xi$, $p\xi$.

When the Lagrangian ℓ is given by the kinetic energy associated to the inner product on $\mathfrak{g}$, the control is given by $\xi=\pm(p\diamond a)^\sharp$, and we get the equations

$$\dot{a}+(p\diamond a)^\sharp a=\sigma^2 p^\sharp,\qquad \dot{p}+(p\diamond a)^\sharp p=0.$$

This is the abstract formulation of the *double bracket equations*, modified by the extra term $\sigma^2 p^\sharp$; see Section 8.1.2 below. Note that in the formula above, there are two different sharp operators, $\sharp:\mathfrak{g}^*\to\mathfrak{g}$ and $\sharp:V\to V^*$, associated to the inner products on $\mathfrak{g}$ and V^*, respectively.

Note that, consistently with Theorem 3.7, if the inner product is G-invariant, then $\nu^\flat\diamond\nu=0$. This has already been noticed in the Remark 2.4 of the introduction, for the case of an isotropic inner product.

8.1.1. Heavy top.   Consider the evolution equations for a state system in the frame of a rotating body

$$\dot{\mathbf{X}}=\mathbf{X}\times\boldsymbol{\Omega}, \qquad (8.8)$$

for vector state and control variables $\mathbf{X},\boldsymbol{\Omega}\in\mathbb{R}^3$ related to the rotation matrix $O\in SO(3)$ by $\mathbf{X}=O^{-1}\hat{\mathbf{z}}$ and $\boldsymbol{\Omega}\times=O^{-1}\dot{O}\in\mathfrak{so}(3)$. These vectors are, respectively, the vertical spatial axis as seen from the rotating body and the body angular velocity vector.

We choose to optimize a cost functional consisting of the difference between the rotational kinetic energy and the gravitational potential energy, subject to a penalty imposed by the state system (8.8). This cost function is

$$S_p=\int_0^T\left(\ell(\boldsymbol{\Omega},\mathbf{X})+\frac{1}{2\sigma^2}\|\dot{\mathbf{X}}+\boldsymbol{\Omega}\times\mathbf{X}\|^2\right)\mathrm{d}t \qquad (8.9)$$

$$=\int_0^T\left(\frac{1}{2}\mathbb{I}\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}-mg\boldsymbol{\chi}\cdot\mathbf{X}+\frac{1}{2\sigma^2}\|\dot{\mathbf{X}}+\boldsymbol{\Omega}\times\mathbf{X}\|^2\right)\mathrm{d}t, \qquad (8.10)$$

where m is the total mass of the body, g is the value of the gravitational acceleration, $\mathbb{I}$ is the real positive definite symmetric matrix of moments of inertia in the body, $\boldsymbol{\chi}$

is the center of mass vector in the body, and σ is a real constant. The variation with respect to $\dot{\mathbf{X}}$ defines the Legendre transform relation (costate variable)

$$\sigma^2 \mathbf{P}^\sharp := \dot{\mathbf{X}} + \mathbf{\Omega} \times \mathbf{X}. \tag{8.11}$$

The variation of the cost functional is given by

$$\delta S = \int_0^T \left[\left(\mathbb{I}\mathbf{\Omega} + \mathbf{X} \times \mathbf{P} \right) \cdot \delta\mathbf{\Omega} - \left(\dot{\mathbf{P}} + \mathbf{\Omega} \times \mathbf{P} + mg\chi \right) \cdot \delta\mathbf{X} \right] dt + \left[\mathbf{P} \cdot \delta\mathbf{X} \right]_0^T. \tag{8.12}$$

The general system (8.7) takes in this case the following *double cross* form, involving the double cross product of vectors $(\mathbf{X}, \mathbf{P}) \in \mathbb{R}^3 \times \mathbb{R}^3$ (cf. equations (2.20)),

$$\begin{cases} \dot{\mathbf{X}} - (\mathbf{X} \times \mathbf{P})^\sharp \times \mathbf{X} = \sigma^2 \mathbf{P}^\sharp, \\[2mm] \dot{\mathbf{P}} - (\mathbf{X} \times \mathbf{P})^\sharp \times \mathbf{P} = -mg\chi, \end{cases} \tag{8.13}$$

with

$$\mathbf{\Omega} = \mathbb{I}^{-1}(\mathbf{P} \times \mathbf{X}) = (\mathbf{P} \times \mathbf{X})^\sharp.$$

These equations correspond to the three equations in the general system (8.7), with the upper sign chosen. After denoting the angular momentum vector $\mathbf{\Pi} \in \mathbb{R}^3$ by

$$\mathbf{\Pi} := \mathbb{I}\mathbf{\Omega} = \mathbf{P} \times \mathbf{X}, \tag{8.14}$$

substitution of equations (8.14) into (8.13) yields

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Pi}^\sharp - mg\chi \times \mathbf{X} + \sigma^2 \mathbf{P} \times \mathbf{P}^\sharp \quad \text{and} \quad \dot{\mathbf{X}} + \mathbf{\Pi}^\sharp \times \mathbf{X} = \sigma^2 \mathbf{P}^\sharp, \tag{8.15}$$

which can be written in matrix form as

$$\begin{bmatrix} \dot{\mathbf{\Pi}} \\ \dot{\mathbf{X}} \\ \dot{\mathbf{P}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi}\times & \mathbf{X}\times & \mathbf{P}\times \\ \mathbf{X}\times & 0 & 1 \\ \mathbf{P}\times & -1 & 0 \end{bmatrix} \begin{bmatrix} \partial h_M/\partial\mathbf{\Pi} \\ \partial h_M/\partial\mathbf{X} \\ \partial h_M/\partial\mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi}\times & \mathbf{X}\times & \mathbf{P}\times \\ \mathbf{X}\times & 0 & 1 \\ \mathbf{P}\times & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{\Pi}^\sharp \\ mg\chi \\ \sigma^2\mathbf{P}^\sharp \end{bmatrix}, \tag{8.16}$$

where

$$h_M(\mathbf{\Pi}, \mathbf{X}, \mathbf{P}) = \frac{1}{2}\mathbf{\Pi} \cdot \mathbf{\Pi}^\sharp + mg\chi \cdot \mathbf{X} + \frac{\sigma^2}{2}\mathbf{P} \cdot \mathbf{P}^\sharp, \tag{8.17}$$

which suggests that one might regard the system (8.16) physically as a model of the motion of an ellipsoidal underwater vehicle, influenced by an external gravitational torque. These are equations (8.2) in this particular case.

REMARK 8.1. The analogous extremal problem for compressible fluids is given by

$$\min_{u,\rho} \int_0^T \left(\ell(u,\rho) + \frac{1}{2\sigma^2} \|\dot{\rho} + \mathrm{div}(\rho u)\|^2 \right) dt,$$

where u is the Eulerian velocity and ρ is the fluid density in spatial representation. The advection law $\dot{\rho} + \mathrm{div}(\rho u) = 0$ (exact matching) is no longer imposed. Instead its expression $\|\dot{\rho} + \mathrm{div}(\rho u)\|_{L^2}^2$ (inexact matching) is used as a penalty. Since this can be treated in a more general case, we defer this discussion to Section 8.3.2.

8.1.2. Adjoint representations. We let G act on on the right on its Lie algebra $\mathfrak{g}$ by the adjoint representation. The infinitesimal generator is thus $\xi_{\mathfrak{g}}(x) = [x,\xi]$, the diamond operator is $\diamond : \mathfrak{g}^* \times \mathfrak{g} \to \mathfrak{g}^*, p \diamond x = -\operatorname{ad}_x^* p$, and the momentum map is $\mathbf{J}(x,p) = \operatorname{ad}_x^* p$.

The Clebsch optimal control approach. The Clebsch optimal control (with condition (A), that is, $\dot{x} = [x,\xi]$) associated to a cost function $\ell = \ell(\xi,x)$ yields the (generalized) Euler-Poincaré equations

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \operatorname{ad}_\xi^* \frac{\delta \ell}{\delta \xi} + \operatorname{ad}_x^* \frac{\partial \ell}{\partial x}.$$

We suppose that $\mathfrak{g}$ is endowed with a bi-invariant inner product γ. This allows us to identify the dual Lie algebra with itself and to write $\operatorname{ad}_x^* p = -[x,p]$. In this case, the motion equations are

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \left[\frac{\delta \ell}{\delta \xi},\xi\right] + \left[\frac{\partial \ell}{\partial x},x\right].$$

These equations are obtained from the stationarity conditions

$$\frac{\delta \ell}{\delta \xi} = [p,x], \quad \dot{x} = [x,\xi], \quad \dot{p} = [p,\xi] + \frac{\partial \ell}{\partial x}.$$

If the Legendre transform associated to ℓ is a diffeomorphism, we can write $\xi = \frac{\delta h}{\delta[p,x]}$ and the equations take the form

$$\dot{x} = \left[x, \frac{\delta h}{\delta[p,x]}\right], \quad \dot{p} = \left[p, \frac{\delta h}{\delta[p,x]}\right] + \frac{\partial \ell}{\partial x}.$$

In the particular case where ℓ is given by the kinetic energy of a bi-invariant inner product, one obtains the control $\xi = [p,x]$ and the *double bracket equations*

$$\dot{x} = [x,[p,x]], \quad \dot{p} = [p,[p,x]].$$

More generally, the Lagrangian $\ell(\xi,x) = \frac{1}{2}\|\xi\|^2 - V(x)$ implies the motion equation $\dot{\xi} = \left[x,\frac{\delta V}{\delta x}\right]$; see [9]. An interesting example is provided by the potential $V(x) = -\frac{1}{2}\|[x,n]\|^2$; see [15]. For more discussion of the history, theoretical developments, and other examples of double bracket equations, see, e.g., [29].

The distributed optimization approach. The penalty functional is defined on $\mathfrak{g} \times T\mathfrak{g}$ and reads

$$\ell(\xi,x) + \frac{1}{\sigma^2}\|\dot{x} - [x,\xi]\|^2,$$

where the norm is associated to an inner product on the Lie algebra $\mathfrak{g}$. The associated equations of motion read

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \operatorname{ad}_\xi^* \frac{\delta \ell}{\delta \xi} + \operatorname{ad}_x^* \frac{\partial \ell}{\partial x} + \frac{1}{\sigma^2}\operatorname{ad}_\nu^* \nu^\flat.$$

As above, we now suppose that $\mathfrak{g}$ is endowed with a bi-invariant inner product γ and we use it to identify the dual Lie algebra $\mathfrak{g}^*$ with $\mathfrak{g}$. In this case, the above equations become

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \left[\frac{\delta \ell}{\delta \xi},\xi\right] + \left[\frac{\partial \ell}{\partial x},x\right].$$

These equations are obtained from the stationarity conditions

$$\frac{\delta\ell}{\delta\xi}=\frac{1}{\sigma^2}[\nu,x], \quad \dot{x}=[x,\xi]+\nu, \quad \dot{\nu}=[\nu,\xi]+\sigma^2\frac{\partial\ell}{\partial x}.$$

As usual, we define the variable $p:=\frac{1}{\sigma^2}\nu$ in order to rewrite these conditions as

$$\frac{\delta\ell}{\delta\xi}=[p,x], \quad \dot{x}=[x,\xi]+\sigma^2p, \quad \dot{p}=[p,\xi]+\frac{\partial\ell}{\partial x}.$$

As before, if the Legendre transform associated to ℓ is a diffeomorphism, we get

$$\dot{x}=\left[x,\frac{\delta h}{\delta[p,x]}\right]+\sigma^2p, \quad \dot{p}=\left[p,\frac{\delta h}{\delta[p,x]}\right]+\frac{\partial\ell}{\partial x}.$$

If the Lagrangian ℓ is given by the kinetic energy of a bi-invariant inner product, we get the control $\xi=[p,x]$. Now the double bracket equations are modified by an extra term:

$$\dot{x}=[x,[p,x]]+\sigma^2p, \quad \dot{p}=[p,[p,x]].$$

Further investigation of this class of equations will be pursued in future research.

8.2. Action by affine representation. We now consider the more general case where G acts on V^* by a left affine representation, $a\mapsto ga+c(g)$, where $c:G\to V^*$ is a group one-cocycle. In this case, the infinitesimal generator is

$$\xi_{V^*}(a)=\xi a+\mathbf{d}c(\xi)$$

and the cotangent bundle momentum map is

$$\mathbf{J}(a,v)=-v\diamond a+\mathbf{d}c^T(v).$$

Affine representations play an important role for a comprehensive approach to the Hamiltonian dynamics of complex fluids. We quickly give below the main equations arising in that case, in order to understand the influence of the cocycle.

The Clebsch optimal control approach for affine action. The Clebsch optimal control problem (with condition (A)$'$) yields the affine Euler-Poincaré equations

$$\frac{d}{dt}\frac{\delta\ell}{\delta\xi}=\mathrm{ad}^*_\xi\frac{\delta\ell}{\delta\xi}+\frac{\delta\ell}{\delta a}\diamond a-\mathbf{d}c^T\left(\frac{\delta\ell}{\delta a}\right).$$

These equations appear naturally in the study of spin systems and complex fluids; see [28].

The distributed optimization approach for affine action. The penalty function in the case of an affine representation is $\|\dot{a}+\xi a+\mathbf{d}c(\xi)\|^2$. The presence of the cocycle c does not modify the tensor field $\mathcal{F}^\nabla$, and one finds the motion equations

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi}=\mathrm{ad}^*_\xi\dfrac{\delta\ell}{\delta\xi}+\dfrac{\delta\ell}{\delta a}\diamond a-\mathbf{d}c^T\left(\dfrac{\delta\ell}{\delta a}\right)+\dfrac{1}{\sigma^2}\nu^\flat\diamond\nu, \\ \dfrac{d}{dt}\nu^\flat-\sigma^2\dfrac{\delta\ell}{\delta a}=-\xi\nu^\flat, \qquad \nu=\dot{a}+\xi a+\mathbf{d}c(\xi). \end{cases} \qquad (8.18)$$

As before, these equations can be obtained either by metamorphosis reduction, or by the stationarity conditions

$$\frac{\delta\ell}{\delta\xi}=\frac{1}{\sigma^2}\left(\nu^\flat\diamond a-\mathbf{d}c^T(\nu^\flat)\right),\quad \dot{a}=-\xi a-\mathbf{d}c(\xi)+\nu,\quad \dot{\nu}^\flat=-\xi\nu^\flat+\sigma^2\frac{\delta\ell}{\delta a}.$$

Defining the variable $p:=\frac{1}{\sigma^2}\nu^\flat$, we can write

$$\frac{\delta\ell}{\delta\xi}=p\diamond a-\mathbf{d}c^T(p),\qquad \dot{a}=-\xi a-\mathbf{d}c(\xi)+\sigma^2 p^\sharp,\qquad \dot{p}=-\xi p+\frac{\delta\ell}{\delta a}.$$

When the affine term is not present, one recovers (8.7). If the Lagrangian ℓ is given by the kinetic energy associated to an inner product, then the control is given by

$$\xi=(p\diamond a-\mathbf{d}c^T(p))^\sharp,$$

and we get the equations

$$\dot{a}+(p\diamond a-\mathbf{d}c^T(p))^\sharp a+\mathbf{d}c\left((p\diamond a-\mathbf{d}c^T(p))^\sharp\right)=\sigma^2 p^\sharp,$$
$$\dot{p}+(p\diamond a-\mathbf{d}c^T(p))^\sharp p=0.$$

8.3. Actions by multiplication on Lie groups.   We now specialize to the case where $Q=H$ is a Lie group, containing G as a subgroup. We will then apply the results to the rigid body and ideal fluids.

Suppose that G acts on H by left (resp. right) multiplication. Given a Lie algebra element $\xi\in\mathfrak{g}$, the infinitesimal generator is

$$\xi_H(h)=TR_h\xi=:\xi h,\quad \text{resp.}\quad \xi_H(h)=TL_h\xi=:h\xi,$$

and the cotangent bundle momentum map $\mathbf{J}:T^*H\to\mathfrak{g}^*$ is

$$\mathbf{J}(\alpha_h)=i^*(T^*R_h\alpha_h)=i^*(\alpha_h h^{-1})\quad \text{resp.}\quad \mathbf{J}(\alpha_h)=i^*(T^*L_h\alpha_h)=i^*(h^{-1}\alpha_h),$$

where $i^*:\mathfrak{h}^*\to\mathfrak{g}^*$ is the dual map to the Lie algebra inclusion $i:\mathfrak{g}\to\mathfrak{h}$.

The Clebsch optimal control approach. Given a cost function $\ell=\ell(\xi,h)$, and assuming the constraint (A)$'$, that is, $\dot{h}=-\xi h$ (resp. $\dot{h}=-h\xi$), the Clebsch optimal control problem yields the (generalized) Euler-Poincaré equations

$$\frac{d}{dt}\frac{\delta\ell}{\delta\xi}=\mathrm{ad}_\xi^*\frac{\delta\ell}{\delta\xi}-i^*\left(\frac{\partial\ell}{\partial h}h^{-1}\right),\quad \text{resp.}\quad \frac{d}{dt}\frac{\delta\ell}{\delta\xi}=-\mathrm{ad}_\xi^*\frac{\delta\ell}{\delta\xi}-i^*\left(h^{-1}\frac{\partial\ell}{\partial h}\right).$$

If the constraint (A) is assumed, that is $\dot{h}=\xi h$ (resp. $\dot{h}=h\xi$), the equations are

$$\frac{d}{dt}\frac{\delta\ell}{\delta\xi}=-\mathrm{ad}_\xi^*\frac{\delta\ell}{\delta\xi}+i^*\left(\frac{\partial\ell}{\partial h}h^{-1}\right),\quad \text{resp.}\quad \frac{d}{dt}\frac{\delta\ell}{\delta\xi}=+\mathrm{ad}_\xi^*\frac{\delta\ell}{\delta\xi}+i^*\left(h^{-1}\frac{\partial\ell}{\partial h}\right).$$

These equations are obtained by inserting the expression of the momentum map in equations (3.3) and (3.5).

The distributed optimization approach. The penalty functional is defined on $\mathfrak{g} \times TH$ and reads

$$\ell(\xi, h) + \frac{1}{2\sigma^2}\|\dot{h} \pm \xi h\|^2, \quad \text{resp.} \quad \ell(\xi, h) + \frac{1}{2\sigma^2}\|\dot{h} \pm h\xi\|^2,$$

relative to a Riemannian metric on H. We will restrict to the case of an *H-invariant metric*. More precisely, given an inner product γ on $\mathfrak{h}$, we consider the associated left (resp. right)-invariant Riemannian metric γ_h on H, that is, we have $\gamma_h(u_h, v_h) := \gamma(h^{-1}u_h, h^{-1}v_h)$, resp. $\gamma_h(u_h, v_h) := \gamma(u_h h^{-1}, v_h h^{-1})$.

Since G acts by isometries, the motion equations are given by (3.34) and (3.35). To compute these equations in our particular case, we need the concrete expression of the Levi-Civita connection associated to the Riemannian metric γ_h on H. It is written in terms of the isomorphism $\psi : \mathcal{F}(H, \mathfrak{h}) \to \mathfrak{X}(H)$ given by $\psi(f)(h) = TL_h(f(h))$ (resp. $\psi(f)(h) = TR_h(f(h))$). For a vector field $X \in \mathfrak{X}(H)$, the Levi-Civita covariant derivative associated to the left (resp. right)-invariant extension of γ to H is given by

$$\nabla_{v_h} X(h) = TL_h\left(\mathbf{d}f(v_h) - \frac{1}{2}\operatorname{ad}_v^\dagger f(h) - \frac{1}{2}\operatorname{ad}_{f(h)}^\dagger v + \frac{1}{2}[v, f(h)]\right),$$

$$\text{resp.} \quad \nabla_{v_h} X(h) = TR_h\left(\mathbf{d}f(v_h) + \frac{1}{2}\operatorname{ad}_v^\dagger f(h) + \frac{1}{2}\operatorname{ad}_{f(h)}^\dagger v - \frac{1}{2}[v, f(h)]\right),$$

where $v := h^{-1}v_h$ (resp. $v := v_h h^{-1}$), $f = \psi^{-1}(X)$, and $\operatorname{ad}_\xi^\dagger$ is the transpose of ad_ξ with respect to the inner product γ on $\mathfrak{h}$; see [50], Section 46.5.

We now specialize these formulas to the case where the vector field X is given by the infinitesimal generator ξ_H. In the case of multiplication in the left, we have $X(h) = \xi_H(h) = \xi h$ and $f(h) = \operatorname{Ad}_{h^{-1}}\xi$. Thus we obtain

$$\nabla_{v_h}\xi_H(h) = TL_h\left(-[h^{-1}v_h, f(h)] - \frac{1}{2}\operatorname{ad}_v^\dagger f(h) - \frac{1}{2}\operatorname{ad}_{f(h)}^\dagger v + \frac{1}{2}[v, f(h)]\right)$$

$$= -\frac{1}{2}TL_h\left([v, f(h)] + \operatorname{ad}_v^\dagger f(h) + \operatorname{ad}_{f(h)}^\dagger v\right).$$

For right-invariant metrics, we have $\xi_H(h) = h\xi$, $f(h) = \operatorname{Ad}_h \xi$, and the previous formula becomes

$$\nabla_{v_h}\xi_H(h) = \frac{1}{2}TR_h\left([v, f(h)] + \operatorname{ad}_v^\dagger f(h) + \operatorname{ad}_{f(h)}^\dagger v\right).$$

When condition (A)′ is assumed, the motion equations are (see (3.35))

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi} = \operatorname{ad}_\xi^* \dfrac{\delta\ell}{\delta\xi} - i^*\left(\dfrac{\partial\ell}{\partial h}h^{-1}\right), \quad \nu_h = \dot{h} + \xi h, \\[3mm] \dfrac{D}{Dt}\nu_h - \sigma^2\dfrac{\partial\ell}{\partial h}^\sharp = \dfrac{1}{2}TL_h\left([\nu, f(h)] + \operatorname{ad}_{f(h)}^\dagger \nu + \operatorname{ad}_\nu^\dagger f(h)\right) \in T_h H, \end{cases} \tag{8.19}$$

resp. for the opposite sign required for right action,

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi} = -\operatorname{ad}_\xi^* \dfrac{\delta\ell}{\delta\xi} - i^*\left(h^{-1}\dfrac{\partial\ell}{\partial h}\right), \quad \nu_h = \dot{h} + h\xi, \\[3mm] \dfrac{D}{Dt}\nu_h - \sigma^2\dfrac{\partial\ell}{\partial h}^\sharp = -\dfrac{1}{2}TR_h\left([\nu, f(h)] + \operatorname{ad}_{f(h)}^\dagger \nu + \operatorname{ad}_\nu^\dagger f(h)\right) \in T_h H, \end{cases} \tag{8.20}$$

and the stationarity condition (3.25) is

$$\frac{\delta\ell}{\delta\xi} = -\frac{1}{\sigma^2}i^*\left(\nu_h^\flat h^{-1}\right), \quad \nu_h = \dot{h} + \xi h, \quad \text{resp.} \quad \frac{\delta\ell}{\delta\xi} = -\frac{1}{\sigma^2}i^*\left(h^{-1}\nu_h^\flat\right), \quad \nu_h = \dot{h} + h\xi.$$

If the constraint (A) is assumed, then we have (see (3.34))

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi} = -\operatorname{ad}_\xi^*\dfrac{\delta\ell}{\delta\xi} + i^*\left(\dfrac{\partial\ell}{\partial h}h^{-1}\right), \quad \nu_h = \dot{h} - \xi h, \\[2ex] \dfrac{D}{Dt}\nu_h - \sigma^2\dfrac{\partial\ell}{\partial h}^\sharp = -\dfrac{1}{2}TL_h\left([\nu, f(h)] + \operatorname{ad}_{f(h)}^\dagger\nu + \operatorname{ad}_\nu^\dagger f(h)\right) \in T_hH, \end{cases} \tag{8.21}$$

resp.

$$\begin{cases} \dfrac{d}{dt}\dfrac{\delta\ell}{\delta\xi} = \operatorname{ad}_\xi^*\dfrac{\delta\ell}{\delta\xi} + i^*\left(h^{-1}\dfrac{\partial\ell}{\partial h}\right), \quad \nu_h = \dot{h} - h\xi, \\[2ex] \dfrac{D}{Dt}\nu_h - \sigma^2\dfrac{\partial\ell}{\partial h}^\sharp = \dfrac{1}{2}TR_h\left([\nu, f(h)] + \operatorname{ad}_{f(h)}^\dagger\nu + \operatorname{ad}_\nu^\dagger f(h)\right) \in T_hH, \end{cases} \tag{8.22}$$

and the stationarity condition (3.25) is

$$\frac{\delta\ell}{\delta\xi} = \frac{1}{\sigma^2}i^*\left(\nu_h^\flat h^{-1}\right), \quad \nu_h = \dot{h} - \xi h, \quad \text{resp.} \quad \frac{\delta\ell}{\delta\xi} = \frac{1}{\sigma^2}i^*\left(h^{-1}\nu_h^\flat\right), \quad \nu_h = \dot{h} - h\xi.$$

From the general theory developed in Section 5, these equations can be obtained by metamorphosis reduction, starting from the G-invariant Lagrangian $L = L\left(g, \dot{g}, f, \dot{f}\right)$: $T(G \times H) \to \mathbb{R}$ given by

$$L\left(g, \dot{g}, f, \dot{f}\right) = \mathcal{L}(g, \dot{g}, f) + \frac{1}{2\sigma^2}\|\dot{f}\|^2,$$

where $\mathcal{L}: TG \times H \to \mathbb{R}$ is the G-invariant function associated to ℓ. One can pass from the Lagrangian variables (g, f) to the reduced variables (ξ, ν_h) via the map

$$\left(g, \dot{g}, f, \dot{f}\right) \to (\xi, \nu_h) := \left(g^{-1}\dot{g}, g^{-1}\dot{f}\right),$$

for example. Note the relation $h = g^{-1}f$.

REMARK 8.2. Note that if the inner product γ on $\mathfrak{h}$ is bi-invariant, then $\operatorname{ad}^\dagger = -\operatorname{ad}$ and the equations above simplify.

8.3.1. Euler fluid equations. Hamilton's principle for ideal fluid flow might be summarized by saying that *water moves as well as possible to get out of its own way* [63]. The question pursued in [37] was whether Euler's fluid equations represent optimal control, or only optimization. As it turned out, the geodesic flow represented by the Euler's fluid equations was found to arise from either formulation. An optimization method used in image-processing (metamorphosis) is found to imply Euler's equations for incompressible flow of an inviscid fluid, without requiring that the Lagrangian particle labels exactly follow the flow lines of the Eulerian velocity vector field. That is, an optimal control formulation and an optimization formulation for incompressible ideal fluid flow both yield the *same* Euler fluid equations, although their Lagrangian parcel dynamics are *different*. This is a result of the *gauge freedom*

in the definition of the fluid pressure for an incompressible flow, in combination with the symmetry of fluid dynamics under relabeling of their Lagrangian coordinates.

We apply here the result of this section to the Lie group $H = \mathrm{Diff}(\mathcal{D})$ of all diffeomorphisms of the manifold $\mathcal{D}$ and its subgroup $G = \mathrm{Diff}_{vol}(\mathcal{D})$ of volume preserving diffeomorphisms. We shall recover and extend the approach given in [37]. Recall that a curve $\eta_t \in \mathrm{Diff}_{vol}(\mathcal{D})$ represents the Lagrangian motion of an ideal fluid in the domain $\mathcal{D}$, that is, the curve $\eta_t(x)$ in $\mathcal{D}$ is the trajectory of the fluid particle located at x at time $t = 0$, assuming that η_0 is the identity; η_t is referred to as the *forward map*. The Lie algebra of G consists of divergence free vector fields on $\mathcal{D}$ parallel to the boundary and is denoted by $\mathfrak{g} = \mathfrak{X}_{vol}(\mathcal{D})$. The curve η_t is the flow of the Eulerian velocity $u_t \in \mathfrak{X}_{vol}(\mathcal{D})$, that is, we have $\dot{\eta}_t = u_t \circ \eta_t$. The curve $l_t := \eta_t^{-1}$ is called the *back-to-labels map*. (See, e.g., [23] where the name "back-to-labels" was introduced and the map was used as a sufficient variable to describe and analyze the incompressible Euler equations.) The back-to-labels map is related to the Eulerian velocity u_t via the relation $\dot{l}_t + T l_t \cdot u_t = 0$.

As is well known, a curve $\eta_t \in \mathrm{Diff}_{vol}(\mathcal{D})$ is a geodesic with respect to the L^2 right invariant Riemannian metric if and only if u_t is a solution of the Euler fluid equations

$$\partial_t u + u \cdot \nabla u = -\mathrm{grad}\, p.$$

In other words, the Euler fluid equation is given by the Euler-Poincaré equation on $\mathfrak{X}_{vol}(\mathcal{D})$ associated to the Lagrangian $\ell(u) = \frac{1}{2} \int_{\mathcal{D}} \|u\|^2 dx$.

First approach - composition on the left: We let the group $G = \mathrm{Diff}_{vol}(\mathcal{D})$ act on $H = \mathrm{Diff}(\mathcal{D})$ by composition on the *left*. The infinitesimal generator is thus given by $u_{\mathrm{Diff}(\mathcal{D})}(\varphi) = u \circ \varphi$, for $\varphi \in \mathrm{Diff}(\mathcal{D})$.

The Clebsch optimal control approach. Using the Lagrangian $\ell(u, \varphi) = \ell(u) = \frac{1}{2} \int_{\mathcal{D}} \|u\|^2 dx = \frac{1}{2} \|u\|_{L^2}^2$ and the constraint $\dot{\varphi} = u \circ \varphi$, the Clebsch optimal control problem yields the Euler equations

$$\partial_t u + \nabla_u u = -\mathrm{grad}\, p.$$

Note that here there is no dependence of ℓ on the variable φ, therefore the Clebsch approach yields the standard Euler-Poincaré equations. The stationarity conditions are

$$u^\flat = \mathbf{J}(\varphi, \pi) = \mathbb{P}(J\varphi^{-1}(\pi \circ \varphi^{-1})), \quad \dot{\varphi} = u \circ \varphi, \quad \dot{\pi} = -(T^* u \circ \varphi) \cdot \pi, \qquad (8.23)$$

where $J\varphi$ is the Jacobian determinant of φ and $\mathbb{P}: \Omega^1(\mathcal{D}) \to \Omega^1_{div}(\mathcal{D})$ is the Hodge projector. Here $\Omega^1(\mathcal{D})$ denotes the space of one-forms on $\mathcal{D}$ and $\Omega^1_{div}(\mathcal{D}) = \{\alpha \in \Omega^1(\mathcal{D}) \,|\, \mathrm{div}(\alpha^\sharp) = 0\}$ is the space of divergence-free one-forms on $\mathcal{D}$. Here we have chosen the L^2 pairing between one-forms and vector fields on $\mathcal{D}$ and the symbols $\flat$ and $\sharp$ denote, respectively, the index lowering and rising operations defined by the Riemannian metric on $\mathcal{D}$.

The distributed optimization approach. The penalty term is $\|\dot{\varphi} - u \circ \varphi\|_{L^2}^2$, where the norm is taken with respect to the *left*-invariant L^2 metric on $\mathrm{Diff}(\mathcal{D})$, and one needs to minimize the functional

$$\int_0^T \left(\frac{1}{2} \|u\|_{L^2}^2 + \frac{1}{2\sigma^2} \|\dot{\varphi} - u \circ \varphi\|_{L^2}^2 \right) dt$$

among all curves $u(t), \varphi(t)$ in $\mathfrak{X}_{div}(\mathcal{D}) \times \mathrm{Diff}(\mathcal{D})$. The stationarity condition (3.25) reads

$$\frac{\delta \ell}{\delta u} = \frac{1}{\sigma^2} \mathbb{P}(J\varphi^{-1}(\nu_\varphi^\flat \circ \varphi^{-1})) = \frac{1}{\sigma^2} J\varphi^{-1}(\nu_\varphi^\flat \circ \varphi^{-1}) - \mathbf{d}k,$$

in which $\mathbf{d}k$ is the differential of an arbitrary scalar function, in the kernel of the projection $\mathbb{P}$ to incompressible flows. In our case, the equations of motion are given by (8.21) and we get

$$\begin{cases} \partial_t u + \nabla_u u = -\operatorname{grad} p, \quad \dot{\varphi} = u \circ \varphi + \nu_\varphi, \\ \dfrac{D}{Dt}\nu_\varphi + T\varphi \circ \nabla_{\varphi^* u}\nu = -T\varphi \circ \mathcal{F}(\varphi^* u, \nu), \end{cases} \tag{8.24}$$

where $\nu := T\varphi^{-1} \circ \nu_\varphi$ and $\mathcal{F}(v, \nu) = \frac{1}{2}(\operatorname{grad} g(v, \nu) + \nu \operatorname{div} v + v \operatorname{div} \nu)$. To see this, we compute the right hand side

$$-\frac{1}{2} T L_h \left([\nu, f(h)] + \mathrm{ad}^\dagger_{f(h)} \nu + \mathrm{ad}^\dagger_\nu f(h)\right)$$

of the second equation in (8.21). We have

$$[\nu, v] + \mathrm{ad}^\dagger_v \nu + \mathrm{ad}^\dagger_\nu v = \nabla_v \nu - \nabla_\nu v + \nabla v^T \nu + \nabla_v \nu + \nu \operatorname{div} v + \nabla \nu^T v + \nabla_\nu v + v \operatorname{div} \nu$$
$$= \operatorname{grad} g(v, \nu) + 2\nabla_v \nu + \nu \operatorname{div} v + v \operatorname{div} \nu,$$

since $\mathrm{ad}^\dagger_u m = \nabla u^T m + \nabla_u m + m \operatorname{div} u$. By choosing $\nu := T L_{\varphi^{-1}}(\nu_\varphi) = T\varphi^{-1} \circ \nu_\varphi$ and $v := \mathrm{Ad}_{\varphi^{-1}} u = T\varphi^{-1} \circ u \circ \varphi = \varphi^* u$, we obtain the result. Note that $\varphi^* u$ is an analogue of the convective velocity, but recall that φ is not the flow of u.

As usual, the stationarity conditions can also be expressed in terms of the variable $\pi := \frac{1}{\sigma^2}\nu_\varphi^\flat$. They can alternatively be written as

$$u^\flat = \mathbf{J}(\varphi, \pi) = \mathbb{P}(J\varphi^{-1}(\pi \circ \varphi^{-1})), \quad \dot{\varphi} = u \circ \varphi + \sigma^2 \pi^\sharp, \quad \dot{\pi} = -(T^* u \circ \varphi) \cdot \pi + \sigma^2 \mathcal{S}(\pi),$$

in order to be compared to (8.23), where $\mathcal{S}$ denotes the geodesic spray of the left invariant Riemannian metric. Here $\sharp := \flat^{-1}$.

The equations (8.24) can be obtained by metamorphosis reduction for the Lagrangian defined on $(u_\eta, u_f) \in T(\mathrm{Diff}_{vol}(\mathcal{D}) \times \mathrm{Diff}(\mathcal{D}))$ by

$$\frac{1}{2}\|u_\eta\|^2_{L^2} + \frac{1}{2\sigma^2}\|u_f\|^2_{L^2},$$

where the L^2 norms are associated to the right and left invariant extension of the L^2 inner product, respectively. This Lagrangian is invariant under the tangent lift of the right Diff_{vol}-action given by

$$(\eta, f) \mapsto (\eta \circ h, h^{-1} \circ f).$$

The link between the Lagrangian variables $(\eta, \dot{\eta}, f, \dot{f})$ and the reduced variables (u, ν_φ) is given by the reduction map

$$(\eta, \dot{\eta}, f, \dot{f}) \mapsto (u, \nu_\varphi) := (\dot{\eta} \circ \eta^{-1}, T\eta \circ \dot{f}).$$

In particular, we have $\varphi = \eta \circ f$.

Note that the operator D/Dt denotes the covariant derivative with respect to the left-invariant L^2 Riemannian metric on $\mathrm{Diff}_{vol}(\mathcal{D})$ and does not have a simple expression, in general, contrary to the covariant derivative associated to the right-invariant L^2 Riemannian metric, which is simply given by functorial lift. As we will see below, for the penalty approach to the Euler equations, it is more convenient to work with the back-to-labels map.

Note that instead of $H = \mathrm{Diff}(\mathcal{D})$, one can use the subgroup $H = \mathrm{Diff}_{vol}(\mathcal{D})$ of volume preserving diffeomorphisms. In this case, the second equation in (8.24) simplifies to

$$\frac{D}{Dt}\nu_\varphi + T\varphi \circ \nabla_{\varphi^* u}\nu = 0.$$

Second approach - composition on the right: We now let the group $G = \mathrm{Diff}_{vol}(\mathcal{D})$ act on $H = \mathrm{Diff}(\mathcal{D})$ by composition on the *right*. The infinitesimal generator is thus given by $u_{\mathrm{Diff}(\mathcal{D})}(l) = Tl \circ u$. Recall that the back-to-labels map l is related to the Eulerian velocity u by the formula $\dot{l} = -Tl \circ u$; therefore, we need to impose condition (A)$'$.

Clebsch optimal control approach. Using the same Lagrangian $\ell(u) = \frac{1}{2}\int_{\mathcal{D}}\|u\|^2 dx$ as before, and imposing the condition $\dot{l} = -Tl \circ u$ (condition (A)$'$), the Clebsch optimal control problem yields the Euler-Poincaré equations on $\mathfrak{X}_{vol}(\mathcal{D})$. We thus recover the Euler fluid equations

$$\partial_t u + u \cdot \nabla u = -\mathrm{grad}\,p.$$

The associated stationarity conditions are ([37])

$$u^\flat = -\mathbb{P}(\pi \circ Tl), \quad \dot{l} = -Tl \circ u, \quad \dot{\pi} = -T\pi \circ u. \tag{8.25}$$

In analogy with the symmetric representation of the rigid body ([9]), these equations are referred to as the *symmetric representation of the Euler fluid equations.*

The distributed optimization approach. The penalty term reads $\|\dot{l} + Tl \circ u\|^2_{L^2}$ where the norm is taken relative to the right-invariant L^2 metric on the group of diffeomorphisms. Therefore, we minimize the functional

$$\int_0^T \left(\frac{1}{2}\|u\|^2_{L^2} + \frac{1}{2\sigma^2}\|\dot{l} + Tl \circ u\|^2_{L^2}\right) dt$$

among all curves $u(t), l(t)$ in $\mathfrak{X}_{div}(\mathcal{D}) \times \mathrm{Diff}(\mathcal{D})$. The stationarity condition (3.25) reads

$$\frac{\delta\ell}{\delta u} = -\frac{1}{\sigma^2}\mathbb{P}\left(\nu_l^\flat \circ Tl\right) = -\frac{1}{\sigma^2}\nu_l^\flat \circ Tl - \mathbf{d}k,$$

where $\nu_l := \dot{l} + Tl \circ u$, $\mathbf{d}k$ is the differential of an arbitrary scalar function, and the associated equations of motion are

$$\begin{cases} \partial_t u + \nabla_u u = -\mathrm{grad}\,p, \\ \dfrac{D}{Dt}\nu_l + \nabla_u \nu_l = -(\mathrm{grad}\,q) \circ l. \end{cases} \tag{8.26}$$

These equations are obtained by computations similar to those above, but using (8.20) instead of (8.21). In particular, the right hand side of the second equation of (8.20) becomes

$$-(\nabla_v \nu) \circ l - (\mathrm{grad}\,q) \circ l = -\nabla_u \nu_l - (\mathrm{grad}\,q) \circ l, \quad q = \frac{1}{2}g\left(l_* u, \nu_\varphi \circ l^{-1}\right),$$

since we need to choose $\nu := \nu_l \circ l^{-1}$ and $v = \mathrm{Ad}_l u = l_* u$. Note that v is the *convective velocity* of the fluid. As usual, the stationarity conditions can also be expressed in terms of the variable $\pi := \frac{1}{\sigma^2} \nu_l^\flat$. They can alternatively be written as

$$u^\flat = -\mathbf{J}(l, \pi) = -\mathbb{P}(\pi \circ Tl), \quad \dot{l} = -Tl \circ u + \sigma^2 \pi^\sharp, \quad \dot{\pi} = -T\pi \circ u + \sigma^2 \mathcal{S}(\pi),$$

in order to be compared to (8.25), where $\mathcal{S}$ denotes the geodesic spray of the right invariant Riemannian metric.

The equations of motion (8.26) can be obtained by metamorphosis reduction of the Lagrangian defined on $(u_\eta, u_f) \in T(\mathrm{Diff}_{vol}(\mathcal{D}) \times \mathrm{Diff}(\mathcal{D}))$ by

$$\frac{1}{2} \|u_\eta\|_{L^2} + \frac{1}{2\sigma^2} \|u_f\|_{L^2}^2,$$

where the L^2 norms are associated to the right invariant extension of the L^2 inner product. This Lagrangian is invariant under the tangent lift of the right Diff_{vol}-action given by

$$(\eta, f) \mapsto (\eta \circ h, f \circ h).$$

The link between the Lagrangian variables $(\eta, \dot{\eta}, f, \dot{f})$ and the reduced variables (u, ν_l) is given by the reduction map

$$(\eta, \dot{\eta}, f, \dot{f}) \mapsto (u, \nu_l) := (\dot{\eta} \circ \eta^{-1}, \dot{f} \circ \eta^{-1}).$$

In particular, we have $l = f \circ \eta^{-1}$.

Working with $H = \mathrm{Diff}_{vol}(\mathcal{D})$, instead of the whole diffeomorphism group, yields the second equation of (8.26) in the simpler form

$$\frac{D}{Dt} \nu_l + \nabla_u \nu_l = 0.$$

These results recover Theorem 10 in [37].

8.3.2. Optimization dynamics of a compressible fluid.   For the compressible fluid, we choose to minimize the functional

$$S_p = \int\limits_0^T \left(\ell(u, \rho) + \frac{1}{2\sigma_1^2} \|\dot{l} + Tl \circ u\|_{L^2}^2 + \frac{1}{2\sigma_2^2} \|\dot{\rho} + \mathrm{div}(\rho u)\|_{L^2}^2 \right) dt$$

over all curves $u(t), l(t), \rho(t)$ in $\mathfrak{X}(\mathcal{D}) \times \mathrm{Diff}(\mathcal{D}) \times \mathrm{Den}(\mathcal{D})$, where $\mathrm{Den}(\mathcal{D})$ denotes the space of densities on $\mathcal{D}$. This minimization involves penalties and tolerances at two levels. We seek the stationarity conditions implied by optimization of the functional S_p, subject to homogeneous endpoint and boundary conditions. We introduce the notation

$$m := \frac{\delta \ell}{\delta u} \in \Omega^1(\mathcal{D}), \quad \varpi := \frac{\delta \ell}{\delta \rho} \in C^\infty(\mathcal{D}), \tag{8.27}$$

$$\pi := \frac{1}{\sigma_1^2} (\dot{l} + Tl \circ u)^\flat \in T_l^* \mathrm{Diff}(\mathcal{D}), \quad \phi := \frac{1}{\sigma_2^2} (\dot{\rho} + \mathrm{div}\,\rho u) \in C^\infty(\mathcal{D}), \tag{8.28}$$

and then write the stationarity conditions associated to the variations δu, δl, and $\delta \rho$, respectively, as:

$$m + \pi \circ Tl - \rho \mathbf{d}\phi = 0;$$

$$\dot{\pi} + \operatorname{div}(\pi u) = 0;$$
$$\dot{\phi} + \mathbf{d}\phi \circ u - \varpi = 0. \tag{8.29}$$

Combining these equations into Hamiltonian form yields (in index notation for clarity) explicitly, in terms of indices and differential operators,

$$\frac{\partial}{\partial t} \begin{bmatrix} m_i \\ \rho \\ \phi \\ l^A \\ \pi_A \end{bmatrix} = -\mathcal{B} \begin{bmatrix} \delta h_M/\delta m_j = u^j \\ \delta h_M/\delta \rho = -\varpi \\ \delta h_M/\delta \phi = \sigma_2^2 \phi \\ \delta h_M/\delta l^B = 0 \\ \delta h_M/\delta \pi_B = \sigma_1^2 \pi^{\sharp B} \end{bmatrix}, \tag{8.30}$$

where

$$\mathcal{B} = \begin{bmatrix} m_j \partial_i + \partial_j m_i & \rho \partial_i & -\phi_{,i} & -l^B_{,i} & \pi_B \partial_i \\ \partial_j \rho & 0 & -1 & 0 & 0 \\ \phi_{,j} & 1 & 0 & 0 & 0 \\ l^A_{,j} & 0 & 0 & 0 & -1 \\ \partial_j \pi_A & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{8.31}$$

Here, the summation convention is enforced on repeated indices. Upper Latin indices refer to the spatial components of the inverse map, lower Latin indices refer to the spatial reference frame, and subscript-comma notation is used for spatial derivatives. The partial derivative $\partial_j = \partial/\partial x_j$, say, acts to the right on all terms in a product by the chain rule. The Hamiltonian whose variations are taken in (8.30) is given by

$$h_M(m, \rho, \phi, l, \pi) = h(m, \rho) + \frac{\sigma_1^2}{2}\|\gamma\|^2 + \frac{\sigma_2^2}{2}\|\phi\|^2.$$

8.4. N-dimensional Camassa-Holm equation.   In this section we apply the distributed optimization method to the N-dimensional Camassa-Holm equations

$$\dot{v} + u\cdot\nabla v + \nabla u^T \cdot v + v\operatorname{div} u = 0, \quad v := (1 - \alpha^2 \Delta)u,$$

which are the spatial representation of the geodesic spray on the group $\mathrm{Diff}(\mathcal{D})$ of all diffeomorphisms of $\mathcal{D}$, relative to a Sobolev H^1 metric; see [40]. They are thus obtained by Euler-Poincaré reduction and represent a particular case of the well known EPDiff equations, to which the approach described here generalizes easily. For simplicity, we assume that $\mathcal{D}$ has no boundary.

By analogy with the Euler equations, we shall give two approaches, namely, by composition on the left and on the right. However, in the case of the Camassa-Holm equations it is convenient to slightly generalize the previous setting by letting the diffeomorphism group act on a space of embeddings. More precisely, we first consider the *left* action of $\mathrm{Diff}(\mathcal{D})$ on the space of embeddings $\mathrm{Emb}(S, \mathcal{D})$ of a manifold S into $\mathcal{D}$ and obtain the distributed optimization for the cost function

$$\int_0^T \left(\frac{1}{2}\|u\|_{H^1}^2 + \frac{1}{2\sigma^2}\|\dot{\mathbf{Q}} - u \circ \mathbf{Q}\|^2 \right) dt, \quad \mathbf{Q} \in \mathrm{Emb}(S, \mathcal{D}).$$

Then, we let $\mathrm{Diff}(\mathcal{D})$ acts on the *right* on the space of embeddings $\mathrm{Emb}(\mathcal{D}, M)$ of a manifold $\mathcal{D}$ into a manifold M and obtain the cost function

$$\int_0^T \left(\frac{1}{2}\|u\|_{H^1}^2 + \frac{1}{2\sigma^2}\|\dot{\mathbf{q}} + T\mathbf{q} \circ u\|^2 \right) dt, \quad \mathbf{q} \in \mathrm{Emb}(\mathcal{D}, M).$$

8.4.1. Left action of diffeomorphisms on embedded subspaces. Consider the left action of the configuration diffeomorphism group $G = \mathrm{Diff}(\mathcal{D})$ on $Q = \mathrm{Emb}(S,\mathcal{D})$. The infinitesimal generator associated to a Lie algebra element $u \in \mathfrak{X}(\mathcal{D})$ reads $u_{\mathrm{Emb}(S,\mathcal{D})}(\mathbf{Q}) = u \circ \mathbf{Q}$ and belongs to the tangent space $T_{\mathbf{Q}}\,\mathrm{Emb}(S,\mathcal{D})$.

The Clebsch optimal control approach. Using the Lagrangian $\ell(u,\mathbf{Q}) = \ell(u) = \frac{1}{2}\int_{\mathcal{D}} \|u\|_{H^1}^2\, dx$ and the constraint $\dot{\mathbf{Q}} = u \circ \mathbf{Q}$, the Clebsch optimal control problem yields the N-Camassa-Holm equation; see Section 4 in [29]. Note that here there is no dependence of ℓ on the variable $\mathbf{Q}$, therefore the Clebsch approach yields the standard Euler-Poincaré equations. The stationarity conditions are

$$\frac{\delta \ell}{\delta u} = \mathbf{J}(\mathbf{Q},\mathbf{P}) = \int_S \mathbf{P}(s)\delta(x - \mathbf{Q}(s))ds \in \Omega^1(\mathcal{D}), \quad \dot{\mathbf{Q}} = u \circ \mathbf{Q}, \quad \dot{\mathbf{P}} = -(T^* u \circ \mathbf{Q}) \cdot \mathbf{P}.$$

The last equation can also be written as

$$\frac{D}{Dt}\mathbf{P} = -\big((\nabla u)^T \circ \mathbf{Q}\big) \cdot \mathbf{P},$$

where D/Dt denotes the covariant derivative associated to the Riemannian metric on $\mathcal{D}$.

The distributed optimization approach. The proposed associated cost function is

$$S_p = \int_0^T \left(\ell(u) + \frac{1}{2\sigma^2}\|\dot{\mathbf{Q}} - u \circ \mathbf{Q}\|_{L^2}^2 \right) dt. \tag{8.32}$$

For definiteness, we rewrite this expression more explicitly as

$$S_p = \int_0^T \left(\ell(u) + \frac{1}{2\sigma^2}\int_S |\dot{\mathbf{Q}}(t,s) - u(t,\mathbf{Q}(t,s))|^2 ds \right) dt, \tag{8.33}$$

in which, for simplicity, $|\cdot|^2$ denotes the norm of vectors in $T\mathcal{D}$ defined by the Riemannian metric on $\mathcal{D}$ and ds denotes the volume form on S. There could also be a sum on integrations over some finite number of embedded submanifolds of various dimensions, but this possibility is unimportant in the subsequent reasoning, so it will be suppressed in the notation.

The choice of the reduced Lagrangian $\ell(u)$ will be left unspecified, except that sufficient smoothness will be assumed for the variational calculations manipulations to make mathematical sense, at least in terms of weak solutions. With these assumptions we have the following result.

THEOREM 8.2. *The extremals of S_p in* (8.33) *are given by*

$$\frac{\delta \ell}{\delta u}(x) = \int_S \mathbf{P}(t,s)\delta(x - \mathbf{Q}(t,s))ds, \quad \dot{\mathbf{Q}} = u \circ \mathbf{Q} + \sigma^2 \mathbf{P}^\sharp, \quad \frac{D}{Dt}\mathbf{P} = -\big((\nabla u)^T \circ \mathbf{Q}\big) \cdot \mathbf{P},$$
$$\tag{8.34}$$

where $\mathbf{Q} \in \mathrm{Emb}(S,\mathcal{D})$, $\mathbf{P}^\sharp \in T_{\mathbf{Q}}\,\mathrm{Emb}(S,\mathcal{D})$, *and* D/Dt *is the covariant derivative of the Levi-Civita connection on* $\mathcal{D}$.

Proof. We can obtain these conditions directly from the general equations (3.28). However, it is also instructive to derive them directly from the variational principle.

Consider the variations $\varepsilon \mapsto u_\varepsilon$ and $\varepsilon \mapsto \mathbf{Q}_\varepsilon$, and define $\mathbf{P}^\sharp$ by

$$\sigma^2 \mathbf{P}^\sharp := \dot{\mathbf{Q}} - u \circ \mathbf{Q}.$$

For $\delta u = \frac{d}{d\varepsilon}\big|_{\varepsilon=0} u_\varepsilon$ and $\delta \mathbf{Q} = \frac{d}{d\varepsilon}\big|_{\varepsilon=0}\mathbf{Q}_\varepsilon$, we have

$$\delta S_p = \int_0^T \left\langle \frac{\delta \ell}{\delta u}, \delta u \right\rangle dt$$

$$+ \int_0^T \int_S \left\langle \mathbf{P}(t,s), \frac{D}{D\varepsilon}\Big|_{\varepsilon=0} \dot{\mathbf{Q}}_\varepsilon(t,s) - \delta u(t,\mathbf{Q}(t,s)) - \frac{D}{D\varepsilon}\Big|_{\varepsilon=0} u(t,\mathbf{Q}_\varepsilon(t,s)) \right\rangle ds dt$$

$$= \int_0^T \left\langle \frac{\delta \ell}{\delta u}, \delta u \right\rangle dt + \int_0^T \int_S \left\langle \mathbf{P}(t,s), \frac{D}{Dt}\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathbf{Q}_\varepsilon(t,s) \right\rangle ds dt$$

$$- \int_0^T \int_{\mathcal{D}} \int_S \langle \mathbf{P}(s)\delta(x-\mathbf{Q}(t,s)), \delta u(t,x) \rangle ds dx dt - \int_0^T \int_S \langle \mathbf{P}(t,s), \nabla_{\delta \mathbf{Q}} u(t,s) \rangle ds dt$$

$$= \int_0^T \left\langle \frac{\delta \ell}{\delta u} - \int_S \mathbf{P}(s)\delta(x-\mathbf{Q}(t,s)) ds, \delta u \right\rangle dt - \int_0^T \left\langle \frac{D}{Dt}\mathbf{P} + \left((\nabla u)^T \circ \mathbf{Q}\right)\cdot \mathbf{P}, \delta \mathbf{Q} \right\rangle dt$$

$$+ \Big[\langle \mathbf{P}, \delta \mathbf{Q} \rangle \Big]_0^T.$$

The stationarity conditions follow immediately, upon noting that $\delta \mathbf{Q}(0,s) = 0 = \delta \mathbf{Q}(T,x)$, so that temporal endpoint terms arising under integrations by parts may be ignored. □

Suppose the reduced Lagrangian defines a velocity norm, $\ell(u) = \frac{1}{2}\|u\|^2 = \frac{1}{2}\langle u, Q_{op}(u)\rangle$. For example, let the norm be a Sobolev H^1 norm, so that it makes sense for its variational derivative in u to result in a singular distribution defined on an embedded subspace. Then, the density equation

$$\frac{\delta \ell}{\delta u}(t,x) = \int_S \mathbf{P}(t,s)\delta(x-\mathbf{Q}(t,s)) ds =: (Q_{op}u)(t,x) \tag{8.35}$$

has a natural dual solution for the velocity, given by

$$u(t,x) = \int_S \mathbf{P}^\sharp(t,s) G(x-\mathbf{Q}(t,s)) ds, \tag{8.36}$$

where G is the Green's function for the positive L^2 self-adjoint operator Q_{op}, that is,

$$Q_{op} G(x-\mathbf{Q}(t,s)) = \delta(x-\mathbf{Q}(t,s)). \tag{8.37}$$

In this situation, we have enough assumptions to obtain a coupled system of equations for the momentum densities $\mathbf{P}(t,s)$ and $m(t,x)$.

THEOREM 8.3. *The system of variational equations (8.34) for the minima of S in (8.33) implies the following dynamics for the momentum densities $\mathbf{P}(t,s)$ and $m(t,x)$,*

$$\partial_t v + \nabla_u v + \nabla u^T \cdot v + v\,\mathrm{div}(u) = -\sigma^2\,\mathrm{Div}\int_S \mathbf{P}^\sharp \otimes \mathbf{P}^\sharp(t,s)\delta(x-\mathbf{Q}(t,s)) ds, \tag{8.38}$$

$$\frac{D}{Dt}\mathbf{P} + \left((\nabla u)^T \circ \mathbf{Q}\right)\cdot \mathbf{P} = 0, \tag{8.39}$$

where Div *denotes the divergence of a contravariant two-tensor field on $\mathcal{D}$. The remaining decoupled equation*

$$\dot{\mathbf{Q}} = u\circ \mathbf{Q} + \sigma^2 \mathbf{P}^\sharp$$

allows reconstruction of the Lagrangian coordinates $\mathbf{Q}(t,s)$ on the embedded surface(s) from the dynamics of the coupled equations for the momentum densities $m(t,x)$ and $\mathbf{P}(t,s)$.

Proof.   Substitution of equations (8.34) and definitions (8.36)-(8.37) into the definition of the momentum m in equation (8.35) verifies its evolution by (8.38), upon pairing with a smooth test function and integrating appropriately by parts.

Alternatively, one can use the abstract formulation of the dynamical equations given in (3.32). As recalled before, the Euler-Poincaré part of these equations gives the N-Camassa-Holm equation

$$\dot{v} + \nabla_u v + \nabla u^T \cdot v + v \operatorname{div}(u) = 0.$$

Thus, it remains to compute the expression of the tensor $\mathcal{F}^\nabla$. Let $\mathbf{P}^\sharp \in T_\mathbf{Q} \operatorname{Emb}(S, \mathcal{D})$ and $u \in \mathfrak{X}(\mathcal{D})$, and choose $X \in \mathfrak{X}(\mathcal{D})$ such that $\mathbf{P}^\sharp(s) = X(\mathbf{Q}(s))$. Using the fact that the covariant derivative on $\operatorname{Emb}(S, \mathcal{D})$ is the functorial lift of the covariant derivative on $\mathcal{D}$, using (3.31) we get

$$\left\langle \mathcal{F}^\nabla(\mathbf{P}, \mathbf{P}^\sharp), u \right\rangle = \left\langle \mathbf{P}, \nabla_{\mathbf{P}^\sharp} u_{\operatorname{Emb}(S, \mathcal{D})}(\mathbf{Q}) \right\rangle = \int_S g\left(\mathbf{P}^\sharp(s), \nabla_{\mathbf{P}^\sharp(s)} u(\mathbf{Q}(s))\right) ds$$

$$= \int_S \int_\mathcal{D} g\left(X(x), \nabla_{X(x)} u(x)\right) \delta(x - \mathbf{Q}(s)) dx ds$$

$$= -\int_\mathcal{D} g\left(\int_S \operatorname{Div}\left(X(x) \otimes X(x) \delta(x - \mathbf{Q}(s))\right), u(x)\right) dx ds,$$

where we make use of the identity

$$\int_\mathcal{D} g(X, \nabla_Y u) dx = -\int_\mathcal{D} g(\operatorname{Div}(Y \otimes X), u) dx, \quad \text{for all} \quad X, Y, u \in \mathfrak{X}(\mathcal{D}),$$

where $\operatorname{Div}(T)^j = \nabla_i T^{ij}$, where $T = T^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$ is a contravariant two-tensor on $\mathcal{D}$. We thus obtain the formula

$$\mathcal{F}^\nabla(\mathbf{P}, \mathbf{P}^\sharp) = -\int_S \operatorname{Div}\left(X(x) \otimes X(x) \delta(x - \mathbf{Q}(s))\right) ds$$

$$= -\operatorname{Div} \int_S X(\mathbf{Q}(s)) \otimes X(\mathbf{Q}(s)) \delta(x - \mathbf{Q}(s)) ds$$

$$= -\operatorname{Div} \int_S \left(\mathbf{P}^\sharp(s) \otimes \mathbf{P}^\sharp(s) \delta(x - \mathbf{Q}(s))\right) ds$$

as required.   □

REMARK 8.3. Equations (8.38) and (8.39) represent a new dynamical system, whose exploration has only just begun and we expect will be a subject of future research.

8.4.2. Back-to-labels map for fluids.   We next present the optimal control derivation of the Camassa-Holm equation using the back-to-labels map. This means that we shall use the right action of $\operatorname{Diff}(\mathcal{D})$ on $\operatorname{Emb}(\mathcal{D}, M)$.

Recall that particles *frozen* into an ideal fluid flow are represented by time-dependent vector labels l_t whose components each satisfy the *advection law* obtained from the time derivative of the *back-to-labels map*, $l_t(x) := \eta_t^{-1}(x) = l(t, x)$, and hence the following equation holds:

$$\dot{l} + Tl \circ u = 0, \tag{8.40}$$

where u is the Eulerian velocity of the fluid.

We shall slightly generalize the back-to-labels map by considering embeddings $\mathbf{q} : \mathcal{D} \to M$, where M is a given Riemannian manifold, instead of diffeomorphisms $l : \mathcal{D} \to \mathcal{D}$.

The Clebsch approach. We recall from [29] how one can obtain the Camassa-Holm equation by Clebsch optimal control via a generalization of the back-to-labels map.

Let the group $G = \text{Diff}(\mathcal{D})$ act freely on the right on the manifold $\text{Emb}(\mathcal{D}, M)$. The associated infinitesimal generator reads $u_{\text{Emb}(\mathcal{D},M)(\mathbf{Q})} = T\mathbf{q} \circ u$. Using the Lagrangian $\ell(u) = \frac{1}{2}\|u_t\|_{H^1}^2$ and the constraint $\dot{\mathbf{q}} + T\mathbf{q} \circ u = 0$ we get the stationarity conditions

$$\frac{\delta\ell}{\delta u} = -\mathbf{p} \cdot T\mathbf{q}, \qquad \dot{\mathbf{q}} + T\mathbf{q} \circ u = 0, \qquad \dot{\mathbf{p}} + T\mathbf{p} \circ u = 0.$$

These equations produce the Camassa-Holm equation if one uses the Hamiltonian

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \iint \mathbf{p}(x) \cdot T\mathbf{q}(x) G(x - x') \mathbf{p}(x') \cdot T\mathbf{q}(x') dx dx'.$$

Distributed optimization. As opposed the Clebsch approach, we do not impose $\dot{\mathbf{q}}_t + T\mathbf{q}_t \circ u_t = 0$. Instead we use $\|\dot{\mathbf{q}}_t + T\mathbf{q}_t \circ u_t\|_{L^2}^2$ as a penalty, that is, we consider the cost functional given by

$$S_p = \int_0^T \left(\ell(u) + \frac{1}{2\sigma^2} \underbrace{\|\dot{\mathbf{q}} + T\mathbf{q} \circ u\|_{L^2}^2}_{\text{Penalty}} \right) dt. \tag{8.41}$$

Thus we need to minimize S_p subject to spatial boundary conditions on u (e.g., periodic, or vanishing at spatial infinity) and endpoint conditions (in which $\mathbf{q}(0, x)$ and $\mathbf{q}(T, x)$ are prescribed), and also subject to the penalty for the error in the L^2 norm,

$$\|\dot{\mathbf{q}} + T\mathbf{q} \circ u\|_{L^2}^2 = \int_{\mathcal{D}} |\dot{\mathbf{q}}(x) + T\mathbf{q}(u(x))|^2 dx, \tag{8.42}$$

in which, for simplicity, $|\cdot|^2$ denotes the norm of vectors in TM defined by a Riemannian metric on M. It is important to note that the L^2 Riemannian metric used in the penalty is *not* invariant under the right $\text{Diff}(\mathcal{D})$-action on itself.

REMARK 8.4 (An alternative penalty term). If $M = \mathcal{D}$, the quantity

$$v := -\dot{l} \circ l^{-1} = Tl \circ u \circ l^{-1} = l_* u = \text{Ad}_l u$$

is called the *convective velocity* [41] of the fluid. This is analogous to the relation $\Omega = \text{Ad}_{O^{-1}}\omega$ for $O \in SO(3)$ satisfied by body angular velocity Ω and spatial angular velocity ω for rigid body motion in $\mathbb{R}^3$, both viewed as elements of $\mathfrak{so}(3)$. Penalizing in (8.41) for $\|v - \text{Ad}_l u\|_{L^2}^2$ is an interesting alternative approach, which will be presented, in general, in Section 8.5.1.

Let $\sigma^2 > 0$ and choose the reduced Lagrangian to be a norm $\ell(u) = \frac{1}{2}\|u\|^2$. Then, when extremals of (8.41) exist, they will be minima.

Later we shall specialize the reduced Lagrangian to the norm $\ell(u) = \frac{1}{2}\|u\|_{H^1}^2$. For the moment, however, we leave the choice arbitrary, only assuming that sufficient smoothness is present for all functions to exist locally and be differentiable in space and time. With these assumptions we have the following result.

THEOREM 8.4. *The extremals of S_p in (8.41) are given by*

$$\frac{\delta\ell}{\delta u} + \pi \circ T\mathbf{q} = 0, \quad \frac{D\pi}{Dt} + \text{Div}(\pi u) = 0, \quad \dot{\mathbf{q}} + T\mathbf{q} \circ u =: \sigma^2 \pi^\sharp, \tag{8.43}$$

where the expression $\mathrm{Div}(\pi u) \in T_{\mathbf{q}}^* \mathrm{Emb}(\mathcal{D}, M)$ *is defined by*

$$\mathrm{Div}(\pi u) := (\mathrm{div}\, u)\pi + \nabla_u \pi, \quad with \quad \nabla_{u_x}\pi := \frac{D}{D\varepsilon}\bigg|_{\varepsilon=0} \pi(c(\varepsilon))$$

for $u_x \in T_x\mathcal{D}$, $\varepsilon \mapsto c(\varepsilon)$ *a curve such that* $\frac{d}{d\varepsilon}\big|_{\varepsilon=0} c(\varepsilon) = u_x$, *and* $\frac{D}{D\varepsilon}$ *is the covariant derivative of the Levi-Civita connection of the Riemannian metric on* M.

Note that if $M = \mathbb{R}^n$ endowed with the Riemannian metric given by the dot product then $\mathrm{Div}(\pi u)_i = \mathrm{div}(\pi_i u)$.

Proof. Define $\pi \in T_{\mathbf{q}}\mathrm{Emb}(\mathcal{D}, M)$ by $\sigma^2 \pi := \dot{\mathbf{q}} + T\mathbf{q} \circ u$. For variations $\varepsilon \mapsto u_\varepsilon$ and $\varepsilon \mapsto \mathbf{q}_\varepsilon$, we compute

$$\delta S_p = \int_0^T \left\langle \frac{\delta\ell}{\delta u}, \delta u \right\rangle dt + \int_0^T \left\langle \pi, \frac{D}{D\varepsilon}\bigg|_{\varepsilon=0} (\dot{\mathbf{q}}_\varepsilon + T\mathbf{q}_\varepsilon \circ u_\varepsilon) \right\rangle$$

$$= \int_0^T \left\langle \frac{\delta\ell}{\delta u}, \delta u \right\rangle dt + \int_0^T \left\langle \pi, \frac{D}{Dt}\delta\mathbf{q} + T\mathbf{q} \circ \delta u + \nabla_u \delta\mathbf{q} \right\rangle$$

$$= \int_0^T \left\langle \frac{\delta\ell}{\delta u} + \pi \circ T\mathbf{q}, \delta u \right\rangle - \int_0^T \left\langle \frac{D}{Dt}\pi + \mathrm{Div}(\pi u), \delta\mathbf{q} \right\rangle,$$

where in the last equality, we used integration by parts and the definition of Div. $\square$

THEOREM 8.5. *The system of variational equations* (8.43) *for the minima of* S_p *yields the following dynamical system for the momentum* π *and momentum 1-form* $v^\flat := \mathrm{d}\ell/\delta u = -\pi \circ T\mathbf{q}$:

$$\partial_t v + \nabla_u v + \nabla u^T \cdot v + v\,\mathrm{div}(u) = \sigma^2 (\nabla\pi)^T \cdot \pi^\sharp, \tag{8.44}$$

$$\partial_t \pi + \mathrm{Div}(\pi u) = 0, \tag{8.45}$$

where $\mathrm{Div}(\pi u)$ *is defined above. The decoupled equation* $\sigma^2 \pi^\sharp = \dot{\mathbf{q}} + T\mathbf{q} \circ u$ *allows reconstruction of the labels* $\mathbf{q}$ *from the dynamics of the coupled equations for* v *and* π.

Proof. One can directly obtain these equations from the stationarity condition given in (8.43). We shall however use the abstract formulation (3.33) and compute the tensor field $\mathcal{F}^\nabla$ defined in (3.31). Given $\pi \in T_{\mathbf{q}}^* \mathrm{Emb}(\mathcal{D}, M)$, $u \in \mathfrak{X}(\mathcal{D})$, and a curve $\varepsilon \mapsto \mathbf{q}_\varepsilon \in \mathrm{Emb}(\mathcal{D}, M)$ such that $\frac{d}{d\varepsilon}\big|_{\varepsilon=0} \mathbf{q}_\varepsilon = \pi^\sharp$, we have

$$\left\langle \mathcal{F}^\nabla(\pi, \pi^\sharp), u \right\rangle = \left\langle \pi, \nabla_{\pi^\sharp} u_{\mathrm{Emb}(\mathcal{D},M)}(\mathbf{q}) \right\rangle = \int_{\mathcal{D}} g\left(\pi^\sharp(x), \frac{D}{D\varepsilon}\bigg|_{\varepsilon=0} T\mathbf{q}_\varepsilon(u(x)) \right) dx$$

$$= \int_{\mathcal{D}} g\left(\pi^\sharp(x), \nabla_u \pi(x) \right) dx = \left\langle (\nabla\pi^T) \cdot \pi^\sharp, u \right\rangle,$$

which proves (8.44). Equation (8.45) is part of the system (8.43). $\square$

REMARK 8.5 (Two-component Camassa-Holm equation). If $\mathcal{D} = \mathbb{R}$, $M = \mathbb{R}$, and we assume appropriate decay properties at infinity such that all boundary terms appearing in integration by parts vanish, specializing the reduced Lagrangian to

$$\ell(u) = \frac{1}{2}\|u\|_{H^1}^2 = \frac{1}{2}\int_{\mathbb{R}} (u^2 + \alpha^2 u_x^2)\, dx,$$

with homogeneous boundary conditions on the infinite real line or on a periodic spatial interval yields the variational derivative $(\mathrm{d}\ell/\delta u)^\sharp = v = u - \alpha^2 u_{xx}$ for a length scale α. In this case, equations (8.44), (8.45) recover the *two-component Camassa-Holm equations*,

$$\partial_t v + (uv)_x + vu_x = \sigma^2 \pi \pi_x, \tag{8.46}$$
$$\partial_t \pi + (u\pi)_x = 0. \tag{8.47}$$

This system forms a completely integrable Hamiltonian system with soliton solutions associated to an isospectral linear eigenvalue problem, so it may be solved analytically by using the inverse scattering transform method [22, 51]. These equations are also known to be the spatial representation of geodesics on the semidirect product $\mathrm{Diff}(\mathbb{R})\circledS\mathcal{F}(\mathbb{R})$; see [46], [31].

8.5. Metamorphosis dynamics. Consider a Lie group G acting on the *left* on a manifold N. The Lie group G is the group of deformations and the manifold N contains what are called "deformable objects". In imaging applications we take G to be the group of diffeomorphisms of N.

DEFINITION 8.6. *A **metamorphosis** ([68, 47]) is a pair of curves $(g_t, \eta_t) \in G \times N$ parameterized by time t, with $g_0 = \mathrm{id}$. Its **image** is the curve $n_t \in N$ defined by the action $n_t = g_t \eta_t$ denoted by concatenation from the left. The quantities g_t and η_t are called, respectively, the **deformation part** of the metamorphosis, and its **template part**. When η_t is constant, the metamorphosis is said to be a **pure deformation**. In the general case, the image is a combination of a deformation and template variation.*

A metamorphosis may be determined as an optimal curve (g_t, η_t), with $g_t \in G$ and $\eta_t \in N$, with respect to a metric that is invariant under the **right action** of G on $G \times N$ defined by

$$(g, \eta)h = (gh, h^{-1}\eta) \tag{8.48}$$

for any $g, h \in G$ and $\eta \in N$. More specifically, a metamorphosis (g, η) may be obtained by seeking a stationary point $\mathrm{d}S = 0$ of a right-invariant cost function S on $T(G \times N)$. This general situation has been considered in detail in the first sections of the paper.

The present conventions are those of equations (3.32) with the upper sign chosen. Recall in particular that we start from a right G-invariant Lagrangian of the form

$$L(g, \dot{g}, \eta, \dot{\eta}) = \mathcal{L}(g, \dot{g}, \eta) + \frac{1}{2\sigma^2}\|g\dot{\eta}\|^2,$$

where the norm involved in the penalty is associated to a Riemannian metric g on N. The corresponding reduced Lagrangians on $\mathfrak{g} \times TN$ read

$$\ell_{EP}(u, \nu_n) = \ell(u, n) + \frac{1}{2\sigma^2}\|\nu_n\|^2, \quad \ell_M(u, n, \dot{n}) = \ell(u, n) + \frac{1}{2\sigma^2}\|\dot{n} - u_N(n)\|^2,$$

where the reduced variables are

$$u = \dot{g}g^{-1} \in \mathfrak{g}, \quad n = g\eta \in N, \quad \nu_n = g\dot{\eta} \in T_n N,$$

with $g \in G$, $\eta \in N$.

We shall discuss the case in which G and N are Lie groups and the action of G on N is a Lie group homomorphism for all $g \in G$, that is,

$$g(n\tilde{n}) = (gn)(g\tilde{n}), \quad \text{for all} \quad n, \tilde{n} \in N.$$

For example, N can be a vector space and the action of G can be linear. In the case of action by homomorphisms we can form the semidirect product Lie group $G \circledS N$ with product defined by

$$(g, n)(\tilde{g}, \tilde{n}) = (g\tilde{g}, (g\tilde{n})n). \tag{8.49}$$

Define on $T(G \circledS N)$ a right-invariant metric whose value at the identity is denoted $\|(\cdot, \cdot)\|_{(\mathrm{id}_G, \mathrm{id}_N)}$. A geodesic for this metric that optimizes the kinetic energy in $T(G \circledS N)$ between $(g_0 = \mathrm{id}_G, n_0)$ and (g_1, n_1) with fixed images n_0 and n_1 and free deformation g_1 yields a particular case of metamorphosis. A right-invariant cost function on $T(G \circledS N)$ may be expressed in terms of the time-dependent quantities

$$u = \dot{g}g^{-1}, \quad n = g\eta, \quad \text{and} \quad \nu = g\dot{\eta}, \tag{8.50}$$

that are each invariant under the right action defined in equation (8.48). For example, a right-invariant cost function on $T(G \circledS N)$ may be expressed as

$$S = \int L(g, \dot{g}, \eta, \dot{\eta})dt = \int \ell(u, n, \nu)dt. \tag{8.51}$$

Right invariance of a metric on $T(G \circledS N)$ implies

$$\|(U_g, U_n)\|_{(g,n)} = \|(U_g\tilde{g}, (U_g\tilde{n})n + (g\tilde{n})U_n\|_{(g\tilde{g}, (g\tilde{n})n)}, \tag{8.52}$$

which, upon choosing $(\tilde{g}, \tilde{n}) = (g^{-1}, g^{-1}n^{-1})$ and denoting $u := U_g g^{-1}$, $\zeta := n^{-1}U_n$, yields

$$\|(U_g, U_n)\|_{(g,n)} = \|(u, (un^{-1})n + \zeta\|_{(\mathrm{id}_G, \mathrm{id}_N)}$$

$$= \|(u, \zeta - \mathrm{Ad}_{n^{-1}}u)\|_{(\mathrm{id}_G, \mathrm{id}_N)}.$$

The last line follows from $0 = u(n^{-1}n) = (un^{-1})n + n^{-1}(un)$. In this notation, the cost function given by the geodesic energy on $T(G \circledS N)$ for a path of unit length is given by

$$\frac{1}{2}\int_0^1 \|(u, \zeta - \mathrm{Ad}_{n^{-1}}u)\|_{(\mathrm{id}_G, \mathrm{id}_N)}^2 dt. \tag{8.53}$$

The definition of the variable n in (8.50) implies $\dot{n} = \dot{g}\eta + g\dot{\eta} = un + \nu$. Upon writing $\zeta = n^{-1}\dot{n}$ as the left-invariant image velocity, one sees the variable transformation

$$\zeta - \mathrm{Ad}_{n^{-1}}u = n^{-1}\nu.$$

Consequently, optimizing the geodesic energy (8.53) with fixed n_0 and n_1 is equivalent to solving the metamorphosis problem formulated in [47] as a stationary principle $\delta S = 0$ with $S = \int l(u, n, \nu)dt$ and

$$l(u, n, \nu) = \|(u, n^{-1}\nu)\|_{(\mathrm{id}_G, \mathrm{id}_N)}^2. \tag{8.54}$$

We shall investigate the additively separated form of the metric,

$$\ell(u, n, \nu) = l(u) + \frac{1}{2\sigma^2}\|n^{-1}\nu\|_{(\mathrm{id}_G, \mathrm{id}_N)}^2$$

$$= l(u) + \frac{1}{2\sigma^2}\|\zeta - \mathrm{Ad}_{n^{-1}}u\|_{(\mathrm{id}_G, \mathrm{id}_N)}^2.$$

8.5.1. Subgroup actions.   We shall discuss in this paragraph the particular case in which N is also a Lie group that contains G as subgroup and on which G acts by multiplication on the left. We also assume that the Riemannian metric g on N is left invariant (relative to left translations by elements of N). In this case, one can make use of left trivialization of the tangent bundle TN to get the diffeomorphism

$$\mathfrak{g} \times TN \to \mathfrak{g} \times N \times \mathfrak{n}, \quad (u,n,\dot{n}) \mapsto (u,n,n^{-1}\dot{n}) =: (u,n,\zeta).$$

The reduced Lagrangian in terms of the new variables is denoted ℓ_L and reads

$$\ell_L(u,n,\zeta) = \ell(u,n) + \frac{1}{2\sigma^2}\|\zeta - \mathrm{Ad}_{n^{-1}} u\|^2$$

since we have the relations

$$n^{-1}\nu_n = n^{-1}(\dot{n} - u_N(n)) = n^{-1}(\dot{n} - un) = \zeta - \mathrm{Ad}_{n^{-1}} u.$$

We now rewrite the stationarity conditions relative to these new variables. Consider variations $\varepsilon \mapsto u_\varepsilon$ and $\varepsilon \mapsto n_\varepsilon$ of the curves u and n. We have as usual

$$\delta\zeta = \dot{\Sigma} + [\zeta, \Sigma], \tag{8.55}$$

where $\Sigma = n^{-1}\delta n$. Likewise,

$$\delta(\mathrm{Ad}_{n^{-1}} u) = \mathrm{Ad}_{n^{-1}}\left(\delta u + [u, \delta n n^{-1}]\right) = \mathrm{Ad}_{n^{-1}}\left(\delta u + [u, \mathrm{Ad}_n \Sigma]\right)$$
$$= \mathrm{Ad}_{n^{-1}}\delta u + [\mathrm{Ad}_{n^{-1}} u, \Sigma].$$

For simplicity, we suppose that ℓ does not depend on n. Substituting these relations into the variation of the action integral we get

$$dS_p = d\int_0^T \ell_L(u,n,\zeta)dt = \delta\int_0^T\left(\ell(u) + \frac{1}{2\sigma^2}\|\zeta - \mathrm{Ad}_{n^{-1}} u\|^2\right)dt$$

$$= \int_0^T\left(\left\langle\frac{\delta\ell}{du}, \delta u\right\rangle + \langle\pi, \delta\zeta - \delta(\mathrm{Ad}_{n^{-1}} u)\rangle\right)dt$$

$$= \int_0^T\left(\left\langle\frac{\delta\ell}{\delta u}, \delta u\right\rangle + \left\langle\pi, \dot{\Sigma} + \mathrm{ad}_\zeta\Sigma - \mathrm{Ad}_{n^{-1}}\delta u - \mathrm{ad}_{(\mathrm{Ad}_{n^{-1}} u)}\Sigma\right\rangle\right)dt$$

$$= \int_0^T\left(\left\langle\frac{\delta\ell}{\delta u} - \mathrm{Ad}^*_{n^{-1}}\pi, \delta u\right\rangle - \left\langle\dot{\pi} - \mathrm{ad}^*_\zeta\pi + \mathrm{ad}^*_{(\mathrm{Ad}_{n^{-1}} u)}\pi, \Sigma\right\rangle\right)dt + \left[\langle\pi,\Sigma\rangle\right]_0^T,$$

where $\pi \in \mathfrak{n}^*$ is the image momentum dual to the left-invariant image velocity $\zeta \in \mathfrak{n}$, that is,

$$\pi := \frac{\delta\ell_L}{\delta\zeta} = \frac{1}{\sigma^2}\left(\zeta - \mathrm{Ad}_{n^{-1}} u\right)^\flat = \frac{1}{\sigma^2}\left(n^{-1}\nu_n^\flat\right) = \frac{1}{\sigma^2}n^{-1}\left(\nu_n^\flat\right) = \in\mathfrak{n}^*.$$

Stationarity $dS = 0$ and $\Sigma(0) = \Sigma(T) = 0$ then imply

$$\frac{\delta\ell}{\delta u} = \mathrm{Ad}^*_{n^{-1}}\pi \quad\text{and}\quad \dot{\pi} = \mathrm{ad}^*_\zeta\pi - \mathrm{ad}^*_{(\mathrm{Ad}_{n^{-1}} u)}\pi = \mathrm{ad}^*_{\sigma^2\pi^\sharp}\pi = \sigma^2\mathrm{ad}^*_{\pi^\sharp}\pi. \tag{8.56}$$

From the general theory, since the G-action on N is by isometries it follows that $\mathcal{F}^\nabla = 0$, and thus these equations imply the Euler-Poincaré equations. It is also

instructive to obtain them directly. Taking the time derivative and using general results relating the Ad^* and ad^* operations yields

$$
\frac{d}{dt}\frac{\delta\ell}{\delta u} = \frac{d}{dt}\left(\mathrm{Ad}^*_{n^{-1}}\pi\right)
$$
$$
= \mathrm{Ad}^*_{n^{-1}}\left(\dot{\pi} - \mathrm{ad}^*_{\zeta}\pi\right)\quad\left(\text{with } \zeta = n^{-1}\dot{n}\right)
$$
$$
\text{by (8.56b)} = -\,\mathrm{Ad}^*_{n^{-1}}\mathrm{ad}^*_{\left(\mathrm{Ad}_{n^{-1}}u\right)}\pi
$$
$$
= -\,\mathrm{ad}^*_u\left(\mathrm{Ad}^*_{n^{-1}}\pi\right)
$$
$$
\text{by (8.56a)} = -\,\mathrm{ad}^*_u\frac{\delta\ell}{\delta u}.
$$

In turn, using $u = \dot{g}g^{-1}$ and $n = g\eta$, from the Euler-Poincaré equation we get the conservation law,

$$
0 = \mathrm{Ad}^*_g\left(\frac{d}{dt}\frac{\delta\ell}{\delta u} + \mathrm{ad}^*_u\frac{\delta\ell}{\delta u}\right) = \frac{d}{dt}\left(\mathrm{Ad}^*_g\frac{\delta\ell}{\delta u}\right) = \frac{d}{dt}\left(\mathrm{Ad}^*_g\mathrm{Ad}^*_{n^{-1}}\pi\right)
$$
$$
= \frac{d}{dt}\left(\mathrm{Ad}^*_{\eta^{-1}}\pi\right) = \mathrm{Ad}^*_{\eta^{-1}}\left(\dot{\pi} - \mathrm{ad}^*_\upsilon\pi\right),
$$

where $\upsilon := \eta^{-1}\dot{\eta}$ is the left-invariant template velocity.

REMARK 8.6 (Interpretation of the equations).

1. The conservation laws for $\mathrm{Ad}^*_g(\delta\ell/\delta u)$ and $\mathrm{Ad}^*_{\eta^{-1}}\pi$ provide the interpretations of the momentum dynamics. Namely, the momentum $\delta\ell/\delta u$ (resp. π) undergoes coadjoint motion with respect to g (resp. η^{-1}).

2. The peculiar form of the momentum equation (8.56b) is then understood, because the template velocity υ is proportional to image momentum π by a factor of the penalty constant, which also maps it from the dual of the Lie algebra, back to Lie algebra, namely,

$$
\upsilon := \eta^{-1}\dot{\eta} = n^{-1}\nu_n = \sigma^2\pi^\sharp.
$$

Perhaps not unexpectedly, when $\sigma^2 \to 0$ the template velocity vanishes and the remaining image motion reduces to a pure deformation governed by the Euler-Poincaré equation.

3. The metamorphosis (g_t, η_t) is determined as an initial value problem, as follows. Given the Lagrangian $\ell(u)$, the Euler-Poincaré equation

$$
\frac{d}{dt}\frac{\delta\ell}{\delta u} + \mathrm{ad}^*_u\frac{\delta\ell}{\delta u} = 0,
$$

determines the velocity $u = \dot{g}g^{-1}$ which then yields g_t by reconstruction from solving $\dot{g}_t = u_t g_t$. Next, the relations

$$
\dot{\pi} = \mathrm{ad}^*_{\left(\zeta - \mathrm{Ad}_{n^{-1}}u\right)}\pi \quad\text{and}\quad \sigma^2\pi^\sharp = \zeta - \mathrm{Ad}_{n^{-1}}u,
$$

with $\zeta = n^{-1}\dot{n}$ and $\dot{n} = un + \nu_n$, need to be negotiated to obtain the image curve n_t. Finally, the template curve is obtained from $\eta_t = g_t^{-1}n_t$. This process is worth discussing in an example.

8.5.2. Example: Metamorphosis equations on $SE(2)$**.**   In $SE(2)$ the manifold of "deformable objects" $N = \mathbb{R}^2$ is acted upon by the Lie group of "deformations" $G = SO(2)$ on the left. The situation simplifies in this case because N is a vector space and we recover the setting described in Section 8.1. Hence,

$$\ell_M(u,n,\nu) = \ell(u) + \frac{1}{2\sigma^2}\|\nu\|^2 = \ell(u) + \frac{1}{2\sigma^2}\|\dot{n} - un\|^2 = \ell_{LP}(u,n,\dot{n}),$$

and the cost function becomes

$$S_p = \int_0^T \left(\ell(u) + \frac{1}{2\sigma^2}\|\nu\|^2 \right) dt = \int_0^T \left(\ell(u) + \frac{1}{2\sigma^2}\|\dot{n} - un\|^2 \right) dt,$$

where the $\mathfrak{se}(2)$ Lie algebra action un may be written on $\mathbb{R}^2$ as a cross product of vectors [36]

$$un = u\hat{z} \times n.$$

Consequently, the SDP metamorphosis equations (see (8.7))

$$\frac{\mathrm{d}\ell}{\mathrm{d}u} + \pi \diamond n = 0,$$
$$\dot{\pi} - u\pi = 0,$$
$$\dot{n} - un = \sigma^2 \pi^\sharp = \nu,$$

which may be written in vector form as

$$\frac{\mathrm{d}\ell}{\mathrm{d}u}\hat{z} + \pi \times n = 0,$$
$$\dot{\pi} - u\hat{z} \times \pi = 0,$$
$$\dot{n} - u\hat{z} \times n = \sigma^2 \pi = \nu.$$

A few statements may be made about the qualitative properties of the solutions of this system.

1. We first observe that $|\pi|$ is constant because by the second equation above, we have $\frac{d}{dt}|\pi|^2 = 2\pi \cdot \dot{\pi} = 2\pi \cdot (u\hat{z} \times \pi) = 0$. So π executes circular motion in the plane at constant rotation frequency $\pi \times \dot{\pi}/|\pi|^2 = u\hat{z}$.

2. Substituting the second and third equations into the time derivative of the first one yields the conservation law

$$\frac{d}{dt}\frac{\delta\ell}{\delta u} = 0$$

   for the planar motion. In particular, we obtain the constant of motion $\pi \times n = $ const.

3. The other two equations are closed provided one may solve $\delta\ell/\delta u$ for u, which of course we shall assume is possible. More precisely, we now assume that the Legendre transformation $u \mapsto \delta\ell/\delta u$ is a diffeomorphism. In this case, since $\delta\ell/\delta u$ is constant, u is also constant.

4. It remains to determine the effects of $\sigma^2 \neq 0$ on the dynamics of n. A short computation shows that

$$\frac{d}{dt}(\pi \cdot n) = \sigma^2|\pi|^2 \quad \text{and} \quad \frac{d}{dt}|n|^2 = 2\sigma^2(\pi \cdot n),$$

so, since $|\pi|^2 =$ const, $|n|^2(t)$ increases quadratically with scaled time $\sigma^2 t$ and the motion may be visualized as taking place in $\mathbb{R}^3$ with coordinates $(x_1, x_2, x_3) = (|\pi|^2, |n|^2, \pi \cdot n)$ along the parabolas formed by intersections of level sets of the two integrals of motion $|\pi|^2 =$ constant and $|\pi \times n|^2 =$ const. The rotation frequency of n is found to be

$$\frac{n \times \dot{n}}{|n|^2} = \hat{z}\left(u + \frac{\sigma^2}{|n|^2}\frac{\mathrm{d}l}{\mathrm{d}u}\right).$$

As $\sigma^2 t \to \infty$, the directions of the vectors π and n tend toward a state of alignment, rotating together at frequency $u\hat{z}$. In contrast, for $\sigma^2 = 0$, the vectors π and n keep their magnitudes and rotate together at frequency $u\hat{z}$ with constant relative orientation.

8.5.3. Lie-Poisson Hamiltonian formulation of metamorphosis for right action.

In this example we particularize the system of motion equations (4.8) to the case of a representation but without imposing the endpoint condition at $t = 1$. The resulting equations are obtained by metamorphosis reduction from an arbitrary Lagrangian $L : T(G \times V) \to \mathbb{R}$, where V is a vector space. Thus, the equations below are more general that those obtained in the penalty approach.

As explained in Section 5, the variational problem optimizes over metamorphoses (g_t, η_t) by minimizing $S = \int_0^1 L\,dt$, for a Lagrangian L of the form

$$L(g_t, \dot{g}_t, \eta_t, \dot{\eta}_t) = \mathcal{L}(g_t, \dot{g}_t, \eta_t) + \frac{1}{2\sigma^2}\|g_t\dot{\eta}_t\|^2,$$

with fixed boundary conditions for the initial and final images n_0 and n_1, with image $n_t = g_t\eta_t$ for template η_t and $g_0 = \mathrm{id}_G$; thus only the images are constrained at the endpoints.

For the concrete metamorphosis example, the group G of diffeomorphisms $\mathrm{Diff}(\mathcal{D}) \ni g$ of the domain $\mathcal{D}$ is taken to act on the space of smooth maps (images) $V = \mathcal{F}(\mathcal{D}) \ni \eta$ by the left action $g\eta := \eta \circ g^{-1}$ of G on V. Therefore, the right action (8.48) of G on $G \times V$ is given in this case by $(g, \eta)h := (g \circ h, \eta \circ h)$ for $g, h \in \mathrm{Diff}(\mathcal{D})$ and $\eta \in \mathcal{F}(\mathcal{D})$. The reduced Lagrangians $\ell_{LP}(u_t, n_t, \dot{n}_t)$ and $\ell_M(u_t, n_t, \nu_t)$ are defined on the space $\mathfrak{g} \times V \times V$. In imaging applications, $u_t = \dot{g}_t g_t^{-1}$ is the velocity along the optimal path g_t sought between two images; $n_t := g_t\eta_t$ is the path in the image space; and $\nu_t := g_t\dot{\eta}_t$ is the image velocity.

From a visual point of view, image metamorphoses are similar to what is usually called "morphing" in computer graphics. The evolution of the image over time, $t \mapsto n_t$, is a combination of deformations and image intensity variation. Algorithms and experimental results for the solution of the boundary value problem (minimize the time-integrated Lagrangian between two images) can be found in [61, 26].

From the general metamorphosis equations (4.8) (with the minus sign corresponding to the right action of G on $G \times V$) we obtain the dynamical system

$$\begin{cases} \dfrac{\partial}{\partial t}\dfrac{\delta \ell_M}{\delta u} + \mathrm{ad}_{u_t}^*\dfrac{\delta \ell_M}{\delta u} + \dfrac{\delta \ell_M}{\delta n} \diamond n_t + \dfrac{\delta \ell_M}{\delta \nu} \diamond \nu_t = 0, \\[2ex] \dfrac{\partial}{\partial t}\dfrac{\delta \ell_M}{\delta \nu} - u_t\dfrac{\delta \ell_M}{\delta \nu} - \dfrac{\delta \ell_M}{\delta n} = 0, \\[2ex] \dot{n}_t = \nu_t + u_t n_t, \\[2ex] \dfrac{\delta \ell_M}{\delta u}(1) + \dfrac{\delta \ell_M}{\delta \nu}(1) \diamond n_1 = 0, \end{cases} \qquad (8.57)$$

where for $n \in V$, $a \in V^*$, and $u \in \mathfrak{X}(\mathcal{D}) = \mathfrak{g}$, the infinitesimal actions and the diamond operators are given by

$$
\begin{aligned}
un &= -\mathbf{d}n \cdot u \in V = \mathcal{F}(\mathcal{D}), \\
ua &= \mathrm{div}(au) \in V^* = \mathcal{F}(\mathcal{D})^* \cong \mathcal{F}(\mathcal{D}), \\
n \diamond a &= -a\mathbf{d}n \in \mathfrak{g}^* = \Omega^1(\mathcal{D}).
\end{aligned}
$$

Even though we fixed the standard volume form on $\mathcal{D} \subset \mathbb{R}^n$, so densities on $\mathcal{D}$ are identified with functions and one-form densities with one-forms, we recall that one should think of ua as a density and $n \diamond a$ as a one-form density.

In contrast to earlier sections, fixed endpoints at $t = 1$ are not assumed in metamorphosis. This difference leads to the last equation in the system (8.57). For details of the derivation of the system (8.57) and discussions of the regularity of its solutions, see [47].

System (8.57) describes coadjoint motion

$$
\frac{\partial}{\partial t}\left(\frac{\delta \ell_M}{\delta u} + \frac{\delta \ell_M}{\delta \nu} \diamond n \right) + \mathrm{ad}^*_{u_t}\left(\frac{\delta \ell_M}{\delta u} + \frac{\delta \ell_M}{\delta \nu} \diamond n \right) = 0, \tag{8.58}
$$

or, equivalently,

$$
\frac{\partial}{\partial t}\left(\mathrm{Ad}^*_{g_t}\left(\frac{\delta \ell_M}{\delta u} + \frac{\delta \ell_M}{\delta \nu} \diamond n \right) \right) = 0, \tag{8.59}
$$

so that

$$
\left. \left(\frac{\delta \ell_M}{\delta u} + \frac{\delta \ell_M}{\delta \nu} \diamond n \right) \right|_t = \mathrm{Ad}^*_{g_t^{-1}}\left. \left(\frac{\delta \ell_M}{\delta u} + \frac{\delta \ell_M}{\delta \nu} \diamond n \right) \right|_{t=0}, \tag{8.60}
$$

for the coadjoint action of the Lie group G on the dual of its Lie algebra $\mathfrak{g}$.

Hamiltonian formulation. One passes from the Euler-Poincaré metamorphosis equations on the Lagrangian side to their Lie–Poisson Hamiltonian formulation via the ***Legendre transformation***; see the presentation and general formulas at the end of Section 6. The Legendre transformation of the reduced Lagrangian $\ell_M(u, n, \nu)$: $\mathfrak{g} \times V \times V \to \mathbb{R}$ in its variables u and ν defines the Hamiltonian,

$$
h(\mu, n, \beta) = \langle \mu, u \rangle + \langle \beta, \nu \rangle - \ell_M(u, n, \nu), \tag{8.61}
$$

on $\mathfrak{g}^* \times V \times V^*$, where

$$
\mu = \frac{\delta \ell_M}{\delta u} \quad \text{and} \quad \beta = \frac{\delta \ell_M}{\delta \nu} \tag{8.62}
$$

are given by the Legendre transformation. The variational derivatives of the Hamiltonian h are

$$
\frac{\delta h}{\delta \mu} = u, \quad \frac{\delta h}{\delta \beta} = \nu, \quad \frac{\delta h}{\delta n} = -\frac{\delta \ell_M}{\delta n}. \tag{8.63}
$$

Consequently, the Euler-Poincaré equations (8.57) for metamorphosis in the Eulerian description imply the following equations for the Legendre-transformed variables (μ, n, β), written symbolically as a matrix operation:

$$
\frac{\partial}{\partial t}\begin{bmatrix} \mu \\ n \\ \beta \end{bmatrix} = -\begin{bmatrix} \mathrm{ad}^*_\square \mu & -\square \diamond n & \beta \diamond \square \\ -\square n & 0 & -1 \\ -\square \beta & 1 & 0 \end{bmatrix}\begin{bmatrix} \delta h/\delta \mu \\ \delta h/\delta n \\ \delta h/\delta \beta \end{bmatrix} =: \mathcal{B}\begin{bmatrix} \delta h/\delta \mu \\ \delta h/\delta n \\ \delta h/\delta \beta \end{bmatrix}, \tag{8.64}
$$

with boxes $\square$ indicating where the substitutions occur. These equations can also be obtained from the system (6.10) (with minus sign chosen in $\mp$) by explicitly computing every term for this situation. The Poisson bracket defined by the L^2 skew-symmetric Hamiltonian matrix $\mathcal{B}$ is given by

$$\{f,h\}(\mu,n,\beta) = \int \begin{bmatrix} \delta f/\delta\mu \\ \delta f/\delta n \\ \delta f/\delta\beta \end{bmatrix}^T \mathcal{B} \begin{bmatrix} \delta h/\delta\mu \\ \delta h/\delta n \\ \delta h/\delta\beta \end{bmatrix} \mathrm{d}x. \qquad (8.65)$$

The pair (n,β) satisfies canonical Poisson-bracket relations. The other parts of the Poisson bracket are linear in the variables (μ,n,β). This linearity is the signature of the Lie-Poisson bracket on the dual of the semidirect product Lie algebra of vector fields $\mathfrak{X}(\mathcal{D})$ acting on functions $\mathcal{F}(\mathcal{D},W)$ and its dual $\mathcal{F}(\mathcal{D},W^*)$ with a canonical cocycle between them. The semidirect product Lie algebra bracket on $\mathfrak{g} \times V \times V$ is

$$[(u,n,\nu),(\bar{u},\bar{n},\bar{\nu})] = ([u,\bar{u}], u\bar{n} - \bar{u}n, u\bar{\nu} - \bar{u}\nu).$$

A similar Lie-Poisson bracket was found for complex fluids in [35]. Ongoing work in this direction includes a Lagrange-Poincaré formulation of these equations ([30]).

9. Conclusions and outlook

This paper has begun the development of the family of dynamical systems associated with optimal control and optimization problems. The theory was developed in the context of many examples inspired by control theory and optimization, particularly in the new area of applications in imaging analysis of the theory of *metamorphosis*, a means of optimally tracking the changes of shape necessary for registration of images of various types, or data structures, without requiring that the transformations of shape be diffeomorphisms. The main idea was to soften the exact dynamical constraint by replacing it with a quadratic penalty term. The resulting *optimization dynamics* was studied by using methods that originated in geometric mechanics. In particular, Lagrange-Poincaré reduction and its associated variational formulations were adapted to this sort of *optimal inexact reduction*. This approach allowed us to obtain the equations of metamorphosis dynamics that are naturally generated by the stationarity conditions, then study their properties from both the Lagrangian and Hamiltonian points of view.

This geometric setup for optimization dynamics was illustrated in diverse examples in Section 8. Besides metamorphosis (Section 8.5), these examples included optimally reduced versions of the heavy top (Section 8.1.1), the double bracket equations (Section 8.1.2), the Euler equations for an inviscid ideal fluid both incompressible and compressible (Section 8.3.1), and the N-dimensional Camassa-Holm equation (Section 8.4). For the one-dimensional Camassa-Holm equation the optimal reduction process produced its integrable Hamiltonian extension, the two-component Camassa-Holm equations in (8.46) and (8.47).

We plan to continue the investigation of the relationships among problems in imaging, optimal control, and geometric mechanics. In particular, we plan to continue developing the dynamical systems framework for designing and interpreting methods of large deformation matching for image registration in computational anatomy.

REFERENCES

[1] V.I. Arnold, *Sur la géométrie différentielle des groupes de Lie de dimenson infinie et ses applications à l'hydrodynamique des fluides parfaits*, Ann. Inst. Fourier, Grenoble, 16, 319–361, 1966.

[2] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, second ed. Springer, 1989.

[3] J. Baillieul and J.C. Willems, *Mathematical Control Theory*, Springer, New York, 1999.

[4] M. Barbero-Liñan, *A Geometric Study of Abnormality in Optimal Control Problems for Control and Mechanical Control Systems*, PhD thesis, Technical University of Catalonia, 2008.

[5] M.F. Beg, *Variational and Computational Methods for Flows of Diffeomorphisms in Image Matching and Growth in Computational Anatomy*, Ph.D. Thesis, Johns Hopkins University, 2003.

[6] M.F. Beg, M.I. Miller, A. Trouvé, and L. Younes, *Computing large deformation metric mappings via geodesic flows of diffeomorphisms*, Int. J. Comput. Vis., 61(2), 139–157, 2005.

[7] A.M. Bloch, J. Baillieul, P.E.Crouch, and J. E. Marsden, *Nonholonomic Mechanics and Control*, Springer, New York, 2004.

[8] A.M. Bloch, R.W. Brockett, and P.E. Crouch, *Double bracket equations and geodesic flows on symmetric spaces*, Commun. Math. Phys., 187, 357–373, 1997.

[9] A.M. Bloch and P.E. Crouch, *Optimal control and geodesic flows*, Syst. & Control Lett., 28, 65–76, 1996.

[10] A.M. Bloch, P.E. Crouch, J.E. Marsden, and T.S. Ratiu, *Discrete rigid body dynamics and optimal control*, Proc. IEEE Conf. on Decision and Control, 37, 2249–2254, 1998.

[11] A.M. Bloch, P.E. Crouch, D.D. Holm, and J.E. Marsden, *An optimal control formulation for inviscid incompressible ideal fluid flow*, Proc. CDC IEEE, 39, 1273–1279, 2000.

[12] A.M. Bloch, P.E. Crouch, J.E. Marsden, and A. K Sanyal, *Optimal control and geodesics on quadratic matrix Lie groups*, Found. Comput. Math., 8, 469–500, 2008.

[13] A.M. Bloch, P.E. Crouch, and A. K Sanyal, *A variational problem on Stiefel manifolds*, Nonlin., 19(10), 2247–2276, 2006.

[14] R. Brockett, *Lie theory and control systems defined on spheres*, SIAM J. Appl. Math., 25(2), 213–225, 1973.

[15] R. Brockett, *The double bracket equation as the solution of a variational problem. Hamiltonian and gradient flows, algorithms and control*, Fields Inst. Commun. Amer. Math. Soc., Providence, RI., 3, 69–76, 1994.

[16] M. Bruveris, F. Gay-Balmaz, D.D. Holm, and T.S. Ratiu, *The momentum map representation of images*, J. Nonlin. Sci., 21, 115–150, 2011.

[17] R. Camassa and D.D. Holm, *An integrable shallow water equation with peaked solitons*, Phys. Rev. Lett., 71, 1661–1664, 1993.

[18] H. Cendra, J.E. Marsden, S. Pekarsky, and T.S. Ratiu, *Variational principles for Lie-Poisson and Hamilton-Poincaré equations*, Mosc. Math. J., 3(3), 833–867, 2003.

[19] H. Cendra, J.E. Marsden, and T.S. Ratiu, *Lagrangian Reduction by Stages*, Mem. Amer. Math. Soc., 152(722), 2001.

[20] S. Chen, C. Foias, D.D. Holm, E.J. Olson, E.S. Titi, and S. Wynne, *The Camassa-Holm equations as a closure model for turbulent channel and pipe flows*, Phys. Rev. Lett., 81, 5338–5341, 1998.

[21] S. Chen, C. Foias, D.D. Holm, E.J. Olson, E.S. Titi, and S. Wynne, *A connection between the Camassa-Holm equations and turbulence in pipes and channels*, Phys. Fluids, 11, 2343–2353, 1999.

[22] M. Chen, S. Liu, and Y. Zhang, *A two-component generalization of the Camassa-Holm equa-*

tion and its solutions, Lett. Math. Phys., 75, 1–15, 2006.

[23] P. Constantin, *An Eulerian-Lagrangian approach for incompressible fluids: Local theory*, J. Amer. Math. Soc., 14, 263–278, 2001.

[24] C.J. Cotter, and D.D. Holm, *Geodesic boundary value problems with symmetry*, J. Geom. Mech., 2(1), 417–444, 2010.

[25] P. Dupuis, U. Grenander, and M.I. Miller, *Variational problems on flows of diffeomorphisms for image matching*, Quart. Appl. Math., 56, 587–600, 1998.

[26] L. Garcin and L. Younes, *Geodesic image matching: A wavelet based energy minimization scheme*, in Proceedings of EMMCVPR 2005, Lecture Notes in Computer Science, 3757, 349–364, 2005.

[27] F. Gay-Balmaz and T.S. Ratiu, *Reduced Lagrangian and Hamiltonian formulations of Euler-Yang-Mills fluids*, J. Symplectic Geom., 6, 189–237, 2008.

[28] F. Gay-Balmaz and T.S. Ratiu, *The geometric structure of complex fluids*, Adv. Appl. Math., 42(2), 176–275, 2008.

[29] F. Gay-Balmaz and T.S. Ratiu, *Clebsch optimal control formulation in mechanics*, J. Geom. Mech., 3(1), 47–79, 2011.

[30] F. Gay-Balmaz and C. Tronci, *Reduction theory for symmetry breaking*, Phys. D, 239(20-22), 1929–1947, 2010.

[31] F. Gay-Balmaz, C. Tronci, and C. Vizman, *Geometric dynamics on the automorphism group of principal bundles: Geodesic flows, dual pairs and chromomorphism groups*, available at `http://arxiv.org/pdf/1006.0650v2`, preprint.

[32] U. Grenander, *General Pattern Theory*, Oxford Science Publications, Oxford University Press, New York, 1993.

[33] R. Hermann and C.F. Martin, *Lie theory of transformation groups and the parameterization and identification of linear systems. Geometry and identification*, (Weston, Mass., 1981), Lie Groups: History, Frontiers and Applications. Ser. B: Systems Inform. Control, Math. Sci. Press, Brookline, MA, 1, 101–121, 1983.

[34] D.D. Holm, *Hamiltonian dynamics of a charged fluid, including electro- and magnetohydro-dynamics*, Phys. Lett. A, 114, 137–141, 1986.

[35] D.D. Holm, *Euler-Poincaré dynamics of perfect complex fluids*, in Geometry, Dynamics and Mechanics: 60th Birthday Volume for J.E. Marsden., P. Holmes, P. Newton, and A. Weinstein (eds.), Springer-Verlag, 113–167, 2002.

[36] D.D. Holm, *Geometric Mechanics II: Rotating, Translating and Rolling*, World Scientific: Imperial College Press, Singapore, 2008.

[37] D.D. Holm, *Euler's fluid equations: Optimal control vs optimization*, Phys. Lett. A, 373, 4354–4359, 2009.

[38] D.D. Holm and R.I. Ivanov, *Two-component CH system: Inverse scattering, peakons and geometry*, Inverse Problems, 27, 045013, 2011.

[39] D.D. Holm and B.A. Kupershmidt, *Poisson brackets and Clebsch representations for magne-tohydrodynamics, multifluid plasmas and elasticity*, Physica D, 6, 347–363, 1983.

[40] D.D. Holm and J.E. Marsden, *Momentum maps and measure-valued solutions (peakons, filaments and sheets) for the EPDiff equation*, in The Breadth of Symplectic and Poisson Geometry, A Festschrift for Alan Weinstein, Progr. Math., J.E. Marsden and T.S. Ratiu, (eds.), Birkhäuser Boston, Boston, MA, 232, 203–235, 2004.

[41] D.D. Holm, J.E. Marsden, and T.S. Ratiu, *Hamiltonian Structure and Lyapunov Stability for Ideal Continuum Dynamics*, University of Montreal Press, Montreal, 1986.

[42] D.D. Holm, J.E. Marsden, and T.S. Ratiu, *The Euler-Poincaré equations and semidirect products with applications to continuum theories*, Adv. in Math., 137, 1–81, 1998.

[43] D.D. Holm, J.E. Marsden, and T.S. Ratiu, *Euler-Poincaré models of ideal fluids with nonlinear dispersion*, Phys. Rev. Lett., 349, 4173–4177, 1998.

[44] D.D. Holm, J.T. Ratnanather, A. Trouvé, and L. Younes, *Soliton dynamics in computational anatomy*, NeuroImage, 23, S170–S178, 2004.

[45] D.D. Holm, T. Schmah, and C. Stoica, *Geometric Mechanics and Symmetry: From Finite to Innite Dimensions*, Oxford University Press, 2009.

[46] D.D. Holm and C. Tronci, *Geodesic flows on semidirect-product Lie groups: Geometry of singular measure-valued solutions*, Proc. R. Soc. A, 465, 457–476, 2008.

[47] D.D. Holm, A. Trouvé, and L. Younes, *The Euler-Poincaré theory of metamorphosis*, Quart. Appl. Math., 67, 661–685, 2009.

[48] S. Joshi and M.I. Miller, *Landmark matching via large deformation diffeomorphisms*, IEEE Trans. Image Proc., 9(8), 1357–1370, 2000.

[49] V. Jurdjevic, *Geometric Control Theory*, Cambridge University Press, 1997.

[50] A Kriegl, and P. Michor, *The Convenient Setting of Global Analysis*, Surveys and Monographs,

American Mathematical Society, 53, 1997.

[51] P.A. Kuz'min, *Two-component generalizations of the Camassa-Holm equation*, Mathematical Notes, 81, 130–134, 2007. (P.A. Kuz'min, Matematicheskie Zametki,in Russian) 81, 149–152, 2007.

[52] C. de Lellis, T. Kappeler, and P. Topalov, *Low-regularity solutions of the periodic Camassa-Holm equation*, Commun. Part. Diff. Eqs., 32, 87–126, 2007.

[53] J.E. Marsden, G. Misiołek, J.P. Ortega, M. Perlmutter, and T.S. Ratiu, *Hamiltonian Reduction by Stages*, Springer Lecture Notes in Mathematics, Springer-Verlag, 1913, 2007.

[54] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry*, Second Edition, Springer, 1999.

[55] J.E. Marsden, T.S. Ratiu, and A. Weinstein, *Semidirect product and reduction in mechanics*, Trans. Amer. Math. Soc., 281, 147–177, 1984.

[56] J.E. Marsden and A. Weinstein, *Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids*, Phys. D, 7, 305–323, 1983.

[57] R.I. McLachlan and S. Marsland, *N-particle dynamics of the Euler equations for planar diffeomorphisms*, Dyn. Syst., 22(3), 269–290, 2007.

[58] M. Micheli, *Effects of curvature on the analysis of landmark shape manifolds*, Proc. of the 2008 IEEE International Conference on Image Processing (ICIP 2008) San Diego, California, October, 12-15, 2008.

[59] P.W. Michor, *Topics in Differential Geometry*, Graduate Studies in Mathematics, American Mathematical Society, Providence, 93, 2008.

[60] M.I. Miller, A. Trouvé, and L. Younes, *On the metrics and Euler-Lagrange equations of computational anatomy*, Ann. Rev. Biomed. Eng., 4, 375–405, 2002.

[61] M.I. Miller and L. Younes, *Group action, diffeomorphism and matching: A general framework*, Int. J. Comput. Vis., 41, 61–84, 2001.

[62] G. Misiołek, *Classical solutions of the periodic Camassa–Holm equation*, Geom. Funct. Anal., 12, 1080–1104, 2002.

[63] J. Monaghan, Private communication.

[64] D. Mumford and A. Desolneux, *The Stochastic Analysis of Real-Worlds Signals*, A.K. Peters Ltd., 2010.

[65] D.W. Thompson, *On Growth and Form*, Dover reprint of 1942 2nd ed. (1st ed., 1917). ISBN 0-486-67135-6.

[66] A. Trouvé, *An Infinite Dimensional Group Approach for Physics Based Model*, Technical report, 1995, electronically available at http://www.cis.jhu.edu.

[67] A. Trouvé, *Diffeomorphic groups and pattern matching in image analysis*, Int. J. Computer Vision, 28, 213–221, 1998.

[68] A. Trouvé and L. Younes, *Metamorphoses through Lie group action*, Found. Comput. Math., 5, 173–198, 2005.

[69] F.-X. Vialard, *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, 2009.

[70] L. Younes, *Computable elastic distances between shapes*, SIAM J. Appl. Math., 58, 565–86, 1998.

[71] L. Younes, *Shapes and Diffeomorphisms*, Springer Applied Mathematical Sciences, Berlin, 2010.

[72] L. Younes, F. Arrate, and M.I. Miller, *Evolutions equations in computational anatomy*, NeuroImage, 45, S40–S50, 2009.