Reduction theory and the Lagrange–Routh equations

Jerrold E. Marsden[a)]
Control and Dynamical Systems 107-81, California Institute of Technology, Pasadena, California 91125

Tudor S. Ratiu[b)]
Département de Mathématiques, École Polytéchnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Jürgen Scheurle[c)]
Zentrum Mathematik, TU München, Arcisstrasse 21, D-80290 München, Germany

(Received 8 November 1999; accepted for publication 10 March 2000)

Reduction theory for mechanical systems with symmetry has its roots in the classical works in mechanics of Euler, Jacobi, Lagrange, Hamilton, Routh, Poincaré, and others. The modern vision of mechanics includes, besides the traditional mechanics of particles and rigid bodies, field theories such as electromagnetism, fluid mechanics, plasma physics, solid mechanics as well as quantum mechanics, and relativistic theories, including gravity. Symmetries in these theories vary from obvious translational and rotational symmetries to less obvious particle relabeling symmetries in fluids and plasmas, to subtle symmetries underlying integrable systems. Reduction theory concerns the removal of symmetries and their associated conservation laws. Variational principles, along with symplectic and Poisson geometry, provide fundamental tools for this endeavor. Reduction theory has been extremely useful in a wide variety of areas, from a deeper understanding of many physical theories, including new variational and Poisson structures, to stability theory, integrable systems, as well as geometric phases. This paper surveys progress in selected topics in reduction theory, especially those of the last few decades as well as presenting new results on non-Abelian Routh reduction. We develop the geometry of the associated Lagrange–Routh equations in some detail. The paper puts the new results in the general context of reduction theory and discusses some future directions.
[S0022-2488(00)02006-5]

CONTENTS

[a)]Electronic mail: marsden@cds.caltech.edu
[b)]Electronic mail: Tudor.Ratiu@epfl.ch
[c)]Electronic mail: scheurle@mathematik.tu-muenchen.de

I. INTRODUCTION

This section surveys some of the literature and basic results in reduction theory. We will come back to many of these topics in ensuing sections.

A. Overview

A brief history of reduction theory. We begin with an overview of progress in reduction theory and some new results in Lagrangian reduction theory. Reduction theory, which has its origins in the classical work of Euler, Lagrange, Hamilton, Jacobi, Routh, and Poincaré, is one of the fundamental tools in the study of mechanical systems with symmetry. At the time of this classical work, traditional variational principles and Poisson brackets were fairly well understood. In addition, several classical cases of reduction (using conservation laws and/or symmetry to create smaller dimensional phase spaces), such as the elimination of cyclic variables as well as Jacobi's elimination of the node in the n-body problem, were developed. The ways in which reduction theory has been generalized and applied since that time has been rather impressive. General references in this area are Abraham and Marsden [1978],[2] Arnold [1989],[12] and Marsden [1992].[103]

Of the above-mentioned classical works, Routh [1860,1884][140,142] pioneered reduction for Abelian groups. Lie [1890][88] discovered many of the basic structures in symplectic and Poisson geometry and their link with symmetry. Meanwhile, Poincaré [1901][133] discovered the generalization of the Euler equations for rigid body mechanics and fluids to general Lie algebras. This was more or less known to Lagrange [1788][80] for SO(3), as we shall explain later in the paper. The modern era of reduction theory began with the fundamental papers of Arnold [1966][8–10] and Smale [1970].[147] Arnold focused on systems on Lie algebras and their duals, as in the works of Lie and Poincaré, while Smale focused on the Abelian case giving, in effect, a modern version of Routh reduction.

With hindsight we now know that the description of many physical systems such as rigid bodies and fluids requires *noncanonical Poisson brackets* and *constrained variational principles* of the sort studied by Lie and Poincaré. An example of a noncanonical Poisson bracket on $\mathfrak{g}^*$, the dual of a Lie algebra $\mathfrak{g}$, is called, following Marsden and Weinstein [1983],[112] the *Lie–Poisson bracket*. These structures were known to Lie around 1890, although Lie seemingly did not

recognize their importance in mechanics. The symplectic leaves in these structures, namely the coadjoint orbit symplectic structures, although implicit in Lie's work, were discovered by Kirillov, Kostant, and Souriau in the 1960s.

To synthesize the Lie algebra reduction methods of Arnold [1966][8] with the techniques of Smale [1970][147] on the reduction of cotangent bundles by Abelian groups, Marsden and Weinstein [1974][110] developed reduction theory in the general context of symplectic manifolds and equivariant momentum maps; related results, but with a different motivation and construction (not stressing equivariant momentum maps) were found by Meyer [1973].[114]

The construction is now standard: Let (P, Ω) be a symplectic manifold and let a Lie group G act freely and properly on P by symplectic maps. The free and proper assumption is to avoid singularities in the reduction procedure, as is discussed later. Assume that this action has an equivariant momentum map $\mathbf{J}:P\to\mathfrak{g}^*$. Then the *symplectic reduced space* $\mathbf{J}^{-1}(\mu)/G_\mu=P_\mu$ is a symplectic manifold in a natural way; the induced symplectic form Ω_μ is determined uniquely by $\pi_\mu^*\Omega_\mu=i_\mu^*\Omega$, where $\pi_\mu:\mathbf{J}^{-1}(\mu)\to P_\mu$ is the projection and $i_\mu:\mathbf{J}^{-1}(\mu)\to P$ is the inclusion. If the momentum map is not equivariant, Souriau [1970][148] discovered how to centrally extend the group (or algebra) to make it equivariant.

Coadjoint orbits were shown to be symplectic reduced spaces by Marsden and Weinstein [1974].[110] In the reduction construction, if one chooses $P=T^*G$, with G acting by (say left) translation, the corresponding space P_μ is identified with the coadjoint orbit $\mathcal{O}_\mu$ through μ together with its coadjoint orbit symplectic structure. Likewise, the Lie–Poisson bracket on $\mathfrak{g}^*$ is inherited from the canonical Poisson structure on T^*G by *Poisson reduction*, that is, by simply identifying $\mathfrak{g}^*$ with the quotient $(T^*G)/G$. It is not clear who first *explicitly* observed this, but it is *implicit* in many works such as Lie [1890],[88] Kirillov [1962,1976],[69,70] Guillemin and Sternberg [1980],[44] and Marsden and Weinstein [1982, 1983],[111,112] but is *explicit* in Marsden *et al.* [1983],[113] and in Holmes and Marsden [1983].[51]

Kazhdan, Kostant, and Sternberg [1978][66] showed that P_μ is symplectically diffeomorphic to an orbit reduced space $P_\mu\cong J^{-1}(\mathcal{O}_\mu)/G$ and from this it follows that P_μ are the symplectic leaves in P/G. This paper was also one of the first to notice deep links between reduction and integrable systems, a subject continued by, for example, Bobenko, Reyman, and Semenov-Tian-Shansky [1989][20] in their spectacular group theoretic explanation of the integrability of the Kowalewski top.

The way in which the *Poisson* structure on P_μ is related to that on P/G was clarified in a generalization of *Poisson reduction* due to Marsden and Ratiu [1986],[97] a technique that has also proven useful in integrable systems (see, e.g., Pedroni [1995][132] and Vanhaecke [1996][151]).

Reduction theory for mechanical systems with symmetry has proven to be a powerful tool enabling advances in stability theory (from the Arnold method to the energy–momentum method) as well as in bifurcation theory of mechanical systems, geometric phases via reconstruction—the inverse of reduction—as well as uses in control theory from stabilization results to a deeper understanding of locomotion. For a general introduction to some of these ideas and for further references, see Marsden and Ratiu [1994].[98]

More about Lagrangian reduction. Routh reduction for Lagrangian systems is classically associated with systems having cyclic variables (this is almost synonymous with having an Abelian symmetry group); modern accounts can be found in Arnold [1988][11] and in Marsden and Ratiu [1994],[98] Sec. 8.9. A key feature of Routh reduction is that when one drops the Euler–Lagrange equations to the quotient space associated with the symmetry, and when the momentum map is constrained to a specified value (i.e., when the cyclic variables and their velocities are eliminated using the given value of the momentum), then the resulting equations are in Euler–Lagrange form not with respect to the Lagrangian itself, but with respect to the *Routhian*. In his classical work, Routh [1877][141] applied these ideas to stability theory, a precursor to the energy–momentum method for stability (Simo, Lewis, and Marsden [1991];[145] see Marsden [1992][103] for an exposition and references). Of course, Routh's stability method is still widely used in mechanics.[31]

Another key ingredient in Lagrangian reduction is the classical work of Poincaré [1901][133] in

which the *Euler–Poincaré equations* were introduced. Poincaré realized that both the equations of fluid mechanics and the rigid body and heavy top equations could all be described in Lie algebraic terms in a beautiful way. The importance of these equations was realized by Hamel [1904,1949][47,48] and Chetayev [1941].[30]

Tangent and cotangent bundle reduction. The simplest case of cotangent bundle reduction is reduction at zero in which case one chooses $P=T^*Q$ and then the reduced space at $\mu=0$ is given by $P_0=T^*(Q/G)$, the latter with the canonical symplectic form. Another basic case is when G is Abelian. Here, $(T^*Q)_\mu \cong T^*(Q/G)$ but the latter has a symplectic structure modified by magnetic terms; that is, by the curvature of the mechanical connection.

The Abelian version of cotangent bundle reduction was developed by Smale [1970][147] and Satzer [1977][143] and was generalized to the non-Abelian case in Abraham and Marsden [1978].[2] Kummer [1981][77] introduced the interpretations of these results in terms of a connection, now called the *mechanical connection*. The geometry of this situation was used to great effect in, for example, Guichardet [1984],[40] Iwai [1987c, 1990],[61,62] and Montgomery [1984, 1990, 1991].[116,120,121] Routh reduction may be viewed as the Lagrangian analog of cotangent bundle reduction.

Tangent and cotangent bundle reduction evolved into what we now term as the ''bundle picture'' or the ''gauge theory of mechanics.'' This picture was first developed by Montgomery, Marsden, and Ratiu [1984][115] and Montgomery [1984, 1986].[116,118] That work was motivated and influenced by the work of Sternberg [1977][149] and Weinstein [1978][155] on a *Yang–Mills construction* that is, in turn, motivated by Wong's equations, that is, the equations for a particle moving in a Yang–Mills field. The main result of the bundle picture gives a structure to the quotient spaces $(T^*Q)/G$ and $(TQ)/G$ when G acts by the cotangent and tangent lifted actions. We shall review this structure in some detail in the following.

Nonabelian Routh reduction. Marsden and Scheurle [1993a, 1993b][107,108] showed how to generalize the Routh theory to the non-Abelian case as well as realizing how to get the Euler–Poincaré equations for matrix groups by the important technique of *reducing variational principles*. This approach was motivated by related earlier work of Cendra and Marsden [1987][27] and Cendra, Ibort, and Marsden [1987].[26] The work of Bloch *et al.* [1996][17] generalized the Euler–Poincaré variational structure to general Lie groups and Cendra, Marsden, and Ratiu [2000][28] carried out a Lagrangian reduction theory that extends the Euler–Poincaré case to arbitrary configuration manifolds. This work was in the context of the Lagrangian analog of Poisson reduction in the sense that no momentum map constraint is imposed.

One of the things that makes the Lagrangian side of the reduction story interesting is the lack of a general category that is the Lagrangian analog of Poisson manifolds. Such a category, that of *Lagrange–Poincaré bundles*, is developed in Cendra, Marsden, and Ratiu [2000],[28] with the tangent bundle of a configuration manifold and a Lie algebra as its most basic example. That work also develops the Lagrangian analog of reduction for central extensions and, as in the case of symplectic reduction by stages (see Marsden *et al.* [1998, 2000][90,91]), cocycles and curvatures enter in this context in a natural way.

The Lagrangian analog of the bundle picture is the bundle $(TQ)/G$, which, as shown later, is a vector bundle over Q/G; this bundle was studied in Cendra, Marsden, and Ratiu [2000].[28] In particular, the equations and variational principles are developed on this space. For $Q=G$ this reduces to Euler–Poincaré reduction and for G Abelian, it reduces to the classical Routh procedure. Given a G-invariant Lagrangian L on TQ, it induces a Lagrangian l on $(TQ)/G$. The resulting equations inherited on this space, given explicitly later, are the *Lagrange–Poincaré equations* (or the *reduced Euler–Lagrange equations*).

Methods of Lagrangian reduction have proven very useful in, for example, optimal control problems. It was used in Koon and Marsden [1997][71] to extend the falling cat theorem of Montgomery [1990][120] to the case of nonholonomic systems as well as nonzero values of the momentum map.

Semidirect product reduction. Recall that in the simplest case of a semidirect product, one has a Lie group G that acts on a vector space V (and hence on its dual V^*) and then one forms the

semidirect product $S=G\circledS V$, generalizing the semidirect product structure of the Euclidean group $\mathrm{SE}(3)=\mathrm{SO}(3)\circledS\mathbb{R}^3$.

Consider the isotropy group G_{a_0} for some $a_0\in V^*$. The *semidirect product reduction theorem* states that each of the *symplectic reduced spaces for the action of* G_{a_0} *on* T^*G *is symplectically diffeomorphic to a coadjoint orbit in* $(\mathfrak{g}\circledS V)^*$, *the dual of the Lie algebra of the semidirect product*. This semidirect product theory was developed by Guillemin and Sternberg [1978, 1980],[43,44] Ratiu [1980, 1981, 1982],[135–139] and Marsden, Ratiu, and Weinstein [1984a, 1984b].[100,101]

This construction is used in applications where one has ''advected quantities'' (such as the direction of gravity in the heavy top, density in compressible flow and the magnetic field in magnetohydrodynamics). Its Lagrangian counterpart was developed in Holm, Marsden, and Ratiu [1998b][54] along with applications to continuum mechanics. Cendra *et al.* [1998][24] applied this idea to the Maxwell–Vlasov equations of plasma physics. Cendra *et al.* [1998][25] showed how Lagrangian semidirect product theory fits into the general framework of Lagrangian reduction.

Reduction by stages and group extensions. The semidirect product reduction theorem can be viewed using reduction by stages: If one reduces T^*S by the action of the semidirect product group $S=G\circledS V$ in two stages, first by the action of V at a point a_0 and then by the action of G_{a_0}. Semidirect product reduction by stages for actions of semidirect products on general symplectic manifolds was developed and applied to underwater vehicle dynamics in Leonard and Marsden [1997].[84] Motivated partly by semidirect product reduction, Marsden *et al.* [1998, 1999][90,91] gave a significant generalization of semidirect product theory in which one has a group M with a normal subgroup $N\subset M$ (so M is a group extension of N) and M acts on a symplectic manifold P. One wants to reduce P in two stages, first by N and then by M/N. On the Poisson level this is easy: $P/M\cong(P/N)/(M/N)$, but on the symplectic level it is quite subtle.

Cotangent bundle reduction by stages is especially interesting for group extensions. An example of such a group, besides semidirect products, is the Bott–Virasoro group, where the Gelfand–Fuchs cocycle may be interpreted as the curvature of a mechanical connection. The work of Cendra, Marsden, and Ratiu [2000][28] briefly described previously, contains a Lagrangian analog of reduction for group extensions and reduction by stages.

Singular reduction. Singular reduction starts with the observation of Smale [1970][147] that $z\in P$ is a regular point of $\mathbf{J}$ iff z has no continuous isotropy. Motivated by this, Arms, Marsden, and Moncrief [1981, 1982][6,7] showed that the level sets $\mathbf{J}^{-1}(0)$ of an equivariant momentum map $\mathbf{J}$ have quadratic singularities at points with continuous symmetry. While such a result is easy for compact group actions on finite dimensional manifolds, the main examples of Arms, Marsden, and Moncrief [1981][6] were, in fact, infinite dimensional—both the phase space and the group. Otto [1987][130] has shown that if G is a compact Lie group, $\mathbf{J}^{-1}(0)/G$ is an orbifold. Singular reduction is closely related to convexity properties of the momenum map.[42,45]

The detailed structure of $\mathbf{J}^{-1}(0)/G$ for compact Lie groups acting on finite dimensional manifolds was developed in Sjamaar and Lerman [1991][146] and extended for proper Lie group actions to $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ by Bates and Lerman [1997],[13] if $\mathcal{O}_\mu$ is locally closed in $\mathfrak{g}^*$. Ortega [1998][129] and Ortega and Ratiu [2000][128] redid the entire singular reduction theory for proper Lie group actions starting with the point reduced spaces $\mathbf{J}^{-1}(\mu)/G_\mu$ and also connected it to the more algebraic approach to reduction theory of Arms, Cushman, and Gotay [1991].[5] Specific examples of singular reduction and further references may be found in Cushman and Bates [1997].[33] This theory is still under development.

The method of invariants. This method seeks to parametrize quotient spaces by group invariant functions. It has a rich history going back to Hilbert's *invariant theory*. It has been of great use in bifurcation with symmetry (see Golubitsky, Stewart, and Schaeffer [1988][39] for instance). In mechanics, the method was developed by Kummer, Cushman, Rod, and co-workers in the 1980s. We will not attempt to give a literature survey here, other than to refer to Kummer [1990],[78] Kirk, Marsden, and Silber [1996],[68] Alber *et al.* [1998],[4] and the book of Cushman and Bates [1997][33] for more details and references.

The new results in this paper. The main new results of the present paper are as follows.

(1) In Sec. III A, a global realization of the reduced tangent bundle, with a momentum map constraint, in terms of a fiber product bundle, which is shown to also be globally diffeomorphic to an associated coadjoint orbit bundle.
(2) Section III E shows how to drop Hamilton's variational principle to these quotient spaces.
(3) We derive, in Sec. III H, the corresponding reduced equations, which we call the *Lagrange–Routh equations*, in an intrinsic and global fashion.
(4) In Sec. IV we give a Lagrangian view of some known and new reconstruction and geometric phase formulas.

The Euler free rigid body, the heavy top, and the underwater vehicle are used to illustrate some of the points of the theory. The main techniques used in this paper build primarily on the work of Marsden and Scheurle [1993a, 1993c][107,108] and of Jalnapurkar and Marsden [2000a][64] on non-Abelian Routh reduction theory, but with the recent developments in Cendra, Marsden, and Ratiu [2000][28] in mind.

B. Bundles, momentum maps, and Lagrangians

The shape space bundle and Lagrangian. We shall be primarily concerned with the following setting. Let Q be a configuration manifold and let G be a Lie group that acts freely and properly on Q. The quotient $Q/G:=S$ is referred to as the *shape space* and Q is regarded as a principal fiber bundle over the base space S. Let $\pi_{Q,G}:Q\to Q/G=S$ be the canonical projection. The theory of quotient manifolds guarantees (because the action is free and proper) that Q/G is a smooth manifold and the map $\pi_{Q,G}$ is smooth. See Abraham, Marsden, and Ratiu [1988][3] for the proof of these statements. We call the map $\pi_{Q,G}:Q\to Q/G$ the *shape space bundle*.

Let $\langle\langle\cdot,\ \cdot\rangle\rangle$ be a G-invariant metric on Q, also called a *mass matrix*. The *kinetic energy* $K:TQ\to\mathbb{R}$ is defined by $K(v_q)=\frac{1}{2}\langle\langle v_q,v_q\rangle\rangle$. If V is a G-invariant *potential* on Q, then the Lagrangian $L=K-V:TQ\to\mathbb{R}$ is also G-invariant. We focus on Lagrangians of this form, although much of what we do can be generalized. We make a few remarks concerning this in the body of the paper.

Momentum map, mechanical connection, and locked inertia. Let G have Lie algebra $\mathfrak{g}$ and $\mathbf{J}_L:TQ\to\mathfrak{g}^*$ be the momentum map on TQ, which is defined by $\mathbf{J}_L(v_q)\cdot\xi=\langle\langle v_q,\xi_Q(q)\rangle\rangle$. Here $v_q\in T_qQ$, $\xi\in\mathfrak{g}$, and ξ_Q denotes the infinitesimal generator corresponding to ξ.

Recall that a *principal connection* $\mathcal{A}:TQ\to\mathfrak{g}$ is an equivariant $\mathfrak{g}$-valued one-form on TQ that satisfies $\mathcal{A}(\xi_Q(q))=\xi$ and its kernel at each point, denoted Hor_q, complements the vertical space, namely the tangents to the group orbits. Let $\mathfrak{A}:TQ\to\mathfrak{g}$ be the *mechanical connection*, namely the principal connection whose horizontal spaces are orthogonal to the group orbits. (Shape space and its geometry also play an interesting and key role in computer vision. See e.g., Le and Kendall [1993].[83]) For each $q\in Q$, the *locked inertia tensor* $\mathfrak{I}(q):\mathfrak{g}\to\mathfrak{g}^*$, is defined by the equation $\langle\mathfrak{I}(q)\xi,\eta\rangle=\langle\langle\xi_Q(q),\eta_Q(q)\rangle\rangle$. The locked inertia tensor has the following equivariance property: $\mathfrak{I}(g\cdot q)=Ad^*_{g^{-1}}\,\mathfrak{I}(q)\mathrm{Ad}_{g^{-1}}$, where the adjoint action by a group element g is denoted Ad_g and $\mathrm{Ad}^*_{g^{-1}}$ denotes the dual of the linear map $\mathrm{Ad}_{g^{-1}}:\mathfrak{g}\to\mathfrak{g}$. The mechanical connection $\mathfrak{A}$ and the momentum map $\mathbf{J}_L$ are related as follows:

$$\begin{aligned}&\mathbf{J}_L(v_q)=\mathfrak{I}(q)\mathfrak{A}(v_q),\\ &\text{i.e.,}\quad \mathfrak{A}(v_q)=\mathfrak{I}(q)^{-1}\mathbf{J}_L(v_q).\end{aligned}\tag{I.1}$$

In particular, or from the definitions, we have that $\mathbf{J}_L(\xi_Q(q))=\mathfrak{I}(q)\xi$. For free actions and a Lagrangian of the form kinetic minus potential, the locked inertia tensor is invertible at each $q\in Q$. Many of the constructions can be generalized to the case of regular Lagrangians, where the locked inertia tensor is the second fiber derivative of L (see Lewis [1992][87]).

Horizontal and vertical decomposition. We use the mechanical connection $\mathfrak{A}$ to express v_q (also denoted $\dot{q}$) as the sum of horizontal and vertical components:

$$v_q = \mathrm{Hor}(v_q) + \mathrm{Ver}(v_q) = \mathrm{Hor}(v_q) + \xi_Q(q),$$

where $\xi = \mathfrak{A}(v_q)$. Thus, the kinetic energy is given by

$$K(v_q) = \tfrac{1}{2}\langle\langle v_q, v_q\rangle\rangle = \tfrac{1}{2}\langle\langle \mathrm{Hor}(v_q), \mathrm{Hor}(v_q)\rangle\rangle + \tfrac{1}{2}\langle\langle \xi_Q(q), \xi_Q(q)\rangle\rangle.$$

Being G-invariant, the metric on Q induces a metric $\langle\langle\cdot,\cdot\rangle\rangle_S$ on S by $\langle\langle u_x, v_x\rangle\rangle_S = \langle\langle u_q, v_q\rangle\rangle$, where $u_q, v_q \in T_qQ$ are horizontal, $\pi_{Q,G}(q) = x$ and $T\pi_{Q,G}\cdot u_q = u_x$, $T\pi_{Q,G}\cdot v_q = v_x$.

Useful formulas for group actions. The following formulas are assembled for convenience (see, e.g., Marsden and Ratiu [1994][98] for the proofs). We denote the action of $g \in G$ on a point $q \in Q$ by $gq = g\cdot q = \Phi_g(q)$, so that $\Phi_g : Q \to Q$ is a diffeomorphism.

(1) Transformations of generators: $T\Phi_g \cdot \xi_Q(q) = (\mathrm{Ad}_g\, \xi)_Q(g\cdot q)$. which we also write, using concatenation notation for actions, as $g\cdot \xi_Q(q) = (\mathrm{Ad}_g\, \xi)_Q(g\cdot q)$.

(2) Brackets of generators: $[\xi_Q, \eta_Q] = -[\xi, \eta]_Q$.

(3) Derivatives of curves. Let $q(t)$ be a curve in Q and let $g(t)$ be a curve in G. Then

$$\frac{d}{dt}(g(t)\cdot q(t)) = (\mathrm{Ad}_{g(t)}\, \xi(t))_Q(g(t)\cdot q(t)) + g(t)\cdot \dot{q}(t) = g(t)\cdot[(\xi(t))_Q(q(t)) + \dot{q}(t)], \quad \text{(I.2)}$$

where $\xi(t) = g(t)^{-1}\cdot \dot{g}(t)$.

It is useful to recall the *Cartan formula*. Let α be a one form and let X and Y be two vector fields on a manifold. Then the exterior derivative $\mathbf{d}\alpha$ of α is related to the Jacobi–Lie bracket of vector fields by $\mathbf{d}\alpha(X,Y) = X[\alpha(Y)] - Y[\alpha(X)] - \alpha([X,Y])$.

C. Coordinate formulas

We next give a few coordinate formulas for the case when G is *Abelian*.

The coordinates and Lagrangian. In a local trivialization, Q is realized as $U \times G$ where U is an open set in shape space $S = Q/G$. We can accordingly write coordinates for Q as x^α, θ^a where x^α, $\alpha = 1,\ldots,n$ are coordinates on S and where θ^a, $a = 1,\ldots,r$ are coordinates for G. In a local trivialization, θ^a are chosen to be cyclic coordinates in the classical sense. We write L (with the summation convention in force) as

$$L(x^\alpha, \dot{x}^\beta, \dot{\theta}^a) = \tfrac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{\theta}^a + \tfrac{1}{2} g_{ab}\dot{\theta}^a \dot{\theta}^b - V(x^\alpha). \quad \text{(I.3)}$$

The momentum conjugate to the cyclic variable θ^a is $J_a = \partial L/\partial \dot{\theta}^a = g_{\alpha a}\dot{x}^\alpha + g_{ab}\dot{\theta}^b$, which are the components of the map $\mathbf{J}_L$.

Mechanical connection and locked inertia tensor. The locked inertia tensor is the matrix $I_{ab} = g_{ab}$ and its inverse is denoted $I^{ab} = g^{ab}$. The matrix I_{ab} is the block in the matrix of the metric tensor g_{ij} associated with the group variables and, of course, I^{ab} need not be the corresponding block in the inverse matrix g^{ij}. The mechanical connection, as a vector valued one form, is given by $\mathfrak{A}^a = d\theta^a + \mathfrak{A}^a_\alpha\, dx^\alpha$, where the components of the mechanical connection are defined by $\mathfrak{A}^b_\alpha = g^{ab} g_{a\alpha}$. Notice that the relation $\mathbf{J}_L(v_q) = \mathfrak{I}(q)\cdot \mathfrak{A}(v_q)$ is clear from this component formula.

Horizontal and vertical projections. For a vector $v = (\dot{x}^\alpha, \dot{\theta}^a)$, and suppressing the base point (x^α, θ^a) in the notation, its horizontal and vertical projections are verified to be

$$\mathrm{Hor}(v) = (\dot{x}^\alpha, -g^{ab} g_{\alpha b}\dot{x}^\alpha), \quad \mathrm{Ver}(v) = (0, \dot{\theta}^a + g^{ab} g_{\alpha b}\dot{x}^\alpha).$$

Notice that $v = \mathrm{Hor}(v) + \mathrm{Ver}(v)$, as it should.

Horizontal metric. In coordinates, the horizontal kinetic energy is

$$\begin{aligned}\tfrac{1}{2} g(\mathrm{Hor}(v), \mathrm{Hor}(v)) &= \tfrac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta - g_{a\alpha} g^{ab} g_{b\beta}\dot{x}^\alpha \dot{x}^\beta + \tfrac{1}{2} g_{a\alpha} g^{ab} g_{b\beta}\dot{x}^\alpha \dot{x}^\beta \\ &= \tfrac{1}{2}(g_{\alpha\beta} - g_{a\alpha} g^{ab} g_{b\beta})\dot{x}^\alpha \dot{x}^\beta. \end{aligned} \quad \text{(I.4)}$$

Thus, the components of the horizontal metric (the metric on shape space) are given by $A_{\alpha\beta} = g_{\alpha\beta} - g_{\alpha d} g^{da} g_{\beta a}$.

D. Variational principles

Variations and the action functional. Let $q:[a,b]\to Q$ be a curve and let $\delta q = (d/d\epsilon)|_{\epsilon=0} q_\epsilon$ be a variation of q. Given a Lagrangian L, let the associated *action functional* $\mathfrak{S}_L(q_\epsilon)$ be defined on the space of curves in Q defined on a fixed interval $[a,b]$ by

$$\mathfrak{S}_L(q_\epsilon) = \int_a^b L(q_\epsilon, \dot{q}_\epsilon)\, dt.$$

The differential of the action function is given by the following theorem.

Theorem I.1: *Given a smooth Lagrangian L, there is a unique mapping* $\mathcal{EL}(L):\ddot{Q}\to T^*Q$, defined on the *second-order submanifold*

$$\ddot{Q} \equiv \left\{ \frac{d^2 q}{dt^2}(0) \,\middle|\, q \text{ a smooth curve in } Q \right\}$$

of TTQ, *and a unique one-form* Θ_L *on* TQ, *such that, for all variations* $\delta q(t)$,

$$\mathbf{d}\mathfrak{S}_L(q(t)) \cdot \delta q(t) = \int_a^b \mathcal{EL}(L)\left(\frac{d^2 q}{dt^2}\right) \cdot \delta q\, dt + \Theta_L\left(\frac{dq}{dt}\right) \cdot \widehat{\delta q}\,\Big|_a^b, \tag{I.5}$$

where

$$\delta q(t) \equiv \frac{d}{d\epsilon}\Big|_{\epsilon=0} q_\epsilon(t), \quad \widehat{\delta q}(t) \equiv \frac{d}{d\epsilon}\Big|_{\epsilon=0} \frac{d}{dt}\Big|_{t=0} q_\epsilon(t).$$

The one-form Θ_L *so defined is called the Lagrange one-form.*

The Lagrange one-form defined by this theorem coincides with the Lagrange one-form obtained by pulling back the canonical form on T^*Q by the Legendre transformation. This term is readily shown to be given by

$$\Theta_L\left(\frac{dq}{dt}\right) \cdot \widehat{\delta q}|_a^b = \langle \mathbb{F}L(q(t)\cdot \dot{q}(t)), \delta q\rangle|_a^b.$$

In verifying this, one checks that the projection of $\widehat{\delta q}$ from TTQ to TQ under the map $T\tau_Q$, where $\tau_Q: TQ\to Q$ is the standard tangent bundle projection map, is δq. Here we use $\mathbb{F}L: TQ\to T^*Q$ for the fiber derivative of L.

E. Euler–Poincaré reduction

In rigid body mechanics, the passage from the attitude matrix and its velocity to the body angular velocity is an example of Euler–Poincaré reduction. Likewise, in fluid mechanics, the passage from the Lagrangian (material) representation of a fluid to the Eulerian (spatial) representation is an example of Euler–Poincaré reduction. These examples are well known and are spelled out in, e.g., Marsden and Ratiu [1994].[98]

For $g\in G$, let $TL_g: TG\to TG$ be the tangent of the left translation map $L_g: G\to G;\ h\mapsto gh$. Let $L: TG\to\mathbb{R}$ be a left invariant Lagrangian. For what follows, L does not have to be purely kinetic energy (any invariant potential would be a constant, so is ignored), although this is one of the most important cases.

Theorem I.2 (Euler–Poincaré reduction): *Let l: $\mathfrak{g}\to\mathbb{R}$ be the restriction of L to $\mathfrak{g}=T_eG$. For a curve $g(t)$ in G, let $\xi(t)=TL_{g(t)^{-1}}\dot{g}(t)$, or using concatenation notation, $\xi=g^{-1}\dot{g}$. The following are equivalent.*

(a) *The curve $g(t)$ satisfies the Euler–Lagrange equations on G.*

(b) *The curve $g(t)$ is an extremum of the action functional*

$$\mathfrak{S}_L(g(\cdot))=\int L(g(t),\dot{g}(t))dt,$$

for variations δg with fixed end points.

(c) *The curve $\xi(t)$ solves the Euler–Poincaré equations*

$$\frac{d}{dt}\frac{\delta l}{\delta\xi}=\mathrm{ad}^*_\xi\frac{\delta l}{\delta\xi}, \tag{I.6}$$

where the coadjoint action ad^*_ξ is defined by $\langle \mathrm{ad}^*_\xi\,\nu,\zeta\rangle=\langle\nu,[\xi,\zeta]\rangle$, where ξ, $\zeta\in\mathfrak{g}$, $\nu\in\mathfrak{g}^*$, $\langle\cdot,\cdot\rangle$ is the pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$, and $[\cdot,\cdot]$ is the Lie algebra bracket.

(d) *The curve $\xi(t)$ is an extremum of the reduced action functional*

$$\mathfrak{s}_l(\xi)=\int l(\xi(t))dt,$$

for variations of the form $\delta\xi=\dot{\eta}+[\xi,\eta]$, where $\eta=TL_{g^{-1}}\delta g=g^{-1}\delta g$ vanishes at the end points.

There is, of course, a similar statement for right invariant Lagrangians; one needs to change the sign on the right-hand side of (I.6) and use variations of the form $\delta\xi=\dot{\eta}-[\xi,\eta]$. See Marsden and Scheurle [1993b][108] and Sec. 13.5 of Marsden and Ratiu [1994][98] for a proof of this theorem for the case of matrix groups and Bloch, Krishnaprasad, Marsden, and Ratiu [1996][17] for the case of general finite dimensional Lie groups. For discussions of the infinite dimensional case, see Kouranbaeva [1999][76] and Marsden, Ratiu, and Shkoller [1999].[99]

F. Lie–Poisson reduction

Lie–Poisson reduction is the Poisson counterpart to Euler–Poincaré reduction. The dual space $\mathfrak{g}^*$ is a Poisson manifold with either of the two *Lie–Poisson* brackets

$$\{f,k\}_\pm(\mu)=\pm\left\langle\mu,\left[\frac{\delta f}{\delta\mu},\frac{\delta k}{\delta\mu}\right]\right\rangle, \tag{I.7}$$

where $\delta f/\delta\mu\in\mathfrak{g}$ is defined by $\langle\nu,\delta f/\delta\mu\rangle=\mathbf{D}f(\mu)\cdot\nu$ for $\nu\in\mathfrak{g}^*$, and where $\mathbf{D}$ denotes the Fréchet derivative. (In the infinite dimensional case one needs to worry about the existence of $\delta f/\delta\mu$. See, for instance, Marsden and Weinstein [1982, 1983][111,112] for applications to plasma physics and fluid mechanics and Marsden and Ratiu [1994][98] for additional references. The notation $\delta f/\delta\mu$ is used to conform to the functional derivative notation in classical field theory.) In coordinates, $(\xi^1,\ldots,\xi^m)$ on $\mathfrak{g}$ relative to a vector space basis $\{e_1,\ldots,e_m\}$ and corresponding dual coordinates $(\mu_1,\ldots,\mu_m)$ on $\mathfrak{g}^*$, the bracket (I.7) is

$$\{f,k\}_\pm(\mu)=\pm\mu_aC^a_{bc}\frac{\partial f}{\partial\mu_b}\frac{\partial k}{\partial\mu_c},$$

where C^a_{bc} are the structure constants of $\mathfrak{g}$ defined by $[e_a,e_b]=C^c_{ab}e_c$. The Lie–Poisson bracket appears explicitly in Lie [1890][88] Sec. 75, see Weinstein [1983].[156]

Which sign to take in (I.7) is determined by understanding how the Lie–Poisson bracket is related to *Lie–Poisson reduction*, which can be summarized as follows. Consider the left and right translation maps to the identity: $\lambda:T^*G\to\mathfrak{g}^*$ defined by $\alpha_g\mapsto(T_eL_g)^*\alpha_g\in T^*_eG=\mathfrak{g}^*$ and

$\rho: T^*G \to \mathfrak{g}^*$, defined by $\alpha_g \mapsto (T_e R_g)^* \alpha_g \in T_e^* G = \mathfrak{g}^*$. Let $\mathfrak{g}_-^*$ denote $\mathfrak{g}^*$ with the minus Lie–Poisson bracket and let $\mathfrak{g}_+^*$ be $\mathfrak{g}^*$ with the plus Lie–Poisson bracket. We use the canonical structure on T^*Q unless otherwise noted.

Theorem I.3 (Lie-Poisson reduction–geometry): *The maps*

$$\lambda: T^*Q \to \mathfrak{g}_-^* \quad \textit{and} \quad \rho: T^*Q \to \mathfrak{g}_+^*$$

are Poisson maps.

This procedure uniquely characterizes the Lie–Poisson bracket and provides a basic example of Poisson reduction. For example, using the left action, λ induces a Poisson diffeomorphism $[\lambda]: (T^*G)/G \to \mathfrak{g}_-^*$.

Every left invariant Hamiltonian and Hamiltonian vector field is mapped by λ to a Hamiltonian and Hamiltonian vector field on $\mathfrak{g}^*$. There is a similar statement for right invariant systems on T^*G. One says that the original system on T^*G has been *reduced* to $\mathfrak{g}^*$. One way to see that λ and ρ are Poisson maps is by observing that they are equivariant momentum maps for the action of G on itself by right and left translations, respectively, together with the fact that equivariant momentum maps are Poisson maps. The fact that equivariant momentum maps are Poisson again has a cloudy history. It was given implicitly in the works of Lie and in Guillemin and Sternberg [1980][44] and explicitly in Marsden *et al.* [1982],[113] and Holmes and Marsden [1983].[51]

If $(P, \{\,,\,\})$ is a Poisson manifold, a function $C \in \mathcal{F}(P)$ satisfying $\{C, f\} = 0$ for all $f \in \mathcal{F}(P)$ is called a *Casimir function*. Casimir functions are constants of the motion for *any* Hamiltonian since $\dot{C} = \{C, H\} = 0$ for any H. Casimir functions and momentum maps play a key role in the stability theory of relative equilibria (see, e.g., Marsden [1992][103] and Marsden and Ratiu [1994][98] and references therein for a discussion of the relation between Casimir functions and momentum maps).

Theorem I.4 (Lie–Poisson reduction–dynamics): *Let $H: T^*G \to \mathbb{R}$ be a left invariant Hamiltonian and $h: \mathfrak{g}^* \to \mathbb{R}$ its restriction to the identity. For a curve $\alpha(t) \in T_{g(t)}^* G$, let $\mu(t) = (T_e^* L_{g(t)}) \cdot \alpha(t) = \lambda(\alpha(t))$ be the induced curve in $\mathfrak{g}^*$. The following are equivalent:*

(i) *$\alpha(t)$ is an integral curve of X_H, i.e., Hamilton's equations on T^*G hold.*

(ii) *For any smooth function $F \in \mathcal{F}(T^*G)$, $\dot{F} = \{F, H\}$ along $\alpha(t)$, where $\{\,,\,\}$ is the canonical bracket on T^*G.*

(iii) *$\mu(t)$ satisfies the Lie–Poisson equations*

$$\frac{d\mu}{dt} = \mathrm{ad}^*_{\delta h/\delta \mu}\, \mu, \tag{I.8}$$

where $\mathrm{ad}_\xi: \mathfrak{g} \to \mathfrak{g}$ is defined by $\mathrm{ad}_\xi\, \eta = [\xi, \eta]$ and ad^*_ξ is its dual.

(iv) *For any $f \in \mathcal{F}(\mathfrak{g}^*)$, we have $\dot{f} = \{f, h\}_-$ along $\mu(t)$, where $\{\,,\,\}_-$ is the minus Lie–Poisson bracket.*

There is a similar statement in the right invariant case with $\{\cdot,\cdot\}_-$ replaced by $\{\cdot,\cdot\}_+$ and a sign change on the right-hand side of (I.8).

The Lie–Poisson equations in coordinates are $\dot{\mu}_a = C_{ba}^d (\delta h/\delta \mu_b) \mu_d$.

Given a reduced Lagrangian $l: \mathfrak{g} \to \mathbb{R}$, when the reduced Legendre transform $\mathbb{F}l: \mathfrak{g} \to \mathfrak{g}^*$ defined by $\xi \mapsto \mu = \delta l/\delta \xi$ is a diffeomorphism (this is the *regular case*), then this map takes the Euler–Poincaré equations to the Lie–Poisson equations. There is, of course a similar inverse map starting with a reduced Hamiltonian.

Additional history. The symplectic and Poisson theory of mechanical systems on Lie groups could easily have been given shortly after Lie's work, but amazingly it was not observed for the rigid body or ideal fluids until the work of Pauli [1953],[131] Martin [1959],[105] Arnold [1966],[8] Ebin and Marsden [1970],[34] Nambu [1973],[125] and Sudarshan and Mukunda [1974],[150] all of whom were apparently unaware of Lie's work on the Lie–Poisson bracket and of Poincaré's [1901][133] work on the Euler–Poincaré equations. One is struck by the large amount of rediscovery and confusion in this subject, which evidently is not unique to mechanics.

Arnold [1988][11] and Chetaev [1989][32] brought Poincaré's work on the Euler–Poincaré equations to the attention of the community. Poincaré [1910][134] goes on to study the effects of the deformation of the earth on its precession—he apparently recognizes the equations as Euler equations on a semidirect product Lie algebra. Poincaré [1901][133] has no bibliographic references, so it is rather hard to trace his train of thought or his sources; in particular, he gives no hints that he understood the work of Lie on the Lie–Poisson structure.

In the dynamics of ideal fluids, the Euler–Poincaré variational principle is essentially that of ''Lin constraints.'' See Cendra and Marsden [1987][27] for a discussion of this theory and for further references. Variational principles in fluid mechanics itself has an interesting history, going back to Ehrenfest, Boltzmann, and Clebsch, but again, there was little, if any, contact with the heritage of Lie and Poincaré on the subject. Interestingly, Seliger and Witham [1968][144] remarked that ''Lin's device still remains somewhat mysterious from a strictly mathematical view.'' See also Bretherton [1970].[22]

Lagrange [1788],[80] Vol. 2, Eq. A on p. 212, are the Euler–Poincaré equations for the rotation group written out explicitly for a reasonably general Lagrangian. Lagrange also developed the key concept of the Lagrangian representation of fluid motion, but it is not clear that he understood that both systems are special instances of one theory. Lagrange spends a large number of pages on his derivation of the Euler–Poincaré equations for SO(3), in fact, it is a good chunk of Vol. 2 of *Mécanique Analytique*.

G. Examples

The free rigid body—The Euler top. Let us first review some basics of the rigid body. We regard an element. $\mathbf{A}\in \mathrm{SO}(3)$, giving the configuration of the body as a map of a reference configuration $\mathcal{B}\subset\mathbb{R}^3$ to the current configuration $\mathbf{A}(\mathcal{B})$; the map $\mathbf{A}$ takes a reference or label point $X\in\mathcal{B}$ to a current point $x=\mathbf{A}(X)\in\mathbf{A}(\mathcal{B})$. When the rigid body is in motion, the matrix $\mathbf{A}$ is time dependent and the velocity of a point of the body is $\dot{x}=\dot{\mathbf{A}}X=\dot{\mathbf{A}}\mathbf{A}^{-1}x$. Since $\mathbf{A}$ is an orthogonal matrix, $\mathbf{A}^{-1}\dot{\mathbf{A}}$ and $\dot{\mathbf{A}}\mathbf{A}^{-1}$ are skew matrices, and so we can write $\dot{x}=\dot{\mathbf{A}}\mathbf{A}^{-1}x=\boldsymbol{\omega}\times x$, which defines the *spatial angular velocity vector* ω. The corresponding *body angular velocity* is defined by $\boldsymbol{\Omega}=\mathbf{A}^{-1}\boldsymbol{\omega}$, i.e., $\mathbf{A}^{-1}\dot{\mathbf{A}}v=\boldsymbol{\Omega}\times v$ so that $\boldsymbol{\Omega}$ is the angular velocity relative to a body fixed frame. The kinetic energy is

$$K=\frac{1}{2}\int_{\mathcal{B}}\rho(X)\|\dot{\mathbf{A}}X\|^2 d^3X, \tag{I.9}$$

where ρ is a given mass density in the reference configuration. Since

$$\|\dot{\mathbf{A}}X\|=\|\boldsymbol{\omega}\times x\|=\|\mathbf{A}^{-1}(\omega\times x)\|=\|\boldsymbol{\Omega}\times X\|,$$

K is a quadratic function of $\boldsymbol{\Omega}$. Writing $K=\frac{1}{2}\boldsymbol{\Omega}^T I\boldsymbol{\Omega}$ defines the *moment of inertia tensor* I, which, if the body does not degenerate to a line, is a positive definite 3×3 matrix, or equivalently, a quadratic form. This quadratic form can be diagonalized, and this defines the principal axes and moments of inertia. In this basis, we write $I=\mathrm{diag}(I_1,I_2,I_3)$.

The function $K(\mathbf{A},\dot{\mathbf{A}})$ is taken to be the Lagrangian of the system on $T\mathrm{SO}(3)$. It is left invariant. The reduced Lagrangian is $k(\boldsymbol{\Omega})=\frac{1}{2}\boldsymbol{\Omega}^T I\boldsymbol{\Omega}$. One checks that the *Euler–Poincaré equations are given by the classical Euler equations for a rigid body:*

$$\dot{\boldsymbol{\Pi}}=\boldsymbol{\Pi}\times\boldsymbol{\Omega}, \tag{I.10}$$

where $\boldsymbol{\Pi}=I\boldsymbol{\Omega}$ is the body angular momentum. The corresponding reduced variational principle is

$$\delta\int_a^b l(\boldsymbol{\Omega}(t))dt=0$$

for variations of the form $\delta\boldsymbol{\Omega}=\dot{\boldsymbol{\Sigma}}+\boldsymbol{\Omega}\times\boldsymbol{\Sigma}$.

By means of the Legendre transformation, we get the corresponding Hamiltonian description on $T^*\mathrm{SO}(3)$. The reduced Hamiltonian is given by $h(\mathbf{\Pi})=\frac{1}{2}\mathbf{\Pi}\cdot(I^{-1}\mathbf{\Pi})$. One can verify directly from the chain rule and properties of the triple product that Euler's equations are also equivalent to the following equation for all $f\in\mathcal{F}(\mathbb{R}^3)$: $\dot{f}=\{f,h\}$, where the corresponding (minus) Lie–Poisson structure on $\mathbb{R}^3$ is given by

$$\{f,k\}(\mathbf{\Pi})=-\mathbf{\Pi}\cdot(\nabla f\times\nabla k). \tag{I.11}$$

Every function $C:\mathbb{R}^3\to\mathbb{R}$ of the form $C(\mathbf{\Pi})=\Phi(\|\mathbf{\Pi}\|^2)$, where $\Phi:\mathbb{R}\to\mathbb{R}$ is a differentiable function, is a Casimir function, as is readily checked. In particular, for the rigid body, $\|\mathbf{\Pi}\|^2$ is a constant of the motion.

In the notation of the general theory, one chooses $Q=G=\mathrm{SO}(3)$ with G acting on itself by left multiplication. The shape space is $Q/G=a$ single point.

As explained previously, the free rigid body kinetic energy is given by the left invariant metric on $Q=\mathrm{SO}(3)$ whose value at the identity is $\langle\langle\mathbf{\Omega}_1,\mathbf{\Omega}_2\rangle\rangle=I\mathbf{\Omega}_1\cdot\mathbf{\Omega}_2$, where $\mathbf{\Omega}_1,\mathbf{\Omega}_2\in\mathbb{R}^3$ are thought of as elements of $\mathfrak{so}(3)$, the Lie algebra of SO(3), via the isomorphism $\mathbf{\Omega}\in\mathbb{R}^3\mapsto\hat{\mathbf{\Omega}}\in\mathfrak{so}(3)$, $\hat{\mathbf{\Omega}}\mathbf{v}:=\mathbf{\Omega}\times\mathbf{v}$. The Lagrangian equals the kinetic energy.

The infinitesimal generator of $\hat{\boldsymbol{\xi}}\in\mathfrak{so}(3)$ for the action of G is, according to the definitions, given by $\hat{\boldsymbol{\xi}}_{\mathrm{SO}(3)}(\mathbf{A})=\hat{\boldsymbol{\xi}}\mathbf{A}\in T_{\mathbf{A}}\mathrm{SO}(3)$. The locked inertia tensor is, for each $\mathbf{A}\in\mathrm{SO}(3)$, the linear map $\mathfrak{J}(\mathbf{A}):\mathfrak{so}(3)\to\mathfrak{so}(3)^*$ given by $\langle\mathfrak{J}(\mathbf{A})\hat{\boldsymbol{\xi}},\hat{\boldsymbol{\eta}}\rangle=\langle\langle\hat{\boldsymbol{\xi}}_Q(\mathbf{A}),\hat{\boldsymbol{\eta}}_Q(\mathbf{A})\rangle\rangle=\langle\langle\hat{\boldsymbol{\xi}}\mathbf{A},\hat{\boldsymbol{\eta}}\mathbf{A}\rangle\rangle$. Since the metric is left SO(3)-invariant, and using the general identity $(\mathbf{A}^{-1}\boldsymbol{\xi})\hat{}=\mathbf{A}^{-1}\hat{\boldsymbol{\xi}}\mathbf{A}$, this equals

$$\langle\langle\mathbf{A}^{-1}\hat{\boldsymbol{\xi}}\mathbf{A},\mathbf{A}^{-1}\hat{\boldsymbol{\eta}}\mathbf{A}\rangle\rangle=\langle\langle\mathbf{A}^{-1}\boldsymbol{\xi},\mathbf{A}^{-1}\boldsymbol{\eta}\rangle\rangle=(\mathbf{A}^{-1}\boldsymbol{\xi})\cdot(I\mathbf{A}^{-1}\boldsymbol{\eta})=(\mathbf{A}I\mathbf{A}^{-1}\boldsymbol{\xi})\cdot\boldsymbol{\eta}.$$

Thus, identifying $\mathfrak{J}(\mathbf{A})$ with a linear map of $\mathbb{R}^3$ to itself, we get $\mathfrak{J}(\mathbf{A})=\mathbf{A}I\mathbf{A}^{-1}$.

Now we use the general definition $\langle\mathbf{J}_L(v_q),\xi\rangle=\langle\langle v_q,\xi_Q(q)\rangle\rangle$ to compute the momentum map $\mathbf{J}_L:T\,\mathrm{SO}(3)\to\mathbb{R}$ for the action of G. Using the definition $\hat{\mathbf{\Omega}}=\mathbf{A}^{-1}\dot{\mathbf{A}}$, we get

$$\langle\mathbf{J}_L(\mathbf{A},\dot{\mathbf{A}}),\hat{\boldsymbol{\xi}}\rangle=\langle\langle\dot{\mathbf{A}},\hat{\boldsymbol{\xi}}\mathbf{A}\rangle\rangle_{\mathbf{A}}=\langle\langle\mathbf{A}^{-1}\dot{\mathbf{A}},\mathbf{A}^{-1}\hat{\boldsymbol{\xi}}\mathbf{A}\rangle\rangle_{\mathbf{I}}=(I\mathbf{\Omega})\cdot(\mathbf{A}^{-1}\boldsymbol{\xi})=(\mathbf{A}I\mathbf{\Omega})\cdot\boldsymbol{\xi}.$$

Letting $\boldsymbol{\pi}=\mathbf{A}\mathbf{\Pi}$, where $\mathbf{\Pi}=I\mathbf{\Omega}$, we get $\mathbf{J}_L(\mathbf{A},\dot{\mathbf{A}})=\boldsymbol{\pi}$, the spatial angular momentum.

According to the general formula $\mathfrak{A}(v_q)=\mathfrak{J}(q)^{-1}\mathbf{J}_L(v_q)$, the mechanical connection $\mathfrak{A}(\mathbf{A}):T_{\mathbf{A}}\mathrm{SO}(3)\to\mathfrak{so}(3)$ is given by $\mathfrak{A}(\mathbf{A},\dot{\mathbf{A}})=\mathbf{A}I^{-1}\mathbf{A}^{-1}\boldsymbol{\pi}=\mathbf{A}\mathbf{\Omega}$. This is $\mathfrak{A}(\mathbf{A})$ regarded as taking values in $\mathbb{R}^3$. Regarded as taking values in $\mathfrak{so}(3)$, the space of skew matrices, we get $\mathfrak{A}(\mathbf{A},\dot{\mathbf{A}})=\widehat{\mathbf{A}\mathbf{\Omega}}=\mathbf{A}\hat{\mathbf{\Omega}}\mathbf{A}^{-1}=\dot{\mathbf{A}}\mathbf{A}^{-1}$, the spatial angular velocity. Notice that the mechanical connection is independent of the moment of inertia of the body.

The heavy top. The system is a spinning rigid body with a fixed point in a gravitational field, as shown in Fig. 1.

One usually finds the equations written as

$$\dot{\mathbf{\Pi}}=\mathbf{\Pi}\times\mathbf{\Omega}+Mgl\mathbf{\Gamma}\times\boldsymbol{\chi},$$

$$\dot{\mathbf{\Gamma}}=\mathbf{\Gamma}\times\mathbf{\Omega}.$$

Here, M is the body's mass, $\mathbf{\Pi}$ is the body angular momentum, $\mathbf{\Omega}$ is the body angular velocity, g is the acceleration due to gravity, $\boldsymbol{\chi}$ is the body fixed unit vector on the line segment connecting the fixed point with the body's center of mass, and l is the length of this segment. Also, I is the (time independent) moment of inertia tensor in body coordinates, defined as in the case of the free rigid body. The body angular momentum and the body angular velocity are related, as before, by $\mathbf{\Pi}=I\mathbf{\Omega}$. Also, $\mathbf{\Gamma}=\mathbf{A}^{-1}\mathbf{k}$, which may be thought of as the (negative) direction of gravity as seen from the body, where $\mathbf{k}$ points upward and $\mathbf{A}$ is the element of SO(3) describing the current configuration of the body.

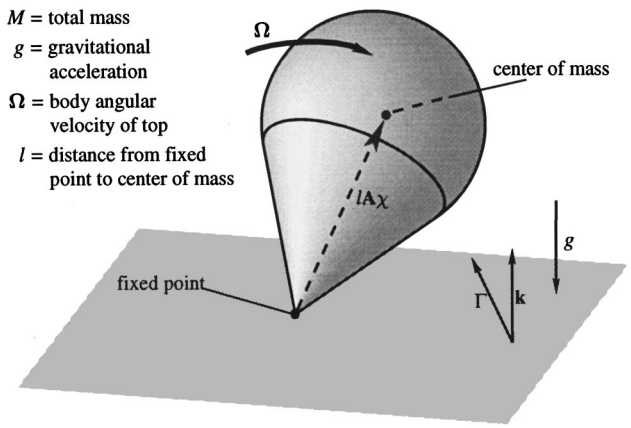

FIG. 1. Heavy top.

For a discussion of the Lie–Poisson nature of these equations on the dual of the Lie algebra $\mathfrak{se}(3)$ of the Euclidean group and for further references, see Marsden and Ratiu [1994].[98] For the Euler–Poincaré point of view, see Holm, Marsden, and Ratiu [1998].[53] These references also discuss this example from the semidirect product point of view, the theory of which we shall present shortly.

Now we discuss the shape space, the momentum map, the locked inertia tensor, and the mechanical connection for this example. We choose $Q=\mathrm{SO}(3)$ and $G=S^1$, regarded as rotations about the spatial z axis, that is, rotations about the axis of gravity.

The shape space is $Q/G=S^2$, the two sphere. Notice that in this case, the bundle $\pi_{Q,G}:\mathrm{SO}(3)\to S^2$ given by $\mathbf{A}\in\mathrm{SO}(3)\mapsto\mathbf{\Gamma}=\mathbf{A}^{-1}\mathbf{k}$ is not a trivial bundle. That is, the angle of rotation ϕ about the z axis is not a *global* cyclic variable. In other words, in this case, Q *cannot* be written as the product $S^2\times S^1$. The classical Routh procedure usually assumes, often implicitly, that the cyclic variables are global.

As with the free rigid body, the heavy top kinetic energy is given by the left invariant metric on $Q=\mathrm{SO}(3)$ whose value at the identity is $\langle\langle\mathbf{\Omega}_1,\mathbf{\Omega}_2\rangle\rangle=I\mathbf{\Omega}_1\cdot\mathbf{\Omega}_2$, where $\mathbf{\Omega}_1,\mathbf{\Omega}_2\in\mathbb{R}^3$ are thought of as elements of $\mathfrak{so}(3)$. This kinetic energy is thus left invariant under the action of the full group SO(3).

The potential energy is given by $Mgl\mathbf{A}^{-1}\mathbf{k}\cdot\boldsymbol{\chi}$. This potential energy is invariant under the group $G=S^1$. As usual, the Lagrangian is the kinetic minus the potential energies.

We next compute the infinitesimal generators for the action of G. We identify the Lie algebra of G with the real line $\mathbb{R}$ and this is identified with the (trivial) subalgebra of $\mathfrak{so}(3)$ by $\xi\mapsto\xi\hat{\mathbf{k}}$. These are given, according to the definitions, by $\xi_{\mathrm{SO}(3)}(\mathbf{A})=\xi\hat{\mathbf{k}}\mathbf{A}\in T_{\mathbf{A}}\mathrm{SO}(3)$.

The locked inertia tensor is, for each $\mathbf{A}\in\mathrm{SO}(3)$, a linear map $\mathfrak{I}(\mathbf{A}):\mathbb{R}\to\mathbb{R}$ which we identify with a real number. According to the definitions, it is given by

$$\mathfrak{I}(\mathbf{A})\xi\eta=\langle\mathfrak{I}(\mathbf{A})\xi,\eta\rangle=\langle\langle\xi_Q(\mathbf{A}),\eta_Q(\mathbf{A})\rangle\rangle=\langle\langle\xi\hat{\mathbf{k}}\mathbf{A},\eta\hat{\mathbf{k}}\mathbf{A}\rangle\rangle.$$

Using the definition of the metric and its left SO(3)-invariance, this equals

$$\langle\langle\xi\hat{\mathbf{k}}\mathbf{A},\eta\hat{\mathbf{k}}\mathbf{A}\rangle\rangle=\xi\eta\langle\langle\mathbf{A}^{-1}\hat{\mathbf{k}}\mathbf{A},\mathbf{A}^{-1}\hat{\mathbf{k}}\mathbf{A}\rangle\rangle=\xi\eta\langle\langle\mathbf{A}^{-1}\mathbf{k},\mathbf{A}^{-1}\mathbf{k}\rangle\rangle=\xi\eta(\mathbf{A}I\mathbf{A}^{-1}\mathbf{k})\cdot\mathbf{k}.$$

Thus, $\mathfrak{I}(\mathbf{A})=(\mathbf{A}I\mathbf{A}^{-1}\mathbf{k})\cdot\mathbf{k}$, that is, the (3,3)-component of the matrix $\mathbf{A}I\mathbf{A}^{-1}$.

Next, we compute the momentum map $\mathbf{J}_L:T\mathrm{SO}(3)\to\mathbb{R}$ for the action of G. According to the general definition, namely, $\langle\mathbf{J}_L(v_q),\xi\rangle=\langle\langle v_q,\xi_Q(q)\rangle\rangle$, we get

$$\langle\mathbf{J}_L(\mathbf{A},\dot{\mathbf{A}}),\xi\rangle=\langle\langle\dot{\mathbf{A}},\xi\hat{\mathbf{k}}\mathbf{A}\rangle\rangle_{\mathbf{A}}=\xi\langle\langle\mathbf{A}^{-1}\dot{\mathbf{A}},\mathbf{A}^{-1}\hat{\mathbf{k}}\mathbf{A}\rangle\rangle_{\mathbf{I}}=\xi\langle\langle\mathbf{\Omega},\mathbf{A}^{-1}\mathbf{k}\rangle\rangle.$$

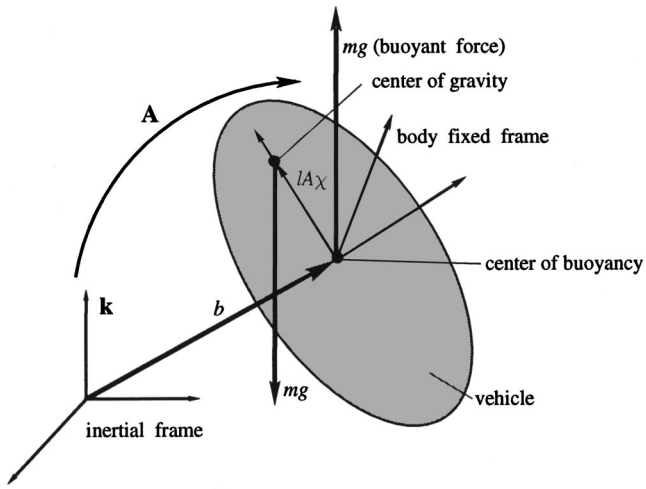

FIG. 2. Schematic of a neutrally buoyant ellipsoidal underwater vehicle.

Using the definition of the metric, we get

$$\xi\langle\langle \mathbf{\Omega}, \mathbf{A}^{-1}\mathbf{k}\rangle\rangle = \xi(I\mathbf{\Omega})\cdot(\mathbf{A}^{-1}\mathbf{k}) = \xi(\mathbf{A}\Pi)\cdot\mathbf{k} = \xi\boldsymbol{\pi}_3,$$

where $\boldsymbol{\pi} = \mathbf{A}\Pi$ is the spatial angular momentum. Thus, $\mathbf{J}_L(\mathbf{A},\dot{\mathbf{A}}) = \boldsymbol{\pi}_3$, the third component of the spatial angular momentum. The mechanical connection $\mathfrak{A}(\mathbf{A}): T_{\mathbf{A}}\mathrm{SO}(3) \to \mathbb{R}$ is given, using the general formula $\mathfrak{A}(v_q) = \mathfrak{I}(q)^{-1}\mathbf{J}_L(v_q)$, by $\mathfrak{A}(\mathbf{A},\dot{\mathbf{A}}) = \boldsymbol{\pi}_3/(\mathbf{A}I\mathbf{A}^{-1}\mathbf{k})\cdot\mathbf{k}$.

Underwater vehicle. The underwater vehicle is modeled as a rigid body moving in ideal potential flow according to Kirchhoff's equations. The vehicle is assumed to be neutrally buoyant (often ellipsoidal), but not necessarily with coincident centers of gravity and buoyancy. The vehicle is free to both rotate and translate in space.

Fix an orthonormal coordinate frame to the body with origin located at the center of buoyancy and axes aligned with the principal axes of the displaced fluid (Fig. 2).

When these axes are also the principal axes of the body and the vehicle is ellipsoidal, the inertia and mass matrices are simultaneously diagonalized. Let the inertia matrix of the body-fluid system be denoted by $I = \mathrm{diag}(I_1, I_2, I_3)$ and the mass matrix by $M = \mathrm{diag}(m_1, m_2, m_3)$; these matrices include the ''added'' inertias and masses due to the fluid. The total mass of the body is denoted m and the acceleration of gravity is g.

The current position of the body is given by a vector $\mathbf{b}$ (the vector from the spatially fixed origin to the center of buoyancy) and its attitude is given by a rotation matrix $\mathbf{A}$ (the center of rotation is the spatial origin). The body fixed vector from the center of buoyancy to the center of gravity is denoted $l\boldsymbol{\chi}$, where l is the distance between these centers.

We shall now formulate the structure of the problem in a form relevant for the present needs, omitting the discussion of how one obtains the equations and the Lagrangian. We refer the reader to Leonard [1997][86] and to Leonard and Marsden [1997][84] for additional details. In particular, these references study the formulation of the equations as Euler–Poincaré and Lie–Poisson equations on a double semidirect product and do a stability analysis.

In this problem, $Q = \mathrm{SE}(3)$, the group of Euclidean motions in space, the symmetry group is $G = \mathrm{SE}(2)\times\mathbb{R}$, and G acts on Q on the left as a subgroup; the symmetries correspond to translation and rotation in a horizontal plane together with vertical translations. Because the centers of gravity and buoyancy are different, rotations around nonvertical axes are not symmetries, as with the heavy top.

The shape space is $Q/G=S^2$, as in the case of the heavy top because the quotient operation removes the translational variables. The bundle $\pi_{Q,G}:\mathrm{SO}(3)\to S^2$ is again given by $\mathbf{A}\in\mathrm{SO}(3)\mapsto\mathbf{\Gamma}=\mathbf{A}^{-1}\mathbf{k}$, where $\mathbf{\Gamma}$ has the same interpretation as it did in the case of the heavy top.

Elements of SE(3) are pairs $(\mathbf{A},\mathbf{b})$ where $\mathbf{A}\in\mathrm{SO}(3)$ and $\mathbf{b}\in\mathbb{R}^3$. If the pair $(\mathbf{A},\mathbf{b})$ is identified with the matrix

$$\begin{bmatrix}\mathbf{A} & \mathbf{b}\\ 0 & 1\end{bmatrix},$$

then, as is well-known, group multiplication in SE(3) is given by matrix multiplication. The Lie algebra of SE(3) is $\mathfrak{se}(3)=\mathbb{R}^3\times\mathbb{R}^3$ with the bracket $[(\mathbf{\Omega},\mathbf{u}),(\mathbf{\Sigma},\mathbf{v})]=(\mathbf{\Omega}\times\mathbf{\Sigma},\mathbf{\Omega}\times\mathbf{v}-\mathbf{\Sigma}\times\mathbf{u})$.

As shown in the cited references, the underwater vehicle kinetic energy is that of the left invariant metric on SE(3) given at the identity as follows:

$$\langle\langle(\mathbf{\Omega}_1,\mathbf{v}_1),(\mathbf{\Omega}_2,\mathbf{v}_2)\rangle\rangle=\mathbf{\Omega}_1\cdot I\mathbf{\Omega}_2+\mathbf{\Omega}_1\cdot D\mathbf{v}_2+\mathbf{v}_1\cdot D^T\mathbf{\Omega}_2+\mathbf{v}_1\cdot M\mathbf{v}_2, \tag{I.12}$$

where $D=m\hat{\boldsymbol{\chi}}$. The kinetic energy is thus the SE(3) invariant function on $T\,\mathrm{SE}(3)$ whose value at the identity is given by $K(\mathbf{\Omega},\mathbf{v})=\frac{1}{2}\mathbf{\Omega}\cdot I\mathbf{\Omega}+\mathbf{\Omega}\cdot D\mathbf{v}+\frac{1}{2}\mathbf{v}\cdot M\mathbf{v}$. The potential energy is given by $V(\mathbf{A},\mathbf{b})=mgl\mathbf{A}^{-1}\mathbf{k}\cdot\boldsymbol{\chi}$ and $L=K-V$.

The momenta conjugate to $\mathbf{\Omega}$ and $\mathbf{v}$ are given by

$$\mathbf{\Pi}=\frac{\partial L}{\partial\mathbf{\Omega}}=I\mathbf{\Omega}+D\mathbf{v},\quad \mathbf{P}=\frac{\partial L}{\partial\mathbf{v}}=M\mathbf{v}+D^T\mathbf{\Omega},$$

the ''angular momentum'' and the ''linear momentum.'' Equivalently, $\mathbf{\Omega}=A\mathbf{\Pi}+B^T\mathbf{P}$ and $\mathbf{v}=C\mathbf{P}+B\mathbf{\Pi}$, where

$$A=(I-DM^{-1}D^T)^{-1},\quad B=-CD^TI^{-1}=-M^{-1}D^TA,\quad C=(M-D^TI^{-1}D)^{-1}.$$

The equations of motion are computed to be

$$\begin{aligned}\dot{\mathbf{\Pi}}&=\mathbf{\Pi}\times\mathbf{\Omega}+\mathbf{P}\times\mathbf{v}-mgl\mathbf{\Gamma}\times\boldsymbol{\chi},\\ \dot{\mathbf{P}}&=\mathbf{P}\times\mathbf{\Omega},\quad \dot{\mathbf{\Gamma}}=\mathbf{\Gamma}\times\mathbf{\Omega},\end{aligned} \tag{I.13}$$

which is the Lie–Poisson (or Euler–Poincaré) form in a double semidirect product.

The Lie algebra of G is $\mathfrak{se}(2)\times\mathbb{R}$, identified with the set of pairs $(\xi,\mathbf{v})$ where $\xi\in\mathbb{R}$ and $\mathbf{v}\in\mathbb{R}^3$ and this is identified with the subalgebra of $\mathfrak{se}(3)$ of elements of the form $(\xi\hat{\mathbf{k}},\mathbf{v})$.

The infinitesimal generators for the action of G are given by

$$(\xi,\mathbf{v})_{\mathrm{SE}(3)}(\mathbf{A},\mathbf{b})=(\xi\hat{\mathbf{k}}\mathbf{A},\xi\mathbf{k}\times\mathbf{b}+\mathbf{v})\in T_{(\mathbf{A},\mathbf{b})}\mathrm{SE}(3).$$

The locked inertia tensor is, for each $(\mathbf{A},\mathbf{b})\in\mathrm{SE}(3)$, a linear map $\mathfrak{I}(\mathbf{A},\mathbf{b}):\mathfrak{so}(2)\times\mathbb{R}\to(\mathfrak{so}(2)\times\mathbb{R})^*$. We identify, as previously, the Lie algebra $\mathfrak{g}$ with pairs $(\xi,\mathbf{v})$ and identify the dual space with the algebra itself using ordinary multiplication and the Euclidean dot product.

According to the definitions, $\mathfrak{I}$ is given by

$$\begin{aligned}\langle\mathfrak{I}(\mathbf{A},\mathbf{b})(\xi,\mathbf{v}),(\eta,\mathbf{w})\rangle&=\langle\langle(\xi,\mathbf{v})_{\mathrm{SE}(3)}(\mathbf{A},\mathbf{b}),(\eta,\mathbf{w})_{\mathrm{SE}(3)}(\mathbf{A},\mathbf{b})\rangle\rangle_{(\mathbf{A},\mathbf{b})}\\ &=\langle\langle(\xi\hat{\mathbf{k}}\mathbf{A},\xi\mathbf{k}\times\mathbf{b}+\mathbf{v}),(\eta\hat{\mathbf{k}}\mathbf{A},\eta\mathbf{k}\times\mathbf{b}+\mathbf{w})\rangle\rangle_{(\mathbf{A},\mathbf{b})}.\end{aligned}$$

The tangent of left translation on the group SE(3) is given by $TL_{(\mathbf{A},\mathbf{b})}(\mathbf{U},\mathbf{w})=(\mathbf{AU},\mathbf{Aw})$. Using the fact that the metric is left SE(3) invariant and formula (I.12) for the inner product, we arrive at

$$\mathfrak{I}(\mathbf{A},\mathbf{b})\cdot(\xi,\mathbf{v})=(\xi(\mathbf{A}I\mathbf{A}^{-1}\mathbf{k})\cdot\mathbf{k}+\xi(\mathbf{A}D\mathbf{A}^{-1}\mathbf{k})\cdot\mathbf{k}$$
$$+[\mathbf{A}D\mathbf{A}^{-1}(\xi\mathbf{k}\times\mathbf{b}+\mathbf{v})]\cdot\mathbf{k}+[\mathbf{A}M\mathbf{A}^{-1}(\xi\mathbf{k}\times\mathbf{b}+\mathbf{v})]\cdot(\mathbf{k}\times\mathbf{b}),$$
$$\mathbf{A}D^T\mathbf{A}^{-1}\mathbf{k}+\mathbf{A}M\mathbf{A}^{-1}(\xi\mathbf{k}\times\mathbf{b}+\mathbf{v})). \tag{I.14}$$

The momentum map $\mathbf{J}_L:T\,\mathrm{SE}(3)\to\mathfrak{se}(2)^*\times\mathbb{R}$ for the action of G is readily computed using the general definition, namely, $\langle\mathbf{J}_L(v_q),\xi\rangle=\langle\langle v_q,\xi_Q(q)\rangle\rangle$; one gets

$$\mathbf{J}_L(\mathbf{A},\mathbf{b},\dot{\mathbf{A}},\dot{\mathbf{b}})=((\mathbf{A}\mathbf{\Pi}+\mathbf{b}\times\mathbf{A}\mathbf{P})\cdot\mathbf{k},\mathbf{A}\mathbf{P}),$$

where, recall, $\mathbf{\Pi}=\partial L/\partial\mathbf{\Omega}=I\mathbf{\Omega}+D\mathbf{v}$ and $\mathbf{P}=\partial L/\partial\mathbf{v}=M\mathbf{v}+D^T\mathbf{\Omega}$.

The mechanical connection $\mathfrak{A}(\mathbf{A},\mathbf{b}):T_{(\mathbf{A},\mathbf{b})}\mathrm{SE}(3)\to\mathfrak{se}(2)^*\times\mathbb{R}$ is therefore given, according to the general formula $\mathfrak{A}(v_q)=\mathfrak{I}(q)^{-1}\mathbf{J}_L(v_q)$, by

$$\mathfrak{A}(\mathbf{A},\mathbf{b},\dot{\mathbf{A}},\dot{\mathbf{b}})=\mathfrak{I}(\mathbf{A},\mathbf{b})^{-1}\cdot((\mathbf{A}\mathbf{\Pi}+\mathbf{b}\times\mathbf{A}\mathbf{P})\cdot\mathbf{k},\mathbf{A}\mathbf{P}),$$

where $\mathfrak{I}(\mathbf{A},\mathbf{b})$ is given by (I.14). We do not attempt to invert the locked inertia tensor explicitly in this case.

II. THE BUNDLE PICTURE IN MECHANICS

A. Cotangent bundle reduction

Cotangent bundle reduction theory lies at the heart of the bundle picture. We will describe it from this point of view in this section.

Some history. We continue the history begun in the Introduction. From the symplectic viewpoint, a principal result is that the *symplectic reduction of a cotangent bundle T^*Q at $\mu\in\mathfrak{g}^*$ is a bundle over $T^*(Q/G)$ with fiber the coadjoint orbit through μ.* This result can be traced back in a preliminary form, to Sternberg [1977][149] and Weinstein [1977].[154] It was refined in Montgomery, Marsden, and Ratiu [1984][115] and in Montgomery [1986].[118] See also Abraham and Marsden [1978][2] and Marsden [1992].[103] It was also shown that the symplectically reduced cotangent bundle can be symplectically embedded in $T^*(Q/G_\mu)$—this is the *injective version of the cotangent bundle reduction theorem*. From the Poisson viewpoint, in which one simply takes quotients by group actions, this reads: $(T^*Q)/G$ is a $\mathfrak{g}^*$-bundle over $T^*(Q/G)$, or a Lie–Poisson bundle over the cotangent bundle of shape space. We shall return to this bundle point of view shortly and sharpen some of these statements.

The bundle point of view. We choose a principal connection $\mathcal{A}$ on the shape space bundle. The general theory, in principle, does not require one to choose a connection. However, there are many good reasons to do so, such as applications to stability theory and geometric phases. Define $\tilde{\mathfrak{g}}=(Q\times\mathfrak{g})/G$, the *associated bundle* to $\mathfrak{g}$, where the quotient uses the given action on Q and the coadjoint action on $\mathfrak{g}$. The connection $\mathcal{A}$ defines a bundle isomorphism $\alpha_{\mathcal{A}}:TQ/G\to T(Q/G)\oplus\tilde{\mathfrak{g}}$ given by $\alpha_{\mathcal{A}}([v_q]_G)=T\pi_{Q,G}(v_q)\oplus[q,\mathcal{A}(v_q)]_G$. Here, the sum is a Whitney sum of vector bundles over Q/G (the fiberwise direct sum) and the symbol $[q,\mathcal{A}(v_q)]_G$ means the equivalence class of $(q,\mathcal{A}(v_q))\in Q\times\mathfrak{g}$ under the G-action. The map $\alpha_{\mathcal{A}}$ is a well-defined vector bundle isomorphism with inverse given by $\alpha_{\mathcal{A}}^{-1}(u_x\oplus[q,\xi]_G)=[(u_x)_q^h+\xi_Q(q)]_G$, where $(u_x)_q^h$ denotes the horizontal lift of u_x to the point q.

Poisson cotangent bundle reduction. The bundle view of Poisson cotangent bundle reduction considers the inverse of the fiberwise dual of $\alpha_{\mathcal{A}}$, which defines a bundle isomorphism $(\alpha_{\mathcal{A}}^{-1})^*:T^*Q/G\to T^*(Q/G)\oplus\tilde{\mathfrak{g}}^*$, where $\tilde{\mathfrak{g}}^*=(Q\times\mathfrak{g}^*)/G$ is the vector bundle over Q/G associated with the coadjoint action of G on $\tilde{\mathfrak{g}}^*$. This isomorphism makes explicit the sense in which $(T^*Q)/G$ is a bundle over $T^*(Q/G)$ with fiber $\mathfrak{g}^*$. The Poisson structure on this bundle is a synthesis of the canonical bracket, the Lie–Poisson bracket, and curvature. The inherited Poisson structure on this space was derived in Montgomery, Marsden, and Ratiu [1984][115] (details were given in Montgomery [1986][118]) and was put into the present context in Cendra, Marsden, and Ratiu [2000a].[28]

Symplectic cotangent bundle reduction. Marsden and Perlmutter [1999][95] show that each symplectic reduced space of T^*Q, which are the symplectic leaves in $(T^*Q)/G \cong T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$, are given by a fiber product $T^*(Q/G) \times_{Q/G} \tilde{\mathcal{O}}$, where $\tilde{\mathcal{O}}$ is the associated coadjoint orbit bundle. This makes precise the sense in which the symplectic reduced spaces are bundles over $T^*(Q/G)$ with fiber a coadjoint orbit. They also give an intrinsic expression for the reduced symplectic form, which involves the canonical symplectic structure on $T^*(Q/G)$, the curvature of the connection, the coadjoint orbit symplectic form, and interaction terms that pair tangent vectors to the orbit with the vertical projections of tangent vectors to the configuration space; see also Zaalani [1999].[160]

As we shall show in the next section, the reduced space P_μ for $P = T^*Q$ is globally diffeomorphic to the bundle $T^*(Q/G) \times_{Q/G} Q/G_\mu$, where Q/G_μ is regarded as a bundle over Q/G. In fact, these results simplify the study of these symplectic leaves. In particular, this makes the injective version of cotangent bundle reduction transparent. Indeed, there is a natural inclusion map $T^*(Q/G) \times_{Q/G} Q/G_\mu \to T^*(Q/G_\mu)$, induced by the dual of the tangent of the projection map $\rho_\mu : Q/G_\mu \to Q/G$. This inclusion map then realizes the reduced spaces P_μ as symplectic subbundles of $T^*(Q/G_\mu)$.

B. Lagrange–Poincaré reduction

In a local trivialization, write $Q = S \times G$ where $S = Q/G$, and TQ/G as $TS \times \mathfrak{g}$. Coordinates on Q are written x^α, s^a and those for $(TQ)/G$ are denoted $(x^\alpha, \dot{x}^\alpha, \xi^a)$. Locally, the connection one-form on Q is written $ds^a + \mathcal{A}^a_\alpha dx^\alpha$ and we let $\Omega^a = \xi^a + \mathcal{A}^a_\alpha \dot{x}^\alpha$. The components of the curvature of $\mathcal{A}$ are

$$B^b_{\alpha\beta} = \left(\frac{\partial A^b_\beta}{\partial x^\alpha} - \frac{\partial A^b_\alpha}{\partial x^\beta} - C^b_{cd} A^c_\alpha A^d_\beta \right),$$

where C^a_{bd} are the structure constants of the Lie algebra $\mathfrak{g}$. Later, in the text, we review the intrinsic definition of curvature.

Let, as explained earlier, $L : TQ \to \mathbb{R}$ be a G-invariant Lagrangian and let $l : (TQ)/G \to \mathbb{R}$ be the corresponding function induced on $(TQ)/G$. The Euler–Lagrange equations on Q induce equations on this quotient space. The connection is used to write these equations intrinsically as a coupled set of Euler–Lagrange type equations and Euler–Poincaré equations. These *reduced Euler–Lagrange* equations, also called the *Lagrange–Poincaré equations* (implicitly contained in Cendra, Ibort, and Marsden [1987][26] and explicitly in Marsden and Scheurle [1993b][108]) are, in coordinates,

$$\frac{d}{dt} \frac{\partial l}{\partial \dot{x}^\alpha} - \frac{\partial l}{\partial x^\alpha} = \frac{\partial l}{\partial \Omega^a} (B^a_{\alpha\beta} \dot{x}^\beta - C^a_{db} A^b_\alpha \Omega^d),$$

$$\frac{d}{dt} \frac{\partial l}{\partial \Omega^b} = \frac{\partial l}{\partial \Omega^a} (C^a_{db} \Omega^d - C^a_{db} A^d_\alpha \dot{x}^\alpha).$$

Using the geometry of the bundle $TQ/G \cong T(Q/G) \oplus \tilde{\mathfrak{g}}$, one can write these equations intrinsically in terms of covariant derivatives (see Cendra, Marsden, and Ratiu [2000a][28]). Namely, they take the form

$$\frac{\partial l}{\partial x}(x, \dot{x}, \overline{v}) - \frac{D}{Dt} \frac{\partial l}{\partial \dot{x}}(x, \dot{x}, \overline{v}) = \left\langle \frac{\partial l}{\partial \overline{v}}(x, \dot{x}, \overline{v}), \mathbf{i}_{\dot{x}} \mathrm{Curv}_{\mathcal{A}}(x) \right\rangle,$$

$$\frac{D}{Dt} \frac{\partial l}{\partial \overline{v}}(x, \dot{x}, \overline{v}) = \mathrm{ad}^*_v \frac{\partial l}{\partial \overline{v}}(x, \dot{x}, \overline{v}).$$

The first of these equations is the *horizontal Lagrange–Poincaré equation* while the second is the *vertical Lagrange–Poincaré equation*. The notation here is as follows. Points in $T(Q/G) \oplus \tilde{\mathfrak{g}}$ are

denoted $(x,\dot{x},\overline{v})$ and $l(x,\dot{x},\overline{v})$ denotes the Lagrangian induced on the quotient space from L. The bundles $T(Q/G)\oplus\tilde{\mathfrak{g}}$ naturally inherit vector bundle connections and D/Dt denotes the associated covariant derivatives. Also, $\mathrm{Curv}_{\mathcal{A}}$ denotes the curvature of the connection $\mathcal{A}$ thought of as an adjoint bundle valued two-form on Q/G—basic definitions and properties of curvature will be reviewed shortly.

Lagrangian reduction by stages. The perspective developed in Cendra, Marsden, and Ratiu [2000a][28] is motivated by *reduction by stages*. In fact, that work develops a context (of Lagrange–Poincaré bundles) in which Lagrangian reduction can be repeated. In particular, this theory treats successive reduction for group extensions. Reduction for group extensions, in turn, builds on semidirect product reduction theory, to which we turn next.

C. Hamiltonian semidirect product theory

Lie–Poisson systems on semidirect products. The study of Lie–Poisson equations for systems on the dual of a semidirect product Lie algebra grew out of the work of many authors including Sudarshan and Mukunda [1974],[150] Vinogradov and Kupershmidt [1977],[153] Ratiu [1980, 1981, 1982],[135,138,139] Guillemin and Sternberg [1980],[44] Marsden [1982],[102] Marsden *et al.* [1983],[113] Holm and Kupershmidt [1983],[49,50] Kupershmidt and Ratiu [1983],[79] Holmes and Marsden [1983],[51] Marsden, Ratiu, and Weinstein [1984a,b],[100,101] Guillemin and Sternberg [1984],[46] Holm *et al.* [1985],[56] Abarbanel *et al.* [1986],[1] Leonard and Marsden [1997],[84] and Marsden *et al.* [1998].[90] As these and related references show, the Lie–Poisson equations apply to a surprisingly wide variety of systems such as the heavy top, compressible flow, stratified incompressible flow, MHD (magnetohydrodynamics), and underwater vehicle dynamics.

In each of the above-mentioned examples as well as in the general theory, one can view the given Hamiltonian in the material representation as a function depending on a parameter; this parameter becomes a dynamic variable when reduction is performed. For example, in the heavy top, the direction and magnitude of gravity, the mass and location of the center of mass may be regarded as parameters, but the direction of gravity becomes the dynamic variable $\mathbf{\Gamma}$ when reduction is performed.

We first recall how the Hamiltonian theory proceeds for systems defined on semidirect products. We present the abstract theory, but of course historically this grew out of the examples, especially the heavy top and compressible flow. When working with various models of continuum mechanics and plasmas one has to keep in mind that many of the actions are right actions, so one has to be careful when employing general theorems involving left actions. We refer to Holm, Marsden, and Ratiu [1998][53] for a statement of some of the results explicitly for right actions.

Generalities on semidirect products. Let V be a vector space and assume that the Lie group G acts *on the left* by linear maps on V (and hence G also acts on the left on its dual space V^*). The semidirect product $S=G\circledS V$ is the set $S=G\times V$ with group multiplication given by

$$(g_1,v_1)(g_2,v_2)=(g_1g_2,v_1+g_1v_2),$$

where the action of $g\in G$ on $v\in V$ is denoted gv. The identity element is $(e,0)$ where e is the identity in G and the inverse of (g,v) is $(g,v)^{-1}=(g^{-1},-g^{-1}v)$. The Lie algebra of S is the semidirect product Lie algebra, $\mathfrak{s}=\mathfrak{g}\circledS V$, whose bracket is

$$[(\xi_1,v_1),(\xi_2,v_2)]=([\xi_1,\xi_2],\xi_1v_2-\xi_2v_1),$$

where we denote the induced action of $\mathfrak{g}$ on V by ξ_1v_2.

The adjoint and coadjoint actions are given by

$$(g,v)(\xi,u)=(g\xi,gu-(g\xi)v),\quad (g,v)(\mu,a)=(g\mu+\rho_v^*(ga),ga),$$

where $(g,v)\in S=G\times V$, $(\xi,u)\in\mathfrak{s}=\mathfrak{g}\times V$, $(\mu,a)\in\mathfrak{s}^*=\mathfrak{g}^*\times V^*$, $g\xi=\mathrm{Ad}_g\,\xi$, $g\mu=\mathrm{Ad}^*_{g^{-1}}\,\mu$, ga denotes the induced *left* action of g on a (the *left* action of G on V induces a *left* action of G on

V^*—the inverse of the transpose of the action on V), $\rho_v:\mathfrak{g}\to V$ is the linear map given by $\rho_v(\xi)=\xi v$, and $\rho_v^*:V^*\to\mathfrak{g}^*$ is its dual. For $a\in V^*$, we write $\rho_v^* a = v\diamond a\in\mathfrak{g}^*$, which is a bilinear operation in v and a. Equivalently, we can write $\langle \eta a, v\rangle = -\langle v\diamond a,\eta\rangle$. Using this notation, the coadjoint action reads $(g,v)(\mu,a)=(g\mu+v\diamond(ga),ga)$.

Lie–Poisson brackets and Hamiltonian vector fields. For a *left* representation of G on V the $\pm$Lie–Poisson bracket of two functions $f,k:\mathfrak{s}^*\to\mathbb{R}$ is given by

$$\{f,k\}_\pm(\mu,a)=\pm\left\langle \mu,\left[\frac{\delta f}{\delta\mu},\frac{\delta k}{\delta\mu}\right]\right\rangle \pm\left\langle a,\frac{\delta f}{\delta\mu}\frac{\delta k}{\delta a}-\frac{\delta k}{\delta\mu}\frac{\delta f}{\delta a}\right\rangle,$$

where $\delta f/\delta\mu\in\mathfrak{g}$, and $\delta f/\delta a\in V$ are the functional derivatives of f. The Hamiltonian vector field of $h:\mathfrak{s}^*\to\mathbb{R}$ has the expression

$$X_h(\mu,a)=\mp\left(\mathrm{ad}^*_{\delta h/\delta\mu}\,\mu-\frac{\delta h}{\delta a}\diamond a,-\frac{\delta h}{\delta\mu}a\right).$$

Thus, Hamilton's equations on the dual of a semidirect product are given by

$$\dot\mu=\mp\,\mathrm{ad}^*_{\delta h/\delta\mu}\,\mu\pm\frac{\delta h}{\delta a}\diamond a,$$

$$\dot a=\pm\frac{\delta h}{\delta\mu}a.$$

Symplectic actions by semidirect products. Consider a *left* symplectic action of S on a symplectic manifold P that has an equivariant momentum map $\mathbf{J}_S:P\to\mathfrak{s}^*$. Since V is a (normal) subgroup of S, it also acts on P and has a momentum map $\mathbf{J}_V:P\to V^*$ given by $\mathbf{J}_V=i_V^*\circ\mathbf{J}_S$, where $i_V:V\to\mathfrak{s}$ is the inclusion $v\mapsto(0,v)$ and $i_V^*:\mathfrak{s}^*\to V^*$ is its dual. We think of $\mathbf{J}_V$ as the second component of $\mathbf{J}_S$. We can regard G as a subgroup of S by $g\mapsto(g,0)$. Thus, G also has a momentum map that is the first component of $\mathbf{J}_S$ but this will play a secondary role in what follows. Equivariance of $\mathbf{J}_S$ under G implies that $\mathbf{J}_V(gz)=g\mathbf{J}_V(z)$. To prove this relation, one uses the fact that for the coadjoint action of S on $\mathfrak{s}^*$ the second component is the dual of the given action of G on V.

The classical semidirect product reduction theorem. In a number of interesting applications such as compressible fluids, the heavy top, MHD, etc., one has two symmetry groups that do not commute and thus the *commuting reduction by stages theorem* of Marsden and Weinstein [1974][110] does not apply. In this more general situation, it matters in what order one performs the reduction, which occurs, in particular for semidirect products. The main result covering the case of semidirect products has a complicated history, with important early contributions by many authors, as we have listed previously. The final version of the theorem as we shall use it, is due to Marsden, Ratiu, and Weinstein [1984a,b].[100,101]

Theorem II.1 (Semidirect product reduction theorem): *Let $S=G\circledS V$, choose $\sigma=(\mu,a)\in\mathfrak{g}^*\times V^*$, and reduce T^*S by the action of S at σ giving the coadjoint orbit $\mathcal{O}_\sigma$ through $\sigma\in\mathfrak{s}^*$. There is a symplectic diffeomorphism between $\mathcal{O}_\sigma$ and the reduced space obtained by reducing T^*G by the subgroup G_a (the isotropy of G for its action on V^* at the point $a\in V^*$) at the point $\mu|\mathfrak{g}_a$ where $\mathfrak{g}_a$ is the Lie algebra of G_a.*

This theorem is a consequence of a more general result given in the next section.

D. Semidirect product reduction by stages

A theorem on reduction by stages for semidirect products acting on a symplectic manifold is due to Leonard and Marsden [1997][84] (where the motivation was the application to underwater vehicle dynamics) and Marsden *et al.* [1998].[91]

Consider a symplectic action of S on a symplectic manifold P that has an equivariant momentum map $\mathbf{J}_S:P\to\mathfrak{s}^*$. As we have explained, the momentum map for the action of V is the map $\mathbf{J}_V:P\to V^*$ given by $\mathbf{J}_V=i_V^*\circ\mathbf{J}_S$.

We carry out the reduction of P by S at a regular value $\sigma=(\mu,a)$ of the momentum map $\mathbf{J}_S$ for S in two stages. First, reduce P by V at the value a (assume it to be a regular value) to get the reduced space $P_a=\mathbf{J}_V^{-1}(a)/V$. Second, form the isotropy group G_a of $a\in V^*$. One shows (this step is not trivial) that the group G_a acts on P_a and has an induced equivariant momentum map $\mathbf{J}_a:P_a\to\mathfrak{g}_a^*$, where $\mathfrak{g}_a$ is the Lie algebra of G_a, so one can reduce P_a at the point $\mu_a:=\mu|\mathfrak{g}_a$ to get the reduced space $(P_a)_{\mu_a}=\mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

Theorem II.2 (reduction by stages for semidirect products): *The reduced space $(P_a)_{\mu_a}$ is symplectically diffeomorphic to the reduced space P_σ obtained by reducing P by S at the point $\sigma=(\mu,a)$.*

Combined with the cotangent bundle reduction theorem, the semidirect product reduction theorem is a useful tool. For example, this shows that the generic coadjoint orbits for the Euclidean group are cotangent bundles of spheres with the associated coadjoint orbit symplectic structure given by the canonical structure plus a magnetic term.

Semidirect product reduction of dynamics. There is a technique for reducing dynamics that is associated with the geometry of the semidirect product reduction theorem. One proceeds as follows.

We start with a Hamiltonian H_{a_0} on T^*G that depends parametrically on a variable $a_0\in V^*$. The Hamiltonian, regarded as a map $H:T^*G\times V^*\to\mathbb{R}$, is assumed to be invariant on $T^*G\times V^*$ under the action of G on $T^*G\times V^*$. One shows that this condition is equivalent to the invariance of the function H defined on $T^*S=T^*G\times V\times V^*$ extended to be constant in the variable V under the action of the semidirect product. By the semidirect product reduction theorem, the dynamics of H_{a_0} reduced by G_{a_0}, the isotropy group of a_0, is symplectically equivalent to Lie–Poisson dynamics on $\mathfrak{s}^*=\mathfrak{g}^*\times V^*$. The Lie–Poisson structure determines the reduced dynamics (given explicitly above) using the function $h(\mu,a)=H(\alpha_g,g^{-1}a)$ where $\mu=g^{-1}\alpha_g$.

E. Lagrangian semidirect product theory

Lagrangian semidirect product reduction is modeled after the reduction theorem for the basic Euler–Poincaré equations, although they are *not* literally special cases of it. To distinguish these, we use phrases like *basic* Euler–Poincaré equations for Eq. (I.6) and simply the Euler–Poincaré equations *with advection* or the Euler–Poincaré equations *with advected parameters*, for the equations that follow.

The main difference between the invariant Lagrangians considered in the Euler–Poincaré reduction theorem earlier and the ones we work with now is that L and l depend on an additional parameter $a\in V^*$, where V is a representation space for the Lie group G and L has an invariance property relative to both arguments.

The parameter $a\in V^*$ acquires dynamical meaning under Lagrangian reduction as it did for the Hamiltonian case: $\dot{a}=\pm(\delta h/\delta\mu)a$. For the heavy top, the parameter is the unit vector $\mathbf{\Gamma}$ in the (negative) direction of gravity, which becomes a dynamical variable in body representation. For compressible fluids, a becomes the density of the fluid in spatial representation, which becomes a dynamical variable (satisfying the continuity equation).

The basic ingredients are as follows. There is a *left* representation of the Lie group G on the vector space V and G acts in the natural way on the *left* on $TG\times V^*$: $h(v_g,a)=(hv_g,ha)$. Assume that the function $L:TG\times V^*\to\mathbb{R}$ is left G-invariant. In particular, if $a_0\in V^*$, define the Lagrangian $L_{a_0}:TG\to\mathbb{R}$ by $L_{a_0}(v_g)=L(v_g,a_0)$. Then L_{a_0} is left invariant under the lift to TG of the left action of G_{a_0} on G, where G_{a_0} is the isotropy group of a_0. Left G-invariance of L permits us to define $l:\mathfrak{g}\times V^*\to\mathbb{R}$ by $l(g^{-1}v_g,g^{-1}a_0)=L(v_g,a_0)$. Conversely, this relation defines for any $l:\mathfrak{g}\times V^*\to\mathbb{R}$ a left G-invariant function $L:TG\times V^*\to\mathbb{R}$. For a curve $g(t)\in G$, let $\xi(t):=g(t)^{-1}\dot{g}(t)$ and define the curve $a(t)$ as the unique solution of the following linear differential equation with time-dependent coefficients $\dot{a}(t)=-\xi(t)a(t)$, with initial condition $a(0)=a_0$. The

solution can be written as $a(t)=g(t)^{-1}a_0$.

Theorem II.3: *With the preceding notation, the following are equivalent:*

(i) *With a_0 held fixed, Hamilton's variational principle*

$$\delta\int_{t_1}^{t_2} L_{a_0}(g(t),\dot{g}(t))dt=0 \tag{II.1}$$

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

(ii) *$g(t)$ satisfies the Euler–Lagrange equations for L_{a_0} on G.*

(iii) *The constrained variational principle,*

$$\delta\int_{t_1}^{t_2} l(\xi(t),a(t))dt=0, \tag{II.2}$$

holds on $\mathfrak{g}\times V^*$, using variations of ξ and a of the form $\delta\xi=\dot{\eta}+[\xi,\eta]$ and $\delta a=-\eta a$, where $\eta(t)\in\mathfrak{g}$ vanishes at the end points.

(iv) *The Euler–Poincaré equations hold on $\mathfrak{g}\times V^*$,*

$$\frac{d}{dt}\frac{\delta l}{\delta\xi}=\mathrm{ad}^*_\xi\frac{\delta l}{\delta\xi}+\frac{\delta l}{\delta a}\diamond a. \tag{II.3}$$

Remarks:

(1) As with the basic Euler–Poincaré equations, this is not strictly a variational principle in the same sense as the standard Hamilton's principle. It is more of a Lagrange–d'Alembert principle, because we impose the stated constraints on the variations allowed.

(2) Note that Eq. (II.3) is not the basic Euler–Poincaré equations because we are not regarding $\mathfrak{g}\times V^*$ as a Lie algebra. Rather, these equations are thought of as a generalization of the classical Euler–Poisson equations for a heavy top, written in body angular velocity variables, as we shall see in the examples. Some authors may prefer the term Euler–Poisson–Poincaré equations for these equations.

We refer to Holm, Marsden, and Ratiu [1998][53] for the proof. It is noteworthy that these Euler–Poincaré equations (II.3) are not the (pure) Euler–Poincaré equations for the semidirect product Lie algebra $\mathfrak{g}\circledS V^*$

The Legendre transformation. Start with a Lagrangian on $\mathfrak{g}\times V^*$ and perform a *partial Legendre transformation* in the variable ξ only, by writing

$$\mu=\frac{\delta l}{\delta\xi},\quad h(\mu,a)=\langle\mu,\xi\rangle-l(\xi,a).$$

Since

$$\frac{\delta h}{\delta\mu}=\xi+\left\langle\mu,\frac{\delta\xi}{\delta\mu}\right\rangle-\left\langle\frac{\delta l}{\delta\xi},\frac{\delta\xi}{\delta\mu}\right\rangle=\xi,$$

and $\delta h/\delta a=-\delta l/\delta a$, we see that (II.3) and $\dot{a}(t)=-\xi(t)a(t)$ imply the Lie–Poisson dynamics on a semidirect product for the *minus* Lie–Poisson bracket. If this Legendre transformation is invertible, then we can also pass from the minus Lie–Poisson equations to the Euler–Poincaré equations (II.3) together with the equations $\dot{a}(t)=-\xi(t)a(t)$.

Relation with Lagrangian reduction. The Euler–Poincaré equations are shown to be a special case of the reduced Euler–Lagrange equations in Cendra *et al.* [1998].[25] We also refer to Cendra *et al.* [1998],[25] who study the Euler–Poincaré formulation of the Maxwell–Vlasov equations for plasma physics.

The Kelvin–Noether theorem. There is a version of the Noether theorem that holds for solutions of the Euler–Poincaré equations. Our formulation is motivated by and designed for ideal continuum theories (and hence the name Kelvin–Noether), but it may be also of interest for finite dimensional mechanical systems. Of course it is well known (going back at least to Arnold [1966][8]) that the Kelvin circulation theorem for ideal flow is closely related to the Noether theorem applied to continua using the particle relabeling symmetry group.

Start with a Lagrangian L_{a_0} depending on a parameter $a_0 \in V^*$ as above and introduce a manifold $\mathcal{C}$ on which G acts (we assume this is also a left action) and suppose we have an equivalent map $\mathcal{K}:\mathcal{C}\times V^* \to \mathfrak{g}^{**}$. In the case of continuum theories, the space $\mathcal{C}$ is usually a loop space and $\langle \mathcal{K}(c,a),\mu\rangle$ for $c\in\mathcal{C}$ and $\mu\in\mathfrak{g}^*$ will be a circulation. This class of examples also shows why we *do not* want to identify the double dual $\mathfrak{g}^{**}$ with $\mathfrak{g}$.

Define the *Kelvin–Noether quantity* $I:\mathcal{C}\times\mathfrak{g}V^* \to \mathbb{R}$ by

$$I(c,\xi,a)=\left\langle \mathcal{K}(c,a),\frac{\delta l}{\delta \xi}\right\rangle. \tag{II.4}$$

Theorem II.4 (Kelvin–Noether): *Fixing $c_0\in\mathcal{C}$, let $\xi(t)$, $a(t)$ satisfy the Euler–Poincaré equations and define $g(t)$ to be the solution of $\dot{g}(t)=g(t)\xi(t)$ and, say, $g(0)=e$. Let $c(t)=g(t)^{-1}c_0$ and $I(t)=I(c(t),\xi(t),a(t))$. Then*

$$\frac{d}{dt}I(t)=\left\langle \mathcal{K}(c(t),a(t)),\frac{\delta l}{\delta a}\diamond a\right\rangle. \tag{II.5}$$

Again, we refer to Holm, Marsden, and Ratiu [1998][53] for the proof.

Corollary II.5: For the basic Euler–Poincaré equations, the Kelvin quantity $I(t)$, defined the same way as above but with $I:\mathcal{C}\times\mathfrak{g}\to\mathbb{R}$, is conserved.

The heavy top. As we explained earlier, the heavy top kinetic energy is given by the left invariant metric on SO(3) whose value at the identity is $\langle \mathbf{\Omega}_1,\mathbf{\Omega}_2\rangle = I\mathbf{\Omega}_1\cdot\mathbf{\Omega}_2$, where $\mathbf{\Omega}_1,\mathbf{\Omega}_2\in\mathbb{R}^3$ are thought of as elements of $\mathfrak{so}(3)$, the Lie algebra of SO(3), via the isomorphism $\mathbf{\Omega}\in\mathbb{R}^3\to\hat{\mathbf{\Omega}}\in\mathfrak{so}(3)$, $\hat{\mathbf{\Omega}}\mathbf{v}:=\mathbf{\Omega}\times\mathbf{v}$.

This kinetic energy is thus left invariant under SO(3). The potential energy is given by $Mgl\mathbf{A}^{-1}\mathbf{k}\cdot\boldsymbol{\chi}$. This potential energy breaks the full SO(3) symmetry and is invariant only under the rotations S^1 about the $\mathbf{k}$ axis.

For the application of Theorem II.3 we think of the Lagrangian of the heavy top as a function on $T\,\mathrm{SO}(3)\times\mathbb{R}^3\to\mathbb{R}$. Define $U(u_{\mathbf{A}},\mathbf{v})=Mgl\mathbf{A}^{-1}\mathbf{v}\cdot\chi$ which is verified to be SO(3)-invariant, so the hypotheses of Theorem II.3 are satisfied. Thus, the heavy top equations of motion in the body representation are given by the Euler–Poincaré equations (II.3) for the Lagrangian $l:\mathfrak{so}(3)\times\mathbb{R}^3\to\mathbb{R}$. defined by $l(\mathbf{\Omega},\mathbf{\Gamma})=L(\mathbf{A}^{-1}u_{\mathbf{A}},\mathbf{A}^{-1}v)=\frac{1}{2}\mathbf{\Pi}\cdot\mathbf{\Omega}-U(\mathbf{A}^{-1}u_{\mathbf{A}},\mathbf{A}^{-1}\mathbf{v})=\frac{1}{2}\mathbf{\Pi}\cdot\mathbf{\Omega}-Mgl\mathbf{\Gamma}\cdot\boldsymbol{\chi}$. It is then straightforward to compute the Euler–Poincaré equations for this reduced Lagrangian and to verify that one gets the usual heavy top equations.

Let $\mathcal{C}=\mathfrak{g}$ and let $\mathcal{K}:\mathcal{C}\times V^*\to\mathfrak{g}^{**}\cong\mathfrak{g}$ be the map $(\mathbf{W},\mathbf{\Gamma})\mapsto\mathbf{W}$. Then the Kelvin–Noether theorem gives the statement $(d/dt)\langle\mathbf{W},\mathbf{\Pi}\rangle=Mgl\langle\mathbf{W},\Gamma\times\boldsymbol{\chi}\rangle$, where $\mathbf{W}(t)=\mathbf{A}(t)^{-1}\mathbf{w}$; in other words, $\mathbf{W}(t)$ is the body representation of a space fixed vector. This statement is easily verified directly. Also, note that $\langle\mathbf{W},\mathbf{\Pi}\rangle=\langle\mathbf{w},\boldsymbol{\pi}\rangle$, with $\boldsymbol{\pi}=\mathbf{A}(t)\mathbf{\Pi}$, so the Kelvin–Noether theorem may be viewed as a statement about the rate of change of the momentum map of the system (the spatial angular momentum) relative to the full group of rotations, not just those about the vertical axis.

F. Reduction by stages

Poisson reduction by stages. Suppose that a Lie group M acts symplectically on a symplectic manifold P. Let N be a normal subgroup of M (so M is an *extension* of N). The problem is to carry out a reduction of P by M in two steps, first a reduction of P by N followed by, roughly speaking, a reduction by the quotient group M/N. On a Poisson level, this is elementary: P/M is Poisson

diffeomorphic to $(P/N)/(M/N)$. However, symplectic reduction is a much deeper question.

Symplectic reduction by stages. We now state the theorem on symplectic reduction by stages regarded as a generalization of the semidirect product reduction theorem. We refer to Marsden *et al.* [1998, 2000][90,91] and Leonard and Marsden [1997][84] for details and applications.

Start with a symplectic manifold (P,Ω) and a Lie group M that acts on P and has an Ad^*-equivariant momentum map $\mathbf{J}_M:P\to\mathfrak{m}^*$, where $\mathfrak{m}$ is the Lie algebra of M. We shall denote this action by $\Phi:M\times P\to P$ and the mapping associated with a group element $m\in M$ by $\Phi_m:P\to P$.

Assume that N is a normal subgroup of M and denote its Lie algebra by $\mathfrak{n}$. Let $i:\mathfrak{n}\to\mathfrak{m}$ denote the inclusion and let $i^*:\mathfrak{m}^*\to\mathfrak{n}^*$ be its dual, which is the natural projection given by restriction of linear functionals. The equivariant momentum map for the action of the group N on P is given by $\mathbf{J}_N(z)=i^*(\mathbf{J}_M(z))$. Let $\nu\in\mathfrak{n}^*$ be a regular value of $\mathbf{J}_N$ and let N_ν be the isotropy subgroup of ν for the coadjoint action of N on its Lie algebra. We suppose that the action of N_ν (and in fact that of M) is free and proper and form the *first symplectic reduced space*: $P_\nu=\mathbf{J}_N^{-1}(\nu)/N_\nu$.

Since N is a normal subgroup, the adjoint action of M on its Lie algebra $\mathfrak{m}$ leaves the subalgebra $\mathfrak{n}$ invariant, and so it induces a dual action of M on $\mathfrak{n}^*$. Thus, we can consider M_ν, the isotropy subgroup of $\nu\in\mathfrak{n}^*$ for the action of M on $\mathfrak{n}^*$. One checks that the subgroup $N_\nu\subset M$ is normal in M_ν, so we can form the quotient group M_ν/N_ν. In the context of semidirect products, with the second factor being a vector space V, M_ν/N_ν reduces to G_a where $\nu=a$ in our semidirect product notation.

Now one shows that there is a well-defined symplectic action of M_ν/N_ν on the reduced space P_ν. In fact, there is a natural sense in which the momentum map $\mathbf{J}_M:P\to\mathfrak{m}^*$ induces a momentum map $\mathbf{J}_\nu:P_\nu\to(\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ for this action. However, this momentum map in general *need not* be equivariant.

However, nonequivariant reduction is a well-defined process and so P_ν can be further reduced by the action of M_ν/N_ν at a regular value $\rho\in(\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$. Let this *second reduced space* be denoted by $P_{\nu,\rho}=\mathbf{J}_{M_\nu/N_\nu}^{-1}(\rho)/(M_\nu/N_\nu)_\rho$ where, as usual, $(M_\nu/N_\nu)_\rho$ is the isotropy subgroup for the action of the group M_ν/N_ν on the dual of its Lie algebra.

Assume that $\sigma\in\mathfrak{m}^*$ is a given regular element of $\mathbf{J}_M$ so that we can form the reduced space $P_\sigma=\mathbf{J}_M^{-1}(\sigma)/M_\sigma$ where M_σ is the isotropy subgroup of σ for the action of M on $\mathfrak{m}^*$. We also require that the relation $(r'_\nu)^*(\rho)=k_\nu^*\sigma-\overline{\nu}$ holds where $r'_\nu:\mathfrak{m}_\nu\to\mathfrak{m}_\nu/\mathfrak{n}_\nu$ is the quotient map, $k_\nu:\mathfrak{m}_\nu\to\mathfrak{m}$ is the inclusion, and $\overline{\nu}$ is some extension of ν to $\mathfrak{m}_\nu$. We assume that the following condition holds.

Stages hypothesis: *For all $\sigma_1,\sigma_2\in\mathfrak{m}^*$ such that $\sigma_1|\mathfrak{m}_\nu=\sigma_2|\mathfrak{m}_\nu$ and $\sigma_1|\mathfrak{n}=\sigma_2|\mathfrak{n}$, there exists $n\in N_\nu$ such that $\sigma_2=\mathrm{Ad}^*_{n^{-1}}\,\sigma_1$.*

Theorem II.6 (symplectic reduction by stages): *Under the above hypotheses, there is a symplectic diffeomorphism between P_σ and $P_{\nu,\rho}$.*

Lagrangian stages. We will just make some comments on the Lagrangian counterpart to Hamiltonian reduction by stages. First of all, it should be viewed as a Lagrangian counterpart to Poisson reduction by stages, which, as we have remarked, is relatively straightforward. What makes the Lagrangian counterpart more difficult is the *a priori* lack of a convenient category, like that of Poisson manifolds, which is stable under reduction. Such a category, which may be viewed as the minimal category satisfying this property and containing tangent bundles, is given in Cendra, Marsden, and Ratiu [2000a].[28] This category must, as we have seen, contain bundles of the form $T(Q/G)\oplus\tilde{\mathfrak{g}}$. This gives a clue as to the structure of the general element of this *Lagrange–Poincaré category*, namely direct sums of tangent bundles with vector bundles with fiberwise Lie algebra structure and certain other (curvature-like) structures. In particular, this theory can handle the case of general group extensions and includes Lagrangian semidirect product reduction as a special case.

The Lagrangian analog of symplectic reduction is non-Abelian Routh reduction to which we turn next. Developing Routh reduction by stages is an interesting and challenging open problem.

III. ROUTH REDUCTION

Routh reduction differs from Lagrange–Poincaré reduction in that the momentum map constraint $\mathbf{J}_L=\mu$ is imposed. Routh dealt with systems having cyclic variables. The heavy top has an Abelian group of symmetries, with a free and proper action, yet it does not have global cyclic variables in the sense that the bundle $Q\to Q/G$ is not trivial; that is, Q is not globally a product $S\times G$. For a modern exposition of Routh reduction in the case when $Q=S\times G$ and G is Abelian, see Marsden and Ratiu [1994],[98] Sec. 8.9, and Arnold [1988].[11]

We shall now embark on a global intrinsic presentation of non-Abelian Routh reduction. Preliminary versions of this theory, which represent our starting point are given in Marsden and Scheurle [1993a][107] and Jalnapurkar and Marsden [2000a].[64]

A. The global realization theorem for the reduced phase space

Let G_μ denote the isotropy subgroup of μ for the coadjoint action of G on $\mathfrak{g}^*$. Because G acts freely and properly on Q and assuming that μ is a regular value of the momentum map $\mathbf{J}_L$, the space $\mathbf{J}_L^{-1}(\mu)/G_\mu$ is a smooth symplectic manifold (by the symplectic reduction theorem). The symplectic structure is not of immediate concern to us.

Fiber products. Given two fiber bundles $f:M\to B$ and $g:N\to B$, the *fiber product* is $M\times_B N=\{(m,n)\in M\times N|f(m)=g(n)\}$. Using the fact that $M\times_B N=(f\times g)^{-1}(\Delta)$ where Δ is the diagonal in $B\times B$, one sees that $M\times_B N$ is a smooth submanifold of $M\times N$ and a smooth fiber bundle over B with the projection map $(m,n)\mapsto f(m)=g(n)$.

Statement of the global realization theorem. Consider the two fiber bundles $\tau_{Q/G}:T(Q/G)\to Q/G$ and $\rho_\mu:Q/G_\mu\to Q/G$. The first is the tangent bundle of shape space, while the second is the map taking an equivalence class with respect to the G_μ group action and mapping it to the larger class (orbit) for the G action on Q. We write the map ρ_μ as $[q]_{G_\mu}\mapsto[q]_G$. The map ρ_μ is smooth being the quotient map induced by the identity. We form the fiber product bundle $p_\mu:T(Q/G)\times_{Q/G}Q/G_\mu\to Q/G$.

A couple of remarks about the bundle structures are in order. The fibers of the bundle $\rho_\mu:Q/G_\mu\to Q/G$ are diffeomorphic to the coadjoint orbit $\mathcal{O}_\mu$ through μ for the G action on $\mathfrak{g}^*$, that is, to the homogeneous quotient space G/G_μ. Also, the space $\mathbf{J}_L^{-1}(\mu)/G_\mu$ is a bundle over both Q/G_μ and Q/G. Namely, we have the smooth maps

$$\sigma^\mu:\mathbf{J}_L^{-1}(\mu)/G_\mu\to Q/G_\mu,\quad [v_q]_{G_\mu}\mapsto[q]_{G_\mu},$$

$$\sigma_\mu:\mathbf{J}_L^{-1}(\mu)/G_\mu\to Q/G,\quad [v_q]_{G_\mu}\mapsto[q]_G.$$

Theorem III.1: *The bundle $\sigma_\mu:\mathbf{J}_L^{-1}(\mu)/G_\mu\to Q/G$ is bundle isomorphic (over the identity) to the bundle $p_\mu:T(Q/G)\times_{Q/G}Q/G_\mu\to Q/G$.*

The maps involved in this theorem and defined in the proof are shown in Fig. 3.

Proof: We first define a bundle map and then check it is a bundle isomorphism by producing an inverse bundle map. We already have defined a map σ^μ that will give the second component of our desired map. To define the first component, we start with the map $T\pi_{Q,G}|\mathbf{J}_L^{-1}(\mu):\mathbf{J}_L^{-1}(\mu)\to T(Q/G)$. This map is readily checked to be G_μ-invariant and so it defines a map of the quotient space $r_\mu:\mathbf{J}_L^{-1}(\mu)/G_\mu\to T(Q/G)$, a bundle map over the base Q/G. The map r_μ is smooth as it is induced by the smooth map $T\pi_{Q,G}|\mathbf{J}_L^{-1}(\mu)$.

The map we claim is a bundle isomorphism is the fiber product $\phi_\mu=r_\mu\times_{Q/G}\sigma^\mu$. This map is smooth as it is the fiber product of smooth maps. Concretely, this bundle map is given as follows. Let $v_q\in\mathbf{J}_L^{-1}(\mu)$. Then $\phi_\mu([v_q]_{G_\mu})=(T_q\pi_{Q,G}(v_q),[q]_{G_\mu})$.

We now construct the inverse bundle map. From the theory of quotient manifolds, recall that one identifies the tangent space $T_x(Q/G)$ at a point $x=[q]_G$ with the quotient space $T_qQ/\mathfrak{g}\cdot q$, where q is a representative of the class x and where $\mathfrak{g}\cdot q=\{\xi_Q(q)|\xi\in\mathfrak{g}\}$ is the tangent space to the group orbit through q. The isomorphism in question is induced by the tangent map $T_q\pi_{Q,G}:T_qQ\to T_x(Q/G)$, whose kernel is exactly $\mathfrak{g}\cdot q$.

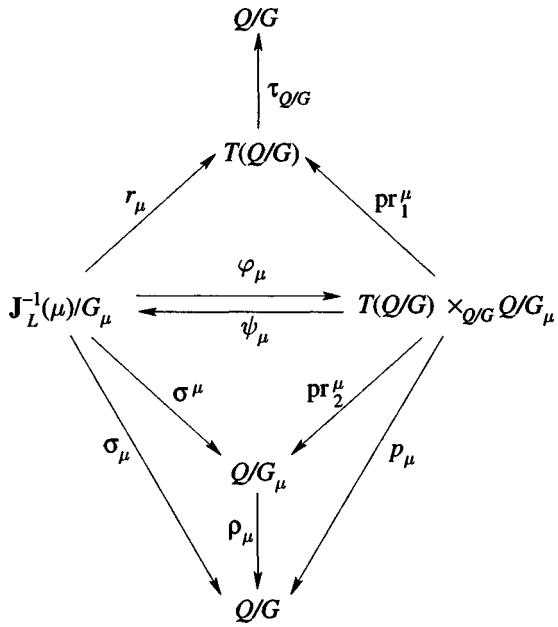

FIG. 3. The maps involved in the proof of the global realization theorem.

Lemma III.2: Let $u_x=[w_q]\in T_qQ/\mathfrak{g}\cdot q$. *There exists a unique* $\xi\in\mathfrak{g}$ *such that* $v_q:=w_q+\xi_Q(q)\in\mathbf{J}_L^{-1}(\mu)$. *In fact,* $\xi=\mathfrak{I}(q)^{-1}(\mu-\mathbf{J}_L(w_q))$.

Proof: The condition that $\mathbf{J}_L(v_q)=\mu$ is equivalent to the following condition for all $\eta\in\mathfrak{g}$:

$$\begin{aligned}\langle\mu,\eta\rangle&=\langle\mathbf{J}_L(w_q),\eta\rangle+\langle\mathbf{J}_L(\xi_Q(q)),\eta\rangle\\&=\langle\mathbf{J}_L(w_q),\eta\rangle+\langle\langle\xi_Q(q),\eta_Q(q)\rangle\rangle=\langle\mathbf{J}_L(w_q),\eta\rangle+\langle\mathfrak{I}(q)\xi,\eta\rangle.\end{aligned}$$

Thus, this condition is equivalent to $\mu=\mathbf{J}_L(w_q)+\mathfrak{I}(q)\xi$. Solving for ξ gives the result. ▽

As a consequence, note that for each $u_x\in T_x(Q/G)$, and each $q\in Q$ with $[q]_G=x$, there is a $v_q\in\mathbf{J}_L^{-1}(\mu)$ such that $u_x=[v_q]$.

We claim that an inverse for ϕ_μ is the map $\psi_\mu:T(Q/G)\times_{Q/G}Q/G_\mu\to\mathbf{J}_L^{-1}(\mu)/G_\mu$ defined by $\psi_\mu(u_x,[q]_{G_\mu})=[v_q]_{G_\mu}$, where $x=[q]_G$ and $u_x=[v_q]$, with $v_q\in\mathbf{J}_L^{-1}(\mu)$ given by the above-mentioned lemma. To show that ψ_μ is well-defined, we must show that if we represent the pair $(u_x,[q]_{G_\mu})$, $x=[q]_G$, in a different way, the value of ψ_μ is unchanged.

Let $u_x=[\overline{v_q}]$, with $[q]_{G_\mu}=[\overline{q}]_{G_\mu}$ and $\overline{v_q}\in\mathbf{J}_L^{-1}(\mu)$. Then we must show that $[v_q]_{G_\mu}=[\overline{v_q}]_{G_\mu}$. Since $[q]_{G_\mu}=[\overline{q}]_{G_\mu}$, we can write $\overline{q}=h\cdot q$ for some $h\in G_\mu$. Consider $h^{-1}\cdot\overline{v_q}\in T_qQ$. By equivariance of $\mathbf{J}_L$, and the fact that $h\in G_\mu$, we have $h^{-1}\cdot\overline{v_q}\in\mathbf{J}_L^{-1}(\mu)$. However,

$$u_x=T_q\pi_{Q,G}(v_q)=T_{\overline{q}}\pi_{Q,G}(\overline{v_q})=T_q\pi_{Q,G}(h^{-1}\cdot\overline{v_q})$$

and therefore, $v_q-h^{-1}\cdot\overline{v_q}\in\mathfrak{g}\cdot q$. In other words, $v_q-h^{-1}\cdot\overline{v_q}=\xi_Q(q)$ for some $\xi\in\mathfrak{g}$. Applying $\mathbf{J}_L$ to each side gives $0=\mathbf{J}_L(\xi_Q(q))=\mathfrak{I}(q)\xi$ and so $\xi=0$. Thus, $v_q=h^{-1}\cdot\overline{v_q}$ and so $[v_q]_{G_\mu}=[\overline{v_q}]_{G_\mu}$. Thus, ψ_μ is a well-defined map.

To show that ψ_μ is smooth, we show that it has a smooth local representative. If we write, locally, $Q=S\times G$ where the action is on the second factor alone, then we identify $Q/G_\mu=S\times\mathcal{O}_\mu$ and $T(Q/G)\times_{Q/G}Q/G_\mu=TS\times\mathcal{O}_\mu$. We identify $\mathbf{J}_L^{-1}(\mu)$ with $TS\times G$ since the level set of the momentum map in local representation is given by the product of TS with the graph of the right invariant vector field on G whose value at e is the vector $\zeta\in\mathfrak{g}$ such that $\langle\langle\zeta,\eta\rangle\rangle=\langle\mu,\eta\rangle$. In

this representation, $\mathbf{J}_L^{-1}(\mu)/G_\mu$ is identified with $TS\times G/G_\mu$ and the map ψ_μ is given by $(u_x,[g]_{G_\mu})\in TS\times G/G_\mu\mapsto(u_x,g\cdot\mu)\in TS\times\mathcal{O}_\mu$. This map is smooth by the construction of the manifold structure on the orbit. Thus, ψ_μ is smooth.

It remains to show that ψ_μ and ϕ_μ are inverses. To do this, note that

$$(\psi_\mu\circ\phi_\mu)([v_q]_{G_\mu})=\psi_\mu(T_q\pi_{Q/G}(v_q),[q]_{G_\mu})=[v_q]_{G_\mu}$$

since v_q is, by assumption, in $\mathbf{J}_L^{-1}(\mu)$. □

Associated bundles. We now show that the bundle $\rho_\mu:Q/G_\mu\to Q/G$ is globally diffeomorphic to an associated coadjoint orbit bundle. Let $\mathcal{O}_\mu\subset\mathfrak{g}^*$ denote the coadjoint orbit through μ. The *associated coadjoint bundle* is the bundle $\tilde{\mathcal{O}}_\mu=(Q\times\mathcal{O}_\mu)/G$, where the action of G on Q is the given (left) action, the action of G on $\mathcal{O}_\mu$ is the left coadjoint action, and the action of G on $Q\times\mathcal{O}_\mu$ is the diagonal action. This coadjoint bundle is regarded as a bundle over Q/G with the projection map given by $\tilde{\rho}_\mu:\tilde{\mathcal{O}}_\mu\to Q/G;[(q,g\cdot\mu)]_G\mapsto[q]_G$.

Theorem III.3: *There is a global bundle isomorphism $\Phi_\mu:\tilde{\mathcal{O}}_\mu\to Q/G_\mu$ covering the identity on the base Q/G.*

Proof: As in the preceding theorem, we construct the map Φ_μ and show it is an isomorphism by constructing an inverse. Define Φ_μ by $[q,g_0\cdot\mu]_G\mapsto[g_0^{-1}\cdot q]_{G_\mu}$. To show that Φ_μ is well defined, suppose that $g_0\cdot\mu=\overline{g}\cdot\mu$ and $g\in G$. We have to show that $[g_0^{-1}\cdot q]_{G_\mu}=[((g\overline{g})^{-1})\cdot(g\cdot q)]_{G_\mu}$, i.e., $[g_0^{-1}\cdot q]_{G_\mu}=[\overline{g}^{-1}\cdot q]_{G_\mu}$, which is true because $g_0^{-1}\overline{g}\in G_\mu$. Define $\Psi_\mu:Q/G_\mu\to\tilde{\mathcal{O}}_\mu$ by $[q]_{G_\mu}\mapsto[q,\mu]_G$. It is clear that Ψ_μ is well defined and is the inverse of Φ_μ. Smoothness of each of these maps follows from general theorems on smoothness of quotient maps (see, e.g., Abraham, Marsden, and Ratiu [1988][3]). □

A consequence of these two theorems is that there are global bundle isomorphisms between the three bundles $\mathbf{J}_L^{-1}(\mu)/G_\mu$, $T(Q/G)\times_{Q/G}Q/G_\mu$, and $T(Q/G)\times_{Q/G}\tilde{\mathcal{O}}_\mu$.

The second space is convenient for analyzing the Routhian and the reduced variational principles, while the third is convenient for making links with the Hamiltonian side.

B. The Routhian

We again consider Lagrangians of the form kinetic minus potential using our earlier notation. Given a fixed $\mu\in\mathfrak{g}^*$, the associated *Routhian* $R^\mu:TQ\to\mathbb{R}$ is defined by

$$R^\mu(v_q)=L(v_q)-\langle\mu,\mathfrak{A}(v_q)\rangle.$$

Letting $\mathfrak{A}_\mu(v_q)=\langle\mu,\mathfrak{A}(v_q)\rangle$, we can write this simply as $R^\mu=L-\mathfrak{A}_\mu$.

Proposition III.4: *For $v_q\in\mathbf{J}_L^{-1}(\mu)$, we have $R^\mu(v_q)=\frac{1}{2}\|\mathrm{Hor}(v_q)\|^2-V_\mu(q)$, where the amended potential V_μ is given by $V_\mu(q)=V(q)+C_\mu(q)$ and $C_\mu=\frac{1}{2}\langle\mu,\mathfrak{I}(q)^{-1}\mu\rangle$ is called the amendment.*

Proof: Because the horizontal and vertical components in the mechanical connection are metrically orthogonal, we have

$$R^\mu(v_q)=\tfrac{1}{2}\|v_q\|^2-V(q)-\langle\mu,\mathfrak{A}(v_q)\rangle=\tfrac{1}{2}\|\mathrm{Hor}(v_q)\|^2+\tfrac{1}{2}\|\mathrm{Ver}(v_q)\|^2-V(q)-\langle\mu,\mathfrak{A}(v_q)\rangle.$$

For $v_q\in\mathbf{J}_L^{-1}(\mu)$, we have

$$\begin{aligned}\|\mathrm{Ver}(v_q)\|^2&=\|(\mathfrak{A}(v_q))_Q(q)\|^2=\langle\mathfrak{I}(q)\mathfrak{A}(v_q),\mathfrak{A}(v_q)\rangle\\&=\langle\mathbf{J}_L(v_q),\mathfrak{A}(v_q)\rangle=\langle\mu,\mathfrak{A}(v_q)\rangle=\langle\mu,\mathfrak{I}(q)^{-1}(\mu)\rangle.\end{aligned}$$ □

Using this, one now verifies the following:

Proposition III.5: *The function R^μ is G_μ-invariant and so it induces, by restriction and quotienting, a function on $\mathbf{J}_L^{-1}(\mu)/G_\mu$ and hence, by the global realization theorem, a function $\mathfrak{R}^\mu: T(Q/G)\times_{Q/G}(Q/G_\mu)\to\mathbb{R}$ called the reduced Routhian; it is given by*

$$\mathfrak{R}^\mu(u_x,[q]_{G_\mu})=\tfrac{1}{2}\|u_x\|_S^2-\mathfrak{V}_\mu([q]_{G_\mu}),$$

where $x=[q]_G$, the metric on $S=Q/G$ is naturally induced from the metric on Q (that is, if $u_x=T_q\pi_{Q,G}(v_q)$ then $\|u_x\|_S=\|\mathrm{Hor}(v_q)\|$), and $\mathfrak{V}_\mu: Q/G_\mu\to\mathbb{R}$ is the reduced amended potential given by $\mathfrak{V}_\mu([q]_{G_\mu})=V_\mu(q)$.

Additional notation will prove useful. Let $\mathfrak{V}$ be the function on Q/G induced by the function V on Q and let $\mathfrak{C}_\mu$ be the *reduced amendment*, the function on Q/G_μ induced by the amendment C_μ. Thus, $\mathfrak{V}_\mu=\mathfrak{V}\circ\rho_\mu+\mathfrak{C}_\mu$. Let the Lagrangian on Q/G be denoted $\mathfrak{L}=\mathfrak{K}-\mathfrak{V}$, where $\mathfrak{K}(u_x)=\|u_x\|_S^2/2$ is the kinetic energy on the shape space Q/G.

C. Examples

Rigid body. Here the shape space is a point since $Q=G$, $\mu=\boldsymbol{\pi}$, the spatial angular momentum, so $T(Q/G)\times_{Q/G}Q/G_{\boldsymbol{\pi}}=S^2_{\|\boldsymbol{\pi}\|}$, the sphere of radius $\|\boldsymbol{\pi}\|$, a coadjoint orbit for the rotation group. The reduced Routhian $\mathfrak{R}^{\boldsymbol{\pi}}: S^2_{\|\boldsymbol{\pi}\|}\to\mathbb{R}$ is the negative of the reduced amendment, namely $-\frac{1}{2}\mathbf{\Pi}\cdot I^{-1}\mathbf{\Pi}$. This is of course the negative of the reduced energy.

Heavy top. In this case $Q=\mathrm{SO}(3)$ and $G=S^1$ is the subgroup of rotations about the vertical axis. Shape space is $Q/G=S_1^2$, the sphere of radius 1. As with any Abelian group, $G_\mu=G$, so $T(Q/G)\times_{Q/G}Q/G_\mu=T(Q/G)$. In the case of the heavy top, we get TS_1^2.

The isomorphism from $\mathbf{J}_L^{-1}(\mu)/G_\mu\to TS^2$ is induced by the map that takes $(\mathbf{A},\dot{\mathbf{A}})$ to $(\mathbf{\Gamma},\dot{\mathbf{\Gamma}}=\mathbf{\Gamma}\times\mathbf{\Omega})$. One checks that the horizontal lift of $(\mathbf{\Gamma},\dot{\mathbf{\Gamma}})$ to the point $\mathbf{A}$ is the vector $(\mathbf{A},\dot{\mathbf{A}}_h)$ satisfying $\mathbf{A}^{-1}\dot{\mathbf{A}}_h=\mathbf{\Omega}_h$, where

$$\mathbf{\Omega}_h:=\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma}-\frac{(\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma})\cdot I\mathbf{\Gamma}}{\mathbf{\Gamma}\cdot I\mathbf{\Gamma}}\mathbf{\Gamma}.$$

In doing this computation, it may be helpful to keep in mind that the condition of horizontality is the same as zero momentum. Thus, the reduced Routhian is given by

$$\mathfrak{R}^\mu(\mathbf{\Gamma},\dot{\mathbf{\Gamma}})=\tfrac{1}{2}\langle\mathbf{\Omega}_h,I\mathbf{\Omega}_h\rangle-Mgl\mathbf{\Gamma}\cdot\boldsymbol{\chi}-\frac{1}{2}\frac{\mu^2}{\mathbf{\Gamma}\cdot I\mathbf{\Gamma}}.$$

Underwater vehicle. As we have seen, $Q=\mathrm{SE}(3)$, $G=\mathrm{SE}(2)\times\mathbb{R}$ and so again $Q/G=S_1^2$. However, because G is non-Abelian, for $\mu\neq 0$, the bundle $Q/G_\mu\to Q/G$ has nontrivial fibers. These fibers are coadjoint orbits for SE(2), namely cylinders. A computation shows that $Q/G_\mu=\mathrm{SO}(3)\times\mathbb{R}$, regarded as a bundle over S_1^2 by sending $(\mathbf{A},\lambda)$ to $\mathbf{A}^{-1}\mathbf{k}$. Thus, $T(Q/G)\times_{Q/G}Q/G_\mu=T(S_1^2)\times_{S_1^2}\mathrm{SO}(3)\times\mathbb{R}$, a six-dimensional space, a nontrivial bundle over the two sphere with fiber the product of the tangent space to the sphere with a cylinder. The reduced Routhian may be computed as in the previous example, but we omit the details.

D. Hamilton's variational principle and the Routhian

Now we shall recast Hamilton's principle for the Lagrangian L in terms of the Routhian. To do so, we shall first work out the expression for $\mathbf{d}\mathfrak{S}_{R^\mu}$.

Recalling that $R^\mu=L-\mathfrak{A}_\mu$ and that on the space of curves parametrized on a fixed interval $[a,b]$, $\mathfrak{S}_L(q(\cdot))=\int_a^b L(q(t),\dot{q}(t))dt$, we see that $\mathfrak{S}_{R^\mu}=\mathfrak{S}_L-\mathfrak{S}_{\mathfrak{A}_\mu}$, and hence that

$$\mathbf{d}\mathfrak{S}_{R^\mu}\cdot\delta q(t)=\mathbf{d}\mathfrak{S}_L\cdot\delta q(t)-\mathbf{d}\mathfrak{S}_{\mathfrak{A}_\mu}\cdot\delta q(t). \tag{III.1}$$

We know from the formula for $\mathbf{d}\mathfrak{S}_L$ given in Proposition I.1 that

$$\mathbf{d}\mathfrak{S}_L(q(t))\cdot\delta q(t)=\int_a^b \mathcal{EL}(L)\left(\frac{d^2q}{dt^2}\right)\cdot\delta q\,dt+\Theta_L\left(\frac{dq}{dt}\right)\cdot\widehat{\delta q}\Big|_a^b. \tag{III.2}$$

To work out the term $\mathbf{d}\mathfrak{S}_{\mathfrak{A}_\mu}\cdot\delta q(t)$ we shall proceed in a more geometric way.

Variations of integrals of forms. We shall pause for a moment to consider the general question of variations of the integrals of differential forms. Consider a manifold M, a k-dimensional compact oriented submanifold S (with boundary) and a k-form ω defined on M. By a *variation* of S we shall mean a vector field δs defined along S in the following way. Let $\varphi_\epsilon:M\to M$ be a family of diffeomorphisms of M with φ_0 the identity. Set

$$\delta s(m)=\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\varphi_\epsilon(m),\quad \delta\int_S\omega=\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\int_{\varphi_\epsilon(S)}\omega.$$

Proposition III.6: The above variation is given by

$$\delta\int_S\omega=\int_S\mathbf{i}_{\delta s}\,\mathbf{d}\omega+\int_{\delta S}\mathbf{i}_{\delta s}\omega,$$

where $\mathbf{i}_{\delta s}\omega$ *denotes the interior product of the vector field* δs *with the k-form* ω.

Proof: We use the definition, the change of variables formula, the Lie derivative, and Stokes' formula as follows:

$$\delta\int_S\omega=\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\int_{\varphi_\epsilon(S)}\omega=\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\int_S\varphi_\epsilon^*\omega=\int_S\pounds_{\delta s}\omega=\int_S\mathbf{i}_{\delta s}\mathbf{d}\omega+\int_S\mathbf{d}\mathbf{i}_{\delta s}\omega=\int_S\mathbf{i}_{\delta s}\mathbf{d}\omega+\int_{\partial S}\mathbf{i}_{\delta s}\omega.$$

□

Application to the mechanical connection. In particular, we can apply the preceding proposition to the variations of the integral of the one-form $\mathfrak{A}_\mu$ over curves. We get

$$\delta\int_a^b\mathfrak{A}_\mu=\int_a^b\mathbf{i}_{\delta q}\mathfrak{B}_\mu+\mathfrak{A}_\mu(\delta q(b))-\mathfrak{A}_\mu(\delta q(a)),$$

where $\mathfrak{B}_\mu=\mathbf{d}\mathfrak{A}_\mu$, the exterior derivative of the one form $\mathfrak{A}_\mu$.

The computation of boundary terms. Summing up what we have proved so far, we write

$$\begin{aligned}\mathbf{d}\mathfrak{S}_{R^\mu}(q(t),\dot q(t))\cdot\delta q&=\mathbf{d}\mathfrak{S}_L(q(t),\dot q(t))\cdot\delta q-\mathbf{d}\mathfrak{S}_{\mathfrak{A}_\mu}(q(t),\dot q(t))\cdot\delta q\\&=\int_a^b\mathcal{EL}(L)\left(\frac{d^2q}{dt^2}\right)\cdot\delta q\,dt+\Theta_L\left(\frac{dq}{dt}\right)\cdot\widehat{\delta q}\Big|_a^b\\&\quad-\int_a^b\mathbf{i}_{\delta q}\mathfrak{B}_\mu-[\mathfrak{A}_\mu(\delta q(b))-\mathfrak{A}_\mu(\delta q(a))].\end{aligned}$$

We now compute the boundary terms in this expression. Recalling the formula for the boundary terms in the variational formula for L, splitting the variation into horizontal and vertical parts, we get

$$\Theta_L\left(\frac{dq}{dt}\right)\cdot\widehat{\delta q}\Big|_a^b=\langle \mathbb{F}L(q(t),\dot q(t)),\delta q\rangle|_a^b=\langle \mathbb{F}L(q(t),\dot q(t)),\mathrm{Hor}\,\delta q\rangle|_a^b+\langle \mathbb{F}L(q(t),\dot q(t)),\mathrm{Ver}\,\delta q\rangle|_a^b$$
$$=\langle\langle \dot q(t),\mathrm{Hor}\,\delta q\rangle\rangle|_a^b+\langle\langle \dot q(t),\mathrm{Ver}\,\delta q\rangle\rangle|_a^b.$$

Assuming the curve $(q(t),\dot q(t))$ lies in the level set of the momentum map, we have

$$\langle\langle \dot q(t),\mathrm{Ver}\,\delta q\rangle\rangle=\langle\langle \dot q(t),[\mathfrak{A}(\delta q)]_Q(q)\rangle\rangle=\langle \mathbf{J}_L(q(t),\dot q(t)),\mathfrak{A}(\delta q)\rangle=\langle \mu,\mathfrak{A}(\delta q)\rangle=\mathfrak{A}_\mu(\delta q).$$

Therefore, we get

$$\Theta_L\left(\frac{dq}{dt}\right)\cdot\widehat{\delta q}\Big|_a^b=\langle\langle \dot q(t),\mathrm{Hor}\,\delta q\rangle\rangle|_a^b+\mathfrak{A}_\mu(\delta q)|_a^b.$$

Noticing that the terms involving $\mathfrak{A}_\mu$ cancel, we can say, in summary, that

$$\mathbf{d}\mathfrak{S}_{R^\mu}(q(t),\dot q(t))\cdot\delta q=\mathbf{d}\mathfrak{S}_L(q(t),\dot q(t))\cdot\delta q-\mathbf{d}\mathfrak{S}_{\mathfrak{A}_\mu}(q(t),\dot q(t))\cdot\delta q$$
$$=\int_a^b\mathcal{EL}(L)\left(\frac{d^2q}{dt^2}\right)\cdot\delta q\,dt-\int_a^b\mathbf{i}_{\delta q}\mathfrak{B}_\mu+\langle\langle \dot q(t),\mathrm{Hor}\,\delta q\rangle\rangle|_a^b.$$

We can conclude the following.

Theorem III.7: *A solution of the Euler–Lagrange equations which lies in the level set* $\mathbf{J}_L=\mu$, *satisfies the following variational principle:*

$$\delta\int_a^b R^\mu(q(t),\dot q(t))dt=-\int_a^b\mathbf{i}_{\delta q}\mathfrak{B}_\mu(q(t),\dot q(t))dt+\langle\langle \dot q(t),\mathrm{Hor}\,\delta q\rangle\rangle|_a^b.$$

It is very important to notice that in this formulation, there are no boundary conditions or constraints whatsoever imposed on the variations. However, we can choose vanishing boundary conditions for δq and derive:

Corollary III.8: *Any solution of the Euler–Lagrange equations which lies in the level set* $\mathbf{J}_L=\mu$, *also satisfies the equations*

$$\mathcal{EL}(R^\mu)\left(\frac{d^2q}{dt^2}(t)\right)=\mathbf{i}_{\dot q(t)}\mathfrak{B}_\mu.$$

Conversely, any solution of these equations that lies in the level set $\mathbf{J}_L=\mu$ *of the momentum map is a solution to the Euler–Lagrange equations for* L.

In deriving these equations, we have interchanged the contractions with δq and $\dot q$ using skew symmetry of the two-form $\mathfrak{B}_\mu$. One can also check this result with a coordinate computation, as was done in Marsden and Scheurle [1993];[108] see also Marsden and Ratiu [1994][98] for this calculation in the case of Abelian groups.

E. The Routh variational principle on quotients

We now show how to drop the variational principle given in Sec. III C to the reduced space $T(Q/G)\times_{Q/G}Q/G_\mu$. An important point is whether or not one imposes constraints on the variations in the variational principle. One of our main points is that such constraints *are not needed*; for a corresponding derivation with the varied curves constrained to lie in the level set of the momentum map, see Jalnapurkar and Marsden [2000a].[64]

Later in this section we illustrate the procedure with the rigid body, which already contains the key to how one relaxes the constraints. Some readers may find it convenient to study that example simultaneously with the general theory.

Our first goal is to show that the variation of the Routhian evaluated at a solution depends only on the quotient variations. Following this, we shall show that the ***gyroscopic terms*** $\mathbf{i}_{\dot{q}(t)}\mathfrak{B}_\mu$ also depend only on the quotient variations.

Analysis of the variation of the Routhian. We begin by writing the Routhian as follows:

$$R^\mu(v_q)=\tfrac{1}{2}\|\mathrm{Hor}(v_q)\|^2+\tfrac{1}{2}\|\mathrm{Ver}(v_q)\|^2-V(q)-\langle\mu,\mathfrak{A}(v_q)\rangle. \tag{III.3}$$

We next analyze the variation of two of the terms in this expression, namely

$$\tfrac{1}{2}\|\mathrm{Ver}(v_q)\|^2-\langle\mu,\mathfrak{A}(v_q)\rangle=\tfrac{1}{2}\langle\mathfrak{I}(q)\mathfrak{A}(v_q),\mathfrak{A}(v_q)\rangle-\langle\mu,\mathfrak{A}(v_q)\rangle.$$

We choose a family of curves $q(t,\epsilon)$ with the property that $q(t,0)$ is a solution of the Euler–Lagrange equations with a momentum value μ and let v_q be the time derivative of $q(t,0)$. As usual, we also let δq be the ϵ derivative evaluated at $\epsilon=0$. Then, the desired variation is given by

$$\begin{aligned}\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}&\left(\frac{1}{2}\left\langle\mathfrak{I}(q)\mathfrak{A}\left(\frac{\partial q}{\partial t}\right),\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle-\left\langle\mu,\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle\right)\\&=\frac{1}{2}\langle(T_q\mathfrak{I}\cdot\delta q)\mathfrak{A}(v_q),\mathfrak{A}(v_q)\rangle+\left\langle\mathfrak{I}(q)\mathfrak{A}(v_q)-\mu,\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle.\end{aligned} \tag{III.4}$$

Here, T_q denotes the tangent map at the point q. Since the curve $q(t,0)$ is assumed to be a solution with momentum value μ and since $\mathfrak{I}(q)\mathfrak{A}(v_q)=\mathbf{J}_L(v_q)$, the second term in the preceding display vanishes. Thus, we conclude that

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\left(\frac{1}{2}\left\langle\mathfrak{I}(q)\mathfrak{A}\left(\frac{\partial q}{\partial t}\right),\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle-\left\langle\mu,\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle\right)=\frac{1}{2}\langle(T_q\mathfrak{I}\cdot\delta q)\mathfrak{A}(v_q),\mathfrak{A}(v_q)\rangle. \tag{III.5}$$

Next, we observe that

$$\mathbf{d}(\tfrac{1}{2}\langle\mu,\mathfrak{I}(q)^{-1}\mu\rangle)\cdot\delta q=-\tfrac{1}{2}\langle\mu,\mathfrak{I}(q)^{-1}(T_q\mathfrak{I}\cdot\delta q)\mathfrak{I}(q)^{-1}\mu\rangle. \tag{III.6}$$

On a solution with momentum value μ, we have $\mu=\mathbf{J}_L(v_q)=\mathfrak{I}(q)\mathfrak{A}(v_q)$. Substituting this into the preceding expression, we get

$$\begin{aligned}\mathbf{d}(\tfrac{1}{2}\langle\mu,\mathfrak{I}(q)^{-1}\mu\rangle)\cdot\delta q&=-\tfrac{1}{2}\langle\mathfrak{I}(q)\mathfrak{A}(v_q),\mathfrak{I}(q)^{-1}(T_q\mathfrak{I}\cdot\delta q)\mathfrak{I}(q)^{-1}\mathfrak{I}(q)\mathfrak{A}(v_q)\rangle\\&=-\frac{1}{2}\langle(T_q\mathfrak{I}\cdot\delta q)\mathfrak{A}(v_q),\mathfrak{A}(v_q)\rangle.\end{aligned} \tag{III.7}$$

Therefore, on a solution with momentum value μ, we have

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\left(\frac{1}{2}\left\langle\mathfrak{I}(q)\mathfrak{A}\left(\frac{\partial q}{\partial t}\right),\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle-\left\langle\mu,\mathfrak{A}\left(\frac{\partial q}{\partial t}\right)\right\rangle\right) \tag{III.8}$$

$$=-\mathbf{d}\left(\frac{1}{2}\langle\mu,\mathfrak{I}(q)^{-1}\mu\rangle\right)\cdot\delta q. \tag{III.9}$$

We conclude that when evaluated on a solution with momentum value μ,

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}R^\mu\left(\frac{\partial q}{\partial t}\right)=\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\left(\frac{1}{2}\left\|\mathrm{Hor}\left(\frac{\partial q}{\partial t}\right)\right\|^2-V_\mu(q)\right)=\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\overline{R}^\mu\left(\frac{\partial q}{\partial t}\right), \tag{III.10}$$

where $\overline{R^\mu}(v_q)=\frac{1}{2}\|\mathrm{Hor}(v_q)\|^2-V_\mu(q)$. Proposition III.4 shows that $\overline{R^\mu}$ agrees with R^μ on $\mathbf{J}_L^{-1}(\mu)$ and, more important, $\overline{R^\mu}=\mathfrak{R}^\mu\circ(T\pi_{Q,G}\times_{Q/G}\pi_{Q,G_\mu})$, where, recall, $\pi_{Q,G}:Q\to Q/G$ and $\pi_{Q,G_\mu}:Q\to Q/G_\mu$ are the projection maps. Thus, $\overline{R^\mu}$ drops to the quotient with no restriction to the level set of the momentum map. Differentiating this relation with respect to ϵ, it follows that the variation of $\overline{R^\mu}$ drops to the variation of $\mathfrak{R}^\mu$.

Analysis of the variation of the gyroscopic terms. Now we shall show how the exterior derivative of the one form $\mathfrak{A}_\mu$ drops to the quotient space. Precisely, this means the following. We consider the one-form $\mathfrak{A}_\mu$ on the space Q and its exterior derivative $\mathfrak{B}_\mu=\mathbf{d}\mathfrak{A}_\mu$. We claim that *there is a unique two-form* β_μ on Q/G_μ such that $\mathfrak{B}_\mu=\pi^*_{Q,G_\mu}\beta_\mu$, where, recall, $\pi_{Q,G_\mu}:Q\to Q/G_\mu$ is the natural projection. To prove this, one must show that for any $u,v\in T_qQ$, the following identity holds:

$$\mathbf{d}\mathfrak{A}_\mu(q)(u,v)=\mathbf{d}\mathfrak{A}_\mu(g\cdot q)(g\cdot u+\xi_Q(g\cdot q),g\cdot v+\eta_Q(g\cdot q)), \tag{III.11}$$

for any $g\in G_\mu$, and $\xi,\eta\in\mathfrak{g}_\mu$. To prove this, one first shows that

$$\mathbf{d}\mathfrak{A}_\mu(g\cdot q)(g\cdot u+\xi_Q(g\cdot q),g\cdot v+\eta_Q(g\cdot q))=\mathbf{d}\mathfrak{A}_\mu(q)(u+(\mathrm{Ad}_{g^{-1}}\xi)_Q(q),v+(\mathrm{Ad}_{g^{-1}}\eta)_Q(q))$$

using the identities $\xi_Q(g\cdot q)=(\mathrm{Ad}_{g^{-1}}\xi)_Q(q)$ and $\Phi_g^*\mathfrak{A}_\mu=\mathfrak{A}_\mu$, where $\Phi_g(q)=g\cdot q$ is the group action. Second, one shows that $\mathbf{d}\mathfrak{A}_\mu(q)(u+\zeta_Q(q),v)=\mathbf{d}\mathfrak{A}_\mu(q)(u,v)$ for any $\zeta\in\mathfrak{g}_\mu$. This holds because $\mathbf{i}_{\zeta_Q}\mathbf{d}\mathfrak{A}_\mu=0$. Indeed, from $\Phi_g^*\mathfrak{A}_\mu=\mathfrak{A}_\mu$ we get $\pounds_{\zeta_Q}\mathfrak{A}_\mu=0$ and hence $\mathbf{i}_{\zeta_Q}\mathbf{d}\mathfrak{A}_\mu+\mathbf{d}\mathbf{i}_{\zeta_Q}\mathfrak{A}_\mu=0$. However, $\mathbf{i}_{\zeta_Q}\mathfrak{A}_\mu=\langle\mu,\zeta\rangle$, a constant, so we get the desired result.

Now we can apply Theorem III.7 to obtain the following result.

Theorem III.9: *If $q(t),a\leq t\leq b$, is a solution of the Euler–Lagrange equations with momentum value $\mu,y(t)=\pi_{Q,G_\mu}(q(t))$, and $x(t)=\pi_{Q,G}(q(t))$, then $y(t)$ satisfies the* ***reduced variational principle***

$$\delta\int_a^b\mathfrak{R}^\mu(x(t),\dot{x}(t),y(t))dt=\int_a^b\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))\cdot\delta y\,dt+\langle\langle\dot{x}(t),\delta x(t)\rangle\rangle_S\big|_a^b.$$

Conversely, if $q(t)$ is a curve such that $\dot{q}(t)\in\mathbf{J}_L^{-1}(\mu)$ and if its projection to $y(t)$ satisfies this reduced variational principle, then $q(t)$ is a solution of the Euler–Lagrange equations.

It is already clear from the case of the Euler–Poincaré equations that dropping the variational principle to the quotient can often be easier than dropping the equations themselves. Notice also that there is a slight abuse of notation, similar to that when one writes a tangent vector as a pair $(q,\dot{q})$. The notation (x,y) is redundant since x can be recovered from y by projection from Q/G_μ to Q/G. Consistent with this convenient notational abuse, we use the notation $(x,\dot{x})$ as an alternative to u_x.

F. Curvature

We pause briefly to recall some key facts about curvatures of connections, and establish our conventions. Then we shall relate β_μ to curvature.

Review of the curvature of a principal connection. Consider a principal connection $\mathcal{A}$ on a principal G bundle $\pi_{Q,G}:Q\to Q/G$. The ***curvature*** $\mathcal{B}$ is the Lie algebra-valued two-form on Q defined by $\mathcal{B}(u_q,v_q)=\mathbf{d}\mathcal{A}(\mathrm{Hor}_q(u_q),\mathrm{Hor}_q(v_q))$, where $\mathbf{d}$ is the exterior derivative.

Using the fact that $\mathcal{B}$ depends only on the horizontal part of the vectors and equivariance, one shows that it defines an adjoint bundle (that is, $\tilde{\mathfrak{g}}$)-valued two-form on the base Q/G by $\mathrm{Curv}_{\mathcal{A}}(x)(u_x,v_x)=[q,\mathbf{d}\mathcal{A}(u_q,v_q)]_G$, where $[q]_G=x\in Q/G$, u_q and v_q are horizontal, $T\pi_{Q,G}\cdot u_q=u_x$, and $T\pi_{Q,G}\cdot v_q=v_x$.

Curvature *measures the lack of integrability of the horizontal distribution* in the sense that on two vector fields u, v on Q one has

$$\mathcal{B}(u,v) = -\mathcal{A}([\mathrm{Hor}(u),\mathrm{Hor}(v)]).$$

The proof uses the Cartan formula relating the exterior derivative and the Jacobi–Lie bracket:

$$\mathcal{B}(u,v) = \mathrm{Hor}(u)[\mathcal{A}(\mathrm{Hor}(v))] - \mathrm{Hor}(v)[\mathcal{A}(\mathrm{Hor}(u))] - \mathcal{A}([\mathrm{Hor}(u),\mathrm{Hor}(v)]).$$

The first two terms vanish since $\mathcal{A}$ vanishes on horizontal vectors.

An important formula for the curvature of a principal connection is given by the ***Cartan structure equations***: for any vector fields u, v on Q one has

$$\mathcal{B}(u,v) = \mathbf{d}\mathcal{A}(u,v) - [\mathcal{A}(u),\mathcal{A}(v)],$$

where the bracket on the right-hand side is the Lie bracket in $\mathfrak{g}$. One writes this equation for short as $\mathcal{B} = \mathbf{d}\mathcal{A} - [\mathcal{A},\mathcal{A}]$. An important consequence of these equations that we will need below is the following identity (often this is a lemma used to prove the structure equations):

$$\mathbf{d}\mathcal{A}(q)(\mathrm{Hor}\, u_q, \mathrm{Ver}\, v_q) = 0 \tag{III.12}$$

for any $u_q, v_q \in T_q Q$.

Recall also that when applied to the left trivializing connection on a Lie group, the structure equations reduce to the Mauer–Cartan equations. We also remark, although we shall not need it, that one has the ***Bianchi identities***: For any vector fields u, v, w on Q, we have

$$\mathbf{d}\mathcal{B}(\mathrm{Hor}(u),\mathrm{Hor}(v),\mathrm{Hor}(w)) = 0.$$

The connection on the bundle ρ_μ. The bundle $\rho_\mu : Q/G_\mu \to Q/G$ has an Ehresmann connection induced from the principal connections on the two bundles $Q \to Q/G_\mu$ and $Q \to Q/G$. However, we can also determine this connection directly by giving its horizontal space at each point $y = [q]_{G_\mu} \in Q/G_\mu$. This horizontal space is taken to be the orthogonal complement within $T_y(Q/G_\mu) \cong T_q Q/[\mathfrak{g}_\mu \cdot q]$ to the vertical space $[\mathfrak{g}\cdot q]/[\mathfrak{g}_\mu \cdot q]$. This latter space inherits its metric from that on $T_q Q$ by taking the quotient metric. As before, since the action is by isometries, this metric is independent of the representatives chosen.

This horizontal space is denoted by Hor_{ρ_μ} and the operation of taking the horizontal part of a vector is denoted by the same symbol. The vertical space is of course the fiber of this bundle. This vertical space at a point $y = [q]_{G_\mu}$ is given by $\ker T_y \rho_\mu$, which is isomorphic to the quotient space $[\mathfrak{g}\cdot q]/[\mathfrak{g}_\mu \cdot q]$. This vertical bundle will be denoted by $\mathrm{Ver}(Q/G_\mu) \subset T(Q/G_\mu)$ and the fiber at the point $y \in Q/G_\mu$ is denoted $\mathrm{Ver}_y(Q/G_\mu) = \ker T_y \rho_\mu$. The projection onto the vertical part defines the analog of the connection form, which we denote $\mathfrak{A}_{\rho_\mu}$. Thus, $\mathfrak{A}_{\rho_\mu} : T(Q/G_\mu) \to \mathrm{Ver}(Q/G_\mu)$, which we think of as a vertical valued one-form.

Compatibility of the three connections. We shall now work toward the computation of β_μ on various combinations of horizontal and vertical vectors relative to the connection $\mathfrak{A}_{\rho_\mu}$. To do this, keep in mind that $\rho_\mu \circ \pi_{Q,G_\mu} = \pi_{Q,G}$ by construction. We shall need the following.

Lemma III.10: Let $u_q \in T_q Q$ and $u_y = T_y \pi_{Q,G_\mu} \cdot u_q$, where $y = \pi_{Q,G_\mu}(q)$. Then

(1) *u_y is ρ_μ-vertical if and only if u_q is $\pi_{Q,G}$-vertical.*
(2) *The identity $T\pi_{Q,G_\mu} \cdot \mathrm{Hor}\, u_q = \mathrm{Hor}_{\rho_\mu}(u_y)$ holds, where* Hor *denotes the horizontal projection for the mechanical connection $\mathfrak{A}$.*
(3) *The following identity holds: $T\pi_{Q,G_\mu} \cdot \mathrm{Ver}\, u_q = \mathrm{Ver}_{\rho_\mu}(u_y)$.*

Proof. (1): Because $\rho_\mu \circ \pi_{Q,G_\mu} = \pi_{Q,G}$, the chain rule gives

$$T_y \rho_\mu \cdot u_y = T_y \rho_\mu \cdot T_q \pi_{Q,G_\mu} \cdot u_q = T_q \pi_{Q,G} \cdot u_q,$$

so $T_y\rho_\mu \cdot u_y = 0$ if and only if $T_q\pi_{Q,G}\cdot u_q = 0$, which is the statement of part 1.
2. Let $v_y \in \ker T_y\rho_\mu \subset T_y(Q/G_\mu)$ and write $v_y = T_q\pi_{Q,G_\mu}\cdot v_q$. By the definition of the metric $\langle\langle\cdot,\cdot\rangle\rangle_{Q/G_\mu}$ on Q/G_μ we have

$$0 = \langle\langle \mathrm{Hor}\, u_q, v_q\rangle\rangle_Q = \langle\langle T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q, T_q\pi_{Q,G_\mu}\cdot v_q\rangle\rangle_{Q/G_\mu} = \langle\langle T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q, v_y\rangle\rangle_{Q/G_\mu}.$$

Hence, since v_y is an arbitrary ρ_μ-vertical vector, we conclude that $T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q$ is ρ_μ-horizontal. Next, write

$$\mathrm{Hor}_{\rho_\mu}(T\pi_{Q,G_\mu}\cdot u_q) = \mathrm{Hor}_{\rho_\mu}(T\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q) + \mathrm{Hor}_{\rho_\mu}(T\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, u_q) = T\pi_{Q,G_\pi}\cdot \mathrm{Hor}\, u_q$$

by assertion 1 of this lemma.
3. As in 1, $T_q\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, u_q$ is ρ_μ-vertical. Therefore,

$$\begin{aligned}\mathrm{Ver}_{\rho_\mu} u_y = \mathrm{Ver}_{\rho_\mu}(T_q\pi_{Q,G_\mu}\cdot u_q) &= \mathrm{Ver}_{\rho_\mu}(T_q\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, u_q) + \mathrm{Ver}_{\rho_\mu}(T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q)\\ &= T_q\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, u_q,\end{aligned}$$

since $T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q$ is ρ_μ-horizontal by part 2. □

The pairing between $\tilde{\mathfrak{g}}$ *and* $\tilde{\mathfrak{g}}^*$. We shall need to define a natural pairing between the adjoint and coadjoint bundles. Recall that, by definition, $\tilde{\mathfrak{g}} = (Q\times\mathfrak{g})/G$ and $\tilde{\mathfrak{g}}^* = (Q\times\mathfrak{g}^*)/G$, where G acts by the given action on Q and by the adjoint action on $\mathfrak{g}$ and the coadjoint action on $\mathfrak{g}^*$. For $[q,\mu]_G \in \tilde{\mathfrak{g}}^*$ and $[q,\xi]_G \in \tilde{\mathfrak{g}}$, the pairing is $\langle [q,\mu]_G, [q,\xi]_G\rangle = \langle \mu,\xi\rangle$. One shows that this pairing is independent of the representatives chosen.

We define, for $y = [q]_{G_\mu} \in Q/G_\mu$, the (y,μ)-component of $\mathrm{Curv}_{\mathcal{A}}$ by

$$\mathrm{Curv}_{\mathcal{A}}^{(y,\mu)}(x)(u_x,v_x) = \langle [q,\mu]_G, \mathrm{Curv}_{\mathcal{A}}(x)(u_x,v_x)\rangle, \tag{III.13}$$

where $[q]_G = x$. One shows that this is independent of the representative q chosen for y.

Hor–hor components of β_μ. Now we compute the horizontal–horizontal components of β_μ as follows. Let $u_q, v_q \in T_qQ$, and

$$u_y = T_q\pi_{Q,G_\mu}\cdot u_q, \quad v_y = T_q\pi_{Q,G_\mu}\cdot v_q.$$

Using Lemma III.10, the definition of curvature, and (III.13), we have

$$\begin{aligned}\beta_\mu(y)(\mathrm{Hor}_{\rho_\mu} u_y, \mathrm{Hor}_{\rho_\mu} v_y) &= \beta_\mu(\pi_{Q,G_\mu}(q))(T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q, T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, v_q)\\ &= (\pi^*_{Q,G_\mu}\beta_\mu)(q)(\mathrm{Hor}\, u_q, \mathrm{Hor}\, v_q) = \mathfrak{B}_\mu(q)(\mathrm{Hor}\, u_q, \mathrm{Hor}\, v_q)\\ &= \langle \mu, \mathbf{d}\mathfrak{A}(q)(\mathrm{Hor}\, u_q, \mathrm{Hor}\, v_q)\rangle\\ &= \langle [q,\mu]_G, [q, \mathbf{d}\mathfrak{A}(q)(\mathrm{Hor}\, u_q, \mathrm{Hor}\, v_q)]_G\rangle\\ &= \langle [q,\mu]_G, \mathrm{Curv}_{\mathcal{A}}(x)(u_x,v_x)\rangle = \mathrm{Curv}_{\mathfrak{A}}^{(y,\mu)}(x)(u_x,v_x),\end{aligned}$$

where $x = \pi_{Q,G}(q) = \rho_\mu(y)$, $u_x = T_q\pi_{Q,G}\cdot u_q = T_y\rho_\mu\cdot u_y$, and similarly for v_x.

We summarize what we have proved in the following lemma.

Lemma III.11: *The two-form β_μ on horizontal vectors is given by*

$$\beta_\mu(y)(\mathrm{Hor}_{\rho_\mu} u_y, \mathrm{Hor}_{\rho_\mu} v_y) = \mathrm{Curv}_{\mathfrak{A}}^{(y,\mu)}(x)(u_x,v_x). \tag{III.14}$$

Hor-ver components of β_μ. Now we compute the horizontal–vertical components of β_μ as follows. Let u_q, $v_q \in T_qQ$, and $u_y = T_q\pi_{Q,G_\mu}\cdot u_q$, $v_y = T_q\pi_{Q,G_\mu}\cdot v_q$. Using Lemma III.10, we have

$$\begin{aligned}\beta_\mu(y)(\mathrm{Hor}_{\rho_\mu}u_y, \mathrm{Ver}_{\rho_\mu}v_y) &= \beta_\mu(\pi_{Q,G_\mu}(q))(T_q\pi_{Q,G_\mu}\cdot \mathrm{Hor}\, u_q, T_q\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, v_q)\\ &= (\pi^*_{Q,G_\mu}\beta_\mu)(q)(\mathrm{Hor}\, u_q, \mathrm{Ver}\, v_q)\\ &= \mathfrak{B}_\mu(q)(\mathrm{Hor}\, u_q, \mathrm{Ver}\, v_q) = \langle \mu, \mathbf{d}\mathfrak{A}(q)(\mathrm{Hor}\, u_q, \mathrm{Ver}\, v_q)\rangle = 0,\end{aligned}$$

by (III.12). We summarize what we have proved in the following lemma.

Lemma III.12: *The two-form* β_μ *on pairs of horizontal and vertical vectors vanishes:*

$$\beta_\mu(y)(\mathrm{Hor}_{\rho_\mu}u_y, \mathrm{Ver}_{\rho_\mu}v_y) = 0. \tag{III.15}$$

Ver–ver components of β_μ. Now we compute the vertical–vertical components of β_μ as follows. As previously, let u_q, $v_q \in T_qQ$, and $u_y = T_q\pi_{Q,G_\mu}\cdot u_q$, $v_y = T_q\pi_{Q,G_\mu}\cdot v_q$. Using Lemma III.10, we have

$$\begin{aligned}\beta_\mu(y)(\mathrm{Ver}_{\rho_\mu}u_y, \mathrm{Ver}_{\rho_\mu}v_y) &= \beta_\mu(\pi_{Q,G_\mu}(q))(T_q\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, u_q, T_q\pi_{Q,G_\mu}\cdot \mathrm{Ver}\, v_q)\\ &= (\pi^*_{Q,G_\mu}\beta_\mu)(q)(\mathrm{Ver}\, u_q, \mathrm{Ver}\, v_q)\\ &= \mathfrak{B}_\mu(q)(\mathrm{Ver}\, u_q, \mathrm{Ver}\, v_q)\\ &= \langle \mu, \mathbf{d}\mathfrak{A}(q)(\mathrm{Ver}\, u_q, \mathrm{Ver}\, v_q)\rangle = \langle \mu, [\mathfrak{A}(q)\mathrm{Ver}\, u_q, \mathfrak{A}(q)\mathrm{Ver}\, v_q]\rangle\end{aligned}$$

by the Cartan structure equations. We now write

$$\mathrm{Ver}\, u_q = \xi_Q(q), \quad \mathrm{Ver}\, v_q = \eta_Q(q),$$

so that the preceding equation becomes

$$\beta_\mu(y)(\mathrm{Ver}_{\rho_\mu}u_y, \mathrm{Ver}_{\rho_\mu}v_y) = \langle \mu, [\xi, \eta]\rangle. \tag{III.16}$$

Now given $\mathrm{Ver}_{\rho_\mu}u_y \in \ker T_y\rho_\mu$, we can represent it as a class $[\xi_Q(q)] \in \mathfrak{g}\cdot q/\mathfrak{g}_\mu\cdot q$. The map $\xi \mapsto \xi_Q(q)$ induces an isomorphism of $\mathfrak{g}/\mathfrak{g}_\mu$ with the ρ_μ-vertical space. Note that the above-mentioned formula depends only on the class of ξ and of η.

We summarize what we have proved in the following lemma.

Lemma III.13: *The two-form* β_μ *on pairs of vertical vectors is given by the following formula:*

$$\beta_\mu(y)(\mathrm{Ver}_{\rho_\mu}u_y, \mathrm{Ver}_{\rho_\mu}v_y) = \langle \mu, [\xi, \eta]\rangle, \tag{III.17}$$

where $\mathrm{Ver}_{\rho_\mu}u_y = [\xi_Q(q)]$ *and* $\mathrm{Ver}_{\rho_\mu}v_y = [\eta_Q(q)]$.

G. Splitting the reduced variational principle

Now we want to take the reduced variational principle, namely

$$\delta\int_a^b \mathfrak{R}^\mu(x(t), \dot{x}(t), y(t))\,dt = \int_a^b \mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))\cdot \delta y\, dt + \langle\langle \dot{x}(t), \delta x(t)\rangle\rangle\big|_a^b$$

and relate it intrinsically to two sets of differential equations corresponding to the horizontal and vertical components of the bundle $\rho_\mu : Q/G_\mu \to Q/G$.

Recall that in this principle, we are considering all curves $y(t)\in Q/G_\mu$ and $x(t)=\rho_\mu(y(t))\in Q/G$. For purposes of deriving the equations, we can restrict ourselves to variations such that δx vanishes at the end points, so that the boundary term disappears.

Now the strategy is to split the variations $\delta y(t)$ of $y(t)$ into horizontal and vertical components relative to the induced connection on the bundle $\rho_\mu: Q/G_\mu \to Q/G$.

Breaking up the variational principle. Now we can break up the variational principle by decomposing variations into their horizontal and vertical pieces, which we shall write

$$\delta y = \mathrm{Hor}_{\rho_\mu}\,\delta y + \mathrm{Ver}_{\rho_\mu}\,\delta y,$$

where

$$\mathfrak{A}_{\rho_\mu}\delta y = \mathrm{Ver}_{\rho_\mu}\,\delta y.$$

We also note that, by construction, the map $T\rho_\mu$ takes δy to δx. Since this map has kernel given by the set of vertical vectors, it defines an isomorphism on the horizontal space to the tangent space to shape space. Thus, we can identify $\mathrm{Hor}_{\rho_\mu}\,\delta y$ with δx.

Horizontal variations. Now we take variations that are purely horizontal and vanish at the end points; that is, $\delta y = \mathrm{Hor}_{\rho_\mu}\,\delta y$. In this case, the variational principle,

$$\delta\int_a^b \mathfrak{R}^\mu(x(t),\dot{x}(t),y(t))dt = \int_a^b \mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))\cdot\delta y\,dt + \langle\langle \dot{x}(t),\delta x(t)\rangle\rangle\big|_a^b \tag{III.18}$$

becomes

$$\delta\int_a^b [\mathfrak{L}(x(t),\dot{x}(t)) - \mathfrak{C}_\mu(y(t))]dt\cdot\mathrm{Hor}_{\rho_\mu}\,\delta y(t) = \int_a^b (\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t)))\cdot\mathrm{Hor}_{\rho_\mu}\,\delta y(t)dt. \tag{III.19}$$

Since, by our general variational formula, for variations vanishing at the end points,

$$\delta\left(\int_a^b \mathfrak{L}(x(t),\dot{x}(t))dt\right)\cdot\delta x = \int_a^b \mathcal{EL}(\mathfrak{L})(x(t),\dot{x}(t),\ddot{x}(t))\cdot\delta x(t)dt,$$

(III.19) is equivalent to

$$\mathcal{EL}(\mathfrak{L})(\ddot{x}) = \mathrm{Hor}_{\rho_\mu}[\mathbf{d}\mathfrak{C}_\mu(y) + \mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))] \tag{III.20}$$

where, for a point $\gamma\in T_y^*(Q/G_\mu)$, we define

$$\mathrm{Hor}_{\rho_\mu}\gamma \in T_x^*(Q/G)$$

by

$$(\mathrm{Hor}_{\rho_\mu}\,\gamma)(T_y\rho_\mu\cdot\delta y) = \gamma(\mathrm{Hor}_{\rho_\mu}\,\delta y).$$

This is well defined because the kernel of $T_y\rho_\mu$ consists of vertical vectors and these are annihilated by the map Hor_{ρ_μ}.

Vertical variations. Now we consider vertical variations; that is, we take variations $\delta y(t) = \mathrm{Ver}_{\rho_\mu}\,\delta y(t)$. The left-hand side of the variational principle (III.18) now becomes

$$\delta\int_a^b \mathfrak{R}^\mu(x(t),\dot{x}(t),y(t))dt = \delta\int_a^b [-\mathfrak{C}_\mu(y(t))]dt = \int_a^b [-\mathbf{d}\mathfrak{C}_\mu(y(t))\cdot\mathrm{Ver}_{\rho_\mu}\,\delta y(t)]dt.$$

As before, the right-hand side is $\int_a^b(\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t)))\cdot \mathrm{Ver}_{\rho_\mu}\,\delta y(t)dt$. Hence, the variational principle (III.18) gives

$$\mathrm{Ver}_{\rho_\mu}[\mathbf{d}\mathfrak{C}_\mu(y)+\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))]=0, \tag{III.21}$$

where, for a point $\gamma\in T_y^*(Q/G_\mu)$, we define

$$\mathrm{Ver}_{\rho_\mu}\gamma\in \mathrm{Ver}_y^*(Q/G_\mu)$$

by

$$(\mathrm{Ver}_{\rho_\mu}\,\gamma)=\gamma|\mathrm{Ver}_y(Q/G_\mu).$$

We can rewrite (III.21) to isolate $\mathrm{Ver}_{\rho_\mu}\dot{y}$ as follows:

$$\mathrm{Ver}_{\rho_\mu}(\mathbf{i}_{\mathrm{Ver}_{\rho_\mu}\dot{y}}\beta_\mu(y))=-\mathrm{Ver}_{\rho_\mu}[\mathbf{d}\mathfrak{C}_\mu(y)+\mathbf{i}_{\mathrm{Hor}_{\rho_\mu}\dot{y}(t)}\beta_\mu(y(t))]. \tag{III.22}$$

H. The Lagrange–Routh equations

We now put together the information on the structure of the two-form β_μ with the reduced equations in the previous section.

The horizontal equation. We begin with the horizontal reduced equation:

$$\mathcal{EL}(\mathfrak{L})(\ddot{x})=\mathrm{Hor}_{\rho_\mu}[\mathbf{d}\mathfrak{C}_\mu(y)+\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))]. \tag{III.23}$$

We now compute the term $\mathrm{Hor}_{\rho_\mu}\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))$. To do this, let $\delta x\in T_x(Q/G)$ and write $\delta x=T_y\rho_\mu\cdot\delta y$. By definition,

$$\begin{aligned}\langle \mathrm{Hor}_{\rho_\mu}\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t)),\delta x\rangle&=\langle \mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t)),\mathrm{Hor}_{\rho_\mu}\,\delta y\rangle\\&=\beta_\mu(y(t))(\dot{y}(t),\mathrm{Hor}_{\rho_\mu}\,\delta y)\\&=\beta_\mu(y(t))(\mathrm{Hor}_{\rho_\mu}\dot{y}(t),\mathrm{Hor}_{\rho_\mu}\,\delta y)\\&\quad+\beta_\mu(y(t))(\mathrm{Ver}_{\rho_\mu}\dot{y}(t),\mathrm{Hor}_{\rho_\mu}\,\delta y).\end{aligned} \tag{III.24}$$

Using Lemmas III.11 and III.12, this becomes

$$\langle \mathrm{Hor}_{\rho_\mu}\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t)),\delta x\rangle=\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t))(T_{y(t)}\rho_\mu\cdot(\mathrm{Hor}_{\rho_\mu}\dot{y}(t)),T_{y(t)}\rho_\mu\cdot(\mathrm{Hor}_{\rho_\mu}\,\delta y)) \tag{III.25}$$

$$=\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t))(T_{y(t)}\rho_\mu\cdot(\mathrm{Hor}_{\rho_\mu}\dot{y}(t)),T_{y(t)}\rho_\mu\cdot\delta y), \tag{III.26}$$

since $T_{y(t)}\rho_\mu$ annihilates the vertical component of δy. Next, we claim that

$$T_{y(t)}\rho_\mu\cdot(\mathrm{Hor}_{\rho_\mu}\dot{y}(t))=\dot{x}(t). \tag{III.27}$$

To see this, we start with the definition of $x(t)=\rho_\mu(y(t))$ and use the chain rule to get $\dot{x}(t)=T_{y(t)}\rho_\mu\cdot\dot{y}(t)=T_{y(t)}\rho_\mu\cdot(\mathrm{Hor}_{\rho_\mu}\dot{y}(t))$ since $T_y\rho_\mu$ vanishes on ρ_μ-vertical vectors. This proves the claim. Substituting (III.27) into (III.26) and using $\delta x=T_y\rho_\mu\cdot\delta y$, we get

$$\langle \mathrm{Hor}_{\rho_\mu}\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t)),\delta x\rangle=\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t))(\dot{x}(t),\delta x). \tag{III.28}$$

Therefore,

$$\mathrm{Hor}_{\rho_\mu}\mathbf{i}_{\dot{y}(t)}\beta_\mu(y(t))=\mathbf{i}_{\dot{x}(t)}\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t)). \tag{III.29}$$

Thus, (III.23) becomes

$$\mathcal{EL}(\mathfrak{L})(\ddot{x})=\mathbf{i}_{\dot{x}(t)}\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t))+\mathrm{Hor}_{\rho_\mu}\mathbf{d}\mathfrak{C}_\mu(y). \tag{III.30}$$

The vertical equation. Now we analyze in a similar manner, the vertical equation. We start with

$$\mathrm{Ver}_{\rho_\mu}(\mathbf{i}_{\mathrm{Ver}_{\rho_\mu}\dot{y}}\beta_\mu(y))=-\mathrm{Ver}_{\rho_\mu}[\mathbf{d}\mathfrak{C}_\mu(y)+\mathbf{i}_{\mathrm{Hor}_{\rho_\mu}\dot{y}(t)}\beta_\mu(y(t))]. \tag{III.31}$$

We pair the left-hand side with a vertical vector, $\mathrm{Ver}_{\rho_\mu}\delta y$ and use the definitions to get

$$\langle\mathrm{Ver}_{\rho_\mu}(\mathbf{i}_{\mathrm{Ver}_{\rho_\mu}\dot{y}}\beta_\mu(y)),\mathrm{Ver}_{\rho_\mu}\delta y\rangle=\beta_\mu(y)(\mathrm{Ver}_{\rho_\mu}\dot{y},\mathrm{Ver}_{\rho_\mu}\delta y)=\langle\mu,[\xi,\eta]\rangle=\langle\mathrm{ad}_\xi^*\mu,\eta\rangle \tag{III.32}$$

by Lemma III.13, where $\mathrm{Ver}_{\rho_\mu}\dot{y}=[\xi_Q(q)]$ and $\mathrm{Ver}_{\rho_\mu}\delta y=[\eta_Q(q)]$.

We can interpret this result by saying that the *vertical–vertical component* of β_μ *is given by the negative of the fiberwise coadjoint orbit symplectic form*.

The second term on the right-hand side of (III.31) is zero by Lemma III.12. The first term on the right-hand side of (III.31) paired with $\mathrm{Ver}_{\rho_\mu}\delta y$ is

$$\langle\mathrm{Ver}_{\rho_\mu}\mathbf{d}\mathfrak{C}_\mu(y),\mathrm{Ver}_{\rho_\mu}\delta y\rangle=\langle\mathbf{d}\mathfrak{C}_\mu(y),\mathrm{Ver}_{\rho_\mu}\delta y\rangle=\langle\mathbf{d}\mathfrak{C}_\mu(y),[\eta_Q(q)]\rangle. \tag{III.33}$$

Now define, by analogy with the definition of the momentum map for a cotangent bundle action, a map $\mathfrak{J}:T^*(Q/G_\mu)\to(\mathfrak{g}/\mathfrak{g}_\mu)^*$ by

$$\langle\mathfrak{J}(\alpha_y),[\xi]\rangle=\langle\alpha_y,[\xi_Q(q)]\rangle,$$

where $y=[q]_{G_\mu}=\pi_{Q,G_\mu}(q)$, $\alpha_y\in T_y^*(Q/G_\mu)$, and where $[\xi]\in\mathfrak{g}/\mathfrak{g}_\mu$. Therefore,

$$\langle\mathrm{Ver}_{\rho_\mu}\mathbf{d}\mathfrak{C}_\mu(y),\mathrm{Ver}_{\rho_\mu}\delta y\rangle=\langle\mathfrak{J}(\mathbf{d}\mathfrak{C}_\mu(y)),\eta\rangle. \tag{III.34}$$

From (III.32) and (III.34), the vertical equation (III.31) is equivalent to

$$\mathrm{ad}_\xi^*(\mu)=-\mathfrak{J}(\mathbf{d}\mathfrak{C}_\mu(y)). \tag{III.35}$$

Thus, the reduced variational principle is equivalent to the following system of ***Lagrange–Routh equations***:

$$\mathcal{EL}(\mathfrak{L})(\ddot{x})=\mathbf{i}_{\dot{x}(t)}\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t))+\mathrm{Hor}_{\rho_\mu}\mathbf{d}\mathfrak{C}_\mu(y), \tag{III.36}$$

$$-\mathrm{ad}_\xi^*(\mu)=\mathfrak{J}(\mathbf{d}\mathfrak{C}_\mu(y)), \tag{III.37}$$

where $\mathrm{Ver}_{\rho_\mu}\dot{y}=[\xi_Q(q)]$.

The first equation may be regarded as a second-order equation for $x\in Q/G$ and the second equation is an equation determining the ρ_μ-vertical component of $\dot{y}$. This can also be thought of as an equation for $[\xi]\in\mathfrak{g}/\mathfrak{g}_\mu$ which in turn determines the vertical component of $\dot{y}$. We also think of these equations as the two components of the equations for the evolution in the fiber product $T(Q/G)\times_{Q/G}Q/G_\mu$.

We can also describe the second equation by saying that the equation for $\mathrm{Ver}_{\rho_\mu}\dot{y}$ is Hamiltonian on the fiber relative to the fiberwise symplectic form and with Hamiltonian given by $\mathfrak{C}_\mu$ restricted to that fiber. This can be formalized as follows. Fix a point $x\in Q/G$ and consider the

fiber $\rho_\mu^{-1}(x)$, which is, as we have seen, diffeomorphic to a coadjoint orbit. Consider the vector field X_x on $\rho_\mu^{-1}(x)$ given by $X_x(y)=\mathrm{Ver}_{\rho_\mu}\dot{y}$. Let ω_x^- denote the pullback of $-\beta_\mu$ to the fiber $\rho_\mu^{-1}(x)$. Then we have $\mathbf{i}_{X_x}\omega_x^-=\mathbf{d}(\mathfrak{C}_\mu|\rho_\mu^{-1}(x))$, which just says that X_x is the Hamiltonian vector field on the fiber with Hamiltonian given by the restriction of the amendment function to the fiber.

We summarize what we have proved with the following.

Theorem III.14: *The reduced variational principle is equivalent to the following system of* ***Lagrange–Routh equations:***

$$\mathcal{EL}(\mathfrak{L})(\ddot{x})=\mathbf{i}_{\dot{x}(t)}\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t))+\mathrm{Hor}_{\rho_\mu}\mathbf{d}\mathfrak{C}_\mu(y), \tag{III.38}$$

$$\mathbf{i}_{\mathrm{ver}_{\rho_\mu}\dot{y}}\omega_x^-=\mathbf{d}(\mathfrak{C}_\mu|\rho_\mu^{-1}(x)). \tag{III.39}$$

For Abelian groups (the traditional case of Routh) the second of the Lagrange–Routh equations disappears and the first of these equations can be rewritten as follows. Recall that the reduced Routhian is given by $\mathfrak{R}^\mu=\mathfrak{L}-\mathfrak{C}_\mu$ and in this case, the spaces Q/G and Q/G_μ are identical and the horizontal projection is the identity. Thus, in this case we get

$$\mathcal{EL}(\mathfrak{R}^\mu)(\ddot{x})=\mathbf{i}_{\dot{x}(t)}\mathrm{Curv}_{\mathfrak{A}}^{(y(t),\mu)}(x(t)). \tag{III.40}$$

Note that this form of the equations agrees with the Abelian case of Routh reduction discussed in Marsden and Ratiu [1994],[98] Sec. 8.9 and in Marsden and Scheurle [1993],[107] namely we start with a Lagrangian of the form

$$L(x,\dot{x},\dot{\theta})=\tfrac{1}{2}g_{\alpha\beta}(x)\dot{x}^\alpha\dot{x}^\beta+g_{a\alpha}(x)\dot{x}^\alpha\dot{\theta}^a+\tfrac{1}{2}g_{ab}(x)\dot{\theta}^a\dot{\theta}^b-V(x),$$

where there is a sum over α, β from 1 to m and over a, b from 1 to k. Here, the θ^a are cyclic variables and the momentum map constraint reads $\mu_a=g_{a\alpha}\dot{x}^\alpha+g_{ab}\dot{\theta}^b$. In this case, the components of the mechanical connection are $A_\alpha^a=g^{ab}g_{b\alpha}$, the locked inertia tensor is $I_{ab}=g_{ab}$, and the Routhian is $R^\mu=\frac{1}{2}(g_{\alpha\beta}-g_{a\alpha}g^{ab}g_{b\beta})\dot{x}^\alpha\dot{x}^\beta-V_\mu(x)$, where the amended potential is $V_\mu(x)=V(x)+\frac{1}{2}g^{ab}\mu_a\mu_b$. The Lagrange–Routh equations are

$$\frac{d}{dt}\frac{\partial R^\mu}{\partial\dot{x}^\alpha}-\frac{\partial R^\mu}{\partial x^\alpha}=B_{\alpha\beta}^a\mu_a\dot{x}^\beta \tag{III.41}$$

(with the second equation being trivial; it simply expresses the conservation of μ_a), where, in this case, the components of the curvature are given by

$$B_{\alpha\beta}^a=\frac{\partial A_\beta^a}{\partial x^\alpha}-\frac{\partial A_\alpha^a}{\partial x^\beta}.$$

I. Examples

The rigid body. In this case, the Lagrange–Routh equations reduced only to a coadjoint orbit equation and simply state that the equations are Hamiltonian on the coadjoint orbit. This same statement is true of course for any system with $Q=G$.

The heavy top. In this case, the coadjoint orbit equation is trivial and so the Lagrange–Routh equations reduce to second-order equations for $\mathbf{\Gamma}$ on S^2. These equations are computed to be as follows:

$$\ddot{\mathbf{\Gamma}}=-\|\dot{\mathbf{\Gamma}}\|^2\mathbf{\Gamma}+\mathbf{\Gamma}\times\mathbf{\Sigma},$$

where

$$\mathbf{\Sigma}=b\dot{\mathbf{\Gamma}}-\nu\dot{\mathbf{\Gamma}}+I^{-1}[(I(\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma})+(\nu-b)I\mathbf{\Gamma})\times((\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma})+(\nu-b)\mathbf{\Gamma})+Mgl\mathbf{\Gamma}\times\boldsymbol{\chi}],$$

$b=[(\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma})\cdot I\mathbf{\Gamma}]/(\mathbf{\Gamma}\cdot I\mathbf{\Gamma})$, and $\nu=\mu/(\mathbf{\Gamma}\cdot I\mathbf{\Gamma})$. Notice that one can write $\mathbf{\Sigma}=\mathbf{\Sigma}_0+\mu\mathbf{\Sigma}_1+\mu^2\mathbf{\Sigma}_2$, which represent the three terms in the Lagrange–Routh equations that are independent of μ, linear in μ, and quadratic in μ. In particular, the term linear in μ is the magnetic term:

$$\mathbf{\Sigma}_1=(\mathbf{\Gamma}\cdot I\mathbf{\Gamma})^{-1}[-\dot{\mathbf{\Gamma}}+I^{-1}(I(\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma})\times\mathbf{\Gamma}+I\mathbf{\Gamma}\times(\dot{\mathbf{\Gamma}}\times\mathbf{\Gamma}))].$$

This of course is the ''curvature term'' in the Lagrange–Routh equations. Notice that it is, according to the general theory, linear in the ''velocity'' $\dot{\mathbf{\Gamma}}$. The remaining terms are the Euler–Lagrange expression of the reduced Routhian with those quadratic in the velocity being the differential of the amendment.

IV. RECONSTRUCTION

A. First reconstruction equation

The local formula. For a curve with known constant value of momentum, the evolution of the group variable can be determined from the shape space trajectory. This ***reconstruction equation*** is usually written in a local trivialization $S\times G$ of the bundle $Q\to Q/G$ in the following way. Given a shape space trajectory $x(t)$, the curve $q(t)=(x(t),g(t))$ has momentum μ (i.e., $\mathbf{J}_L(q(t),\dot{q}(t))=\mu$) if and only if $g(t)$ solves the differential equation

$$\dot{g}=g\cdot[\mathfrak{I}_{\mathrm{loc}}(x)^{-1}\mathrm{Ad}_g^*\,\mu-\mathfrak{A}_{\mathrm{loc}}(x)\dot{x}]. \tag{IV.1}$$

Here, $\mathfrak{I}_{\mathrm{loc}}$ is the local representative of the locked inertia tensor and $\mathfrak{A}_{\mathrm{loc}}$ is the local representative of the mechanical connection. This equation is one of the central objects in the study of phases and locomotion and has an analog for nonholonomic systems (see Marsden, Montgomery, and Ratiu [1990][92] and Bloch *et al.* [1996][16]).

The intrinsic equation. We will now write this equation in an intrinsic way without choosing a local trivialization.

Let $x(t)\in S=Q/G$ be a given curve and let μ be a given value of the momentum map. We want to find a curve $q(t)\in Q$ that projects to $x(t)$ and such that its tangent $\dot{q}(t)$ lies in the level set $\mathbf{J}_L^{-1}(\mu)$. We first choose any curve $\tilde{q}(t)\in Q$ that projects to $x(t)$. For example, in a local trivialization, it could be the curve $t\mapsto(x(t),e)$ or it could be the horizontal lift of the base curve. Now we write $q(t)=g(t)\cdot\tilde{q}(t)$.

We shall now make use of the following formula for the derivatives of curves that was given in Eq. (I.2):

$$\dot{q}(t)=(\mathrm{Ad}_{g(t)}\,\xi(t))_Q(q(t))+g(t)\cdot\dot{\tilde{q}}(t),$$

where $\xi(t)=g(t)^{-1}\cdot\dot{g}(t)$. Applying the mechanical connection $\mathfrak{A}$ to both sides, using the identity $\mathfrak{A}(v_q)=\mathfrak{I}(q)^{-1}\cdot\mathbf{J}_L(v_q)$, the fact that $\mathfrak{A}(\eta_Q(q))=\eta$, equivariance of the mechanical connection, and assuming that $\dot{q}(t)\in\mathbf{J}_L^{-1}(\mu)$ gives

$$\mathfrak{I}(q)^{-1}\mu=\mathrm{Ad}_{g(t)}\,\xi(t)+\mathrm{Ad}_{g(t)}\,\mathfrak{A}(\dot{\tilde{q}}(t)).$$

Solving this equation for $\xi(t)$ gives

$$\xi(t)=\mathrm{Ad}_{g(t)^{-1}}\,\mathfrak{I}(q(t))^{-1}\mu-\mathfrak{A}(\dot{\tilde{q}}(t)).$$

Using equivariance of $\mathfrak{I}$ leads to the ***first reconstruction equation***:

$$g(t)^{-1}\dot{g}(t)=\mathfrak{I}(\tilde{q}(t))^{-1}\mathrm{Ad}_{g(t)}^*\,\mu-\mathfrak{A}(\dot{\tilde{q}}(t)). \tag{IV.2}$$

Notice that this reproduces the local equation (IV.1).

Example—The rigid body. In this example, there is no second term in the preceding equation since the bundle has a trivial base, so we choose $\tilde{q}(t)$ to be the identity element. Thus, this reconstruction process amounts to the following equation for the attitude matrix $\mathbf{A}(t)$:

$$\dot{\mathbf{A}}(t)=\mathbf{A}(t)I^{-1}\mathbf{A}(t)^{-1}\boldsymbol{\pi}.$$

This is the method that Whittaker [1907][159] used to integrate for the attitude matrix.

B. Second reconstruction equation

In symplectic reconstruction, one needs only solve a differential equation on the subgroup G_μ instead of on G since the reduction bundle $\mathbf{J}^{-1}(\mu)\to P_\mu=\mathbf{J}^{-1}(\mu)/G_\mu$ is one that quotients only by the subgroup G_μ. See Marsden, Montgomery, and Ratiu [1990][92] for details. This suggests that one can do something similar from the Lagrangian point of view.

Second reconstruction equation. Given a curve $y(t)\in Q/G_\mu$, we find a curve $\overline{q}(t)\in Q$ that projects to $y(t)$. We now write $q(t)=g(t)\cdot\overline{q}(t)$ where $g(t)\in G_\mu$ and require that $\dot{q}(t)\in\mathbf{J}_L^{-1}(\mu)$. Now we use the same formula for derivatives of curves as above and again apply the mechanical connection for the G-action to derive the ***second reconstruction equation***

$$g(t)^{-1}\dot{g}(t)=\mathfrak{I}(\overline{q}(t))^{-1}\mu-\mathfrak{A}(\dot{\overline{q}}(t)). \tag{IV.3}$$

Notice that we have $\mathrm{Ad}^*_{g(t)}\mu=\mu$ since $g(t)\in G_\mu$.

This ***second reconstruction equation*** (IV.3) is now a differential equation on G_μ, which normally would be simpler to integrate than its counterpart equation on G. The reason we are able to get an equation on a smaller group is because we are using more information, namely that of $y(t)$ as opposed to $x(t)$.

The Abelian case. For generic $\mu\in\mathfrak{g}^*$, the subgroup G_μ is Abelian by a theorem of Duflo and Vergne. In this Abelian case, the second reconstruction equation reduces to a quadrature. One has, in fact,

$$g(t)=g(0)\exp\left[\int_0^t(\mathfrak{I}(\overline{q}(s))^{-1}\mu-\mathfrak{A}(\dot{\overline{q}}(s))ds\right]. \tag{IV.4}$$

Example—The rigid body. In the case of the free rigid body, $G=\mathrm{SO}(3)$ and thus if $\boldsymbol{\pi}\neq 0$, we have $G_{\boldsymbol{\pi}}=S^1$, the rotations about the axis $\boldsymbol{\pi}$. The above formula leads to an expression for the attitude matrix that depends only on a quadrature as opposed to nonlinear differential equations to be integrated. The curve $y(t)$ is the body angular momentum $\mathbf{\Pi}(t)$ and the momentum is the spatial angular momentum $\boldsymbol{\pi}$. The curve $\overline{q}(t)$ is the choice of a curve $\overline{\mathbf{A}}(t)$ in SO(3) such that it rotates the vector $\mathbf{\Pi}(t)$ to the vector $\boldsymbol{\pi}$. For example, one can choose this rotation to be about the axis $\mathbf{\Pi}(t)\times\boldsymbol{\pi}$ through the angle given by the angle between the vectors $\mathbf{\Pi}(t)$ and $\boldsymbol{\pi}$. Explicitly,

$$\overline{\mathbf{A}}(t)=\exp\left[\varphi(t)\frac{\widehat{\mathbf{\Pi}(t)\times\boldsymbol{\pi}}}{\|\mathbf{\Pi}(t)\times\boldsymbol{\pi}\|}\right], \tag{IV.5}$$

where $\cos\varphi(t)=\mathbf{\Pi}(t)\cdot\boldsymbol{\pi}/\|\boldsymbol{\pi}\|^2$.

The group element $g(t)$ now is an angle $\alpha(t)$ that represents a rotation about the axis $\boldsymbol{\pi}$ through the angle $\alpha(t)$. Then (IV.4) becomes

$$\begin{aligned}\alpha(t)&=\alpha(0)+\left[\int_0^t\frac{\boldsymbol{\pi}}{\|\boldsymbol{\pi}\|}\cdot(\mathfrak{I}(\overline{\mathbf{A}}(s))^{-1}\boldsymbol{\pi}-\mathfrak{A}(\dot{\overline{\mathbf{A}}}(s))^{\vee})ds\right]\\&=\alpha(0)+\left[\int_0^t\frac{\boldsymbol{\pi}}{\|\boldsymbol{\pi}\|}\cdot(\overline{\mathbf{A}}(s)I^{-1}\overline{\mathbf{A}}(s)^{-1}\boldsymbol{\pi}-[\dot{\overline{\mathbf{A}}}(s)\overline{\mathbf{A}}(s)^{-1}]^{\vee})ds\right].\end{aligned}$$

Some remarks are in order concerning this formula. We have used the hat map and its inverse, the check map, to identify $\mathbb{R}^3$ with $\mathfrak{so}(3)$. In this case, the group elements in S^1 are identified with real numbers, namely, the angles of rotations about the axis $\boldsymbol{\pi}$. Thus, the product in the general formula (IV.4) becomes a sum and the integral over the curve in $\mathfrak{g}_\mu$ becomes an ordinary integral. The integrand at first glance, is an element of $\mathfrak{g}$, but, of course, it actually belongs to $\mathfrak{g}_\mu$. For the example of the rigid body, we make this explicit by taking the inner product with a unit vector along $\boldsymbol{\pi}$.

C. Third reconstruction equation

The second reconstruction equation used the information on a curve $y(t)$ in Q/G_μ as opposed to a curve $x(t)$ in Q/G in order to enable one to integrate on the smaller, often Abelian, group G_μ. However, it still used the mechanical connection associated with the G-action. We can derive yet a third reconstruction equation by using the mechanical connection associated with the G_μ-action.

The momentum map for the G_μ-action on TQ is given by $\mathbf{J}_L^{G_\mu}=\iota_\mu^*\circ\mathbf{J}_L$ where $\iota_\mu:\mathfrak{g}_\mu\to\mathfrak{g}$ is the inclusion and where $\iota_\mu^*:\mathfrak{g}^*\to\mathfrak{g}_\mu^*$ is its dual (the projection, or restriction map). We can also define the locked inertia tensor and mechanical connection for the G_μ-action, in the same way as was done for the G-action. We denote these by

$$\mathfrak{I}^{G_\mu}(q)=\iota_\mu^*\circ\mathfrak{I}(q)\circ\iota_\mu:\mathfrak{g}_\mu\to\mathfrak{g}_\mu^*,\quad \mathfrak{A}^{G_\mu}:TQ\to\mathfrak{g}_\mu.$$

In the third reconstruction equation we organize the logic a little differently and in effect, *take dynamics into account*. Namely, we assume we have a curve $q(t)\in\mathbf{J}_L^{-1}(\mu)$, e.g., a solution of the Euler–Lagrange equations with initial conditions in $\mathbf{J}_L^{-1}(\mu)$. We now let $y(t)\in Q/G_\mu$ be the projection of $q(t)$. We also let $\overline{\mu}=\iota_\mu^*\mu=\mu|\mathfrak{g}_\mu$. We first choose any curve $\overline{q}(t)\in Q$ that projects to $y(t)$. For example, as before, in a local trivialization, it could be the curve $t\mapsto(y(t),e)$ or it could be the horizontal lift of $y(t)$ relative to the connection $\mathfrak{A}^{G_\mu}$. Now we write $q(t)=g(t)\cdot\overline{q}(t)$, where $g(t)\in G_\mu$.

As before, we use the following formula for the derivatives of curves:

$$\dot{q}(t)=(\mathrm{Ad}_{g(t)}\,\xi(t))_Q(q(t))+g(t)\cdot\dot{\overline{q}}(t),$$

where $\xi(t)=g(t)^{-1}\cdot\dot{g}(t)\in\mathfrak{g}_\mu$. Applying the mechanical connection $\mathfrak{A}^{G_\mu}$ to both sides, using the identity $\mathfrak{A}^{G_\mu}(v_q)=\mathfrak{I}^{G_\mu}(q)^{-1}\cdot\mathbf{J}_L^{G_\mu}(v_q)$, the fact that $\mathfrak{A}^{G_\mu}(\eta_Q(q))=\eta$, equivariance of the mechanical connection gives

$$\mathfrak{I}^{G_\mu}(q)^{-1}\overline{\mu}=\mathrm{Ad}_{g(t)}\,\xi(t)+\mathrm{Ad}_{g(t)}\,\mathfrak{A}^{G_\mu}(\dot{\overline{q}}(t)).$$

Solving this equation for $\xi(t)$ gives $\xi(t)=\mathrm{Ad}_{g(t)^{-1}}\,\mathfrak{I}^{G_\mu}(q(t))^{-1}\overline{\mu}-\mathfrak{A}^{G_\mu}(\dot{\overline{q}}(t))$. Using equivariance of $\mathfrak{I}^{G_\mu}$ leads to $g(t)^{-1}\dot{g}(t)=\mathfrak{I}^{G_\mu}(\overline{q}(t))^{-1}\mathrm{Ad}^*_{g(t)}\,\overline{\mu}-\mathfrak{A}^{G_\mu}(\dot{\overline{q}}(t))$, where in the last equation, $\mathrm{Ad}^*_{g(t)}$ is the coadjoint action for G_μ. One checks that $\mathrm{Ad}^*_{g(t)}\,\overline{\mu}=\overline{\mu}$, using the fact that $g(t)\in G_\mu$, so this equation becomes

$$g(t)^{-1}\dot{g}(t)=\mathfrak{I}^{G_\mu}(\overline{q}(t))^{-1}\overline{\mu}-\mathfrak{A}^{G_\mu}(\dot{\overline{q}}(t)). \tag{IV.6}$$

The same remarks as before apply concerning the generic Abelian nature of G_μ applied to this equation. In particular, when G_μ is Abelian, we have the formula

$$g(t)=g(0)\exp\left[\int_0^t(\mathfrak{I}^{G_\mu}(\overline{q}(s))^{-1}\overline{\mu}-\mathfrak{A}^{G_\mu}(\dot{\overline{q}}(s)))ds\right]. \tag{IV.7}$$

Example—The rigid body: Here we start with a solution of the Euler–Lagrange equations $\mathbf{A}(t)$ and we let $\boldsymbol{\pi}$ be the spatial angular momentum and $\boldsymbol{\Pi}(t)$ be the body angular momentum. We

choose the curve $\overline{\mathbf{A}}(t)$ using formula (IV.5). We now want to calculate the angle $\alpha(t)$ of rotation around the axis $\boldsymbol{\pi}$ such that $\mathbf{A}(t)=\mathbf{R}_{\alpha,\boldsymbol{\pi}}\overline{\mathbf{A}}(t)$, where $\mathbf{R}_{\alpha,\boldsymbol{\pi}}$ denotes the rotation about the axis $\boldsymbol{\pi}$ through the angle α. In this case, we get

$$\alpha(t)=\alpha(0)+\left[\int_0^t(\mathfrak{I}^{G_\mu}(\overline{\mathbf{A}}(s))^{-1}\overline{\mu}-\mathfrak{A}^{G_\mu}(\dot{\overline{\mathbf{A}}}(s)))ds\right]. \tag{IV.8}$$

Now we identify $\mathfrak{g}_\mu$ with $\mathbb{R}$ by the isomorphism $a\mapsto a\boldsymbol{\pi}/\|\boldsymbol{\pi}\|$. Then, for $\mathbf{B}\in\mathrm{SO}(3)$,

$$\mathfrak{I}^{G_\mu}(\mathbf{B})=\frac{\boldsymbol{\pi}\cdot(\mathbf{B}I\mathbf{B}^{-1})\boldsymbol{\pi}}{\|\boldsymbol{\pi}\|^2}.$$

Taking $\mathbf{B}=\overline{\mathbf{A}}(s)$, and using the fact that $\overline{\mathbf{A}}(s)$ maps $\mathbf{\Pi}(s)$ to $\boldsymbol{\pi}$, we get

$$\mathfrak{I}^{G_\mu}(\overline{\mathbf{A}}(s))^{-1}=\frac{\|\boldsymbol{\pi}\|^2}{\boldsymbol{\pi}\cdot(\overline{\mathbf{A}}(s)I\overline{\mathbf{A}}(s)^{-1})\boldsymbol{\pi}}=\frac{\|\boldsymbol{\pi}\|^2}{\mathbf{\Pi}(s)\cdot I\mathbf{\Pi}(s)}.$$

The element $\overline{\mu}$ is represented, according to our identifications, by the number $\|\boldsymbol{\pi}\|$, so

$$\mathfrak{I}^{G_\mu}(\overline{\mathbf{A}}(s))^{-1}\overline{\mu}=\frac{\|\boldsymbol{\pi}\|^3}{\mathbf{\Pi}(s)\cdot I\mathbf{\Pi}(s)}.$$

Thus, (IV.8) becomes

$$\alpha(t)=\alpha(0)+\left[\int_0^t\left(\frac{\|\boldsymbol{\pi}\|^3}{\mathbf{\Pi}(s)\cdot I\mathbf{\Pi}(s)}-\mathfrak{A}^{G_\mu}(\dot{\overline{\mathbf{A}}}(s))\right)ds\right]. \tag{IV.9}$$

This formula agrees with that found in Marsden, Montgomery, and Ratiu [1990],[92] Sec. 5.1.2.

D. The vertical Killing metric

For some calculations as well as a deeper insight into geometric phases studied in the next section, it is convenient to introduce a modified metric.

Definition of the vertical Killing metric. First, we assume that the Lie algebra $\mathfrak{g}$ has an inner product which we shall denote $\langle\cdot,\cdot\rangle^\circ$, with the property that $\mathrm{Ad}_g:\mathfrak{g}\to\mathfrak{g}$ is orthogonal for every g. For example, if G is compact, the negative of the Killing form is such a metric. For SO(3), we shall use the standard dot product as this metric. For convenience, we shall refer to the inner product $\langle\cdot,\cdot\rangle^\circ$ as the ***Killing metric***.

Now we use the Killing metric on $\mathfrak{g}$ to define a new metric on Q by using the same horizontal and vertical decomposition given by the mechanical connection of the original (kinetic energy) metric. On the horizontal space we use the given inner product while on the vertical space, we take the inner product of two vertical vectors, say $\xi_Q(q)$ and $\eta_Q(q)$ to be $\langle\xi,\eta\rangle^\circ$. Finally, in the new metric we declare the horizontal and vertical spaces to be orthogonal. These properties define the new metric, which we shall call the ***vertical Killing metric***. This metric has been used by a variety of authors, such as Montgomery [1990, 1991].[120,121] Related-modifications of the kinetic energy metric are used by Bloch, Leonard, and Marsden [1998,1999][18,19,63,65] for the stabilization of relative equilibria of mechanical control systems and we shall denote it $\langle\langle\cdot,\cdot\rangle\rangle^\circ$.

The metric $\langle\langle\cdot,\cdot\rangle\rangle^\circ$ is easily checked to be G-invariant, so we can repeat the previous constructions for it. In particular, since the horizontal spaces are unchanged, the mechanical connection on the bundle $Q\to Q/G$ is identical to what it was before. However, for our purposes, we are more interested in the connection on the bundle $Q\to Q/G_\mu$; here the connections need not be the same.

The mechanical connection in terms of the vertical Killing metric. We now compute the momentum map $\mathbf{J}_\circ$ and the locked inertia tensor $\mathfrak{I}_\circ$ for the metric $\langle\langle\cdot,\cdot\rangle\rangle^\circ$ associated with the G-action on Q. Notice that by construction, the mechanical connection associated with this metric is identical to that for the kinetic energy metric.

First of all, the locked inertia tensor $\mathfrak{I}_\circ(q):\mathfrak{g}\to\mathfrak{g}^*$ is given by

$$\langle\mathfrak{I}_\circ(q)\xi,\eta\rangle=\langle\langle\xi_Q(q),\eta_Q(q)\rangle\rangle^\circ=\langle\xi,\eta\rangle^\circ.$$

In other words, the locked inertia tensor for the vertical Killing metric is simply the map associated with the Killing metric on the Lie algebra.

Next, we compute the momentum map $\mathbf{J}_\circ:TQ\to\mathfrak{g}^*$ associated with the vertical Killing metric. For $\eta\in\mathfrak{g}_\mu$, we have, by definition,

$$\langle\mathbf{J}_\circ(v_q),\eta\rangle=\langle\langle v_q,\eta_Q(q)\rangle\rangle^\circ=\langle\langle\mathrm{Hor}(v_q)+\mathrm{Ver}(v_q),\eta_Q(q)\rangle\rangle^\circ=\langle\mathfrak{A}(v_q),\eta\rangle^\circ,$$

where $\mathfrak{A}$ is the mechanical connection for the G-action.

Notice that these quantities are related by

$$\mathfrak{A}(v_q)=\mathfrak{I}_\circ(q)^{-1}\mathbf{J}_\circ(v_q). \tag{IV.10}$$

It is interesting to compare this with the similar formula (I.1) for $\mathfrak{A}$ using the kinetic energy metric.

The G_μ-connection in the vertical Killing metric. We now compute the momentum map $\mathbf{J}_\circ^{G_\mu}$, the locked inertia tensor $\mathfrak{I}_\circ^{G_\mu}$, and the mechanical connection $\mathfrak{A}_\circ^{G_\mu}$ for the metric $\langle\langle\cdot,\cdot\rangle\rangle^\circ$ and the G_μ-action on Q.

First of all, the locked inertia tensor $\mathfrak{I}_\circ^{G_\mu}(q):\mathfrak{g}_\mu\to\mathfrak{g}_\mu^*$ is given by

$$\langle\mathfrak{I}_\circ^{G_\mu}(q)\xi,\eta\rangle=\langle\langle\xi_Q(q),\eta_Q(q)\rangle\rangle^\circ=\langle\xi,\eta\rangle^\circ.$$

Next, we compute $\mathbf{J}_\circ^{G_\mu}:TQ\to\mathfrak{g}_\mu^*$; for $\eta\in\mathfrak{g}_\mu$, we have

$$\langle\mathbf{J}_\circ^{G_\mu}(v_q),\eta\rangle=\langle\langle v_q,\eta_Q(q)\rangle\rangle^\circ=\langle\langle\mathrm{Hor}(v_q)+\mathrm{Ver}(v_q),\eta_Q(q)\rangle\rangle^\circ=\langle\mathfrak{A}(v_q),\eta\rangle^\circ=\langle\mathrm{pr}_\mu\,\mathfrak{A}(v_q),\eta\rangle^\circ,$$

where $\mathfrak{A}$ is the mechanical connection for the G-action (for either the original or the modified metric) and where $\mathrm{pr}_\mu:\mathfrak{g}\to\mathfrak{g}_\mu$ is the orthogonal projection with respect to the metric $\langle\langle\cdot,\cdot\rangle\rangle^\circ$ onto $\mathfrak{g}_\mu$.

As before, these quantities are related by $\mathfrak{A}_\circ^{G_\mu}(v_q)=\mathfrak{I}_\circ^{G_\mu}(q)^{-1}\mathbf{J}_\circ^{G_\mu}(v_q)$, and so from the preceding two relations, it follows that $\mathfrak{A}_\circ^{G_\mu}(v_q)=\mathrm{pr}_\mu\,\mathfrak{A}(v_q)$.

The connection on the bundle ρ_μ. We just computed the mechanical connection on the bundle $\pi_{Q,G_\mu}:Q\to Q/G_\mu$ associated with the vertical Killing metric. There is a similar formula for that associated with the kinetic energy metric. In particular, it follows that in general, *these two connections are different*. This difference is important in the next section on geometric phases.

Despite this difference, it is interesting to note that each of them induces *the same Ehresmann connection* on the bundle $\rho_\mu:Q/G_\mu\to Q/G$. Thus, in splitting the Lagrange–Routh equations into horizontal and vertical parts, there is no difference between using the kinetic energy metric and the vertical Killing metric.

E. Fourth reconstruction equation

There is yet a fourth reconstruction equation that is based on a different connection. The new connection will be that associated with the vertical Killing metric.

As before, we first choose any curve $\overline{q}(t)\in Q$ that projects to $y(t)$. For example, in a local trivialization, it could be the curve $t\mapsto(y(t),e)$ or it could be the horizontal lift of $y(t)$ relative to

the connection $\mathfrak{A}_\circ^{G_\mu}$. Now we write $q(t)=g(t)\cdot\overline{q}(t)$, where $g(t)\in G_\mu$. Again, we use the following formula for the derivatives of curves:

$$\dot{q}(t)=(\mathrm{Ad}_{g(t)}\,\xi(t))_Q(q(t))+g(t)\cdot\dot{\overline{q}}(t), \tag{IV.11}$$

where $\xi(t)=g(t)^{-1}\cdot\dot{g}(t)\in\mathfrak{g}_\mu$.

Now we assume that $\dot{q}(t)\in\mathbf{J}_L^{-1}(\mu)$ and apply the connection $\mathfrak{A}_\circ^{G_\mu}$ to both sides. The left-hand side of (IV.11) then becomes

$$\mathfrak{A}_\circ^{G_\mu}(\dot{q}(t))=\mathrm{pr}_\mu\,\mathfrak{A}(\dot{q}(t))=\mathrm{pr}_\mu\,\mathfrak{I}(q(t))^{-1}\mathbf{J}_L(\dot{q}(t))=\mathrm{pr}_\mu\,\mathfrak{I}(q(t))^{-1}\mu.$$

The right-hand side of (IV.11) becomes $\mathrm{Ad}_{g(t)}\,\xi(t)+\mathrm{Ad}_{g(t)}\,\mathfrak{A}_\circ^{G_\mu}(\dot{\overline{q}}(t))$. Thus, we have proved that

$$\mathrm{pr}_\mu\,\mathfrak{I}(q(t))^{-1}\mu=\mathrm{Ad}_{g(t)}(\xi(t)+\mathfrak{A}_\circ^{G_\mu}(\dot{\overline{q}}(t))).$$

Solving this equation for $\xi(t)$ and using the fact that $\mathrm{Ad}_{g(t)}$ is orthogonal in the Killing inner product on $\mathfrak{g}$ gives

$$\xi(t)=\mathrm{Ad}_{g(t)^{-1}}[\mathrm{pr}_\mu\,\mathfrak{I}(q(t))^{-1}\mu]-\mathfrak{A}_\circ^{G_\mu}(\dot{\overline{q}}(t))=\mathrm{pr}_\mu[\mathrm{Ad}_{g(t)^{-1}}\,\mathfrak{I}(q(t))^{-1}\mu]-\mathfrak{A}_\circ^{G_\mu}(\dot{\overline{q}}(t)).$$

Using equivariance of $\mathfrak{I}$ leads to the ***fourth reconstruction equation*** for $q(t)=g(t)\cdot\overline{q}(t)\in\mathbf{J}_L^{-1}(\mu)$ given $y(t)\in Q/G_\mu$:

$$g(t)^{-1}\dot{g}(t)=\mathrm{pr}_\mu[\mathfrak{I}(\overline{q}(t))^{-1}\mu]-\mathfrak{A}_\circ^{G_\mu}(\dot{\overline{q}}(t)), \tag{IV.12}$$

where, recall, $\overline{q}(t)$ is any curve in Q such that $[\overline{q}(t)]_{G_\mu}=y(t)$.

When G_μ is Abelian, we have, as with the other reconstruction equations,

$$g(t)=g(0)\exp\left[\int_0^t(\mathrm{pr}_\mu[\mathfrak{I}(\overline{q}(s))^{-1}\mu]-\mathfrak{A}_\circ^{G_\mu}(\dot{\overline{q}}(s)))ds\right]. \tag{IV.13}$$

F. Geometric phases

Once one has formulas for the reconstruction equation, one gets formulas for geometric phases as special cases. Recall that geometric phases are important in a wide variety of phenomena such as control and locomotion generation (see Marsden and Ostrowski [1998][93] and Marsden [1999][104] for accounts and further literature).

The way one proceeds in each case is similar. We consider a *closed curve* $y(t)$ in Q/G_μ, with, say, $0\leqslant t\leqslant T$, and lift it to a curve $q(t)$ according to one of the reconstruction equations in the preceding sections. Then we can write the final point $q(T)$ as $q(T)=g_{\mathrm{tot}}q(0)$, which defines the ***total phase***, g_{tot}. The group element g_{tot} will be in G or in G_μ according to which reconstruction formula is used.

For example, suppose that one uses Eq. (IV.12) with $\overline{q}(t)$ chosen to be the horizontal lift of $y(t)$ with respect to the connection $\mathfrak{A}_\circ^{G_\mu}$ with initial conditions q_0 covering $y(0)$. Then $\overline{q}(T)=g_{\mathrm{geo}}q_0$, where g_{geo} is the holonomy of the base curve $y(t)$. This group element is called the ***geometric phase***. Then we get $q(T)=g_{\mathrm{dyn}}g_{\mathrm{geo}}q(0)$ where $g_{\mathrm{dyn}}=g(T)$, and $g(t)$ is the solution of $g(t)^{-1}\dot{g}(t)=\mathfrak{I}(\overline{q}(t))^{-1}\mu$ in the group G_μ with $g(0)$ the identity. The group element g_{dyn} is often called the ***dynamic phase***. Thus, we have $g_{\mathrm{tot}}=g_{\mathrm{dyn}}g_{\mathrm{geo}}$. Of course in case G_μ is Abelian, this group multiplication is given by addition and the dynamic phase is given by the explicit integral

$$g_{\mathrm{dyn}}=\int_0^T\mathrm{pr}_\mu[\mathfrak{I}(\overline{q}(s))^{-1}\mu]ds.$$

Example: The rigid body. In the case of the rigid body, the holonomy is simply given by the symplectic area on the coadjoint orbit S^2 since the curvature, as we have seen, is, in this case, the symplectic form and since the holonomy is given by the integral of the curvature over a surface bounding the given curve (see, e.g., Kobayashi and Nomizu [1963][75] or Marsden, Montgomery, and Ratiu [1990][92] for this classical formula for holonomy).

We now compute the dynamic phase. Write the horizontal lift as $\overline{\mathbf{A}}$ so that we have, as before, $\mathbf{A}(t)\mathbf{\Pi}(t)=\boldsymbol{\pi}$, $\overline{\mathbf{A}}(t)\mathbf{\Pi}(t)=\boldsymbol{\pi}$ and $\mathbf{A}(t)=\mathbf{R}_{\alpha,\boldsymbol{\pi}}(t)\overline{\mathbf{A}}(t)$.

Now $\mathfrak{I}(\overline{\mathbf{A}}(t))=\overline{\mathbf{A}}(t)I\overline{\mathbf{A}}(t)^{-1}$. Therefore,

$$\mathfrak{I}(\overline{\mathbf{A}}(t))^{-1}\boldsymbol{\pi}=\overline{\mathbf{A}}(t)I^{-1}\overline{\mathbf{A}}^{-1}(t)\boldsymbol{\pi}=\overline{\mathbf{A}}(t)I^{-1}\mathbf{\Pi}(t)=\overline{\mathbf{A}}(t)\mathbf{\Omega}(t).$$

But then

$$\begin{aligned}\mathrm{pr}_{\mu}[\mathfrak{I}(\overline{q}(s))^{-1}\mu]&=\mathrm{pr}_{\boldsymbol{\pi}}[\mathfrak{I}(\overline{\mathbf{A}}(s))^{-1}\boldsymbol{\pi}]=\mathfrak{I}(\overline{\mathbf{A}}(s))^{-1}\boldsymbol{\pi}\cdot\frac{\boldsymbol{\pi}}{\|\boldsymbol{\pi}\|}\\&=\overline{\mathbf{A}}(s)\mathbf{\Omega}(s)\cdot\frac{\boldsymbol{\pi}}{\|\boldsymbol{\pi}\|}=\mathbf{\Omega}\cdot\frac{\overline{\mathbf{A}}(s)^{-1}\boldsymbol{\pi}}{\|\boldsymbol{\pi}\|}=\frac{\mathbf{\Omega}\cdot\mathbf{\Pi}}{\|\boldsymbol{\pi}\|}=\frac{2E}{\|\boldsymbol{\pi}\|},\end{aligned}$$

where E is the energy of the trajectory. Thus, the dynamic phase is given by

$$g_{\mathrm{dyn}}=\frac{2ET}{\|\boldsymbol{\pi}\|},$$

which is the rigid body phase formula of Montgomery [1991b][122] and Marsden, Montgomery, and Ratiu [1990].[92]

V. FUTURE DIRECTIONS AND OPEN QUESTIONS

The Hamiltonian bundle picture. As we have described earlier, on the Lagrangian side, we choose a connection on the bundle $\pi_{Q,G}:Q\to Q/G$ and realize TQ/G as the Whitney sum bundle $T(Q/G)\oplus\tilde{\mathfrak{g}}$ over Q/G. Correspondingly, on the Hamiltonian side we realize T^*Q/G as the Whitney sum bundle $T^*(Q/G)\oplus\tilde{\mathfrak{g}}^*$ over Q/G. The reduced Poisson structure on this space, as we have mentioned already, has been investigated by Montgomery, Marsden, and Ratiu [1984],[115] Montgomery [1986],[118] Cendra, Marsden, and Ratiu [2000a],[28] and Zaalani [1999].[160] See also Guillemin, Lerman, and Sternberg [1996],[41] and references therein.

The results of the present paper on Routh reduction show that on the Lagrangian side, the reduced space $\mathbf{J}_L^{-1}(\mu)/G_\mu$ is $T(Q/G)\times_{Q/G}Q/G_\mu$. This is consistent (by taking the dual of our isomorphism of bundles) with the fact that the symplectic leaves of $(T^*Q)/G$ can be identified with $T^*(Q/G)\times_{Q/G}Q/G_\mu$. The symplectic structure on these leaves has been investigated by Marsden and Perlmutter [1999][95] and Zaalani [1999].[160] It would be interesting to see if the techniques of the present paper shed any further light on these constructions.

In the way we have set things up, we ***conjecture*** that the symplectic structure on $T^*(Q/G)\times_{Q/G}Q/G_\mu$ is the canonical cotangent symplectic form on $T^*(Q/G)$ plus β_μ (that is, the canonical cotangent symplectic form plus $\mathrm{Curv}_{\mathfrak{A}}^{(x,\mu)}$, the (x,μ)-component of the curvature of the mechanical connection, $x\in Q/G$, pulled up from Q/G to $T^*(Q/G)$) plus the coadjoint orbit symplectic form on the fibers.

It would also be of interest to see to what extent one can derive the symplectic (and Poisson) structures directly from the variational principle as boundary terms, as in Marsden, Patrick, and Shkoller [1998].[94]

Singular reduction and bifurcation. We mentioned the importance of singular reduction in Sec. I. However, almost all of the theory of singular reduction is confined to the general symplec-

tic category, with little attention paid to the tangent and cotangent bundle structure. However, explicit examples, as simple as the spherical pendulum (see Lerman, Montgomery, and Sjamaar [1993])[85] show that this cotangent bundle structure together with a ''stitching construction'' is important.

As was mentioned already in Marsden and Scheurle [1993][107] in connection with the double spherical pendulum, it would be interesting to develop the general theory of *singular Lagrangian reduction* using, amongst other tools, the techniques of blow up. In addition, this should be dual to a similar effort for the general theory of symplectic reduction of cotangent bundles. We believe that the general bundle structures in this paper will be useful for this endeavor. The links with bifurcation with symmetry are very interesting; see Golubitsky and Schaeffer [1985],[37] Golubitsky *et al.* [1995],[36] Golubitsky and Stewart [1987],[38] and Ortega and Ratiu [1997],[127] for instance.

Groupoids. There is an approach to Lagrangian reduction using groupoids and algebroids due to Weinstein [1996][157] (see also Martinez [1999][106]). It would of course be of interest to make additional links between these approaches and the present ones.

Quantum systems. The bundle picture in mechanics is clearly important in understanding quantum mechanical systems, and the quantum-classical relationship. For example, the mechanical connection has already proved useful in understanding the relation between vibratory and rotational modes of molecules. This effort really started with Guichardet [1984][40] and Iwai [1987c].[61] See also Iwai [1982, 1985, 1987a].[57–60] Littlejohn and Reinch [1997][89] (and other recent references as well) have carried on this work in a very interesting way. Landsman [1995, 1998][81,82] also uses reduction theory in an interesting way.

Multisymplectic geometry and variational integrators. There have been significant developments in *multisymplectic geometry* that have led to interesting integration algorithms, as in Marsden, Patrick, and Shkoller [1998][94] and Marsden and Shkoller [1999].[109] There is also all the work on *reduction for discrete mechanics* which also takes a variational view, following Veselov [1988].[124,152] These variational integrators have been important in numerical integration of mechanical systems, as in Kane *et al.* [2000],[67] Wendlandt and Marsden [1997],[158] and references therein. Discrete analogs of reduction theory have begun in Ge and Marsden [1988],[35] Marsden, Pekarsky, and Shkoller [1999],[96] and Bobenko and Suris [1998].[21] We expect that one can generalize this theory from the Euler–Poincaré and semidirect product context to the context of general configuration spaces using the ideas of Lagrange–Routh reduction in the present work.

Geometric phases. In this paper we have begun the development of the theory of geometric phases in the Lagrangian context building on work of Montgomery [1985, 1988, 1993][117,119,123] and Marsden, Montgomery, and Ratiu [1990].[92] In fact, the Lagrangian setting also provides a natural setting for averaging which is one of the basic ingredients in geometric phases. We expect that our approach will be useful in a variety of problems involving control and locomotion.

Nonholonomic mechanics. Lagrangian reduction has had a significant impact on the theory of nonholonomic systems, as in Bloch *et al.* [1996][16] and Koon and Marsden [1997a,b,c, 1998].[71–74] The almost symplectic analog was given in Bates and Sniatycki [1993].[14] These references also develop Lagrangian reduction methods in the context of nonholonomic mechanics with symmetry (such as systems with rolling constraints). These methods have also been quite useful in many control problems and in robotics; see, e.g., Bloch and Crouch [1999].[15] One of the main ingredients in these applications is the fact that one no longer gets conservation laws, but rather one replaces the momentum map constraint with a momentum equation. It would be of considerable interest to extend the reduction ideas of the present paper to that context. A Lagrange–d'Alembert–Poincaré reduction theory, the nonholonomic version of Lagrange–Poincaré reduction, is considered in Cendra, Marsden, and Ratiu [2000b].[29]

Stability and block diagonalization. Further connections and development of stability and bifurcation theory on the Lagrangian side (also in the singular case) would also be of interest. Already a start on this program is done by Lewis [1992].[87] Especially interesting would be to reformulate Lagrangian block diagonalization in the current framework. We ***conjecture*** that the structure of the Lagrange–Routh equations given in the present paper is in a form for which block diagonalization is automatically and naturally achieved.

Fluid theories. The techniques of Lagrangian reduction have been very useful in the study of interesting fluid theories, as in Holm, Marsden, and Ratiu [1986, 1998, 1999][52,54,55] and plasma theories, as in Cendra *et al.* [1998],[24] including interesting analytical tools (as in Cantor [1975][23] and Nirenberg and Walker [1973]).[126] Amongst these, the averaged Euler equations are especially interesting; see Marsden, Ratiu, and Shkoller [1999].[99]

Routh by stages. In the text we discussed the current state of affairs in the theory of reduction by stages, both Lagrangian and Hamiltonian. The Lagrangian counterpart of symplectic reduction is of course what we have developed here, namely Lagrange–Routh reduction. Naturally then, the development of this theory for reduction by stages for group extensions would be very interesting.

ACKNOWLEDGMENTS

We thank our many collaborators and students for their help, direct or indirect, with this paper. In particular, we would like to single out Anthony Blaom, Anthony Bloch, Sameer Jalnapurkar, Hans-Peter Kruse, Melvin Leok, Naomi Leonard, Sergey Pekarsky, Matt Perlmutter, Steve Shkoller, Alan Weinstein, and Matt West. Research by J.E.M. was partially supported by the National Science Foundation, the Humboldt Foundation, and the California Institute of Technology. Research by T.S.R. was partially supported by the US and Swiss National Science Foundations and the Humboldt Foundation.

REFERENCES

This order of references, which departs from the usual practice of the Journal of Mathematical Physics, was permitted on an exceptional basis by the editor for the convenience of the readers.

[1] Abarbanel, H. D. I., D. D. Holm, J. E. Marsden, and T. S. Ratiu [1986], ''Nonlinear stability analysis of stratified fluid equilibria,'' Philos. Trans. R. Soc. London, Ser. A **318**, 349–409; see also Richardson number criterion for the nonlinear stability of three-dimensional stratified flow, Phys. Rev. Lett. **52**, 2552–2555 (1984).

[2] Abraham, R. and J. E. Marsden [1978], *Foundations of Mechanics*, 2nd ed. (Addison–Wesley, Reading, MA).

[3] Abraham, R., J. E. Marsden, and T. S. Ratiu [1988], *Manifolds, Tensor Analysis and Applications*, Applied Mathematical Sciences Vol. 75, 2nd ed. (Springer, Berlin).

[4] Alber, M. S., G. G. Luther, J. E. Marsden, and J. M. Robbins [1988], ''Geometric phases, reduction and Lie–Poisson structure for the resonant three-wave interaction,'' Physica D **123**, 271–290.

[5] Arms, J. M., R. H. Cushman, and M. Gotay [1991], ''A universal reduction procedure for Hamiltonian group actions,'' in *The Geometry of Hamiltonian systems*, edited by T. Ratiu (Springer, Berlin), Vol. 22, pp. 33–52.

[6] Arms, J. M., J. E. Marsden, and V. Moncrief [1981], ''Symmetry and bifurcations of momentum mappings,'' Commun. Math. Phys. **78**, 455–478.

[7] Arms, J. M., J. E. Marsden, and V. Moncrief [1982], ''The structure of the space solutions of Einstein's equations. II. Several Killings fields and the Einstein–Yang–Mills equations,'' Ann. Phys. (N.Y.) **144**, 81–106.

[8] Arnold, V. I. [1966a], ''Sur la géométrie differentielle des groupes de Lie de dimenson infinie et ses applications à l'hydrodynamique des fluidsparfaits,'' Ann. Inst. Fourier, Grenoble, **16**, 319–361.

[9] Arnold, V. I. [1966b], ''On an *a priori* estimate in the theory of hydrodynamical stability,'' Izv. Vyssh. Uchebn. Zaved. Mat. Nauk **54**, 3–5; English Translation: Am. Math. Soc. Trans. **79**, 267–269 (1969).

[10] Arnold, V. I. [1966c], ''Sur un principe variationnel pour les découlements stationnaires des liquides parfaits et ses applications aux problemes de stabilité non linéaires,'' J. Mec. **5**, 29–43.

[11] Arnold, V. I. [1988], *Dynamical Systems III*, in Encyclopedia of Mathematics Vol. 3 (Springer, Berlin).

[12] Arnold, V. I. [1989], *Mathematical Methods of Classical Mechanics*, in Graduate Texts in Mathematics Vol. 60, 2nd ed. (Springer, Berlin).

[13] Bates, L. and E. Lerman [1997], ''Proper group actions and symplectic stratified spaces,'' Pac. J. Math. **181**, 201–229.

[14] Bates, L. and J. Sniatycki [1993], ''Nonholonomic reduction,'' Rep. Math. Phys. **32**, 99–115.

[15] Bloch, A. M. and P. E. Crouch [1999], ''Optimal control, optimization, and analytical mechanics,'' in *Mathematical Control Theory* (Peters, New York), pp. 268–321.

[16] Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and R. Murray [1996], ''Nonholonomic mechanical systems with symmetry,'' Arch. Ration. Mech. Anal. **136**, 21–99.

[17] Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and T. S. Ratiu [1996], ''The Euler–Poincaré equations and double bracket dissipation,'' Commun. Math. Phys. **175**, 1–42.

[18] Bloch, A. M., N. Leonard, and J. E. Marsden [1998], ''Matching and stabilization by the method of controlled Lagrangians,'' in Proc. CDC **37**, 1446–1451.

[19] Bloch, A. M., N. Leonard, and J. E. Marsden, ''Controlled Lagrangians and the stabilization of mechanical systems. I. The first matching theorem,'' IEEE Trans. Autom. Control. (to be published).

[20] Bobenko, A. I., A. G. Reyman, and M. A. Semenov-Tian-Shansky [1989], ''The Kowalewski top 99 years later: A Lax pair, generalizations and explicit solutions,'' Commun. Math. Phys. **122**, 321–354.

[21] Bobenko, A. I. and Y. B. Suris [1998], ''Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top'' (unpublished).
[22] Bretherton, F. P. [1970], ''A note on Hamilton's principle for perfect fluids,'' J. Fluid Mech. **44**, 19–31.
[23] Cantor, M. [1975], ''Spaces of functions with asymptotic conditions on $\mathbb{R}^n$,'' Indiana. Univ. Math. J. **24**, 897–902.
[24] Cendra, H., D. D. Holm, M. J. W. Hoyle, and J. E. Marsden [1998], ''The Maxwell–Vlasov equations in Euler–Poincaré form,'' J. Math. Phys. **39**, 3138–3157.
[25] Cendra, H., D. D. Holm, J. E. Marsden, and T. S. Ratiu [1998], ''Lagrangian reduction, the Euler–Poincaré equations and semidirect products,'' Am. Math. Soc. Trans. **186**, 1–25.
[26] Cendra, H., A. Ibort, and J. E. Marsden [1987], ''Variational principal fiber bundles: A geometric theory of Clebsch potentials and Lin constraints,'' J. Geom. Phys. **4**, 183–206.
[27] Cendra, H. and J. E. Marsden [1987], ''Lin constraints, Clebsch potentials and variational principles,'' Physica D **27**, 63–89.
[28] Cendra, H., J. E. Marsden, and T. S. Ratiu [2000a], ''Lagrangian reduction by stages,'' Mem. Am. Math. Soc. (to be published).
[29] Cendra, H., J. E. Marsden, and T. S. Ratiu [2000b], ''Geometric mechanics, Lagrangian reduction and nonholonomic systems,'' in *Mathematics Unlimited* (Springer, New York, to be published).
[30] Chetayev, N. G. [1941], ''On the equations of Poincaré,'' J. Appl. Math. Mech. **5**, 253–262.
[31] Chetayev, N. G. [1961], *The Stability of Motion* (Pergamon, New York).
[32] Chetayev, N. G. [1989], *Theoretical Mechanics* (Springer, Berlin).
[33] Cushman, R. and L. Bates [1997], *Global Aspects of Classical Integrable Systems* (Birkhäuser, Boston).
[34] Ebin, D. G. and J. E. Marsden [1970], ''Groups of diffeomorphisms and the motion of an incompressible fluid,'' Ann. Math. **92**, 102–163.
[35] Ge, Z. and J. E. Marsden [1988], ''Lie–Poisson integrators and Lie–Poisson Hamilton–Jacobi theory,'' Phys. Lett. A **133**, 134–139.
[36] Golubitsky, M., J. E. Marsden, I. Stewart, and M. Dellnitz [1995], ''The constrained Liapunov Schmidt procedure and periodic orbits,'' Fields Inst. Commun. **4**, 81–127.
[37] Golubitsky, M. and D. Schaeffer [1985], *Singularities and Groups in Bifurcation Theory*, Applied Mathematical Sciences Vol. 69 (Springer, Berlin). Vol. 1.
[38] Golubitsky, M. and I. Stewart [1987], ''Generic bifurcation of Hamiltonian systems with symmetry,'' Physica D **24**, 391–405.
[39] Golubitsky, M., I. Stewart, and D. Schaeffer [1988], *Singularities and Groups in Bifurcation Theory*. Applied Mathematical Sciences Vol. 69 (Springer, Berlin), Vol. 2.
[40] Guichardet, A. [1984], ''On rotation and vibration motions of molecules,'' Ann. Inst. Henri Poincaré Phys. Theor. **40**, 329–342.
[41] Guillemin, V., E. Lerman, and S. Sternberg [1996], ''Symplectic fibrations and multiplicity diagrams,'' *LMS Lecture Note Series* (Cambridge University Press, Cambridge).
[42] Guillemin, V. and E. Prato 1990, ''Heckman, Kostant and Steinberg formulas for symplectic manifolds,'' Adv. Math. **82**, 160–179.
[43] Guillemin, V. and S. Sternberg [1978], ''On the equations of motions of a classic particle in a Yang–Mills field and the principle of general covariance,'' Hadronic J. **1**, 1–32.
[44] Guillemin, V. and S. Sternberg [1980], ''The moment map and collective motion,'' Ann. Phys. (Leipzig) **1278**, 220–253.
[45] Guillemin, V. and S. Sternberg [1982], ''Convexity properties of the moment map,'' Invent. Math. **67**, 491–513; **77**, 533–546.
[46] Guillemin, V. and S. Sternberg [1984], *Symplectic Techniques in Physics* (Cambridge University Press, Cambridge).
[47] Hamel, G. [1904], ''Die Lagrange–Eulerschen gleichungen der mechanik,'' Z. Mathematik Physik **50**, 1–57.
[48] Hamel, G. [1949], *Theoretische Mechanik* (Springer, Berlin).
[49] Holm, D. D. and B. A. Kuperschmidt [1983], ''Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas and elasticity,'' Physica D **6**, 347–363.
[50] Holm, D. D. and B. A. Kupershmidt [1983], ''Poisson structures and superconductors,'' Phys. Lett. **93A**, 177–181.
[51] Holmes, P. J. and J. E. Marsden [1983], ''Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups,'' Indiana Univ. Math. J. **32**, 273–310.
[52] Holm, D. D., J. E. Marsden, and T. S. Ratiu [1986], ''The Hamiltonian structure of continuum mechanics in material, spatial and convective representations,'' in *Séminaire de Mathématiques Supérieurs*, 100, 11–122 (Les Presses de L'Univ. de Montréal), Vol. 100, pp. 11–122.
[53] Holm, D. D., J. E. Marsden, and T. S. Ratiu [1998a], ''The Euler–Poincaré equations and semidirect products with applications to continuum theories,'' Adv. Math. **137**, 1–8.
[54] Holm, D. D., J. E. Marsden, and T. S. Ratiu 1998b, ''Euler–Poincaré models of ideal fluids with nonlinear dispersion,'' Phys. Rev. Lett. **349**, 4173–4177.
[55] Holm, D. D., J. E. Marsden, and T. Ratiu ''The Euler–Poincaré equations in geophysical fluid dynamics,'' in *Proceedings of the Isaac Newton Institute Program on the Mathematics of Atmospheric and Ocean Dynamics* (Cambridge University Press, to be published).
[56] Holm, D. D., J. E. Marsden, T. S. Ratiu, and A. Weinstein [1985], ''Nonlinear stability of fluid and plasma equilibria,'' Phys. Rep. **123**, 1–6.
[57] Iwai, T. [1982], ''The symmetry group of the harmonic oscillator and its reduction,'' J. Math. Phys. **23**, 1088–1092.

[58] Iwai, T. [1985], ''On reduction of two degrees of freedom Hamiltonian systems by an S^1 action and $SO_0(1,2)$ as a dynamical group,'' J. Math. Phys. **26**, 885–893.
[59] Iwai, T. [1987a], ''A gauge theory for the quantum planar three-body system,'' J. Math. Phys. **28**, 1315–1326.
[60] Iwai, T. [1987b], ''A geometric setting for internal motions of the quantum three-body system,'' J. Math. Phys. **28**, 1315–1326.
[61] Iwai, T. [1987c], ''A geometric setting for classical molecular dynamics,'' Ann. Inst. Henri Poincaré Phys. Theor. **47**, 199–219.
[62] Iwai, T. [1990], ''On the Guichardet/Berry connection,'' Phys. Lett. A **149**, 341–344.
[63] Jalnapurkar, S. M. and J. E. Marsden, ''Stabilization of relative equilibria. II,'' Regul. Chaotic Dyn. **3**, 161–179.
[64] Jalnapurkar, S. M. and J. E. Marsden [2000a], ''Reduction of Hamilton's variational Principle,'' Dynam. Stability Systems (to be published).
[65] Jalnapurkar, S. M. and J. E. Marsden [2000b], ''Stabilization of Relative Equilibria,'' IEEE Trans. Automat. Control (to be published).
[66] Kazhdan, D., B. Kostant, and S. Sternberg [1978], ''Hamiltonian group actions and dynamical systems of Calogero type,'' Commun. Pure Appl. Math. **31**, 481–508.
[67] Kane, C., J. E. Marsden, M. Ortiz, and M. West [2000], ''Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems,'' Int. J. Num. Math. Eng. (to be published).
[68] Kirk, V., J. E. Marsden, and M. Silber [1996], ''Branches of stable three-tori using Hamiltonian methods in Hopf bifurcation on a rhombic lattice,'' Dyn. Stab. Systems **11**, 267–302.
[69] Kirillov, A. A. [1962], ''Unitary representations of nilpotent Lie groups,'' Russ. Math. Surveys **17**, 53–104.
[70] Kirillov, A. A. [1976], *Elements of the Theory of Representations*, Grundlehren der Math. Wiss. Vol. 220 (Springer, Berlin).
[71] Koon, W. S. and J. E. Marsden [1997a], ''Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction,'' SIAM J. Control Optim. **35**, 901–929.
[72] Koon, W. S. and J. E. Marsden [1997b], ''The geometric structure of nonholonomic mechanics,'' in Proc. CDC **36**, 4856–4862.
[73] Koon, W. S. and J. E. Marsden [1997c], ''The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems,'' Rep. Math. Phys. **40**, 21–62.
[74] Koon, W. S. and J. E. Marsden [1998], ''The Poisson reduction of nonholonomic mechanical systems,'' Rep. Math. Phys. **42**, 101–134.
[75] Kobayashi, S. and K. Nomizu [1963], *Foundations of Differential Geometry* (Wiley, New York).
[76] Kouranbaeva, S. [1999], ''The Camassa–Holm equation as a geodesic flow on the diffeomorphism group,'' J. Math. Phys. **40**, 857–868.
[77] Kummer, M. [1981], ''On the construction of the reduced phase space of a Hamiltonian system with symmetry,'' Indiana Univ. Math. J. **30**, 281–291.
[78] Kummer, M. [1990], ''On resonant classical Hamiltonians with n frequencies,'' J. Diff. Eqns. **83**, 220–243.
[79] Kuperschmidt, B. A. and T. Ratiu [1983], ''Canonical maps between semidirect products with applications to elasticity and superfluids,'' Commun. Math. Phys. **90**, 235–250.
[80] Lagrange, J. L. [1788], *Mécanique Analytique* (Chez la Veuve Desaint).
[81] Landsman, N. P. [1995], ''Rieffel induction as generalized quantum Marsden–Weinstein reduction,'' J. Geom. Phys. **15**, 285–319.
[82] Landsman, N. P. [1998], ''Mathematical topics between classical and quantum mechanics,'' *Springer Monographs in Mathematics* (Springer, New York).
[83] Le, H., and D. G. Kendall [1993], ''The Riemannian structure of Euclidean shape spaces: A novel environment for statistics,'' Ann. Statistics **21**, 1225–1271.
[84] Leonard, N. E. and J. E. Marsden [1997], ''Stability and drift of underwater vehicle dynamics: Mechanical systems with rigid motion symmetry,'' Physica D **105**, 130–162.
[85] Lerman, E., R. Montgomery, and R. Sjamaar [1993], ''Examples of singular reduction, in *Symplectic Geometry*,'' London Mathematical Society Lecture Note Series Vol. 192 (Cambridge University Press, Cambridge), pp. 127–155.
[86] Leonard, N. E. [1997], ''Stability of a bottom-heavy underwater vehicle,'' Automatica J. IFAC **33**, 331–346.
[87] Lewis, D. [1992], ''Lagrangian block diagonalization,'' Dyn. Diff. Eqns. **4**, 1–42.
[88] Lie, S. [1890], *Theorie der Transformationsgruppen* (Zweiter Abschnitt, Teubner).
[89] Littlejohn, R. and M. Reinch [1997], ''Gauge fields in the separation of rotations and internal motions in the n-body problem,'' Rev. Mod. Phys. **69**, 213–275.
[90] Marsden, J., G. Misiolek, M. Perlmutter, and T. Ratiu [1998], ''Symplectic reduction for semidirect products and central extensions,'' Diff. Geom. Appl. **9**, 173–212.
[91] Marsden, J. E., G. Misiolek, M. Perlmutter, and T. S. Ratiu [2000], ''Reduction by stages and group extensions'' (preprint).
[92] Marsden, J. E., R. Montgomery, and T. S. Ratiu [1990], ''Reduction, symmetry and phases in mechanics,'' *Memoirs* (American Mathematical Society, Providence, RI), Vol. 436.
[93] Marsden, J. E. and J. Ostrowski [1998], ''Symmetries in Motion: Geometric Foundations of Motion Control,'' Nonlinear Sci. Today (http://link.springer-ny.com).
[94] Marsden, J. E., G. W. Patrick, and S. Shkoller [1998], ''Multisymplectic geometry, variational integrators and Nonlinear PDEs,'' Commun. Math. Phys. **199**, 351–395.
[95] Marsden, J. E. and M. Perlmutter [1999], ''The orbit bundle picture of cotangent bundle reduction'' Comptes Rendu, Roy. Soc. Canada (to appear).

[96] Marsden, J. E., S. Pekarsky, and S. Shkoller [1999], ''Discrete Euler–Poincaré and Lie–Poisson equations,'' Nonlinearity **12**, 1647–1662.
[97] Marsden, J. E. and T. Ratiu [1986], ''Reduction of Poisson Manifolds,'' Lett. Math. Phys. **11**, 161–170.
[98] Marsden, J. E. and T. S. Ratiu [1994], *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics Vol 17, 2nd ed. (Springer, New York, 1999).
[99] Marsden, J. E., T. Ratiu, and S. Shkoller, ''The geometry and analysis of the averaged Euler equations and a new diffeomorphism group,'' Geom. Funct. Anal. (to be published).
[100] Marsden, J. E., T. S. Ratiu, and A. Weinstein [1984a], ''Semi-direct products and reduction in mechanics,'' Trans. Am. Math. Soc. **281**, 147–177.
[101] Marsden, J. E., T. S. Ratiu, and A. Weinstein [1984b], ''Reduction and Hamiltonian structures on duals of semidirect product Lie Algebras,'' Contemp. Math. **28**, 55–100.
[102] Marsden, J. E. [1982], ''A group theoretic approach to the equations of plasma physics,'' Can. Math. Bull. **25**, 129–142.
[103] Marsden, J. E. [1992], Lectures on Mechanics, London Mathematical Society Lecture Note Series Vol. 174 (Cambridge University Press, Cambridge).
[104] Marsden, J. E. [1999], ''Park City lectures on mechanics, dynamics and symmetry,'' in *Symplectic Geometry and Topology*, edited by Y. Eliashberg and L. Traynor, IAS/Park City Mathematical Series Vol. 7 (American Mathematical Society, Providence, RI).
[105] Martin, J. L. [1959], ''Generalized classical dynamics and the 'classical analogue' of a Fermi oscillation,'' Proc. R. Soc. London, Ser. A **251**, 536.
[106] Martinez, E. [1999], ''Lagrangian mechanics on Lie algebroids'' (preprint).
[107] Marsden, J. E. and J. Scheurle [1993a], ''Lagrangian reduction and the double spherical pendulum,'' ZAMP **44**, 17–43.
[108] Marsden, J. E. and J. Scheurle [1993b], ''The Lagrange–Poincaré equations,'' Fields Inst. Commun. **1**, 139–164.
[109] Marsden, J. E. and S. Shkoller [1999], ''Multisymplectic geometry, covariant Hamiltonians and water waves,'' Math. Proc. Cambridge Philos. Soc. **125**, 553–575.
[110] Marsden, J. E. and A. Weinstein [1974], ''Reduction of symplectic manifolds with symmetry,'' Rep. Math. Phys. **5**, 121–130.
[111] Marsden, J. E. and A. Weinstein [1982], ''The Hamiltonian structure of the Maxwell–Vlasov equations,'' Physica D **4**, 394–406.
[112] Marsden, J. E. and A. Weinstein [1983], ''Coadjoint orbits, vortices and Clebsch variables for incompressible fluids,'' Physica D **7**, 305–323.
[113] Marsden, J. E., A. Weinstein, T. S. Ratiu, R. Schmid, and R. G. Spencer [1983], ''Hamiltonian systems with symmetry, coadjoint orbits and plasma physics,'' in *Proceedings IUTAM-IS1MM Symposium on Modern Developments in Analytical Mechanics, Torino, 1982* (Atti della Acad. della Sc. di Torino), Vol. 117, pp. 289–340.
[114] Meyer, K. R. [1973], ''Symmetries and integrals in mechanics,'' in *Dynamical Systems*, edited by M. Peixoto (Academic, New York), pp. 259–273.
[115] Montgomery, R., J. E. Marsden, and T. S. Ratiu [1984], ''Gauged Lie–Poisson structures,'' Contemp. Math. **28**, 101–114.
[116] Montgomery, R. [1984], ''Canonical formulations of a particle in a Yang–Mills field,'' Lett. Math. Phys. **8**, 59–67.
[117] Montgomery, R. [1985], ''Analytic proof of chaos in the Leggett equations for superfluid ^{3}He,'' J. Low Temp. Phys. **58**, 417–453.
[118] Montgomery, R. [1986], ''The bundle picture in mechanics,'' Ph.D. thesis, University of California, Berkeley.
[119] Montgomery, R. [1988], ''The connection whose holonomy is the classical adiabatic angles of Hannay and Berry and its generalization to the non-integrable case,'' Commun. Math. Phys. **120**, 269–294.
[120] Montgomery, R. [1990], ''Isoholonomic problems and some applications,'' Commun. Math. Phys. **128**, 565–592.
[121] Montgomery, R. [1991a], ''Optimal control of deformable bodies and its relation to gauge theory,'' in *The Geometry of Hamiltonian Systems*, edited by T. Ratiu (Springer, New York).
[122] Montgomery, R. [1991b], ''How much does a rigid body rotate? A Berry's phase from the 18th century,'' Am. J. Phys. **59**, 394–398.
[123] Montgomery, R. [1993], ''Gauge theory of the falling cat,'' Fields Inst. Commun. **1**, 193–218.
[124] Moser, J. and A. P. Veselov [1991], ''Discrete versions of some classical integrable systems and factorization of matrix polynomials,'' Commun. Math. Phys. **139**, 217–243.
[125] Nambu, Y. [1973], ''Generalized Hamiltonian dynamics,'' Phys. Rev. D **7**, 2405–2412.
[126] Nirenberg, L. and H. F. Walker [1973], ''The null spaces of elliptic partial differential operators in $\mathbb{R}^n$,'' J. Math. Anal. Appl. **42**, 271–301.
[127] Ortega, J.-P. and T. S. Ratiu [1997], ''Persistence and smoothness of critical relative elements in Hamiltonian systems with symmetry,'' C. R. Acad. Sci. Sér. I; Math **325**, 1107–1111.
[128] Ortega, J.-P. and T. S. Ratiu [2000], ''*Hamiltonian Singular Reduction*,'' Progress in Math (Birkhäuser, to appear).
[129] Ortega, J.-P. [1998], ''Symmetry, reduction and stability in Hamiltonian systems,'' Ph.D. thesis, University of California, Santa Cruz.
[130] Otto, M. [1987], ''A reduction scheme for phase spaces with almost Kähler symmetry. Regularity results for momentum level sets,'' J. Geom. Phys. **4**, 101–118.
[131] Pauli, W. [1953], ''On the Hamiltonian structure of non-local field theories,'' Nuovo Cimento **X**, 648–667.
[132] Pedroni, M. [1995], ''Equivalence of the Drinfelćd–Sokolov reduction to a bi-Hamiltonian reduction,'' Lett. Math. Phys. **35**, 291–302.
[133] Poincaré, H. [1901], ''Sur une forme nouvelle des équations de la méchanique,'' C. R. Acad. Sci. **132**, 369–371.
[134] Poincaré, H. [1910], ''Sur la precession des corps deformables,'' Bull. Astron. **27**, 321–356.

[135] Ratiu, T. S. [1980a], ''The Euler–Poisson equations and integrability,'' Ph.D. thesis, University of California at Berkeley.
[136] Ratiu, T. S. [1980b], ''The motion of the free *n*-dimensional rigid body,'' Indiana Univ. Math. J. **29**, 609–629.
[137] Ratiu, T. S. [1980c], ''Involution theorems,'' in *Geometric Methods in Mathematical Physics*, edited by G. Kaiser and J. Marsden, Springer Lecture Notes Vol. 775 (Springer, New York), pp. 219–257.
[138] Ratiu, T. S. [1981], ''Euler–Poisson equations on Lie algebras and the *N*-dimensional heavy rigid body,'' Proc. Natl. Acad. Sci. USA **78**, 1327–1328.
[139] Ratiu, T. S. [1982], ''Euler–Poisson equations on Lie algebras and the *N*-dimensional heavy rigid body,'' Am. J. Math. **104**, 409–448; **104**, 1337.
[140] Routh, E. J. [1860], *Treatise on the Dynamics of a System of Rigid Bodies* (MacMillan, London).
[141] Routh, E. J. [1877], *Stability of a Given State of Motion* [Halsted, New York; reprinted in *Stability of Motion* (1975), edited by A. T. Fuller].
[142] Routh, E. J. [1884], *Advanced Rigid Dynamics* (MacMillian, London).
[143] Satzer, W. J. [1977], ''Canonical reduction of mechanical systems invariant under Abelian group actions with an application to celestial mechanics,'' Ind. Univ. Math. J. **26**, 951–976.
[144] Seliger, R. L. and G. B. Whitham [1968], ''Variational principles in continuum mechanics,'' Proc. R. Soc. London, Ser. A **305**, 1–25.
[145] Simo, J. C., D. R. Lewis, and J. E. Marsden [1991], ''Stability of relative equilibria. I. The reduced energy momentum method,'' Arch. Ration. Mech. Anal. **115**, 15–59.
[146] Sjamaar, R. and E. Lerman [1991], ''Stratified symplectic spaces and reduction,'' Ann. Math. **134**, 375–422.
[147] Smale, S. [1970], ''Topology and Mechanics,'' Invent. Math. **10**, 305–331; **11**, 45–64.
[148] Souriau, J. M. [1970], *Structure des Systemes Dynamiques* (Dunod, Paris).
[149] Sternberg, S. [1977], ''Minimal coupling and the symplectic mechanics of a classical particle in the presence of a Yang–Mills field,'' Proc. Natl. Acad. Sci. USA **74**, 5253–5254.
[150] Sudarshan, E. C. G. and N. Mukunda [1974], *Classical Mechanics: A Modern Perspective* (Wiley, New York; 2nd ed., Krieber, Melbourne-FL, 1983).
[151] Vanhaecke, P. [1996], *Integrable Systems in the Realm of Algebraic Geometry*, Lect. Notes Math. **1638** (Springer-Verlag, New York).
[152] Veselov, A. P. [1988], ''Integrable discrete-time systems and difference operators,'' Funct. Anal. Appl. **22**, 83–94.
[153] Vinogradov, A. M. and B. A. Kuperschmidt [1977], ''The structures of Hamiltonian mechanics,'' Russ. Math. Surv. **32**, 177–243.
[154] Weinstein, A. [1977], *Lectures on symplectic manifolds*, CBMS Regional Conference Series in Mathematics, Vol. 29 (American Mathematical Society Providence, RI).
[155] Weinstein, A. [1978], ''A universal phase space for particles in Yang–Mills fields,'' Lett. Math. Phys. **2**, 417–420.
[156] Weinstein, A. [1983], ''Sophus Lie and symplectic geometry,'' Expo. Math. **1**, 95–96.
[157] Weinstein, A. [1996], ''Lagrangian mechanics and groupoids,'' Fields Inst. Commun. **7**, 207–231.
[158] Wendlandt, J. M. and J. E. Marsden [1997], ''Mechanical integrators derived from a discrete variational principle,'' Physica D **106**, 223–246.
[159] Whittaker, E. T. [1907], *A Treatise on the Analytical Dynamics of Particles and Rigid bodies* (Cambridge University Press, Cambridge, 1938 4th ed.; Reprinted by Dover 1944 and Cambridge University 1988).
[160] Zaalani, N. [1999], ''Phase space reduction and Poisson structure,'' J. Math. Phys. **40**, 3431–3438.