Physica 27D (1987) 63–89
North-Holland, Amsterdam

LIN CONSTRAINTS, CLEBSCH POTENTIALS AND VARIATIONAL PRINCIPLES

Hernan CENDRA*
Universidad Nacional del sur, Departimento de Matemática, Ave. Alem 1253, 8000 Bahia Blanca, Argentina

and

Jerrold E. MARSDEN‡
Department of Mathematics, University of California, Berkeley, CA 94720, USA

The Poisson bracket formulation of fluid, plasma and rigid body type systems has undergone considerable recent development using techniques of symmetry group reduction. The relationship between this approach and that using Lin constraints and Clebsch potentials is established. The connection is made in the setting of abstract Clebsch variables as well as that of variational principles on reduced spaces. Variational principles for both the Clebsch and reduced form (such as fluids in spatial representation) are derived from the standard variational principle of Hamilton in material (Lagrangian) representation using reduction theory.

Received 1 August 1986
Revised manuscript received 4 November 1986

1. Introduction

Variational principles in continuum mechanics have a complex history, going back at least to Walter [1; 1868] and Kirchhoff [2; 1876]. For a historical survey with many references up to 1960. see Truesdell and Toupin [3] and Serrin [4]. In material representation the appropriate variational principles are essentially the same as the classical variational principles of mechanics. However, since the equations in spatial or convective representation are not in canonical Hamiltonian form, difficulties arise. This point was investigated by Lin [5] and applied to many examples in the work of Seliger and Whitham [6]; this approach is closely related to the introduction of Clebsch potentials (Clebsch [7; 1857, 1859] also had variational principles in view). In Seliger and Whitham [6; p. 6] it is stated that "Lin's device still remains somewhat mysterious from a strictly mathematical view". We hope the results of this paper will help to answer this type of concern by providing a suitable abstract and precise framework.

In recent years a Hamiltonian formulation of continuum systems written entirely in spatial representation and its relationship with the material and Clebsch representations has been established. Some of the relevant references are Iwinski and Turski [8], Dzyolshinski and Volovick [9], Morrison and Green [10], Marsden and Weinstein [11, 12], Holm and Kuperschmit [13], Marsden et al. [14], Benjamin [15], and Marsden, Ratiu and Weinstein [16, 17].

The purpose of this paper is to formulate precise variational principles in two contexts:

a) an *abstract Clebsch variable setting*, and

b) a *variational formulation on reduced spaces*, representing an abstraction of variational principles in spatial representation.

*Present address: Department of Mathematics, University of California, Santa Cruz, CA 95064, USA.
‡Research partially supported by DOE contract DE-AT-085-ER12097.

In both cases, *the validity of the relevant variational principle is not postulated in an ad hoc way but rather is derived from the standard canonical variational principles in material representation using the geometric methods of reduction and canonical maps*. This provides a coherent setting for the variational principles in common use as well as giving a single abstract proof for many diverse principles. In addition we give new direct variational principles in the spatial representation which are a natural outgrowth of this approach.

In section 2, we give the abstract variational principle for Lin constraints and Clebsch potentials for systems whose configuration space is a Lie group G. The idea is as follows. Let L: $TG \to \mathbb{R}$ be a left invariant Lagrangian, and ρ: $G \times E \to E$ be a linear representation. Elements $a \in E$ and $b \in E^*$ are called *Clebsch potentials*. Given curves $g \in \Omega(G)$ (the space of curves on G), $a \in \Omega(E)$ and $b \in \Omega(E^*)$, we will prove later that

$$\int_{t_1}^{t_2} \left[L(g, g') + \left\langle \frac{\mathrm{d}\rho(g, a)}{\mathrm{d}t}, \rho^*(g, b) \right\rangle \right] \mathrm{d}t = \int_{t_1}^{t_2} \left[L(g, g') + \langle a' + v_E(a), b \rangle \right] \mathrm{d}t,$$

where $v \in \Omega(\mathfrak{g})$ is the body expression of the velocity and v_E is the infinitesimal generator of the action ρ. Lin constraints as defined in this paper (see Definition 2.2) are the submanifolds of $\Omega(\mathfrak{g}) \times \Omega(E)$ defined by the conditions: $a' + v_E(a) = 0$, $a(t_1) = a_1$, $a(t_2) = a_2$, with given $a_1, a_2 \in E$. We will prove that under certain conditions on ρ, a_1, a_2, critical points (v, a) of the last integral, with these constraints, are such that the component v is a solution to the equations of motion. The rigid body and homogeneous incompressible fluids are presented as examples.

Section 3 generalizes the results of section 2 to systems whose configuration space is $G \times V$ where V is a vector space. Compressible flow is given as an example. Section 4 generalizes this situation further to the case where the configuration space is $G \times V \times W$, where V and W are vector spaces, G acts trivially on W but nontrivially on V. Section 5 deals with variational principles on reduced spaces. Here the context is the reduction setting of Marsden and Weinstein [18]. On an exact symplectic manifold $(P, \omega = -\mathrm{d}\theta)$ one normally forms a variational principle in the usual sense of Hamilton's principle:

$$\delta \int_{t_1}^{t_2} [\theta - H\,\mathrm{d}t] = 0.$$

Reduced spaces, such as those for rigid body motion have symplectic forms which are not exact. Nevertheless, we shall show that this variational principle still makes sense. This section sets up the abstract results first that are suitable for examples like the rigid body (where the reduced space is a sphere in body angular momentum space) and incompressible fluids (where the reduced space consists of "isovortical surfaces"). Then a more general reduction result is given. Finally the end of the section links the approach of reduced variational principles with the Clebsch–Lin approach using the idea of a dual pair.

While a number of examples are treated in this paper to illustrate the theory, we do not attempt an exhaustive treatment of them. However, from the examples presented, it is clear that one could treat many others as well. For example, one could do plasmas (see Low [19]), superconductors and superfluids (see Holm and Kuperschmit [20, 13]), elasticity (see Seliger and Whitham [6], Holm and Kuperschmit [13], Marsden, Ratiu, and Weinstein [16, 17], Krishnaprasad and Marsden [21], and Simo, Marsden and Krishnaprasad [22]), stratified flow (see Long [23], and Abarbanel, Holm, Marsden, and Ratiu [24]), free boundary problems (see Lewis, Marsden, Montgomery, and Ratiu [25]) and relativistic systems, including fluids, plasmas, and MHD (see Tam [26], Tam and O'Hanlon [27], Bao, Marsden and Walton [28], and

Holm [29]) and references therein. A few additional examples are treated in a slightly expanded version of the present paper, which we shall provide on request. We also remark that one can generalize many of the constructions in this paper to the case of systems on principal bundles in the context of Montgomery, Marsden, and Ratiu [30]. This is the planned subject of a future paper (of Cendra, Ibort and Marsden [31]).

2. Clebsch potentials and Lin constraints

Lin [5] pointed out that for ideal compressible flow, a variational principle with the Lagrangian density given by the kinetic energy minus the potential energy,

$$L(v, \rho, \sigma) = \tfrac{1}{2}\rho\|v\|^2 - \rho w(\rho, \sigma), \tag{2.1}$$

where v is the Eulerian velocity, ρ is the density, σ is the entropy, w is the internal energy, and with constraints given by conservation of mass and entropy, gives a velocity field of the form

$$v = \nabla\varphi + \eta\nabla\sigma. \tag{2.2}$$

This is an unreasonable restriction, since for isentropic flow, it implies that v is irrotational. To avoid this difficulty, Lin gave a variational principle in terms of the Clebsch representation,

$$v = \nabla\varphi + \eta\nabla\sigma + \boldsymbol{\beta}\cdot\nabla\boldsymbol{\alpha}. \tag{2.3}$$

He imposed the condition that the $\mathbb{R}^3$-valued function $\boldsymbol{\alpha}$ be a function of the original position of the particles only (which amounts to assuming that $\mathrm{d}\boldsymbol{\alpha}/\mathrm{d}t = \boldsymbol{0}$) as a further constraint with the Lagrange multiplier $\boldsymbol{\beta}$.

These ideas of Lin were extended and applied to several cases in continuum mechanics by Seliger and Whitham [6]. For the case of a fluid in a simply connected region, they used the pressure as a Lagrangian density regarded as a functional of the other variables, to produce the desired equations of motion. A Clebsch representation of the velocity of the type (2.3), with α and β real valued rather than $\mathbb{R}^3$-valued appears among the equations.

In this section, by using the geometric approach to Clebsch variables given in Marsden and Weinstein [12], we will generalize those ideas from fluid mechanics to a general abstract context, thereby making them useful in other examples of physical interest. Our approach also clarifies how the variational principles written in terms of Eulerian quantities and Clebsch potentials, like those considered in Seliger and Whitham's paper, are reformulations of the standard Hamilton's principle written in terms of material (Lagrangian) coordinates. We do this using a systematic reduction procedure that takes into account the system's symmetry, such as particle relabelling symmetry for fluid mechanics. A result of this procedure is a sufficient condition (assumption 2.3 below) which ensures existence of a Clebsch potential representation of the type considered in Seliger and Whitham [6].

To start, we recall that according to Marsden and Weinstein [12], Clebsch potentials are given by momentum mappings $J\colon S \to \mathfrak{g}^*$, where S is a symplectic manifold on which a Lie group G acts by symplectic diffeomorphisms and $\mathfrak{g}$ is the Lie algebra of G. We are interested in the following particular case. Let E be a vector space and $\rho\colon G \times E \to E$ be a left (or for some examples a right) representation of G on E. [For the fluid mechanics example above, G will be the group of diffeomorphisms of the fluid particles and E will be a vector space of α's.] The elements of E will be denoted by "a" and the representation corresponding to a group element g acting on a will be denoted $\rho(g,a)=\rho_g(a)=\rho_a(g)=ga$. Let E^* be a space in duality with E and $\langle\ ,\ \rangle$: $E\times E^*\to\mathbb{R}$ be the associated pairing. For each $g\in G$, $\rho^*_{g^{-1}}$: $E^*\to E^*$ is the dual of the linear isomorphism $\rho_{g^{-1}}$: $E\to E$. Since E is a vector space, $T^*E=E\times E^*$ and the map $(\rho_g, \rho^*_{g^{-1}}$ is the cotangent lift of ρ_g for each $g\in G$ and therefore is a symplectic map (Abraham and Marsden [32; p. 180]). If we write $\mathrm{T}\mathrm{T}^*E=E\times E\times E^*\times E^*$, the canonical form θ_0 on T^*E becomes

$$\theta_0(a,\dot a,b,\dot b)=\langle\dot a,b\rangle \tag{2.4}$$

and the momentum mapping induced by the action is

$$J(a,b)(\xi)=\langle\xi_E(a),b\rangle, \tag{2.5}$$

where ξ_E is the infinitesimal generator of the action of G on E and $\xi\in\mathfrak{g}$ (see Abraham and Marsden [32; p. 283]). The body coordinate map B: $\mathrm{T}G\to\mathfrak{g}$ is defined by

$$B(v_g)=\mathrm{T}_gL_{g^{-1}}v_g=v, \tag{2.6}$$

where L_g is left translation by g on G.

For a given manifold M and c_1 and $c_2\in M$, we write

$$\begin{aligned}
&\Omega(M)=\mathrm{C}^\infty([t_1,t_2],M),\\
&\Omega_{c_1}(M)=\{c\in\Omega(M)\,|\,c(t_1)=c_1\},\\
&\Omega_{c_1,c_2}(M)=\{c\in\Omega(M)\,|\,c(t_1)=c_1 \text{ and } c(t_2)=c_2\}.
\end{aligned} \tag{2.7}$$

Also, if f: $M\to N$ is a C^∞ map of manifolds, we define

$$\Omega(f)\colon \Omega(M)\to\Omega(N) \text{ by } \Omega(f)(c)=f\circ c. \tag{2.8}$$

Lemma 2.1. Let $g\in\Omega(G)$, $a\in\Omega(E)$, $b\in\Omega(E^*)$ and write $v=B(\mathrm{d}g/\mathrm{d}t)$. Then

$$\begin{aligned}
\left\langle\frac{\mathrm{d}\rho(g,a)}{\mathrm{d}t},\rho^*(g^{-1},b)\right\rangle&=\langle\dot a+v_E(a),b\rangle\\
&=\theta_0(a,a,b,b)+J(a,b)(v).
\end{aligned} \tag{2.9}$$

Proof. The last equality is a consequence of the expressions (2.4) and (2.5) for θ_0 and J. From the identity

$$\frac{\partial\rho(g,a)}{\partial g}\frac{\mathrm{d}g}{\mathrm{d}t}=\frac{\partial\rho(g,a)}{\partial g}\cdot\mathrm{T}_eL_gv=\rho_g\left[\frac{\partial\rho(e,a)}{\partial g}\cdot v\right]=\rho_g(v_E(a))$$

we get

$$\begin{aligned}
\left\langle\frac{\mathrm{d}\rho(g,a)}{\mathrm{d}t},\rho^*(g^{-1},b)\right\rangle&=\left\langle\frac{\partial\rho(g,a)}{\partial g}\frac{\mathrm{d}g}{\mathrm{d}t}+\rho_g(\dot a),\rho^*_{g^{-1}}(b)\right\rangle\\
&=\langle\rho_g(\dot a+v_E(a)),\rho^*_{g^{-1}}b\rangle=\langle\dot a+v_E(a),b\rangle.
\end{aligned}$$

■

Definition 2.2. a) A given Lagrangian L: $TG \to \mathbb{R}$ is called *left invariant* if $L(TL_h v_g) = L(v_g)$ for all $v_g \in T_g G$ and $h \in G$. In body coordinates this is equivalent to $L^B(g, v) = L^B(e, v) := L(v)$, where $L^B(g, v) = L(TL_g \cdot v)$. By abuse of notation, we shall often write $L(g, v)$ for $L^B(g, v)$.

b) For a given Lagrangian L: $TG \to \mathbb{R}$, define a new Lagrangian L^E: $T(G \times E \times E^*) \to \mathbb{R}$ and action φ_g^E: $G \times E \times E^* \to G + E \times E^*$ of G on $G \times E \times E^*$ by

$$\begin{aligned} L^E\left(v_g, a, \dot a, b, \dot b\right) &= L(v_g) + \langle \mathrm{d}\rho(g, a)/\mathrm{d}t, \rho^*_{g^{-1}} b\rangle \\ &= L(v_g) + \langle \dot a + v_E(a), b\rangle \\ &= L(v_g) + \theta_0(a, \dot a, b, \dot b) + J(a, b)(v) \end{aligned} \tag{2.10}$$

and

$$\varphi_g^E(h, a, b) = (gh, a, b).$$

c) For given $a_1, a_2 \in E$, the *Lin constraints* are defined by

$$\left\{(v, a) \in \Omega(\mathfrak{g}) \times \Omega(E):\ \dot a + v_E(a) = 0,\ a(t_1) = a_1, a(t_2) = a_2\right\}.$$

Remarks. 1) If L is left invariant, then L^E is invariant under the action φ^E and

$$\begin{aligned} L^E\left(v_g, a, \dot a, b, \dot b\right) &= L(v) + \langle \dot a + v_E(a), b\rangle \\ &:= L^E(v, a, \dot a, b, \dot b). \end{aligned} \tag{2.11}$$

2) According to the momentum lemma (Abraham and Marsden [32; p. 288]),

$$J(a, b)(v) + \theta_0(a, \dot a, b, \dot b) = (\Phi^*\theta_0)\left(v_g, a, \dot a, b, \dot b\right) \tag{2.12}$$

where Φ^* denotes the pull back by the action Φ: $G \times T^*E \to T^*E$, the lifted action of ρ. Notice that the expression (2.11) is valid for any element g of G, for example the identity. In particular, we get the following expression for L^E:

$$L^E = L \circ T\pi_1 + \Phi^*\theta_0, \tag{2.13}$$

where π_1: $G \times T^*E \to G$ is the projection.

Consider the following assumption:

Assumption 2.3. There is an open set $U \subset E$ such that the following conditions hold:
1) $G \subset E$;
2) $\rho(G \times U) \subset U$;
3) For each a in U, ρ_a: $G \to E$ is a diffeomorphism onto its image.

We will assume that 2.3 holds from now on. The assumption that U is open should be properly interpreted in each example, especially in infinite dimensional cases. This involves issues of functional analysis that we will not detail here. The point to keep in mind is that U should allow enough variations of curves to apply the usual variational techniques. Similarly, the assertion that ρ_a is a diffeomorphism onto its image should be properly interpreted in examples. It should be remarked that in many examples we can find E, U, and ρ satisfying the assumption 2.3 by realizing G as a matrix group of $n\times n$ matrices, choosing $U=\mathrm{GL}(n)$, $E=\mathrm{L}(\mathbb{R}^n,\mathbb{R}^n)$, the vector space of linear maps of $\mathbb{R}^n$ to itself, and $\rho(g,a)=ga$.

The following observation will be useful. *For given $g\in G$, $a\in E$ and $b\in E^*$, there is a $b'\in E^*$ such that for all $\delta g\in \mathrm{T}_gG$,*

$$\langle \delta g, b\rangle=\left\langle \frac{\partial\rho(g,a)}{\partial g}\delta g, b'\right\rangle. \tag{2.14}$$

Notice that since $G\subset E$, we can think of δg as being an element of E to make sense of the left-hand side of (2.14). To prove the assertion, first note that by using left translations, it is enough to prove (2.14) at $g=e$, the identity. Thus what we must prove is that

$$\langle \xi, b\rangle=\langle \xi_E(a), b'\rangle \tag{2.14a}$$

for all $\xi\in\mathfrak{g}$ determines b'. Since ρ_a is a diffeomorphism onto its image, the map

$$A_a:\ \xi\mapsto\xi_E(a) \tag{2.15}$$

from $\mathfrak{g}\subset E$ to E is one to one. Letting A_a^{-1} denote its inverse, from range (A_a) to $\mathfrak{g}$, we can let b' be any extension of $(A_a^{-1})^*(b|\mathfrak{g})$. (In infinite dimensions, we assume that range (A_a) is closed and that the adjoint exists.)

The main results of this section can now be stated. First an observation to get rid of some technicalities in the proof: for fixed $u_0\in U$ we have an embedding ρ_{u_0}: $G\to E$. Thus we can identify G with $\rho_{u_0}(G)$. Under this identification the action L_g of G on itself by left translations can be extended to an action of G on E by $L_g(a)=\rho_g(a)$. Since we have the natural inclusion $\mathrm{T}\rho_{u_0}$: $\mathrm{T}G\to G\times E$, we can extend every left invariant Lagrangian L: $\mathrm{T}G\to\mathbb{R}$ to an invariant function L: $G\times E\to\mathbb{R}$. In examples this usually comes about in a natural way. We will always assume such an extension has been chosen to make sense of expressions like $\partial L(g,\dot g)/\partial g$.

Theorem 2.4. Let $a_1, a_2\in U$, and $g_1, g_2\in G$ be such that $\rho_{g_1}(a_1)=\rho_{g_2}(a_2)$. Let L be a Lagrangian on $\mathrm{T}G$, $g(t)$ be a curve in $\Omega_{g_1,g_2}(G)$ and $v_g=\mathrm{d}g/\mathrm{d}t$. The following assertions are equivalent:

i) $g(t)$ is a critical point of

$$\int_{t_1}^{t_2}L(v_g)\,\mathrm{d}t=\int_{t_1}^{t_2}L(g,v)\,\mathrm{d}t$$

on $\Omega_{g_1,g_2}(G)$.

ii) There are curves a in $\Omega_{a_1,a_2}(U)$ and b in $\Omega(E^*)$ such that (g,a,b) is a critical point of

$$\int_{t_1}^{t_2}L^E\left(v_g, a, \dot a, b, \dot b\right)\mathrm{d}t$$

on $\Omega_{g_1}(G)\times\Omega_{a_1,a_2}(U)\times\Omega(E^*)$.

If in addition, we suppose that L is left invariant, then either condition (i) or (ii) is equivalent to:
iii) There are curves a in $\Omega_{a_1,a_2}(U)$ and b in $\Omega(E^*)$ such that (v, a, b) is a critical point of

$$\int_{t_1}^{t_2} L^E(v, a, \dot{a}, b, \dot{b})\,\mathrm{d}t$$

on $\Omega(\mathfrak{g}) \times \Omega_{a_1,a_2}(U) \times \Omega(E^*)$.

Example 2.5. *The free rigid body.* (See Abraham and Marsden [32], and Marsden, Ratiu, and Weinstein [16, 17].) Here $G = \mathrm{SO}(3)$, with the group elements denoted $A \in \mathrm{SO}(3)$. The Lie algebra is written

$$\mathfrak{g} = \mathrm{SO}(3) = \{\hat{v} \mid v \in \mathbb{R}^3\},$$

where for $v \in \mathbb{R}^3$,

$$\hat{v} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}.$$

Traditionally the angular velocity in body coordinates v is denoted "ω_B" We shall follow this notation and shall write the components of ω_B as $\omega_B = (\omega_1, \omega_2, \omega_3)$. Note that $\hat{\omega}_B = A^{-1}\dot{A}$. Define $E := \mathrm{L}(\mathbb{R}^3, \mathbb{R}^3)$ which we identify with E^* using the pairing $\langle a, b\rangle = -\frac{1}{2}\mathrm{Tr}(ba)$ and we let $\rho(A, a) = Aa$, so that $\rho^*(A^{-1}, b) = bA^{-1}$ and we let $U = \mathrm{GL}(3)$. We easily check that $\langle \rho_A a, \rho^*_{A^{-1}} b\rangle = \langle a, b\rangle$, and assumption 2.3 is readily verified. We consider the left invariant Lagrangian given by the kinetic energy:

$$L(\omega_B) = \left[I_1(\omega_1)^2 + I_2(\omega_2)^2 + I_3(\omega_3)^2\right]/2,$$

where a body frame is chosen in which the moment of inertia tensor I is diagonal. Therefore using (2.11), and writing $\dot{A}$ for a tangent vector to SO(3) at A, we get

$$L^E(\dot{A}, a, \dot{a}, b, \dot{b}) = L(\omega_B) + \langle \dot{a} + \hat{\omega}_B a, b\rangle,$$

where, as in the general theory, $\hat{\omega}_B = A^{-1}\dot{A}$.

Since rigid body dynamics is in canonical Hamiltonian form on TSO(3), we know from Hamilton's principle that critical points of

$$\int_{t_1}^{t_2} L(\dot{A})\,\mathrm{d}t \quad \text{on} \quad \Omega_{A_1,A_2}(\mathrm{SO}(3))$$

are solutions to the equations of motion (see Arnold [33]). Therefore using 2.4, we conclude that $\omega_B(t)$ is a solution to the classical Euler equations if and only if there exist curves a and b such that (ω_B, a, b) is a critical point of the integral

$$\int_{t_1}^{t_2} \left\{\tfrac{1}{2}\left[I_1(\omega_1)^2 + I_2(\omega_2)^2 + I_3(\omega_3)^2\right] + \langle \dot{a} + \hat{\omega}_B a, b\rangle\right\}\mathrm{d}t$$

on $\Omega(\mathfrak{g}) \times \Omega_{a_1,a_2}(U) \times \Omega(E^*)$. If we apply the usual techniques of the calculus of variations, considering arbitrary variations $\delta\hat{\omega}_B$ in $\Omega(\mathfrak{g})$, δa in $\Omega_{a_1,a_2}(U)$ and δb in $\Omega(E^*)$, we get the system in Clebsch representation

$$2I_1\omega_1 = (ab)_{23} - (ab)_{32},$$
$$2I_2\omega_2 = (ab)_{31} - (ab)_{13},$$
$$2I_3\omega_3 = (ab)_{12} - (ab)_{21},$$

with the evolution equations

$$\dot{a} + \hat{\omega}_B a = 0, \quad \dot{b} - b\hat{\omega}_B = 0.$$

Observe that we can replace the velocity from the first three equations into the last two, to give the equations of motion written in terms of Clebsch potentials.

Let us now turn to the proof of 2.4.

Lemma 2.6. Let B: $TG \to \mathfrak{g}$ be the body coordinate map defined by (2.6). Then for each fixed g_1 in G, the map

$$\Omega(B)\circ(\mathrm{d}/\mathrm{d}t): \Omega_{g_1}(G) \to \Omega(\mathfrak{g}) \tag{2.16}$$

is an isomorphism.

Proof. Let v be a given curve in $\mathfrak{g}$. Define the time dependent vector field X on G by $X(g,t) = T_eL_g v(t)$. This vector field has a unique integral curve $g(t)$ satisfying $g(t_1) = g_1$. The map so defined is the inverse of the map (2.16), so it is an isomorphism. ■

Proof of 2.4. Let $(g_s, a_s, b_s) \in \Omega_{g_1}(G) \times \Omega_{a_1,a_2}(U) \times \Omega(E^*)$ be a curve with $(g_0, a_0, b_0) = (g, a, b)$. Integration by parts and using the functional derivative notation on L regarded as a function on $G \times E$ gives:

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{t_1}^{t_2}\left[L(g_s,\dot{g}_s) + \left\langle \frac{\mathrm{d}\rho(g_s,a_s)}{\mathrm{d}t}, \rho^*(g_s^{-1},b_s)\right\rangle\right]\mathrm{d}t\bigg|_{s=0} = \mathrm{I} + \mathrm{II}$$

where

$$\mathrm{I} = \int_{t_1}^{t_2}\left[\left\langle \frac{\mathrm{d}g_s}{\mathrm{d}s}, \frac{\delta L(g_s,\dot{g}_s)}{\delta g} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta L(g_s,\dot{g}_s)}{\delta \dot{g}_s}\right\rangle\bigg|_{s=0}\right.$$
$$\left. + \left\langle \frac{\mathrm{d}\rho(g_s,a_s)}{\mathrm{d}t}, \frac{\mathrm{d}}{\mathrm{d}s}\rho^*(g_s^{-1},b_s)\right\rangle\bigg|_{s=0} - \left\langle \frac{\mathrm{d}\rho(g_s,a_s)}{\mathrm{d}s}, \frac{\mathrm{d}}{\mathrm{d}t}\rho^*(g_s^{-1},b_s)\right\rangle\bigg|_{s=0}\right]\mathrm{d}t$$

and

$$\mathrm{II} = \left\langle \frac{\mathrm{d}g_s}{\mathrm{d}s}, \frac{\delta L(g, \dot{g})}{\delta \dot{g}} \right\rangle \bigg|_{s=0,\, t=t_2} + \left\langle \frac{\partial \rho(g, a)}{\partial g} \frac{\mathrm{d}g_s}{\mathrm{d}s}, \rho^*(g^{-1}, b) \right\rangle \bigg|_{s=0,\, t=t_2}.$$

Proof of (ii) $\Rightarrow$ (i). Assumption (ii) gives $\mathrm{I} + \mathrm{II} = 0$ for any choice of curve (g_s, a_s, b_s). If we chose the curve such that $g_s = g$, $a_s = a$ (independent of s, but still depending on t), and b_s is arbitrary, $\mathrm{d}g_s/\mathrm{d}s = 0$, and so $\mathrm{II} = 0$. Thus $\mathrm{I} = 0$ as well. Using this and the fact that $\mathrm{d}\rho^*(g_s^{-1}, b_s)/\mathrm{d}s = \mathrm{d}\rho^*(g^{-1}, b_s)/\mathrm{d}s$ becomes an arbitrary variation and also that $\mathrm{d}\rho(g_s, a_s)/\mathrm{d}s = \mathrm{d}\rho(g, a)/\mathrm{d}s = 0$, we conclude that

$$\frac{\mathrm{d}\rho(g, a)}{\mathrm{d}t} = 0. \tag{2.17}$$

Now choosing $g_s = g$, $b_s = b$, but a_s arbitrary, one shows using a similar argument that

$$\frac{\mathrm{d}\rho^*(g^{-1}, b)}{\mathrm{d}t} = 0. \tag{2.18}$$

If g_s is a curve in $\Omega_{g_1, g_2}(G)$ then since the endpoint $g_s(t_2)$ is fixed, $\mathrm{d}g_s/\mathrm{d}s|_{t=t_2, s=0} = 0$ and so $\mathrm{II} = 0$. From this and (2.17) and (2.18) we can conclude that

$$\int_{t_1}^{t_2} \left[\left\langle \frac{\mathrm{d}g_s}{\mathrm{d}s}, \frac{\delta L(g_s, \dot{g}_s)}{\delta g} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\delta L(g_s, \dot{g}_s)}{\delta \dot{g}_s} \right\rangle \bigg|_{s=0} \right] \mathrm{d}t = 0. \tag{2.19}$$

Integrating the left-hand side of (2.19) by parts, assuming that $g_s(t_2) = g_2$ is fixed, we get

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_{t_1}^{t_2} L(g_s, \dot{g}_s)\, \mathrm{d}t \Big|_{s=0}$$

and so (i) is proved.

Proof of (i) $\Rightarrow$ (ii). The usual integration by parts argument shows that (i) gives the Euler–Lagrange equations:

$$\frac{\delta L(g, \dot{g})}{\delta g} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\delta L(g, \dot{g})}{\delta \dot{g}} = 0. \tag{2.20}$$

Thus (2.19) holds. Define $a(t) = \rho_{g^{-1}} \circ \rho_{g_1} a_1$. Using 2.3, it follows that $a(t) \in U$ and that $a(t_2) = a_2$. Thus $a \in \Omega_{a_1, a_2}(U)$. It is also obvious that $\rho_g a = \rho_{g_1} a_1$ does not depend on t, so that $\mathrm{d}\rho(g, a)/\mathrm{d}t = 0$ and thus (2.17) holds. Now use the observation following assumption 2.3 to prove that there exists $b_2 = b(t_2)$ such that $\mathrm{II} = 0$. Then define $b(t) = \rho_g^* \rho_{g_2^{-1}}^* b_2$. It follows that $\mathrm{d}\rho^*(g^{-1}, b)/\mathrm{d}t = 0$, so that (2.18) holds. From (2.17–2.19), it follows that $\mathrm{I} = 0$ for any curve $(g_s, a_s, b_s) \in \Omega_{g_1}(G) \times \Omega_{a_1, a_2}(U) \times \Omega(E^*)$ such that $(g_0(t), a_0(t), b_0(t)) = (g(t), a(t), b(t))$. Thus for those (g_s, a_s, b_s), we have $\mathrm{I} + \mathrm{II} = 0$ which is equivalent to (ii). This finishes the proof of the equivalence of (i) and (ii).

Proof of (ii) ⇔ (iii). This is a consequence of the commutativity of the diagram:

$$\begin{tikzcd}
\Omega_{g_1}(G)\times\Omega_{a_1,a_2}(U)\times\Omega(E^*) \arrow[r, "\varphi"] \arrow[dr, "\int_{t_1}^{t_2} L^E(v_g, a, \dot a, b, \dot b)\,\mathrm{d}t"'] & \Omega(\mathfrak{g})\times\Omega_{a_1,a_2}(U)\times\Omega(E^*) \arrow[d, "\int_{t_1}^{t_2} L^E(v, a, \dot a, b, \dot b)\,\mathrm{d}t"] \\
 & \mathbb{R}
\end{tikzcd}$$

where $\varphi=\Omega(B)\circ \mathrm{d}/\mathrm{d}t\times$ (identity on $\Omega_{a_1,a_2}(U)$) $\times$ (identity on $\Omega(E^*)$), which is an isomorphism by (1.5). ■

Remarks. 1) The previous results can be generalized without difficulty to the case of a right representation. Accordingly, we should replace L_g by R_g. This is an important consideration in many examples.

2) Theorem 2.4 may be regarded as a particular case of the Lagrange multiplier theorem, using the constraint $\mathrm{d}\rho(g,a)/\mathrm{d}t=0$ in the space $\Omega_{g_1}(G)\times\Omega_{a_1,a_2}(U)\times\Omega(E^*)$ with the Lagrange multiplier $\rho^*(g^{-1},b)$. One checks that a curve (g,a,b) satisfying this constraint is such that $g(t_2)=g_2$, as in part (i) of the theorem. In fact, $\mathrm{d}\rho(g,a)/\mathrm{d}t=0$ implies that $\rho(g(t_1),a(t_1))=\rho(g(t_2),a(t_2))$ and since $g(t_1)=g_1$, $a(t_1)=a_1$ and $a_2\in U$, we can use injectivity of ρ_{a_2} to get $g_2=g(t_2)$.

Example 2.7. *Homogeneous incompressible fluids.* Let D be the domain in which the fluid is moving, $G=\mathrm{Diff}_{\mathrm{vol}}(D)$, the group of volume-preserving diffeomorphisms, and $\mathfrak{g}$ be the Lie algebra of G, which consists of all divergence free vector fields on D parallel to ∂D. Let $g\in\Omega(G)$ be a given motion of the fluid. Using the homogeneity assumption (density = 1) the kinetic energy becomes:

$$\tfrac{1}{2}\langle\dot g,\dot g\rangle=\frac{1}{2}\int_D\left[\frac{\partial g(X,t)}{\partial t}\right]^2\mathrm{d}^3X. \tag{2.21}$$

Let $v=\mathrm{T}_gR_{g^{-1}}\dot g$ be the Eulerian velocity. Then the change of variables formula, using the fact that the Jacobian determinant of g is 1, implies that

$$\tfrac12\langle\dot g,\dot g\rangle=\tfrac12\langle v,v\rangle \tag{2.22}$$

which shows that the kinetic energy is right invariant. With the potential energy set equal to zero, Hamilton's principle will say that solutions g to the equations of motion are critical points of

$$\tfrac12\int_{t_1}^{t_2}\langle\dot g,\dot g\rangle\,\mathrm{d}t \tag{2.23}$$

on $\Omega_{g_1,g_2}(G)$. (See Arnold [33] and Ebin and Marsden [34].) Let E be the space of functions of D to $\mathbb{R}^3$ (these functions really should be of a specific differentiability class such as those found in Ebin and

Marsden [34]). Identify E^* with E using the pairing

$$\langle a, b\rangle = \int_D a(X)\cdot b(X)\,\mathrm{d}^3X.$$

Using the right actions

$$\rho(g, a) = a\circ g, \quad \text{and} \quad \rho^*(g^{-1}, b) = b\circ g,$$

the change of variables formula gives

$$\langle \rho(g, a), \rho^*(g^{-1}, b)\rangle = \langle a, b\rangle.$$

We let U be the set of all $a \in E$ such that a is an embedding. Then the conditions of assumption 2.3 are readily checked. Thus we can apply theorem 2.4 to obtain the following variational principle:

A curve v of vector fields on D is a solution to Euler's equations for a perfect fluid if and only if there are curves a and b in U such that (v, a, b) is a critical point of

$$\int_{t_1}^{t_2}\int_D\left\{\frac{1}{2}\|v(X,t)\|^2 + \left[\frac{\partial a(X,t)}{\partial t} + v_i(X,t)\frac{\partial a(X,t)}{\partial X^i}\right]\cdot b(X,t)\right\}\mathrm{d}^3X\,\mathrm{d}t$$

on $\Omega(\mathfrak{g})\times\Omega_{a_1,a_2}(U)\times\Omega(E^*)$. By the usual procedure of the calculus of variations, we get for the corresponding Euler–Lagrange equations:

$$\delta v:\quad v^b = -J(a,b), \quad \text{where} \quad v^b(\delta v) = \langle v, \delta v\rangle \text{ defines } v^b$$

$$\text{i.e.}\quad v_i = -\frac{\partial a(X,t)}{\partial x^i}\cdot b(X,t) + \frac{\partial p}{\partial X^i},$$

where p is a function determined in the usual way by the incompressibility condition,

$$\delta b:\quad \partial a/\partial t + v_i \partial a/\partial X^i = 0 \quad \text{i.e. } \rho(g,a) \quad \text{is a constant},$$

$$\delta a:\quad \partial b/\partial t + v_i \partial b/\partial X^i = 0 \quad \text{i.e. } \rho^*(g^{-1},b) \quad \text{is a constant}.$$

These last two equations just say that the vector quantities a and b are advected by the flow. (This example can also be done in terms of the vorticity instead of the velocity as in Marsden and Weinstein [12].)

3. Lagrangians depending on a parameter

For compressible flow, the Lagrangian in material representation depends parametrically on the fluid density; only in spatial representation does the density become time dependent. We shall generalize the results of the previous section to cover cases like this. We shall also generalize the configuration space somewhat to be of the form $G\times W$, where W is a vector space. The Lagrangian will be taken to be a map L_{α_0}: $\mathrm{T}(G\times W)\to\mathbb{R}$, depending on the parameter α_0 in a vector space V. We assume that there is a left (or right) representation r: $G\times V\to V$. For compressible flow, G will be the group of diffeomorphisms of the region D containing the fluid (we assume that the fluid has fixed boundaries), W is absent, V is the space of densities on D and the representation r is by push forward of densities (corresponding to the advection of the density by the flow.)

The cotangent lift of r_g: $V \to V$ is the map $(r_g, r^*_{g^{-1}})$: $V \times V^* \to V \times V^*$, where $r^*_{g^{-1}}$ is the dual of the map $r_{g^{-1}}$: $V \to V$. If we denote the natural pairing of V with V^* by $\langle\ ,\ \rangle$: $V \times V^* \to \mathbb{R}$, then we have $\langle r_g\alpha, r^*_{g^{-1}}\beta\rangle = \langle \alpha, \beta\rangle$, for $\alpha \in V$ and $\beta \in V^*$. Now define the Lagrangian L^V: $\mathrm{T}(G \times W \times V \times V^*) \to \mathbb{R}$ by

$$\begin{aligned} L^V(g, \dot{g}, w, \dot{w}, \alpha, \dot{\alpha}, \beta, \dot{\beta}) &= L_{r(g,\alpha)}(g, \dot{g}, w, \dot{w}) + \langle \mathrm{d}r(g,\alpha)/\mathrm{d}t, r^*_{g^{-1}}\beta\rangle \\ &= L_{r(g,\alpha)}(g, \dot{g}, w, \dot{w}) + \langle \dot{\alpha} + v_V(\alpha), \beta\rangle, \end{aligned} \tag{3.1}$$

where the last equality comes from eq. (2.9). The following proposition relates the dynamics on $G \times W$ described by L_{α_0} for fixed α_0 and the dynamics on $G \times W \times V \times V^*$ described by L^V.

Proposition 3.1. Let $g_1, g_2 \in G$, $w_1, w_2 \in W$ and $\alpha_1, \alpha_2 \in V$ be given and assume that $r(g_1, \alpha_1) = r(g_2, \alpha_2) = \alpha_0$ is fixed. Then the following conditions on a curve $(g, w) \in \Omega_{g_1, g_2}(g) \times \Omega_{w_1, w_2}(W)$ are equivalent:

i) (g, w) is a critical point of

$$\int_{t_1}^{t_2} L_{\alpha_0}(g, \dot{g}, w, \dot{w})\,\mathrm{d}t \tag{3.2}$$

on the space $\Omega_{g_1, g_2}(G) \times \Omega_{w_1, w_2}(W)$.

ii) There are curves α in $\Omega_{\alpha_1, \alpha_2}(V)$ and β in $\Omega(V^*)$ such that (g, w, α, β) is a critical point of

$$\int_{t_1}^{t_2} L^V(g, \dot{g}, w, \dot{w}, \alpha, \dot{\alpha}, \beta, \dot{\beta})\,\mathrm{d}t \tag{3.3}$$

on the space $\Omega_{g_1, g_2}(G) \times \Omega_{w_1, w_2}(W) \times \Omega_{\alpha_1, \alpha_2}(V) \times \Omega(V^*)$.

Proof. We will give the proof for the case $W = \{0\}$, so that L_{α_0}: $\mathrm{T}G \to \mathbb{R}$ and L^V: $\mathrm{T}(G \times V \times V^*) \to \mathbb{R}$. The proof for the case of a general W is entirely similar. First assume that (ii) holds, and let

$$(g_s, \alpha_s, \beta_s) \in \Omega_{g_1, g_2}(G) \times \Omega_{\alpha_1, \alpha_2}(V) \times \Omega(V^*)$$

be a variation of (g, α, β) such that for $s = 0$,

$$(g_0(t), \alpha_0(t)\beta_0(t)) = (g(t), \alpha(t), \beta(t)).$$

Then we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}s} \int_{t_1}^{t_2} \left[L_{r(g_s, \alpha_s)}(g_s, \dot{g}_s) + \left\langle \frac{\mathrm{d}r(g_s, \alpha_s)}{\mathrm{d}t}, r^*(g_s^{-1}, \beta_s) \right\rangle \right] \mathrm{d}t \bigg|_{s=0}. \tag{3.4}$$

Choose g_s in $\Omega_{g_1, g_2}(G)$ arbitrarily and choose $\alpha_s = r(g_s^{-1}, \alpha_0)$ so that $\mathrm{d}r(g_s, \alpha_s)/\mathrm{d}t = 0$. Therefore

$$0 = \frac{\mathrm{d}}{\mathrm{d}s} \int_{t_1}^{t_2} L_{\alpha_0}(g_s, \dot{g}_s)\,\mathrm{d}t\bigg|_{s=0}$$

for any given variation g_s, which is equivalent to (i).

Next we assume that (i) holds. Choose $\alpha = r(g^{-1}, \alpha_0)$, so that $\mathrm{d}r(g,\alpha)/\mathrm{d}t = 0$. Let us denote $\mathrm{d}L_{\gamma+\varepsilon\xi}(g,\dot g)/\mathrm{d}\varepsilon|_{\varepsilon=0} := L'_\gamma(g,\dot g)(\xi)$. Integrating (3.4) by parts gives

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{t_1}^{t_2} L^V(g_s, \dot g_s, \alpha_s, \dot\alpha_s, \beta_s, \dot\beta_s)\,\mathrm{d}t\Big|_{s=0}$$

$$= \int_{t_1}^{t_2} L'_{r(g,\alpha)}(g,\dot g)\left[\frac{\mathrm{d}r(g_s,\alpha_s)}{\mathrm{d}s}\right]\Bigg|_{s=0}\mathrm{d}t + \int_{t_1}^{t_2}\frac{\mathrm{d}}{\mathrm{d}s}L_{\alpha_0}(g_s,\dot g_s)|_{s=0}\,\mathrm{d}t$$

$$+ \int_{t_1}^{t_2}\left\langle \frac{\mathrm{d}}{\mathrm{d}t}r(g,\alpha),\ \frac{\mathrm{d}r^*(g_s^{-1},\beta_s)}{\mathrm{d}s}\Big|_{s=0}\right\rangle \mathrm{d}t - \int_{t_1}^{t_2}\left\langle \frac{\mathrm{d}r(g_s,\alpha_s)}{\mathrm{d}s}\Big|_{s=0},\ \frac{\mathrm{d}r^*(g^{-1},\beta)}{\mathrm{d}t}\right\rangle \mathrm{d}t. \tag{3.5}$$

The third term on the right side of (3.5) is obviously zero and the second is zero by assumption. Now write

$$L'_{r(g,\alpha)}(g,\dot g)\left[\frac{\mathrm{d}r(g_s,\alpha_s)}{\mathrm{d}s}\right]\Bigg|_{s=0} = \left\langle \frac{\mathrm{d}r(g_s,\alpha_s)}{\mathrm{d}s}\Big|_{s=0},\ \frac{\delta}{\delta\alpha_0}L_{r(g,\alpha)}(g,\dot g)\right\rangle.$$

If we choose $\beta \in V^*$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}r^*(g^{-1},\beta) = \frac{\delta}{\delta\alpha_0}L_{r(g,\alpha)}(g,\dot g),$$

then the proof will be finished. But we can choose

$$\beta = r_g^*\int^t \frac{\delta}{\delta\alpha}L_{r(g,\alpha)}(g,\dot g)\,\mathrm{d}t. \qquad \blacksquare$$

Our Lagrangian will be said to be *left invariant* if

$$L_{r(h,\alpha_0)}(\mathrm{T}_g L_h \cdot v_g, w, \dot w) = L_{\alpha_0}(v_g, w, \dot w) \tag{3.6}$$

for all g and h in G, α_0 in V, v_g in T_gG and where we take r to be a left representation. Likewise, we will say that L is *right invariant* if we take r to be a right representation and

$$L_{r(\alpha_0,h)}(\mathrm{T}_g R_h \cdot v_g, w, \dot w) = L_{\alpha_0}(v_g, w, \dot w). \tag{3.7}$$

The following proposition is readily verified:

Proposition 3.2. Let φ be the action of G on $G \times W \times V \times V^*$ given by

$$\varphi_h(g, w, \alpha_0) = (hg, w, \alpha_0) \tag{3.8}$$

if r is a left representation and

$$\varphi_h(g, w, \alpha_0) = (gh, w, \alpha_0) \tag{3.9}$$

if r is a right representation. Then in both cases, L_{α_0} is invariant if and only if L^V is invariant under the action of φ. In this case, L^V can be written as

$$L^V(g, \dot g, w, \dot w, \alpha, \dot\alpha, \beta, \dot\beta) = L_\alpha(e, v, w, \dot w) + \langle \dot\alpha + v_V(\alpha), \beta\rangle, \tag{3.10}$$

where, as usual, v denotes $\dot g$ in body (respectively space) coordinates.

Remarks. 1) Consider the integral (2.3). Variations $\delta\beta$ produce the equation

$$\mathrm{d}r(g, \alpha)/\mathrm{d}t = 0 \quad \text{i.e. } \alpha = r_{g^{-1}}(\alpha_0) \text{ for some } \alpha_0 \in V \tag{3.11}$$

or equivalently $\dot\alpha + v_V(\alpha) = 0$. This means effectively that the parameter α is "Lie dragged" along by the motion g on G. The same thing happens to β as one sees by taking variations in α.

2) If L_{α_0} is invariant, then L_{α_0} is invariant in the usual sense under the action of the isotropy subgroup

$$G_{\alpha_0} = \{g \in G \mid r(g, \alpha_0) = \alpha_0\}.$$

The converse need not be true, however.

3) If L_{α_0} is invariant, the integral (2.3) becomes

$$\int_{t_1}^{t_2} [L_\alpha(e, v, w, \dot w) + \langle \dot\beta + v_V(\alpha), \beta\rangle]\, \mathrm{d}t. \tag{3.12}$$

The variational principle allows arbitrary variations of $w \in \Omega_{w_1, w_2}(W)$, $\alpha \in \Omega_{\alpha_1, \alpha_2}(V)$ and $\beta \in \Omega(V^*)$. However, variations of v are constrained by the condition

$$v = \mathrm{T}_g L_{g^{-1}}(g, \dot g) \quad \text{and} \quad g \in \Omega_{g_1, g_2}(G);$$

arbitrary variations of g with fixed endpoints will not produce arbitrary variation of v. We will avoid this difficulty in the next section by using Lin constraints. This will be a slight but useful generalization of the results of section 1.

4) The roles of V and V^* in the previous definitions and results are interchangeable, and in some examples such as fluids, in which the parameter is the density which is naturally an element of the dual to the space of functions (see examples D and E below), it seems quite natural to do so. This interchange is also consistent with the useage in the general theory of semidirect products in Marsden, Ratiu, and Weinstein [16, 17].

Suppose that $V, V^*, \langle\ ,\ \rangle$, and r are as before and that we have a Lagrangian depending on a parameter $\beta_0 \in V^*$, say

$$L_{\beta_0}: \mathrm{T}(G \times W) \to \mathbb{R}.$$

Then define

$$L^{V^*}: \mathrm{T}(G \times W \times V \times V^*) \to \mathbb{R}$$

by

$$L^{V^*}(g, \dot g, w, \dot w, \alpha, \dot\alpha, \beta, \dot\beta) = L_{r_g^*{}^{-1}\beta}(g, \dot g, w, \dot w) + \left\langle r_g\alpha, \frac{\mathrm{d}r_g^*{}^{-1}\beta}{\mathrm{d}t}\right\rangle$$
$$= L_{r_g^*{}^{-1}\beta}(g, \dot g, w, \dot w) + \langle \alpha, \dot\beta + v_{V^*}(\beta)\rangle.$$

The Lagrangian L_{β_0} is invariant in the sense that

$$L_{r_g^*{}^{-1}\beta_0}(\mathrm{T}_h L_g v_h, w, \dot w) = L_{\beta_0}(v_h, w, \dot w)$$

if and only if L^{V^*} is invariant under the lifted action as before. In that case we have

$$L^{V^*}(g, \dot g, w, \dot w, \alpha, \dot\alpha, \beta, \dot\beta) = L^{V^*}(e, v, w, \dot w, \alpha, \dot\alpha, \beta, \dot\beta)$$
$$= L_\beta(e, v, w, \dot w) + \langle \alpha, \dot\beta + v_{V^*}(\beta)\rangle.$$

Examples 3.3. A) *The heavy top in body representation*. (See Marsden, Ratiu, and Weinstein [16, 17] for a general reference; we will, to a large extent, follow the notation of this reference.) A heavy top is a rigid body moving in three dimensional Euclidean space about a fixed point and under the influence of gravity. The configuration space is $G = \mathrm{SO}(3)$ and the motion of the body is a curve $A \in \Omega(G)$. Let M be the total mass of the top and let $l\chi$ denote the vector determining the center of mass, where χ is a unit vector along the line from the fixed point to the center of mass. Also let α_0 be a vector representing the force of gravity (usually α_0 is taken to be a vector pointing vertically downward in spatial representation), with magnitude $\|\alpha_0\|$ the acceleration due to gravity. Thus the potential energy is

$$V(A, \alpha_0) = Ml\alpha_0 \cdot A\chi = MlA^{-1}\alpha_0 \cdot \chi. \tag{3.13}$$

The kinetic energy, being as in the example of the rigid body, we have the Lagrangian L_{α_0}: $TG \to \mathbb{R}$ given by

$$L_{\alpha_0}(A, \dot A) = \tfrac12\left[I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2\right] - Ml\alpha_0 \cdot A\chi. \tag{3.14}$$

We regard L_{α_0} as a Lagrangian depending on a parameter (even though α_0 is of course fixed) and take the action of G on the parameter space $V = \mathbb{R}^3$ to be the usual left action. The vector χ and the moment of inertia tensor I are held fixed under this action. We also identify V^* with V using the standard dot product. Since the kinetic energy is left invariant, as in the free rigid body, the Lagrangian L_{α_0} is left invariant in the sense of Lagrangians depending on parameters. We also compute using (3.1) or (3.10) with $\alpha_0 = A\alpha$, that

$$L^V(A, \dot A, \alpha, \dot\alpha, \beta, \dot\beta) = L_{A\alpha}(\dot A) + \langle \dot\alpha + \omega_B \times \alpha, \beta\rangle$$
$$= \tfrac12\left[I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2\right] - Ml\alpha \cdot \chi + \langle \dot\alpha + \omega_B \times \alpha, \beta\rangle. \tag{3.15}$$

Thus, by proposition 3.1, solutions of the heavy top equations are given by critical points of the integral (2.3). (We accept as known that the heavy top equations satisfy Hamilton's principle on TSO(3) for the Lagrangian (3.14) with α held fixed–see for example, Marsden, Ratiu, and Weinstein [16, 17].)

B) *Compressible isentropic fluids*. Here we start with the Lagrangian

$$L_{\rho_0}: \mathrm{T}(\mathrm{Diff}(D)) \to \mathbb{R}$$

given by

$$L_{\rho_0}(\varphi,\dot\varphi)=\tfrac{1}{2}\int_D\rho_0(X)\|v\circ\varphi(X)\|^2\,\mathrm{d}^3X-\int_D\rho_0(X)w\big(\rho_0(X)\big(J_{\varphi^{-1}}\circ\varphi(X)\big)\big)\,\mathrm{d}^3X, \tag{3.16}$$

where $J_{\varphi^{-1}}$ is the Jacobian of φ^{-1}. Choose

$$V=\mathscr{F}(D)=V^*,$$

$$\langle\alpha,\rho\rangle=\int_D\alpha(X)\rho(X)\,\mathrm{d}^3X,$$

and

$$r(\varphi,\alpha)=\alpha\circ\varphi,\quad r^*(\varphi,\rho)=\rho\circ\varphi^{-1}J_{\varphi^{-1}}$$

so that we have the usual invariance property

$$\langle r(\varphi,\alpha),r^*(\varphi^{-1},\rho)\rangle=\langle\alpha,\rho\rangle.$$

The Lagrangian L_{ρ_0} is invariant as is readily verified, and so we get

$$L^{V^*}(\varphi,\dot\varphi,\alpha,\dot\alpha,\rho,\dot\rho)=\tfrac{1}{2}\int_D\rho(x)\|v(x)\|^2\,\mathrm{d}^3x-\int_D\rho(x)w(\rho(x))\,\mathrm{d}^3x$$
$$+\int_D\alpha(x)(\dot\rho+\nabla(\rho v))\,\mathrm{d}^3x, \tag{3.17}$$

and thus the continuity constraint $\dot\rho+\nabla(\rho v)=0$ appears in a natural way.

Additional examples that one can treat in a similar fashion are the following: *the heavy top in spatial representation*, *incompressible inhomogeneous fluids*, *compressible nonisentropic fluids*, *and fluids in the convective picture* (see Holm, Marsden, and Ratiu [35]). The reader who is interested in these examples can receive from us a longer version of the present paper.

4. Variational principles and Lin constraints for parameter dependent Lagrangians

In this section we will introduce Lin constraints for variational principles involving Lagrangians of type $L\colon \mathrm{T}(G\times F)\to\mathbb{R}$, where F is a vector space. For example, the Lagrangians considered in section 3 are of this type, with $F=W\times V\times V^*$.

The main result of this section is a generalization of theorem 2.4. The result is similar enough so that we may omit the proof; however, it is still useful in examples. Another interesting case of this type is that of a Lagrangian $L\colon \mathrm{TT}^*G\to\mathbb{R}$, $L=\theta_0-H$, where θ_0 is the canonical 1-form regarded as a function on TT^*G and H is a given Hamiltonian. In this case we identify $\mathrm{T}^*G=G\times\mathfrak{g}^*$ (using body or space coordinates) so that we can take $\mathfrak{g}^*=F$. More generally, we may consider Lagrangians like $L\colon \mathrm{T}(\mathrm{T}^*(G\times W))\to\mathbb{R}$, where W is a given vector space. Then $F=\mathfrak{g}^*\times\mathrm{T}^*W=\mathfrak{g}^*\times W\times W^*$.

Let φ be the (left) action of G on $G\times F$ given by $\varphi_g(h,f)=(gh,f)=(L_gh,f)$. We will often represent $T(G\times F)=G\times\mathfrak{g}\times F\times F$ where $\mathrm{T}G$ is identified with $G\times\mathfrak{g}$ using the body coordinate representation. A typical element of $T(G\times F)$ will be denoted by $(g, v, f, \dot f)$ where $v = TL_{g^{-1}}V_g$, $V_g \in T_gG$. Now assume that E, E^*, ρ, ρ^*, $\langle\ ,\ \rangle$, U, θ_0, and J are as in section 2. Note that these can be chosen once the group G is given, no matter what the space F is. For a given Lagrangian L: $T(G\times F)\to \mathbb{R}$, define the map L^E: $T(G\times F\times E\times E^*)\to\mathbb{R}$ by

$$\begin{aligned} L^E(g, v, f, \dot f, a, \dot a, b, \dot b) &= L(g, v, f, \dot f) + \left\langle \frac{d\rho(g,a)}{dt}, \rho^*(g^{-1}, b)\right\rangle \\ &= L(g, v, f, \dot f) + \langle \dot a + v_E(a), b\rangle \\ &= L(g, v, f, \dot f) + \theta_0(a, \dot a, b, \dot b) + J(a,b)(v). \end{aligned} \tag{4.1}$$

Proposition 4.1. Suppose that $(g_i, f_i)\in G\times F$, $i=1,2$, $\alpha_1, a_2 \in U$, and $\rho_{g_1}a_1 = \rho_{g_1}a_2$. Then the following assertions are equivalent:

i) (g, f) is a critical point of

$$\int_{t_1}^{t_2} L(g, v, f, \dot f)\,dt$$

on

$$\Omega_{g_1,g_2}(G)\times\Omega_{f_1 f_2}(F).$$

ii) There are curves a, b such that (g, f, a, b) is a critical point of

$$\int_{t_1}^{t_2} L^E(g, v, f, \dot f, a, \dot a, b, \dot b)\,dt$$

on

$$\Omega_{g_1}(G)\times\Omega_{f_1,f_2}(F)\times\Omega_{a_1,a_2}(U)\times\Omega(V^*).$$

Assume, in addition, that L is invariant under the (lifted) action of the left action φ on $G\times F$. Then, either of conditions (i) or (ii) are equivalent to:

iii) There are curves a, b such that (v, f, a, b) is a critical point of

$$\int_{t_1}^{t_2} L^E(v, f, \dot f, a, \dot a, b, \dot b)\,dt$$

on

$$\Omega(\mathfrak{g})\times\Omega_{f_1,f_2}(F)\times\Omega_{a_1,a_2}(U)\times\Omega(V^*),$$

where $L^E(v, f, \dot f, a, \dot a, b, \dot b)$ stands for $L^E(e, v, f, \dot f, a, \dot a, b, \dot b)$.

Remark. The previous proposition and definition can be readily modified for the case of a right action

$$\varphi_g(h, f) = (hg, f) = (R_g h, f).$$

Examples. The following examples correspond to those in section 3. In each case, we first choose E, E^*, $\langle\ ,\ \rangle, \rho, \rho^*, U$ and then, as a direct application of (4.1), we write the variational principle involving Lin constraints. Each time we start with the Lagrangian $L = L^V$ obtained in section 3.

A) *Heavy top in body coordinates.* We choose $F = V \times V^*$, with $V = \mathbb{R}^3$ and L given by

$$L(A, \omega_B, \alpha, \dot{\alpha}, \beta, \dot{\beta}) = \tfrac{1}{2} \sum_{i=1}^{3} I_i \omega_i^2 - Ml\alpha \cdot \chi + \langle \dot{\alpha} + (\hat{\omega}_B)_V(\alpha), \beta \rangle. \tag{4.2}$$

Since $G = \mathrm{SO}(3)$, we may choose $E, E^*, \rho, \rho^*, U, \langle\ ,\ \rangle$ as in example 2.5. Then we get the Lagrangian

$$L^E(g, \omega_B, \alpha, \dot{\alpha}, \beta, \dot{\beta}, a, \dot{a}, b, \dot{b}) = \tfrac{1}{2} \sum_{i=1}^{3} I_i \omega_i^2 - Ml\alpha \cdot \chi + \langle \dot{\alpha} + \omega_B \times \alpha, \beta \rangle + \langle \dot{a} + \hat{\omega}_B a, b \rangle. \tag{4.3}$$

Since the variations $\delta\omega_B$, δa, $\delta\beta$, δa, and δb are arbitrary (except for a fixed endpoint condition in some cases) we can apply the usual techniques of the calculus of variations to the integral

$$\int_{t_1}^{t_2} L^E \, \mathrm{d}t$$

and obtain the following equations:

$$\begin{aligned}
\delta\omega_B:\quad & I_1\omega_1 = \tfrac{1}{2}((ab)_{23} - (ab)_{32}) - \beta_3\alpha_2 + \beta_2\alpha_3, \\
& I_2\omega_2 = \tfrac{1}{2}((ab_{31} - (ab)_{13}) - \beta_1\alpha_3 + \beta_3\alpha_1, \\
& I_3\omega_3 = \tfrac{1}{2}((ab)_{12} - (ab)_{21}) - \beta_2\alpha_1 + \beta_1\alpha_2
\end{aligned}$$

$$\left(\text{or } I\omega = \xi + \beta \times \alpha, \text{ where } \hat{\xi} = \frac{(ab)^{\mathrm{T}} - (ab)}{2}\right).$$

$$\begin{aligned}
\delta\beta:\quad & \dot{\alpha} + \omega_B \times \alpha = 0. \\
\delta\alpha:\quad & \dot{\beta} + \omega_B \times \beta + Ml\chi = 0. \\
\delta b:\quad & \dot{a} + \hat{\omega}_B a = 0. \\
\delta a:\quad & \dot{b} - b\hat{\omega}_B = 0.
\end{aligned}$$

B) *Compressible isentropic fluids.* Let E, E^*, ρ, ρ^*, U be chosen in a way analogous to example 2.7, and let L^{V^*} be the Lagrangian obtained in section 3, example B. Then we get

$$\begin{aligned}
L^E(\varphi, v, \alpha, \dot{\alpha}, \rho, \dot{\rho}, a, \dot{a}, b, \dot{b}) &= \tfrac{1}{2} \int_D \rho(x) \|v(x)\|^2 \,\mathrm{d}^3x - \int_D \rho(x) w(\rho(x)) \,\mathrm{d}^3x \\
&\quad + \int_D \alpha(x)(\dot{\rho} + \nabla \cdot (\rho v))(x) \,\mathrm{d}^3(x) + \int_D \left(a + \frac{\partial a}{\partial x_i} v_i(x) \right) \cdot b(x) \,\mathrm{d}^3x.
\end{aligned}$$

Variations $\delta v \| \partial D, \delta\alpha, \delta\rho, \delta a, \delta b$ can be chosen arbitrarily (except for the usual fixed endpoint conditions imposed in (4.1)) and we get

$$\begin{array}{lll}\delta v: & v_i = \dfrac{\partial \alpha}{\partial x_i} - \dfrac{b}{\rho}\cdot\dfrac{\partial a}{\partial x_i} & \text{[Clebsch representation]} \\ \delta\alpha: & \dot\rho + \nabla\cdot(\rho v) = 0 & \text{[Continuity]} \\ \delta b: & \dot a_i + v\cdot\nabla a_i = 0, \quad i=1,2,3 & \\ \delta a: & \dot b_i + v\cdot\nabla b_i + b_i\nabla\cdot v = 0, \quad i=1,2,3 & \multirow{2}{*}{\text{[Lin constraint]}} \\ \delta\rho: & \dot\alpha = \tfrac{1}{2}\|v\|^2 - \dfrac{\partial}{\partial\rho}(\rho w(\rho)) - v\cdot\nabla\alpha. & \end{array}$$

As in section 3, we remark that the additional examples of the heavy top in spatial coordinates, incompressible inhomogeneous fluids, compressible nonisentropic fluids, and fluids in the convective picture are available in a longer version of the present paper.

5. Variational principles and reduction

Let (P,ω) be a symplectic manifold where the symplectic form ω is exact, say $\omega = -\mathrm{d}\theta$, where θ is a 1-form. Then solutions to the equations of motion for a given Hamiltonian $H: P \to \mathbb{R}$ are critical points of the function

$$\int_{t_1}^{t_2}\left[\theta\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right) - H(z)\right]\mathrm{d}t, \tag{5.1}$$

where

$$z \in \Omega_{z_1,z_2}(P)$$

and $z_1, z_2 \in P$. Now assume that G acts on P (say, by a left action) by θ-preserving diffeomorphisms of P, and let J be the associated momentum mapping. According to reduction theory (see Marsden and Weinstein [18] and Abraham and Marsden [32]), there is a symplectic structure ω_μ on the reduced space $P_\mu = J^{-1}(\mu)/G_\mu$, where G_μ is the isotropy group of the coadjoint action that leaves $\mu \in \mathfrak{g}^*$ fixed. Note that ω_μ is not necessarily exact. In other words, one cannot expect to obtain a reduced version of the 1-form θ, except in some particular cases.

The purpose of this section is to show that, nevertheless, we can still give a reduced version of the variational principle. Let G be a group. Then T^*G is a symplectic manifold and G acts on T^*G by either (lifted) left or right translations.

Remark. Some of the concepts on this section are related to results of Novikov [36]. There the functional on curves given by

$$\int_{t_1}^{t_2}\theta(\mathrm{d}z/\mathrm{d}t)\,\mathrm{d}t$$

is studied as a locally defined object when ω is not exact, giving rise to a further definition of an infinitely sheeted covering and a related global extension of the previous functional. However, the main purpose of Novikov's paper is to study some generalization of Morse theory, while the role of symmetry is not considered; see also Weinstein [37]. There are some remarks in Novikov's paper, concerning Clebsch potentials as locally defined objects and the question of conditions for its global existence is raised. Our assumption 2.3 partially answers this question.

5.1. *Cotangent bundle case*

We will begin with the case $P = \mathrm{T}^*G$ and will leave considering a more general situation for later. Notice, however, that many interesting examples (compressible flow, MHD, etc.) are still included in the case $P = \mathrm{T}^*G$ by taking G to be a semidirect product (see Marsden, Ratiu, and Weinstein [16]).

The momentum mapping corresponding to left translation, say J: $\mathrm{T}^*G \to \mathfrak{g}^*$, is well-known to be

$$J(\alpha_g) = \alpha_g \circ \mathrm{T}_e R_g.$$

In body coordinates α_g becomes $\nu = \mathrm{T}_e^* L_g \alpha_g$, so we identify α_g with $(g, \nu) \in G \times \mathfrak{g}^* \equiv \mathrm{T}^*G$, and so J becomes

$$J(g, \nu) = \mathrm{Ad}^*_{g^{-1}} \nu.$$

The set $J^{-1}(\mu)$ is the graph of the 1-form α_μ given by

$$\alpha_\mu(g) = \mu \circ \mathrm{T}_g R_{g^{-1}} \tag{5.2}$$

which is invariant under the (lifted) action of G by right translations, and is also left invariant under the (lifted) action of G_μ by left translations. Let H: $\mathrm{T}^*G \to \mathbb{R}$ be a given Hamiltonian, and let $\mathscr{L}_\mu$: $\Omega(G) \to \mathbb{R}$ be the functional

$$\begin{aligned} \mathscr{L}_\mu(g) &= \int_{t_1}^{t_2} \left[\theta \circ \left(\frac{\mathrm{d}\alpha_\mu(g)}{\mathrm{d}t} \right) - H(\alpha_\mu(g)) \right] \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left[\alpha_\mu(g)(\dot{g}) - H(\alpha_\mu(g)) \right] \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left[\mathrm{Ad}^*_g \mu(v) - H(\alpha_\mu(g)) \right] \mathrm{d}t, \end{aligned} \tag{5.3}$$

where $v = \mathrm{T}_g L_{-1} \dot{g}$ is the body representation of the velocity. Observe that if H is (left) invariant then, in addition to the equalities (5.3), we have

$$\mathscr{L}_\mu(g) = \int_{t_1}^{t_2} \left[(\mathrm{Ad}^*_g \mu)(v) - H(\mathrm{Ad}^*_g \mu) \right] \mathrm{d}t. \tag{5.4}$$

Since the map α_μ: $G \to J^{-1}(\mu)$ is a diffeomorphism, to find critical points of $\mathscr{L}_\mu(g)$ on $g \in \Omega_{g_1, g_2}(G)$ amounts to "restricting" the variational principle (5.1) to $J^{-1}(\mu) \subset \mathrm{T}^*G$. Critical points of (5.1) are solutions to equations of motion that stay on level sets $J^{-1}(\mu)$, and they project under τ_G^*: $\mathrm{T}^*G \to G$ onto critical points of $\mathscr{L}_\mu(g)$ on $\Omega_{g_1, g_2}(G)$ for some g_1, g_2. However, it can be readily verified in simple examples that critical points of $\mathscr{L}_\mu(g)$ on $\Omega_{g_1,g_2}(G)$ are usually not unique. Moreover, by applying the usual techniques of the calculus of variations we get part (a) of the following proposition:

Proposition 5.1. a) Critical points of $\mathscr{L}_\mu(g)$ on $\Omega_{g_1,g_2}(G)$ are curves g such that

$$-\mathrm{d}\alpha_\mu(\dot{g},\delta g)-\mathrm{d}\tilde{H}(g)(\delta g)=0 \quad \text{for all } \delta g,$$

where $\tilde{H}(g)=H(\alpha_\mu(g))$.

b) A curve g is a critical point of $\mathscr{L}_\mu(g)$ on $\Omega_{g_1,g_2}(G)$ if and only if the projection $\pi_\mu(\alpha_\mu(g))$ is a solution of the equations of motion of the reduced system, where π_μ: $J^{-1}(\mu)\to P_\mu$ is the natural projection.

We can prove part b) of proposition 5.1 by using part a) and techniques from Abraham and Marsden [32; p. 302].

Let us observe that the map π_μ is given by $\pi_\mu(\alpha_\mu(g))=\mathrm{Ad}_g^*\mu$ and that P_μ coincides with the coadjoint orbit $G\cdot\mu\approx P_\mu$.

As an immediate consequence of proposition 5.1b and the definition of the isotropy group $G_\mu=\{h\in G|\mathrm{Ad}_h^*\mu=\mu\}$, we have:

Corollary 5.2. The curves g and $\bar{g}\in\Omega_{g_1,g_2}(G)$ are critical points of $\mathscr{L}_\mu(g)$ that project onto the same solution of the reduced system on P_μ if and only if there exists $h\in\Omega(G_\mu)$ such that $h(t)g(t)=\bar{g}(t)$.

The quantity

$$\int_{t_1}^{t_2}(\mathrm{Ad}_g^*\mu)(v)\,\mathrm{d}t$$

that appears in (5.3) and (5.4) clearly depends on the curve $g\in\Omega_{g_1,g_2}(G)$. However, it can be shown after some calculation that it only depends on the curve

$$\mathrm{Ad}_g^*\mu:=\nu\in\Omega_{\nu_1,\nu_2}(P_\mu),$$

where $\nu_i=\mathrm{Ad}_g^*\mu$, $i=1,2$, at least for small variations of the curve ν about a fixed position. This amounts to showing that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{t_1}^{t_2}\mathrm{Ad}_{h_\lambda g}^*\mu(v_\lambda)\,\mathrm{d}t=0,$$

where $h_\lambda(t)\in G_\mu$, $h_0(t)=e$ for $t\in[t_1,t_2]$, $h_\lambda(t_i)=e$, $i=1,2$, and $v_\lambda=TL_{(h_\lambda g)^{-1}}(\mathrm{d}h_\lambda g/\mathrm{d}t)$. To be precise, let us introduce the following notion. Define on $\Omega(G)$ the equivalence relation $g\sim\bar{g}$ if and only if there exists $h_\lambda \in \Omega(G_\mu)$ such that $h_\lambda(t_i) = e$, for $\lambda \in [0,1]$, $h_0 = e$ for $t \in [t_1, t_2]$, and $h_1 g = g$. Then $\mathscr{L}_\mu(g)$ induces a function $\tilde{\mathscr{L}}_\mu$: $\Omega(G)/\sim \to \mathbb{R}$. In other words, we have a commutative diagram

$$\begin{tikzcd}
\mathbb{R} & & \Omega(P_\mu) \arrow[ll, "\overline{\tilde{\mathscr{L}}}_\mu"'] \\
& \Omega(G)/\sim \arrow[ul, "\tilde{\mathscr{L}}_\mu"] \arrow[ur, "\beta"'] & \\
& \Omega(G) \arrow[u, "\alpha"] \arrow[uul, "\mathscr{L}_\mu"] \arrow[uur, "\Omega(J \circ \alpha_\mu)"'] &
\end{tikzcd}$$

where α, β are the obvious natural maps.

An important observation is that the restriction $\beta | \Omega_{g_1, g_2}(G)/\sim$ is a local diffeomorphism onto $\Omega_{\nu_1, \nu_2}(P_\mu)$ where $\nu_i = \mathrm{Ad}^*_{g_i} \mu$, $i = 1, 2$ (this is readily verified and it is also a consequence of the corresponding result in part B, below), though it is not a diffeomorphism unless G_μ is simply connected. Thus $\tilde{\mathscr{L}}_\mu$ induces, in general, a multivalued function, say $\overline{\mathscr{L}}_\mu$ on $\Omega_{\nu_1, \nu_2}(P_\mu)$, and this will be single valued if G_μ is simply connected. Another important fact is that the variation

$$\delta \overline{\mathscr{L}}_\mu : \Omega_{\nu_1, \nu_2}(P_\mu) \to \mathrm{T}^* \Omega_{\nu_1, \nu_2}(P_\mu)$$

is given by

$$\delta \overline{\mathscr{L}}_\mu(\nu)(\delta\nu) = \int_{t_1}^{t_2} \left[\omega_\mu(\nu, \delta\nu) - \mathrm{d}H(\nu)(\delta\nu) \right] \mathrm{d}t. \tag{5.5}$$

In particular, this implies that $\delta \overline{\mathscr{L}}_\mu$ is a well defined single valued 1-form on $\Omega_{\nu_1, \nu_1}(P_\mu)$.

The proof of these facts will be given in (B) in a more general situation.

Remarks 5.3. 1) Given a symplectic manifold (P, ω) and a Hamiltonian H: $P \to \mathbb{R}$, *solutions to the equations of motions are critical points of*

$$\int_{t_1}^{t_2} [\omega(\dot{z}, \delta z) - \mathrm{d}H(z)(\delta z)] \, \mathrm{d}t \tag{5.6}$$

on $\Omega_{z_1, z_2}(P)$. In other words, the integral (4.6) is 0 for all δz if and only if $z(t)$ is a solution to the equations of motion. This is an obvious fact.

2) All the previous results remain valid, after appropriate modifications, if we work with right invariant systems rather than left invariant systems.

3) To determine

$$\int_{t_1}^{t_2} \mathrm{Ad}^*_g \mu(v) \, \mathrm{d}t$$

as a locally defined function on $\Omega_{\nu_1,\nu_2}(P_\mu)$ amounts to choosing a map

$$\Omega_{\nu_1,\nu_2}(P_\mu) \to \Omega_{g_1,g_2}(G)$$

such that, composed with $g \mapsto \mathrm{Ad}_g^* \mu$, it gives the identity. In many cases this can be done in a natural way.

5.2. *Exact symplectic manifold case*

Let (P, ω) be a symplectic manifold with $\omega = -\mathrm{d}\theta$.

Lemma 5.4. Let $z \in \Omega_{z_1,z_2}(P)$ and let z_λ be a curve on $\Omega_{z_1,z_2}(P)$ at z, meaning $z_\lambda(t_i) = z(t_i)$, $i = 1,2$, and $z_0(t) = z(t)$, $t \in [t_1, t_2]$. Then

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{t_1}^{t_2}\theta\left(\frac{\partial z_\lambda}{\partial t}\right)\mathrm{d}t\bigg|_{\lambda=0} = \int_{t_1}^{t_2}\omega\left(\frac{\partial z}{\partial t}, \frac{\partial z}{\partial \lambda}\right)\mathrm{d}t\bigg|_{\lambda=0}.$$

In other words,

$$\delta\int_{t_1}^{t_2}\theta(\dot z)\,\mathrm{d}t = \int_{t_1}^{t_2} -\mathrm{d}\theta(\dot z, \delta z)\,\mathrm{d}t = \int_{t_1}^{t_2} i_{\dot z}\omega(\delta z)\,\mathrm{d}t. \tag{5.7}$$

Proof. Consider a partition $t_1 = a_1 < a_2 \ldots a_n = t_2$ such that, for each $i = 1,2,\ldots, n-1$, the curve $z(t)$, $t \in [a_i, a_{i+1}]$ belongs to the domain of a chart of P. Therefore, in local coordinates, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{a_i}^{a_{i+1}}\left[\theta_i(z(t,\lambda))\cdot\frac{\partial z^i}{\partial t}(t,\lambda)\,\mathrm{d}t\right]\bigg|_{\lambda=0} &= \int_{a_i}^{a_{i+1}}\left[\frac{\partial\theta_i}{\partial z^j}\cdot\frac{\partial z^j}{\partial\lambda}\cdot\frac{\partial z^i}{\partial t} + \theta_i\frac{\partial^2 z^i}{\partial t\,\partial\lambda}\right]\mathrm{d}t\bigg|_{\lambda=0}\\
&= \int_{a_i}^{a_{i+1}}\left[\frac{\partial\theta_i}{\partial z^j}\cdot\frac{\partial z^j}{\partial\lambda}\cdot\frac{\partial z^i}{\partial t} - \frac{\partial\theta_i}{\partial z^j}\cdot\frac{\partial z^j}{\partial t}\cdot\frac{\partial z^i}{\partial\lambda}\right] + \left[\theta_i\frac{\partial z^i}{\partial\lambda}\right]\bigg|_{\lambda=0}\bigg|_{a_i}^{a_{i+1}}\\
&= \int_{a_i}^{a_{i+1}} -\mathrm{d}\theta\left(\frac{\partial z}{\partial t}, \frac{\partial z}{\partial\lambda}\right)\mathrm{d}t\bigg|_{\lambda=0} + \left[\theta\left(\frac{\partial z}{\partial\lambda}\right)\right]\bigg|_{\lambda=0}\bigg|_{a_i}^{a_{i+1}}.
\end{aligned} \tag{5.8}$$

Adding up expression (5.8) over $i = 1,2,\ldots, n-1$, we get

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{t_1}^{t_2}\theta\left(\frac{\partial z}{\partial t}\right)\mathrm{d}t\bigg|_{\lambda=0} = \int_{t_1}^{t_2} -\mathrm{d}\theta\left(\frac{\partial z}{\partial t}, \frac{\partial z}{\partial\lambda}\right)\mathrm{d}t\bigg|_{\lambda=0}, \tag{5.9}$$

since $(\partial z/\partial\lambda)(t_i) = 0$, $i = 1,2$. ∎

Proposition 5.5. Let (P, ω) be a symplectic manifold such that $\omega = -\mathrm{d}\theta$, and let G be a group of θ-preserving diffeomorphisms of P. Let $\mu \in \mathfrak{g}^*$ be chosen and define the equivalence relation $\sim$ on $\Omega(P)$ as follows: $z \sim z'$ if and only if there exists a curve h_λ on $\Omega(G_\mu)$, $\lambda \in [0,1]$ such that $h_\lambda(t_i) = e$ for $i = 1,2$

and $\lambda \in [0,1]$, $h_0(t) = e$ for $t \in [t_1, t_2]$ and $z' = hz$. Denote by $\bar{z}$ the equivalence class of $z \in \Omega(P)$, and, for $z_i \in J^{-1}(\mu)$, denote $\pi_\mu(z_i) = \bar{z}_i$.

Let $H\colon P \to \mathbb{R}$ be a G-invariant Hamiltonian and define $\mathscr{L}\colon \Omega(P) \to \mathbb{R}$ as follows:

$$\mathscr{L}(z) = \int_{t_1}^{t_2} [\theta(\dot{z}) - H(z)]\, \mathrm{d}t. \tag{5.10}$$

Then

a) $\mathscr{L}$ induces a function on $\Omega(P)/\sim$, the restriction of which to $\Omega(J^{-1}(\mu))/\sim$ will be denoted $\mathscr{L}_\mu$. The map $\beta\colon \Omega_{z_1,z_2}(J^{-1}(\mu))/\sim \to \Omega_{\bar{z}_1,\bar{z}_2}(P_\mu)$ is a local diffeomorphism and $\mathscr{L}_\mu$ induces a multivalued function $\bar{\mathscr{L}}_\mu\colon \Omega_{\bar{z}_1,\bar{z}_2}(P_\mu) \to \mathbb{R}$. We have the commutative diagram

$$\begin{tikzcd}
\mathbb{R} & & \Omega_{\bar{z}_1,\bar{z}_2}(P_\mu) \arrow[ll, "\bar{\mathscr{L}}_\mu"'] \\
& \Omega_{z_1,z_2}(J^{-1}(\mu))/\sim \arrow[ul, "\mathscr{L}_\mu"'] \arrow[ur, "\beta"] & \\
& \Omega_{z_1,z_2}(J^{-1}(\mu)) \arrow[u, "\alpha"'] \arrow[uul, "\mathscr{L}"] \arrow[uur, "\Omega_{z_1,z_2}(\pi_\mu)"'] &
\end{tikzcd}$$

b) The variation

$$\delta\bar{\mathscr{L}}_\mu\colon \Omega_{\bar{z}_1,\bar{z}_2}(P_\mu) \to \mathrm{T}^*\Omega_{\bar{z}_1,\bar{z}_2}(P_\mu) \tag{5.11}$$

is a well defined (single valued) function.

c) The critical points of $\delta\bar{\mathscr{L}}_\mu$, i.e., the points $\bar{z} \in \Omega_{\bar{z}_1,\bar{z}_2}(P_\mu)$ such that $\delta\bar{\mathscr{L}}_\mu(\bar{z})(\delta\bar{z}) = 0$ for all $\delta\bar{z} \in \mathrm{T}\Omega_{\bar{z}_1,\bar{z}_2}(P_\mu)$, are exactly the solutions of the equations of motion on the reduced space.

d) The functional $\delta\bar{\mathscr{L}}_\mu(\bar{z})(\delta\bar{z})$ on $\Omega_{\bar{z}_1,\bar{z}_2}(P_\mu)$ has the expression

$$\begin{aligned}
\delta\bar{\mathscr{L}}_\mu(\bar{z})(\delta\bar{z}) &= \int_{t_1}^{t_2} \left[\omega_\mu\left(\dot{\bar{z}}, \delta(\bar{z}) - \mathrm{d}H_\mu(\bar{z})(\delta\bar{z})\right)\right] \mathrm{d}t \\
&= \int_{t_1}^{t_2} [\omega(\dot{z}, \delta z) - \mathrm{d}H(z)(\delta z)]\, \mathrm{d}t.
\end{aligned} \tag{5.12}$$

Proof. a) The assertion about β being a local diffeomorphism amounts to construction of a map

$$\Omega_{\bar{z}_1,\bar{z}_2}(P_\mu) \to \Omega_{z_1,z_2}(J^{-1}(\mu))$$

in a neighborhood of each curve $\beta(z)$ such that, composed with β, this map gives the identity map. This can be readily achieved by using the fact that π_μ is a submersion.

To see that $\mathscr{L}_\mu$ is well defined, we should check that, for a curve $h_\lambda \in \Omega(G_\mu)$ such that $h_0(t) = e$ for $t \in [t_1, t_2]$ and $h_\lambda(t_i) = e$, $i = 1, 2$, and, say $z_\lambda = h_\lambda z$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \int_{t_1}^{t_2} [\theta(\dot{z}_\lambda) - H(z_\lambda)]\, \mathrm{d}t \bigg|_{\lambda=\lambda_0} = 0$$

for small values of λ_0. It is enough to check the case $\lambda_0 = 0$ since we can always replace z by $h_{\lambda_0}z$. Then, by using the lemma, we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{t_1}^{t_2}[\theta(\dot{z}_\lambda) - H(z_\lambda)]\,\mathrm{d}t\bigg|_{\lambda=0} = \int_{t_1}^{t_2}[\omega(\dot{z},\delta z) - \mathrm{d}H(z)(\delta z)],$$

where $\delta z = (\partial z/\partial\lambda)(t,0)$.

Now, from the reduction procedure we have

$$\omega(\dot{z},\delta z) - \mathrm{d}H(z)(\delta z) = \omega_\mu(\mathrm{T}\pi_\mu\dot{z}, \mathrm{T}\pi_\mu\delta z) - \mathrm{d}H_\mu\big(\pi_\mu(z)\big)(\mathrm{T}\pi_\mu\delta z).$$

But

$$\mathrm{T}\pi_\mu\delta z = \frac{\mathrm{d}}{\mathrm{d}\lambda}\pi_\mu(h_\lambda z)\bigg|_{\lambda=0} = \frac{\mathrm{d}}{\mathrm{d}\lambda}\pi_\mu(z)\bigg|_{\lambda=0} = 0,$$

which proves the assertion.

b) and c) follow from d)

d) As in the proof of a), we get

$$\begin{aligned}\delta\overline{\mathscr{L}}_\mu(\bar{z})(\delta\bar{z}) &= \delta\mathscr{L}(z)(\delta z)\\ &= \delta\int_{t_1}^{t_2}[\theta(\dot{z}) - H(z)]\,\mathrm{d}t\\ &= \int_{t_1}^{t_2}[\omega(\dot{z},\delta z) - \mathrm{d}H(z)(\delta z)]\,\mathrm{d}t\\ &= \int_{t_1}^{t_2}\big[\omega_\mu(\dot{\bar{z}},\delta\bar{z}) - \mathrm{d}H_\mu(\bar{z})(\delta\bar{z})\big]\,\mathrm{d}t.\end{aligned}$$ ∎

Remarks. 1) It may be of interest to compare the previous procedure with the variational principle given in Balachandran et al. [38]. There the action for a system (P, ω, H) is defined by

$$\int_\Delta\left[\omega\left(\frac{\partial z}{\partial t}, \frac{\partial z}{\partial\lambda}\right) - H(z(t,\lambda))\right]\mathrm{d}t\wedge\mathrm{d}\lambda,$$

where $\Delta\subset P$ is the surface defined by $z = z(t,\lambda)$, $t_1\le t\le t_2$, $0\le\lambda\le 1$, constrained as follows: $\delta z(t_1,\lambda) = 0$, $\delta z(t_2,\lambda) = 0$ for $\lambda\in[0,1]$. This amounts to considering the action as a functional on the "path space" of curves with fixed origin z_0. Thus $z(t,\lambda)$ is a curve on that path space and the condition $\delta z(t_i,\lambda) = 0$, $i = 1,2$, $\lambda\in[0,1]$ is the usual fixed endpoint condition.

2) The approach of Capriz [39] appears to be closely related to the Clebsch potential approach of Seliger and Whitham [6].

Remark. We will combine some cases of Clebsch representation, like these described in previous sections, with the reduction of a variational principle. Thus, let G be a group acting on a vector space V by a representation ρ: $G\times V\to V$ as in section 2. In particular, we have a momentum mapping J_G: $V\times V^*\to$

$\mathfrak{g}^*$. If K is the group of canonical transformations that leave J_G invariant, i.e., K is the Gauge group of J_G, then we have a dual pair (see Marsden and Weinstein [12] and also Weinstein [40])

$$\begin{tikzcd} & V \times V^* \arrow[dl, "J_G"'] \arrow[dr, "J_K"] & \\ \mathfrak{g}^* & & \mathfrak{k}^* \end{tikzcd}$$

The reduced spaces for the action of K are the symplectic leaves of $\mathfrak{g}^*$, i.e., the coadjoint orbits. Now suppose that a Hamiltonian is defined on $\mathfrak{g}^*$, giving rise to a Hamiltonian H on $V \times V^*$ by composition with J_G. Then H is K-invariant and we can apply the methods described in this section to get a variational principle on coadjoint orbits of $\mathfrak{g}^*$ starting with the variational principle with Lagrangian given by $\theta - H$ on $V \times V^*$, and reducing by the action of K.

This also gives an *interpretation of Lin constraints in intrinsic terms*. In fact, Lin constraints are restrictions on curves $(\boldsymbol{a}_t, \boldsymbol{b}_t) \in V \times V^*$ of the type $(\rho(g_t, \boldsymbol{a}_t), \rho^*(g_t^{-1}, \boldsymbol{b}_t)) = (\boldsymbol{a}_0, \boldsymbol{b}_0)$ fixed. Since J_K is G-invariant, this implies that $J_K(\boldsymbol{a}_t, \boldsymbol{b}_t) = \lambda$ is fixed. Now assume that, given $(\boldsymbol{a}_1, \boldsymbol{b}_1), (\boldsymbol{a}_2, \boldsymbol{b}_2) \in J_K^{-1}(\lambda)$, there exists $g \in G$ such that $(\rho(g, \boldsymbol{a}_1), \rho^*(g^{-1}, \boldsymbol{b}_1)) = (\boldsymbol{a}_2, \boldsymbol{b}_2)$. Then *Lin constraints would be the level sets of J_K, which in turn are in one to one correspondence with coadjoint orbits of G.*

Acknowledgements

We thank Darryl Holm, Aberto Ibort, Richard Montgomery, Tudor Ratiu, and Alan Weinstein for their valuable comments on various aspects of this work.

References

[1] A. Walter, Anwendung der Methode Hamiltons auf die Grundgleichungen der Mathematischen Theorie der Elasticitat, Dissertation, Berlin (1868).
[2] G. Kirchhoff, Vorlesungen Uber Mathematische Physik: Mechanik (Leipzig. 1876, third ed. 1883).
[3] C. Truesdell and R. Toupin, The classical field theories, Handbuch der Physik III/1, S. Flugge, ed. (Springer, Berlin, 1960).
[4] J. Serrin, Mathematical principles of classical fluid mechanics, in: Handbuch der Physik VII/1 (Spinger, Berlin, 1959).
[5] C.C. Lin, Liquid Helium, Proc. Int. School of Physics, Course XXI (Academic Press, New York, 1963).
[6] R.L. Seliger and G.B. Whitham, Variational principles in continuum mechanics, Proc. Roy. Soc. Lond. 305 (1968) 1–25.
[7] A. Clebsch, Uber eine allgemeine Transformation der hydrodynamischen Gleichungen, J. reine angew. Math. 54 (1857) 293–312. A. Clebsch, Uber die Integration der hydrodynamischen Gleichungen, J. reine angew. Math 56 (1859) 1–10.
[8] Z.R. Iwinski and L.A. Turski, Canonical theories of systems interacting electromagnetically, Lett. in Appl. and Eng. Sci. 4 (1976) 179–191.
[9] I.E. Dzyaloshinskii and G.E. Volovick, Poisson brackets in condensed matter physics, Ann. Phys. 125 (1980) 67–97.
[10] P.J. Morrison and J.M. Greene, Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics, Phys. Rev. Lett. 45 (1980) 790–794.
[11] J.E. Marsden and A. Weinstein, The Hamiltonian structure of the Maxwell-Vlasov equations, Physica 4D (1982) 394–406.
[12] J.E. Marsden and A. Weinstein, Coadjoint orbits, vortices and Clebsch variables for incompressible fluids, Physica 7D (1983) 305–323.

[13] D.D. Holm and B.A. Kuperschmidt, Poisson brackets dn Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity, Physica 6D (1983) 347–363.
D.D. Holm and B.A. Kuperschmidt, Poisson structures of superconductors. Phys. Lett. A 93 (1983) 177–181.
[14] J.E. Marsden, A. Weinstein, T. Ratiu, R. Schmid and R.G. Spencer, Hamiltonian Systems with Symmetry, Coadjoint Orbits and Plasma Physics, Proc. IUTAM-ISIMM Symposium on Modern Developments in Analytical Mechanics, Torino, June 7–11, 1982, Atti della Academia della Scienze di Torino 117 (1983) 289–340.
[15] T.B. Benjamin, Impulse, flow force and variational principles, IMA J. of Appl. Math. 32 (1984) 3–68.
[16] J.E. Marsden, T. Ratiu and A. Weinstein, Semi-direct products and reduction in mechanics, Trans. Am. Math. Soc. 281 (1984) 147–177.
[17] J.E. Marsden, T. Ratiu and A. Weinstein, Reduction and Hamiltonian structures on duals of semidirect product Lie algebras, Cont. Math. AMS 28 (1984) 55–100.
[18] J.E. Marsden and A. Weinstein, Reduction of symplectic manifolds with symmetry, Rep. Math. Phys. 5 (1974) 121–130.
[19] F.E. Low, A Lagrangian formulation of the Boltzmann–Vlasov equation for plasmas, Proc. Roy. Soc. Lond. A248 (1958) 282–287.
[20] D.D. Holm and B.A. Kuperschmidt, Poisson structures of superfluids, Phys. Lett. A 91 (1982) 425–430.
[21] P.S. Krishnaprasad and J.E. Marsden, Hamiltonian structure and stability for rigid bodies with flexible attachments, Arch. Rat. Mech. An. 98 (1987) 71–93.
E.Z. Kuznetzov and A.V. Mikhailov, On the topological meaning of canonical Clebsch variables, Phys. Lett. A 77 (1980) 37–38.
[22] J.C. Simo, J. Marsden, and P.S. Krishnaprasad, The Hamiltonian structure of nonlinear elasticity: the convective representation of solids, rods, and plates, to appear.
[23] R.R. Long, Some aspects of the flow of stratified fluids I. A. theoretical investigation, Tellus 5 (1953) 42–57.
[24] H. Abarbanel, D. Holm, J. Marsden and T. Ratiu, Nonlinear stability of stratified flow, Phys. Rev. Lett. 52 (1986) 2352–2355; Phil. Trans. Roy. Soc. Lond. A318 (1986) 349–409.
[25] D. Lewis, J.E. Marsden, R. Montgomery and T. Ratiu, The Hamiltonian structure for dynamic free boundary problems, Physica 18D (1986) 391–404.
[26] K.-K. Tam, A variational principle in relativistic magnetohydrodynamics, Can. J. Phys. 44 (1966) 2403–2409.
[27] K.-K. Tam and J. O'Hanlon, Relativistic magnetohydrodynamics of a gravitating fluid, Il Nuovo Cimento 62B (1969) 351–359.
[28] D. Bao, J. Marsden and R. Walton, The Hamiltonian structure of general relativistic perfect fluids, Comm. Math. Phys. 99 (1985) 319–345.
[29] D. Holm, Hamiltonian formalism of general relativistic adiabatic fluids, Physica 17D (1985) 1–36.
[30] Montgomery, Marsden and Ratiu, Gauged Lie–Poisson structures, Cont. Math. AMS 28 (1984) 101–114.
[31] Cendra, Ibort and Marsden, Variational principles on principal fiber bundles: a geometric theory of Clebsch potentials and Lin constraints, to be publ.
[32] R. Abraham and J. Marsden, Foundations of Mechanics, second ed. (Addison-Wesley, New York, 1978).
V. Arnold, Variational principle for three dimensional steady-state flow of an ideal fluid, J. Appl. Math. Mech. 29 (1965) 1002–1008.
[33] V. Arnold, Sur la geometrie differentielle des groupes de Lie de dimension infinie et ses aplications a l'hydrodynamique des fluids parfaits, Ann. Inst. Fourier, Grenoble 16 (1966) 319–361.
[34] D. Ebin and J.E. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, Ann. Math. 92 (1970) 102–163.
[35] D.D. Holm, J.E. Marsden and Ratiu, The Hamiltonian structure of continuum mechanics in material, inverse material, spatial and convective representations, seminaire de Mathematiques Superieures, Les Presses de L'Universite de Montreal, 100 (1986) 11–124.
[36] S.P. Novikov, The Hamiltonian formalism and many-valued analogue of Morse theory, Russian Math Surveys 37 (1982) 1–56.
[37] A. Weinstein, Bifurcations and Hamilton's principle, Math. Zeit. 159 (1978) 235–248.
[38] A.P. Balachandran, G. Marmo, B.S. Skagertan and A. Stern, Gauge Symmetry and Fiber Bundles, Springer Lect. Notes in Phys. 188 (Springer, Berlin, 1983).
[39] G. Capriz, Spatial variational principles in continuum mechanics, Arch. Rat. Mech. An. 85 (1984) 99–109.
[40] A. Weinstein, The local structure of Poisson manifolds, J. Diff. Geom. 18 (1983) 523–557.