13
Euler–Poincaré and Lie–Poisson Reduction

Besides the Poisson structure on a symplectic manifold, the Lie–Poisson bracket on $\mathfrak{g}^*$, the dual of a Lie algebra, is perhaps the most fundamental example of a Poisson structure. We shall obtain it in the following manner. Given two smooth functions $F, H \in \mathcal{F}(\mathfrak{g}^*)$, we extend them to functions, F_L, H_L (respectively, F_R, H_R) on all T^*G by left (respectively, right) translations. The bracket $\{F_L, H_L\}$ (respectively, $\{F_R, H_R\}$) is taken in the canonical symplectic structure Ω on T^*G. The result is then restricted to $\mathfrak{g}^*$ regarded as the cotangent space at the identity; this defines $\{F, H\}$. We shall prove that one get the Lie–Poisson bracket this way. In §14.6 we show that the symplectic leaves of this bracket are the coadjoint orbits in $\mathfrak{g}^*$.

There is another side to the story too, where the basic objects that are reduced are not Poisson brackets, but rather are variational principles. This aspect of the story, which takes place on $\mathfrak{g}$ rather than on $\mathfrak{g}^*$, will be told as well.

13.1 The Lie–Poisson Reduction Theorem

We begin by studying the way the canonical Poisson bracket on T^*G is related to the Lie–Poisson bracket on $\mathfrak{g}^*$.

Theorem 13.1.1 (The Lie–Poisson Reduction Theorem). *Identifying the set of functions on $\mathfrak{g}^*$ with the set of left (respectively, right) in-*

variant functions on T^*G endows $\mathfrak{g}^*$ with Poisson structures given by

$$\{F, H\}_{\pm}(\mu) = \pm\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\right\rangle. \tag{13.1.1}$$

The space $\mathfrak{g}^*$ with this Poisson structure is denoted $\mathfrak{g}^*_-$ (respectively, $\mathfrak{g}^*_+$). In contexts where the choice of left or right is clear, we shall drop the "$-$" or "$+$" from $\{F, H\}_-$ and $\{F, H\}_+$.

Following Marsden and Weinstein [1983], this bracket on $\mathfrak{g}^*$ is called the ***Lie–Poisson bracket*** after Lie [1890], p. 204, where the bracket is given explicitly. See Weinstein [1983a] and §13.8 below for more historical information. In fact, there are already some hints of this structure in Jacobi [1866], p. 7. It was rediscovered several times since Lie's work. For example, it appears explicitly in Berezin [1967]. It is closely related to results of Arnold, Kirillov, Kostant, and Souriau in the 1960s.

Before proving the theorem, we explain the terminology used in its statement. First, recall from Chapter 9 how the Lie algebra of a Lie group G is constructed. We define $\mathfrak{g} = T_eG$, the tangent space at the identity. For $\xi \in \mathfrak{g}$, we define a left invariant vector field $\xi_L = X_\xi$ on G by setting

$$\xi_L(g) = T_eL_g \cdot \xi \tag{13.1.2}$$

where $L_g : G \to G$ denotes left translation by $g \in G$ and is defined by $L_gh = gh$. Given $\xi, \eta \in \mathfrak{g}$, define

$$[\xi, \eta] = [\xi_L, \eta_L](e), \tag{13.1.3}$$

where the bracket on the right-hand side is the Jacobi–Lie bracket on vector fields. The bracket (13.1.3) makes $\mathfrak{g}$ into a Lie algebra, that is, $[\,,]$ is bilinear, antisymmetric, and satisfies Jacobi's identity. For example, if G is a subgroup of $\mathrm{GL}(n)$, the group of invertible $n \times n$ matrices, we identify $\mathfrak{g} = T_eG$ with a vector space of matrices and then, as we calculated in Chapter 9,

$$[\xi, \eta] = \xi\eta - \eta\xi \tag{13.1.4}$$

is the usual commutator of matrices.

A function $F_L : T^*G \to \mathbb{R}$ is called ***left invariant*** if, for all $g \in G$,

$$F_L \circ T^*L_g = F_L, \tag{13.1.5}$$

where T^*L_g denotes the cotangent lift of L_g, so T^*L_g is the pointwise adjoint of TL_g. Let $\mathcal{F}_L(T^*G)$ denote the space of all smooth left invariant functions on T^*G. One similarly defines ***right invariant*** functions on T^*G and the space $\mathcal{F}_R(T^*G)$. Given $F : \mathfrak{g}^* \to \mathbb{R}$ and $\alpha_g \in T^*G$, set

$$F_L(\alpha_g) = F(T_e^*L_g \cdot \alpha_g) = (F \circ \mathbf{J}_R)(\alpha_g), \tag{13.1.6}$$

where $\mathbf{J}_R : T^*G \to \mathfrak{g}^*$, $\mathbf{J}_R(\alpha_g) = T_e^*L_g \circ \alpha_g$ is the momentum map of the lift of right translation on G (see (12.2.8). $F_L = F \circ \mathbf{J}_R$ is called the ***left invariant extension*** of F from $\mathfrak{g}^*$ to T^*G. One similarly defines the ***right invariant extension*** by

this sentence about left inavriant extension should not begin with a math symbol. Can you modify it?

$$F_R(\alpha_g) = F(T_e^*R_g \cdot \alpha_g) = (F \circ \mathbf{J}_L)(\alpha_g), \tag{13.1.7}$$

where $\mathbf{J}_L : T^*G \to \mathfrak{g}^*$, $\mathbf{J}_L(\alpha_g) = T_e^*R_g$ is the momentum map of the lift of left translation on G (see (12.2.7)).

Right composition with $\mathbf{J}_R$ (respectively, $\mathbf{J}_L$) thus defines a ring isomorphism $\mathcal{F}(\mathfrak{g}^*) \to \mathcal{F}_L(T^*G)$ (respectively, $\mathcal{F}(\mathfrak{g}^*) \to \mathcal{F}_R(T^*G)$) whose inverse is restriction to the fiber $T_e^*G = \mathfrak{g}^*$.

Since T^*L_g and T^*R_g are symplectic maps on T^*G, it follows that $\mathcal{F}_L(T^*G)$ and $\mathcal{F}_R(T^*G)$ are closed under the usual Poisson bracket on T^*G. Thus, one way of rephrasing the Lie–Poisson reduction theorem (we will see another way, using quotients, in §13.5) is to say that the above ring isomorphisms of $\mathcal{F}(\mathfrak{g}^*)$ with $\mathcal{F}_L(T^*G)$ and $\mathcal{F}_R(T^*G)$ respectively, are also isomorphisms of Lie algebras, that is, the following pair fromulae are valid.

$$\{F, H\}_- = \{F_L, H_L\} \,|\mathfrak{g}^* \tag{13.1.8}$$

and

$$\{F, H\}_+ = \{F_R, H_R\} \,|\mathfrak{g}^*, \tag{13.1.9}$$

where $\{\,,\}_\pm$ is the Lie–Poisson bracket on $\mathfrak{g}^*$ and $\{\,,\}$ is the canonical bracket on T^*G.

Proof of the Theorem. $\mathbf{J}_R : T^*G \to \mathfrak{g}^*_-$ is a Poisson map by Theorem 12.4.1. Therefore,

$$\{F, H\}_- \circ \mathbf{J}_R = \{F \circ \mathbf{J}_R, H \circ \mathbf{J}_R\} = \{F_L, H_L\}.$$

Restriction of this relation to $\mathfrak{g}^*$ gives (13.1.8) One similarly proves (13.1.9) using the Poisson property of the map $\mathbf{J}_L : T^*G \to \mathfrak{g}^*_+$. ■

The proof above was *a posteriori*, that is, one had to already know the formula for the Lie–Poisson bracket. In §13.5 we will prove this theorem again using momentum functions and quotienting by G (see §10.7). This will represent an *a priori* proof, in the sense that the formula for the Lie–Poisson bracket will be deduced as part of the proof. To gain further insight into this, the next rhee sections will give constructive prrofs of this theorem in three special cases; the abstract proof can be found in §13.5

13.2 Proof of the Lie–Poisson Reduction Theorem for GL(n)

We now prove the Lie–Poisson reduction theorem for the special case of the matrix group $G = \mathrm{GL}(n)$ of real invertible $n \times n$ matrices. Left translation by $\mathbf{U} \in G$ is given by matrix multiplication: $L_{\mathbf{U}}\mathbf{A} = \mathbf{U}\mathbf{A}$. Identify the tangent space to G at $\mathbf{A}$ with the vector space of all $n \times n$ matrices, so for $\mathbf{B} \in T_{\mathbf{A}}G$,

$$T_{\mathbf{A}}L_{\mathrm{U}} \cdot \mathbf{B} = \mathbf{U}\mathbf{B}$$

as well, since $L_{\mathrm{U}}\mathbf{A}$ is linear in $\mathbf{A}$. The cotangent space is identified with the tangent space via the pairing

$$\langle \pi, \mathbf{B} \rangle = \operatorname{trace}(\pi^T \mathbf{B}), \tag{13.2.1}$$

where π^T is the transpose of π. The cotangent lift of L_{U} is thus given by

$$\langle T^* L_{\mathbf{U}} \pi, \mathbf{B} \rangle = \langle \pi, T L_{\mathbf{U}} \cdot \mathbf{B} \rangle = \operatorname{trace}(\pi^T \mathbf{U}\mathbf{B}),$$

that is,

$$T^* L_{\mathrm{U}} \pi = \mathbf{U}^T \pi. \tag{13.2.2}$$

Given functions $F, G : \mathfrak{g}^* \to \mathbb{R}$, let

$$F_L(\mathbf{A}, \pi) = F(\mathbf{A}^T \pi) \quad \text{and} \quad G_L(\mathbf{A}, \pi) = G(\mathbf{A}^T \pi) \tag{13.2.3}$$

be their left invariant extensions. By the chain rule, letting $\mu = \mathbf{A}^T \pi$, we get

$$\begin{aligned}
\mathbf{D}_{\mathrm{A}} F_L(\mathbf{A}, \pi) \cdot \delta\mathbf{A} &= \mathbf{D}F(\mathbf{A}^T \pi) \cdot (\delta\mathbf{A})^T \pi \\
&= \left\langle \frac{\delta F}{\delta \mu}, (\delta\mathbf{A})^T \pi \right\rangle \\
&= \operatorname{trace}\left(\pi^T \delta\mathbf{A} \frac{\delta F}{\delta \mu} \right).
\end{aligned} \tag{13.2.4}$$

The canonical bracket is therefore

$$\begin{aligned}
\{F_L, G_L\} &= \left\langle \frac{\delta F_L}{\delta \mathbf{A}}, \frac{\delta G_L}{\delta \pi} \right\rangle - \left\langle \frac{\delta G_L}{\delta \mathbf{A}}, \frac{\delta F_L}{\delta \pi} \right\rangle \\
&= \mathbf{D}_{\mathrm{A}} F_L(\mathbf{A}, \pi) \cdot \frac{\delta G_L}{\delta \pi} - \mathbf{D}_{\mathrm{A}} G_L(\mathbf{A}, \pi) \cdot \frac{\delta F_L}{\delta \pi}.
\end{aligned} \tag{13.2.5}$$

Since $\delta F_L / \delta \pi = \delta F / \delta \mu$ at the identity $\mathbf{A} = \mathrm{Id}$, where $\pi = \mu$, using (13.2.4), the Poisson bracket (13.2.5) becomes

$$\begin{aligned}
\{F_L, G_L\}(\mu) &= \operatorname{trace}\left(\mu^T \frac{\delta G}{\delta \mu} \frac{\delta F}{\delta \mu} - \mu^T \frac{\delta F}{\delta \mu} \frac{\delta G}{\delta \mu} \right) \\
&= -\left\langle \mu, \frac{\delta F}{\delta \mu} \frac{\delta G}{\delta \mu} - \frac{\delta G}{\delta \mu} \frac{\delta F}{\delta \mu} \right\rangle \\
&= -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta G}{\delta \mu} \right] \right\rangle,
\end{aligned} \tag{13.2.6}$$

which is the $(-)$Lie–Poisson bracket. This derivation can be adapted for other matrix groups, including the rotation group SO(3) as special cases. However, in the latter case, one has to be extremely careful to treat the orthogonality constraint properly.

13.3 Proof of the Lie–Poisson Reduction Theorem for $\mathrm{Diff}_{\mathrm{vol}}(M)$

Another special case is $G = \mathrm{Diff}_{\mathrm{vol}}(\Omega)$, the subgroup of the group of diffeomorphisms $\mathrm{Diff}(\Omega)$ of a region $\Omega \subset \mathbb{R}^3$, consisting of the volume-preserving diffeomorphisms. We shall treat $\mathrm{Diff}(\Omega)$ and $\mathrm{Diff}_{\mathrm{vol}}(\Omega)$ formally, although it is known how to handle the functional analysis issues involved (see Ebin and Marsden [1970] and Adams, Ratiu, and Schmid [1986a,b] and references therein). We shall prove (13.1.9) for this case.

For $\eta \in \mathrm{Diff}(\Omega)$, the tangent space at η is given by the set of maps $V : \Omega \to T\Omega$ satisfying $V(X) \in T_{\eta(X)}\Omega$, that is, vector fields over η. We think of V as a material velocity field. Thus, the tangent space at the identity is the space of vector fields on Ω (tangent to $\partial\Omega$). Given two such vector fields, their left Lie algebra bracket is related to the Jacobi–Lie bracket by (see Chapter 9):

$$[V, W]_{LA} = -\,[V, W]_{JL}\,,$$

that is,

$$[V, W]_{LA} = (W \cdot \nabla)V - (V \cdot \nabla)W, \tag{13.3.1}$$

as one finds using the definitions. Let us compute the *right* Lie–Poisson bracket on $\mathfrak{g}^*$. Right translation by φ on G is given by

$$R_\varphi \eta = \eta \circ \varphi. \tag{13.3.2}$$

Differentiating (13.3.2) with respect to η gives

$$TR_\varphi \cdot V = V \circ \varphi. \tag{13.3.3}$$

Identify $T_\eta G$ with those V's such that the vector field on $\mathbb{R}^3$ given by $\mathbf{v} = V \circ \eta^{-1}$, is divergence-free and identify $T^*_\eta G$ with $T_\eta G$ via the pairing

$$\langle \pi, V \rangle = \int_\Omega \pi \cdot V \, dx\, dy\, dz, \tag{13.3.4}$$

where $\pi \cdot V$ is the dot product on $\mathbb{R}^3$. By the change of variables formula, and the fact that $\varphi \in G$ has unit Jacobian,

$$\begin{aligned}\langle T^* R_\varphi \cdot \pi, V \rangle &= \langle \pi, TR_\varphi \cdot V \rangle \\ &= \int_\Omega \pi \cdot (V \circ \varphi)\, dx\, dy\, dz = \int_\Omega (\pi \circ \varphi^{-1}) \cdot V \, dx\, dy\, dz,\end{aligned}$$

so

$$T^* R_\varphi \cdot \pi = \pi \circ \varphi^{-1}. \tag{13.3.5}$$

If $F : \mathfrak{g}^* \to \mathbb{R}$ is given, its right invariant extension is

$$F_R(\eta, \pi) = F(\pi \circ \eta^{-1}). \tag{13.3.6}$$

Let us denote elements of $\mathfrak{g}^*$ by $\mathbf{M}$, so we are investigating the relation between the canonical bracket of F_R and H_R and the Lie–Poisson bracket of F and H via the relation

$$\mathbf{M} \circ \eta = \pi.$$

From (13.3.6) and the chain rule, we get

$$\begin{aligned} \mathbf{D}_\eta F_R(\mathrm{Id}, \pi) \cdot \mathbf{v} &= -\mathbf{D}_{\mathrm{M}} F(\mathbf{M}) \cdot \mathbf{D}_\eta \pi(\mathrm{Id}) \cdot \mathbf{v} \\ &= -\int_\Omega ((\mathbf{v} \cdot \nabla)\mathbf{M}) \cdot \frac{\delta F}{\delta \mathbf{M}}\, dx\, dy\, dz, \end{aligned} \tag{13.3.7}$$

where $\delta F/\delta \mathbf{M}$ is a divergence-free vector field parallel to the boundary. Since T^*G is not given as a product space, one has to worry about what it means to hold π constant in (13.3.7). We leave it to the ambitious reader to justify this formal calculation. Thus, the canonical bracket at the identity becomes

$$\begin{aligned} \{F_R, H_R\}(\mathrm{Id}, \pi) &= \int_\Omega \left(\frac{\delta F_R}{\delta \eta} \frac{\delta H_R}{\delta \pi} - \frac{\delta H_R}{\delta \eta} \frac{\delta F_R}{\delta \pi} \right) dx\, dy\, dz \\ &= \mathbf{D}_\eta F_R(\mathrm{Id}, \pi) \cdot \frac{\delta H_R}{\delta \pi} \\ &\qquad - \mathbf{D}_\eta H_R(\mathrm{Id}, \pi) \cdot \frac{\delta F_R}{\delta \pi}. \end{aligned} \tag{13.3.8}$$

At the identity, $\pi = \mathbf{M}$ and $\delta F_R/\delta\pi = \delta F/\delta\mathbf{M}$, so substituting this and (13.3.7) into (13.3.8), we get

$$\begin{aligned} &\{F_R, H_R\}(\mathrm{Id}, \mathbf{M}) \\ &= -\int_\Omega \left[\left(\frac{\delta H}{\delta \mathbf{M}} \cdot \nabla \right) \mathbf{M} \cdot \frac{\delta F}{\delta \mathbf{M}} - \left(\frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \right) \mathbf{M} \cdot \frac{\delta H}{\delta \mathbf{M}} \right] dx\, dy\, dz. \end{aligned} \tag{13.3.9}$$

Equation (13.3.9) may be integrated by parts to give

$$\begin{aligned} &\{F_R, H_R\}(\mathrm{Id}, \mathbf{M}) \\ &\quad = \int \mathbf{M} \cdot \left[\left(\frac{\delta H}{\delta \mathbf{M}} \cdot \nabla \right) \frac{\delta F}{\delta \mathbf{M}} - \left(\frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \right) \frac{\delta H}{\delta \mathbf{M}} \right] dx\, dy\, dz \\ &\quad = \int \mathbf{M} \cdot \left[\frac{\delta F}{\delta \mathbf{M}}, \frac{\delta H}{\delta \mathbf{M}} \right]_{LA} dx\, dy\, dz \end{aligned} \tag{13.3.10}$$

which is the "+" Lie–Poisson bracket. In doing this step note $\mathrm{div}(\delta H/\delta\mathbf{M}) = 0$ and since $\delta H/\delta\mathbf{M}$ and $\delta F/\delta\mathbf{M}$ are parallel to the boundary, no boundary term appears. When doing free boundary problems, these boundary terms are essential to retain (see Lewis, Marsden, Montgomery, and Ratiu [1986]).

For other diffeormorphism groups, it may be convenient to treat $\mathbf{M}$ as a one-form density rather than a vector field.

13.4 The Lie–Poisson Reduction Theorem using Momentum Functions

Now we turn to a constructive proof of the Lie–Poisson reduction theorem. We begin by observing that T^*G/G is diffeomorphic to $\mathfrak{g}^*$. To see this, note that the trivialization of T^*G by left translations given by

$$\lambda : \alpha_g \in T_g^*G \mapsto (g, T_e^*L_g(\alpha_g)) = (g, \mathbf{J}_R(\alpha_g)) \in G \times \mathfrak{g}^*$$

transofrms the usual cotangent lift of left translation on G into the G-action on $G \times g^*$ given by

$$g \cdot (h, \mu) = (gh,\, u), \tag{13.4.1}$$

for $g, h \in G$and $\mu \in \mathfrak{g}^*$. Therefore, T^*G/G is diffeomorphic to $(G \times \mathfrak{g}^*)/G$ which in turn equals $\mathfrak{g}^*$, since G does not act on $\mathfrak{g}$ (see (13.4.1)). Thus, we can regard $\mathcal{T}_R : T^*G \to \mathfrak{g}^*$ as the canonical projection $T^*G \to T^*G/G$ and, as a consequence of Theorem **??**, $\mathfrak{g}^*$ inherits a Poisson bracket, which we will call $\{\,,\,\}$—for the time being, uniquely characterized by the relation:

$$\{F, H\}_- \circ \mathbf{J}_R = \{F \circ \mathbf{J}_R, H \circ \mathbf{J}_R\} \tag{13.4.2}$$

for any functions $F, H \in \mathcal{F}(\mathfrak{g}^*)$. The goal of this section is to explicitly compute this bracket $\{,\}_-$ and to discover at the end that it equals the $(-)$ Lie–Poisson bracket.

using properties of momentum functions. It will be useful to recall that the Poisson bracket $\{F, H\}$ depends only on the linearization of F and H at each point, so in determining the canonical bracket on T^*G, we can assume the functions in question are linear on fibers.

Proof of the Lie–Poisson Reduction Theorem. Recall that the space $\mathcal{F}_L(T^*G)$ of left invariant functions on T^*G is isomorphic (as a vector space) to $\mathcal{F}(\mathfrak{g}^*)$, the space of all functions on the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$ of G. This isomorphism is given by $F \in \mathcal{F}(\mathfrak{g}^*) \leftrightarrow F_L \in \mathcal{F}_L(T^*G)$, where

overflow---needs to be reworded

$$F_L(\alpha_g) = F(T^*L_g \cdot \alpha_g). \tag{13.4.3}$$

Since $\mathcal{F}_L(T^*G)$ is closed under bracketing (T^*L_g is a symplectic map), $\mathcal{F}(\mathfrak{g}^*)$ gets endowed with a unique Poisson structure. As we remarked before, it is enough to consider the case in which F is replaced by its linearization at a particular point. This means it is enough to prove the Lie–Poisson reduction theorem for linear functions on $\mathfrak{g}^*$. If F is linear, we can write $F(\mu) = \langle \mu, \delta F/\delta\mu \rangle$, where $\delta F/\delta\mu$ is a constant on $\mathfrak{g}$, so that letting $\mu = T^*L_g \cdot \alpha_g$, we get

$$\begin{aligned} F_L(\alpha_g) = F(T^*L_g \cdot \alpha_g) &= \left\langle T^*L_g \cdot \alpha_g, \frac{\delta F}{\delta \mu} \right\rangle \\ &= \left\langle \alpha_g, TL_g \cdot \frac{\delta F}{\delta \mu} \right\rangle = \mathcal{P}\left(\left(\frac{\delta F}{\delta \mu} \right)_L \right)(\alpha_g), \end{aligned} \tag{13.4.4}$$

where $\xi_L(g) = T_e L_g(\xi)$ is the left invariant vector field on G whose value at e is $\xi \in \mathfrak{g}$. Thus, by (12.1.2), (13.4.4), and the definition of the Lie algebra bracket, we have

$$\begin{aligned}
\{F_L, H_L\}(\mu) &= \left\{\mathcal{P}\left(\left(\frac{\delta F}{\delta \mu}\right)_L\right), \mathcal{P}\left(\left(\frac{\delta H}{\delta \mu}\right)_L\right)\right\}(\mu) \\
&= -\mathcal{P}\left(\left[\left(\frac{\delta F}{\delta \mu}\right)_L, \left(\frac{\delta H}{\delta \mu}\right)_L\right]\right)(\mu) \\
&= -\mathcal{P}\left(\left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]_L\right)(\mu) \\
&= -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\right\rangle
\end{aligned} \tag{13.4.5}$$

as required.

The formula with "+" follows by using right invariant extensions of linear functions since the Lie bracket of two right invariant vector fields equals minus the Lie algebra bracket of their generators.

Finally, we observe that since T^*G is a Poisson manifold, formulae (13.1.8) and (13.1.9) show that $\mathfrak{g}^*_-$ and $\mathfrak{g}^*_+$ inherit the same properties, so they are Poisson as well. ■

13.5 Reduction and Reconstruction of Dynamics

In the examples in subsequent sections, we will use the Lie–Poisson reduction theorem in the following way:

Theorem 13.5.1 (Lie–Poisson Reduction of Dynamics). *Let G be a Lie group and $H : T^*G \to \mathbb{R}$. Assume H is left (respectively, right) invariant. Then the function $H^- := H|\mathfrak{g}^*$ (respectively, $H^+ := H|\mathfrak{g}^*$) on $\mathfrak{g}^*$ satisfies*

$$H(\alpha_g) = H^-(\pi_L(\alpha_g)) \quad \textit{for all } \alpha_g \in T^*_g G \tag{13.5.1}$$

where $\pi_L : T^*G \to \mathfrak{g}^*_-$ is given by $\pi_L(\alpha_g) = T^*L_g \cdot \alpha_g$ (respectively,

$$H(\alpha_g) = H^+(\pi_R(\alpha_g)) \quad \textit{for all } \alpha_g \in T^*_g G, \tag{13.5.2}$$

where $\pi_R : T^*G \to \mathfrak{g}^*_+$ is given by $\pi_R(\alpha_g) = T^*R_g \cdot \alpha_g$).

The flow F_t of H on T^*G and the flow F_t^- (respectively, F_t^+) of H^- (respectively, H^+) on $\mathfrak{g}^*_-$ (respectively, $\mathfrak{g}^*_+$) are related by

$$\pi_L(F_t(\alpha_g)) = F_t^-(\pi_L(\alpha_g)), \tag{13.5.3}$$

$$\pi_R(F_t(\alpha_g)) = F_t^+(\pi_R(\alpha_g)). \tag{13.5.4}$$

In other words, a left invariant Hamiltonian on T^*G induces Lie–Poisson dynamics on $\mathfrak{g}^*_-$, while a right invariant one induces Lie–Poisson dynamics on $\mathfrak{g}^*_+$. The result is a direct consequence of the Lie–Poisson reduction theorem and the fact that a Poisson map relates Hamiltonian systems and their integral curves to Hamiltonian systems. As we shall see in Volume II, this is a special case of the reduction procedure.

As we have remarked, the maps λ and ρ induce Poisson isomorphisms between $(T^*G)/G$ and $\mathfrak{g}^*$ (with the $-$ and $+$ brackets, respectively) and this is a special instance of Poisson reduction. The following result is one useful way of formulating the general relation between T^*G and $\mathfrak{g}^*$. We treat the left invariant case for simplicity.

Theorem 13.5.2. *Let G be a Lie group and let $H : T^*G \to \mathbb{R}$ be a left invariant Hamiltonian. Let $h : \mathfrak{g}^* \to \mathbb{R}$ be the restriction of H to T^*_eG. For a curve $p(t) \in T^*_{g(t)}G$, let $\mu(t) = (T^*_{g(t)}L) \cdot p(t) = \lambda(p(t))$ be the induced curve in $\mathfrak{g}^*$. Assuming that $g(t)$ satisfies the differential equation*

$$\dot{g} = T_eL_g \frac{\delta h}{\delta \mu},$$

where $\mu = p(0)$, the following are equivalent:

(i) *$p(t)$ is an integral curve of X_H; that is, Hamilton's equations on T^*G hold;*

(ii) *for any $F \in \mathcal{F}(T^*G)$, $\dot{F} = \{F, H\}$, where $\{\,,\}$ is the canonical bracket on T^*G;*

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

$$\frac{d\mu}{dt} = \mathrm{ad}^*_{\delta h/\delta\mu}\mu, \tag{13.5.5}$$

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

$$\dot{\mu}_a = C^d_{ba} \frac{\delta h}{\delta \mu_b} \mu_d; \tag{13.5.6}$$

(iv) *for any $f \in \mathcal{F}(\mathfrak{g}^*)$, we have*

$$\dot{f} = \{f, h\}_-, \tag{13.5.7}$$

where $\{\,,\}_-$ is the minus Lie–Poisson bracket.

Proof. First of all, the equivalence of (i) and (ii) is general for any cotangent bundle, as we know. The equivalence of (ii) and (iv) follows from the fact that λ is a Poisson map and $H = h \circ \lambda$. Finally, we establish the equivalence of (iii) and (iv). Indeed, $\dot{f} = \{f, h\}_-$ means

$$\begin{aligned}\left\langle \dot{\mu}, \frac{\delta f}{\delta \mu} \right\rangle &= -\left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta h}{\delta \mu}\right] \right\rangle \\ &= \left\langle \mu, \operatorname{ad}_{\delta h/\delta \mu} \frac{\delta f}{\delta \mu} \right\rangle \\ &= \left\langle \operatorname{ad}^*_{\delta h/\delta \mu} \mu, \frac{\delta f}{\delta \mu} \right\rangle .\end{aligned}$$

Since f is arbitrary, this is equivalent to (iii). ■

It is useful to keep in mind that the Hamiltonian H on T^*G generally arises from a Lagrangian $L : TG \to \mathbb{R}$ via a Legendre transform $\mathbb{F}L$. In fact, many of the constructions and verifications are simpler using the Lagrangian. Assume that L is left invariant (respectively, right invariant); that is,

$$L(TL_g \cdot v) = L(v), \tag{13.5.8}$$

respectively,

$$L(TR_g \cdot v) = L(v) \tag{13.5.9}$$

for all $g \in G$ and $v \in T_hG$. Differentiating (13.5.8) and (13.5.9), we find

$$\mathbb{F}L(TL_g \cdot v) \cdot (TL_g \cdot w) = \mathbb{F}L(v) \cdot w, \tag{13.5.10}$$

respectively,

$$\mathbb{F}L(TR_g \cdot v) \cdot (TR_g \cdot w) = \mathbb{F}L(v) \cdot w \tag{13.5.11}$$

for all $v, w \in T_hG$ and $g \in G$. In other words,

$$T^*L_g \circ \mathbb{F}L \circ TL_g = \mathbb{F}L, \tag{13.5.12}$$

respectively,

$$T^*R_g \circ \mathbb{F}L \circ TR_g = \mathbb{F}L. \tag{13.5.13}$$

Note that the action of L is left (respectively, right) invariant

$$A(TL_g \cdot v) = A(v), \tag{13.5.14}$$

respectively,

$$A(TR_g \cdot v) = A(v) \tag{13.5.15}$$

since $A(TL_g \cdot v) = \mathbb{F}L(TL_g \cdot v) \cdot (TL_g \cdot v) = \mathbb{F}L(v) \cdot v = A(v)$ by (13.5.10). Thus, the energy $E = A - L$ is left (respectively, right) invariant on TG. If L is hyperregular, so $\mathbb{F}L : TG \to T^*G$ is a diffeomorphism, then $H = E \circ (\mathbb{F}L)^{-1}$ is left (respectively, right) invariant on T^*G.

The Lagrangian formalism is also useful for the reconstruction process in which we reconstruct the dynamics on T^*G or TG from that on $\mathfrak{g}^*$:

$$T^*G \underset{\text{reconstruction}}{\overset{\text{Lie–Poisson reduction}}{\rightleftarrows}} \mathfrak{g}^*$$

Theorem 13.5.3 (Lie–Poisson Reconstruction Theorem). *Let $L : TG \to \mathbb{R}$ be a hyperregular Lagrangian which is left (respectively, right) invariant on TG. Let $H : T^*G \to \mathbb{R}$ be the associated Hamiltonian and $H^- : \mathfrak{g}^*_- \to \mathbb{R}$ (respectively, $H^+ : \mathfrak{g}^*_+ \to \mathbb{R}$) be the induced Hamiltonian on $\mathfrak{g}^*$. Let $\mu(t) \in \mathfrak{g}^*$ be an integral curve for H^- (respectively, H^+) with initial condition $\mu(0) = T_e^*L_{g_0} \cdot \alpha_{g_0}$ (respectively, $\mu(0) = T_e^*R_{g_0} \cdot \alpha_{g_0}$) and let $\xi(t) = \mathbb{F}L^{-1}\mu(t) \in \mathfrak{g}$. Let*

$$v_0 = TL_{g_0} \cdot \xi(0) \in T_{g_0}G.$$

Then the integral curve for the Langangian L with initial condition (g_0, v_0) is given by

$$V_L(t) = TL_{g(t)} \cdot \xi(t), \tag{13.5.16}$$

respectively,

$$V_R(t) = TR_g(t) \cdot \xi(t), \tag{13.5.17}$$

where $g(t)$ solves the equation

$$\frac{dg}{dt} = TL_{g(t)} \cdot \xi(t), \tag{13.5.18}$$

respectively,

$$\frac{dg}{dt} = TR_{g(t)} \cdot \xi(t). \tag{13.5.19}$$

The corresponding integral curve of X_H on T^*G with initial condition α_{g_0} and covering $\mu(t)$ is

$$\alpha(t) = \mathbb{F}L(V_L(t)) = T^*L_{(g_0 g(t))^{-1}}\mu(t), \tag{13.5.20}$$

respectively,

$$\alpha(t) = \mathbb{F}L(V_R(t)) = T^*R_{(g_0(g(t))^{-1}}\mu(t). \tag{13.5.21}$$

Proof. According to Lie–Poisson reduction of dynamics, the integral curve of X_{H^-} on $\mathfrak{g}^*_-$ associated to an integral curve $\alpha(t) \in T^*_{g(t)}G$ of X_H is

$$\mu(t) = T^* L_{g(t)} \cdot \alpha(t). \tag{13.5.22}$$

Applying $\mathbb{F}L^{-1}$ to (13.5.22) gives

$$\xi(t) = TL_{g(t)^{-1}} \cdot V(t) \tag{13.5.23}$$

for the corresponding integral curve of the Lagrangian, that is, $V(t) = TL_{g(t)} \cdot \xi(t)$. Since X_E is a second-order equation, $dg/dt = V$, so we get the result. ■

Thus, given $\xi(t)$, one solves (13.5.18) for $g(t)$ and then constructs $V(t)$ or $\alpha(t)$ from (13.5.16) and (13.5.20). As we shall see in the examples, this procedure has a natural physical interpretation. The previous theorem generalizes to general Lagrangian systems in the following way. In fact, Theorem 13.6.3 is a corollary of the next theorem.

Theorem 13.5.4 (Lagrangian Lie–Poisson Reconstruction). *Let $L : TG \to \mathbb{R}$ be a left invariant Lagrangian such that its Lagrangian vector field $Z \in \mathfrak{X}(TG)$ is a second-order equation and is left invariant. Let $Z_G \in \mathfrak{X}(\mathfrak{g})$ be the induced vector field on $(TG)/G \approx \mathfrak{g}$ and let $\xi(t)$ be an integral curve of Z_G. If $g(t) \in G$ is the solution of the nonautonomous ordinary differential equation $\dot{g}(t) = T_e L_{g(t)}\xi(t)$, $g(0) = e$, and $g \in G$, then $V(t) = T_e L_{gg(t)}\xi(t)$ is the integral curve of Z satisfying $V(0) = T_e L_g \xi(0)$ and $V(t)$ projects to $\xi(t)$, that is, $TL_{\tau(V(t))^{-1}} V(t) = \xi(t)$, where $\tau : TG \to G$ is the tangent bundle projection.*

Proof. Let $V(t)$ be the integral curve of Z satisfying $V(0) = T_e L_g \xi(0)$ for a given element $\xi(0) \in \mathfrak{g}$. Since $\xi(t)$ is the integral curve of Z_G whose flow is conjugated to the flow of Z by left translation, we have $TL_{\tau(V(t))^{-1}} V(t) = \xi(t)$. If $h(t) = \tau(V(t))$, since Z is a second-order equation, we have

$$V(t) = \dot{h}(t) = T_e L_{h(t)}\xi(t), \quad h(0) = \tau(V(0)) = g,$$

so that letting $g(t) = g^{-1}h(t)$ we get $g(0) = e$ and

$$\dot{g}(t) = TL_{g^{-1}}\dot{h}(t) = TL_{g^{-1}} TL_{h(t)}\xi(t) = TL_{g(t)}\xi(t).$$

This determines $g(t)$ uniquely from $\xi(t)$ and so

$$V(t) = T_e L_{h(t)}\xi(t) = T_e L_{gg(t)}\xi(t). \qquad ■$$

These calculations suggest rather strongly that one should examine the Lagrangian (rather than the Hamiltonian) side of the story on an independent footing. We will do exactly that shortly.

Since Poisson brackets and Hamilton's equations naturally drop from T^*G to $\mathfrak{g}^*$, it is natural to ask if other structures do too, such as Hamilton–Jacobi theory. We investigate this question now. We shall just state the results of Ge and Marsden [1988], omitting the proofs.

Let H be a G invariant function on T^*G and let H_L be the corresponding left reduced Hamiltonian on $\mathfrak{g}^*$. (To be specific, we deal with left actions; of course, there are similar statements for right reduced Hamiltonians). If S is invariant, there is a unique function S_L such that $S(g, g_0) = S_L(g^{-1}g_0)$. (One gets a slightly different representation for S by writing $g_0^{-1}g$ in place of $g^{-1}g_0$.)

Proposition 13.5.5. *The* ***left*** *reduced Hamilton–Jacobi equation is the following equation for a function $S_L : G \to \mathbb{R}$:*

$$\frac{\partial S_L}{\partial t} + H_L(-TR_g^* \cdot \mathbf{d}S_L(g)) = 0, \tag{13.5.24}$$

which we call the ***Lie–Poisson Hamilton–Jacobi equation****. The Lie–Poisson flow of the Hamiltonian H_L is generated by the solution S_L of* (13.5.24) *in the sense that the flow is given by the Poisson transformation of $\mathfrak{g}^* : \Pi_0 \mapsto \Pi$ defined as follows. Define $g \in G$ by solving the equation*

$$\Pi_0 = -TL_g^* \cdot \mathbf{d}_g S_L \tag{13.5.25}$$

for $g \in G$ and then set

$$\Pi = \mathrm{Ad}_{g^{-1}}^* \, \Pi_0. \tag{13.5.26}$$

Here Ad denotes the adjoint action and so the action in (13.5.26) is the coadjoint action. Note that (13.5.26) and (13.5.25) give $\Pi = -TR_g^* \cdot \mathbf{d}S_L(g)$.

13.6 The Linearized Lie–Poisson Bracket

Here we show that the equations linearized about an equilibrium solution of a Lie–Poisson system (such as the ideal fluid equations) are Hamiltonian with respect to a "constant coefficient" Lie–Poisson bracket. The Hamiltonian for these linearized equations is $\frac{1}{2}\delta^2 (H + C)|_e$, the quadratic functional obtained by taking one-half of the second variation of the Hamiltonian plus conserved quantities and evaluating it at the equilibrium solution where the conserved quantity C (often a Casimir) is chosen so that the first variation $\delta(H + C)$ vanishes at the equilibrium. A consequence is that the linearized dynamics preserves $\frac{1}{2}\delta^2 (H + C)|_e$. This is useful for studying stability of the linearized equations.

For a Lie algebra $\mathfrak{g}$, the Lie–Poisson bracket is defined on $\mathfrak{g}^*$, the dual of $\mathfrak{g}$ with respect to a weakly nondegenerate pairing $\langle \, , \rangle$ between $\mathfrak{g}^*$ and $\mathfrak{g}$ by the usual formula

$$\{F, G\}(\mu) = \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta G}{\delta \mu} \right] \right\rangle, \tag{13.6.1}$$

where $\delta F/\delta\mu \in \mathfrak{g}$ is determined by

$$\mathbf{D}F(\mu) \cdot \delta\mu = \left\langle \delta\mu, \frac{\delta F}{\delta \mu} \right\rangle \tag{13.6.2}$$

when such an element $\delta F/\delta\mu$ exists, for any μ, $\delta\mu \in \mathfrak{g}^*$. The equations of motion are

$$\frac{d\mu}{dt} = -\operatorname{ad}\left(\frac{\delta H}{\delta \mu}\right)^* \mu, \tag{13.6.3}$$

where $H : \mathfrak{g}^* \to \mathbb{R}$ is the Hamiltonian, $\operatorname{ad}(\xi) : \mathfrak{g} \to \mathfrak{g}$ is the adjoint action, $\operatorname{ad}(\xi) \cdot \eta = [\xi, \eta]$ for $\xi, \eta \in \mathfrak{g}$, and $\operatorname{ad}(\xi)^* : \mathfrak{g}^* \to \mathfrak{g}^*$ is its dual. Let $\mu_e \in \mathfrak{g}^*$ be an equilibrium solution of (13.6.3). The linearized equations of (13.6.3) at μ_e are obtained by expanding in a Taylor expansion with small parameter ε using $\mu = \mu_e + \varepsilon\delta\mu$, and taking $(d/d\varepsilon)|_{\varepsilon=0}$ of the resulting equations. This gives

$$\frac{\delta H}{\delta \mu} = \frac{\delta H}{\delta \mu_e} + \varepsilon \mathbf{D}\left(\frac{\delta H}{\delta \mu}\right)(\mu_e) \cdot \delta\mu + O(\varepsilon^2), \tag{13.6.4}$$

where $\langle \delta H/\delta\mu_e, \delta\mu \rangle := \mathbf{D}H(\mu_e) \cdot \delta\mu$, and the derivative $\mathbf{D}(\delta H/\delta\mu)(\mu_e) \cdot \delta\mu$ is the linear functional

$$\nu \in \mathfrak{g}^* \mapsto \mathbf{D}^2 H(\mu_e) \cdot (\delta\mu, \nu) \in \mathbb{R} \tag{13.6.5}$$

by using the definition (13.6.2). Since $\delta^2 H := \mathbf{D}^2 H(\mu_e) \cdot (\delta\mu, \delta\mu)$, it follows that the functional (13.6.5) equals $\frac{1}{2}\delta(\delta^2 H)/\delta(\delta\mu)$. Consequently, (13.6.4) becomes

$$\frac{\delta H}{\delta \mu} = \frac{\delta H}{\delta \mu_e} + \frac{1}{2}\varepsilon \frac{\delta(\delta^2 H)}{\delta(\delta\mu)} + O(\varepsilon^2) \tag{13.6.6}$$

and the Lie–Poisson equations (13.6.3) yield

$$\begin{aligned} \frac{d\mu_e}{dt} + \varepsilon \frac{d(\delta\mu)}{dt} = &- \operatorname{ad}\left(\frac{\delta H}{\delta \mu_e}\right)^* \mu_e \\ &- \frac{1}{2}\varepsilon \left[\operatorname{ad}\left(\frac{\delta(\delta^2 H)}{\delta(\delta\mu)}\right)^* \mu_e - \operatorname{ad}\left(\frac{\delta H}{\delta \mu_e}\right)^* \delta\mu \right] + O(\varepsilon^2). \end{aligned}$$

Thus, the linearized equations are

$$\frac{d(\delta\mu)}{dt} = -\frac{1}{2}\operatorname{ad}\left(\frac{\delta(\delta^2 H)}{\delta(\delta\mu)}\right)^* \mu_e - \operatorname{ad}\left(\frac{\delta H}{\delta \mu_e}\right)^* \delta\mu. \tag{13.6.7}$$

If H is replaced by $H_C := H + C$, with the function C chosen to satisfy $\delta H_C/\delta\mu_e = 0$, we get $\mathrm{ad}(\delta H_C/\delta\mu_e)^*\delta\mu = 0$, and so

$$\frac{d(\delta\mu)}{dt} = -\frac{1}{2}\,\mathrm{ad}\left(\frac{\delta(\delta^2 H_C)}{\delta(\delta\mu)}\right)^*\mu_e. \tag{13.6.8}$$

Equation (13.6.8) is Hamiltonian with respect to the linearized Poisson bracket (see Example (f) of §10.2):

$$\{F,G\}(\mu) = \left\langle \mu_e, \left[\frac{\delta F}{\delta\mu}, \frac{\delta G}{\delta\mu}\right]\right\rangle. \tag{13.6.9}$$

Ratiu [1982] interprets this bracket in terms of a Lie–Poisson structure of a loop extension of $\mathfrak{g}$. The Poisson bracket (13.6.9) differs from the Lie–Poisson bracket (13.6.1) in that it is *constant* in μ. With respect to the Poisson bracket (13.6.9), Hamilton's equations given by $\delta^2 H_C$ are (13.6.8), as an easy verification shows. Note that the critical points of $\delta^2 H_C$ are stationary solutions of the linearized equation (13.6.8), that is, they are *neutral modes* for (13.6.8).

If $\delta^2 H_C$ is definite, then either $\delta^2 H_C$ or $-\delta^2 H_C$ is positive-definite and hence defines a norm on the space of perturbations $\delta\mu$ (which is $\mathfrak{g}^*$). Being twice the Hamiltonian function for (13.6.8), $\delta^2 H_C$ is conserved. So, any solution of (13.6.8) starting on an energy surface of $\delta^2 H_C$ (that is, on a sphere in this norm) stays on it and hence the zero solution of (13.6.8) is (Liapunov) stable. *Thus, formal stability, (that is, $\delta^2 H_C$ definite) implies linearized stability.* It should be noted, however, that the conditions for definiteness of $\delta^2 H_C$ are entirely different from the conditions for "normal mode stability," that is, that the operator acting on $\delta\mu$ given by (13.6.8) have a purely imaginary spectrum. In particular, having a purely imaginary spectrum for the linearized equation does *not* produce Liapunov stability of the linearized equations. The difference between $\delta^2 H_C$ and the operator in (13.6.8) can be made explicit, as follows. Assume that the pairing $\langle\,,\rangle$ identifies the dual $\mathfrak{g}^*$ with $\mathfrak{g}$ itself, that is, there is a weak Ad-invariant metric $\langle\langle\,,\rangle\rangle$ on $\mathfrak{g}$ and a linear operator $L : \mathfrak{g} \to \mathfrak{g}$ such that

$$\delta^2 H_C = \langle\langle \delta\mu, L\delta\mu\rangle\rangle\,; \tag{13.6.10}$$

L is symmetric with respect to the metric $\langle\langle\,,\rangle\rangle$, that is, $\langle\langle \xi, L\eta\rangle\rangle = \langle\langle L\xi, \eta\rangle\rangle$ for all $\xi, \eta \in \mathfrak{g}$. Then the linear operator in (13.6.8) becomes

$$\delta\mu \mapsto [L\delta\mu, \mu_e] \tag{13.6.11}$$

which, of course, differs from L, in general. However, note that the kernel of L is included in the kernel of the linear operator (13.6.11), that is, the zero eigenvalues of L give rise to "neutral modes" in the spectral analysis of (13.6.11). There is a remarkable coincidence of the zero-eigenvalue equations for these operators in fluid mechanics: for the Rayleigh equation describing plane-parallel shear flow in an inviscid homogeneous fluid, taking normal modes makes the zero-eigenvalue equations corresponding to L and to (13.6.11) coincide (see Abarbanel, Holm, Marsden, And Ratiu [1986]).

For additional applications of the stability method, see the Introduction and Holm, Marsden, Ratiu, and Weinstein [1985], Abarbanel and Holm [1987], Simo, Posbergh, and Marsden [1990, 1991], and Simo, Lewis, and Marsden [1991]. For a more general treatment of the linearization process, see Marsden, Ratiu, and Raugel [1991].

Some History of Lie–Poisson and Euler–Poincaré Equations. We continue with some comments on the history of Poisson structures that we began in §10.3. Recall that we pointed out how Lie, in his work up to 1890 on function groups, had many of the essential ideas of general Poisson manifolds and, in particular, had explicitly studied the Lie–Poisson bracket on duals of Lie algebras.

The theory developed so far in this chapter describes the adaptation of the concepts of Hamiltonian mechanics to the context of the duals of Lie algebras. This theory could easily have been given shortly after Lie's work, but evidently it was not observed for the rigid body or ideal fluids until the work of Pauli [1953], Martin [1959], Arnold [1966a], Ebin and Marsden [1970], Nambu [1973], and Sudarshan and Mukunda [1974], all of whom were apparently unaware of Lie's work on the Lie–Poisson bracket. It seems that even Elie Cartan was unaware of this aspect of Lie's work, which does seem surprising. Perhaps it is less surprising when one thinks for a moment about how many other things Cartan was involved in at the time. Nevertheless, one is struck by the amount of rediscovery and confusion in this subject. Evidently, this situation is not unique to mechanics.

Meanwhile, as Arnold [1988] and Chetaev [1989] pointed out, one can also write the equations directly on the Lie algebra, bypassing the Lie–Poisson equations on the dual. The resulting equations were first written down on a general Lie algebra by Poincaré [1901b]; we refer to these as the Euler–Poincaré equations. We shall develop them from a modern point of view in the next section. Poincaré [1910] 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. In general, the command that Poincaré had of the subject is most impressive, and is hard to match in his near contemporaries, except perhaps Riemann [1860, 1861] and Routh [1877, 1884]. It is noteworthy that Poincaré [1901b] has no references, so it is rather hard to trace his train of thought or his sources; compare this style with that of Hamel [1904]! In particular, he gives no hint that he understood the work of Lie on the Lie–Poisson structure, but, of course, Poincaré understood the Lie group and the Lie algebra machine very well indeed.

Our derivation of the Euler–Poincaré equations in the next section is based on a reduction of variational principles, not on a reduction of the symplectic or Poisson structure, which is natural for the dual. We also show that the Lie–Poisson equations are related to the Euler–Poincaré equations by the "fiber derivative," in the same way as one gets from the ordinary Euler–Lagrange equations to the Hamilton equations. Even though this is relatively trivial, it does not appear to have been written down before. In the dynamics of ideal fluids, the resulting variational principle is related to what has been known as "Lin constraints" (see also Newcomb [1962] and Bretherton [1970].) This itself has an interesting history, going back to Ehrenfest, Boltzman, and Clebsch, but again, there was little if any contact with the heritage of Lie and Poincaré on the subject. One person who was well aware of the work of both Lie and Poincaré was Hamel.

How does Lagrange fit into this story? In *Mecanique Analytique*, Volume 2, equations A on page 212 are the Euler–Poincaré equations for the rotation group written out explicitly for a reasonably general Lagrangian. He eventually specializes them to the rigid body equations of course. We should remember that 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, a good chunk of Volume 2. His derivation is not as clean as we would give today, but it seems to have the right spirit of a reduction method. That is, he tries to get the equations from the Euler–Lagrange equations on $T\,\mathrm{SO}(3)$ by passing to the Lie algebra.

In view of the historical situation described above, one might argue that the term "Euler–Lagrange-Poincaré" equations is right for these equations. Since Poincaré noted the generalization to arbitrary Lie algebras, and applied it to interesting fluid problems, it is clear that his name belongs, but in light of other uses of the term "Euler–Lagrange," it seems that "Euler–Poincaré" is a reasonable choice.

Marsden and Scheurle [1993a,b] and Weinstein [1994] have studied a more general version of Lagrangian reduction whereby one drops the Euler–Lagrange equations from TQ to TQ/G. This is a nonabelian generalization of the classical Routh method, and leads to a very interesting coupling of the Euler–Lagrange and Euler–Poincaré equations that we shall briefly sketch in the next section. This problem was also studied by Hamel [1904] in connection with his work on nonholonomic systems (see Koiller [1992] and Bloch, Krishnaprasad, Marsden, and Murray [1994] for more information).

The current vitality of mechanics, including the investigation of fundamental questions, is quite remarkable, given its long history and development. This vitality comes about through rich interactions with both pure mathematics (from topology and geometry to group representation theory), and through new and exciting applications to areas like control theory. It is perhaps even more remarkable that absolutely fundamental points, such as a clear and unambiguous linking of Lie's work on the Lie–Poisson bracket on the dual of a Lie algebra and Poincaré's work on the Euler–Poincaré equations on the Lie algebra itself, with the most basic of examples in mechanics, such as the rigid body and the motion of ideal fluids, took nearly a century to complete. The attendant lessons to be learned about communication between pure mathematics and the other mathematical sciences are, hopefully, obvious.

Exercises

◇ **Exercise 13.6-1.** Write out the linearized rigid body equations about an equilibrium explicitly.

◇ **Exercise 13.6-2.** Let $\mathfrak{g}$ be finite dimensional. Let $e_1, \ldots, e_n$ be a basis for $\mathfrak{g}$ and $e^1, \ldots, e^n$ a dual basis for $\mathfrak{g}^*$. Let $\mu = \mu_a e^a \in \mathfrak{g}^*$ and $H(\mu) = H(\mu_1, \ldots, \mu_n) : \mathfrak{g}^* \to \mathbb{R}$. Let $[\mu_a, \mu_b] = C^d_{ab}\mu_d$. Derive a co-ordinate expression for the linearized equations (13.6.7):

$$\frac{d(\delta\mu)}{dt} = -\frac{1}{2}\,\mathrm{ad}\left(\frac{\delta(\delta^2 H)}{\delta\mu}\right)^* \mu_e - \mathrm{ad}\left(\frac{\delta H}{\delta\mu_e}\right)^* \delta\mu.$$

13.7 The Euler–Poincaré Equations

To understand this section, it will be helpful to develop some more of the basics about rigid body dynamics from the Introduction (further details are given in Chaper 15). We regard an element $\mathbf{R} \in \mathrm{SO}(3)$ giving the configuration of the body as a map of a reference configuration $\mathcal{B} \subset \mathbb{R}^3$ to the current configuration $\mathbf{R}(\mathcal{B})$; the map $\mathbf{R}$ takes a reference or label point $X \in \mathcal{B}$ to a current point $x = \mathbf{R}(X) \in \mathbf{R}(\mathcal{B})$. When the rigid body is in motion, the matrix $\mathbf{R}$ is time-dependent and the velocity of a point of the body is $\dot{x} = \dot{\mathbf{R}}X = \dot{\mathbf{R}}\mathbf{R}^{-1}x$. Since $\mathbf{R}$ is an orthogonal matrix, $\mathbf{R}^{-1}\dot{\mathbf{R}}$ and $\dot{\mathbf{R}}\mathbf{R}^{-1}$ are skew matrices, and so we can write

$$\dot{x} = \dot{\mathbf{R}}\mathbf{R}^{-1}x = \omega \times x, \tag{13.7.1}$$

which defines the ***spatial angular velocity vector*** ω. Thus, ω is essentially given by *right* translation of $\dot{\mathbf{R}}$ to the identity.

The corresponding body angular velocity is defined by

$$\Omega = \mathbf{R}^{-1}\omega, \tag{13.7.2}$$

so that Ω is the angular velocity relative to a body fixed frame. Notice that

$$\begin{aligned} \mathbf{R}^{-1}\dot{\mathbf{R}}X &= \mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}x = \mathbf{R}^{-1}(\omega \times x) \\ &= \mathbf{R}^{-1}\omega \times \mathbf{R}^{-1}x = \Omega \times X \end{aligned} \tag{13.7.3}$$

so that Ω is given by *left* translations of $\dot{\mathbf{R}}$ to the identity. The kinetic energy is obtained by summing up $m\|\dot{x}\|^2/2$ over the body:

$$K = \frac{1}{2}\int_{\mathcal{B}} \rho(X)\|\dot{\mathbf{R}}X\|^2 \, d^3X, \tag{13.7.4}$$

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

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

K is a quadratic function of Ω. Writing

$$K = \frac{1}{2}\Omega^T \mathbb{I}\Omega \tag{13.7.5}$$

defines the ***moment of inertia tensor*** $\mathbb{I}$, which, if the body does not degenerate to a line, is a positive-definite (3×3)-matrix, or better, a quadratic form. This quadratic form can be diagonalized, and this defines the principal axes and moments of inertia. In this basis, we write $\mathbb{I} = \operatorname{diag}(I_1, I_2, I_3)$. The function K is taken to be the Lagrangian of the system on $T\,\mathrm{SO}(3)$ (and by means of the Legendre transformation we get the corresponding Hamiltonian description on $T^*\,\mathrm{SO}(3)$). Notice directly from (13.7.4) that K is *left* (not right) invariant on $T\,\mathrm{SO}(3)$. It follows that the corresponding Hamiltonian is also *left* invariant.

From the Lagrangian point of view, the relation between the motion in $\mathbf{R}$ space and that in body angular velocity (or Ω) space is as follows:

Theorem 13.7.1. *The curve $\mathbf{R}(t) \in \mathrm{SO}(3)$ satisfies the Euler–Lagrange equations for*

$$L(\mathbf{R}, \dot{\mathbf{R}}) = \frac{1}{2}\int_{\mathcal{B}} \rho(X)\|\dot{\mathbf{R}}X\|^2 \, d^3X, \tag{13.7.6}$$

if and only if $\Omega(t)$ defined by $\mathbf{R}^{-1}\dot{\mathbf{R}}v = \Omega \times v$ for all $v \in \mathbb{R}^3$ satisfies Euler's equations

$$\mathbb{I}\dot{\Omega} = \mathbb{I}\Omega \times \Omega. \tag{13.7.7}$$

One instructive way to prove this *indirectly* is to pass to the Hamiltonian formulation and use Lie–Poisson reduction, as outlined above. One way to do it *directly* is to use variational principles. By Hamilton's principle, $R(t)$ satisfies the Euler–Lagrange equations if and only if

$$\delta \int L \, dt = 0.$$

Let $l(\Omega) = \frac{1}{2}(\mathbb{I}\Omega)\cdot\Omega$, so that $l(\Omega) = L(\mathbf{R}, \dot{\mathbf{R}})$ if $\mathbf{R}$ and Ω are related as above. To see how we should transform Hamilton's principle, we differentiate the relation $\mathbf{R}^{-1}\dot{\mathbf{R}} = \hat{\Omega}$ with respect to $\mathbf{R}$ to get

$$-\mathbf{R}^{-1}(\delta\mathbf{R})\mathbf{R}^{-1}\dot{\mathbf{R}} + \mathbf{R}^{-1}(\delta\dot{\mathbf{R}}) = \widehat{\delta\Omega}. \tag{13.7.8}$$

Let the skew matrix $\hat{\Sigma}$ be defined by

$$\hat{\Sigma} = \mathbf{R}^{-1}\delta\mathbf{R} \tag{13.7.9}$$

and define the vector Σ by

$$\hat{\Sigma}v = \Sigma \times v. \tag{13.7.10}$$

Note that

$$\dot{\hat{\Sigma}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}\delta\mathbf{R} + \mathbf{R}^{-1}\delta\dot{\mathbf{R}},$$

so

$$\mathbf{R}^{-1}\delta\dot{\mathbf{R}} = \dot{\hat{\Sigma}} + \mathbf{R}^{-1}\dot{\mathbf{R}}\hat{\Sigma} \tag{13.7.11}$$

substituting (13.7.11) and (13.7.9) into (13.7.8) gives

$$-\hat{\Sigma}\hat{\Omega} + \dot{\hat{\Sigma}} + \hat{\Omega}\hat{\Sigma} = \widehat{\delta\Omega},$$

that is,

$$\widehat{\delta\Omega} = \dot{\hat{\Sigma}} + [\hat{\Omega}, \hat{\Sigma}]. \tag{13.7.12}$$

The identity $[\hat{\Omega}, \hat{\Sigma}] = (\Omega \times \Sigma)\hat{}$ holds by Jacobi's identity for the cross product, and so

$$\delta\Omega = \dot{\Sigma} + \Omega \times \Sigma. \tag{13.7.13}$$

These calculations prove the following:

Theorem 13.7.2. *Hamilton's variational principle*

$$\delta \int_a^b L\, dt = 0 \tag{13.7.14}$$

*on $T\,\mathrm{SO}(3)$ is equivalent to the **reduced variational principle***

$$\delta \int_a^b l\, dt = 0 \tag{13.7.15}$$

on $\mathbb{R}^3$ where the variations $\delta\Omega$ are of the form (13.7.13) with $\Sigma(a) = \Sigma(b) = 0$.

. To complete the proof of Theorem 13.7.1, it suffices to work out the equations equivalent to the reduced variational principle (13.7.15). Since

$l(\Omega) = \frac{1}{2}\langle \mathbb{I}\Omega, \Omega\rangle$, and $\mathbb{I}$ is symmetric, we get

$$\begin{aligned}\delta \int_a^b l\, dt &= \int_a^b \langle \mathbb{I}\Omega, \delta\Omega\rangle\, dt \\ &= \int_a^b \langle \mathbb{I}\Omega, \dot{\Sigma} + \Omega \times \Sigma\rangle\, dt \\ &= \int_a^b \left[\left\langle -\frac{d}{dt}\mathbb{I}\Omega, \Sigma \right\rangle + \langle \mathbb{I}\Omega, \Omega \times \Sigma\rangle\right] \\ &= \int_a^b \left\langle -\frac{d}{dt}\mathbb{I}\Omega + \mathbb{I}\Omega \times \Omega, \Sigma \right\rangle dt,\end{aligned}$$

where we have integrated by parts and used the boundary conditions $\Sigma(b) = \Sigma(a) = 0$. Since Σ is otherwise arbitrary, (13.7.15) is equivalent to

$$-\frac{d}{dt}(\mathbb{I}\Omega) + \mathbb{I}\Omega \times \Omega = 0,$$

which are Euler's equations. ■

We now generalize this procedure to an arbitrary Lie group and later will make the direct link with the Lie–Poisson equations.

Theorem 13.7.3. *Let G be a Lie group and let $L : TG \to \mathbb{R}$ be a left invariant Lagrangian. Let $l : \mathfrak{g} \to \mathbb{R}$ be its restriction to the identity. For a curve $g(t) \in G$, let $\xi(t) = g(t)^{-1} \cdot \dot{g}(t)$; that is, $\xi(t) = T_{g(t)} L_{g(t)^{-1}} \dot{g}(t)$. Then the following are equivalent:*

(i) *$g(t)$ satisfies the Euler–Lagrange equations for L on G;*

(ii) *the variational principle*

$$\delta \int L(g(t), \dot{g}(t))\, dt = 0 \tag{13.7.16}$$

holds, for variations with fixed endpoints;

(iii) *the* ***Euler–Poincaré equations*** *hold:*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \operatorname{ad}^*_\xi \frac{\delta l}{\delta \xi}; \tag{13.7.17}$$

(iv) *the variational principle*

$$\delta \int l(\xi(t))\, dt = 0 \tag{13.7.18}$$

holds on $\mathfrak{g}$, using variations of the form

$$\delta \xi = \dot{\eta} + [\xi, \eta], \tag{13.7.19}$$

where η vanishes at the endpoints.

Proof. First of all, the equivalence of (i) and (ii) holds on the tangent bundle of any configuration manifold Q, as we know from Chapter 8. To see that (ii) and (iv) are equivalent, one needs to compute the variations $\delta\xi$ induced on $\xi = g^{-1}\dot{g} = TL_{g^{-1}}\dot{g}$ by a variation of g. We will do this for matrix groups; see Bloch, Krishnaprasad, Marsden, and Ratiu [1994b] for the general case. To calculate this, we need to differentiate $g^{-1}\dot{g}$ in the direction of a variation δg. If $\delta g = dg/d\epsilon$ at $\epsilon = 0$, where g is extended to a curve g_ϵ, then,

$$\delta\xi = \frac{d}{d\epsilon} g^{-1} \frac{d}{dt} g,$$

while if $\eta = g^{-1}\delta g$, then

$$\dot{\eta} = \frac{d}{dt} g^{-1} \frac{d}{d\epsilon} g.$$

The difference $\delta\xi - \dot{\eta}$ is thus the commutator $[\xi, \eta]$.

To complete the proof, we show the equivalence of (iii) and (iv). Indeed, using the definitions and integrating by parts,

$$\begin{aligned}
\delta \int l(\xi) dt &= \int \frac{\delta l}{\delta \xi} \delta\xi \, dt \\
&= \int \frac{\delta l}{\delta \xi} (\dot{\eta} + \operatorname{ad}_\xi \eta) \, dt \\
&= \int \left[-\frac{d}{dt} \left(\frac{\delta l}{\delta \xi} \right) + \operatorname{ad}_\xi^* \frac{\delta l}{\delta \xi} \right] \eta \, dt
\end{aligned}$$

so the result follows. ■

There is of course a right invariant version of this theorem in which $\xi = \dot{g}g^{-1}$ and when (13.7.17), (13.7.19) aquire minus signs.

In coordinates, (13.7.17), reads as follows

$$\frac{d}{dt} \frac{\partial l}{\partial \xi^a} = C^b_{da} \xi^d \frac{\partial l}{\partial x^b}. \tag{13.7.20}$$

Since the Euler–Lagrange and Hamilton equations on TQ and T^*Q are equivalent, it follows that the Lie–Poisson and Euler–Poincaré equations are also equivalent. To see this *directly*, we make the following Legendre transformation from $\mathfrak{g}$ to $\mathfrak{g}^*$:

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

Note that

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

and so it is now clear that the Lie–Poisson and Euler–Poincaré equations are equivalent.

We close this section by showing that the periodic KdV equation, (see Example (c) in §3.2)

$$u_t + 6uu_x + u_{xxx} = 0$$

is an Euler–Poincaré equation on a certain Lie algebra called the ***Virasoro algebra*** $\mathfrak{v}$. These results were obtained in the Lie–Poisson context by Gelfand and Dorfman [1979], Kirillov [1981], Ovsienko and Khesin [1987], and Segal [1991]. See also Pressley and Segal [1986] and references therein.

We begin with the construction of the Virasoro algebra $\mathfrak{v}$. If one identifies elements of $\mathfrak{X}(S^1)$ with periodic functions of period 1 endowed with the Jacobi–Lie bracket

$$[u, v] = uv' - u'v,$$

the ***Gelfand-Fuchs cocycle*** is defined by the expression

$$\Sigma(u, v) = \gamma \int_0^1 u'(x)v''(x)dx,$$

where $\gamma \in \mathbb{R}$ is an arbitrary constant. The Lie algebra $\mathfrak{X}(S^1)$ of vector fields on the circle has a unique central extension by $\mathbb{R}$ determined by the Gelfand-Fuchs cocycle. Therefore, (see (12.3.22) in remark 5 of §12.4), the Lie algebra bracket on

$$\mathfrak{v} := \{(u, a) \mid u \in \mathfrak{X}(S^1), \quad a \in \mathbb{R}\}$$

is given by

$$[(u, a), (v, b)] = \left(-uv' + u'v, \gamma \int_0^1 u'(x)v''(x)\, dx \right)$$

since the *left* Lie bracket on $\mathfrak{X}(S^1)$ is given by the negative of the Jacobi–Lie bracket for vector fields. Identify the dual of $\mathfrak{v}$ with $\mathfrak{v}$ by the L^2-inner product

$$\langle (u, a), (v, b) \rangle = ab + \int_0^1 u(x)v(x)\, dx.$$

We claim that the coadjoint action $\mathrm{ad}^*_{(u,a)}$ is given by

$$\mathrm{ad}^*_{(u,a)}(v, b) = (b\gamma u''' + 2u'v + uv', 0).$$

Indeed, if $(u,a),(v,b),(w,c) \in \mathfrak{v}$, we have

$$\begin{aligned}
&\left\langle \operatorname{ad}^*_{(u,a)}(v,b),(w,c)\right\rangle \\
&\quad = \langle (v,b), [(u,a),(w,c)]\rangle \\
&\quad = \left\langle (v,b), \left(-uw' + u'w, \gamma \int_0^1 u'(x)w''(x)\,dx\right)\right\rangle \\
&\quad = b\gamma \int_0^1 u'(x)w''(x)\,dx - \int_0^1 v(x)u(x)w'(x)\,dx + \int_0^1 v(x)u'(x)w(x)\,dx.
\end{aligned}$$

Integrating the first term twice and the second term once by parts and remembering that the boundary terms vanish by periodicity, this expresison becomes

$$\begin{aligned}
&b\gamma \int_0^1 u'''(x)w(x)dx + \int_0^1 (v(x)u(x))'w(x)dx + \int_0^1 v(x)u'(x)w(x)dx \\
&\qquad = \int_0^1 (b\gamma u'''(x) + 2u'(x)v(x) + u(x)v'(x))w(x)dx \\
&\qquad = \langle (b\gamma u''' + 2u'v + uv', 0), (w,c)\rangle \,.
\end{aligned}$$

If $F : \mathfrak{v} \to \mathbb{R}$, its functional derivative relative to the L^2-pairing is given by

$$\frac{\delta F}{\delta(u,a)} = \left(\frac{\delta F}{\delta u}, \frac{\partial F}{\partial a}\right)$$

where $\delta F/\delta u$ is the usual L^2-functional derivative of F keeping $a \in \mathbb{R}$ fixed and $\partial F/\partial a$ is the standard partial derivative of F keeping u fixed. The Euler–Poincaré equations for right invariant systems given by $l : \mathfrak{v} \to \mathbb{R}$ becomes

$$\frac{d}{dt}\frac{\delta l}{\delta(u,a)} = -\operatorname{ad}^*_{(u,a)} \frac{\delta l}{\delta(u,a)}.$$

However,

$$\begin{aligned}
\operatorname{ad}^*_{(u,a)} \frac{\delta l}{\delta(u,a)} &= \operatorname{ad}^*_{(u,a)} \left(\frac{\delta l}{\delta u}, \frac{\partial l}{\partial a}\right) \\
&= \left(\gamma \frac{\partial l}{\partial a} u''' + 2u' \frac{\delta l}{\delta u} + u \left(\frac{\delta l}{\delta u}\right)', 0\right),
\end{aligned}$$

so that we get the system

$$\begin{aligned}
&\frac{d}{dt}\frac{\partial l}{\partial a} = 0 \\
&\frac{d}{dt}\frac{\delta l}{\delta u} = -\gamma \frac{\partial l}{\partial a} u''' - 2u' \frac{\delta l}{\delta u} - u \left(\frac{\delta l}{\delta u}\right)'.
\end{aligned}$$

If

$$l(u,a) = \frac{1}{2}\left(a^2 + \int_0^1 u^2(x)\,dx\right),$$

then $\partial l/\partial a = a, \delta l/\delta u = u$ and the above equations become

$$\begin{aligned} \frac{da}{dt} &= 0 \\ \frac{du}{dt} &= -\gamma a u''' - 3u'u. \end{aligned} \tag{13.7.21}$$

Since a is constant, we get

$$u_t + 3u_x u + \gamma a u''' = 0. \tag{13.7.22}$$

This equation is equivalent to the KdV equation upon rescaling time and choosing the constant a appropriately. Indeed, let $u(t,x) = v(\tau(t),x)$ for $\tau(t) = t/2$. Then $u_x = v_x$ and $u_t = -v_\tau/2$ so that (13.7.22) becomes

$$v_\tau + 6vv_x + 2\gamma a v_{xxx} = 0,$$

which becomes the KdV equation (see §3.2) if we choose $a = 1/2\gamma$.

The Lie–Poisson formulation goes the following way. The (+) Lie–Poisson bracket is given by

$$\begin{aligned} \{f,h\}(u,a) &= \left\langle (u,a), \left[\frac{\delta f}{\delta(u,a)}, \frac{\delta h}{\delta(u,a)}\right]\right\rangle \\ &= \int \left[u\left(\left(\frac{\delta f}{\delta u}\right)' \frac{\delta h}{\delta u} - \frac{\delta f}{\delta u}\left(\frac{\delta h}{\delta u}\right)'\right)\right. \\ &\qquad \left. + a\gamma \left(\frac{\delta f}{\delta u}\right)' \left(\frac{\delta h}{\delta u}\right)''\right] dx \end{aligned}$$

so that the Lie–Poisson equations $\dot f = \{f,h\}$ become

$$\begin{aligned} \frac{da}{dt} &= 0 \\ \frac{du}{dt} &= -u'\left(\frac{\delta h}{\delta u}\right) - 2u\left(\frac{\delta h}{\delta u}\right)' - a\gamma\left(\frac{\delta h}{\delta u}\right)'''. \end{aligned} \tag{13.7.23}$$

Taking

$$h(u,a) = \frac{1}{2}a^2 + \frac{1}{2}\int_0^1 u^2(x)\,dx,$$

we get $\partial h/\partial a = a, \delta h/\delta u = u$ and so (13.7.23) becomes (13.7.22) as was to be expected and could have directly obtained by a Legendre transform.

The conclusion is that the KdV equation is the expression in space coordinates of the geodesic equations on the Virasoro group V endowed with the right invariant metric whose value at the identity is the L^2-inner product. We shall not describe here the Virasoro group which is a central extension of the diffeomorphism group on S^1; we refer the reader to Pressley and Segal [1986].

Exercises

◇ **Exercise 13.7-1.** Verify the coordinate form of the Euler–Poincaré equations.

◇ **Exercise 13.7-2.** Show that the Euler equations for a perfect fluid are Euler–Poincaré equations. Find the variational principle (3) in Newcomb [1962] and Bretherton [1970].

13.8 The Reduced Euler–Lagrange Equations

As we have mentioned, the Lie–Poisson and Euler–Poincaré equations occur for many systems besides the rigid body equations. They include the equations of fluid and plasma dynamics, for example. For many other systems, such as a rotating molecule or a spacecraft with movable internal parts, one has a combination of equations of Euler–Poincaré type and Euler–Lagrange type. Indeed, on the Hamiltonian side, this process has undergone development for quite some time, and is discussed at length in Volume II. On the Lagrangian side, this process is also very interesting, and has been recently developed by, amongst others, Marsden and Scheurle [1993a,b]. The general problem is to drop Euler–Lagrange equations and variational principles from a general velocity phase-space TQ to the quotient TQ/G by a Lie group action of G on Q. If L is a G-invariant Lagrangian on TQ, it induces a reduced Lagrangian l on TQ/G. We give a brief *preview* of the general theory in this section. In fact, the material below can also act as motivation for the general theory of connections, also introduced in Volume II.

An important ingredient in this work is to introduce a connection A on the principal bundle $Q \to S = Q/G$, assuming that this quotient is nonsingular. For example, the mechanical connection (see Kummer [1981], Marsden [1992] and references therein), may be chosen for A. This connection allows one to split the variables into a horizontal and vertical part. Let x^α, also called "internal variables," be coordinates for shape-space Q/G, let η^a be coordinates for the Lie algebra $\mathfrak{g}$ relative to a chosen basis, let l be the Lagrangian regarded as a function of the variables $x^\alpha, \dot{x}^\alpha, \eta^a$, and let C^a_{db} be the structure constants of the Lie algebra $\mathfrak{g}$ of G.

If one writes the Euler–Lagrange equations on TQ in a local principal bundle trivialization, with coordinates x^α on the base and η^a in the fiber, then one gets the following system of ***Hamel equations***:

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{x}^\alpha} - \frac{\partial l}{\partial x^\alpha} = 0, \tag{13.8.1}$$

$$\frac{d}{dt}\frac{\partial l}{\partial \eta^b} - \frac{\partial l}{\partial \eta^a} C^a_{db}\eta^d = 0. \tag{13.8.2}$$

However, this representation of the equations does not make global intrinsic sense (unless $Q \to S$ admits a global flat connection). The introduction of a connection overcomes this and one can intrinsically and globally split the original variational principle relative to horizontal and vertical variations. One gets from one-form to the other by means of the velocity shift given by replacing η by the vertical part relative to the connection

$$\xi^a = A^a_\alpha \dot{x}^\alpha + \eta^a.$$

Here, A^d_α are the local coordinates of the connection A. This change of coordinates is motivated from the mechanical point of view since the variables ξ have the interpretation of the locked angular velocity. The resulting ***reduced Euler–Lagrange equations*** have the following form:

$$\frac{d}{dt}\frac{\partial l}{\partial \dot{x}^\alpha} - \frac{\partial l}{\partial x^\alpha} = \frac{\partial l}{\partial \xi^a}\left(B^a_{\alpha\beta}\dot{x}^\beta + B^a_{\alpha d}\xi^d\right), \tag{13.8.3}$$

$$\frac{d}{dt}\frac{\partial l}{\partial \xi^b} = \frac{\partial l}{\partial \xi^a}(B^a_{b\alpha}\dot{x}^\alpha + C^a_{db}\xi^d). \tag{13.8.4}$$

In these equations, $B^a_{\alpha\beta}$ are the coordinates of the curvature B of A, $B^a_{d\alpha} = C^a_{bd}A^b_\alpha$ and $B^a_{b\alpha} = -B^a_{\alpha b}$.

It is interesting to note that the matrix

$$\begin{bmatrix} B^a_{\alpha\beta} & B^a_{\alpha d} \\ B^a_{d\alpha} & C^a_{bd} \end{bmatrix}$$

is itself the curvature of the connection regarded as residing on the bundle $TQ \to TQ/G$. Regarding the structure constants as a curvature tensor in the special case $Q = G$ may be regarded as a reformulation of the Mauer-Cartan equations (see Theorem 9.1.11).

The variables ξ^a may be regarded as the rigid part of the variables on the original configuration space, while x^α are the internal variables. As in Simo, Lewis, and Marsden [1991], the division of variables into internal and rigid parts has deep implications for both stability theory and for bifurcation theory, again, continuing along lines developed originally by Riemann, Poincaré, and others. The main way this new insight is achieved is through a careful split of the variables, using the (mechanical) connection as one of the main ingredients. This split puts the second variation of the augmented Hamiltonian at a relative equilibrium as well as the symplectic form into "normal form." It is somewhat remarkable that they are *simultaneously* put into a simple form. This link helps considerably with an eigenvalue analysis of the linearized equations, and in Hamiltonian bifurcation theory; see, for example, Bloch, Krishnaprasad, Marsden, and Ratiu [1994a].

One of the key results in Hamiltonian reduction theory says that the reduction of a cotangent bundle T^*Q by a symmetry group G is a bundle over T^*S, where $S = Q/G$ is shape-space, and where the fiber is either $\mathfrak{g}^*$, the dual of the Lie algebra of G, or is a coadjoint orbit, depending on whether one is doing Poisson or symplectic reduction. We refer to Montgomery, Marsden, and Ratiu [1984] and Marsden [1992] and Volume II for details and references. The reduced Euler–Lagrange equations give the analogue of this structure on the tangent bundle.

Remarkably, equations (13.8.3) are very close in form to the equations for a mechanical system with classical nonholonomic velocity constraints (see Naimark and Fufaev [1972] and Koiller [1992].) The connection chosen in that case is the one-form that determines the constraints. This link is made precise in Bloch, Krishnaprasad, Marsden, and Murray [1994]. In addition, this structure appears in several control problems, especially the problem of stabilizing controls considered by Bloch, Krishnaprasad, Marsden, and Sanchez [1992].

For systems with a momentum map $\mathbf{J}$ constrained to a specific value μ, the key to the construction of a reduced Lagrangian system is the modification of the Lagrangian L to the Routhian R^μ, which is obtained from the Lagrangian by subtracting off the mechanical connection paired with the constraining value μ of the momentum map. On the other hand, a basic ingredient needed for the reduced Euler–Lagrange equations is a velocity shift in the Lagrangian, the shift being determined by the connection, so this velocity-shifted Lagrangian plays the role that the Routhian does in the constrained theory.