Symplectic Reduction for Semidirect Products and Central Extensions.

Jerrold E. Marsden
Control and Dynamical Systems
California Institute of Technology 107-81
Pasadena, CA 91125
marsden@cds.caltech.edu

Gerard Misiołek
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556
gmisiole@hilbert.helios.nd.edu

Matthew Perlmutter
Department of Mathematics
University of California
Berkeley, CA 94720
perl@math.berkeley.edu

Tudor S. Ratiu
Department of Mathematics
University of California
Santa Cruz, CA 95064
ratiu@math.ucsc.edu

March, 1994; this version May 24, 1998
to appear in *Differential Geometry and its Applications*
MS2000 Classification 37J15, 53D20

Abstract

This paper proves a symplectic reduction by stages theorem in the context of geometric mechanics on symplectic manifolds with symmetry groups that are group extensions. We relate the work to the semidirect product reduction theory developed in the 1980's by Marsden, Ratiu, Weinstein, Guillemin and Sternberg as well as some more recent results and we recall how semidirect product reduction finds use in examples, such as the dynamics of an underwater vehicle.

We shall start with the classical cases of commuting reduction (first appearing in Marsden and Weinstein [1974]) and present a new proof

and approach to semidirect product theory. We shall then give an idea of how the more general theory of group extensions proceeds (the details of which are given in Marsden, Misiołek, Perlmutter and Ratiu [1998]). The case of central extensions is illustrated in this paper with the example of the Heisenberg group. The theory, however, applies to many other interesting examples such as the Bott-Virasoro group and the KdV equation.

Contents

1 Introduction and Background

The Problem Setting. Consider a Lie group M that acts on a symplectic manifold (P, Ω) by symplectic transformations and that this action has an equivariant momentum map $\mathbf{J}_M : P \to \mathfrak{m}^*$, where $\mathfrak{m}$ is the Lie algebra of M. Let N be a normal subgroup of M. The problem is to carry out the symplectic reduction of P by M in two steps, first a reduction of P by N and followed by, roughly speaking, a reduction by the quotient group M/N.

This seemingly straightforward problem is remarkably subtle and when properly understood, has a surprising number of consequences, applications and links with other subjects. These include subjects as diverse as classifying coadjoint orbits, applications to underwater vehicle dynamics, and links to induced representations and quantization. We shall survey some of the literature in this area below.

It is a great pleasure and honor to dedicate this paper to one of the giants in this field, Victor Guillemin. We hope it will do justice to his love of symplectic geometry and symmetry and in particular, his work with Shlomo Sternberg on semidirect products.

Poisson Reduction. Perhaps the simplest context in which one can understand a result of this sort is that of "easy Poisson reduction". That is, assuming that the group actions are free and proper, one simply forms the quotient manifold P/M with its natural quotient Poisson structure (see Marsden and Ratiu [1994] for an exposition of this standard theory). An easy result that the reader can readily prove is that

P/M is Poisson diffeomorphic to $(P/N)/(M/N)$.

A more ambitious task, which we undertake in this paper, is to keep track of the symplectic leaves in this process.

We *do not attempt* to carry out a reduction by stages in the more sophisticated context of Poisson reduction of Marsden and Ratiu [1986] (see also Vanhaecke [1996]). This would be an interesting problem, but it is not addressed here.

Lagrangian Reduction by Stages. The companion paper of Cendra, Marsden and Ratiu [1998] carries out the analogue of this program for the reduction of Lagrangian systems by stages (following the development of Lagrangian reduction of Cendra and Marsden [1987], Cendra, Ibort and Marsden [1987], and Marsden and Scheurle [1993a,b]). In Lagrangian reduction, it is variational principles that are reduced, as opposed to symplectic or Poisson structures. In the context of *Lagrangian reduction of semidirect products*, this is closely connected with the beautiful variational theory of the Euler–Poincaré equations; see Holm, Marsden and Ratiu [1998a,b].

Symplectic Reduction. Recall that for a group G (whose Lie algebra is denoted $\mathfrak{g}$) acting on a symplectic manifold (P, Ω) and with an Ad-equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$, the *symplectic reduced space* is defined by $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$, where $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}$, G_μ is the isotropy subgroup for the coadjoint action of G on $\mathfrak{g}^*$, and we assume that G_μ acts freely and properly on the level set $\mathbf{J}^{-1}(\mu)$. The reduced manifold is a symplectic manifold in a natural way: the pull back of the given symplectic structure to the level set of $\mathbf{J}$ equals the pull back of the reduced symplectic structure by the natural projection map $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ to the quotient space. This is the well known Marsden-Weinstein-Meyer symplectic reduction theorem (see, e.g., Abraham and Marsden [1978] or Guillemin and Sternberg [1984] for expositions).

The regularity and free and proper assumptions can be somewhat weakened (for example, to weakly regular values) as is well known. On the other hand, *we do not treat the singular case in this paper* — it would of course be of interest to do so, following Arms, Marsden and Moncrief [1981], Sjamaar and Lerman [1991], Ortega and Ratiu [1998], and various subsequent papers.

We also need to keep in mind that one can discuss the *reduction of dynamics*; thus, if H is a Hamiltonian on P that is G invariant, it induces a Hamiltonian H_μ on each of the reduced spaces, and the corresponding Hamiltonian vector fields X_H and X_{H_μ} are π_μ-related. The reverse of reduction is reconstruction and this leads one to the theory of classical geometric phases (Hannay-Berry phases); see Marsden, Montgomery and Ratiu [1990]. There are of course many important results related to reduction and to the structure of reduced spaces some of which we will encounter later; we refer to Abraham and Marsden [1978] and Marsden [1992] for some of the general results and for additional references.

The Organization of the Paper. The strategy of the paper is to proceed step by step as follows:

1. The case of the direct product of two groups (commuting reduction).

2. The case of the semidirect product of a group with a vector space (the Euclidean group, groups relevant for compressible flow, etc.).

3. The case of group extensions (when a given group has a normal subgroup).

We do this in a way where the result at one step leads naturally to the approach at the next step. We also mention a number of interesting examples along the way. Many of the deeper and more extensive parts of the work will be left to another paper, Marsden, Misiołek, Perlmutter, and Ratiu [1998].

2 Commuting Reduction

Theorems on reduction by stages have been given in various special instances by a number of authors, starting with time-honored observations in mechanics such as the following: When you want to reduce the dynamics of a rigid body moving in space, *first* you can pass to center of mass coordinates (that is, reduce by translations) and *second* you can pass to body coordinates (that is, reduce by the rotation group). For other problems, such as a rigid body in a fluid (see Leonard and Marsden [1997]) this process is not so simple; one does *not* simply pass to center of mass coordinates to get rid of translations. This shows that the general problem of reducing by the Euclidean group is a bit more subtle than one may think at first. In any case, such procedures correspond to a reduction by stages result for *semidirect products*. But we are getting ahead of ourselves; we need to step back and look first at an even simpler case, namely the case of *direct products*.

The early version of Marsden and Weinstein [1974, p. 127] states that for two commuting group actions, one could reduce by them in succession and in either order and the result is the same as reducing by the *direct* product group. One version of this result is the following theorem.

Theorem 2.1 (Commuting Reduction Theorem). *Let P be a symplectic manifold, K be a Lie group (with Lie algebra $\mathfrak{k}$) acting symplectically on P and having an equivariant momentum map $\mathbf{J}_K : P \to \mathfrak{k}^*$. Assume that $\nu \in \mathfrak{k}^*$ is a regular value of $\mathbf{J}_K$ and that the action of K_ν is free and proper, so that the symplectic reduced space $P_\nu = \mathbf{J}_K^{-1}(\nu)/K_\nu$ is a smooth manifold. Let G be another group (with Lie algebra $\mathfrak{g}$) acting on P with an equivariant momentum map $\mathbf{J}_G : P \to \mathfrak{g}^*$. Assume μ is a regular value for the G-action. Suppose that the actions of G and K on P commute. Then $\mathbf{J}_G \times \mathbf{J}_K$ is a momentum map for the action of $G \times K$ on P and*

- **i** *If $\mathbf{J}_K$ is G-invariant and K is connected, then $\mathbf{J}_G$ is K-invariant and $\mathbf{J}_G \times \mathbf{J}_K$ is equivariant. Moreover, G induces a symplectic action on P_ν, and the map $\mathbf{J}_\nu : P_\nu \to \mathfrak{g}^*$ induced by $\mathbf{J}_G$ is an equivariant momentum map for this action.*

- **ii** *The (symplectic) reduced space for the action of G on P_ν at μ is symplectically diffeomorphic to the reduction of P at the point (μ, ν) by the action of $G \times K$.*

For example, in the dynamics of a rigid body with a fixed point and two equal moments of inertia moving in a gravitational field, there are two commuting S^1 symmetry groups acting on the phase space $T^*\mathrm{SO}(3)$. These actions commute since one is given by (the cotangent lift of) left translation and the other by right translation. The corresponding integrals of motion lead to the complete integrability of the problem. One can reduce by the action of these groups either together or one following the other with the same final reduced space.

This result may be viewed in the general context discussed in the introduction by taking $M = G \times K$ with the normal subgroup being chosen to be either G or K, so that the quotient group of M is the other group.

Proof of the Commuting Reduction Theorem. It is instructive to build up to the general reduction by stages theorem by giving direct proofs of some simpler special cases; these special cases not only point the way to the general case, but contain interesting constructions that are relevant to the cases treated. The general case has some subtleties not shared by these simple cases, which will be spelled out as we proceed.

i. First of all, note that the G-invariance of $\mathbf{J}_K$ implies that $\mathbf{d}\langle \mathbf{J}_K, \eta\rangle \cdot \xi_P = 0$ for all $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{k}$. However,

$$\begin{aligned} \mathbf{d}\langle \mathbf{J}_G, \xi\rangle \cdot \eta_P &= \mathbf{d}\langle \mathbf{J}_G, \xi\rangle \cdot X_{\langle \mathrm{J}_K, \eta\rangle} = \{\langle \mathbf{J}_G, \xi\rangle, \langle \mathbf{J}_K, \eta\rangle\} \\ &= -\mathbf{d}\langle \mathbf{J}_K, \eta\rangle \cdot X_{\langle \mathrm{J}_G, \xi\rangle} = -\mathbf{d}\langle \mathbf{J}_K, \eta\rangle \cdot \xi_P = 0, \end{aligned}$$

from which we conclude K-invariance of $\mathbf{J}_G$.

Also, for all $z \in P$ and $(g, k) \in G \times K$ we have

$$\begin{aligned} (\mathbf{J}_G \times \mathbf{J}_K)((g,k)\cdot z) &= (\mathbf{J}_G(g\cdot k\cdot z),\, \mathbf{J}_K(g\cdot k\cdot z)) \\ &= (g\cdot \mathbf{J}_G(z),\, k\cdot \mathbf{J}_K(z)) \\ &= (g,k)\cdot (\mathbf{J}_G \times \mathbf{J}_K)(z), \end{aligned} \tag{2.1}$$

where we have used the invariance of $\mathbf{J}_G$ and $\mathbf{J}_K$.

Let the action of $g \in G$ on P be denoted by $\Psi_g : P \to P$. Since these maps commute with the action of K and leave the momentum map $\mathbf{J}_K$ invariant by hypothesis, there are well defined induced maps $\Psi_g^\nu : \mathbf{J}_K^{-1}(\nu) \to \mathbf{J}_K^{-1}(\nu)$ and $\Psi_{g,\nu} : P_\nu \to P_\nu$, which then define actions of G on $\mathbf{J}_K^{-1}(\nu)$ and on P_ν.

Let $\pi_\nu : \mathbf{J}_K^{-1}(\nu) \to P_\nu$ denote the natural projection and $i_\nu : \mathbf{J}_K^{-1}(\nu) \to P$ be the inclusion. We have by construction, $\Psi_{g,\nu}\circ\pi_\nu = \pi_\nu\circ\Psi_g^\nu$ and $\Psi_g\circ i_\nu = i_\nu\circ\Psi_g^\nu$. The symplectic form on the reduced space is characterized by $i_\nu^*\Omega = \pi_\nu^*\Omega_\nu$. Therefore,

$$\pi_\nu^*\Psi_{g,\nu}^*\Omega_\nu = (\Psi_g^\nu)^*\pi_\nu^*\Omega_\nu = (\Psi_g^\nu)^* i_\nu^*\Omega = i_\nu^*\Psi_g^*\Omega = i_\nu^*\Omega = \pi_\nu^*\Omega_\nu.$$

Since π_ν is a surjective submersion, we may conclude that

$$\Psi_{g,\nu}^*\Omega_\nu = \Omega_\nu.$$

Thus, we have a symplectic action of G on P_ν.

Since $\mathbf{J}_G$ is invariant under K and hence under K_ν, there is an induced map $\mathbf{J}_\nu : P_\nu \to \mathfrak{g}^*$ satisfying $\mathbf{J}_\nu \circ \pi_\nu = \mathbf{J}_G \circ i_\nu$. We now check that this is the momentum map for the action of G on P_ν. To do this, first note that for all $\xi \in \mathfrak{g}$, the vector fields ξ_P and ξ_{P_ν} are π_ν-related. Denoting the interior product of a vector field X and a form α by $\mathbf{i}_X\alpha$, we have

$$\pi_\nu^*\left(\mathbf{i}_{\xi_{P_\nu}}\Omega_\nu\right) = \mathbf{i}_{\xi_P} i_\nu^*\Omega = i_\nu^*\left(\mathbf{i}_{\xi_P}\Omega\right) = i_\nu^*\left(\mathbf{d}\langle \mathbf{J}_G, \xi\rangle\right) = \pi_\nu^*\left(\mathbf{d}\langle \mathbf{J}_\nu, \xi\rangle\right).$$

Again, since π_ν is a surjective submersion, we may conclude that

$$\mathbf{i}_{\xi_{P_\nu}} \Omega_\nu = \mathbf{d} \langle \mathbf{J}_\nu, \xi \rangle$$

and hence $\mathbf{J}_\nu$ is the momentum map for the G action on P_ν. Equivariance of $\mathbf{J}_\nu$ follows from that for $\mathbf{J}_G$, by a diagram chasing argument as above, using the relation $\mathbf{J}_\nu \circ \pi_\nu = \mathbf{J}_G \circ i_\nu$ and the relations between the actions of G on P, $\mathbf{J}_K^{-1}(\nu)$, and on P_ν.

ii. The equivariant momentum map for the action of the product group $G \times K$ is verified to be $\mathbf{J}_G \times \mathbf{J}_K : P \to \mathfrak{g}^* \times \mathfrak{k}^*$. We begin with the inclusion map

$$j : (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu, \nu) \to \mathbf{J}_K^{-1}(\nu).$$

Composing this map with π_ν gives the map

$$\pi_\nu \circ j : (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu, \nu) \to P_\nu.$$

This map takes values in $\mathbf{J}_\nu^{-1}(\mu)$ because of the relation $\mathbf{J}_\nu \circ \pi_\nu = \mathbf{J}_G \circ i_\nu$. Using the same name, we get a map:

$$\pi_\nu \circ j : (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu, \nu) \to \mathbf{J}_\nu^{-1}(\mu).$$

This map is equivariant with respect to the action of $G_\mu \times K_\nu$ on the domain and G_μ on the range. Thus, it induces a map

$$[\pi_\nu \circ j] : P_{(\mu,\nu)} \to (P_\nu)_\mu.$$

An argument like that in **i** shows that this map is symplectic.

We will show that this map is a diffeomorphism by constructing an inverse. We begin with the map

$$\phi : \mathbf{J}_\nu^{-1}(\mu) \to P_{(\mu,\nu)}$$

defined as follows. Choose an equivalence class $[p]_\nu \in \mathbf{J}_\nu^{-1}(\mu) \subset P_\nu$ for $p \in \mathbf{J}_K^{-1}(\nu)$. The equivalence relation is that associated with the map π_ν; that is, with the action of K_ν. For each such point, we have $p \in (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu, \nu)$ since by construction $p \in \mathbf{J}_K^{-1}(\nu)$ and also

$$\mathbf{J}_G(p) = (\mathbf{J}_G \circ i_\nu)(p) = \mathbf{J}_\nu([p]_\nu) = \mu.$$

Hence, it makes sense to consider the class $[p]_{(\mu,\nu)} \in P_{(\mu,\nu)}$. The result is independent of the representative, since any other representative of the same class has the form $k \cdot p$ where $k \in K_\nu$. This produces the same class in $P_{(\mu,\nu)}$ since for this latter space, the quotient is by $G_\mu \times K_\nu$. The map ϕ is therefore well defined.

This map ϕ is G_μ–invariant, and so it defines a quotient map

$$[\phi] : (P_\nu)_\mu \to P_{(\mu,\nu)}.$$

Chasing the definitions shows that this map is the inverse of the map $[\pi_\nu \circ j]$ constructed above. Thus, both are symplectic diffeomorphisms. ■

3 Semidirect Products

3.1 Summary

Background and Literature. In some applications one has two symmetry groups that do not commute and thus the commuting reduction by stages theorem does not apply. In this more general situation, it matters in what order one performs the reduction.

The main result covering the case of semidirect products is due to Marsden, Ratiu and Weinstein [1984a,b] with important previous versions (more or less in chronological order) due to Sudarshan and Mukunda [1974], Vinogradov and Kupershmidt [1977], Ratiu [1980], Guillemin and Sternberg [1980], Ratiu [1981], [1982], Marsden [1982], Marsden, Weinstein, Ratiu, Schmidt and Spencer [1983], Holm and Kupershmidt [1983] and Guillemin and Sternberg [1984].

The general theory of semidirect products was motivated by several examples of physical interest, such as the Poisson structure for compressible fluids and magnetohydrodynamics. These examples are discussed in the original papers. For some additional (very useful!) concrete applications of this theory, we refer to the literature already cited and, for underwater vehicle dynamics, to Leonard and Marsden [1997].

We shall first state the "classical" result and then shall give a more general one concerning actions by semidirect products on general symplectic manifolds.

Statement of the Theorem. The semidirect product reduction theorem states, roughly speaking, that for the semidirect product $S = G \circledS V$ where G is a group acting on a vector space V and S is the semidirect product, one can first reduce T^*S by V and then by G and one gets the same result as reducing by S. We will let $\mathfrak{s}$ denote the Lie algebra of S so that $\mathfrak{s} = \mathfrak{g} \circledS V$. Below we shall review the relevant conventions and notations for semidirect products in detail.

We now state the classical semidirect product reduction theorem precisely.

Theorem 3.1 (Semidirect Product Reduction Theorem). *As above, let $S = G \circledS V$ and choose $\sigma = (\mu, a) \in \mathfrak{g}^* \times V^*$ and reduce T^*S by the action of S at σ giving the coadjoint orbit $\mathcal{O}_\sigma$ through $\sigma \in \mathfrak{s}^*$. There is a symplectic diffeomorphism between $\mathcal{O}_\sigma$ and the reduced space obtained by reducing T^*G by the subgroup G_a (the isotropy of G for its action on V^* at the point $a \in V^*$) at the point $\mu|\mathfrak{g}_a$ where $\mathfrak{g}_a$ is the Lie algebra of G_a.*

Remarks. Note that in the semidirect product reduction theorem, only a and $\mu|\mathfrak{g}_a$ are used. Thus, one gets, as a corollary, the interesting fact that the semidirect product coadjoint orbits through $\sigma_1 = (\mu_1, a_1)$ and $\sigma_2 = (\mu_2, a_2)$

are symplectically diffeomorphic whenever $a_1 = a_2 = a$ and $\mu_1|\mathfrak{g}_a = \mu_2|\mathfrak{g}_a$. We shall see a similar phenomenon in more general situations of group extensions later.

Semidirect Product Actions. The preceding result is a special case of a theorem we shall prove on reduction by stages for semidirect products acting on a symplectic manifold. To state this more general result, consider a symplectic action of S on a symplectic manifold P and assume that this action has an equivariant momentum map $\mathbf{J}_S : P \to \mathfrak{s}^*$. Since V is a (normal) subgroup of S, it also acts on P and has a momentum map $\mathbf{J}_V : P \to V^*$ given by

$$\mathbf{J}_V = i_V^* \circ \mathbf{J}_S \,,$$

where $i_V : V \to \mathfrak{s}$ is the inclusion $v \mapsto (0, v)$ and $i_V^* : \mathfrak{s}^* \to V^*$ is its dual. We carry out the reduction of P by S at a regular value $\sigma = (\mu, a)$ of the momentum map $\mathbf{J}_S$ for S in two stages using the following procedure. First, reduce P by V at the value a (assume it to be a regular value) to get the reduced manifold $P_a = \mathbf{J}_V^{-1}(a)/V$. Second, form the group G_a consisting of elements of G that leave the point a fixed, using the action of G on V^*. We shall show shortly that the group G_a acts on P_a and has an induced equivariant momentum map $\mathbf{J}_a : P_a \to \mathfrak{g}_a^*$, where $\mathfrak{g}_a$ is the Lie algebra of G_a. Third, reduce P_a at the point $\mu_a := \mu|\mathfrak{g}_a$ to get the reduced space $(P_a)_{\mu_a} = \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

Theorem 3.2 (Reduction by Stages for Semidirect Product Actions). *The reduced space $(P_a)_{\mu_a}$ is symplectically diffeomorphic to the reduced space P_σ obtained by reducing P by S at the point $\sigma = (\mu, a)$.*

We recover the preceding theorem by choosing $P = T^*S$. The commuting reduction theorem for the case in which K is a vector space results from semidirect product reduction when we take the action of G on K to be trivial. This already suggests that there is a generalization of the semidirect product reduction theorem to the case in which V is replaced by a general Lie group. We give this more general result later. Note that in the commuting reduction theorem, what we called ν is called a in the semidirect product reduction theorem.

The original papers of Marsden, Ratiu and Weinstein [1984a,b] give a direct proof of Theorem 3.1 along lines somewhat different than we shall present here. The proofs we give in this paper have the advantage that they work for the more general reduction by stages theorem.

Classifying Orbits. Combined with the cotangent bundle reduction theorem (see Marsden [1992] for an exposition and references), the semidirect product reduction theorem is a very useful tool. For example, using these techniques, one sees readily that the generic coadjoint orbits for the Euclidean group are cotangent bundles of spheres with the associated coadjoint orbit symplectic structure given by the canonical structure plus a magnetic term. We shall discuss this example in detail later.

Reducing Dynamics. There is a method for reducing dynamics that is associated with the geometry of the semidirect product reduction theorem. In effect, one can start with a Hamiltonian on either of the phase spaces and induce one (and hence its associated dynamics) on the other space in a natural way. For example, in many applications, one starts with a Hamiltonian H_a on T^*G that depends parametrically on a variable $a \in V^*$; this parametric dependence identifies the space V^* and hence V. The Hamiltonian, regarded as a map $H : T^*G \times V^* \to \mathbb{R}$ should be invariant on T^*G under the action of G on $T^*G \times V^*$. This condition is equivalent to the invariance of the corresponding function on $T^*S = T^*G \times V \times V^*$ extended to be constant in the variable V under the action of the semidirect product. This observation allows one to identify the reduced dynamics of H_a on T^*Q reduced by G_a with a Hamiltonian system on $\mathfrak{s}^*$ or on the coadjoint orbits of $\mathfrak{s}^*$. For example, this observation is extremely useful in underwater vehicle dynamics (again, see Leonard and Marsden [1997]).

3.2 Generalities on Semidirect Products

Now we embark on the proof of the semidirect product reduction theorem. We begin by recalling some definitions and properties of semidirect products. Let V be a vector space and assume that the Lie group G acts (on the left) by linear maps on V, and hence G also acts on its dual space V^*. As sets, the semidirect product $S = G \circledS V$ is the Cartesian product $S = G \times V$ and the group multiplication is given by

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

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

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

where we denote the induced action of $\mathfrak{g}$ on V by concatenation, as in $\xi_1 v_2$.

Below we will need the formulas for the adjoint and the coadjoint actions for semidirect products. Denoting these and other actions by simple concatenation, they are given by (see, e.g.,, Marsden, Ratiu and Weinstein [1984a,b]):

$$(g, v)(\xi, u) = (g\xi, gu - \rho_v(g\xi))$$

and

$$(g,v)(\mu,a) = (g\mu + \rho_v^*(ga), ga),$$

where $(g,v) \in S = G \times V$, $(\xi, u) \in \mathfrak{s} = \mathfrak{g} \times V$, $(\mu, a) \in \mathfrak{s}^* = \mathfrak{g}^* \times V^*$, and where $\rho_v : \mathfrak{g} \to V$ is defined by $\rho_v(\xi) = \xi v$, the infinitesimal action of ξ on v. The map $\rho_v^* : V^* \to \mathfrak{g}^*$ is the dual of the map ρ_v. The symbol ga denotes the (left) dual action of G on V^*, that is, the inverse of the dual isomorphism induced by $g \in G$ on V. The corresponding (left) action on the dual space is denoted by ξa for $a \in V^*$, that is,

$$\langle \xi a, v \rangle := - \langle a, \xi v \rangle \ .$$

Lie-Poisson Brackets and Hamiltonian Vector Fields. For completeness, we give the formula for the $\pm$ Lie-Poisson bracket of $F, K : s^* \to \mathbb{R}$:

$$\{F, K\}_{\pm}(\mu, a) \quad = \quad \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta K}{\delta \mu} \right] \right\rangle \pm \left\langle a, \frac{\delta F}{\delta \mu} \cdot \frac{\delta K}{\delta a} - \frac{\delta K}{\delta \mu} \cdot \frac{\delta F}{\delta a} \right\rangle, \quad (3.1)$$

where $\delta F/\delta \mu \in \mathfrak{g}$, $\delta F/\delta a \in V$ are the functional derivatives. Also, one verifies that the Hamiltonian vector field of $H : s^* \to \mathbb{R}$ is given by

$$X_H(\mu, a) = \mp \left(\mathrm{ad}^*_{\delta H/\delta \mu} \mu - \rho^*_{\delta H/\delta a} a, \ \frac{\delta H}{\delta \mu} \cdot a \right). \quad (3.2)$$

Symplectic Actions by Semidirect Products. Next we consider a symplectic action of S on a symplectic manifold P and assume that this action has an equivariant momentum map $\mathbf{J}_S : P \to \mathfrak{s}^*$. Since V is a (normal) subgroup of S, it also acts on P and has a momentum map $\mathbf{J}_V : P \to V^*$ given by

$$\mathbf{J}_V = i_V^* \circ \mathbf{J}_S,$$

where $i_V : V \to \mathfrak{s}$ is the inclusion $v \mapsto (0, v)$ and $i_V^* : \mathfrak{s}^* \to V^*$ is its dual. We think of this merely as saying that $\mathbf{J}_V$ is the second component of $\mathbf{J}_S$.

We can regard G as a subgroup of S by $g \mapsto (g, 0)$. Thus, G also has an equivariant momentum map $\mathbf{J}_G : P \to \mathfrak{g}^*$ that is the first component of $\mathbf{J}_S$ but this will play a secondary role in what follows. On the other hand, equivariance of $\mathbf{J}_S$ under G implies the following relation for $\mathbf{J}_V$:

$$\mathbf{J}_V(gz) = g\mathbf{J}_V(z), \quad (3.3)$$

where $z \in P$ and we denote the appropriate action of $g \in G$ on an element by concatenation, as before. To prove (3.3), one uses the fact that for the coadjoint action of S on $\mathfrak{s}^*$ the second component is just the dual of the given action of G on V.

3.3 The Reduction by Stages Construction

We carry out reduction of P by S at a regular value $\sigma = (\mu, a)$ of the momentum map $\mathbf{J}_S$ for S in two stages (see figure 3.1).

- First, reduce P by V at the value $a \in V^*$ (assume it to be a regular value) to get the reduced manifold $P_a = \mathbf{J}_V^{-1}(a)/V$. Since the reduction is by an Abelian group, the quotient is done by the whole of V. We will denote the projection to the reduced by

$$\pi_a : \mathbf{J}_V^{-1}(a) \to P_a.$$

- Second, form the group G_a consisting of elements of G that leave the point a fixed using the induced action of G on V^*. As we show below, the group G_a acts on P_a and has an induced equivariant momentum map $\mathbf{J}_a : P_a \to \mathfrak{g}_a^*$, where $\mathfrak{g}_a$ is the Lie algebra of G_a.
- Third, using this action of G_a, reduce P_a at the point $\mu_a := \mu|\mathfrak{g}_a$ to get the reduced manifold $(P_a)_{\mu_a} = \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

We next check these claims.

Inducing an Action. We first check that we get a symplectic action of G_a on the V–reduced space P_a. We do this in the following lemmas.

Lemma 3.3. *The group G_a leaves the set $\mathbf{J}_V^{-1}(a)$ invariant.*

Proof. Suppose that $\mathbf{J}_V(z) = a$ and that $g \in G$ leaves a invariant. Then by the equivariance relation noted above, we have $\mathbf{J}_V(gz) = g\mathbf{J}_V(z) = ga = a$. Thus, G_a acts on the set $\mathbf{J}_V^{-1}(a)$. ▼

Lemma 3.4. *The action of G_a on $\mathbf{J}_V^{-1}(a)$ constructed in the preceding lemma, induces an action Ψ^a on the quotient space $P_a = \mathbf{J}_V^{-1}(a)/V$.*

Proof. If we let elements of the quotient space be denoted by $[z]_a$, regarded as equivalence classes, then we claim that $g[z]_a = [gz]_a$ defines the action. We only need to show that it is well defined. Indeed, for any $v \in V$ we have $[z]_a = [vz]_a$, so that identifying $v = (e, v)$ and $g = (g, 0)$ in the semidirect product, it follows that

$$[gvz]_a = [(g,0)(e,v)z]_a = [(g,gv)z]_a = [(e,gv)(g,0)z]_a = [(gv)(gz)]_a = [gz]_a.$$

Thus, the action $\Psi^a : (g, [z]_a) \in G_a \times P_a \mapsto [gz]_a \in P_a$ of G_a on the V–reduced space P_a is well defined. ▼

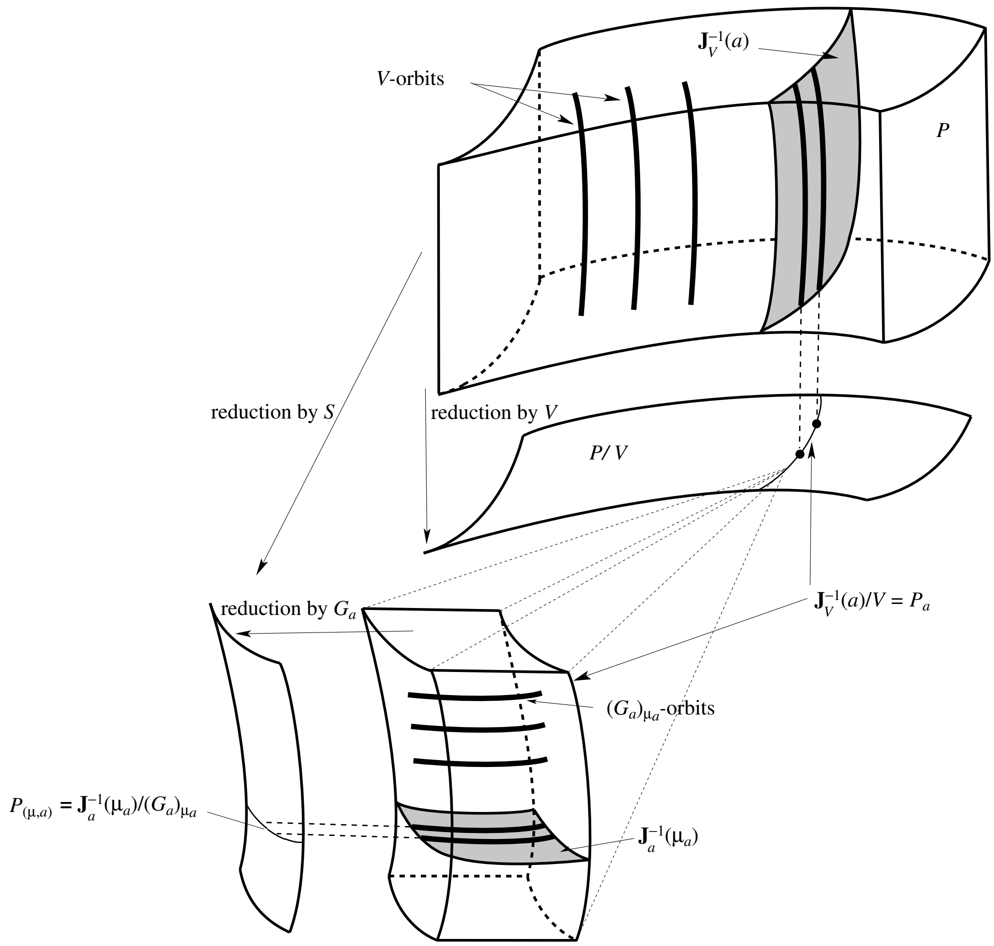

Figure 3.1: A schematic of reduction by stages for semidirect products.

The Induced Action is Symplectic. Our next task is to show that the induced action just obtained is symplectic.

Lemma 3.5. *The action Ψ^a of G_a on the quotient space $P_a = \mathbf{J}_V^{-1}(a)/V$ constructed in the preceding lemma, is symplectic.*

Proof. Let $\pi_a : \mathbf{J}_V^{-1}(a) \to P_a$ denote the natural projection and $i_a : \mathbf{J}_V^{-1}(a) \to P$ be the inclusion. Denote by $\Psi_g : P \to P$ the action of $g \in G$ on P. The preceding lemma states that $(i_a \circ \Psi_g)|\mathbf{J}_V^{-1}(a) = \Psi_g \circ i_a$ for any $g \in G_a$. By construction, $\Psi_g^a \circ \pi_a = (\pi_a \circ \Psi_g)|\mathbf{J}_V^{-1}(a)$. The characterization $i_a^*\Omega = \pi_a^*\Omega_a$ of the reduced symplectic form Ω_a on P_a yields then

$$\pi_a^*(\Psi_g^a)^*\Omega_a = \Psi_g^*\pi_a^*\Omega_a = \Psi_g^* i_a^*\Omega = i_a^*\Psi_g^*\Omega = i_a^*\Omega = \pi_a^*\Omega_a.$$

Since π_a is a surjective submersion, we conclude that

$$(\Psi_g^a)^* \Omega_a = \Omega_a .$$

Thus, we have a symplectic action of G_a on P_a. ▼

An Induced Momentum Map. We next check that the symplectic action obtained in the preceding lemma has an equivariant momentum map. As we shall see later, in more general cases, this turns out to be a critical step; in fact, even for central extensions, the momentum map induced at this step *need not be equivariant*—it is a special feature of semidirect products, about which we shall have more to say later.

Lemma 3.6. *The symplectic action Ψ^a on the quotient space $P_a = \mathbf{J}_V^{-1}(a)/V$ has an equivariant momentum map.*

Proof. We first show that the composition of the restriction $\mathbf{J}_S | \mathbf{J}_V^{-1}(a)$ with the projection to $\mathfrak{g}_a^*$ induces a well defined map $\mathbf{J}_a : P_a \to \mathfrak{g}_a^*$. To check this, note that for $z \in \mathbf{J}_V^{-1}(a)$, and $\xi \in \mathfrak{g}_a$, equivariance gives

$$\langle \mathbf{J}_S(vz), \xi \rangle = \langle v \mathbf{J}_S(z), \xi \rangle = \langle (e, v) \mathbf{J}_S(z), \xi \rangle = \left\langle \mathbf{J}_S(z), (e, v)^{-1}(\xi, 0) \right\rangle .$$

Here, the symbol $(e, v)^{-1}(\xi, 0)$ means the adjoint action of the group element $(e, v)^{-1} = (e, -v)$ on the Lie algebra element $(\xi, 0)$. Thus, $(e, v)^{-1}(\xi, 0) = (\xi, \xi v)$, and so, continuing the above calculation, and using the fact that $\mathbf{J}_V(z) = a$, we get

$$\begin{aligned} \langle \mathbf{J}_S(vz), \xi \rangle &= \langle \mathbf{J}_S(z), (\xi, \xi v) \rangle = \langle \mathbf{J}_G(z), \xi \rangle + \langle \mathbf{J}_V(z), \xi v \rangle \\ &= \langle \mathbf{J}_G(z), \xi \rangle - \langle \xi a, v \rangle = \langle \mathbf{J}_G(z), \xi \rangle . \end{aligned}$$

In this calculation, the term $\langle \xi a, v \rangle$ is zero since $\xi \in \mathfrak{g}_a$. Thus, we have shown that the expression

$$\langle \mathbf{J}_a([z]_a), \xi \rangle = \langle \mathbf{J}_G(z), \xi \rangle$$

for $\xi \in \mathfrak{g}_a$ is well defined. Here, $[z]_a \in P_a$ denotes the V–orbit of $z \in \mathbf{J}_V^{-1}(a)$. This expression may be written as

$$\mathbf{J}_a \circ \pi_a = \iota_a^* \circ \mathbf{J}_G \circ i_a ,$$

where $\iota_a : \mathfrak{g}_a \to \mathfrak{g}$ is the inclusion map and $\iota_a^* : \mathfrak{g}^* \to \mathfrak{g}_a^*$ is its dual.

Next, we show that the map $\mathbf{J}_a$ is the momentum map of the G_a–action on P_a. Since the vector fields $\xi_P | (\mathbf{J}_V^{-1}(a))$ and ξ_{P_a} are π_a–related for all $\xi \in \mathfrak{g}_a$, we have

$$\pi_a^* \left(\mathbf{i}_{\xi_{P_a}} \Omega_a \right) = \mathbf{i}_{\xi_P} i_a^* \Omega = i_a^* \left(\mathbf{i}_{\xi_P} \Omega \right) = i_a^* \left(\mathbf{d} \left\langle \mathbf{J}_G, \xi \right\rangle \right) = \pi_a^* \left(\mathbf{d} \left\langle \mathbf{J}_a, \xi \right\rangle \right) .$$

Again, since π_a is a surjective submersion, we may conclude that

$$\mathbf{i}_{\xi_{P_a}} \Omega_a = \mathbf{d} \langle \mathbf{J}_a, \xi \rangle$$

and hence $\mathbf{J}_a$ is the momentum map for the G_a action on P_a.

Equivariance of $\mathbf{J}_a$ follows from that for $\mathbf{J}_G$, by a diagram chasing argument as above, using the identity $\mathbf{J}_a \circ \pi_a = \iota_a^* \circ \mathbf{J}_G \circ i_a$ and the relations between the actions of G on $\mathbf{J}_V^{-1}(a)$ and of G_a on P_a. ▼

Proof of Theorem 3.2. Having established these preliminary facts, we are ready to prove the main reduction by stages theorem for semidirect products.

Let $\sigma = (\mu, a)$. Start with the inclusion map

$$j : \mathbf{J}_S^{-1}(\sigma) \to \mathbf{J}_V^{-1}(a)$$

which makes sense since the second component of σ is a. Composing this map with π_a, we get the smooth map

$$\pi_a \circ j : \mathbf{J}_S^{-1}(\sigma) \to P_a.$$

This map takes values in $\mathbf{J}_a^{-1}(\mu_a)$ because of the relation $\mathbf{J}_a \circ \pi_a = \iota_a^* \circ \mathbf{J}_G \circ i_a$ and $\mu_a = \iota_a^*(\mu)$. Thus, we can regard it as a map

$$\pi_a \circ j : \mathbf{J}_S^{-1}(\sigma) \to \mathbf{J}_a^{-1}(\mu_a).$$

There is a smooth Lie group homomorphism $\psi : S_\sigma \to (G_a)_{\mu_a}$ defined by projection onto the first factor. The first component g of $(g, v) \in S_\sigma$ lies in $(G_a)_{\mu_a}$ because

$$(\mu, a) = (g, v)(\mu, a) = (g\mu + \rho_v^*(ga), ga)$$

implies, from the second component, that $g \in G_a$ and from the first component, the identity $\iota_a^* \rho_v^* a = 0$, and the G_a–equivariance of the map ι_a, that g also leaves μ_a invariant.

The map $\pi_a \circ j$ is equivariant with respect to the action of S_σ on the domain and $(G_a)_{\mu_a}$ on the range via the homomorphism ψ. Thus, $\pi_a \circ j$ induces a smooth map

$$[\pi_a \circ j] : P_\sigma \to (P_a)_{\mu_a}.$$

Diagram chasing, as above, shows that this map is symplectic and hence an immersion.

We will show that this map is a diffeomorphism by finding an inverse. We begin with the construction of a map (see figure 3.2)

$$\phi : \mathbf{J}_a^{-1}(\mu_a) \to P_\sigma.$$

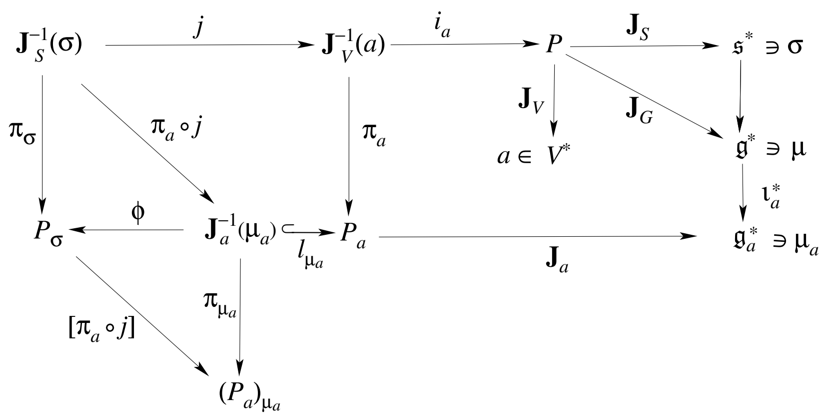

Figure 3.2: The maps in the proof of the semidirect product reduction theorem.

To do this, take an equivalence class $[z]_a \in \mathbf{J}_a^{-1}(\mu_a) \subset P_a$ for $z \in \mathbf{J}_V^{-1}(a)$, that is, the V–orbit of z. For each such point, we will try to choose some $v \in V$ such that $vz \in \mathbf{J}_S^{-1}(\sigma)$. For this to hold, we must have

$$(\mu, a) = \mathbf{J}_S(vz).$$

By equivariance, the right hand side equals

$$\begin{aligned} v\mathbf{J}_S(z) &= (e,v)(\mathbf{J}_G(z), \mathbf{J}_V(z)) \\ &= (e,v)(\mathbf{J}_G(z), a) \\ &= (\mathbf{J}_G(z) + \rho_v^*(a), a). \end{aligned}$$

Thus, we require that

$$\mu = \mathbf{J}_G(z) + \rho_v^*(a).$$

That this is possible follows from the next lemma.

Lemma 3.7. *If* $\mathfrak{g}_a^\circ = \{\nu \in \mathfrak{g}^* \mid \nu|\mathfrak{g}_a = 0\}$ *denotes the annihilator of* $\mathfrak{g}_a$ *in* $\mathfrak{g}^*$, *we have*

$$\mathfrak{g}_a^\circ = \{\rho_v^* a \mid v \in V\}.$$

Proof. The identity we showed above, namely $\iota_a^* \rho_v^* a = 0$, shows that

$$\mathfrak{g}_a^\circ \supset \{\rho_v^* a \mid v \in V\}.$$

Now we use the following elementary fact from linear algebra. Let E and F be vector spaces, and $F_0 \subset F$ be a subspace. Let $T : E \to F^*$ be a linear map whose range lies in the annihilator F_0° of F_0 and such that every element $f \in F$ that annihilates the range of T is in F_0. Then T maps *onto* F_0°. [1]

In our case, we choose $E = V$, $F = \mathfrak{g}$, $F_0 = \mathfrak{g}_a$, and we let $T : V \to \mathfrak{g}^*$ be defined by $T(v) = \rho_v^*(a)$. To verify the hypothesis, note that we have already shown that the range of T lies in the annihilator of $\mathfrak{g}_a$. Let $\xi \in \mathfrak{g}$ annihilate the range of T. Thus, for all $v \in V$,

$$0 = \langle \xi, \rho_v^* a \rangle = \langle \rho_v \xi, a \rangle = \langle \xi v, a \rangle = - \langle v, \xi a \rangle,$$

and so $\xi \in \mathfrak{g}_a$ as required. Thus, the lemma is proved. ▼

We apply the lemma to $\mu - \mathbf{J}_G(z)$, which is in the annihilator of $\mathfrak{g}_a$ because $[z]_a \in \mathbf{J}_a^{-1}(\mu_a)$ and hence $\iota_a^*(\mathbf{J}_G(p)) = \mu_a$. Thus there is a v such that $\mu - \mathbf{J}_G(z) = \rho_v^* a$.

The above argument shows how to construct v so that $vz \in \mathbf{J}_S^{-1}(\sigma)$. We define the map

$$\phi : [z]_a \in \mathbf{J}_a^{-1}(\mu_a) \mapsto [vz]_\sigma \in P_\sigma,$$

where $v \in V$ has been chosen as above and $[vz]_\sigma$ is the S_σ–equivalence class in P_σ of vz.

To show that the map ϕ so constructed is well defined, we replace z by another representative uz of the same class $[z]_a$; here u is an arbitrary element of V. Then one chooses v_1 so that $\mathbf{J}_S(v_1 uz) = \sigma$. Now we must show that $[vz]_\sigma = [v_1 uz]_\sigma$. In other words, we must show that there is a group element $(g, w) \in S_\sigma$ such that $(g, w)(e, v)z = (e, v_1)(e, u)z$. This will hold if we can show that $(g, w) := (e, v_1)(e, u)(e, v)^{-1} \in S_\sigma$. However, by construction, $\mathbf{J}_S(vz) = \sigma = \mathbf{J}_S(v_1 uz)$; in other words, we have

$$\sigma = (\mu, a) = (e, v)\mathbf{J}_S(z) = (e, v_1)(e, u)\mathbf{J}_S(z).$$

Thus, by isolating $\mathbf{J}_S(z)$, we get $(e, v)^{-1}\sigma = (e, u)^{-1}(e, v_1)^{-1}\sigma$ and so the element $(g, w) = (e, v_1)(e, u)(e, v)^{-1}$ belongs to S_σ. Thus, the map ϕ is well defined.

The strategy for proving smoothness of ϕ is to choose a local trivialization of the V bundle $\mathbf{J}_V^{-1}(a) \to \mathbf{J}_a^{-1}(\mu_a)$ and define a local section which takes values in the image of $\mathbf{J}_S^{-1}(\sigma)$ under the embedding j. Smoothness of the local section follows by using a complement to the kernel of the linear map $v \mapsto \rho_v^*(a)$ that defines the solution v of the equation $\rho_v^*(a) = \mu - \mathbf{J}_G(z)$. Using such a complement depending smoothly on the data creates a uniquely defined smooth selection of a solution.

[1] We are phrasing things this way so that the basic framework will also apply in the infinite dimensional case, with the understanding that at this point one would invoke arguments used in the Fredholm alternative theorem. In the finite dimensional case, the result may be proved by a dimension count.

Next, we must show that the map ϕ is $(G_a)_{\mu_a}$–invariant. To see this, let $[z]_a \in \mathbf{J}_a^{-1}(\mu_a)$ and let $g_0 \in (G_a)_{\mu_a}$. Choose $v \in V$ so that $vz \in \mathbf{J}_S^{-1}(\sigma)$ and let $u \in V$ be chosen so that $ug_0z \in \mathbf{J}_S^{-1}(\sigma)$. We must show that $[vz]_\sigma = [ug_0z]_\sigma$. Thus, we must find an element $(g, w) \in S_\sigma$ such that $(g, w)(e, v)z = (e, u)(g_0, 0)z$. This will hold if we can show that $(g, w) := (e, u)(g_0, 0)(e, v)^{-1} \in S_\sigma$. Since $\sigma = \mathbf{J}_S(vz) = \mathbf{J}_S(ug_0z)$, by equivariance of $\mathbf{J}_S$ we get,

$$\sigma = (e, v)\mathbf{J}_S(z) = (e, u)(g_0, 0)\mathbf{J}_S(z).$$

Isolating $\mathbf{J}_S(z)$, this implies that $(e, v)^{-1}\sigma = (g_0, 0)^{-1}(e, u)^{-1}\sigma$ which means that indeed $(g, w) = (e, u)(g_0, 0)(e, v)^{-1} \in S_\sigma$. Hence ϕ is $(G_a)_{\mu_a}$–invariant, and so induces a well defined map

$$[\phi] : (P_a)_{\mu_a} \to P_\sigma.$$

Chasing the definitions shows that $[\phi]$ is the inverse of the map $[\pi_a \circ j]$.

Smoothness of $[\phi]$ follows from smoothness of ϕ since the quotient by the group action, π_a is a smooth surjective submersion. Thus, both $[\pi_a \circ j]$ and ϕ are symplectic diffeomorphisms. ■

In this framework, one can also, of course, reduce the dynamics of a given invariant Hamiltonian as was done for the case of reduction by T^*S by stages.

Remarks.

1. Choose $P = T^*S$ in the preceding theorem, with the cotangent action of S on T^*S induced by left translations of S on itself. Reducing T^*S by the action of V gives a space naturally isomorphic to T^*G—this may be checked directly, but we will detail the *real* reason this is so in the next section. Thus, the reduction by stages theorem gives as a corollary, the semidirect product reduction Theorem 3.1.

2. The original proof of this result in Marsden, Ratiu and Weinstein [1984a,b] essentially used the map $[\phi]$ constructed above to obtain the required symplectic diffeomorphism. However, the generalization presented here to obtain reduction by stages for semidirect product actions, required an essential modification of the original method.

3. In the following section we shall give some details for reduction by stages for SE(3), the Euclidean group of $\mathbb{R}^3$. This illustrates the classical Semidirect Product Reduction Theorem 3.1. We now briefly describe two examples which require the more general result of Theorem 3.2. First, consider a pseudo-rigid body in a fluid; that is, a body which can undergo linear deformations and moving through perfect potential flow,

as in Leonard and Marsden [1997]. Here the phase space is $P = T^* \mathrm{GE}(3)$ (where GE(3) is the semidirect product $\mathrm{GL}(3) \circledS \mathbb{R}^3$) and the symmetry group we want to reduce by is SE(3); it acts on GE(3) on the left by composition and hence on $T^* \mathrm{GE}(3)$ by cotangent lift. According to the general theory, we can reduce by the action of $\mathbb{R}^3$ first and then by SO(3). This example has the interesting feature that the center of mass need not move uniformly along a straight line, so the first reduction by translations is not trivial. The same thing happens for a rigid body moving in a fluid.

A second, more sophisticated example is a fully elastic body, in which case, P is the cotangent bundle of the space of all embeddings of a reference configuration into $\mathbb{R}^3$ (as in Marsden and Hughes [1983]) and we take the group again to be SE(3) acting by composition on the left. Again, one can perform reduction in two stages.

We comment that the reduction by stages philosophy is quite helpful in understanding the dynamics and stability of underwater vehicle dynamics, as in Leonard and Marsden [1997].

4. The next level of generality is the case of the semidirect product of two *nonabelian groups*. Namely, in the preceding case we replace the vector space (thought of as an Abelian group) V by a general Lie group N, with G acting on N by group homomorphisms. Already this case is quite interesting and nontrivial. We shall not discuss this situation in detail here, leaving it for Marsden, Misiołek, Perlmutter and Ratiu [1998]. However, later on we shall briefly discuss the *even more general* case of *group extensions*, of which this is a special case.

5. An example involving the semidirect product of two nonabelian groups is the following. We consider the semidirect product $\mathrm{SU}(2) \circledS \mathrm{SU}(2)$ where the first factor acts on the second by conjugation. Each factor acts on $\mathbb{C}^2 \cong \mathbb{R}^4$ in the obvious way and therefore the semidirect product does as well. The cotangent lift of this action gives a symplectic action on the cotangent bundle $T^*\mathbb{R}^4$. We can then study the reduction of this action by stages, reducing first by the normal subgroup SU(2) and then by the relevant quotient group.

   We believe that this example is related to the fact that the Laplace (also known as Runge–Lenz) vector, a conserved quantity in addition to the usual angular momentum for the Kepler problem, is connected with an SO(4) symmetry (modern references are Moser [1970], Souriau [1973], Bates [1988], and Guillemin and Sternberg [1990]).

4 Cotangent Bundle Reduction by Stages for Semidirect Products

In this section we will couple the semidirect product reduction theorem with cotangent bundle reduction theory in order to study the reduced spaces for the right cotangent lifted action of $G \circledS V$ on $T^*(G \circledS V)$ in more detail. To carry this out, we first construct a mechanical connection on the bundle $G \circledS V \to G$ and prove that this connection is flat. This will allow us to identify (equivariantly) the first (V–reduced) space with $(T^*G, \Omega_{\mathrm{can}})$. We will then be in a position to apply the cotangent bundle reduction theory again to complete the orbit classification.

4.1 Flatness of the First Connection

As in Section 3, let $S = G \circledS V$ be the semidirect product of a Lie group G and a vector space V with multiplication

$$(g, v)(h, w) = (gh, v + gw), \tag{4.1}$$

where $g, h \in G$ and $v, w \in V$. Recall that the Lie algebra of S is the semidirect product $\mathfrak{s} = \mathfrak{g} \circledS V$ with the commutator

$$[(\xi, v), (\eta, w)] = ([\xi, \eta],\, \xi w - \eta v)\,, \tag{4.2}$$

where $\xi, \eta \in \mathfrak{g}$ and $v, w \in V$. In what follows it is convenient to explicitly introduce the homomorphism $\phi : G \to \mathrm{Aut}(V)$ defining the given G–representation on V.

The Mechanical Connection. For the construction that follows we will recall the relevant definitions and basic properties of *mechanical connections* and *locked inertia tensor* (see Marsden [1992]).

Let $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$ and $\langle \cdot, \cdot \rangle_V$ be two inner products on the Lie algebra $\mathfrak{g}$ and on the vector space V, respectively. Then

$$\langle (\xi, v), (\eta, w) \rangle_{\mathfrak{s}} = \langle \xi, \eta \rangle_{\mathfrak{g}} + \langle v, w \rangle_V, \tag{4.3}$$

for any $(\xi, v), (\eta, w) \in \mathfrak{s}$ defines an inner product on $\mathfrak{s}$. Extend it to a right–invariant Riemannian metric on S by

$$\langle (X, u), (Y, w) \rangle_{(g,v)} = \langle T_{(g,v)} R_{(g,v)^{-1}} (X, u), T_{(g,v)} R_{(g,v)^{-1}} (Y, w) \rangle_{\mathfrak{s}}, \tag{4.4}$$

where $(g, v) \in S$, $(X, u), (Y, w) \in T_{(g,v)} S$, and $R_{(g,v)}$ is right translation on S.[2] The derivative of $R_{(h,w)}$ is readily computed from (4.1) to be

$$T_{(g,v)} R_{(h,w)} (Y, u) = (Y \cdot h,\; u + T_g \phi^w (Y)), \tag{4.5}$$

[2] Our choice of *right translations* is motivated by infinite dimensional applications to diffeomorphism groups etc. Of course, there is a left invariant analogue.

where $(Y, w) \in T_{(g,v)}S$, $Y \cdot h = T_g R_h(Y)$, R_h is the right translation on G, and $\phi^w : G \to V$ is given by $\phi^w(g) = gw$.

Since, by construction, the Riemannian metric (4.4) is also right V–invariant, the mechanical connection $\mathcal{A}^V$ on S is defined by the formula

$$\mathcal{A}^V_{(g,v)}(X, u) = \left(\mathbb{I}^{-1}_{(g,v)} \circ \mathbf{J}_V\right) \left(\langle (X, u), \cdot \rangle_{(g,v)}\right), \tag{4.6}$$

where $\mathbb{I}_{(g,v)} : V \to V^*$ is the locked inertia tensor[3] and $\mathbf{J}_V : T^*S \to V^*$ is the momentum map for the action of V on T^*S.

The Flatness Calculation. The "reason" why the first reduced space is so simple is that the mechanical connection $\mathcal{A}^V$ is flat. However, the proof of this is unexpectedly tricky. We begin by explicitly determining the locked inertia tensor and the mechanical connection.

Given $u \in V$, from (4.1) we find the infinitesimal generator for the right V–action on S:

$$u_S(g, v) = \left.\frac{d}{dt}\right|_{t=0} \big((g, v) \cdot (e, tu)\big) = (0, gu). \tag{4.7}$$

Thus, the locked inertia tensor has the expression

$$\begin{aligned} \mathbb{I}_{(g,v)}(u)(w) &= \langle u_S(g, v), w_S(g, v)\rangle_{(g,v)} \\ &= \langle gu, gv\rangle_V = \langle \phi(g)^*\phi(g)u, w\rangle_V, \end{aligned} \tag{4.8}$$

for any $(g, v) \in S$ and any $u, v \in V$. Similarly, for any $(X, u) \in T_{(g,v)}S$ and any $w \in V$, we get from (4.4), (4.5), and $(g, v)^{-1} = (g^{-1}, -g^{-1}v)$,

$$\begin{aligned} \left\langle \mathbf{J}_V \left(\langle (X, u), \cdot \rangle_{(g,v)}\right), w \right\rangle &= \langle (X, u), w_S(g, v)\rangle_{(g,v)} \\ &= \langle T_{(g,v)}R_{(g,v)^{-1}}(X, u), T_{(g,v)}R_{(g,v)^{-1}}(0, gw)\rangle_{\mathfrak{s}} \\ &= \langle (X \cdot u, u - T_g\phi^{g^{-1}v}(X)), (0, gw)\rangle_{\mathfrak{s}} \\ &= \langle \phi(g)^*\big(u - T_g\phi^{g^{-1}v}(X)\big), w\rangle_V. \end{aligned} \tag{4.9}$$

However, taking the derivative relative to h at e in the direction $\xi \in \mathfrak{g}$ in the identity $\phi^{g^{-1}v}(hg) = \phi^v(h)$ yields

$$T_g\phi^{g^{-1}v}(T_eR_g\xi) = T_e\phi^v(\xi) = \xi v,$$

that is,

$$T_g\phi^{g^{-1}v} = T_e\phi^v \circ T_gR_{g^{-1}}. \tag{4.10}$$

[3] The locked inertia tensor (see, for example, Marsden [1992]) for the isometric action of a Lie group K on a manifold M is, for each $x \in M$, the linear map $\mathbb{I}_x : \mathfrak{k} \to \mathfrak{k}^*$ defined in terms of the Riemannian inner product $\langle\, , \rangle$ by $\mathbb{I}_x(\xi, \eta) = \langle \xi_M(x), \eta_M(x)\rangle$.

Therefore

$$\begin{aligned}\mathbf{J}_V\left(\langle (X,u),\cdot\rangle_{(g,v)}\right) &= \phi(g)^*\big(u - T_e\phi^v(T_gR_{g^{-1}}X)\big)\\ &= \phi(g)^*\big(u-(X\cdot g^{-1})v\big).\end{aligned} \tag{4.11}$$

Combining formulas (4.6), (4.8), and (4.11), we get

$$\mathcal{A}^V_{(g,v)}(X,u) = g^{-1}\Big(u-(X\cdot g^{-1})v\Big). \tag{4.12}$$

The following formula which was proved along the way will be useful later on:

$$T_{(g,v)}R_{(g,v)^{-1}}(X,u) = \big(X\cdot g^{-1},\, u-(X\cdot g^{-1})v\big)\,. \tag{4.13}$$

Theorem 4.1. *The mechanical connection $\mathcal{A}^V$ defined on the principal V–bundle $S\to G$ by formula* (4.6) *is flat.*

Remarks. If one's goal is simply to pick a connection on the the principal V–bundle $S\to G$ in order to realize the first reduced space as T^*G with the canonical structure, then one may use the *trivial* connection associated with the product structure $S = G Ⓢ V$, so that the connection form is simply projection to V. This connection has the needed equivariance properties to realize the reduced space as T^*G and identifies the resulting action of G_a as the right action on T^*G. On the other hand, in more general situations in which the bundles may not be trivial, it is the mechanical connection which is used in the construction and so it is of interest to use it here as well. In particular, in the second stage of reduction, one needs a connection on the (generally) nontrivial bundle $G\to G/G_a$ and such a connection is naturally induced by the mechanical connection.

Proof of Theorem 4.1. To calculate the curvature of $\mathcal{A}^V$ it suffices to compute the exterior derivative $\mathbf{d}\mathcal{A}^V$. For this purpose, we shall extend $(X,u),(Y,w)\in T_{(g,v)}S$ to right invariant vector fields $(\bar X,\bar u)$ and $(\bar Y,\bar w)$ on S. Concretely, if we define, $\xi = X\cdot g^{-1}, \zeta = Y\cdot g^{-1}\in\mathfrak{g}$, we have by (4.5) and (4.13)

$$(\bar X,\bar u)(\bar g,\bar v) := T_{(e,0)}R_{(\bar g,\bar v)}(\xi, u-\xi v) = (\xi\cdot\bar g,\, u+\xi(\bar v - v))$$

$$(\bar Y,\bar w)(\bar g,\bar v) := T_{(e,0)}R_{(\bar g,\bar v)}(\zeta, w-\zeta v) = (\zeta\cdot\bar g,\, w+\zeta(\bar v - v))\,.$$

Therefore, by (4.12) we get

$$
\begin{aligned}
(X,u)\big[\big(\mathcal{A}^V(\bar{Y},\bar{w})\big)\big] &= \frac{d}{dt}\bigg|_{t=0} \mathcal{A}^V_{(\exp t\xi g,\, v+tu)}(\bar{Y},\bar{w})(\exp t\xi, v+tu) \\
&= \frac{d}{dt}\bigg|_{t=0} \mathcal{A}^V_{(\exp t\xi g,\, v+tu)}\left(\zeta\cdot(\exp t\xi g),\ w+t\zeta u\right) \\
&= \frac{d}{dt}\bigg|_{t=0} g^{-1}\exp(-t\xi)\big(w+t\zeta u-\zeta(v+tu)\big) \\
&= -g^{-1}\xi(w-\zeta v). \qquad (4.14)
\end{aligned}
$$

and similarly

$$
(Y,w)\big[\big(\mathcal{A}^V(\bar{X},\bar{u})\big)\big] = -g^{-1}\zeta(u-\xi v). \qquad (4.15)
$$

Since

$$
\begin{aligned}
[(\bar{X},\bar{u}),(\bar{Y},\bar{w})](g,v) &= -[(\xi,\ u-\xi v),\ (\zeta,\ w-\zeta v)]\cdot(g,v) \\
&= -\big([\xi,\zeta],\ \xi w-\zeta u-[\xi,\zeta]v\big)\cdot(g,v)
\end{aligned}
$$

from (4.5) and (4.12) it follows that

$$
\begin{aligned}
\mathcal{A}^V_{(g,v)}\big([(\bar{X},\bar{u}),(\bar{Y},\bar{w})](g,v)\big) &= -\mathcal{A}^V_{(g,v)}\big(([\xi,\zeta],\ \xi w-\zeta u-[\xi,\zeta]v\big)\cdot(g,v)\big) \\
&= -\mathcal{A}^V_{(g,v)}([\xi,\zeta]\cdot g,\ \xi w-\zeta u) \\
&= -g^{-1}(\xi w-\zeta u-[\xi,\zeta]v). \qquad (4.16)
\end{aligned}
$$

Using (4.14), (4.15), and (4.16), we get

$$
\begin{aligned}
\mathbf{d}\mathcal{A}^V_{(g,v)}\big((X,u),(Y,w)\big) &= (X,u)\big[\mathcal{A}^V(\bar{Y},\bar{w})\big]-(Y,w)\big[\mathcal{A}^V(\bar{X},\bar{u})\big] \\
&\qquad -\mathcal{A}^V_{(g,v)}\big([(\bar{X},\bar{u}),(\bar{Y},\bar{w})]\big) \\
&= -g^{-1}\xi(w-\zeta v)+g^{-1}\zeta(u-\xi w) \\
&\qquad +g^{-1}(\xi w-\zeta u-[\xi,\zeta]v)=0.
\end{aligned}
$$

This proves that the curvature is zero. ■

Remark. The connection $\mathcal{A}^V$ is *not* S–invariant. In contrast, the same construction for central extensions yields an invariant but nonflat mechanical connection, as we shall see shortly.

4.2 Cotangent Bundle Structure of the Orbits

We now establish the extent to which coadjoint orbits are cotangent bundles (possibly with magnetic terms). We will illustrate the methods with SE(3).

The Structure Theorem. The strategy is to combine the reduction by stages theorem with the cotangent bundle reduction theorem. We shall recover below a result of Ratiu [1980, 1981, 1982] regarding the embedding of the semidirect product coadjoint orbits into cotangent bundles with magnetic terms but will provide a different proof based on connections. We consider here the lifted action of S on T^*S (see Theorem 3.1).

Theorem 4.2. *Let $S = G \circledS V$ be as above. Let $a \in \mathbf{J}_V(T^*S)$ and reduce T^*S by the action of V at a. There is a right G_a–equivariant symplectic diffeomorphism between*

$$(T^*S)_a = \mathbf{J}_V^{-1}(a)/V \simeq (T^*G, \Omega_a)\,, \tag{4.17}$$

where $\Omega_a = \Omega_{\mathrm{can}}$ is the canonical symplectic form. Furthermore, let $\sigma = (\mu, a) \in \mathfrak{s}^* \times V^*$ and reduce T^*S by the action of S at σ obtaining the coadjoint orbit $\mathcal{O}_\sigma$ through σ. Then there is a symplectic diffeomorphism

$$\mathcal{O}_\sigma \simeq \mathbf{J}_a^{-1}(\mu_a)/\left(G_a\right)_{\mu_a} \tag{4.18}$$

where $\mu_a = \mu|\mathfrak{g}_a$. Letting $(T^*G)_{\mu_a}$ be the reduced space for the action of G_a on T^*G, there is a symplectic embedding

$$(T^*G)_{\mu_a} \hookrightarrow \left(T^*\left(G/\left(G_a\right)_{\mu_a}\right), \Omega_{\mu_a}\right),$$

where $\Omega_{\mu_a} = \Omega_{\mathrm{can}} - \pi^*\mathcal{B}_{\mu_a}$ with $\mathcal{B}_{\mu_a}$ a closed two form on $G/\left(G_a\right)_{\mu_a}$. The image of this embedding covers the base $G/\left(G_a\right)_{\mu_a}$. This embedding is a diffeomorphism onto $T^*\left(G/\left(G_a\right)_{\mu_a}\right)$ if G_a is Abelian, in which case $G/(G_a)_{\mu_a} = G/G_a$.

Proof. The fact that the spaces in (4.17) are symplectomorphic is a consequence of the standard cotangent bundle reduction theorem for abelian symmetry groups (see Abraham and Marsden [1978] and Marsden [1992]) combined with Theorem 4.1. As these references show, the symplectomorphism is induced by the shift map

$$\mathrm{hor} : \mathbf{J}_V^{-1}(a) \to \mathbf{J}_V^{-1}(0), \qquad \mathrm{hor}\left(p_{(g,v)}\right) = p_{(g,v)} - \langle a, \mathcal{A}^V_{(g,v)}\rangle.$$

To show the equivariance it only suffices to check that (again using concatenation notation for actions)

$$\mathrm{hor}\left(p_{(g,v)} \cdot (h, 0)\right) = \left(\mathrm{hor}(p_{(g,v)})\right) \cdot (h, 0), \tag{4.19}$$

for any $h \in G_a$. However, if $(X, u) \in T_{(gh,v)}S$, formulas (4.12), (4.5), and $ha = a$ imply

$$\begin{aligned}\langle a, \mathcal{A}^V_{(g,v)(h,0)}(X,u)\rangle &= \langle a, (gh)^{-1}\big(u - (X \cdot (gh)^{-1})v\big)\rangle \\ &= \langle ha, g^{-1}\big(u - (X \cdot (gh)^{-1})v\big)\rangle \\ &= \langle a, g^{-1}\big(u - ((X \cdot h^{-1}) \cdot g^{-1})v\big)\rangle \\ &= \langle a, \mathcal{A}^V_{(g,v)}(X \cdot h^{-1}, u)\rangle \\ &= \langle a, \mathcal{A}^V_{(g,v)}\big((X,u) \cdot (h,0)^{-1}\big)\rangle,\end{aligned}$$

which proves (4.19).

The fact that the map in (4.18) is a symplectomorphism follows from Theorem 3.1 and the G_a–equivariance in (4.17). The rest of the theorem is a direct consequence of the cotangent bundle reduction theorem once we compute the μ_a–component of the curvature the mechanical connection $\mathcal{A}^{G_a}$ on $G \to G/(G_a)_{\mu_a}$. This connection is constructed analogously to $\mathcal{A}^V$ using the metric on G induced from (4.4).

The two form $\mathcal{B}_{\mu_a}$ is then obtained by dropping the exterior derivative of $\mathcal{A}^{G_a}_{\mu_a}$ to the quotient $G/\left(G_a\right)_{\mu_a}$. ■

4.3 Calculation of $\mathcal{A}^{G_a}$ and $\mathbf{d}\mathcal{A}^{G_a}$

We derive formulas for the connection and its curvature in a special case in which these formulas are particularly nice. We assume that $\mathrm{Ad}_g^T \circ \mathrm{Ad}_g$ leaves $\mathfrak{g}_a$ invariant, where $\mathrm{Ad}_g^T : \mathfrak{g} \to \mathfrak{g}$ is the transpose (adjoint) of Ad_g relative to the given metric $\langle\, , \rangle_{\mathfrak{g}}$ on $\mathfrak{g}$. This assumption holds, in particular, when $\langle\, , \rangle_{\mathfrak{g}}$ is Ad-invariant.

Theorem 4.3. *Under the preceding assumption, we have*

$$\mathcal{A}^{G_a} = \mathbb{P}_a \circ \theta^{\mathrm{L}}$$

and

$$\mathbf{d}\mathcal{A}^{G_a}(g)(X_g, Y_g) = -\mathbb{P}_a\left([T_g L_{g^{-1}} X_g, T_g L_{g^{-1}} Y_g]\right),$$

where θ^{L} is the left-invariant Maurer-Cartan form on G (given by $\theta^{\mathrm{L}}(X_g) = T_g L_{g^{-1}} X_g$) and $\mathbb{P}_a : \mathfrak{g} \to \mathfrak{g}_a$ is orthogonal projection relative to the metric $\langle\, , \rangle_{\mathfrak{g}}$.

Proof. We first compute the locked inertia tensor for the right action of G_a on G. Let $\langle\, , \rangle_g$ denote the right invariant extension of the inner product $\langle\, , \rangle_{\mathfrak{g}}$ to an inner product on T_gG, so that $\langle\, , \rangle_e = \langle\, , \rangle_{\mathfrak{g}}$ and let $\xi, \eta \in \mathfrak{g}_a$. By definition, the locked inertia tensor is given by

$$\begin{aligned}\mathbb{I}_g(\xi)(\eta) &= \langle \xi_G(g), \eta_G(g)\rangle_g \\ &= \langle T_eL_g\xi, T_eL_g\eta\rangle_g \\ &= \langle \mathrm{Ad}_g\,\xi, \mathrm{Ad}_g\,\eta\rangle_e \\ &= \langle \mathrm{Ad}_g^T \circ \mathrm{Ad}_g\,\xi, \eta\rangle_e \\ &= \langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(\xi), \eta\rangle_e.\end{aligned}$$

Thus,

$$\mathbb{I}_g(\xi) = \langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(\xi), \cdot\rangle_e \in \mathfrak{g}_a^*. \tag{4.20}$$

Next, we compute $\mathbf{J}(\langle X_g, \cdot\rangle_g) \in \mathfrak{g}_a^*$:

$$\begin{aligned}\left\langle \mathbf{J}(\langle X_g, \cdot\rangle_g), \xi\right\rangle &= \langle X_g, \xi_G(g)\rangle = \langle X_g, g\cdot\xi\rangle_g \\ &= \langle X_g\cdot g^{-1}, \mathrm{Ad}_g\,\xi\rangle_e \\ &= \langle \mathrm{Ad}_g^T\, X_g\cdot g^{-1}, \xi\rangle_e \\ &= \left\langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(g^{-1}\cdot X_g), \xi\right\rangle_e.\end{aligned}$$

We conclude that

$$\mathbf{J}(\langle X_g, \cdot\rangle_g) = \left\langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(g^{-1}\cdot X_g), \cdot\right\rangle_e. \tag{4.21}$$

By hypothesis, $\mathrm{Ad}_g^T \circ \mathrm{Ad}_g$ leaves $\mathfrak{g}_a$ invariant, and since it is symmetric, it also leaves its orthogonal complement invariant and so it commutes with the orthogonal projection $\mathbb{P}_a$. Thus, we get

$$\mathbf{J}(\langle X_g, \cdot\rangle_g) = \left\langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(\mathbb{P}_a(g^{-1}\cdot X_g)), \cdot\right\rangle_e. \tag{4.22}$$

Therefore, combining (4.20) and (4.22), we get

$$\begin{aligned}\mathcal{A}^{G_a}(g)(X_g) &= (\mathbb{I}_g^{-1} \circ \mathbf{J})(\langle X_g, \cdot\rangle_g) \\ &= (\mathbb{P}_a \circ \theta^{\mathrm{L}})(X_g).\end{aligned}$$

To compute $\mathbf{d}\mathcal{A}^{G_a}(g)(X_g, Y_g)$ extend X_g, Y_g to left invariant vector fields $\bar{X}, \bar{Y}$. Then,

$$\begin{aligned}\mathbf{d}\mathcal{A}^{G_a}(g)(X_g, Y_g) &= X_g[\mathcal{A}^{G_a}(\bar{Y})] - Y_g[\mathcal{A}^{G_a}(\bar{X})] - \mathcal{A}^{G_a}(g)([\bar{X}, \bar{Y}]) \\ &= -\mathbb{P}_a([g^{-1}\cdot X_g, g^{-1}\cdot Y_g]),\end{aligned} \tag{4.23}$$

since the first two terms vanish. ■

4.4 Example: SE(3)

Following the earlier results of this section, we will use reduction by stages theory in the context of cotangent bundle reduction to classify the coadjoint orbits of SE(3). We will also make use of mechanical connections and their curvatures to compute the the coadjoint orbit symplectic forms.

We begin by defining a right invariant metric on SE(3). Identify

$$\mathfrak{se}(3) \simeq \mathfrak{so}(3) \oplus \mathbb{R}^3$$

and define the natural inner product at the identity (see (4.3))

$$\langle (X, a), (Y, b) \rangle_{e,o} = -\frac{1}{2}\mathrm{tr}(XY) + (a, b),$$

where $(\cdot, \cdot)$ denotes the Euclidean inner product. Right invariance of the metric and (4.13) gives

$$\begin{aligned}
&\langle (X_A, a_A), (Y_A, b_A) \rangle_{(A,\alpha)} \\
&= \langle (X_A \cdot A^{-1}, a_A - (X_A \cdot A^{-1})\alpha),\ (Y_A \cdot A^{-1}, b_A - (Y_A \cdot A^{-1})\alpha) \rangle_{(I,0)} \\
&= -\frac{1}{2}\mathrm{tr}(X_A \cdot A^{-1} \cdot Y_A \cdot A^{-1}) + ((X_A \cdot A^{-1})\alpha, (Y_A \cdot A^{-1})\alpha) \\
&\quad - ((X_A \cdot A^{-1})\alpha, b_A) - ((Y_A \cdot A^{-1})\alpha, a_A) + (a_A, b_A). \qquad (4.24)
\end{aligned}$$

The mechanical connection for the principal $\mathbb{R}^3$ bundle $\mathrm{SE}(3) \to \mathrm{SO}(3)$, is given by (4.12):

$$\mathcal{A}^V(A, \alpha)(X_A, a_A) = A^{-1}\left(a_A - (X_A \cdot A^{-1})\alpha\right)$$

and from Theorem 4.1, $\mathrm{curv}\mathcal{A}^V = \mathbf{d}\mathcal{A}^V = 0$.

We first reduce by the $\mathbb{R}^3$ cotangent lifted action. Let $a \in \mathbb{R}^{3*}$. By the cotangent bundle reduction theorem, we have $\mathbf{J}_{\mathbb{R}^3}^{-1}(a)/\mathbb{R}^3_a = T^*(\mathrm{SE}(3)/\mathbb{R}^3) = T^*(\mathrm{SO}(3))$ with symplectic form $\Omega_a = \Omega_{\mathrm{can}}$. Next, suppose $a = 0$. Then $G_a = \mathrm{SO}(3)$. Reduction by the SO(3) action gives now coadjoint orbits in SO(3). Thus $\mathcal{O}_{(a=0,\mu)} = S^2_\mu$.

Assume $a \neq 0$. Depending on whether $\mu = 0$ or $\mu \neq 0$, we will now consider two cases. Suppose $\mu = 0$, then the group $\mathrm{SE}(3)_a/\mathbb{R}^3_a \simeq \mathrm{SO}(3)_a \simeq S^1$ now acts on the first reduced space. Note that $\mathrm{SO}(3)/\mathrm{SO}(3)_a \simeq S^2_a$ is the sphere through $a \in \mathbb{R}^3$. Thus, reducing by this $\mathrm{SO}(3)_a$ action at $\mu_a = 0$ gives, by another application of the cotangent bundle reduction theorem for Abelian groups, $(T^*S^2_a, \Omega_{\mathrm{can}})$.

Next consider the case $\mu \neq 0$. Thus, after the first reduction, by the $\mathbb{R}^3$ action at the point $a \neq 0$, we have $P_a = (T^*\mathrm{SO}(3), \Omega_{\mathrm{can}})$ as before. Next, we form the group $G_a = \mathrm{SO}(3)_a$ which acts by cotangent lift on $T^*\mathrm{SO}(3)$. Now, consider the bundle $\mathrm{SO}(3) \to \mathrm{SO}(3)/\mathrm{SO}(3)_a \simeq S^2_a$. This S^1 bundle inherits a metric from the bundle $\mathrm{SE}(3) \to \mathrm{SO}(3)$, which is SO(3) invariant.

Let us now investigate the connection and curvature on this bundle. It is convenient to use the Lie algebra isomorphism $x \mapsto \widehat{x} : \mathbb{R}^3 \to \mathfrak{so}(3)$ defined by the cross product: $\widehat{x}u = x \times u$. We then define the metric

$$\langle \widehat{x}, \widehat{y} \rangle_I = (x, y) = -\frac{1}{2}\mathrm{tr}(\widehat{x}\,\widehat{y})$$

and for $X_A, Y_A \in T_A\mathrm{SO}(3)$,

$$\langle X_A, Y_A \rangle_A = -\frac{1}{2}\mathrm{tr}(X_A \cdot A^{-1} \cdot Y_A \cdot A^{-1}).$$

The Lie algebra of $\mathrm{SO}(3)_a \simeq S^1$ is $\mathrm{span}\{a\}$, so for $u, v \in \mathbb{R}$, we have

$$u_{\mathrm{SO}(3)}(A) = \left.\frac{d}{dt}\right|_{t=0} A\exp(tu\widehat{a}) = A\widehat{ua}, \tag{4.25}$$

so that by right invariance and the identity $A\widehat{a}A^{-1} = \widehat{Aa}$ we get

$$\begin{aligned} \langle \mathbb{I}(A)ua, va \rangle &= \langle A\widehat{ua},\ A\widehat{va} \rangle_A \\ &= uv\langle A\widehat{a}, A\widehat{a} \rangle_A \\ &= uv\langle A\widehat{a}A^{-1}, A\widehat{a}A^{-1} \rangle_I \\ &= uv\langle \widehat{Aa}, \widehat{Aa} \rangle_I = (ua, va). \end{aligned} \tag{4.26}$$

We conclude that $\mathbb{I}(A)(ua) = ua$ and that $\mathbb{I}^{-1}(A)(va) = va$. Next, taking $u \in \mathbb{R}$, the $\mathrm{SO}(3)_a$ momentum map $\mathbf{J} : T^*\mathrm{SO}(3) \to \mathrm{span}\{a\} \cong \mathbb{R}$ is given by

$$\begin{aligned} \mathbf{J}\left(\langle X_A, \cdot \rangle_A\right)(\widehat{ua}) &= \left\langle X_A, u_{\mathrm{SO}(3)}(A) \right\rangle_A \\ &= \langle X_A, A\widehat{ua} \rangle_A \\ &= \left\langle X_A \cdot A^{-1}, A\widehat{a}uA^{-1} \right\rangle_I \\ &= \left\langle \mathrm{Ad}_{A^{-1}}(X_A \cdot A^{-1}), \widehat{a} \right\rangle_I u \\ &= \langle A^{-1} \cdot X_A, \widehat{ua} \rangle_I, \end{aligned}$$

so that

$$\mathbf{J}\left(\langle X_A, \cdot \rangle_A\right) = \frac{1}{\|a\|^2}\langle A^{-1} \cdot X_A, \widehat{a} \rangle_I \widehat{a}. \tag{4.27}$$

Therefore,

$$\mathcal{A}^{G_a}(A)(X_A) := \left(\mathbb{I}_A^{-1} \circ \mathbf{J}\right)\left(\langle X_A, \cdot \rangle_A\right) = \frac{1}{\|a\|^2}\langle A^{-1} \cdot X_A, \widehat{a} \rangle_I \widehat{a}. \tag{4.28}$$

To find the curvature of $\mathcal{A}^{G_a}$, we compute $\mathbf{d}\mathcal{A}^{G_a}(A)(X_A, Y_A)$. Let $X_A = \widehat{x} \cdot A, Y_A = \widehat{y} \cdot A \in T_A\mathrm{SO}(3)$. Denote by $\bar{X}, \bar{Y}$ the right invariant vector fields whose values at I are $\widehat{x}$ and $\widehat{y}$ respectively.

Then (4.28) and the identities $A^{-1}\widehat{x}A = \widehat{A^{-1}x}$, $\langle \widehat{x}, \widehat{y} \rangle_I = (x, y)$, imply

$$
\begin{aligned}
Y_A[\mathcal{A}^{G_a}(\bar{X})] &= \left.\frac{d}{dt}\right|_{t=0} \mathcal{A}^{G_a}(\bar{X})((\exp t\widehat{y})A) \cdot \widehat{a} \\
&= \left.\frac{d}{dt}\right|_{t=0} \mathcal{A}^{G_a}((\exp t\widehat{y})A)(\widehat{x} \cdot (\exp t\widehat{y})A) \cdot \widehat{a} \\
&= \left.\frac{d}{dt}\right|_{t=0} \frac{1}{\|a\|^2} \langle A^{-1} \exp(-t\widehat{y})\widehat{x}(\exp t\widehat{y})A,\ \widehat{a} \rangle_I \cdot \widehat{a} \\
&= \left.\frac{d}{dt}\right|_{t=0} \frac{1}{\|a\|^2} \big(A^{-1} \exp(-t\widehat{y})x, a \big) \cdot \widehat{a} \\
&= \frac{1}{\|a\|^2} \big(A^{-1}(x \times y), a \big) \cdot \widehat{a}. \qquad (4.29)
\end{aligned}
$$

Similarly

$$
X_A[\mathcal{A}^{G_a}(\bar{Y})] = -\frac{1}{\|a\|^2} \big(A^{-1}(x \times y), a \big) \widehat{a}. \qquad (4.30)
$$

Finally,

$$
\begin{aligned}
\mathcal{A}^{G_a}([\bar{X}, \bar{Y}])(A) &= \mathcal{A}^{G_a}(A)(-[\widehat{x}, \widehat{y}] \cdot A)\widehat{a} \\
&= -\frac{1}{\|a\|^2} \langle A^{-1}[\widehat{x}, \widehat{y}]A, \widehat{a} \rangle_I \widehat{a} \\
&= -\frac{1}{\|a\|^2} \big(A^{-1}(x \times y), a \big) \widehat{a} \qquad (4.31)
\end{aligned}
$$

so that (4.29), (4.30), and (4.31)

$$
\begin{aligned}
\mathbf{d}\mathcal{A}^{G_a}(A)(X_A, Y_A) &= X_A[\mathcal{A}^{G_a}(\bar{Y})] - Y_A[\mathcal{A}^{G_a}(\bar{X})] - \mathcal{A}^{G_a}([\bar{X}, \bar{Y}])(A) \\
&= -\frac{1}{\|a\|^2} \big(A^{-1}(x \times y), a \big) \widehat{a}, \qquad (4.32)
\end{aligned}
$$

where $X_A = \widehat{x} \cdot A$, $Y_A = \widehat{y} \cdot A \in T_A\mathrm{SO}(3)$. Note that this equation agrees with the result of Theorem 4.3. This two–form on SO(3) clearly induces a two form $\mathcal{B}$ on the sphere S_a^2 of radius $\|a\|$ by

$$
\mathcal{B}(Aa)(x \times Aa, y \times Aa) = -\frac{1}{\|a\|^2} \big(A^{-1}(x \times y), a \big) \widehat{a}. \qquad (4.33)
$$

Given $\mu \in \mathfrak{so}(3)^* \cong \mathbb{R}^3$, its restriction to $\mathfrak{so}(3)_a \cong \mathbb{R}$ equals (μ, a). Therefore, in this case $\mathcal{B}_{\mu_a} = (\mu, a)\mathcal{B}$, that is,

$$
\mathcal{B}_{\mu_a}(x \times Aa, y \times Aa) = -\frac{(\mu, a)}{\|a\|^2} \big(A^{-1}(x \times y), a \big). \qquad (4.34)
$$

Invoking the cotangent bundle reduction theorem we can conclude:

Theorem 4.4. *The coadjoint orbits of* SE(3) *are of the following types.*

- $O_{(a=0,\mu)} \simeq (S^2_\mu, \omega_\mu)$
- $O_{(a\neq 0,\mu=0)} \simeq (T^*S^2_a, \Omega_{\mathrm{can}})$
- $O_{(a\neq 0,\mu\neq 0)} \simeq (T^*S^2_a, \Omega_{\mathrm{can}} - \pi^*\mathcal{B}_{\mu_a})$

where ω_μ is the orbit symplectic form on the sphere S^2_μ of radius $\|\mu\|$, μ_a is the orthogonal projection of μ to $\mathrm{span}\{a\}$, *$\pi : T^*S^2_a \to S^2_a$ is the cotangent bundle projection, Ω_{can} is the canonical symplectic structure on $T^*S^2_a$, and the two–form $\mathcal{B}_{\mu_a}$ on the sphere S^2_a of radius $\|a\|$ is given by formula* (4.34).

5 Reduction by Stages for Group Extensions

We now embark on extending the preceding theory to the case of arbitrary group extensions. Some results in this general direction are those of Landsman [1995], Sjamaar and Lerman [1991], and Ziegler [1996]. The results in Landsman [1995] make many interesting links with quantization. Duval, Elhadad, Gotay, Śniatycki and Tuynman [1991] give a nice interpretation of semidirect products in the context of BRST theory and quantization and apply it to the pseudo rigid body. The results in Sjamaar and Lerman [1991] deal with general extensions, but only at zero levels of the momentum map and only for compact groups. Unfortunately this does not cover the case of semidirect products and their proofs do not seem to generalize, so it overlaps very little with the work here. Ziegler [1996] (see also Baguis [1998]) makes a lot of nice links with the orbit method and symplectic versions of Mackey's induced representations, amongst other things.

5.1 The Heisenberg Group

To motivate the more general theory for group extensions, we first consider one of the basic examples, namely the Heisenberg group (see for example, Guillemin and Sternberg [1984]). While this example is quite simple, it illustrates nicely what some of the issues are in the general theory that were not encountered in the theory for semidirect products. Thus, we present this example in a direct way before presenting the general construction.

Definitions and Cocycles. We start with the commutative group $\mathbb{R}^2$ with its standard symplectic form ω, the usual area form on the plane. It is easy to show that ω satisfies the group cocycle law so that we can centrally extend $\mathbb{R}^2$ by $\mathbb{R}$ to form the group $H = \mathbb{R}^2 \oplus \mathbb{R}$ with multiplication

$$(u, \alpha)(v, \beta) = (u + v, \alpha + \beta + \omega(u, v)). \tag{5.1}$$

The identity element is $(0, 0)$ and the inverse of (u, α) is given by $(u, \alpha)^{-1} = (-u, -\alpha)$. We compute the corresponding Lie algebra cocycle, C, from the Lie algebra bracket as follows. Conjugation is given by

$$(u, \alpha)(v, \beta)(-u, -\alpha) = (v, \beta + 2\omega(u, v)).$$

Thus,

$$\mathrm{Ad}_{(u,\alpha)}(Y, b) = (Y, b + 2\omega(u, Y))$$

and

$$[(X, a), (Y, b)] = (0, 2\omega(X, Y)),$$

where $(X, a), (Y, b) \in \mathfrak{h} = \mathbb{R}^2 \oplus \mathbb{R}$. Using this, we read off the Lie algebra cocycle,

$$C(X, Y) = 2\omega(X, Y).$$

Coadjoint Orbits. Identify $\mathfrak{h}^*$ with $\mathbb{R}^3$ via the Euclidean inner product. The previous formulas imply then

$$\mathrm{Ad}^*{}_{(u,\alpha)^{-1}}(\mu, \nu) = (\mu + 2\nu \mathbb{J}u, \nu),$$

where $\mu, u \in \mathbb{R}^2$, $\alpha, \nu \in \mathbb{R}$, and $\mathbb{J}(u_1, u_2) = (u_2, -u_1)$ is the matrix of the standard symplectic structure on $\mathbb{R}^2$. Therefore, the coadjoint orbits of the Heisenberg group are:

- $\mathcal{O}_{(\mu,0)} = \{(\mu, 0)\}$
- $\mathcal{O}_{(\mu,\nu\neq 0)} = \mathbb{R}^2 \times \{\nu\}$.

The Mechanical Connection. Next, consider the right principal $\mathbb{R}$ bundle $H \to \mathbb{R}^2$. Following the exposition in the case of semidirect products, we construct a right H invariant metric on H from which we can derive a mechanical connection on the $\mathbb{R}$ bundle. Set

$$\langle (X, a), (Y, b) \rangle = (X, Y) + ab, \tag{5.2}$$

for $(X, a), (Y, b) \in \mathfrak{h}$ and where the Euclidean inner product, $(\cdot, \cdot)$ in $\mathbb{R}^2$ is used in the first summand. If $(X_{(u,\alpha)}, a_{(u,\alpha)}) \in T_{(u,\alpha)}H$, then

$$T_{(u,\alpha)}R_{(v,\beta)}(X_{(u,\alpha)}, a_{(u,\alpha)}) = (X_{(u,\alpha)}, a_{(u,\alpha)} + \omega(X_{(u,\alpha)}, v)) \in T_{(u,\alpha)(v,\beta)}H$$

and, in particular,

$$T_{(u,\alpha)}R_{(u,\alpha)^{-1}}(X_{(u,\alpha)}, a_{(u,\alpha)}) = \big(X_{(u,\alpha)}, a_{(u,\alpha)} - \omega(X_{(u,\alpha)}, u)\big) \in \mathfrak{h}.$$

Thus, the right invariant metric on H is given by

$$\begin{aligned}\big\langle(X_{(u,\alpha)}, a_{(u,\alpha)}),\ (Y_{(u,\alpha)}, b_{(u,\alpha)})\big\rangle_{(u,\alpha)} &= (X_{(u,\alpha)}, Y_{(u,\alpha)}) + a_{(u,\alpha)}b_{(u,\alpha)}\\ &\quad - a_{(u,\alpha)}\omega(Y_{(u,\alpha)}, u) - b_{(u,\alpha)}\omega(X_{(u,\alpha)}, u) + \omega(X_{(u,\alpha)}, u)\omega(Y_{(u,\alpha)}, u).\end{aligned}$$

Given $a \in \mathbb{R}$, the infinitesimal generator for the right $\mathbb{R}$ action on H is

$$a_H(v,\alpha) = \left.\frac{d}{dt}\right|_{t=0}(v,\alpha)(0,ta) = (0,a).$$

Combining these formulas yields the expression of the associated locked inertia tensor:

$$\mathbb{I}_{(v,\alpha)}(a)(b) = \langle a_H(v,\alpha), b_H(v,\alpha)\rangle_{(v,\alpha)} = ab \tag{5.3}$$

For any $(X_{(u,\alpha)}, a_{(u,\alpha)}) \in T_{(u,\alpha)}H$ and $b \in \mathbb{R}$, we get

$$\begin{aligned}\big\langle \mathbf{J}_{\mathbb{R}}\left(\langle(X_{(u,\alpha)}, a_{(u,\alpha)}), \cdot\rangle_{(u,\alpha)}\right), b\big\rangle &= \big\langle(X_{(u,\alpha)}, a_{(u,\alpha)}), (0,b)\big\rangle_{(u,\alpha)}\\ &= \big(a_{(u,\alpha)} - \omega(X_{(u,\alpha)}, u)\big)\, b.\end{aligned} \tag{5.4}$$

Thus, the mechanical connection has the expression

$$\mathcal{A}(u,\alpha)(X_{(u,\alpha)}, a_{(u,\alpha)}) = a_{(u,\alpha)} - \omega(X_{(u,\alpha)}, u). \tag{5.5}$$

Proceeding as in the previous section, an easy calculation shows that

$$\mathbf{d}\mathcal{A}(u,\alpha)((X_{(u,\alpha)}, a_{(u,\alpha)}),\ (Y_{(u,\alpha)}, b_{(u,\alpha)})) = 2\omega(X_{(u,\alpha)}, Y_{(u,\alpha)}). \tag{5.6}$$

This two–form induces a closed two–form $\mathcal{B}$, the curvature form, on the quotient $H/\mathbb{R} \simeq \mathbb{R}^2$ by

$$\mathcal{B}(u)(X,Y) = 2\omega(X,Y),$$

for $u, X, Y \in \mathbb{R}^2$.

The First Reduced Space. Reducing T^*H by the central $\mathbb{R}$ action at a point $\nu \in \mathbb{R}^* \cong \mathbb{R}$ gives

$$\mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R} \simeq (T^*\mathbb{R}^2, \Omega - \pi^*\mathcal{B}_\nu),$$

where $\simeq$ is a symplectic diffeomorphism equivariant with respect to the remaining $\mathbb{R}^2$ action and $\mathcal{B}_\nu = \nu\mathcal{B}$, that is,

$$\mathcal{B}_\nu(X,Y) = 2\nu\omega(X,Y), \tag{5.7}$$

for $u, X, Y \in \mathbb{R}^2$.

We now have a first reduced space and an action on it by cotangent lift. It remains to compute the reduced spaces for this action. To do this we will need to use non–equivariant reduction, since *the momentum map for this remaining action is no longer equivariant*. Equivariance is lost precisely because of the presence of the magnetic term in the first reduced space. *This lack of equivariance is the first major difference with the semidirect product case.*

Calculation of the Momentum Map. Given a cotangent lifted action of G on $(T^*Q, \Omega_{\rm can} - \pi^* B)$ where B is a closed two form on Q, suppose that there is a linear map $\xi \mapsto \varphi^\xi$ from $\mathfrak{g}$ to functions on Q such that for all $\xi \in \mathfrak{g}$, we have $\mathbf{i}_{\xi_Q} B = \mathbf{d}\varphi^\xi$ (where $\mathbf{i}_\xi$ denotes the interior product). In these circumstances, a momentum map is verified to be $\mathbf{J} = \mathbf{J}_{\rm can} - \pi_Q^* \varphi$, where $\pi_Q : T^*Q \to Q$ is cotangent bundle projection.

In the case of the Heisenberg group, let ϕ_ν be the momentum map for the translation action of $\mathbb{R}^2$ on $(\mathbb{R}^2, 2\nu\omega)$, that is,

$$\phi_\nu(x, y) = 2\nu(y, -x) \tag{5.8}$$

where we identified $\mathbb{R}^{2*}$ and $\mathbb{R}^2$ by means of the Euclidean inner product. Denote coordinates on $T^*\mathbb{R}^2$ by (x, y, p_x, p_y). Again using the above identification of $\mathbb{R}^2$ with its dual, the canonical momentum is given by

$$\mathbf{J}_{\rm can}(x, y, p_x, p_y) = (p_x, p_y) \tag{5.9}$$

and hence the momentum map of the lifted $\mathbb{R}^2$ action on the first reduced space $(T^*\mathbb{R}^2, \, \Omega_{\rm can} - 2\nu\pi^*_{\mathbb{R}^2}\omega)$ is the map $\mathbf{J}_\nu : T^*\mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathbf{J}_\nu(x, y, p_x, p_y) = (p_x - 2\nu y, p_y + 2\nu x). \tag{5.10}$$

This formula shows that for $\rho = (\rho_1, \rho_2) \in \mathbb{R}^2$,

$$\mathbf{J}_\nu^{-1}(\rho) = \{(x, y, p_x, p_y) \mid p_x = 2\nu y + \rho_1, p_y = -2\nu x + \rho_2\}. \tag{5.11}$$

Clearly the map $(x, y) \in \mathbb{R}^2 \mapsto (x, y, p_x, p_y) \in \mathbf{J}_\nu^{-1}(\rho)$ defined by these equations for p_x, p_y is a diffeomorphism. The level sets of $\mathbf{J}_\nu$ are therefore two–dimensional planes. We will next compute the subgroup that we quotient these sets by to complete the second stage reduction.

It is a well known result of Souriau that one can modify the action with a cocycle so that the momentum map becomes equivariant relative to this new affine action on the dual of the Lie algebra (see, e.g., Marsden and Ratiu [1994]). This affine G action on $\mathfrak{g}^*$ with respect to which the momentum map becomes equivariant is given by

$$g \cdot \mu = \mathrm{Ad}^*_{g^{-1}} \mu + \sigma(g) \tag{5.12}$$

where $\sigma(g)$ is the group 1–cocycle associated with the non-equivariance of the momentum map.

Lemma 5.1. *Let $\mu \in \mathbb{R}^2$. The isotropy subgroup for μ, using this affine action, is $(0,0)$ if $\nu \neq 0$ and is $\mathbb{R}^2$ if $\nu = 0$.*

Proof. Recalling the definition of the group one–cocycle $\sigma^\nu : \mathbb{R}^2 \to \mathbb{R}^{2*} \cong \mathbb{R}^2$ (see Marsden and Ratiu [1994], Section 12.4) and using the fact that the coadjoint action of $\mathbb{R}^2$ is trivial (since $\mathbb{R}^2$ is Abelian), we get

$$\begin{aligned}
&\big(\sigma^\nu(a,b),(\xi_1,\xi_2)\big)\\
&\quad= \langle \mathbf{J}_\nu((a,b)(x,y,p_x,p_y)),(\xi_1,\xi_2)\rangle - \langle \mathrm{Ad}^*_{(a,b)^{-1}}\mathbf{J}_\nu(x,y,p_x,p_y),(\xi_1,\xi_2)\rangle\\
&\quad= \langle \mathbf{J}_\nu(a+x,b+y,p_x,p_y),(\xi_1,\xi_2)\rangle - \langle \mathbf{J}_\nu(x,y,p_x,p_y),(\xi_1,\xi_2)\rangle\\
&\quad= \langle\big(p_x - 2\nu(b+y), p_y + 2\nu(a+x)\big),(\xi_1,\xi_2)\rangle - \langle\big(p_x - 2\nu y, p_y + 2\nu x\big),(\xi_1,\xi_2)\rangle\\
&\quad= 2\nu\big((-b,a),(\xi_1,\xi_2)\big),
\end{aligned}$$

that is,

$$\sigma^\nu(a,b) = 2\nu(-b,a). \tag{5.13}$$

Therefore, the affine action is

$$(a,b)\cdot(\rho_1,\rho_2) = (\rho_1,\rho_2) + 2\nu(-b,a) \tag{5.14}$$

and the isotropy of (ρ_1,ρ_2) consists of all $(a,b) \in \mathbb{R}^2$ such that $(\rho_1,\rho_2) = (\rho_1,\rho_2) + 2\nu(-b,a)$, from which the conclusion follows trivially. ■

The Second Reduced Space. First consider the case $\nu = 0$. The first reduced space is $(T^*\mathbb{R}^2, \Omega_{\mathrm{can}})$. Thus, reduction by the remaining $\mathbb{R}^2$ action gives single points.

For the case $\nu \neq 0$ reduction at any point ρ is a plane since we quotient the set

$$\mathbf{J}_\nu^{-1}(\rho) = \{(x,y,p_x,p_y) \mid p_x = 2\nu y + \rho_1, p_y = -2\nu x + \rho_2\}.$$

by the identity. We next calculate the reduced symplectic forms on these planes. This is done by restricting the symplectic form on $T^*\mathbb{R}^2$ to the level sets of $\mathbf{J}_\nu$.

Proposition 5.2. *The coadjoint orbit symplectic form for the orbit through the point $(\nu,(\rho_1,\rho_2))$ is given by $\omega_{O_{\nu,\rho}}(x,y)(X,Y) = 2\nu\omega(X,Y)$.*

Proof. Let (x,y) be coordinates of the coadjoint orbit through $(\nu,(\rho_1,\rho_2))$. The embedding of the plane $\mathbf{J}_\nu^{-1}(\rho)$ into $T^*\mathbb{R}^2$ is given by

$$\psi : (x,y) \mapsto (x, y, 2\nu y + \rho_1, -2\nu x + \rho_2).$$

We then have

$$
\begin{aligned}
&\omega_{\mathrm{red}}(x,y)(X,Y)\\
&= \Omega_{\mathrm{can}}(\psi(x,y))\left(X_1\frac{\partial}{\partial x}+X_2\frac{\partial}{\partial y}+2\nu X_2\frac{\partial}{\partial p_x}-2\nu X_1\frac{\partial}{\partial p_y},\right.\\
&\quad\left. Y_1\frac{\partial}{\partial x}+Y_2\frac{\partial}{\partial y}+2\nu Y_2\frac{\partial}{\partial p_x}-2\nu Y_1\frac{\partial}{\partial p_y}\right)-\psi^*\pi^* d\mathcal{A}_\nu(x,y)(X,Y)\\
&= X_1\cdot 2\nu Y_2 - Y_1\cdot 2\nu X_2 + X_2\cdot(-2\nu Y_1) - Y_2\cdot(-2\nu X_1)\\
&\quad -\psi^*\pi^* d\mathcal{A}_\nu(x,y)(X,Y)\\
&= 4\nu\omega(x,y)(X,Y) - 2\nu\omega(x,y)(X,Y)\\
&= 2\nu\omega(x,y)(X,Y), \qquad (5.15)
\end{aligned}
$$

where we have used the fact that $\pi\circ\psi = \mathrm{id}$. ■

Although one can check it directly in this case, the fact that the reduction by stages procedure gives the coadjoint orbits of the Heisenberg group is a consequence of the general theory of the next subsection.

5.2 The Main Steps in the General Construction

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

Assume that N is a normal subgroup of M and denote its Lie algebra by $\mathfrak{n}$. Let $i : \mathfrak{n} \to \mathfrak{m}$ denote the inclusion and let $i^* : \mathfrak{m}^* \to \mathfrak{n}^*$ be its dual, which is the natural projection given by restriction of linear functionals. The equivariant momentum map for the action of the group N on P is given by

$$\mathbf{J}_N(z) = i^*(\mathbf{J}_M(z)) \qquad (5.16)$$

as is well known and easily verified. Let $\nu \in \mathfrak{n}^*$ be a regular value of $\mathbf{J}_N$ and let N_ν be the isotropy subgroup of ν for the coadjoint action of N on its Lie algebra. We suppose that the action of N_ν (and in fact that of M) is free and proper and form the ***first symplectic reduced space***:

$$P_\nu = \mathbf{J}_N^{-1}(\nu)/N_\nu.$$

Since N is a normal subgroup, the adjoint action of M on its Lie algebra $\mathfrak{m}$ leaves the subalgebra $\mathfrak{n}$ invariant, and so it induces a dual action of M on $\mathfrak{n}^*$. By construction, the inclusion map $i : \mathfrak{n} \to \mathfrak{m}$ is equivariant with respect to the action of M on the domain and range. Thus, the dual $i^* : \mathfrak{m}^* \to \mathfrak{n}^*$ is equivariant with respect to the dual action of M.

Because N is a subgroup of M, the adjoint action of N on $\mathfrak{n}$ coincides with the restriction of the action of M on $\mathfrak{n}$ to the subgroup N. Dualizing this, we obtain:

Lemma 5.3. *The restriction of the action of M on $\mathfrak{n}^*$ to the subgroup N coincides with the coadjoint action of N on $\mathfrak{n}^*$.*

Let M_ν denote the isotropy subgroup of $\nu \in \mathfrak{n}^*$ for the action of M on $\mathfrak{n}^*$. It follows from the preceding lemma and normality of N in M that

$$N_\nu = M_\nu \cap N. \tag{5.17}$$

To see that $M_\nu \cap N \subset N_\nu$, let $n \in M_\nu \cap N$ so that, regarded as an element of M, it fixes ν. But since the action of N on $\mathfrak{n}^*$ induced by the action of M on $\mathfrak{n}^*$ coincides with the coadjoint action by the above lemma, this means that n fixes ν using the coadjoint action. The other inclusion is obvious.

Caution. In the case of semidirect products, where we can regard $\mathfrak{n}^*$ as a subspace of $\mathfrak{m}^*$, *the action of a group element $n \in N$ regarded as an element of M on the space $\mathfrak{m}^*$ need not leave the subspace $\mathfrak{n}^*$ invariant.* That is, its coadjoint action regarded as an element of M need not restrict to the coadjoint action regarded as an element of N. Rather than *restricting*, one must *project* the actions using the map i^*, as we have described. Thus, one has to be careful about the space in which one is computing the isotropy of an element ν.

Induced Actions of Quotient Groups. It is an elementary fact that the intersection of a normal subgroup N with another subgroup is normal in that subgroup, so we get:

Lemma 5.4. *The subgroup $N_\nu \subset M$ is normal in M_ν.*

Thus, we can form the quotient group M_ν/N_ν. In the context of semidirect products, with the second factor being a vector space V, M_ν/N_ν reduces to G_a where we have written $\nu = a$, as before. However, if the second factor is nonabelian, M_ν/N_ν *need not be* G_ν. (It is another group G^ν that is computed in Marsden, Misiołek, Perlmutter and Ratiu [1998]).

Lemma 5.5. *There is a well defined symplectic action of M_ν/N_ν on the reduced space P_ν. This action will be denoted Ψ_ν.*

Proof. First of all, using equivariance of $\mathbf{J}_M$ and i^*, we note that the action of M_ν on P leaves the set $\mathbf{J}_N^{-1}(\nu)$ invariant. The action of a group element $m \in M_\nu$ on this space will be denoted $\Phi^\nu_m : \mathbf{J}_N^{-1}(\nu) \to \mathbf{J}_N^{-1}(\nu)$.

It is a general fact that when a group K acts on a manifold Q, and $L \subset K$ is a normal subgroup, then the quotient group K/L acts on the quotient space Q/L. If $\pi_L : Q \to Q/L$ denotes the projection, $\Psi_k : Q \to Q$ denotes the given action of a group element $k \in K$, and $\Psi^L_{[k]} : Q/L \to Q/L$ denotes the quotient action of an element $[k] \in K/L$, then we have

$$\Psi^L_{[k]} \circ \pi_L = \pi_L \circ \Psi_k,$$

that is, the projection onto the quotient is equivariant with respect to the two actions via the group projection.

These general considerations show that the group M_ν/N_ν has a well defined action on the space P_ν. The action of a group element $[m] \in M_\nu/N_\nu$ will be denoted by $\Psi_{[m],\nu} : P_\nu \to P_\nu$. We shall now show that this action is symplectic.

Let $\pi_\nu : \mathbf{J}_N^{-1}(\nu) \to P_\nu$ denote the natural projection and $i_\nu : \mathbf{J}_N^{-1}(\nu) \to P$ be the inclusion. By the equivariance of the projection, we have,

$$\Psi_{[m],\nu} \circ \pi_\nu = \pi_\nu \circ \Phi^\nu_m,$$

for all $m \in M_\nu$. Since the action Φ^ν is the restriction of the action Φ of M, we get

$$\Phi_m \circ i_\nu = i_\nu \circ \Phi^\nu_m$$

for each $m \in M_\nu$.

Recall from the reduction theorem that $i^*_\nu \Omega = \pi^*_\nu \Omega_\nu$. Therefore,

$$\pi^*_\nu \Psi^*_{[m],\nu} \Omega_\nu = (\Phi^\nu_m)^* \pi^*_\nu \Omega_\nu = (\Phi^\nu_m)^* i^*_\nu \Omega = i^*_\nu \Phi^*_m \Omega = i^*_\nu \Omega = \pi^*_\nu \Omega_\nu.$$

Since π_ν is a surjective submersion, we may conclude that

$$\Psi^*_{[m],\nu} \Omega_\nu = \Omega_\nu.$$

Thus, we have a symplectic action of M_ν/N_ν on P_ν. ■

An Induced Momentum Map. We now generalize the argument given for the case of the semidirect products to show that there is a momentum map for the action of M_ν/N_ν on P_ν that we just defined. The Heisenberg example already shows that we should not expect this momentum map to be equivariant in general.

Before doing this, we prepare the following elementary but useful lemma.

Lemma 5.6. *Let M be a Lie group and let N be a normal subgroup with corresponding Lie algebras $\mathfrak{m}$ and $\mathfrak{n}$. For $n \in N$ and for $\xi \in \mathfrak{m}$, we have*

$$\mathrm{Ad}_n \xi - \xi \in \mathfrak{n}$$

Proof Let $I_n : M \to M$ denote the inner automorphism for $n \in N$, defined by

$$I_n(m) = nmn^{-1}.$$

Since the map Ad_n is the derivative of the inner automorphism with respect to m at the identity, we get

$$\begin{aligned} \mathrm{Ad}_n \xi - \xi &= \left.\frac{d}{dt}\right|_{t=0} [I_n(\exp(t\xi))] \exp(-t\xi) \\ &= \left.\frac{d}{dt}\right|_{t=0} (n[\exp(t\xi)]n^{-1}) \exp(-t\xi) \\ &= \left.\frac{d}{dt}\right|_{t=0} n[\exp(t\xi))n^{-1} \exp(-t\xi)]. \end{aligned} \tag{5.18}$$

Since N is a normal subgroup, $\exp(t\xi))n^{-1}\exp(-t\xi)$ is a curve in N (passing through the point n^{-1} at $t = 0$), so the result is some element in $\mathfrak{n}$. ■

The next task is to establish the manner in which the momentum map $\mathbf{J}_M : P \to \mathfrak{m}^*$ induces a map $\mathbf{J}_\nu : P_\nu \to (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$.

Lemma 5.7. *Suppose N_ν is connected. Then a map $\mathbf{J}_\nu : P_\nu \to (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ is well defined by the relation*

$$(r'_\nu)^* \circ \mathbf{J}_\nu \circ \pi_\nu = k^*_\nu \circ \mathbf{J}_M \circ i_\nu - \bar{\nu} \tag{5.19}$$

where

$$r_\nu : M_\nu \to M_\nu/N_\nu$$

is the canonical projection,

$$r'_\nu : \mathfrak{m}_\nu \to \mathfrak{m}_\nu/\mathfrak{n}_\nu$$

is the induced Lie algebra homomorphism,

$$k_\nu : \mathfrak{m}_\nu \to \mathfrak{m}$$

is the inclusion,

$$\pi_\nu : \mathbf{J}_N^{-1}(\nu) \to P_\nu$$

is the projection,

$$i_\nu : \mathbf{J}_N^{-1}(\nu) \to P$$

is the inclusion, and $\bar{\nu}$ is some chosen extension of $\nu|\mathfrak{n}_\nu$ to $\mathfrak{m}_\nu$. Equivalently, we have

$$\langle \mathbf{J}_\nu([z]), [\xi] \rangle = \langle \mathbf{J}_M(z), \xi \rangle - \langle \bar{\nu}, \xi \rangle \tag{5.20}$$

where $z \in \mathbf{J}_N^{-1}(\nu)$, $\xi \in \mathfrak{m}_\nu$, $\nu \in \mathfrak{n}^*$, $[z] = \pi_\nu(z)$ denotes the equivalence class of z in $P_\nu = \mathbf{J}_N^{-1}(\nu)/N_\nu$ and $[\xi] = r_\nu(\xi)$ denotes the equivalence class of ξ in $\mathfrak{m}_\nu/\mathfrak{n}_\nu$.

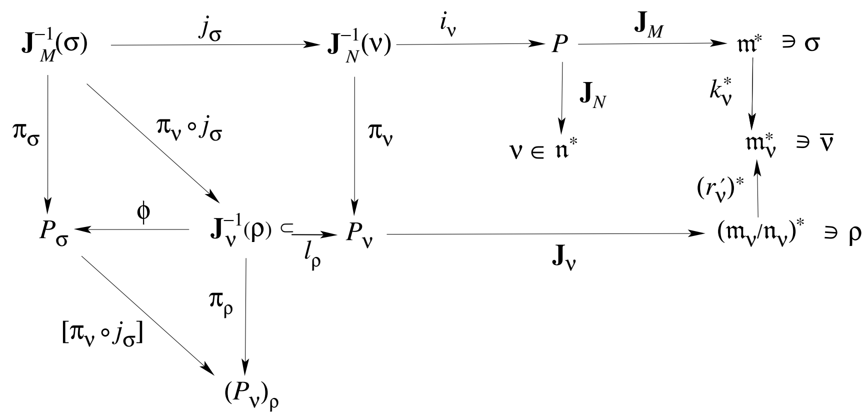

Figure 5.1: Some of the maps involved in reduction by stages.

Proof. It may be helpful to refer to Figure 5.1

First of all, we show that the definition is independent of the representative of $[\xi]$. To do this, it suffices to show that the right hand side of (5.20) vanishes when $\xi \in \mathfrak{n}_\nu$. However, for $\xi \in \mathfrak{n}$, we have

$$\langle \mathbf{J}_M(z), \xi \rangle = \langle \mathbf{J}_N(z), \xi \rangle = \langle \nu, \xi \rangle \, ,$$

since $\mathbf{J}_N(z) = \nu$. Therefore, in this case,

$$\langle \mathbf{J}_M(z), \xi \rangle - \langle \bar{\nu}, \xi \rangle = \langle \mathbf{J}_N(z), \xi \rangle - \langle \nu, \xi \rangle = \langle \nu, \xi \rangle - \langle \nu, \xi \rangle = 0.$$

Next, we must show that the right hand side is independent of the representative of $[z]$. Let $n \in N_\nu$. We must show that

$$\langle \mathbf{J}_M(nz), \xi \rangle - \langle \bar{\nu}, \xi \rangle$$

is independent of n. This is clearly equivalent to showing that

$$\langle \mathbf{J}_M(nz), \xi \rangle = \langle \mathbf{J}_M(z), \xi \rangle$$

for all $n \in N_\nu$. By equivariance of $\mathbf{J}_M$, this in turn is equivalent to

$$\left\langle \mathbf{J}_M(z), \mathrm{Ad}_n^{-1} \xi \right\rangle = \langle \mathbf{J}_M(z), \xi \rangle$$

for all $n \in N_\nu$; i.e., the vanishing of $f(n)$, where

$$f(n) = \left\langle \mathbf{J}_M(z), \mathrm{Ad}_n^{-1} \xi - \xi \right\rangle = \left\langle \nu, \mathrm{Ad}_n^{-1} \xi - \xi \right\rangle$$

by Lemma 5.6, for $z \in \mathbf{J}_N(\nu)$ and $\xi \in \mathfrak{m}_\nu$ fixed and for $n \in N_\nu$.

To show that f vanishes, first of all, note that $f(e) = 0$. Second, we note that the differential of f at the identity in the direction $\eta \in \mathfrak{n}_\nu$ is given by

$$\mathbf{d}f(e) \cdot \eta = \langle \nu, -\mathrm{ad}_\eta \xi \rangle = -\left\langle \mathrm{ad}_\eta^* \nu, \xi \right\rangle = 0$$

since $\mathbf{J}_N(z) = \nu$ and since $\mathrm{ad}_\eta^* \nu = 0$ because $\eta \in \mathfrak{n}_\nu$.

Next, we show that $f(n_1 n_2) = f(n_1) + f(n_2)$. To do this, we write

$$f(n_1 n_2) = \left\langle \nu, \mathrm{Ad}_{n_1 n_2}^{-1} \xi - \xi \right\rangle = \left\langle \nu, \mathrm{Ad}_{n_2}^{-1} \mathrm{Ad}_{n_1}^{-1} \xi - \mathrm{Ad}_{n_2}^{-1} \xi + \mathrm{Ad}_{n_2}^{-1} \xi - \xi \right\rangle .$$

However,

$$\left\langle \nu, \mathrm{Ad}_{n_2}^{-1} \mathrm{Ad}_{n_1}^{-1} \xi - \mathrm{Ad}_{n_2}^{-1} \xi \right\rangle = \left\langle \mathrm{Ad}_{n_2^{-1}}^* \nu, \mathrm{Ad}_{n_1}^{-1} \xi - \xi \right\rangle = \left\langle \nu, \mathrm{Ad}_{n_1}^{-1} \xi - \xi \right\rangle .$$

This calculation shows that $f(n_1 n_2) = f(n_1) + f(n_2)$ as we desired.

Differentiating this relation with respect to n_1 at the identity gives

$$\mathbf{d}f(n_2) \circ T_e R_{n_2} = df(e) = 0$$

and hence $\mathbf{d}f = 0$ on N_ν. Since N_ν is connected we conclude that $f = 0$, which is what we desired to show. ■

Verifying $\mathbf{J}_\nu$ is a Momentum Map. Now we show that $\mathbf{J}_\nu$ is a momentum map, ignoring questions of equivariance for the moment. We shall compute its cocycle shortly.

Proposition 5.8. *The map $\mathbf{J}_\nu$ in the preceding lemma is a momentum map for the action of M_ν/N_ν on P_ν.*

Proof. We first observe that $\mathbf{J}_\nu$ *does depend on the extension $\bar{\nu}$ of $\nu|\mathfrak{n}_\nu$*. If $\bar{\nu}_1$ and $\bar{\nu}_2$ are two such extensions then $(\bar{\nu}_1 - \bar{\nu}_2)|\mathfrak{n}_\nu = 0$ and so it equals $(r_\nu')^*(\rho)$ for $\rho \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$. Formula (5.19) shows that the difference of the two corresponding momentum maps equals ρ (which is precisely the ambiguity in the definition of the momentum map; recall that momentum maps are defined only up to the addition of constant elements in the dual of the Lie algebra).

Secondly, we compute the infinitesimal generator given by $[\xi] = r_\nu'(\xi) \in \mathfrak{m}_\nu/\mathfrak{n}_\nu$. Since

$$\exp_\nu t r_\nu'(\xi) = r_\nu(\exp t\xi),$$

where $\exp_\nu : \mathfrak{m}_\nu/\mathfrak{n}_\nu \to M_\nu/N_\nu$ is the exponential map of the Lie group M_ν/N_ν and exp is that of M_ν, we get for $z \in \mathbf{J}_N^{-1}(\nu)$ using the definition of the

M_ν/N_ν–action on P_ν,

$$
\begin{aligned}
[\xi]_{P_\nu}([z]) &= \left.\frac{d}{dt}\right|_{t=0} \exp_\nu t r'_\nu(\xi) \cdot \pi_\nu(z) \\
&= \left.\frac{d}{dt}\right|_{t=0} r_\nu(\exp t\xi) \cdot \pi_\nu(z) \\
&= \left.\frac{d}{dt}\right|_{t=0} \pi_\nu(\exp t\xi \cdot z) \\
&= T_z\pi_\nu(\xi_P(z)), \qquad (5.21)
\end{aligned}
$$

that is,

$$[\xi]_{P_\nu}([z]) = T_z\pi_\nu(\xi_P(z)). \qquad (5.22)$$

Thirdly, denote by $J^\xi_M : P \to \mathbb{R}$, the map $J^\xi_M(z) = \langle \mathbf{J}_M(z), \xi\rangle$ and similarly for $J^{[\xi]}_\nu : P \to \mathbb{R}$ and note that (5.19) can be written as

$$J^{[\xi]}_\nu(\pi_\nu(z)) = J^\xi_M(z) - \langle \bar{\nu}, \xi\rangle.$$

Taking the z–derivative of this relation in the direction $v \in T_z\mathbf{J}_N^{-1}(\nu)$, we get

$$\mathbf{d}J^{[\xi]}_\nu(\pi_\nu(z)) \cdot T_z\pi_\nu(v) = \mathbf{d}J^\xi_M(z) \cdot v \qquad (5.23)$$

Letting Ω_ν denote the symplectic form on P_ν, for $z \in \mathbf{J}_N^{-1}(\nu)$, $\xi \in \mathfrak{m}_\nu$, $v \in T_z\mathbf{J}_N^{-1}(\nu)$, we get from (5.22) and (5.23)

$$
\begin{aligned}
\Omega_\nu([z])\left([\xi]_{P_\nu}([z]), T_z\pi_\nu(v)\right) &= \Omega_\nu(\pi_\nu(z))\left(T_z\pi_\nu(\xi_P(z)), T_z\pi_\nu(v)\right) \\
&= (\pi_\nu^*\Omega_\nu)(z)\left(\xi_P(z), v\right) \\
&= (i_\nu^*\Omega)(z)(\xi_P(z), v) = \mathbf{d}J^\xi_M(z) \cdot v \\
&= dJ^{[\xi]}_\nu(\pi_\nu(z)) \cdot T_z\pi_\nu(v), \qquad (5.24)
\end{aligned}
$$

which proves that $\mathbf{J}_\nu$ given by (5.19) is a momentum map for the M_ν/N_ν–action on P_ν. ■

Computing the Cocycle of $\mathbf{J}_\nu$. We are now ready to analyze the extent to which we have a lack of equivariance of $\mathbf{J}_\nu$ by computing the associated cocycle.

Proposition 5.9. *The $(\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$–valued one–cycle ϖ of the momentum map $\mathbf{J}_\nu$ is determined by*

$$(r'_\nu)^*(\varpi([m])) = \mathrm{Ad}^*_{m^{-1}}\,\bar{\nu} - \bar{\nu},$$

where $(r'_\nu)^* : (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^* \to \mathfrak{m}^*_\nu$ is the dual of r'_ν.

Proof. Since

$$\mathrm{Ad}_{r_\nu(m)} r'_\nu(\xi) = r'_\nu(\mathrm{Ad}_m \xi) \tag{5.25}$$

for any $m \in M_\nu$, $\xi \in \mathfrak{m}_\nu$, and $z \in \mathbf{J}_N^{-1}(\nu)$, we have

$$\begin{aligned}\langle \mathbf{J}_\nu([m][z]) - &\mathrm{Ad}^*_{[m]^{-1}} \mathbf{J}_\nu([z]), [\xi]\rangle \\ &= \langle \mathbf{J}_\nu([mz]), [\xi]\rangle - \langle \mathbf{J}_\nu([z]), \mathrm{Ad}_{[m]^{-1}}[\xi]\rangle \\ &= \langle \mathbf{J}_M(mz), \xi\rangle - \langle \bar{\nu}, \xi\rangle - \langle \mathbf{J}_\nu([z]), [\mathrm{Ad}_{m^{-1}} \xi]\rangle \\ &= \langle \mathrm{Ad}^*_{m^{-1}} \mathbf{J}_M(z), \xi\rangle - \langle \bar{\nu}, \xi\rangle - \langle \mathbf{J}_M(z), \mathrm{Ad}_{m^{-1}} \xi\rangle + \langle \bar{\nu}, \mathrm{Ad}_{m^{-1}} \xi\rangle \\ &= \langle \bar{\nu}, \mathrm{Ad}_{m^{-1}} \xi - \xi\rangle \\ &= \langle \mathrm{Ad}^*_{m^{-1}} \bar{\nu} - \bar{\nu}, \xi\rangle. \end{aligned} \tag{5.26}$$

Note that if $\xi \in \mathfrak{n}_\nu$, then $\mathrm{Ad}_{m^{-1}} \xi \in \mathfrak{n}_\nu$, since N_ν is a normal subgroup of M_ν. Therefore, denoting by $m\nu$ the action of $m \in M_\nu \subset M$ on $\nu \in \mathfrak{n}^*$, we have

$$\begin{aligned}\langle \mathrm{Ad}^*_{m^{-1}} \bar{\nu}, \xi\rangle &= \langle \bar{\nu}, \mathrm{Ad}_{m^{-1}} \xi\rangle \\ &= \langle \nu, \mathrm{Ad}_{m^{-1}} \xi\rangle \\ &= \langle m\nu, \xi\rangle = \langle \nu, \xi\rangle, \end{aligned} \tag{5.27}$$

since $m \in M_\nu$. This shows that

$$\mathrm{Ad}^*_{m^{-1}} \bar{\nu} - \bar{\nu} \in \mathfrak{n}^0_\nu$$

where $\mathfrak{n}^0_\nu = \{\lambda \in \mathfrak{m}^*_\nu \mid \lambda|\mathfrak{n}_\nu = 0\}$ is the annihilator of $\mathfrak{n}_\nu$ in $\mathfrak{m}^*_\nu$.

However, since $r'_\nu : \mathfrak{m}_\nu \to \mathfrak{m}_\nu/\mathfrak{n}_\nu$ is surjective, its dual $(r'_\nu)^* : (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^* \to \mathfrak{m}^*_\nu$ is injective and it is easy to verify that

$$(r'_\nu)^* \left((\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*\right) \subset \mathfrak{n}^0_\nu.$$

Since

$$\dim \left((r'_\nu)^* \left((\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*\right)\right) = \dim \left(\mathfrak{m}_\nu/\mathfrak{n}_\nu\right) = \dim \mathfrak{m}_\nu - \dim \mathfrak{n}_\nu = \dim \mathfrak{n}^0_\nu$$

it follows that

$$(r'_\nu)^* \left((\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*\right) = \mathfrak{n}^0_\nu. \tag{5.28}$$

(We take the expedient view that in infinite dimensions, this needs to be proven on a case by case basis.) Because of (5.28) it follows that there is a unique $\varpi(m) \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ such that

$$\mathrm{Ad}^*_{m^{-1}} \bar{\nu} - \bar{\nu} = (r'_\nu)^*(\varpi(m)).$$

Let $m_1, m_2 \in M_\nu$. Dualizing relation (5.25) we get

$$\begin{aligned}
(r'_\nu)^*(\varpi(m_1m_2)) &= \mathrm{Ad}^*_{(m_1m_2)^{-1}} \bar\nu - \bar\nu \\
&= \mathrm{Ad}^*_{m_1^{-1}} \mathrm{Ad}^*_{m_2^{-1}} \bar\nu - \mathrm{Ad}^*_{m_1^{-1}} \bar\nu + \mathrm{Ad}^*_{m_1^{-1}} \bar\nu - \bar\nu \\
&= \mathrm{Ad}^*_{m_1^{-1}} \left(\mathrm{Ad}^*_{m_2^{-1}} \bar\nu - \bar\nu \right) + \mathrm{Ad}^*_{m_1^{-1}} \bar\nu - \bar\nu \\
&= \mathrm{Ad}^*_{m_1^{-1}} (r'_\nu)^*(\varpi(m_2)) + (r'_\nu)^*(\varpi(m_1)) \\
&= (r'_\nu)^*(\varpi(m_1)) + (r'_\nu)^* \left(\mathrm{Ad}^*_{[m_1]^{-1}} \varpi(m_2) \right) \\
&= (r'_\nu)^* \left(\varpi(m_1) + \mathrm{Ad}^*_{[m_1]^{-1}} \varpi(m_2) \right). \qquad (5.29)
\end{aligned}$$

Injectivity of $(r'_\nu)^*$ implies that

$$\varpi(m_1m_2) = \varpi(m_1) + \mathrm{Ad}^*_{[m_1]^{-1}} \varpi(m_2). \qquad (5.30)$$

In particular, if $m \in M_\nu$ and $n \in N_\nu$, this relation yields

$$\varpi(nm) = \varpi(n) + \mathrm{Ad}^*_{[n]^{-1}} \varpi(m) = \varpi(n) + \varpi(m),$$

since $[n] = e$. Now we show that $\varpi(n) = 0$. Indeed, if $\xi \in \mathfrak{m}_\nu$,

$$\begin{aligned}
\langle (r'_\nu)^*(\varpi(n)), \xi \rangle &= \langle \mathrm{Ad}^*_{n^{-1}} \bar\nu - \bar\nu, \xi \rangle \\
&= \langle \bar\nu, \mathrm{Ad}_{n^{-1}} \xi - \xi \rangle \\
&= \langle \nu, \mathrm{Ad}_{n^{-1}} \xi - \xi \rangle, \qquad (5.31)
\end{aligned}$$

since by Lemma 5.6, $\mathrm{Ad}_{n^{-1}} \xi - \xi \in \mathfrak{n}_\nu$. However, we already showed in the previous lemma that for N_ν connected

$$\langle \nu, \mathrm{Ad}_{n^{-1}} \xi - \xi \rangle = 0.$$

Thus, for any $n \in N_\nu$, $m \in M_\nu$, we have $\varpi(nm) = \varpi(m)$, which proves that $\varpi(n)$ does depend on $[m]$ and not on m. Denoting this map by the same letter $\varpi : M_\nu/N_\nu \to (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ the formula (5.30) shows that it is a one-cocycle on M_ν/N_ν. ■

Computation of the Isotropy Group. As we discussed with the Heisenberg example, in the case of non-equivariant momentum maps reduction may be carried out by modifying the coadjoint action with a cocycle. Namely, for

$$\lambda \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*,$$

we consider the modified action

$$[m]\lambda = \mathrm{Ad}^*_{[m]^{-1}} \lambda + \varpi([m]).$$

Given $\sigma \in \mathfrak{m}^*$ define $\nu = \sigma|\mathfrak{n} \in \mathfrak{n}^*$ and $\rho \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ by

$$(r'_\nu)^*(\rho) = \sigma|\mathfrak{m}_\nu - \bar{\nu},$$

where $\bar{\nu}$ is an arbitrary extension of $\nu|\mathfrak{n}_\nu$ to $\mathfrak{m}_\nu$. This is makes sense since, for $\eta \in \mathfrak{n}_\nu$, the right hand side satisfies

$$\langle \sigma|\mathfrak{m}_\nu - \bar{\nu}, \eta \rangle = \langle \sigma, \eta \rangle - \langle \bar{\nu}, \eta \rangle = \langle \nu, \eta \rangle - \langle \nu, \eta \rangle = 0,$$

that is, $\sigma|\nu - \bar{\nu} \in \mathfrak{n}_\nu^0$. Observe that ρ depends on the choice of extension $\bar{\nu}$ of ν.

Proposition 5.10. $(M_\nu/N_\nu)_\rho = r_\nu \left((M_\nu)_{\sigma|\mathfrak{m}_\nu} \right)$.

Proof. Note that $[m] \in (M_\nu/N_\nu)_\rho$ if and only if

$$\begin{aligned}(r'_\nu)^*(\rho) &= (r'_\nu)^*([m]\rho) = (r'_\nu)^* \left(\operatorname{Ad}^*_{[m]^{-1}} \rho + \varpi([m]) \right) \\ &= \operatorname{Ad}^*_{m^{-1}} (r'_\nu)^* \rho + (r'_\nu)^*(\varpi([m])) \\ &= \operatorname{Ad}^*_{m^{-1}} (\sigma|\mathfrak{m}_\nu - \bar{\nu}) + \operatorname{Ad}^*_{m^{-1}} \bar{\nu} - \bar{\nu} = \operatorname{Ad}^*_{m^{-1}} (\sigma|\mathfrak{m}_\nu) - \bar{\nu} \\ &= \left(\operatorname{Ad}^*_{m^{-1}} \sigma \right) |\mathfrak{m}_\nu - \bar{\nu}, \end{aligned} \tag{5.32}$$

since $m \in M_\nu$. This is equivalent to

$$\begin{aligned}\langle \sigma, \xi \rangle - \langle \bar{\nu}, \xi \rangle &= \langle \rho, [\xi] \rangle \\ &= \langle [m]_\rho, [\xi] \rangle \\ &= \langle \left(\operatorname{Ad}^*_{m^{-1}} \sigma \right) |\mathfrak{m}_\nu, \xi \rangle - \langle \bar{\nu}, \xi \rangle, \end{aligned} \tag{5.33}$$

for all $\xi \in \mathfrak{m}_\nu$, which says that

$$\left(\operatorname{Ad}^*_{m^{-1}} \sigma \right) |\mathfrak{m}_\nu = \sigma|\mathfrak{m}_\nu,$$

that is $m \in (M_\nu)_{\sigma|\mathfrak{m}_\nu}$. Therefore, we showed that $[m] \in (M_\nu/N_\nu)_\rho$ if and only if $m \in (M_\nu)_{\sigma|\mathfrak{m}_\nu}$, which proves the statement. ■

5.3 The Main Reduction by Stages Theorem

Having established these preliminary facts, we can state the main reduction by stages theorem for group extensions. So far, we have reduced P by the action of N at the point ν to obtain P_ν. Now P_ν can be further reduced by the action of M_ν/N_ν at a regular value $\rho \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$. Let this ***second reduced space*** be denoted by

$$P_{\nu,\rho} = \mathbf{J}^{-1}_{M_\nu/N_\nu}(\rho)/(M_\nu/N_\nu)_\rho$$

where, as usual, $(M_\nu/N_\nu)_\rho$ is the isotropy subgroup for the action of the group M_ν/N_ν on the dual of its Lie algebra.

Assume that $\sigma \in \mathfrak{m}^*$ is a given regular element of $\mathbf{J}_M$ so that we can form the reduced space

$$P_\sigma = \mathbf{J}_M^{-1}(\sigma)/M_\sigma$$

where M_σ is the isotropy subgroup of σ for the action of M on $\mathfrak{m}^*$. We also require that the relation

$$(r'_\nu)^*(\rho) = k_\nu^*\sigma - \bar{\nu}.$$

holds. We assume that the following condition holds:

Hypothesis. *For all $\sigma_1, \sigma_2 \in \mathfrak{m}^*$ such that*

$$\sigma_1|_{\mathfrak{m}_\nu} = \sigma_2|_{\mathfrak{m}_\nu} \quad \textit{and} \quad \sigma_1|_{\mathfrak{n}} = \sigma_2|_{\mathfrak{n}},$$

there exists $n \in N_\nu$ such that $\sigma_2 = \mathrm{Ad}^*_{n^{-1}} \sigma_1$.

This hypothesis holds for semidirect products, central extensions, and, more generally semidirect products with cocycles.

With this assumption, we can state the main theorem.

Theorem 5.11 (Reduction by Stages for Group Extensions). *Using the notations and hypotheses introduced above, there is a (natural) symplectic diffeomorphism between*

$$P_\sigma \quad \textit{and} \quad P_{\nu,\rho}$$

We have set things up so that the proof now proceeds just as in the case of semidirect products. We refer the reader to Marsden, Misiołek, Perlmutter and Ratiu [1998] for details as well as for further exploration of (amongst other things):

1. the case of the semidirect product of two nonabelian groups
2. the use of the cotangent bundle reduction theorem in the group extension context
3. the interpretation of cocycles as curvatures of mechanical connections
4. exploration of specific examples such as the Bott-Virasoro group and applications to the KdV equation.

Acknowledgments. We thank Hernan Cendra, Mark Gotay and the referee for several helpful suggestions. We also thank the Erwin Schrödinger Institute for Mathematical Physics for providing the opportunity to meet in pleasant surroundings during 1994 at which time some of the ideas in the paper were first worked out. The research of JEM was partially supported by NSF grant DMS 96–33161, that of GM by the faculty leave program at Notre Dame, that of MP by by DOE contract DE–FG0395–ER25251 and that of TSR by NSF Grant DMS-9503273, DOE contract DE-FG03-95ER25245-A000, the Erwin Schrödinger Institute, and the Miller Institute of the University of California.

References

Abraham, R. and J.E. Marsden [1978] *Foundations of Mechanics.* Second Edition, Addison-Wesley.

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

Bates, L. [1988] Symmetry preserving deformations of the Kepler problem. *Rep. Math. Phys.* **26**, 413–428.

Bott, R. [1977] On the characteristic classes of groups of diffeomorphisms, *Enseign. Math.* **23**.

Baguis, P. [1998] Semidirect products and the Pukanszky condition, *J. Geom. and Phys.* **25**, 245–270.

Cartan, H. and S. Eilenberg [1956] *Homological algebra*, Princeton University Press.

Cendra, H. and J.E. Marsden [1987] Lin constraints, Clebsch potentials and variational principles, *Physica D* **27**, 63–89.

Cendra, H., A. Ibort, and J.E. Marsden [1987] Variational principal fiber bundles: a geometric theory of Clebsch potentials and Lin constraints, *J. Geom. Phys.* **4**, 183–206.

Cendra, H. J.E. Marsden and T.S. Ratiu [1998] Lagrangian reduction by stages, in preparation.

Duval, C., J. Elhadad, M.J. Gotay, J. Śniatycki and G.M. Tuynman [1991] Quantization and Bosonic BRST Theory. *Ann. Phys.* **206**, 1–26.

Guillemin, V. and S. Sternberg [1980] The moment map and collective motion, *Ann. of Phys.* **1278**, 220–253.

Guillemin, V. and S. Sternberg [1984] *Symplectic Techniques in Physics.* Cambridge University Press.

Guillemin, V. and S. Sternberg [1990] *Variations on a Theme by Kepler.* Colloquium Publications **42**, Amer. Math. Soc.

Harnad, J., J. Hurtubise, and J.E. Marsden [1991] Reduction of Hamiltonian systems with discrete symmetry; see Chapter 8 of Marsden [1992].

Holm, D.D. and B.A. Kuperschmidt [1983] Poisson brackets and clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity. *Physica D* **6**, 347–363

Holm, D. D., Marsden, J. E. and Ratiu, T. [1998a] The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Adv. in Math.* (to appear).

Holm, D. D., J.E. Marsden and T. Ratiu [1998b] The Euler-Poincaré equations in geophysical fluid dynamics, in *Proceedings of the Isaac Newton Institute Programme on the Mathematics of Atmospheric and Ocean Dynamics*, Cambridge University Press (to appear).

Kobayashi, S. and K. Nomizu [1963] *Foundations of Differential Geometry.* Wiley

Krishnaprasad, P.S. and J.E. Marsden [1987] Hamiltonian structure and stability for rigid bodies with flexible attachments, *Arch. Rat. Mech. Anal.* **98**, 137–158.

Kuperschmidt, B.A. and T. Ratiu [1983] Canonical maps between semidirect products with applications to elasticity and superfluids *Comm. Math. Phys.* **90**, 235–250.

Landsman, N.P. [1995] Rieffel induction as generalized quantum Marsden-Weinstein reduction. *J. Geom. and Phys.* **15**, 285–319.

Leonard, N.E. and J.E. Marsden [1997] Stability and drift of underwater vehicle dynamics: mechanical systems with rigid motion symmetry, *Physica D* **105**, 130–162.

Marsden, J.E. [1982] A group theoretic approach to the equations of plasma physics, *Can. Math. Bull.* **25**, 129–142.

Marsden, J.E. [1992], *Lectures on Mechanics* London Mathematical Society Lecture note series, **174**, Cambridge University Press.

Marsden, J.E. and T.J.R. Hughes [1983] *Mathematical Foundations of Elasticity.* Prentice Hall, reprinted by Dover Publications, N.Y., 1994.

Marsden, J.E., G. Misiołek, M. Perlmutter and T. Ratiu [1998] Symplectic Reduction by Stages, *in preparation.*

Marsden, J.E., R. Montgomery, and T.S. Ratiu [1990] *Reduction, symmetry, and phases in mechanics.* Memoirs AMS **436**.

Marsden, J.E. and T.S. Ratiu [1986] Reduction of Poisson manifolds, *Lett. in Math. Phys.* **11**, 161–170.

Marsden, J.E. and T.S. Ratiu [1994] *Introduction to Mechanics and Symmetry.* Texts in Applied Mathematics, **17**, Second Edition, 1998, Springer-Verlag.

Marsden, J.E., T.S. Ratiu, and A. Weinstein [1984a] Semi-direct products and reduction in mechanics, *Trans. Am. Math. Soc.* **281**, 147–177.

Marsden, J.E., T.S. Ratiu, and A. Weinstein [1984b] Reduction and Hamiltonian structures on duals of semidirect product Lie Algebras, *Cont. Math. AMS* **28**, 55–100.

Marsden, J.E. and J. Scheurle [1993a] Lagrangian reduction and the double spherical pendulum, *ZAMP* **44**, 17–43.

Marsden, J.E. and J. Scheurle [1993b] The reduced Euler-Lagrange equations, *Fields Institute Comm.* **1**, 139–164.

Marsden, J.E. and A. Weinstein [1974] Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.* **5**, 121–130.

Marsden, J.E., A. Weinstein, T.S. Ratiu, R. Schmid, and R.G. Spencer [1983] Hamiltonian systems with symmetry, coadjoint orbits and plasma physics, in Proc. IUTAM-IS1MM Symposium on *Modern Developments in Analytical Mechanics*, Torino 1982, *Atti della Acad. della Sc. di Torino* **117**, 289–340.

Moser, J. [1970] Regularization of Kepler's problem and the averaging method on a manifold *Comm. Pure. Appl. Math.* **23**, 609–636.

Ortega J.-P. and T.S. Ratiu [1998] *Symmetry, Reduction, and Stability in Hamiltonian Systems*, in preparation.

Ratiu, T.S. [1980] *Thesis*. Department of Mathematics, University of California at Berkeley.

Ratiu, T.S. [1981] Euler-Poisson equations on Lie algebras and the N-dimensional heavy rigid body. *Proc. Natl. Acad. Sci. USA* **78**, 1327–1328.

Ratiu, T.S. [1982] Euler-Poisson equations on Lie algebras and the N-dimensional heavy rigid body, *Am. J. Math.* **104**, 409–448, 1337.

Sjamaar, R. and E. Lerman [1991] Stratified symplectic spaces and reduction, *Ann. of Math.* **134**, 375–422.

Souriau, J-M. [1973] Sur la variété de Kepler. *Symposia Mathematica* **XIV** (Convegno di Geometria Simplettica e Fisica Matematica, INDAM, Rome, 1973), 343–360. Academic Press, London, 1974.

Sudarshan, E.C.G. and N. Mukunda [1974] *Classical Mechanics: A Modern Perspective.* Wiley, New York, 1974; Second Edition, Krieber, Melbourne–Florida, 1983.

Vanhaecke, P. [1996] *Integrable Systems in the Realm of Algebraic Geometry*, Springer Lecture Notes in Mathematics, **1638**.

Vinogradov, A.M. and B.A. Kuperschmidt [1977] The structures of Hamiltonian mechanics. *Russ. Math. Surv.* **32**, 177–243.

Ziegler, F. [1996] *Méthode des Orbites et Reprśentations Quantiques*, Thesis, Université de Provence.